

¹ **Searches for Supersymmetric signatures in
2 all hadronic final states with the α_T
3 variable.**

⁴ Darren Burton

⁵ Imperial College London
⁶ Department of Physics

⁷ A thesis submitted to Imperial College London
⁸ for the degree of Doctor of Philosophy

Abstract

A search for supersymmetric particles is presented, using the Compact Muon Solenoid detector at the Large Hadron Collider, with a signature of missing energy in events with high p_T jets is presented. The analysis is performed with 11.7 fb^{-1} of data, collected at a center-of-mass energy of 8 TeV during the 2012 run period. The dimensionless kinematic variable α_T is used to select events with genuine missing energy signatures, while Standard Model backgrounds in the signal region estimated using data driven control samples. 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 presented. The α_T search is used as a vehicle to demonstrate proof of principle and as a search region for this technique. Additionally the efficiency of the hadronic Level-1 single jet triggers are measured throughout the 2012 run period. Results are presented with a view to comparing L1 jet performance, before and after, a change to the jet seed algorithm implemented during data taking. No excess of events is found over Standard Model expectations in the α_T search. Exclusion limits are set at the 95% confidence level in the parameter space of simplified models, with special emphasis on compressed spectra and natural SUSY scenarios.

30

Declaration

31

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

37

Darren Burton

38

39

Acknowledgements

40 Of the many people who deserve thanks, some are particularly prominent.... Thank Rob
41 Thank Jad

42 Contents

43	List of Figures	viii
44	List of Tables	xiv
45	1. Introduction	2
46	2. A Theoretical Overview	5
47	2.1. The Standard Model	5
48	2.1.1. Gauge Symmetries of the SM	7
49	2.1.2. The Electroweak Sector and Electroweak Symmetry Breaking	9
50	2.2. Motivation for Physics Beyond the Standard Model	13
51	2.3. Supersymmetry Overview	14
52	2.3.1. R-Parity	16
53	2.4. Experimental Signatures of SUSY at the LHC	16
54	2.4.1. Simplified models	18
55	3. The LHC And The CMS Detector	20
56	3.1. The LHC	20
57	3.2. The CMS Detector	23
58	3.2.1. Detector subsystems	23
59	3.2.2. Tracker	24
60	3.2.3. Electromagnetic calorimeter	25
61	3.2.4. Hadronic calorimeter	26
62	3.2.5. Muon systems	28
63	3.3. Event Reconstruction and Object Definition	28
64	3.3.1. Jets	28
65	3.3.2. B-tagging	30
66	3.4. Triggering System	33
67	3.4.1. The level-1 trigger	34

68	3.4.2. The L1 trigger jet algorithm	35
69	3.4.3. Measuring L1 jet trigger efficiencies	37
70	3.4.4. Effects of the L1 jet seed	38
71	3.4.5. Robustness of L1 jet performance against pile-up	40
72	3.4.6. Summary	43
73	4. SUSY Searches In Hadronic Final States	45
74	4.1. An Introduction to the α_T Search	46
75	4.1.1. The α_T variable	48
76	4.2. Search Strategy	50
77	4.2.1. Physics objects	53
78	4.2.2. Event selection	55
79	4.2.3. Control sample definition and background estimation	58
80	4.2.4. Estimating the QCD multi-jet background	65
81	4.3. Trigger Strategy	67
82	4.4. Measuring MC Normalisation Factors via H_T Sidebands	68
83	4.5. Determining MC Simulation Yields with Higher Statistical Precision	69
84	4.5.1. The formula method	70
85	4.5.2. Establishing proof of principle	71
86	4.5.3. Correcting measured efficiencies in simulation to data	71
87	4.6. Systematic Uncertainties on Transfer Factors	74
88	4.6.1. Determining systematic uncertainties from closure tests	78
89	4.7. Simplified Models, Efficiencies and Systematic Uncertainties	80
90	4.7.1. Signal efficiency	81
91	4.7.2. Applying b-tag scale factor corrections in signal samples	82
92	4.7.3. Experimental uncertainties	83
93	4.8. Statistical Framework	86
94	4.8.1. Hadronic sample	86
95	4.8.2. H_T evolution model	87
96	4.8.3. Electroweak Sector (EWK) control samples	87
97	4.8.4. Contributions from signal	90
98	4.8.5. Total likelihood	91
99	5. Results and Interpretation	92
100	5.1. Standard Model	92
101	5.2. SUSY	101
102	5.2.1. The CL_s method	101

103	5.2.2. Interpretation in simplified signal models	102
104	6. Searches For Natural SUSY With B-tag Templates.	106
105	6.1. Concept	106
106	6.2. Application to the α_T Search	108
107	6.2.1. Proof of principle in simulation	109
108	6.2.2. Results in a data control sample	112
109	6.2.3. Application to the α_T hadronic search region	114
110	6.3. Summary	116
111	7. Conclusions	117
112	A. Miscellaneous	119
113	A.1. Noise Filters	119
114	A.2. Primary Vertices	121
115	B. L1 Jets	123
116	B.1. Jet matching efficiencies	123
117	B.2. Leading Jet Energy Resolution	124
118	B.3. Resolution for Energy Sum Quantities	127
119	C. Additional material on background estimation methods	132
120	C.1. Determination of k_{QCD}	132
121	C.2. Effect of varying background cross sections on closure tests	133
122	D. Additional Material For B-tag Template Method	135
123	D.1. Templates Fits in Simulation	135
124	D.2. Pull Distributions for Template Fits	138
125	D.3. Templates Fits in Data Control Sample	139
126	D.4. Templates Fits in Data Signal Region	141
127	Bibliography	144

¹²⁸ List of Figures

¹²⁹ 2.1.	One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion.	¹³⁰ 14
¹³¹ 2.2.	Two example simplified model decay chains.	¹³¹ 19
¹³² 3.1.	A top down layout of the LHC, with the position of the four main detectors labelled.	¹³³ 21
¹³⁴ 3.2.	The total integrated luminosity delivered to and collected by Compact Muon Solenoid (CMS) during the 2012 8 TeV pp runs	¹³⁵ 22
¹³⁶ 3.3.	A pictorial depiction of the CMS detector.	¹³⁶ 24
¹³⁷ 3.4.	Illustration of the CMS Electromagnetic CALorimeter (ECAL).	¹³⁷ 26
¹³⁸ 3.5.	Schematic of the CMS Hadronic CALorimeter (HCAL).	¹³⁸ 27
¹³⁹ 3.6.	Combined Secondary Vertex (CSV) algorithm discriminator values in enriched ttbar and inclusive multi jet samples	¹⁴⁰ 31
¹⁴¹ 3.7.	Data/MC b-tag scale factors derived using the Combined Secondary Vertex Medium Working Point (CSVM) tagger.	¹⁴² 32
¹⁴³ 3.8.	Data/MC mis-tag scale factors derived using the CSVM tagger.	¹⁴³ 33
¹⁴⁴ 3.9.	The CMS Level 1 Trigger (L1) Trigger system.	¹⁴⁴ 34
¹⁴⁵ 3.10.	Illustration of the Level-1 jet finding algorithm.	¹⁴⁵ 36
¹⁴⁶ 3.11.	L1 jet efficiency turn-on curves as a function of the offline CaloJet and PFJet E_T	¹⁴⁷ 38
¹⁴⁸ 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.	¹⁴⁹ 39

150	3.13. Trigger cross section for the L1HTT150 trigger path.	40
151	3.14. L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T	41
152	3.15. L1 jet efficiency turn-on curves as a function of the leading offline E_T Calo	
153	(left) and PF (right) jet, for low, medium and high pile-up conditions.	42
154	3.16. Fit values from an Exponentially Modified Gaussian (EMG) function fitted	
155	to the resolution plots of leading Calo jet E_T measured as a function of	
156	$\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and high pile-up conditions.	43
157	3.17. Fit values from an EMG function fitted to the resolution plots of leading	
158	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	
159	high pile-up conditions.	44
160	4.1. Reconstructed offline H_T for 11.7fb^{-1} of data after a basic pre-selection.	47
161	4.2. The event topologies of background QCD dijet events (right) and a generic	
162	SUSY signature with genuine Z_T (left).	48
163	4.3. The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet	
164	multiplicities after a full analysis selection and shown for $H_T > 375$	50
165	4.4. Pictorial depiction of the analysis strategy employed by the α_T search to	
166	increase sensitivity to a wide spectra of SUSY models.	52
167	4.5. Data/MC comparisons of key variables for the hadronic signal region.	58
168	4.6. Data/MC comparisons of key variables for the $\mu + \text{jets}$ selection.	61
169	4.7. Data/MC comparisons of key variables for the $\mu\mu + \text{jets}$ selection.	63
170	4.8. Data/MC comparisons of key variables for the $\gamma + \text{jets}$ selection.	64
171	4.9. QCD sideband regions, used for determination of k_{QCD}	66
172	4.10. Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets determined	
173	from all jets within each individual analysis H_T bin.	73
174	4.11. Sets of closure tests overlaid on top of the systematic uncertainty used for	
175	each of the five H_T regions.	79
176	4.12. Signal efficiencies for the Simplified Model Spectra (SMS) models (a) T1	
177	and (b) T2.	81

178	5.1. Comparison of the observed yields and Standard Model (SM) expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \leq 3$	94
182	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 $n_{jet} \leq 3$	95
185	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 $n_{jet} \leq 3$	96
188	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 $n_{jet} \geq 4$	97
191	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 $n_{jet} \geq 4$	98
194	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 $n_{jet} \geq 4$	99
197	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 $n_{jet} \geq 4$	99
200	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 $n_{jet} \geq 4$	100
203	5.9. Production and decay modes for the various SMS models interpreted within the analysis.	104
205	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.	105

207	6.1. The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging rate (c) as measured in simulation after the α_T analysis, $\mu + \text{jets}$ control sample selection in the region $H_T > 375$	108
210	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
213	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
216	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
219	B.1. Leading jet matching efficiency as a function of the offline CaloJet E_T	123
220	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
222	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
224	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
226	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
228	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
230	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
232	B.8. H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile up conditions.	130
234	B.9. H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions.	130

236	B.10. \mathcal{H}_T resolution parameters in bins of \mathcal{H}_T measured for the defined low,	
237	medium and high pile up conditions.	131
238	B.11. \mathcal{H}_T resolution parameters in bins of PF \mathcal{H}_T measured for the defined low,	
239	medium and high pile up conditions.	131
240	C.1. $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is	
241	conducted between the H_T region $275 < H_T < 575$	132
242	C.2. Sets of closure tests overlaid on top of the systematic uncertainty used for	
243	each of the five H_T regions.	133
244	C.3. Sets of closure tests overlaid on top of the systematic uncertainty used for	
245	each of the five H_T regions.	133
246	D.1. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
247	bins taken directly from simulation in the region $H_T > 375$ GeV, for the	
248	$n_{jet} = 3$ category.	136
249	D.2. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
250	bins taken directly from simulation in the region $H_T > 375$ GeV, for the	
251	$n_{jet} = 4$ category.	137
252	D.3. Pull distributions of $\frac{(\theta - \hat{\theta})}{\sigma}$ for 10^4 pseudo-experiments generated from a	
253	gaussian distribution centred on the n_b^{reco} template values from simulation	
254	with width σ	138
255	D.4. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
256	bins taken directly from data, for the $n_{jet} = 3$ category and medium CSV	
257	working point.	139
258	D.5. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
259	bins taken directly from data, for the $n_{jet} = 4$ category and medium CSV	
260	working point.	140
261	D.6. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$	
262	bins taken from data, in the $n_{jet} = 3$ and $H_T > 375$ category for all CSV	
263	working points.	141

264	D.7. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, in the $n_{jet} = 4$ and $H_T > 375$ category for all CSV working points.	142
265		
266		

²⁶⁷ List of Tables

²⁶⁸	2.1. The fundamental particles of the SM, with spin, charge and mass displayed.	6
²⁶⁹	3.1. Results of a cumulative EMG function fit to the turn-on curves for L1	
²⁷⁰	single jet triggers in 2012 Run Period C.	38
²⁷¹	3.2. Results of a cumulative EMG function fit to the turn-on curves for L1	
²⁷²	single jet triggers in the 2012 run period B and C.	40
²⁷³	3.3. Results of a cumulative EMG function fit to the turn-on curves for H_T in	
²⁷⁴	2012 run period B and C.	41
²⁷⁵	3.4. Results of a cumulative EMG function fit to the efficiency turn-on curves	
²⁷⁶	for L1 single jet triggers in the 2012 run period C, for low,medium and	
²⁷⁷	high pile-up conditions.	41
²⁷⁸	3.5. Results of a cumulative EMG function fit to the efficiency turn-on curves	
²⁷⁹	for Level-1 single jet triggers in the 2012 run period C, for low,medium	
²⁸⁰	and high pile-up conditions.	42
²⁸¹	4.1. A summary of the SMS models interpreted in this analysis, involving both	
²⁸²	direct (D) and gluino-induced (G) production of squarks and their decays.	46
²⁸³	4.2. Muon Identification criteria used within the analysis for selection/veto	
²⁸⁴	purposes in the muon control/signal selections.	53
²⁸⁵	4.3. Photon Identification criteria used within the analysis for selection/veto	
²⁸⁶	purposes in the $\gamma +$ jets control/signal selections.	54
²⁸⁷	4.4. Electron Identification criteria used within the analysis for veto purposes.	55
²⁸⁸	4.5. Jet thresholds used in the three H_T regions of the analysis.	56

289	4.6. Best fit values for the parameters k_{QCD} obtained from sideband regions B,C ₁ ,C ₂ ,C ₃	67			
291	292	4.7. Measured efficiencies of the H_T and α_T legs of the HT and HT_alphaT triggers in independent analysis bins.	68		
293	4.8. k-factors calculated for different EWK processes.	69			
294	295	4.9. Comparing yields in simulation within the $\mu + \text{jets}$ selection determined from the formula method described in Equation (4.11), and that taken directly from simulation	72		
297	298	299	300	4.10. The absolute change in the Transfer Factor (TF)'s used to predict the entire signal region SM background, using the $\mu + \text{jets}$ control sample when the systematic uncertainties of the data to simulation scale factors are varied by $\pm 1\sigma$	74
301	302	303	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	
304	305	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		
306	307	308	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	
309	310	4.14. Calculated systematic uncertainties for the five H_T regions, determined from the closure tests.	78		
311	312	313	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	
314	315	316	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	
317	4.17. The systematic parameters used in H_T bins.	89			

318	4.18. Nuisance parameters used within the different hadronic signal bins of the analysis	91
320	5.1. Summary of control samples used by each fit results, and the Figures in which they are displayed.	93
322	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
326	5.3. A table representing the SMS models interpreted within the analysis.	102
327	6.1. Typical underlying b-quark content of different SM processes which are common to many SUSY searches.	107
329	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 fb^{-1}	110
332	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 fb^{-1} of data at $\sqrt{s} = 8\text{TeV}$	112
335	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 fb^{-1} of data at $\sqrt{s} = 8\text{TeV}$	114
338	A.1. Criteria for a reconstructed jet to pass the loose calorimeter jet id.	119
339	A.2. Criteria for a reconstructed jet to pass the loose PF jet id.	120
340	A.3. Noise filters that are applied to remove spurious and non-physical \cancel{E}_T signatures within the CMS detector.	121
342	A.4. Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex.	122

344	B.1. Results of a cumulative EMG function fit to the turn-on curves for the	
345	matching efficiency of the leading jet in an event to a Level-1 jet in run	
346	2012C and 2012B data.	124
347	C.1. Translation factors constructed from the $\mu +$ jets control sample and	
348	signal selection MC, to predict yields for the $W +$ jets and $t\bar{t}$ back-grounds	
349	in the signal region.	134

350 Acronyms

351	ALICE	A Large Ion Collider Experiment
352	ATLAS	A Toroidal LHC ApparatuS
353	APD	Avalanche Photo-Diodes
354	BSM	Beyond Standard Model
355	CERN	European Organization for Nuclear Research
356	CMS	Compact Muon Solenoid
357	CMSSM	Compressed Minimal SuperSymmetric Model
358	CSC	Cathode Stripe Chamber
359	CSV	Combined Secondary Vertex
360	CSVM	Combined Secondary Vertex Medium Working Point
361	DT	Drift Tube
362	ECAL	Electromagnetic CALorimeter
363	EB	Electromagnetic CALorimeter Barrel
364	EE	Electromagnetic CALorimeter Endcap
365	ES	Electromagnetic CALorimeter pre-Shower
366	EMG	Exponentially Modified Gaussian
367	EPJC	European Physical Journal C
368	EWK	Electroweak Sector
369	GCT	Global Calorimeter Trigger
370	GMT	Global MuonTrigger
371	GT	Global Trigger
372	HB	Hadron Barrel
373	HCAL	Hadronic CALorimeter

374	HE	Hadron Endcaps
375	HF	Hadron Forward
376	HLT	Higher Level Trigger
377	HO	Hadron Outer
378	HPD	Hybrid Photo Diode
379	ISR	Initial State Radiation
380	LUT	Look Up Table
381	L1	Level 1 Trigger
382	LHC	Large Hadron Collider
383	LHCb	Large Hadron Collider Beauty
384	LSP	Lightest Supersymmetric Partner
385	NLL	Next to Leading Logarithmic Order
386	NLO	Next to Leading Order
387	NNLO	Next to Next Leading Order
388	POGs	Physics Object Groups
389	PS	Proton Synchrotron
390	QED	Quantum Electro-Dynamics
391	QCD	Quantum Chromo-Dynamics
392	QFT	Quantum Field Theory
393	RBXs	Readout Boxes
394	RPC	Resistive Plate Chamber
395	RCT	Regional Calorimeter Trigger
396	RMT	Regional Muon Trigger
397	SUSY	SUperSYmmetry
398	SM	Standard Model

399	SMS	Simplified Model Spectra
400	SPS	Super Proton Synchrotron
401	TF	Transfer Factor
402	TP	Trigger Primitive
403	VEV	Vacuum Expectation Value
404	VPT	Vacuum Photo-Triodes
405	WIMP	Weakly Interacting Massive Particle

⁴⁰⁶

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

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

⁴⁰⁷

Chapter 1.

⁴⁰⁸ Introduction

⁴⁰⁹ During the 20th century great advances have been made in the human understanding
⁴¹⁰ of the universe, it's origins, it's future and its composition. The Standard Model (**SM**)
⁴¹¹ first formulated in the 1960's is one of the crowning achievements in science's quest to
⁴¹² explain the most fundamental processes and interactions that make up our universe. It
⁴¹³ has provided a highly successful explanation of a wide range of phenomena in Particle
⁴¹⁴ Physics and has stood up to extensive experimental scrutiny [1].

⁴¹⁵ Despite it's successes it is not a complete theory, with significant questions remaining
⁴¹⁶ unanswered. It describes only three of the four known forces with gravity not incorpo-
⁴¹⁷ rated within the framework of the **SM**. Cosmological experiments infer that just $\sim 4\%$
⁴¹⁸ of the observable universe exists as matter, with elusive "Dark Matter" accounting for a
⁴¹⁹ further $\sim 23\%$ [2]. However no particle predicted by the **SM** is able to account for it. At
⁴²⁰ higher energy scales and small distances the (non-)unification of the fundamental forces
⁴²¹ point to problems with the **SM** at least at higher energies not yet probed experimentally.

⁴²² Many theories exist as extensions to the **SM**, which predict a range of observables that
⁴²³ can be detected at the Large Hadron Collider (**LHC**) of which SUperSYmmetry (**SUSY**)
⁴²⁴ is one such example. It predicts a new symmetry of nature in which all current particles
⁴²⁵ in the **SM** would have a corresponding supersymmetric partner. Common to most
⁴²⁶ Supersymmetric theories is a stable, weakly interacting Lightest Supersymmetric Partner
⁴²⁷ (**LSP**), which has the properties of a possible dark matter candidate. The **SM** and the
⁴²⁸ main principles of Supersymmetric theories are outlined in Chapter 2, with emphasis
⁴²⁹ placed on how experimental signatures of **SUSY** may reveal themselves at the **LHC**.

⁴³⁰ The experimental goal of the **LHC** is to further test the framework of the **SM**, exploring the
⁴³¹ TeV mass scale for the first time, and to seek a connection between the particles produced

432 in proton collisions and dark matter. The first new discovery by this extraordinary
433 machine was announced on the 4th of July 2012. The long-awaited discovery was the
434 culmination decades of experimental endeavours in the search for the Higgs boson,
435 providing an answer to the mechanism of electroweak symmetry breaking within the **SM**
436 [3][4].

437 This discovery was made possible through data taken by the two multi purpose detectors
438 (**CMS** and A Toroidal LHC ApparatuS (**ATLAS**)) located on the **LHC** ring. An experi-
439 mental description of the **CMS** detector and the **LHC** is described in Chapter 3, including
440 some of the object reconstruction used by **CMS** in searches for **SUSY** signatures. The
441 performance of the **CMS** Level-1 calorimeter trigger, benchmarked by the author is also
442 included within this chapter.

443 The analysis conducted by the author is detailed within Chapter 4. This chapter contains
444 a description of the search for evidence of the production of Supersymmetric particles
445 at the **LHC**. The main basis of the search centres around the kinematic dimensionless
446 α_T variable, which provides strong rejection of backgrounds with fake missing energy
447 signatures whilst maintaining good sensitivity to a variety of **SUSY** topologies. The
448 author's work (as an integral part of the analysis group) is documented in detail, which
449 has culminated in numerous publications over the past two years. The latest of which
450 was published in the European Physical Journal C (**EPJC**) [5] and contains the results
451 which are discussed within this and the sequential Chapters.

452 The author in particular has played a major role in the extension of the α_T analysis into
453 the additional b-tagged and jet multiplicity dimensions increasing the sensitivity of the
454 analysis to a range of **SUSY** topologies. Additionally the author has worked extensively
455 in both increasing the statistical precision of electroweak predictions measured from
456 simulation through analytical techniques, and the derivation of a data driven systematic
457 uncertainty through the establishment of closure tests within the control samples of the
458 analysis.

459 Additionally a method to search for **SUSY** signatures which are rich in top and bottom
460 flavoured jet final states is dissussed in Chapter 6. A parametrisation of the b-tagging
461 distribution for different Electroweak processes is used to establish templates, which
462 are then used to estimate the expected number of 3 or 4 b-tagged jet events from **SM**
463 processes. The α_T search is used as a cross check for this template method to establish
464 it's functionality.

465 Finally the interpretation of such results within the framework of a variety of Simplified
466 Model Spectra (**SMS**), which describe an array of possible **SUSY** event topologies is
467 documented in Chapter 5. A description of the statistical model used to derive these
468 interpretations and the possible implications of the results presented in this thesis is
469 discussed within this Chapter. Natural units are used throughout this thesis in which \hbar
470 $= c = 1$.

Chapter 2.

⁴⁷¹ A Theoretical Overview

⁴⁷² Within this chapter, a brief introduction and background to the **SM** is given. Its success
⁴⁷³ as a rigorously tested and widely accepted theory is discussed as are its deficiencies,
⁴⁷⁴ leading to the argument that this theory is not a complete description of our universe.
⁴⁷⁵ The motivations for new physics at the TeV scale and in particular Supersymmetric
⁴⁷⁶ theories are outlined within Section (2.3), with the chapter concluding with how an
⁴⁷⁷ experimental signature of such theories can be produced and observed at the **LHC**,
⁴⁷⁸ Section (2.4).

⁴⁷⁹ 2.1. The Standard Model

⁴⁸⁰ The **SM** is the name given to the relativistic Quantum Field Theory (**QFT**), where
⁴⁸¹ particles are represented as excitations of fields, which describe the interactions and
⁴⁸² properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field
⁴⁸³ theory which contains three symmetries: $SU(3)$ for colour charge, $SU(2)$ for weak isospin
⁴⁸⁴ and $U(1)$ relating to weak hyper charge, which require its Lagrangian \mathcal{L}_{SM} to be invariant
⁴⁸⁵ under local gauge transformation.

⁴⁸⁶ Within the **SM** theory, matter is composed of spin $\frac{1}{2}$ fermions, which interact with each
⁴⁸⁷ other via the exchange of spin-1 gauge bosons. A summary of the known fundamental
⁴⁸⁸ fermions and bosons is given in Table 2.1.

⁴⁸⁹ Fermions are separated into quarks and leptons of which only quarks interact with the
⁴⁹⁰ strong nuclear force. Quarks unlike leptons are not seen as free particles in nature, but
⁴⁹¹ rather exist only within baryons, composed of three quarks with an overall integer charge,
⁴⁹² and quark-anti-quark pairs called mesons. Both leptons and quarks are grouped into

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

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

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

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

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

506 2.1.1. Gauge Symmetries of the SM

- 507 Symmetries are of fundamental importance in the description of physical phenomena.
 508 Noether's theorem states that for a dynamical system, the consequence of any symmetry
 509 is an associated conserved quantity [13]. Invariance under translations, rotations, and
 510 Lorentz transformations in physical systems lead to conservation of momentum, energy
 511 and angular momentum.
- 512 In the **SM**, a quantum theory described by Lagrangian formalism, the weak, strong and
 513 electromagnetic interactions are described in terms of “gauge theories”. A gauge theory
 514 possesses invariance under a set of “local transformations”, which are transformations
 515 whose parameters are space-time dependent. The requirement of gauge invariance within
 516 the **SM** necessitates the introduction of force-mediating gauge bosons and interactions
 517 between fermions and the bosons themselves. Given the nature of the topics covered by
 518 this thesis, the formulation of **EWK** within the **SM** Lagrangian is reviewed within this
 519 section.
- 520 The simplest example of the application of the principle of local gauge invariance within
 521 the **SM** is in Quantum Electro-Dynamics (**QED**), the consequences of which require a
 522 massless photon field [14][15].
- 523 Starting from the free Dirac Lagrangian written as

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

- 524 where ψ represents a free non interacting fermionic field, with the matrices $\gamma^\mu, \mu \in 0, 1, 2, 3$
 525 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
 526 flat space-time metric $(+, -, -, -)$ and I_4 the 4×4 identity matrix.
- 527 Under a local U(1) abelian gauge transformation in which ψ transforms as:

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

- 528 the kinetic term of the Lagrangian does not remain invariant, due to the partial derivative
 529 interposed between the $\bar{\psi}$ and ψ 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_μ , that transforms covariantly under phase transformations is introduced. In doing this a vector field A_μ 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 ∂_μ by D_μ :

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

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.

2.1.2. The Electroweak Sector and Electroweak Symmetry Breaking

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

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

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

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

566 where the left handed fermions form isospin doubles χ_L and the right handed fermions
 567 are isosinglets ψ_R . For the first generation of leptons and quarks this represents

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_L$$

$$\psi_R = e_R, \quad u_R, d_R \quad (2.9)$$

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

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

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

$$\mathcal{L}_1 = \bar{\chi}_L \gamma^\mu [i\partial_\mu - g \frac{1}{2} \tau \cdot W_\mu - g' (-\frac{1}{2}) B_\mu] \chi_L$$

$$+ \bar{e}_R \gamma^\mu [i\partial_\mu - g' (-1) B_\mu] e_R - \frac{1}{4} W_{\mu\nu} \cdot W^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (2.11)$$

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

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

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

575 corresponding to the kinetic energy and self coupling of the W_μ fields and the kinetic
 576 energy term of the B_μ field.

- 577 None of these gauge bosons are physical particles, and instead linear combinations of
 578 these gauge bosons make up γ and the W and Z bosons, defined as

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

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

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

586 The Higgs mechanism induces spontaneous symmetry breaking through the introduction
 587 of a complex scalar SU(2) doublet field ϕ which attains a non-zero Vacuum Expectation
 588 Value (VEV) [17][18][19][20].

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

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

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

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

590 where the covariant derivative D_μ is that defined in Equation (2.10). The last two terms
 591 of \mathcal{L}_{Higgs} correspond to the Higgs potential, in which real positive values of μ^2 and λ are
 592 required to ensure the generation of masses for the bosons and leptons. The minimum of

593 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
594 the **VEV**.

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

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

598 where the fluctuations from the vacuum ϕ_0 are parametrized in terms of four real fields,
599 $\theta_1, \theta_2, \theta_3$ and $h(x)$.

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

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

603 where the relations between the physical and electroweak gauge fields from Equation
604 (2.14) are used. The W^\pm and Z bosons can then be determined to be

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

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

611 2.2. Motivation for Physics Beyond the Standard 612 Model

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

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

624 Some other deficiencies with the **SM** include the fact that the predicted rate of Charge-
625 Parity violation does not account for the matter dominated universe which we inhabit,
626 and the **SM** prediction of zero neutrino mass conflicts with the observation of neutrino
627 flavour mixing, attributed to mixing between neutrino mass eigenstates [23][24].

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

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

638 In the absence of new physics between the TeV and Planck scale, calculating beyond
639 tree-level contributions to the Higgs mass term given by its self interaction, result in
640 divergent terms that push the Higgs mass up to the planck mass M_{planck} .

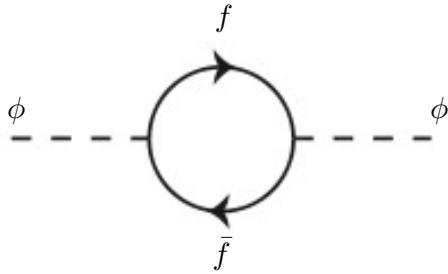


Figure 2.1.: One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion.

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

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

644 where λ_f represents the coupling strength for each type of fermion $\propto m_f$, and Λ the
 645 cutoff energy scale at which the **SM** ceases to be a valid theory.

646 To recover the mass of the now discovered Higgs boson would require a fine-tuning of
 647 the parameters to cancel out these mass corrections of the Higgs mass to the scale of
 648 30 orders of magnitude. This appears as an unnatural solution to physicists and it is
 649 this hierarchy problem that provides one of the strongest motivations for the theory of
 650 SUperSYmmetry (**SUSY**).

651 2.3. Supersymmetry Overview

652 Supersymmetry provides potential solutions to many of the issues raised in the previous
 653 section. It provides a dark matter candidate, can explain baryogenesis in the early
 654 universe and also provides an elegant solution to the hierarchy problem [26][27][28][29].
 655 At its heart it represents a new space-time symmetry that relates fermions and bosons.
 656 This symmetry converts bosonic states into fermionic states, and vice versa, see Equation
 657 (2.21) ,

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

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

660 This symmetry operator therefore acts upon a particles spin altering it by a half integer
 661 value. The consequences of the introduction of this additional space-time symmetry
 662 introduce a new rich phenomenology. For example in supersymmetric theories, both
 663 the left handed $SU(2)$ doublet and right handed singlet of fermions will have a spin-0
 664 superpartner, containing the same electric charge, weak isospin, and colour as its **SM**
 665 partner. In the case of the leptons $(\nu_l, l)_L$, they will have two superpartners, a sneutrino
 666 $\tilde{\nu}_l{}_L$ and a slepton \tilde{l}_L , whilst the singlet l_R also has a superpartner slepton \tilde{l}_R .

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

673 One of the simplest forms of **SUSY**, is to simply have a set of **SM** supersymmetric partners
 674 with the same mass and interactions as their counterparts. However the currently lack
 675 of any experimental evidence for the predicted sparticle spectrum implies **SUSY** must
 676 be a broken symmetry in which any sparticle masses must be greater than their SM
 677 counterparts.

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

684 2.3.1. R-Parity

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

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

694 R-parity is described by the equation

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

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

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

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

707 2.4. Experimental Signatures of **SUSY** at the **LHC**

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

- 710 • squark/anti-squark and gluino pairs can be produced via both gluon fusion and
711 quark/anti-quark scattering.
- 712 • a gluino and squark produced together via quark-gluon scattering
- 713 • squark pairs produced via quark-quark scattering

714 Whilst most **SUSY** searches invoke the requirement of R-parity to explore parameter
715 phase space, there still exist a whole plethora of possible **SUSY** model topologies which
716 are still to be discovered at the **LHC**.

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

725 However recent results from the **LHC** now strongly disfavour large swathes of **CMSSM**
726 parameter space [39][40][41]. In the face of such results a more pragmatic model indepen-
727 dent search strategy is now applied across most **SUSY** searches at the **LHC**, see Section
728 (2.4.1).

729 As previously stated, a stable **LSP** that exhibits the properties of a dark matter candidate
730 would be weakly interacting and therefore will not be directly detected in a detector envi-
731 ronment. Additionally the cascade decays of supersymmetric particles to the **LSP** would
732 also result in significant hadronic activity. These signatures can then be characterised
733 through large amounts of hadronic jets (see Section (3.3.1)), leptons and a significant
734 amount of missing energy dependent upon the size of the mass splitting between the
735 **LSP** and the supersymmetric particle it has decayed from.

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

⁷⁴² 2.4.1. Simplified models

⁷⁴³ With such a variety of different ways for a **SUSY** signal to manifest itself, it is necessary
⁷⁴⁴ to be able to interpret experimental reach through the masses of gluinos and squarks
⁷⁴⁵ which can be excluded by experimental searches rather than on a model specific basis.

⁷⁴⁶ This is accomplished through **SMS** models, which are defined by a set of hypothetical
⁷⁴⁷ particles and a sequence of their production and decay [42][43]. In the **SMS** models
⁷⁴⁸ considered within this thesis, only the production process for the two primary particles
⁷⁴⁹ are considered. Each primary particle can undergo a direct or a cascade decay through
⁷⁵⁰ an intermediate new particle. At the end of each decay chain there remains a neutral,
⁷⁵¹ undetected **LSP** particle, denoted $\tilde{\chi}_{LSP}$ which can represent a neutralino or gravitino.
⁷⁵² Essentially it is easier to consider each **SMS** with branching ratios set to 100% The
⁷⁵³ masses of the primary particle and the **LSP** remain as free parameters, in which the
⁷⁵⁴ absolute value and relative difference between the primary and **LSP** particle alter the
⁷⁵⁵ kinematics of the event.

⁷⁵⁶ Different **SMS** models are denoted with a T-prefix, with a summary of the types interpreted
⁷⁵⁷ within this thesis listed below [44].

- ⁷⁵⁸ • **T1,T1xxxx**, models represent a simplified version of gluino pair production with
⁷⁵⁹ each gluino (superpartner to the gluon) undergoing a three-body decay to a quark-
⁷⁶⁰ antiquark pair and the **LSP** (i.e. $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{LSP}$). The resultant final state from this
⁷⁶¹ decay is typically 4 jets + \cancel{E}_T in the absence of initial/final state radiation and
⁷⁶² detector effects. xxxx denotes models in which the quarks are of a specific flavour,
⁷⁶³ typically t or b quark-antiquarks.

- ⁷⁶⁴ • **T2,T2xx**, models represent a simplified version of squark anti-squark production
⁷⁶⁵ with each squark undergoing a two-body decay into a light-flavour quark and **LSP**
⁷⁶⁶ (i.e. $\tilde{q} \rightarrow q\tilde{\chi}_{LSP}$). This results in final states with less jets than gluino mediated
⁷⁶⁷ production, typically 2 jets + \cancel{E}_T when again ignoring the effect of initial/final state
⁷⁶⁸ radiation and detector effects. xx models again represent decays in which both the
⁷⁶⁹ quark and the squark within the decay is of a specific flavour, typically \tilde{t}/t or \tilde{b}/b .

⁷⁷⁰ Models rich in b and t quarks are interpreted within this thesis as they remain of
⁷⁷¹ particular interest within “Natural **SUSY**” scenarios [45][46]. The largest contribution
⁷⁷² to the quadratic divergence in the Higgs mass parameter comes from a loop of top
⁷⁷³ quarks via the Yukawa coupling. Cancellation of these divergences can be achieved in

⁷⁷⁴ supersymmetric theories by requiring a light right handed top squark, \tilde{t}_R , and left-handed
⁷⁷⁵ double $SU(2)_L$ doublet containing top and bottom squarks, $(\tilde{\tilde{t}}_b)_L$ [47].

⁷⁷⁶ These theories therefore solve the hierarchy problem by predicting light \sim EWK scale
⁷⁷⁷ third generation sleptons, to be accessible at the LHC. Search strategies involving the
⁷⁷⁸ requirement of b-tagging (see Section (3.3.2)) are used to give sensitivity to these type of
⁷⁷⁹ SUSY scenarios and are discussed in greater detail within Chapter 4.

⁷⁸⁰ Two example decay chains are shown in Figure 2.2; the pair production of gluinos (T1)
⁷⁸¹ and the pair production of squarks (T2) decaying into SM particles and LSP's.

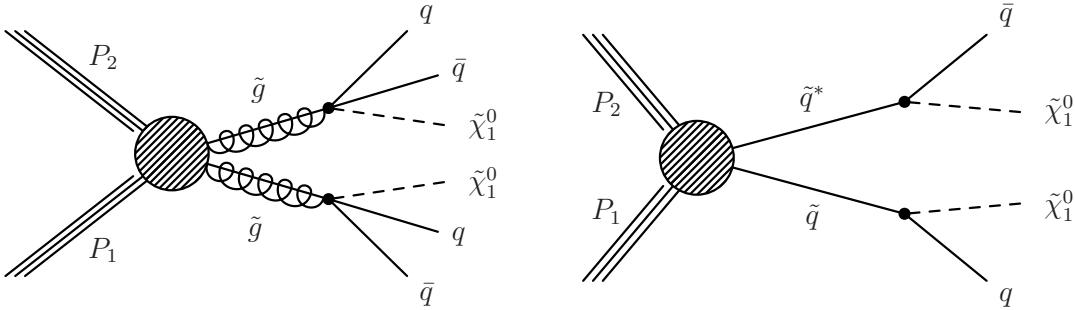


Figure 2.2.: Two example SMS model decays (T1 (left), T2 (right)), which are used in interpretations of physics reach by CMS.

Chapter 3.

⁷⁸² The LHC And The CMS Detector

⁷⁸³ Probing the SM for signs of new physics would not be possible without the immensely
⁷⁸⁴ complex electronics and machinery that makes the TeV energy scale accessible for the
⁷⁸⁵ first time. This chapter will describe both the LHC based at European Organization
⁷⁸⁶ for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS) detector, being
⁷⁸⁷ the experiment the author is a member of. Section (3.2) serves to introduce an overview
⁷⁸⁸ of the different components of the CMS detector, with specific components relevant to
⁷⁸⁹ the search for supersymmetric particles described in greater detail. Section (3.3) will
⁷⁹⁰ focus on event and object reconstruction again with more emphasis on jet level quantities
⁷⁹¹ which are most relevant to the author's analysis research. Finally Section (3.4) will
⁷⁹² cover work performed by the author, as service to the CMS Collaboration, in measuring
⁷⁹³ the performance of the Global Calorimeter Trigger (GCT) component of the L1 trigger
⁷⁹⁴ during the 2012-2013 run period.

⁷⁹⁵ 3.1. The LHC

⁷⁹⁶ The LHC is a storage ring, accelerator, and collider of circulating beams of protons or
⁷⁹⁷ ions. Housed in the tunnel dug for the Large Electron-Positron collider (LEP), it is
⁷⁹⁸ approximately 27 km in circumference, 100 m underground, and straddles the border
⁷⁹⁹ between France and Switzerland outside of Geneva. It is currently the only collider
⁸⁰⁰ in operation that is able to study physics at the TeV scale. A double-ring circular
⁸⁰¹ synchrotron, it was designed to collide both proton-proton (pp) and heavy ion (PbPb)
⁸⁰² with a centre of mass energy $\sqrt{s} = 14$ TeV at a final design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$.

⁸⁰³

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.

815

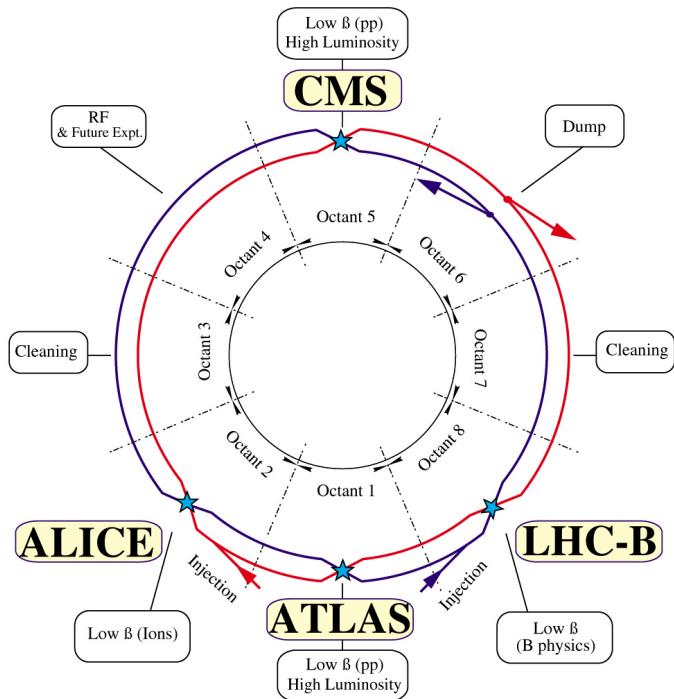


Figure 3.1.: A top down layout of the LHC. [52], with the position of the four main detectors labelled.

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 1.1×10^{11} protons localized into less than 1 ns in the direction of motion. Before collision the beams are ramped to 4

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

826

827 Colliding the beams produced an instantaneous luminosity of approximately 5×10^{33}
 828 $\text{cm}^{-2}\text{s}^{-1}$ during the 2012 run. The high number of protons in each bunch increases
 829 the likelihood of multiple interactions with each crossing of the counter-circulating
 830 beams. This leads to isotropic energy depositions within the detectors positioned at these
 831 interaction points, increasing the energy scale of the underlying event. This is known as
 832 pile-up and the counteracting of it's effects are important to the many measurements
 833 performed at the **LHC**.

834 In the early phase of prolonged operation after the initial shutdown the machine operated
 835 in 2010-2011 at 3.5 TeV per beam, $\sqrt{s} = 7$ TeV, delivering 6.13 fb^{-1} of data [53]. During
 836 the 2012-2013 run period, data was collected at an increased $\sqrt{s} = 8$ TeV improving the
 837 sensitivity of searches for new physics. Over the whole run period 23.3 fb^{-1} of data was
 838 delivered of which 21.8 fb^{-1} was recorded by the **CMS** detector as shown in Figure 3.2
 839 [53]. A total of 12 fb^{-1} of 8 TeV certified data was collected by October 2012, and it is
 840 this data which forms the basis of the results discussed within this thesis.

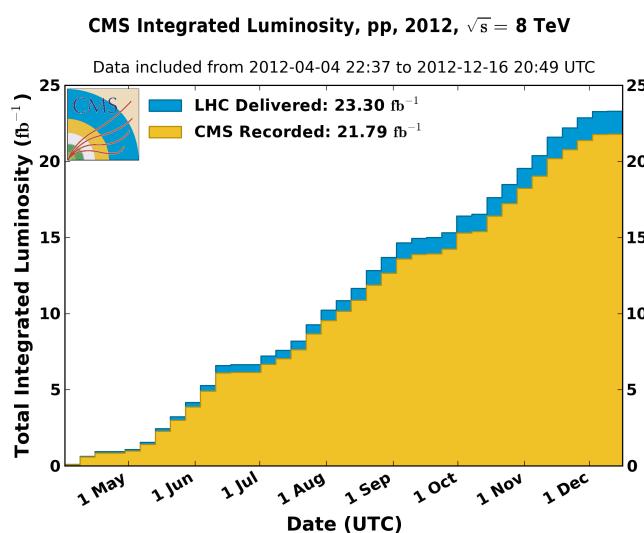


Figure 3.2.: The total integrated luminosity delivered to and collected by **CMS** during the 2012 8 TeV pp runs.

⁸⁴¹ 3.2. The CMS Detector

⁸⁴² The Compact Muon Solenoid (**CMS**) detector is one of two general purpose detectors
⁸⁴³ at the **LHC** designed to search for new physics. The detector is designed to provide
⁸⁴⁴ efficient identification and measurement of many physics objects including photons,
⁸⁴⁵ electrons, muons, taus, and hadronic showers over wide ranges of transverse momentum
⁸⁴⁶ and direction. Its nearly 4π coverage in solid angle allows for accurate measurement of
⁸⁴⁷ global transverse momentum imbalance. These design factors give **CMS** the ability to
⁸⁴⁸ search for direct production of **SUSY** particles at the TeV scale, making the search for
⁸⁴⁹ Supersymmetric particles one of the highest priorities among the wide range of physics
⁸⁵⁰ programmes at **CMS**.

⁸⁵¹

⁸⁵² **CMS** uses a right-handed Cartesian coordinate system with the origin at the interaction
⁸⁵³ point and the z-axis pointing along the beam axis, the x-axis points radially inwards to
⁸⁵⁴ the centre of the collider ring, with the y-axis points vertically upward. The azimuthal
⁸⁵⁵ angle, ϕ ranging between $[-\pi, \pi]$ is defined in the x-y plane starting from the x-axis. The
⁸⁵⁶ polar angle θ is measured from the z axis. The common convention in particle physics is
⁸⁵⁷ to express an out going particle in terms of ϕ and its pseudorapidity defined as

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

⁸⁵⁸ The variable $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is commonly used to define angular distance between
⁸⁵⁹ objects within the detector and additionally energy and momentum is typically measured
⁸⁶⁰ in the transverse plane perpendicular to the beam line. These values are calculated
⁸⁶¹ from the x and y components of the object and are denoted as $E_T = E \sin \theta$ and
⁸⁶² $p_T = \sqrt{p_x^2 + p_y^2}$.

⁸⁶³ 3.2.1. Detector subsystems

⁸⁶⁴ As the range of particles produced in pp collisions interact in different ways with mat-
⁸⁶⁵ ter, **CMS** is divided into subdetector systems, which perform complementary roles to
⁸⁶⁶ identify the identity, mass and momentum of the different physics objects present in
⁸⁶⁷ each event. These detector sub-systems contained within **CMS** are wrapped in layers

868 around a central 13 m long 4 T super conducting solenoid as shown in Figure 3.3. With
 869 the endcaps closed , CMS is a cylinder of length 22 m, diameter 15 m, and mass 12.5
 870 kilotons. A more detailed complete description of the detector can be found elsewhere [50].
 871

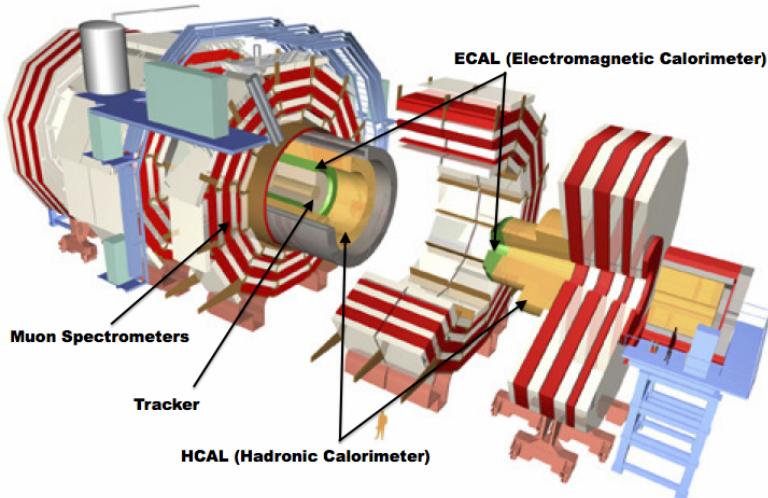


Figure 3.3.: A pictorial depiction of the CMS detector with the main detector subsystems labelled. [54]

872 3.2.2. Tracker

873 The inner-most subdetector of the barrel is the multi-layer silicon tracker, formed of a
 874 pixel detector component encased by layers of silicon strip detectors. The pixel detector
 875 consists of three layers of silicon pixel sensors providing measurements of the momentum,
 876 position coordinates of the charged particles as they pass, and the location of primary
 877 and secondary vertices between 4cm and 10cm transverse to the beam. Outside the pixel
 878 detector, ten cylindrical layers of silicon strip detectors extend the tracking system out to
 879 a radius of 1.20m from the beam line. The tracking system provides efficient and precise
 880 determination of the charges, momenta, and impact parameters of charged particles with
 881 the geometry of the tracker extending to cover a rapidity range up to $|\eta| < 2.5$.
 882

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

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

888 **3.2.3. Electromagnetic calorimeter**

889 Immediately outside of the tracker, but still within the magnet core, sits the Electromag-
890 netic CALorimeter (**ECAL**). Covering a pseudorapidity up to $|\eta| < 3$ and comprising
891 of over 75,000 PbWO₄ (lead tungstate) crystals that scintillate as particles deposit energy,
892 the **ECAL** provides high resolution measurements of the electromagnetic showers from
893 photons, electrons in the detector.

894

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

902

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

911

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

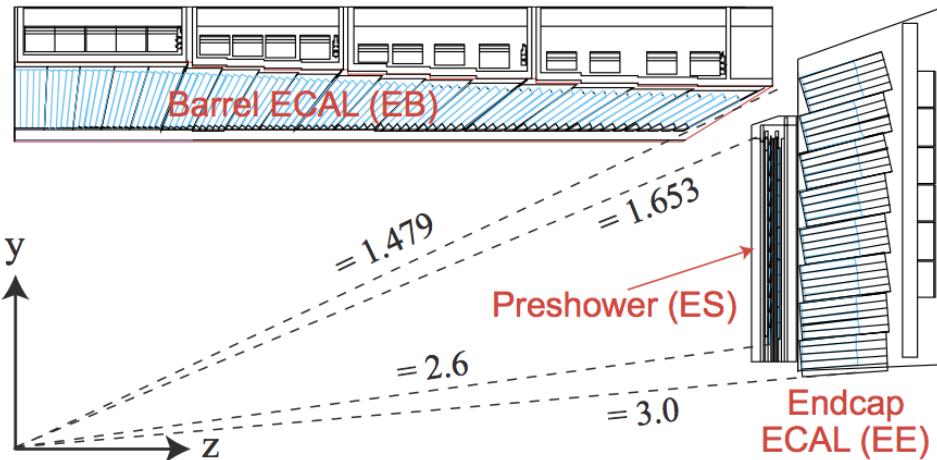


Figure 3.4.: Illustration of the CMS ECAL showing the arrangement of the lead tungstate crystals in the EB and EE. The ES is also shown and is located in front of the EE [55].

918 of temperature and so cooling systems continually maintain an overall constant ECAL
 919 temperature $\pm 0.05^\circ C$.

920 3.2.4. Hadronic calorimeter

921 Beyond the ECAL lies the Hadronic CALorimeter (HCAL) which is responsible for
 922 the accurate measurement of hadronic showers, crucial for analyses involving jets or
 923 missing energy signatures. The HCAL is a sampling calorimeter which consists of al-
 924 ternating layers of brass absorber and plastic scintillator, except in the hadron forward
 925 ($3.0 < |\eta| < 5.0$) region in which steel absorbers and quartz fibre scintillators are used
 926 because of their increased radiation tolerance. Hadron showers are initiated in the
 927 absorber layers inducing scintillation in the plastic scintillator tiles. These scintillation
 928 photons are converted by wavelength shifting fibres for read-out by hybrid photodiodes.
 929

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

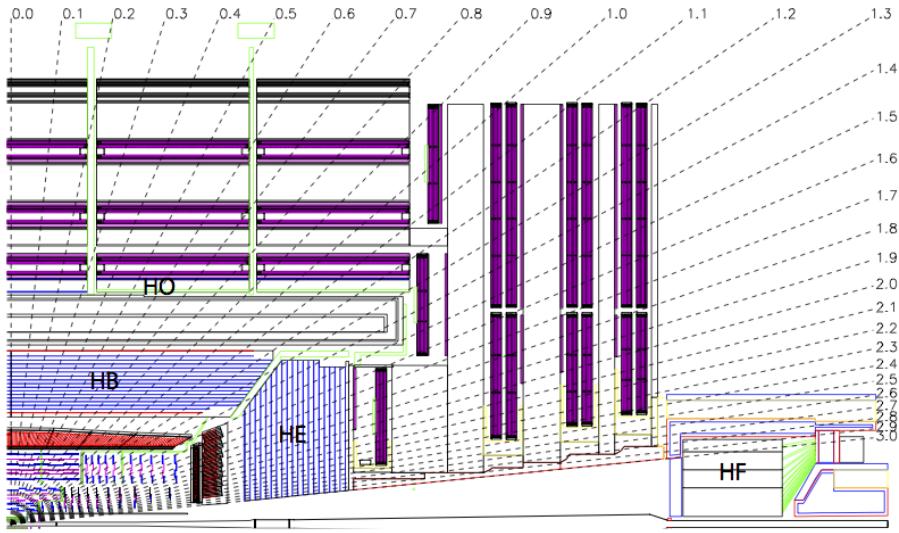


Figure 3.5.: Schematic of the hadron calorimeters in the r-z plane, showing the locations of the **HCAL** components and the **HF**. [50].

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

943 The relatively short number of interaction lengths (λ_l , the distance a hadron will travel
 944 through the absorber material before it has lost $\frac{1}{e}$ of its energy) within the **HB**, the lowest
 945 being $\lambda_l = 5.82$ for $|\eta| = 0$, facilitates the need for the ‘tail catching’ **HO** to increase the
 946 sampling depth in the central barrel rapidity region $|\eta| < 1.3$ to 11 interaction lengths .
 947 Significant fractions of the hadrons energy will be deposited in the **ECAL** as it passed
 948 through the detector. Therefore measurements of hadron energies in the central regions
 949 $|\eta| < 3.0$ use both the **ECAL** and **HCAL** to reconstruct the true energy from showering
 950 hadrons.

951 **3.2.5. Muon systems**

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

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

962

963 **3.3. Event Reconstruction and Object Definition**

964 The goal of event reconstruction is to take the raw information recorded by the detector
965 and to compute from it higher-level quantities which can be used at an analysis level.
966 These typically correspond to an individual particle's energy and momenta, or groups of
967 particles which shower in a narrow cone and the overall global energy and momentum
968 balance of the event. The reconstruction of these objects are described in great detail in
969 [57], however covered below are brief descriptions of those which are most relevant to the
970 analysis detailed in Chapter 4.

971 **3.3.1. Jets**

972 Quarks and gluons are produced copiously at the LHC in the hard scattering of partons.
973 As these quarks and gluons fragment, they hadronise and decay into a group of strongly
974 interactive particles and their decay products. These streams of particles travel in the
975 same direction, as they have been "boosted" by the momentum of the primary hadron.
976 These collections of decay products are reconstructed and identified together as a "jet".

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

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

983

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

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

999 The jets reconstructed by the clustering algorithm in **CMS** typically have an energy
1000 that differs to the ‘true’ energy measured by a perfect detector. This stems from the
1001 non-linear and nonuniform response of the calorimeters as well as other residual effects
1002 including pile-up and underlying events, and therefore additional corrections are applied
1003 to recover a uniform relative response as a function of pseudo-rapidity. These are applied
1004 as separate sub corrections [62].

- 1005 • A PU correction is first applied to the jet. It subtracts the average extra energy
1006 deposited in the jet that comes from other vertices present in the event and is
1007 therefore not part of the hard jet itself.
- 1008 • p_T and η dependant corrections derived from Monte Carlo simulations are used to
1009 account for the non-uniform response of the detector.
- 1010 • p_T and η residual corrections are applied to data only to correct for difference
1011 between data and Monte Carlo. The residual is derived from QCD dijet samples
1012 and the p_T residual from $\gamma+$ jet and $Z+$ jets samples in data.

1013 **3.3.2. B-tagging**

1014 The decays of b quarks are suppressed by small CKM matrix elements. As a result, the
1015 lifetimes of b-flavoured hadrons, produced in the fragmentation of b quarks, are relatively
1016 long; \mathcal{O} 1ps. The identification of jets originating from b quarks is very important for
1017 searches for new physics and for measurements of standard model processes.

1018

1019 Many different algorithms developed by CMS select b-quark jets based on variables such
1020 as the impact parameters of the charged-particle tracks, the properties of reconstructed
1021 decay vertices, and the presence or absence of a lepton, or combinations thereof [63]. One
1022 of the most efficient of which is the Combined Secondary Vertex (CSV) which operates
1023 based on secondary vertex and track-based lifetime information, benchmarked in ‘Loose’,
1024 ‘Medium’ and ‘Tight’ working points, of which the medium point is the tagger used
1025 within the α_T search detailed in Section (4.1).

1026 Using the CSV tagger, a likelihood-based discriminator distinguishes between jets from
1027 b-quarks, and those from charm or light quarks and gluons, which is shown in Figure 3.6.
1028 The minimum thresholds on the discriminator for each working point correspond to the
1029 misidentification probability for light-parton jets of 10%, 1%, and 0.1%, respectively, in
1030 jets with an average p_T of about 80 GeV.

1031 The b-tagging performance is evaluated to measure the b-jet tagging efficiency ϵ_b , and the
1032 misidentification probability of charm ϵ_c and light-parton jets ϵ_s . The tagging efficiencies
1033 for each of these three jet flavours are compared between data and MC simulation, from
1034 which a series of p_T and $|\eta|$ binned jet corrections are determined,

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

1035 These are collectively named ‘Btag Scale Factors’ and allow MC simulation to accu-
1036 rately reflect the running conditions and performance of the tagging algorithm in data.
1037 Understanding of the b-tagging efficiency is essential in order to minimise systematic
1038 uncertainties in physics analyses that employ b-tagging.

1039

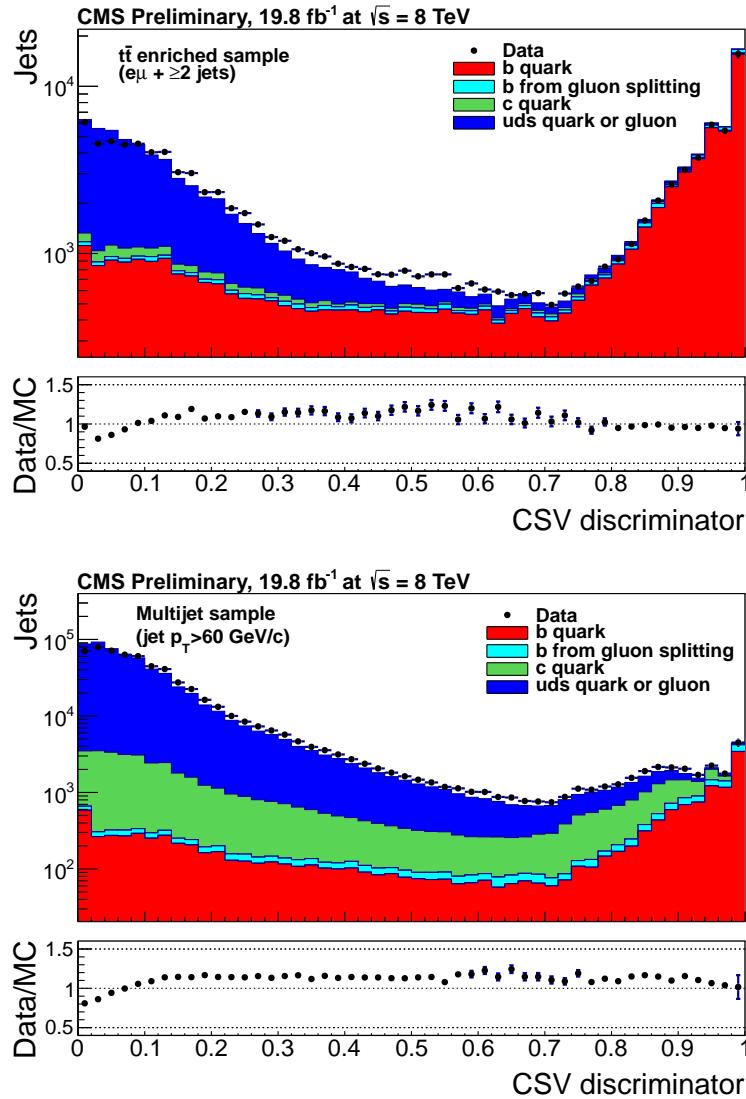


Figure 3.6.: CSV algorithm discriminator values in enriched ttbar (top) and inclusive multi jet samples (bottom) for b,c and light flavoured jets [64]. Working points are determined from the misidentification probability for light-parton jets to be tagged as a b-jet and are given as 0.244, 0.679 and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

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

1047 Additionally the performance of the tagger can also be benchmarked in $t\bar{t}$ events where
1048 in the SM, the top quark is expected to decay to a W boson and a b quark about 99.8%
1049 of the time [1]. Further selection criteria is applied to these events to further enrich the
1050 b quark content of these events. The methods to identify b-jets in data are discussed
1051 in great detail at [65]. The jet flavours are determined in simulation using truth level
1052 information and are compared to data to determine the correction scale factors (SF_b),
1053 which are displayed for the CSVM tagger in Figure 3.7.

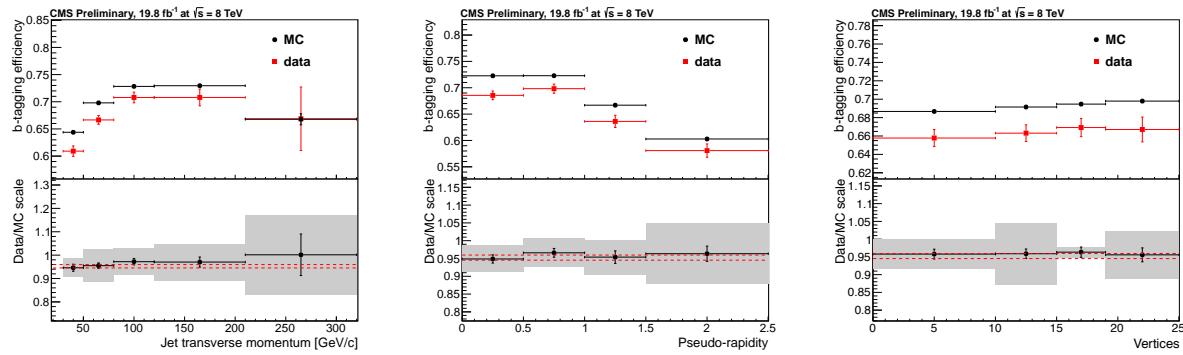


Figure 3.7.: Measured in $t\bar{t} \rightarrow \text{di-lepton}$ events using the CSVM tagger: (upper panels) b-tagging efficiencies and (lower panels) data/MC scale factor SF_b as a function of (left) jet p_T , (middle) jet $|\eta|$ and (right) number of primary vertices. In the lower panels, the grey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

1054 The measurement of the misidentification probability for light-parton jets relies on the
1055 inversion of tagging algorithms, selecting non-b jets using the same variables and tech-
1056 niques used in benchmarking the b-tagging efficiency. The scale factors (SF_s) to be
1057 applied to MC are shown in Figure 3.8 for the CSVM tagger.

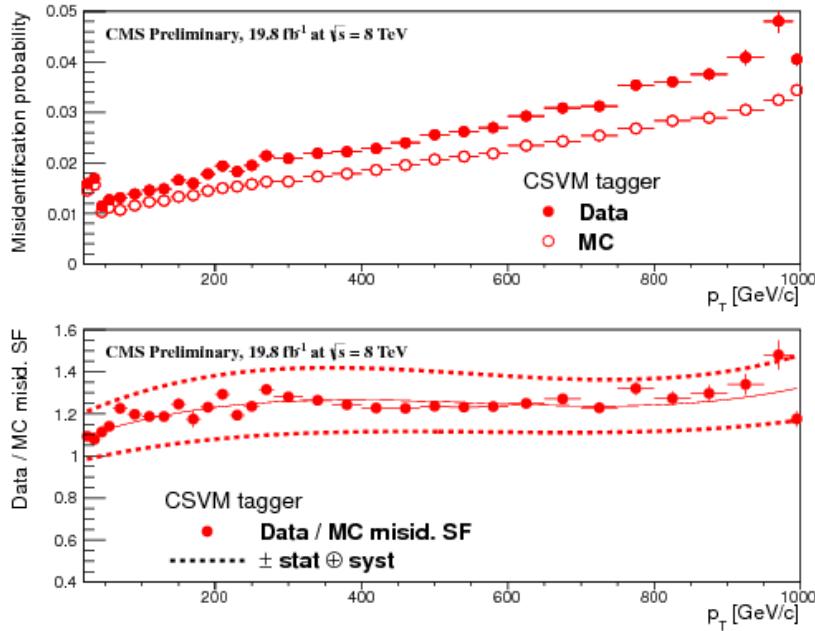


Figure 3.8.: For the **CSVM** tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last p_T bin in each plot includes all jets with $p_T > 1000 \text{ GeV}$. The solid curve is the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

1058 3.4. Triggering System

1059 With bunch crossings separated by just 25 ns, the rate at which data from all collisions
1060 would have to be written out and processed would be unfeasible. A two-tiered triggering
1061 system is applied at **CMS** in order to cope with the high collision rate of protons. The
1062 **CMS** trigger is designed to use limited information from each event to determine whether
1063 to record the event, reducing the rate of data taking to manageable levels whilst ensuring
1064 a high efficiency of interesting physics object events are selected.

1065 The **L1** is a pipelined, dead-timeless system based on custom-built electronics [66], and is
1066 a combination of several sub systems which is shown pictorially in Figure 3.9. The L1
1067 system is covered in more detail within the following section along with a description
1068 of the service work undertaken by the author to benchmark the performance of the L1
1069 calorimeter trigger during the 2012 8 TeV run period.

1070 The Higher Level Trigger (**HLT**) is a large farm of commercial computers [67]. The **HLT**
1071 processes events with software reconstruction algorithms that are more detailed, giving
1072 performance more similar to the reconstruction used offline. The **HLT** reduces the event

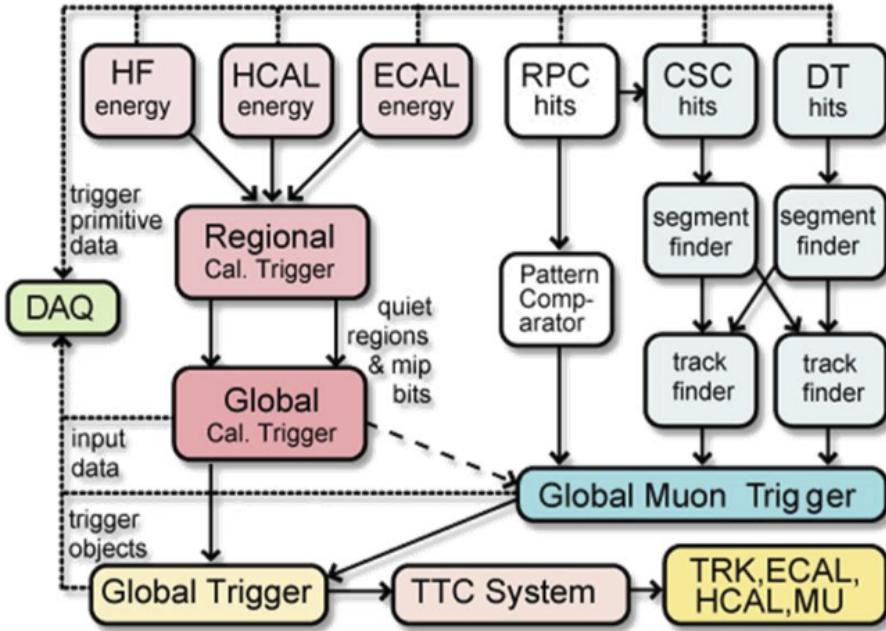


Figure 3.9.: The CMS L1 Trigger system.

1073 rate written to disk by a factor of ~ 500 ($\sim 200\text{Hz}$). The recorded events are transferred
 1074 from CMS to the CERN computing centre, where event reconstruction is performed, and
 1075 then distributed to CMS computing sites around the globe for storage and analysis.

1076 3.4.1. The level-1 trigger

1077 The L1 trigger reduces the rate of events collected from 40 MHz to $\sim 100\text{ kHz}$ using
 1078 information from the calorimeters and muon chambers, but not the tracker. A tree
 1079 system of triggers is used to decide whether to pass on an event to the HLT for further
 1080 reconstruction. Firstly the calorimeter and muon event information is kept separate,
 1081 with local reconstruction of objects ($\mu, e, \gamma, \text{jets}$) performed by the Regional Calorimeter
 1082 Trigger (RCT) and Regional Muon Trigger (RMT) respectively. The RCT generates up to
 1083 72 isolated and non-isolated electromagnetic objects. These are sorted by rank, which is
 1084 equivalent to transverse energy E_T , with the four highest ranked electromagnetic objects
 1085 being passed via the Global Calorimeter Trigger (GCT) and Global MuonTrigger (GMT)
 1086 to the Global Trigger (GT).

1087 In the L1 **GCT**, coarse measurements of the energy deposited in the electromagnetic and
1088 hadronic calorimeters are combined and by using sophisticated algorithms the following
1089 physics objects are formed:

- 1090 • isolated and non-isolated electromagnetic objects (e and γ);
- 1091 • hadronic jets in the central and forward sections of the hadronic calorimeters;
- 1092 • hadronically decaying tau leptons;
- 1093 • total transverse energy (E_T), the scalar sum of the energy measured at L1, and
1094 missing transverse energy (\cancel{E}_T), defined as the vector sum of the energy of L1
1095 objects;
- 1096 • total transverse jet energy (H_T), the scalar sum of the energy of all L1 jet objects,
1097 and missing transverse jet energy (\cancel{H}_T), defined as the vector sum of the energy of
1098 L1 jets, are calculated from uncorrected L1 jets.

1099 In addition quantities suitable for triggering minimum bias events, forward physics and
1100 beam background events are calculated. Additionally relevant muon isolation information
1101 is also passed on to the **GMT** for decisions involving the muon triggers where it is
1102 combined with information from across the three muon sub-systems. The resultant final
1103 accept/reject decision at **L1** is then performed by the **GT** based on the objects received
1104 from the **GCT** and **GMT** (e/γ , μ , jets, E_T , \cancel{E}_T , H_T).

1105 The L1 trigger is therefore of upmost importance to the functioning of the detector.
1106 Without a high-performing trigger and a good understanding of its performance, there
1107 would be no data to analyse. Observations of how the L1 trigger performance is affected
1108 by changing **LHC** running conditions over the 2012 run period and also the introduction
1109 of a jet seed threshold to the L1 jet trigger algorithm is presented in the following Sections
1110 (3.4.2 - 3.4.6).

1111 3.4.2. The L1 trigger jet algorithm

1112 The L1 jet trigger uses the transverse energy sums computed in the calorimeter (both
1113 hadronic and electromagnetic) trigger regions. Each region consists of 4×4 trigger tower
1114 windows, spanning a region of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in pseudorapidity-azimuth. The
1115 jet trigger uses a 3×3 calorimeter region (112 trigger towers) sliding window technique
1116 which spans the full (η, ϕ) coverage of the **CMS** calorimeter as shown in Figure 3.10.

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

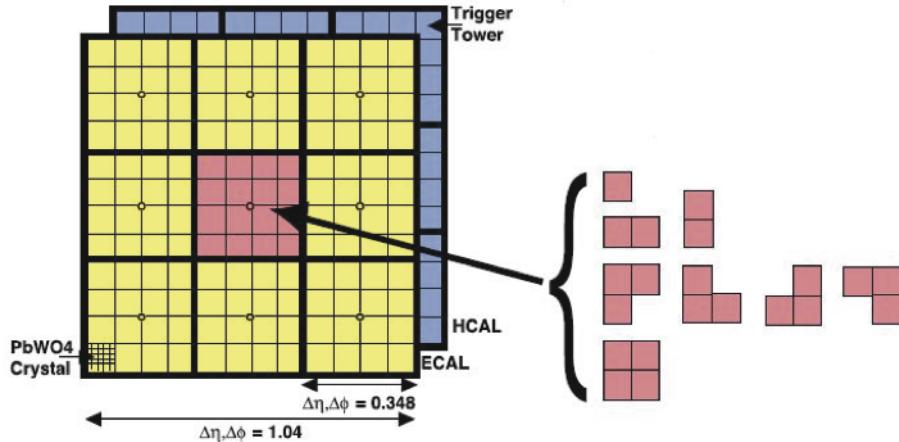


Figure 3.10.: Illustration of the Level-1 jet finding algorithm.

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

¹¹³⁵ 3.4.3. Measuring L1 jet trigger efficiencies

¹¹³⁶ The L1 jet efficiency is defined as the fraction of leading offline jets which were matched
¹¹³⁷ with a L1 tau or central jet above a certain trigger threshold, divided by all the leading
¹¹³⁸ offline jets in the event. This quantity is then plotted as a function of the offline jet E_T ,
¹¹³⁹ η and ϕ .

¹¹⁴⁰ The efficiency is determined by matching the L1 and reconstructed offline jets spatially
¹¹⁴¹ in $\eta - \phi$ space. This is done by calculating the minimum separation in ΔR between the
¹¹⁴² highest offline reconstructed jet in E_T ($E_T > 10$ GeV, $|\eta| < 3$) and any L1 jet. A jet will
¹¹⁴³ be matched if this value is found to be < 0.5 . Should more than one jet satisfy this, the
¹¹⁴⁴ jet closest in ΔR is taken as the matched jet. The matching efficiency is close to 100%,
¹¹⁴⁵ above 30(45) GeV for run 2012B(C) data (see Appendix B.1).

¹¹⁴⁶ Each efficiency curve is fitted with a function which is the cumulative distribution function
¹¹⁴⁷ 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:

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

¹¹⁴⁹ In this functional form, the parameter μ determines the point of 50% of the plateau
¹¹⁵⁰ efficiency and the σ gives the resolution. This parametrisation is used to benchmark
¹¹⁵¹ the efficiency at the plateau, the turn-on points and resolution for each L1 Jet trigger.
¹¹⁵² The choice of function is purely empirical. Previous studies used the error function
¹¹⁵³ alone, which described the data well at high threshold values but could not describe the
¹¹⁵⁴ efficiencies well at lower thresholds [68].

¹¹⁵⁵ The efficiency turn-on curves for various L1 jet thresholds are evaluated as a function of
¹¹⁵⁶ the offline reconstructed jet E_T for central jets with $|\eta| < 3$. These are measured using
¹¹⁵⁷ single isolated μ triggers which have high statistics, and are orthogonal and therefore
¹¹⁵⁸ unbiased to the hadronic triggers under study. The efficiency is calculated with respect to
¹¹⁵⁹ offline Calo and PF Jets in Figure 3.11. Table 3.1 shows the values of these parameters,
¹¹⁶⁰ calculated for three example L1 single jet triggers taken from 2012 8 TeV data.

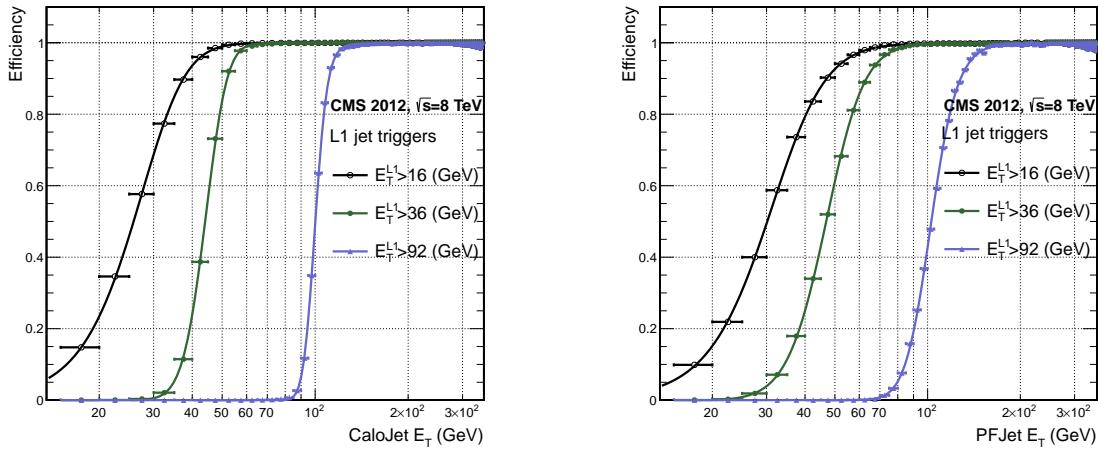


Figure 3.11.: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T (left) and PFJet E_T (right), measured in 2012 Run Period C data and collected with an isolated single μ data sample.

Trigger	Calo		PF	
	μ	σ	μ	σ
L1_SingleJet16	21.09 ± 0.03	7.01 ± 0.02	22.17 ± 0.04	7.83 ± 0.03
L1_SingleJet36	41.15 ± 0.05	5.11 ± 0.02	39.16 ± 0.06	8.04 ± 0.03
L1_SingleJet92	95.36 ± 0.13	5.62 ± 0.03	90.85 ± 0.19	11.30 ± 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 μ data sample. The turn-on point, μ , and resolution, σ , 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

1171 counteract the effects of high pile up running conditions which create a large number of
1172 soft non-collimated jets, that are then added to the jets from the primary interaction or
1173 other soft jets from other secondary interactions [69]. This in turn causes a large increase
1174 in trigger rate due to the increase in the likelihood that the event causes the L1 trigger to
1175 fire. This was implemented to maintain trigger thresholds by cutting the rate of events
1176 recorded without significant reduction in the efficiency of physics events of interest.

1177 The effect of the introduction of this jet seed threshold between these two run periods is
1178 benchmarked through a comparison of the efficiency of the L1 jet triggers with respect
1179 to offline Calo jets shown in Figure 3.12, and the L1 H_T trigger efficiency in Figure 3.14
1180 which is compared to offline H_T constructed from Calo jets with $E_T \geq 40\text{GeV}$.

1181 To negate any effects from different pile-up conditions in the run periods, the efficiencies
1182 are measured in events which contain between 15 and 20 primary vertices as defined in
1183 Appendix (A.2).

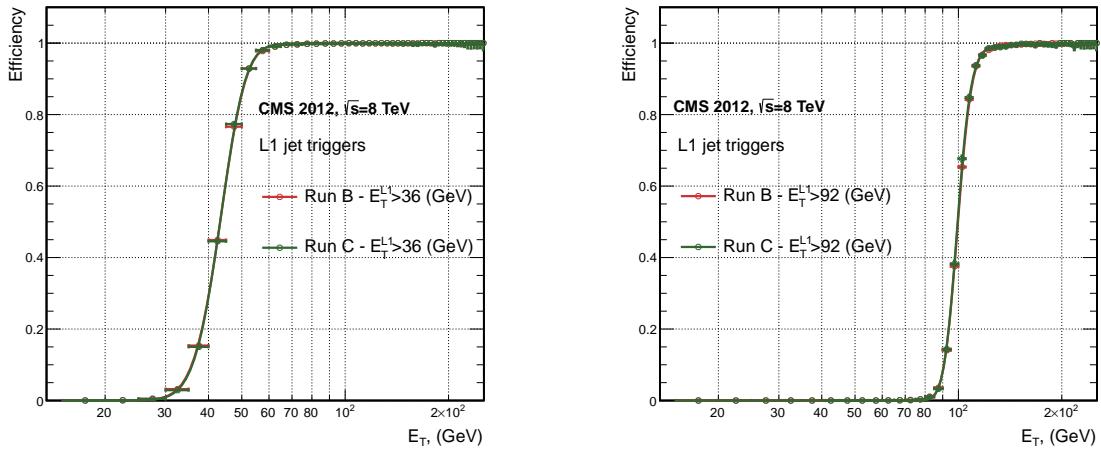


Figure 3.12.: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T , measured for the L1 SingleJet 36 and 92 trigger in 2012 run period B and C collected with an isolated single μ' sample.

1184 It can be seen that the performance of the $E_T > 36, 92$ single jet are almost identical,
1185 with the jet seed having no measurable effect on these triggers as shown in Table 3.2 .
1186 For the H_T triggers, a large increase in rate during high pile-up conditions is expected.
1187 This is due to the low energy threshold required for a jet to be added to the L1 H_T sum,
1188 which is compiled from all uncorrected L1 jets formed in the RCT. The introduction
1189 of the jet seed threshold removes the creation of many of these soft low E_T jets, thus

Trigger	2012B		2012C	
	μ	σ	μ	σ
L1_SingleJet36	40.29 ± 0.04	5.34 ± 0.02	40.29 ± 0.11	5.21 ± 0.05
L1_SingleJet92	94.99 ± 0.09	5.93 ± 0.06	94.82 ± 0.29	5.74 ± 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 μ and resolution σ of the L1 jet triggers are measured with respect to offline Calo Jets in run B (left) and run C (right).

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

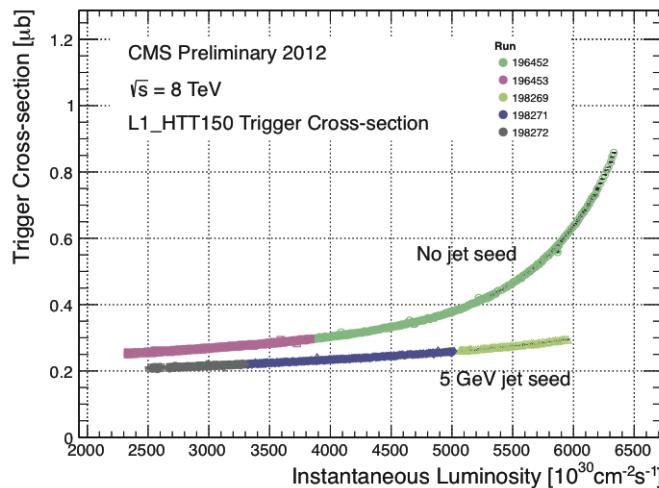


Figure 3.13.: Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1 H_T triggers [70].

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

3.4.5. Robustness of L1 jet performance against pile-up

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

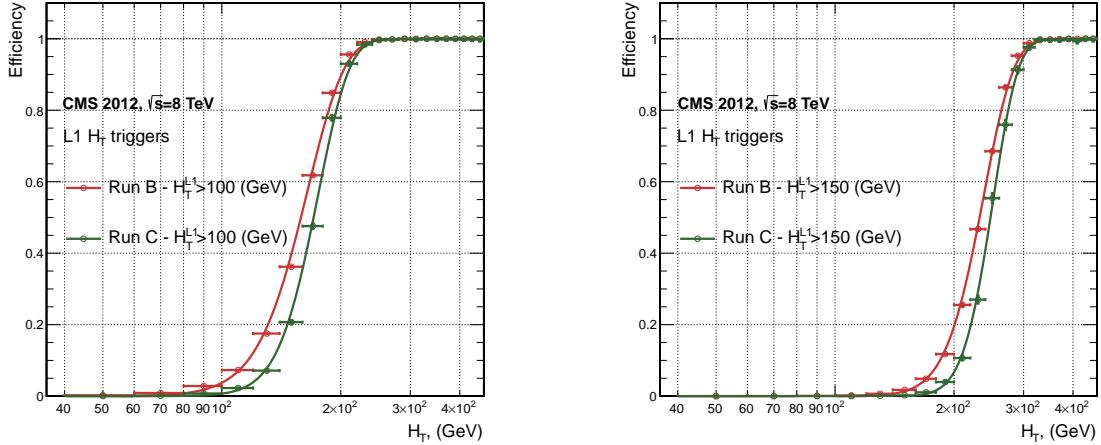


Figure 3.14.: L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T , measured for the L1 H_T 100 and 150 trigger during the run 2012 B and C collected using an isolated single μ triggered sample.

2012B			2012C		
Trigger	μ	σ	μ	σ	
L1 HT-100	157.5 ± 0.08	32.9 ± 0.08	169.8 ± 0.08	28.7 ± 0.03	
L1 H1-150	230.9 ± 0.02	37.3 ± 0.01	246.4 ± 0.16	31.8 ± 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 μ trigger. The turn-on point μ , resolution σ 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	
	μ	σ	μ	σ	μ	σ
L1_SingleJet16	19.9 ± 0.1	6.1 ± 0.3	20.8 ± 0.1	6.5 ± 0.1	22.3 ± 0.2	7.5 ± 0.1
L1_SingleJet36	41.8 ± 0.1	4.6 ± 0.1	40.9 ± 0.1	5.1 ± 0.1	40.6 ± 0.6	5.9 ± 0.2
L1_SingleJet92	95.9 ± 0.2	5.4 ± 0.1	95.2 ± 0.2	5.6 ± 0.1	94.5 ± 0.6	6.2 ± 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 μ triggered data. The turn-on point, μ , and resolution, σ , of the L1 jet triggers are measured with respect to offline Calo jets in low (left), medium (middle) and high (right) pile-up conditions.

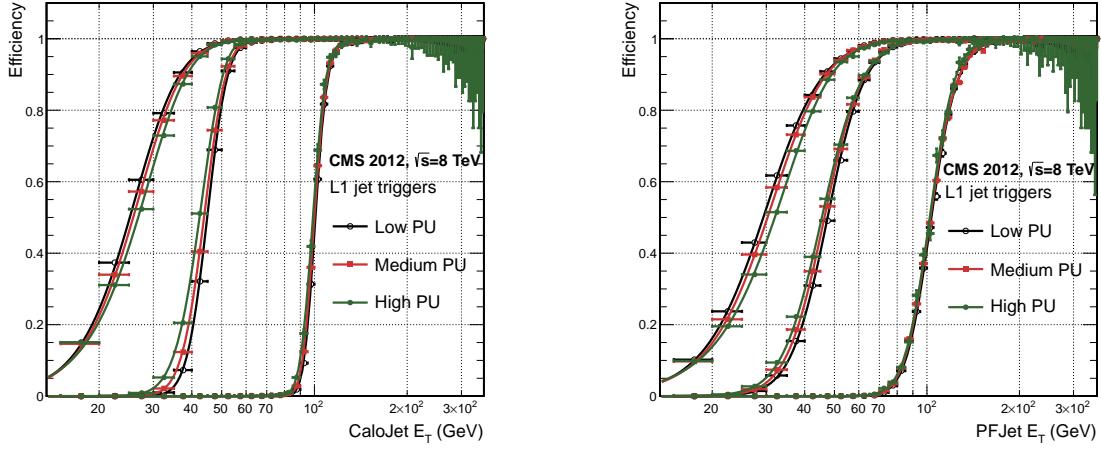


Figure 3.15.: L1 jet efficiency turn-on curves as a function of the leading offline E_T Calo (left) and PF (right) jet, for low, medium and high pile-up conditions.

Vertices	0-10		11-20		> 20	
	μ	σ	μ	σ	μ	σ
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 μ triggered data. The turn-on point, μ , and resolution, σ , of the L1 jet triggers are measured with respect to offline PF jets in low (left), medium (middle) and high (right) pile-up conditions.

1204 No significant drop in efficiency is observed in the presence of a high number of primary
 1205 vertices. The increase in hadronic activity in higher pile-up conditions, combined with
 1206 the absence of pile-up subtraction for L1 jets, results in the expected observation of
 1207 a decrease in the μ value of the efficiency turn-ons as a function of pile-up, while the
 1208 resolution, σ of the turn-ons are found to gradually worsen as expected with increasing
 1209 pile-up.

1210 These features are further emphasised when shown as a function of

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

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

The μ , σ and λ values extracted for the low, medium and high pile-up conditions are shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central value 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 is also observed to improve at higher offline jet E_T .

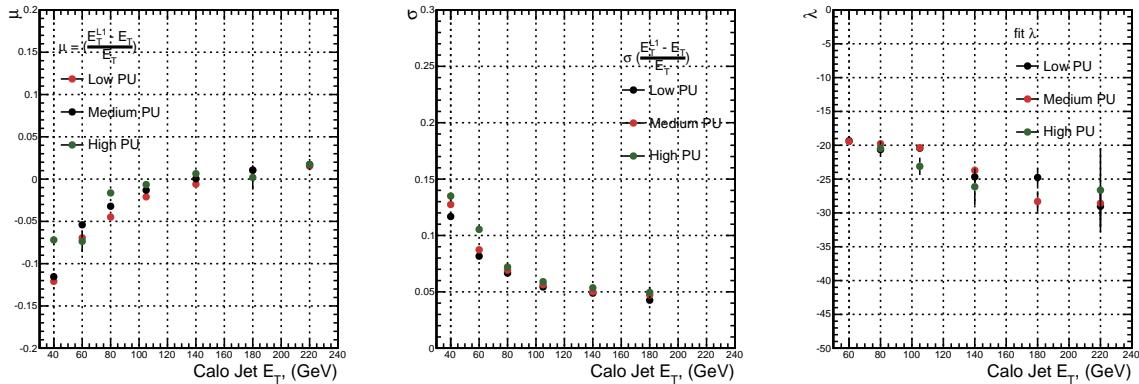


Figure 3.16.: Fit values from an EMG function fitted to the resolution plots of leading Calo jet E_T measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low, medium and high pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian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

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

3.4.6. Summary

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

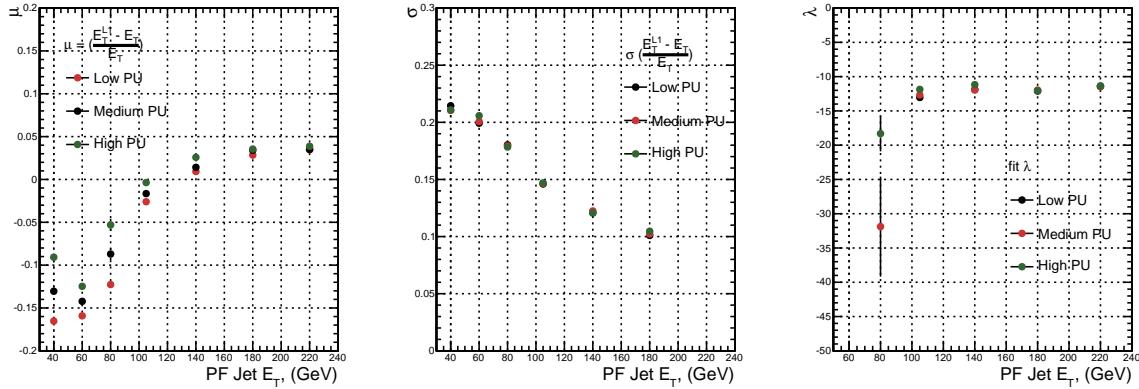


Figure 3.17.: Fit values from an **EMG** function fitted to the resolution plots of leading PF jet E_T measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low and medium pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian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

1229 change in performance is observed with this change and good performance is observed
 1230 for a range of L1 quantities.

Chapter 4.

1231 SUSY Searches In Hadronic Final States 1232

1233 In this chapter a model independent search for **SUSY** in hadronic final states with \cancel{E}_T 1234 using the α_T variable and b-quark multiplicity is introduced and described in detail. The 1235 results presented are based on a data sample of pp collisions collected in 2012 at $\sqrt{s} = 8$ 1236 TeV, corresponding to an integrate luminosity of $11.7 \pm 0.5 \text{ fb}^{-1}$.

1237 The kinematic variable α_T is motivated as a variable to provide strong rejections of QCD 1238 backgrounds, whilst maintaining sensitivity to a possible **SUSY** signal within Section 1239 (4.1). The search and trigger strategy in addition to the event reconstruction and selection 1240 are outlined within Sections (4.2-4.3).

1241 The method in which the **SM** background is estimated using an analytical technique to 1242 improve statistical precision at higher b-tag multiplicities is detailed within Section (4.5), 1243 with a discussion on the impact of b-tagging and mis-tagging scale factors between data 1244 and MC on any background predictions.

1245 A description of the formulation of appropriate systematic uncertainties applied to the 1246 background predictions to account for theoretical uncertainties and limitations in the 1247 simulation modelling of event kinematics and instrumental effects is covered in Section 1248 (4.6).

1249 Finally the statistical likelihood model to interpret the observations in the signal and 1250 control samples is described in Section (4.8). The experimental reach of the analysis 1251 discussed within this thesis is interpreted in two classes of **SMS** models, the topologies 1252 of which are detailed in Section (2.4.1). The **SMS** models considered in this analysis

1253 are summariesed in Table 4.1. For each model, the **LSP** is assumed to be the lightest
1254 neutralino.

1255 Within the table are also defined reference points, parameterised in terms of parent
1256 gluino/squark and **LSP** sparticle masses, m_{parent} and m_{LSP} , respectively, which are used
1257 within the following two chapters to demonstrate potential yields within the signal region
1258 of the search. The masses are chosen to reflect parameter space which is within the
1259 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

1260 4.1. An Introduction to the α_T Search

1261 The experimental signature of **SUSY** signal in the hadronic channel would manifest
1262 itself as a final state containing energetic jets and \cancel{E}_T . The search focuses on topologies
1263 where new heavy supersymmetric, R-parity conserving particles are pair-produced in pp
1264 collisions. These particles decaying to a **LSP** escape the detector undetected, leading to
1265 significant missing energy and missing hadronic transverse energy,

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

1266 defined as the vector sum of the transverse energies of jets selected in an event. Energetic
1267 jets produced in the decay of these supersymmetric particles also can produce significant
1268 visible transverse energy,

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

1269 defined as the scalar sum of the transverse energies of jets selected in an event.

1270 A search within this channel is greatly complicated in a hadron collider environment,
1271 where the overwhelming background comes from inherently balanced multi-jet (“QCD”)
1272 events which are produced with an extremely large cross section as demonstrated within
1273 Figure 4.1. \cancel{E}_T can appear in such events with a substantial mis-measurement of jet
1274 energy or missed objects due to detector miscalibration or noise effects.

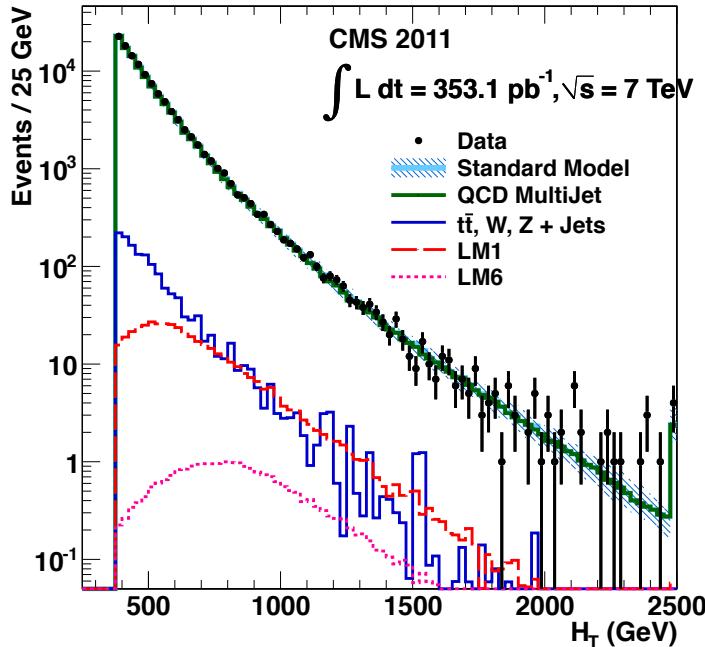


Figure 4.1.: Reconstructed offline H_T for 11.7fb^{-1} of data after a basic pre-selection. Sample is collected from prescaled H_T triggers. Overlaid are expectations from MC simulation of EWK processes as well as a reference signal model (labelled D2 from Table.4.1).

1275 Additional SM background from EWK processes with genuine \cancel{E}_T from escaping neutrinos
1276 comprise the irreducible background within this search and come mainly from:

1277 • $Z \rightarrow \nu\bar{\nu} + \text{jets}$,

- $W \rightarrow l\nu + \text{jets}$ in which a lepton falls outside of detector acceptance, or the lepton decays hadronically $\tau \rightarrow \text{had}$,
- $t\bar{t}$ with at least one leptonic W decay,
- small background contributions from DY, single top and Diboson (WW,ZZ,WZ) processes.

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

4.1.1. The α_T variable

For a perfectly measured di-jet QCD event, conservation laws dictate that they must be 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.

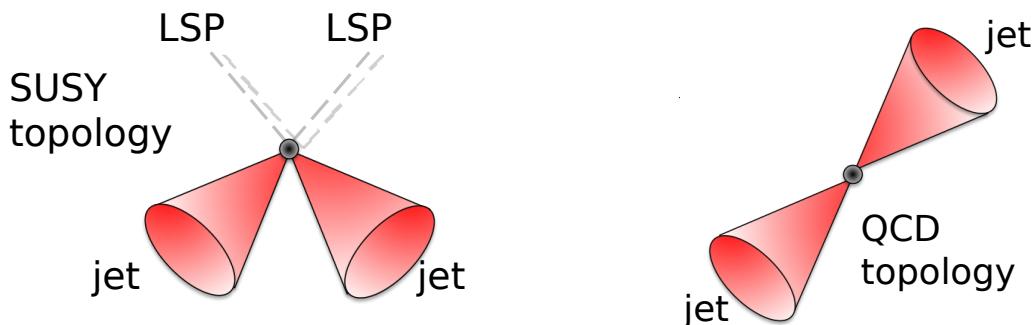


Figure 4.2.: The event topologies of background QCD di-jet events (right) and a generic SUSY signature with genuine \cancel{E}_T (left).

Exploiting this feature leads to the formulation of α_T in di-jet systems defined as,

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

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

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

1293 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
1294 with jets which are not back-to-back, for example in events in which a W or Z recoils off
1295 a system of jets, α_T can achieve values in excess of 0.5.

1296 α_T can be extended to apply to any arbitrary number of jets, undertaken by modelling a
1297 system of n jets as a di-jet system, through the formation of two pseudo-jets [73]. The two
1298 pseudo-jets are built by merging the jets present in the event such that the 2 pseudo-jets
1299 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
1300 between the two pseudo jets. Using Equation (4.4), α_T can be rewritten as,

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

1301 The distribution of α_T for the two jet categories used within this analysis, 2,3 and ≥ 4 jets,
1302 is shown in the Figure 4.3, demonstrating the ability of the α_T variable to discriminate
1303 between multi jet events and EWK processes with genuine \cancel{E}_T in the final state.

1304 The α_T requirement used within the search is chosen to be $\alpha_T > 0.55$ to ensure that
1305 the QCD multijet background is negligible even in the presence of moderate jet mis-
1306 measurement. There still remains other effects which can cause multijet events to
1307 artificially have a large α_T value, which are discussed in detail in Section (4.2.2).

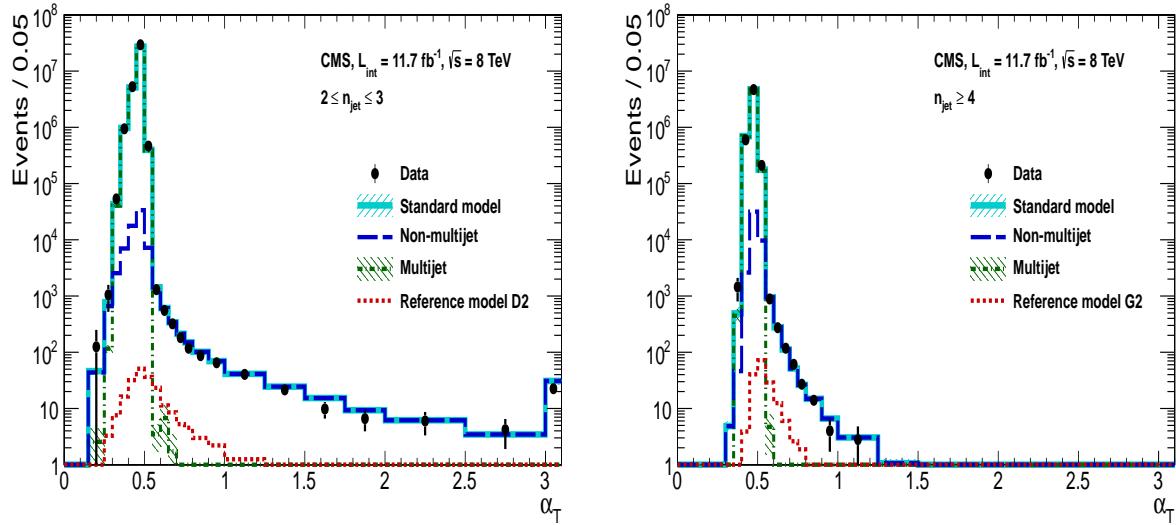


Figure 4.3.: The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet multiplicities after a full analysis selection and shown for $H_T > 375$. Data is collected using both prescaled H_T triggers and dedicated α_T triggers for below and above $\alpha_T = 0.55$ respectively. . Expected yields 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).

4.2. Search Strategy

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

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

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

1321 To remain inclusive to a large range of possible **SUSY** models, the signal region is binned
1322 in the following categories to allow for increased sensitivity in the interpretation of results
1323 for different **SUSY** topologies:

1324 **Sensitivity to a range of SUSY mass splittings**

1325 The hadronic signal region is defined by $H_T > 275$, divided into eight bins in H_T .

- 1326 – Two bins of width 50 GeV in the range $275 < H_T < 375$ GeV,
1327 – five bins of width 100 GeV in the range $375 < H_T < 875$ GeV,
1328 – and a final open bin, $H_T > 875$ GeV.

1329 The choice at low H_T is driven primarily by trigger constraints. The mass difference
1330 between the **LSP** and the particle that it decays from is an important factor in the
1331 amount of hadronic activity in the event.

1332 A large mass splitting will lead to hard high p_T jets which contribute to the H_T
1333 sum. From Figure 4.1 it can be seen that the **SM** background falls sharply at high
1334 H_T values, therefore a large number of H_T bins will lead to easier identification
1335 of such signals. Conversely smaller mass splittings lead to softer jet p_T 's which will
1336 subsequently fall into the lower H_T range.

1337 **Sensitivity to production method of SUSY particles**

1338 The production mechanism of any potential **SUSY** signal can lead to different event
1339 topologies. One such way to discriminate between gluino ($g\tilde{g}$ - “high multiplicity”),
1340 and direct squark ($q\tilde{q}$ - “low multiplicity”) induced production of **SUSY** particles is
1341 realised through the number of reconstructed jets in the final state.

1342 The analysis is thus split into two jet categories : 2-3 jets , ≥ 4 jets to give sensitivity
1343 to both of these mechanisms.

1344 **Sensitivity to “Natural SUSY” via tagging jets from b-quarks**

1345 Jets originating from bottom quarks (b-jets) are identified through vertices that
1346 are displaced with respect to the primary interaction. The algorithm used to tag
1347 b-jets is the Combined Secondary Vertex Medium Working Point (**CSV**) tagger,
1348 described within Section (3.3.2). A cut is placed on the discriminator variable of
1349 > 0.679 , leading to a gluon/light-quark mis-tag rate of 1% and a jet p_T dependant
1350 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, ≥ 4 b-tag categories . In the highest ≥ 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, ≥ 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.

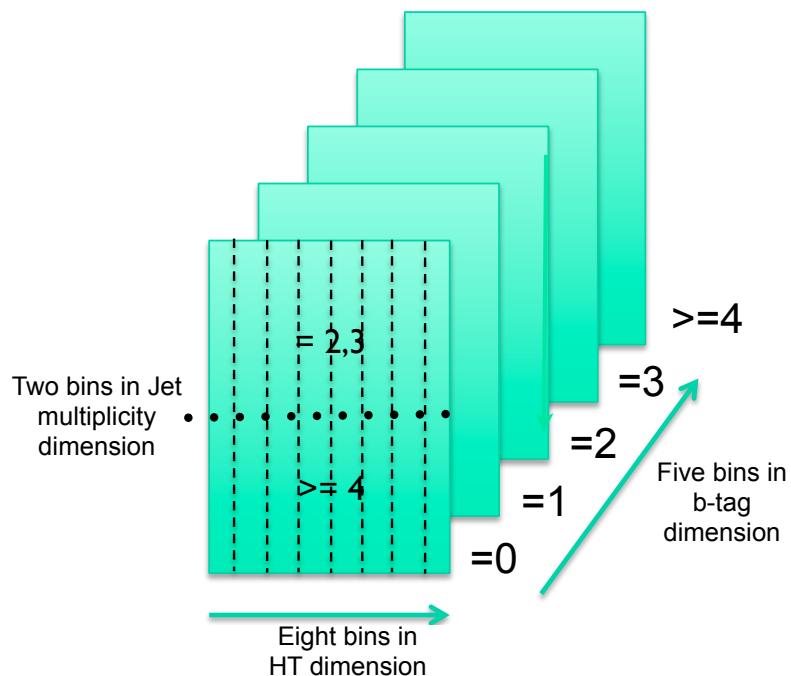


Figure 4.4.: Pictorial depiction of the analysis strategy employed by the α_T search to increase sensitivity to a wide spectra of SUSY models.

1367 4.2.1. Physics objects

1368 The physics objects used in the analysis defined below, follow the recommendation of
1369 the various CMS Physics Object Groups (POGs).

1370 • Jets

1371 The jets used in this analysis are CaloJets, reconstructed as described in Section
1372 ([3.3.1](#)) using the anti- k_T jet clustering algorithm.

1373 To ensure the jet object falls within the calorimeter systems a pseudo-rapidity
1374 requirement of $|\eta| < 3$ is applied. Each jet must pass a “loose” identification criteria
1375 to reject jets resulting from unphysical energy, the criteria of which are detailed in
1376 Table A.1 of Appendix A [[74](#)].

1377 • Muons

1378 Muons are selected in the $\mu +$ jets and $\mu\mu +$ jets control samples, and vetoed in
1379 the signal region. The same cut based identification criteria is applied to muons in
1380 both search regions and is summarised in Table 4.2 [[75](#)].

Variable	Definition
Is Global Muon	Muon contains both a hit in the muon chamber and a matched track in the inner tracking system [76].
$\chi^2 < 10$	χ^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 ΔR 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.

1381 Additionally muons are required to be within the acceptance of the muon tracking
 1382 systems. For the muon control samples, trigger requirements necessitate a $|\eta| <$
 1383 2.1 for the selection of muons. In the signal region where muons are vetoed these
 1384 conditions are relaxed to $|\eta| < 2.5$ and a minimum threshold of $p_T > 10$ GeV is
 1385 required of muon objects.

1386 • Photons

1387 Photons are selected within the $\gamma +$ jets control sample and vetoed in all other
 1388 selections. Photons are identified in both cases according to the cut based criteria
 1389 listed in Table 4.3 [77].

Variable	Definition
H/E < 0.05	The ratio of hadronic energy in the HCAL tower directly behind the ECAL super-cluster and the ECAL super-cluster itself.
$\sigma_{in\eta} < 0.011$	The log energy weighted width (σ), of the extent of the shower in the η dimension.
R9 < 1.0	The ratio of the energy of the 3×3 crystal core of the super-cluster compared to the total energy stored in the 5×5 super-cluster.
Combined Isolation < 6 GeV	The photons are required to be isolated with no electromagnetic or hadronic activity within a radius $\Delta R = 0.3$ of the photon object. A combination of the pileup subtracted [78], ECAL , HCAL 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.

1390 Photon objects are also required to have a minimum momentum of $p_T > 25$ GeV.

1391 • Electrons

1392 Electron identification is defined for veto purposes. They are selected according to
 1393 the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

1394 Electrons are required to be identified at $|\eta| < 2.5$, with a minimum $p_T > 10$ GeV
 1395 threshold to ensure that the electron falls within the tracking system of the detector.

1396 • Noise and E_T Filters

1397 A series of Noise filters are applied to veto events which contain spurious non-physical
 1398 jets that are not picked up by the jet id, and events which give large unphysical \cancel{E}_T
 1399 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 η 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

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

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

1413 At low H_T , the jet threshold requirements applied to be considered as part of the analysis
1414 and enter the H_T sum are scaled downwards. These are scaled down in order to not
1415 restrict phase space, preserving jet multiplicities and background admixture in the lower
1416 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.

1417 Within the signal region to suppress **SM** processes with genuine \cancel{E}_T from neutrinos,
1418 events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure
1419 multi-jet topology, events are vetoed if an isolated photon is found with $p_T > 25$ GeV.
1420 An α_T requirement of > 0.55 is required to reduce the QCD multi-jet background
1421 to a negligible amount. Finally additional cleaning cuts are applied to protect against
1422 pathological deficiencies such as reconstruction failures or severe energy mis-measurements
1423 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 α_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,

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

1424 • Fake \cancel{E}_T and \cancel{H}_T can arise due to significant jet mis-measurements cause by a small
1425 number of non-functioning **ECAL** regions. These regions absorb electromagnetic
1426 showers which are subsequently not added to the jet energy sum. To circumvent
1427 this problem the following procedure is employed : For each jet in the event, the
1428 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 ΔR 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 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.

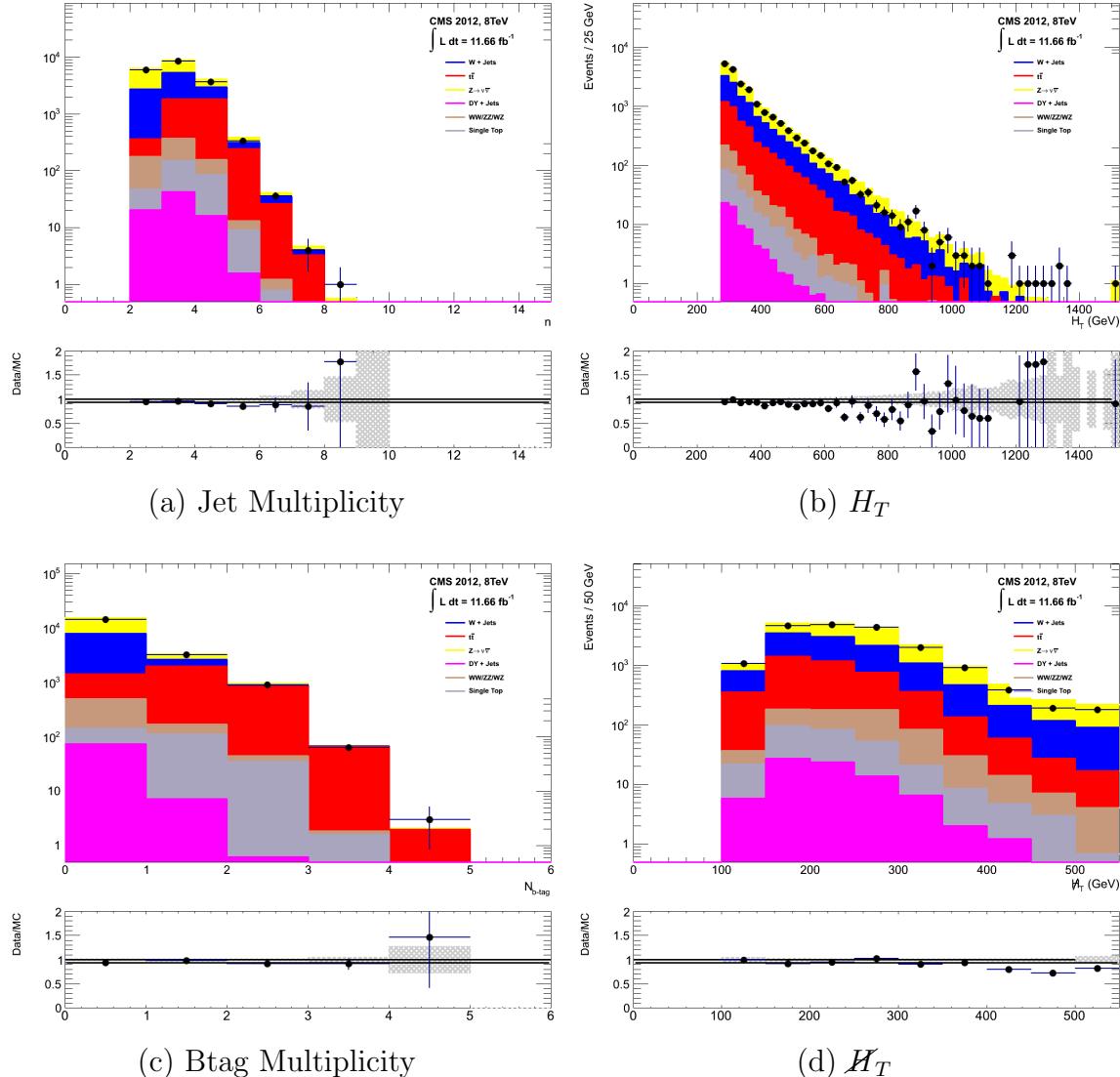


Figure 4.5.: Data/MC comparisons of key variables for the hadronic signal region, following the application of the hadronic selection criteria and the requirements of $H_T > 275$ GeV and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement 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}^{control} , and signal, N_{MC}^{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),

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

1451 All MC samples are normalised to the luminosity of the data samples, 11.7 fb^{-1} . Through
1452 this method, “vanilla” predictions for the **SM** background in the signal region can be
1453 made by considering separately the sum of the prediction from either the $\mu + \text{jets}$ and γ
1454 + jets or $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples. However the final background estimation from
1455 which results are interpreted, is calculated via a fitting procedure defined formally by
1456 the likelihood model described in Section (4.8).

1457 The sum of the expected yields from all MC processes, in each control sample enter the
1458 denominator, $N_{\text{MC}}^{\text{control}}$, of the **TF** defined in Eq (4.7). However for the numerator, $N_{\text{MC}}^{\text{signal}}$,
1459 only the relevant processes that the control sample is used in estimating a background
1460 for, enter into the **TF**.

1461 For the $\mu + \text{jets}$ sample the simulated MC processes which enter the numerator of the
1462 **TF** are,

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

1463 whilst for both the $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ samples the only MC process used in the
1464 numerator is,

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

1465 The control samples and the **EWK** processes they are specifically tuned to select are
1466 defined below, with distributions of key variables for each of the control samples shown
1467 for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed upon
1468 the number of b-tagged jets or jet multiplicity in the distributions shown. The MC
1469 distributions highlight the background compositions of each control sample, where in
1470 general, good agreement is observed between data and simulation, giving confidence

1471 that the samples are well understood. The contribution from QCD multi-jet events is
1472 expected to be negligible :

1473 **The $\mu +$ jets control sample**

1474 Events from $W +$ jets and $t\bar{t}$ processes enter into the hadronic signal sample due to
1475 unidentified leptons from acceptance or threshold effects and hadronic tau decays.
1476 These leptons originate from the decay of high p_T W bosons.

1477 The control samples specifically identifies $W \rightarrow \mu\bar{\nu}$ decays within the same phase-
1478 space of the signal region, where the muon is subsequently ignored in the calculation
1479 of event level variables, i.e. H_T , \mathcal{H}_T , α_T . All kinematic jet-based cuts are identical
1480 to those applied in the hadronic search region detailed in Section (4.2.2), with the
1481 same H_T , jet multiplicity and b-jet multiplicity binning described above.

- 1482 – Muons originating from W boson decays are selected by requiring one tightly
1483 isolated muon defined in Table 4.2, with a $p_T > 30$ GeV and $|\eta| < 2.1$. Both of
1484 these threshold arise from trigger restrictions.
- 1485 – The transverse mass of the W candidate must satisfy $M_T(\mu, \cancel{E}_T) < 30$ GeV (to
1486 suppress QCD multi-jet events).
- 1487 – Events which contain a jet overlapping with a muon $\Delta R(\mu, \text{jet}) < 0.5$ are vetoed
1488 to remove events from muons produced as part of a jet’s hadronisation process.
- 1489 – Events containing a second muon candidate which has failed id, but passed p_T
1490 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.

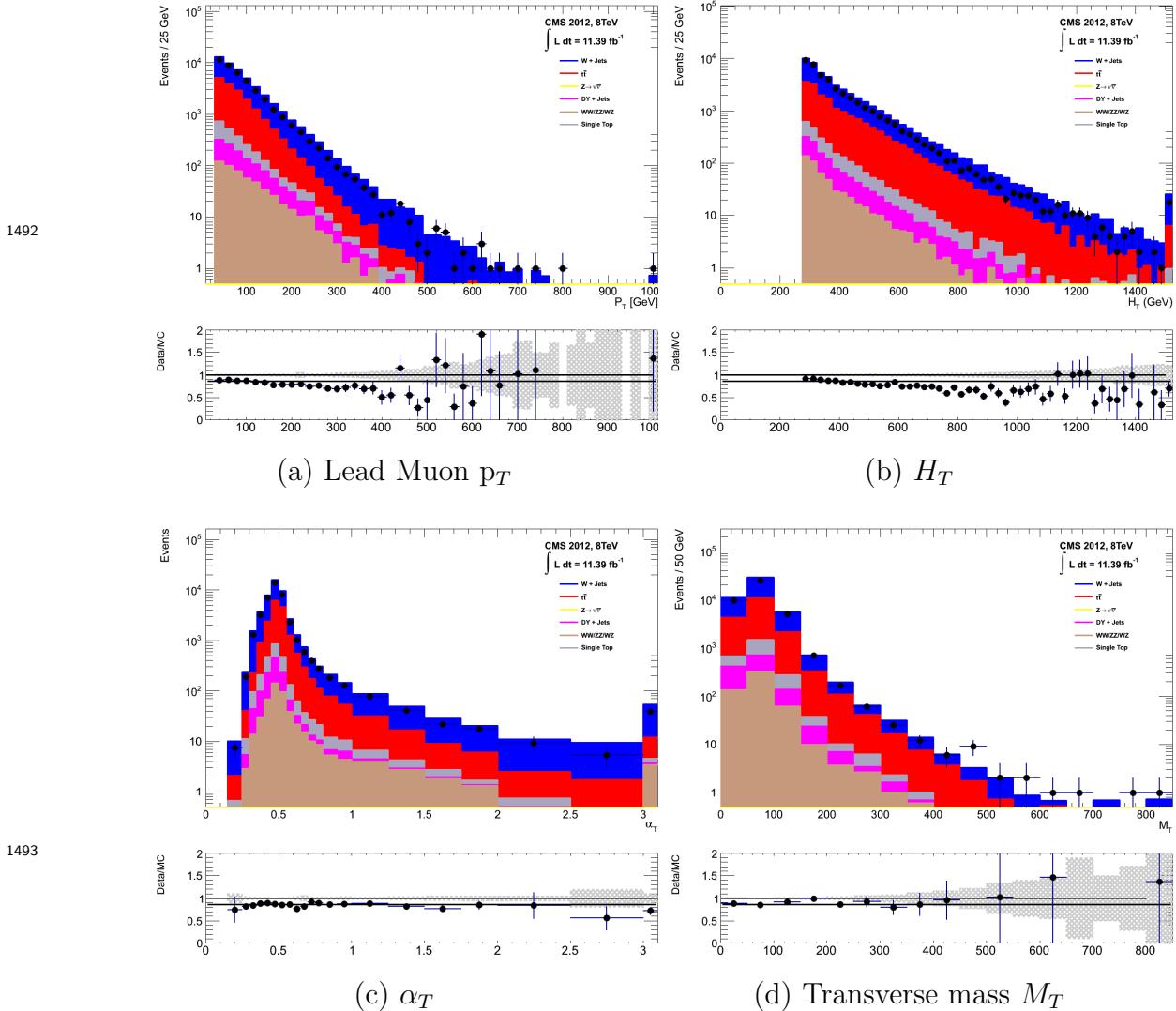


Figure 4.6.: Data/MC comparisons of key variables for the $\mu + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of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

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

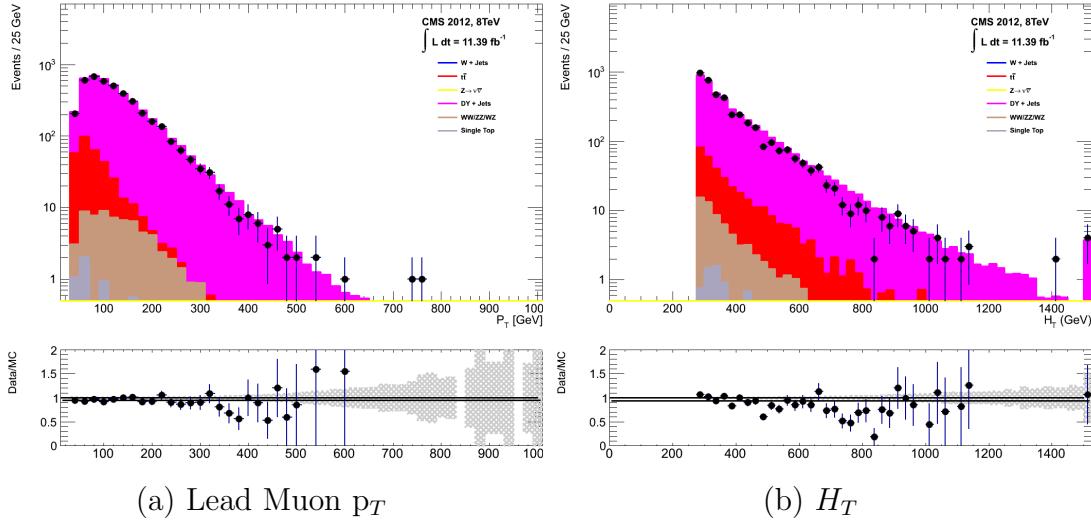
The $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background enters into the signal region from genuine E_T from the escaping neutrinos. This background is estimated using two control samples, the first of which is the $Z \rightarrow \mu\bar{\mu} + \text{jets}$ process, which posses identical kinematic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

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

calculation of event level variables. 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.



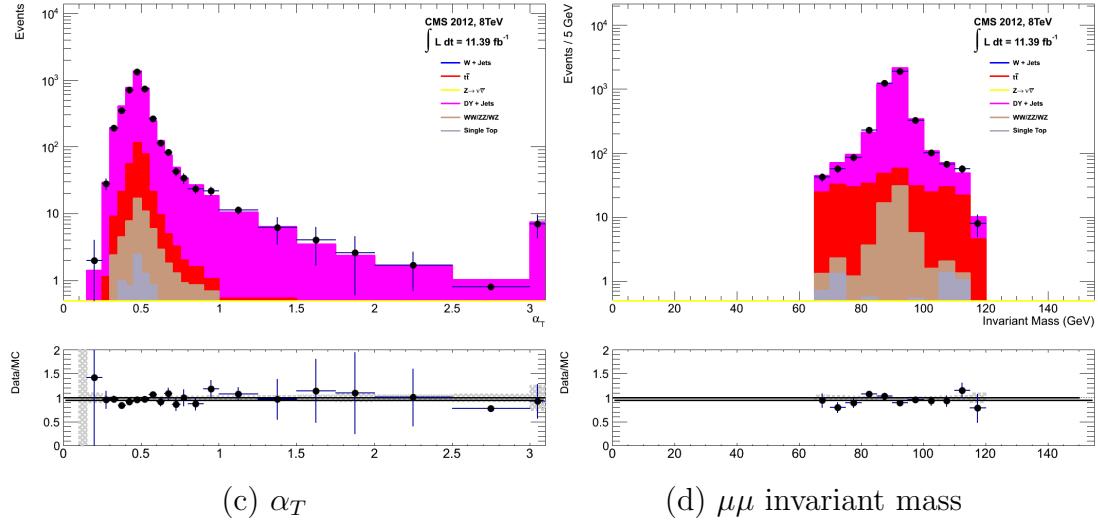


Figure 4.7.: Data/MC comparisons of key variables for the $\mu\mu + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number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

The $\gamma + \text{jets}$ 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 [79][80]. 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 $\Delta R(\gamma, \text{jet}) < 1.0$, between the photon and all jets is applied to ensure the acceptance of only well isolated $\gamma + \text{jets}$ events.
- Given that the photon is ignored, this control sample can only be applied in the H_T region > 375 GeV, due to the trigger thresholds on the minimum p_T of the photon, and the H_T requirement of an $\alpha_T > 0.55$ cut from Equation (4.5).

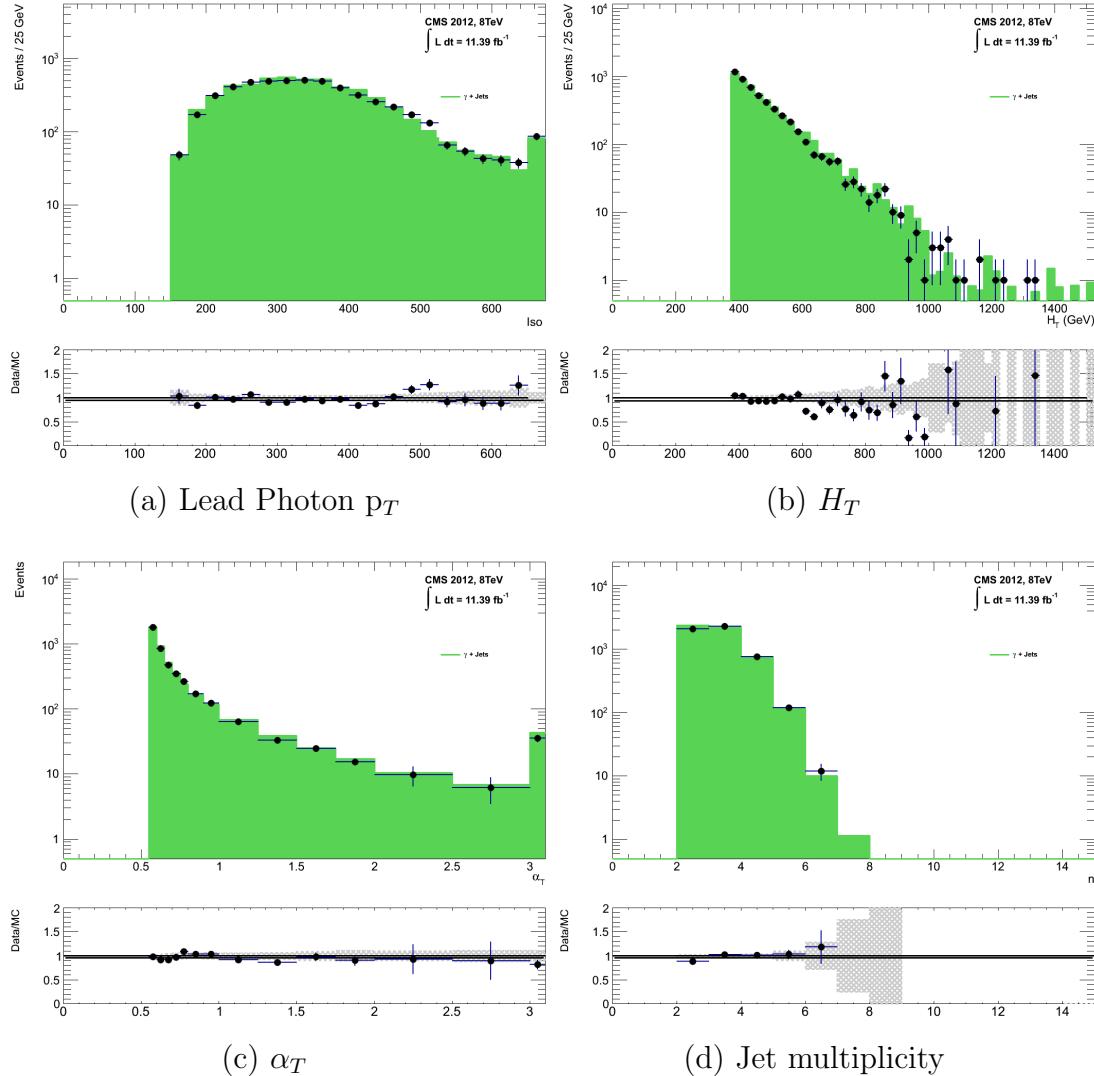


Figure 4.8.: Data/MC comparisons of key variables for the $\gamma + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement 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 α_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,

1540 the acceptance of the two muon control samples can be significantly increased, which
1541 simultaneously improves their predictive power and further reduces the effect of any
1542 potential signal contamination.

1543 The modelling of the α_T variable is probed through a dedicated set of closure tests,
1544 described in Section (4.6), which demonstrate that the different α_T acceptances for the
1545 control and signal samples have no significant systematic bias on the prediction.

1546 **4.2.4. Estimating the QCD multi-jet background**

1547 A negligible background from QCD multi-jet events within the hadronic signal region
1548 is expected due to the selection requirement, and additional cleaning filters applied.
1549 However a conservative approach is still adopted and the likelihood model, see Section
1550 (4.8.2), is given the freedom to estimate any potential QCD multi-jet contamination.

1551 Any potential contamination can be identified through the variable R_{α_T} , defined as the
1552 ratio of events above and below the α_T threshold value used in the analysis. This is
1553 modelled by a H_T dependant falling exponential function which takes the form,

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

1554 where the parameters A and k_{QCD} are the normalisation and exponential decay constants
1555 respectively.

1556 For QCD event topologies this exponential behaviour is expected as a function of H_T for
1557 several reasons. The improvement of jet energy resolution at higher H_T due to higher p_T
1558 jets leads to a narrower peaked distribution, causing R_{α_T} to fall. Similarly at higher H_T
1559 values > 375 GeV, the jet multiplicity rises slowly with H_T . As shown in Figure 4.3, at
1560 higher jet multiplicities, the result of the combinatorics used in the determination of α_T ,
1561 also lead to a narrower α_T distribution.

1562 The value of the decay constant k_{QCD} is constrained via measurements within data
1563 sidebands to the signal region. This is also done to validate the falling exponential
1564 assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet
1565 background and defined as regions where α_T is relaxed or that the R_{miss} cut is inverted.

1566 Figure 4.9 depicts the definition of these data sidebands used to constrain the value of
1567 k_{QCD} .

1568

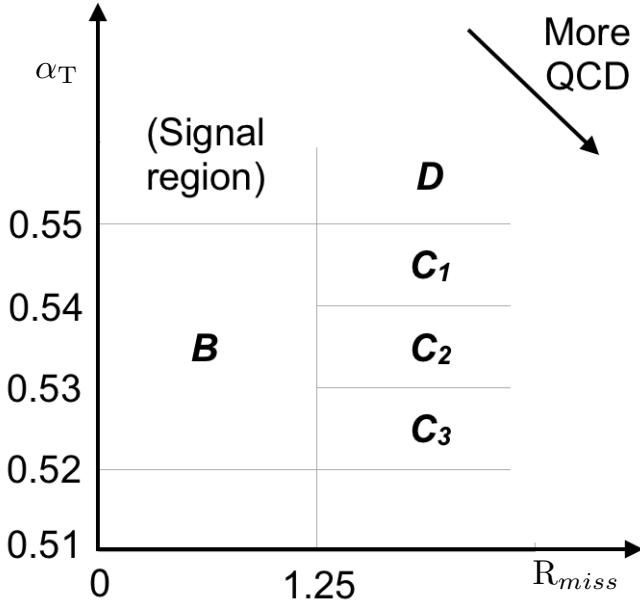


Figure 4.9.: QCD sideband regions, used for determination of k_{QCD} .

1569 The fits to determine the value of k_{QCD} are shown in Appendix (C.1), for which the best
1570 fit value obtained from sideband region B is determined to be $k_{QCD} = 2.96 \pm 0.64 \times 10^{-2}$
1571 GeV^{-1} .

1572 The best fit values of the remaining three C sideband regions are used to estimate
1573 the systematic uncertainty on the central value obtained from sideband region B. The
1574 variation of these measured values is used to determine the error on the determined
1575 central value, and is calculated to be $1.31 \pm 0.26 \times 10^{-2} \text{ GeV}^{-1}$. This relative error of \sim
1576 20% gives an estimate of the systematic uncertainty of the measurement to be applied to
1577 k_{QCD} .

1578 Finally the same procedure is performed for sideband region D to establish that the
1579 value of k_{QCD} extracted from a lower α_T slice can be applied to the signal region $\alpha_T >$
1580 0.55. The likelihood fit is performed across all H_T bins within the QCD enriched region
1581 with no constraint applied to k_{QCD} . The resulting best fit value for k_{QCD} shows good
1582 agreement between that and the weighted mean determined from the three C sidebands
1583 regions. This demonstrates that the assumption of using the central value determined
1584 from sideband region B, to provide an unbiased estimator for k_{QCD} in the signal region
1585 ($\alpha_T > 0.55$) is valid.

1586 Table 4.6, summarises the best fit k_{QCD} values determined for each of the sideband
1587 regions to the signal region.

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

Table 4.6.: Best fit values for the parameters k_{QCD} obtained from sideband regions B,C₁,C₂,C₃. 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.

1588 4.3. Trigger Strategy

1589 A cross trigger based on the quantities H_T and α_T , labelled is used with varying thresholds
1590 across H_T bins to record the events used in the hadronic signal region. The α_T legs of the
1591 HT_alphaT triggers used in the analysis are chosen to fully suppress QCD multi-jet events,
1592 whilst maintaining a sustainable trigger rate. To further maintain an acceptable rate for
1593 these analysis specific triggers, only calorimeter information is used in the reconstruction
1594 of the H_T sum, leading to the necessity for Calo jets to be used within the analysis.

1595 A single object prescaled H_T trigger is used to collect events for the hadronic control
1596 region described above in Section (4.2.4).

1597 The performance of the α_T and H_T triggers used to collect data for the signal and
1598 hadronic control region is measured with respect to a reference sample collected using the
1599 muon system. This allows measurement of both the Level 1 seed and higher level triggers
1600 simultaneously, as the reference sample is collected independent of any jet requirements.

1601 The selection for the trigger efficiency measurement is identical to that described in
1602 Section (4.2.2), with the requirement of exactly one well identified muon with $p_T > 30$
1603 GeV which is subsequently ignored.

1604 The efficiencies measured for the HT_alphaT triggers in bins individual H_T and α_T legs, is
1605 summarised in Table 4.7.

H_T range (GeV)	ϵ on H_T leg (%)	ϵ on α_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- ∞	$100.0^{+0.0}_{-0.0}$	$100.0^{+0.0}_{-4.8}$

Table 4.7.: Measured efficiencies of the H_T and α_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.

1606 Data for the control samples of the analysis, detailed in Section (4.2.3), are collected
 1607 using single object photon trigger for the $\gamma +$ jets sample, and a single object muon
 1608 trigger for both the $\mu +$ jets and $\mu\mu +$ jets control samples. The photon trigger is
 1609 measured to be full efficient for the threshold $p_T^{photon} > 150 GeV$, whilst the single muon
 1610 efficiency satisfying $p_T^{muon} > 30 GeV$ is measured to have an efficiency of $(88 \pm 2)\%$ that
 1611 is independent of H_T . In the case of the $\mu\mu +$ jets control sample, the efficiency is
 1612 measured to be $(95 \pm 2)\%$ for the lowest H_T bin, rising to $(98 \pm 2)\%$ for the highest H_T
 1613 bin.

1614 4.4. Measuring MC Normalisation Factors via H_T 1615 Sidebands

1616 The theoretical cross sections of different SM processes at Next to Next Leading Order
 1617 (NNLO) and the number of MC simulated events generated for that particular process,
 1618 is typically used to determine the appropriate normalisation for a MC sample. However
 1619 within the particular high- H_T and high- \cancel{E}_T corners of kinematic phase space probed
 1620 within this search, the theoretical cross sections for various processes are far less well
 1621 understood.

1622 To mitigate the problem of theoretical uncertainties and arbitrary choices of cross sections,
 1623 the normalisation of MC samples used in the analysis are determined through the use
 1624 data sidebands. The sidebands are used to calculate sample specific correct factors
 1625 (k-factors) that are appropriate for the H_T - \cancel{E}_T phase space covered by this analysis.

1626 They are defined within the $\mu +$ jets and $\mu\mu +$ jets control sample, by the region $200 <$
 1627 $H_T < 275$, using the same jet p_T thresholds as the adjacent first analysis bin. Individual
 1628 EWK processes are isolated within each of these control samples via requirements on

1629 jet multiplicity and the requirement on b-tags, summarised in Table 4.8. The purity of
1630 the samples are typically $> 90\%$ with any residual contamination corrected for. The
1631 resultant k-factor for each process is determined by then taking ratio of the data yield
1632 over the MC expectation in the sideband. Subsequently these k-factors are then applied
1633 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 ± 650.1	0.90 ± 0.02
$Z \rightarrow \mu\mu + \text{jets}$	$\mu\mu + \text{jets, } n_b=0, n_{jet} = 2,3$	3141	3402.0 ± 43.9	0.92 ± 0.02
$t\bar{t}$	$\mu + \text{jets, } n_b=2, n_{jet} = \geq 4$	2190	1967.8 ± 25.1	1.11 ± 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.

1634 4.5. Determining MC Simulation Yields with 1635 Higher Statistical Precision

1636 Reconstructing events from EWK processes with many b-tagged jets (≥ 3), n_b^{reco} , is largely
1637 driven by the mis-tagging of light jets within the event. This is clear when considering
1638 the main EWK backgrounds in the analysis, such as $t\bar{t} + \text{jets}$ events, which typically
1639 contain two b-flavoured jets from the decay of the top quarks, whilst W + jets and
1640 $Z \rightarrow \mu\mu + \text{jets}$ events will typically contain no b-flavoured jets.

1641 When the expectation for the number of n_b^{reco} is taken directly from simulation, the
1642 statistical uncertainty at large b-tag multiplicities becomes relatively large. In order to
1643 reduce this uncertainty one approach is to use the information encoded throughout all
1644 events in the simulation sample, to measure each of the four ingredients:

- 1645 1. the b-tagging efficiency in the event selection,
- 1646 2. the charm-tagging efficiency in the event selection
- 1647 3. the mis-tagging rate in the event selection,
- 1648 4. the underlying flavour distribution of the jets in the events,

1649 that determine the n_b^{reco} distribution of the process being measured. This method allows
1650 the determination of higher b-tag multiplicities to a higher degree of accuracy reducing

1651 the statical uncertainties of the MC which enter into the TF's. For the discussion that
1652 follows, these predictions are determined on average (i.e not on an event-by-event basis),
1653 and is known as the formula method.

1654 **4.5.1. The formula method**

1655 The assigning of jet flavours to reconstruction level jets in simulation is achieved via an
1656 algorithmic method defined as:

- 1657 • Try to find the parton that most likely determines the properties of the jet and
1658 assign that flavour as true flavour,
- 1659 • Here, the “final state” partons (after showering, radiation) are analysed (also within
1660 $\Delta R < 0.3$ of reconstructed jet cone),
- 1661 • Jets from radiation are matched with full efficiency,
- 1662 • If there is a b/c flavoured parton within the jet cone: label as b/c flavoured jet,
- 1663 • Otherwise: assign flavour of the hardest parton.

1664 Within each individual MC process and each H_T - n_{jet} bin in the analysis, the n_b^{reco}
1665 distribution is constructed in the following way:

1666 Let $N(n_b^{gen}, n_c^{gen}, n_q^{gen})$ represent the yield in simulation of events with b underlying
1667 b-quarks, c underlying c-quarks and q underlying light quarks which are matched to
1668 reconstructed jets as detailed above. Light quarks defined as those which originate from
1669 a u , d , s , g and τ jets, which having similar mis-tagging rates are grouped together.
1670 Similarly ϵ , β and m represent the measured b-tagging, c-tagging and mis-tagging
1671 efficiency averaged over all the jets within that particular analysis bin.

1672 Using this information the expected n_b^{reco} distribution can be analytically calculated
1673 using the formula :

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

1674 with $N(n_b)$ representing the yield where n_b jets have been b-tagged. The variables
1675 n_b^{tag} , n_c^{tag} and n_q^{tag} represent the number of times that a particular jet flavour results
1676 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
1677 binomial probabilities for that to happen.

1678 This approach ultimately results in a more precise n_b^{reco} distribution prediction as
1679 information from throughout the entire MC sample is used to estimate the underlying
1680 n_b^{reco} distribution.

1681 4.5.2. Establishing proof of principle

1682 In order to validate the procedure, the predictions obtained from the formula method
1683 summarised in Equation (4.11), are compared directly to those obtained directly from
1684 simulation. Resultantly no simulation to data correction factors are applied when making
1685 this comparison

1686 This sanity check for the $\mu +$ jets control sample is presented in Table 4.9, for all n_b^{reco}
1687 and H_T bins with no requirement placed upon the jet multiplicity of the events.

1688 It can be seen as expected, that there is good consistency between the results determined
1689 via the formula method and ‘raw’ simulation yields. Similarly the power of this approach
1690 can be seen in the reduction of this statistical error in the prediction across all H_T and
1691 n_b^{reco} bins. In particular the statistical uncertainty is reduced by several factors in the
1692 highest $n_b^{reco} \geq 4$ category.

1693 4.5.3. Correcting measured efficiencies in simulation to data

1694 As detailed in Section (3.3.2), it is necessary for certain p_T and η dependant corrections,
1695 to be applied to both the b-tagging efficiency and mis-tagging rates in order correct the
1696 efficiencies from simulation to the distributions seen in data. These corrections factors
1697 are considered when determining the simulation yields for each selection, which are used
1698 to construct the TF’s of the analysis.

1699 Each of the corrections factors for the b, c and light flavoured jets come with an associated
1700 systematic uncertainty. The uncertainties across different jet p_T and η bins, are considered
1701 as fully correlated. When computing the magnitude of the effect of this systematic
1702 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 11.4fb^{-1} . No simulation to data corrections are applied.

1703 simultaneously within H_T bin of the analysis for each of the $SF_{b,c,light}$ scale factors. The
 1704 magnitude of this correction is shown for each H_T bin within Figure 4.10.

1705 Varying the scale factor corrections by their systematic uncertainty will change the
 1706 absolute yields within each n_b^{reco} bin of all selections. However, ultimately it is the change
 1707 in the TF's which influences the final background prediction from each of the control
 1708 samples. The magnitude of these change is shown in Table 4.10, highlighting the absolute
 1709 change in each TF, constructed from when the $\mu +$ jets control sample is used to predict
 1710 the entire hadronic signal region background.

1711 It can be seen that the TF's are found to be relatively insensitive to the systematic
 1712 uncertainty of the b-tag scale factors (showing typically less than $\sim 2\%$ change). This can

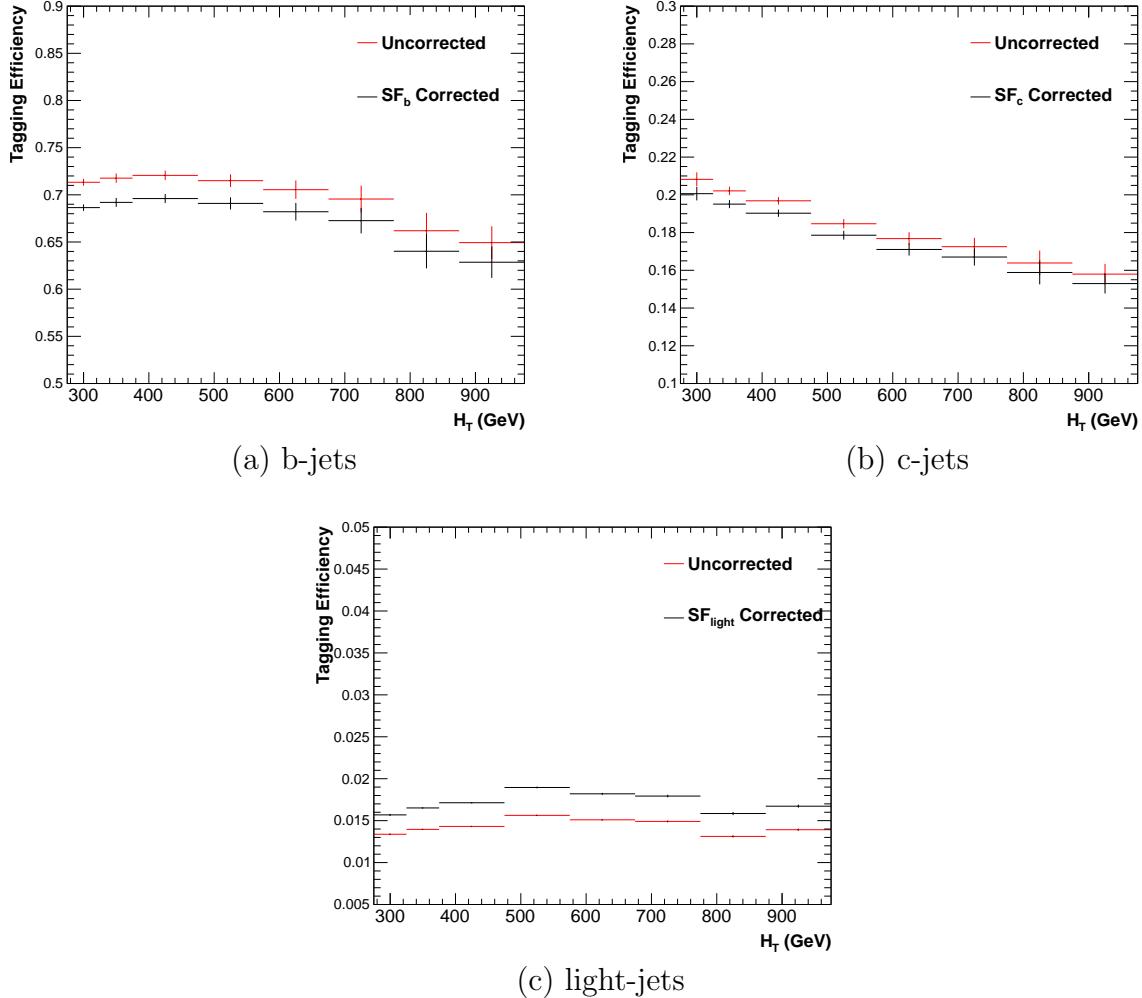


Figure 4.10.: Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets as a function all jets within each individual analysis H_T bin. Efficiencies measured directly from simulation (black) and with data to simulation $SF_{b,c,light}$ correction factors (black) are applied.

1713 be accounted for by the similar composition of the signal and control sample backgrounds,
 1714 such that any change in the underlying n_b^{reco} distribution will be reflected in both signal
 1715 and control regions and cancel out in the TF.

1716 Any overall systematic effect on the overall background prediction of the analysis from
 1717 these b-tag scale factor uncertainties is incorporated within the data driven systematics
 1718 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
≥ 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	-
<hr/>				
	575–675	675–775	775–875	≥ 875
<hr/>				
= 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)

¹⁷¹⁹ 4.6. Systematic Uncertainties on Transfer Factors

¹⁷²⁰ Since the **TF**'s used to establish the background prediction are obtained from simulation,
¹⁷²¹ an appropriate systematic uncertainty is assigned to each factor to account for theoretical
¹⁷²² uncertainties [81] and limitations in the simulation modelling of event kinematics and
¹⁷²³ instrumental effects.

¹⁷²⁴ The magnitudes of these systematic uncertainties are established through a set of data
¹⁷²⁵ driven method, in which the three independent control samples of the analysis ($\mu +$ jets,
¹⁷²⁶ $\mu\mu +$ jets, $\gamma +$ jets) are used to in a series of closure tests. The yields from one of these
¹⁷²⁷ control samples, along with the corresponding **TF** obtained from simulation, are used to
¹⁷²⁸ predict the yields in another control sample, using the same method of establishing a
¹⁷²⁹ background prediction for the signal region as described in Section (4.2.3).

¹⁷³⁰ The level of agreement between the predicted and observed yields is expressed as the
¹⁷³¹ ratio

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

1732 while considering only the statistical uncertainties on N_{pred} , the prediction, and N_{obs} , the
 1733 observation. No systematic uncertainty is assigned to the prediction, and resultantly the
 1734 level of closure is defined by the statistical significance of a deviation from the ratio from
 1735 zero.

1736 This ratio is measured for each H_T bin in the analysis, allowing these closure tests to be
 1737 sensitive to both the presence of any significant biases or any possible H_T dependence on
 1738 the level of closure.

1739 Eight sets of closure tests are defined between the three data control samples, conducted
 1740 independently between the two jet multiplicity ($2 \leq n_{jets} \leq 3$, $n_{jet} \geq 4$) bins. Each of
 1741 these tests are specifically chosen to probe each of the different key ingredients of the
 1742 simulation modelling that can affect the background prediction.

1743 Each of the different modelling components and the relevant closure tests are described
 1744 below :

1745 **α_T modelling**

1746 The modelling of the α_T distribution in genuine \cancel{E}_T events is probed with the $\mu +$
 1747 jets control sample. This test is important to verify the approach of remove the α_T
 1748 > 0.55 requirement from the $\mu +$ jets and $\mu\mu +$ jets samples to increase the precision
 1749 of the background prediction. The test uses the $\mu +$ jets sample without an α_T cut
 1750 to make a prediction into the $\mu +$ jets sample defined with the requirement $\alpha_T >$
 1751 0.55.

1752 **Background admixture**

1753 The sensitivity of the translation factors to the relative admixture of events from
 1754 $W +$ jets and $t\bar{t}$ processes is probed by two closure tests. These tests represent
 1755 an extremely conservative approach as the admixture of the background remains
 1756 similar between the $\mu +$ jets sample and the signal region, contrary to the defined
 1757 closure tests which make predictions between two very different admixtures of $W +$
 1758 jets and $t\bar{t}$ events.

1759 Within the $\mu + \text{jets}$ sample, a W boson enriched sub-sample ($n_b = 0$) is used to
1760 predict yields in a $t\bar{t}$ enriched sub-sample ($n_b = 1$). Similarly the
1761 $t\bar{t}$ enriched sub-sample ($n_b = 1$) is also used to predict yields for a further enriched
1762 $t\bar{t}$ sub-sample ($n_b = 2$).

1763 Similarly a further closure test probes the relative contribution of $Z + \text{jets}$ to $W + \text{jets}$
1764 and $t\bar{t}$ events, through the use of the $\mu + \text{jets}$ sample to predict yields for the $\mu\mu +$
1765 jets control sample. This closure test, also at some level probes the muon trigger
1766 and reconstruction efficiencies, given that exactly one and two muons are required
1767 by the different selections.

1768 Consistency between control samples

1769 An important consistency check between the $\mu\mu + \text{jets}$ jets and $\gamma + \text{jets}$, which are
1770 both used in the prediction of the $Z \rightarrow \nu\bar{\nu}$ in the signal region, is measured by using
1771 the $\gamma + \text{jets}$ sample to predict yields for the $\mu\mu + \text{jets}$ control sample.

1772 Modelling of jet multiplicity

1773 The simulation modelling of the jet multiplicity within each control sample is
1774 important due to the exclusive jet multiplicity binning within the analysis. This is
1775 probed via the use of each of the three control samples to independently predict
1776 from the lower jet multiplicity category $2 \leq n_{jet} \leq 3$, to the high jet category ≥ 4 .

1777 For the case of the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples this test is also a
1778 further probe of the admixture between $W + \text{jets}/Z + \text{jets}$ and $t\bar{t}$.

1779 To test for the assumption that no H_T dependences exist within the background predic-
1780 tions of the analysis, the first five closure tests defined above are taken, with zeroeth and
1781 first order polynomial fits are applied to each. This is summarised in Table 4.11 and
1782 Table 4.12 which show the results for both the $2 \leq n_{jet} \leq 3$ and ≥ 4 jet multiplicity bins
1783 respectively.

1784 Table 4.13 shows the same fits applied to the three closure tests that probe the modelling
1785 between the different n_{jet} bins. The best fit value and its uncertainty is listed for each
1786 set of closure tests in all three tables, along with the p-value of the constant and linear
1787 fits applied.

1788 The best fit value for the constant parameter is indicative of the level of closure, averaged
1789 across the full range of H_T bins in the analysis, and the p-value an indicator of any
1790 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 ± 0.02	0.93	-1.3 ± 2.2	0.91
$0 \text{ b-jets} \rightarrow 1 \text{ b-jet } (\mu + \text{jets})$	Square	0.07 ± 0.02	0.98	-1.6 ± 1.6	1.00
$1 \text{ b-jets} \rightarrow 2 \text{ b-jet } (\mu + \text{jets})$	Triangle	-0.07 ± 0.03	0.76	-2.7 ± 3.0	0.76
$\mu + \text{jets} \rightarrow \mu\mu + \text{jets}$	Cross	0.10 ± 0.03	0.58	-1.1 ± 2.3	0.49
$\mu\mu + \text{jets} \rightarrow \gamma + \text{jets}$	Star	-0.06 ± 0.04	0.31	4.2 ± 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 ± 0.03	0.21	3.0 ± 2.9	0.21
$0 \text{ b-jets} \rightarrow 1 \text{ b-jet } (\mu + \text{jets})$	Square	-0.03 ± 0.03	0.55	-1.0 ± 1.9	0.47
$1 \text{ b-jets} \rightarrow 2 \text{ b-jet } (\mu + \text{jets})$	Triangle	-0.02 ± 0.03	0.39	1.1 ± 2.2	0.31
$\mu + \text{jets} \rightarrow \mu\mu + \text{jets}$	Cross	0.08 ± 0.07	0.08	4.8 ± 4.3	0.07
$\mu\mu + \text{jets} \rightarrow \gamma + \text{jets}$	Star	-0.03 ± 0.10	0.72	-4.0 ± 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 ± 0.02	0.02	0.0 ± 1.0	0.01
$\mu + \text{jets}$ (outlier removed)	Inverted triangle	-0.04 ± 0.01	0.42	-1.4 ± 1.1	0.49
$\gamma + \text{jets}$	Diamond	0.12 ± 0.05	0.79	6.0 ± 4.7	0.94
$\mu\mu + \text{jets}$	Asterisk	-0.04 ± 0.07	0.20	4.9 ± 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 2σ 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 ± 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 .

1799 In addition the best fit values for the slope terms of the linear fits in all three tables are
1800 of the order 10^{-4} , which corresponds to a percent level change per 100 GeV. However in
1801 all cases, the best fit values are fully compatible with zero (within 1σ) once again with
1802 the exception detailed above, indicating that the level of closure is H_T independent.

1803 4.6.1. Determining systematic uncertainties from closure tests

1804 Once it has been established that no significant bias or trend has been exist within
1805 the closure tests, systematic uncertainties are determined. The statistical precision
1806 of the closure tests is considered a suitable benchmark for determining the systematic
1807 uncertainties that are assigned to the TF's, which are propagated through to the likelihood
1808 fit.

1809 The systematic uncertainty band is split into five separate regions of H_T . Within each
1810 region the square root of the sample variance, σ^2 , is taken over the eight closure tests to
1811 determine the systematic uncertainties to be applied within that region.

1812 Using this procedure the systematic uncertainties for each region are calculated and are
1813 shown in Table 4.14, with the systematic uncertainty to be used in the likelihood model
1814 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.

1815 Figure 4.11 shows the sets of closure tests overlaid on top of grey bands that represent
1816 the H_T dependent systematic uncertainties. These systematic uncertainties are assumed
1817 to fully uncorrelated between the different n_b multiplicity categories and across the five
1818 H_T regions. This can be considered a more conservative approach given that some
1819 correlations between adjacent H_T bins could be expected due to comparable kinematics.

1820 As already referenced. These closure tests represent a conservative estimate of the
1821 systematic uncertainty in making a background prediction for the signal region. This

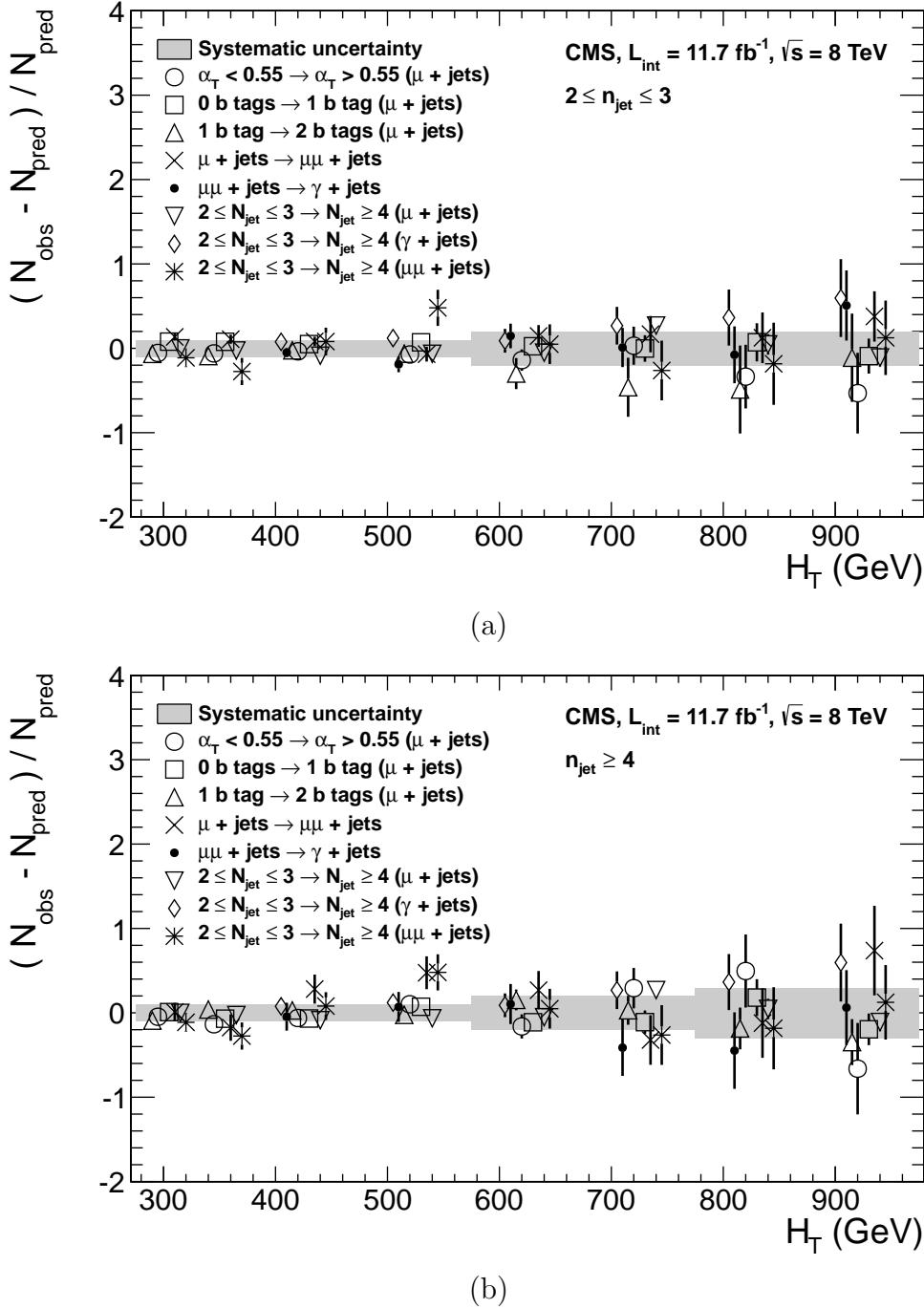


Figure 4.11.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins: (a) $2 \leq n_{\text{jet}} \leq 3$ and (b) $n_{\text{jet}} \geq 4$.

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

1825 admixture and similar kinematics owing to the fact that the predictions are always made
1826 using the same (n_{jet}, n_b, H_T) bin.

1827 This point is emphasised when we examine the sensitivity of the **TF**'s to a change in the
1828 admixture of W + jets and $t\bar{t}$ with the control and signal samples. This is accomplished
1829 by varying the cross sections of the W + jets and $t\bar{t}$ by +20% and -20%, respectively.
1830 Figures C.2 and C.3 within Appendix C, show the effect upon the closure tests for both
1831 jet multiplicity categories. Given these variations in cross sections, the level of closure is
1832 found to be significantly worse, with biases as large as $\sim 30\%$, most apparent in the
1833 lowest H_T bins. However the **TF**'s used to extrapolate from control to signal are seen to
1834 change only at the percent level by this large change in cross section, shown in Table C.1.

1835 Given the robust behaviour of the translation factors with respect to large (and opposite)
1836 variations in the W + jets and $t\bar{t}$ cross sections, one can assume with confidence that
1837 any bias in the translation factors is adequately (and conservatively) covered by the
1838 systematic uncertainties used in the analysis.

1839 4.7. Simplified Models, Efficiencies and Systematic 1840 Uncertainties

1841 The results of the analysis are interpreted using various **SMS** signal models as already
1842 introduced in Section (2.4.1). Each model is parameterised in a two dimensional parameter
1843 space, ($m_{\tilde{q}/\tilde{g}}, m_{LSP}$), from which upper limits on the production cross sections of the
1844 various **SMS** models can be set.

1845 Each signal sample is generated at Next to Leading Order (**NLO**) and Next to Leading
1846 Logarithmic Order (**NLL**) [82] using the **Fastsim** framework. This framework represents
1847 a simplified simulation of the **CMS** detector, but allows for faster production of various
1848 signal topologies with different mass parameters. A series of correction factors are applied
1849 to account for the effects on the b-tagging rate between **Fastsim** [83] and **Fullsim** [84]
1850 and are detailed in Section (4.7.2).

1851 **4.7.1. Signal efficiency**

1852 The analysis selection efficiency, ϵ , is measured for each mass point of the interpreted
 1853 model, this serves as a measure of the sensitivity of the signal selection for that particular
 1854 sparticle and LSP mass . The signal yield is then given by

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

1855 where σ represents the model's cross section and \mathcal{L} the luminosity. An upper limit on σ
 1856 taken from theory can then allow for the setting of limits in terms of the particle mass.

1857 Figure 4.12 shows the expected signal efficiency of the signal selection for the T1 and
 1858 T2 SMS models interpreted in this analysis. The efficiency maps are produced with the
 1859 requirement $H_T > 275$ GeV (i.e., no binning in H_T) and requirements on n_{jet} and n_b^{reco}
 1860 that are appropriate for the model in question.

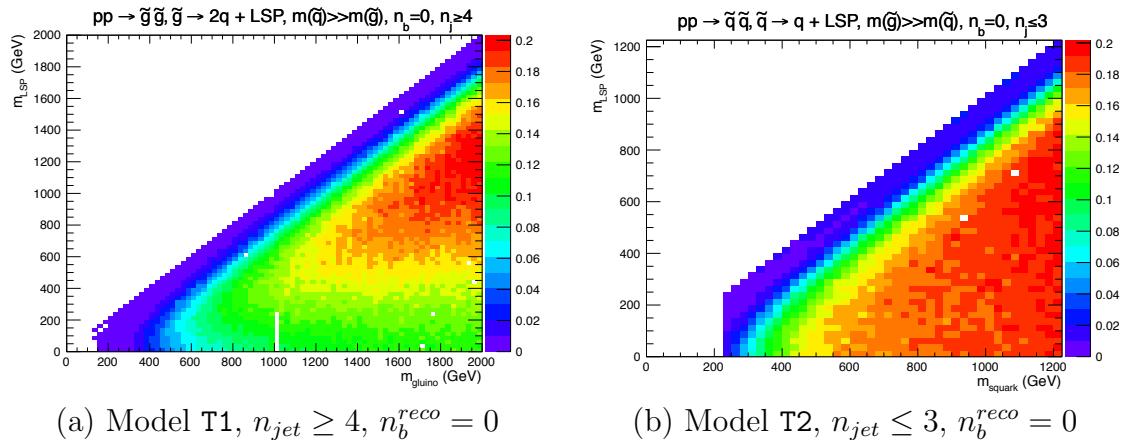


Figure 4.12.: Signal efficiencies for the SMS models (a) T1 ($\tilde{g}\tilde{g}^* \rightarrow q\tilde{\chi}_1^0 q\tilde{\chi}_1^0$) and (b) T2 ($\tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 \bar{q}\tilde{\chi}_1^0$) when requiring $n_{jet} \geq 4$ and ≤ 3 respectively, and $n_b^{reco} = 0$.

1861 The same procedure is conducted in the analysis control samples. It is found in the μ
 1862 + jets control samples, that the S/B ratios for the expected signal yields in each of the
 1863 SMS models are many times (~ 40 -100) smaller than in the hadronic signal region. The
 1864 relative contamination for the $\mu\mu$ + jets sample is smaller still due to the requirement of
 1865 a second muon. The relative contamination for the γ + jets sample is expected to be
 1866 zero for the models under consideration. These small, relative levels of contamination
 1867 are accounted for in the fitting procedure, as described in Section (4.8.4).

1868 4.7.2. Applying b-tag scale factor corrections in signal samples

1869 High-statistic **FastSim** signal simulation samples are unavailable for each signal signal
 1870 point, which means that a different procedure to the formula method described in Section
 1871 (4.5) is employed. Furthermore, the use of the **FastSim** framework in the reconstruction
 1872 introduces an extra set of scale-factor corrections, to be applied simultaneously with
 1873 those correcting the full-simulation to the data.

1874 For these signal models, an event-by-event re-weighting procedure is applied. This applied
 1875 weight depends on both the flavour content and the b-tagging status of the reconstruction
 1876 level jets in the event.

1877 The re-weighting procedure can be described by first considering a single jet in an signal
 1878 sample event. The flavour of the jet is determined using the method described in Section
 1879 (4.5.1).

1880 Taking the flavour, p_T and η values of the jet, the expected tagging efficiency, $\epsilon_{MC}(p_T, \eta, f)$,
 1881 in simulation is retrieved from a map of tagging efficiencies determined from the **FullSim**
 1882 **SM** simulation samples, and binned as a function of jet p_T , η and flavour after the
 1883 application of the hadronic signal selection. The binning is chosen to reflect the set of p_T
 1884 and η dependant corrections of simulation to data defined by [85].

1885 The actual tagging efficiency of the **FastSim** jet, $\epsilon_{\text{FastSim}}(p_T, \eta, f)$, differs from that
 1886 measured in **FullSim**, $\epsilon_{MC}(p_T, \eta, f)$ and is related via an additional correction factor,

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

1887 where $SF_{\text{Fast} \rightarrow \text{Full}}(p_T, \eta, f)$ represents a set of p_T and η dependant corrections determined
 1888 from the ratio between the efficiency and mis-tagging rates of a $t\bar{t}$ **FullSim** and $t\bar{t}$
 1889 **FastSim** sample. The central value for these corrections is the same for all signal samples.
 1890 Similarly the tagging efficiencies measured in data [64], $\epsilon_{Data}(p_T, \eta, f)$, are further related
 1891 to $\epsilon_{\text{FastSim}}(p_T, \eta, f)$ by the equation,

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

1892 For each jet, the weight of the event is re-weighted according to whether the jet fires the
1893 b-tagger. In the instance that the jet *is* b-tagged the event weight is modified by,

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

1894 and in the case that the jet is *not* tagged,

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

1895 Once all events have been reweighted this way, the yields in each n_b^{reco} bin represent the
1896 corrected MC yields.

1897 4.7.3. Experimental uncertainties

1898 The systematic uncertainty on the expected signal acceptance times analysis efficiency is
1899 determined independently for each **SMS** model considered. These systematics stem
1900 from uncertainties on the parton distribution functions, the luminosity measurement,
1901 jet energy scale, b-tag scale factor measurements and the efficiencies of various cuts used
1902 in the signal selection, including the H_T / E_T , dead **ECAL** cleaning filter and lepton /
1903 photon event vetoes.

1904 Rather than trying to estimate the level of systematic that is applicable point-by-point in
1905 a model space, general behaviours are considered and constant systematics are estimated
1906 in two regions of the **SMS** models parameter space. These two regions are defined as
1907 near to (small mass splittings) and far (large mass splittings) from the diagonal, where
1908 far is realised by the condition

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

1909 The total systematics in each region are evaluated in the following ways:

1910 **Jet energy scale** : The relative change in the signal efficiency is gauged by varying
1911 the energy of all jets in an event up or down according to a p_T and η dependent jet
1912 energy scale uncertainty. Within the two systematic regions, the resulting systematic
1913 uncertainties for each **SMS** model are determined by taking the value of the 68th
1914 percentile for the distributions of the relative change in the signal efficiency.

1915 **Luminosity measurement** : The measurement of luminosity taken propagates
1916 through to an uncertainty on the signal event yield when considering any new
1917 physics model, which is currently 4.4% [86].

1918 **Parton density function** : The effects of varying the set of parton distribution
1919 functions used and changing the renormalization/factorization scale used to compute
1920 the cross sections by a factor of two up or down have been found to affect the
1921 efficiency by up to 10%.

1922 **H_T/E_T cleaning cut** : The ratio of the efficiencies of the cleaning cut are compared
1923 in simulation and data after application of the $\mu +$ jets control sample selection.
1924 No α_T cut or further event cleaning filters are applied. The ratio of the efficiencies
1925 observed in data and simulation for a cut value of $H_T/E_T < 1.25$ and the two jet
1926 multiplicity bins, $2 \leq n_{jet} \leq 3$ and $n_{jet} \geq 4$ are 1.028 ± 0.007 and 1.038 ± 0.015
1927 respectively. These deviations are taken to represent the systematic uncertainty on
1928 the simulation modelling of this variable.

1929 **Deal ECAL cleaning filter** : The ratio of the efficiencies observed in data and
1930 simulation for this filter in the two jet multiplicity bins, $2 \leq n_{jet} \leq 3$ and $n_{jet} \geq 4$,
1931 are 0.961 ± 0.008 and 0.961 ± 0.009 , respectively. These deviations from unity
1932 are taken to represent the systematic uncertainties in the modelling in simulation of
1933 this filter.

1934 **Lepton and photon vetoes** : The uncertainty on the efficiency of the lepton and
1935 photon vetoes is established by considering the efficiency of the vetoes after applying
1936 filters with identical logic but based on truth information. If the efficiency is not
1937 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 η 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 68th 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.

¹⁹⁵² 4.8. Statistical Framework

¹⁹⁵³ For a given category of events satisfying requirements on both n_{jet} and n_b^{reco} , a likelihood
¹⁹⁵⁴ model of the observations in multiple data samples is used to gauge agreement between
¹⁹⁵⁵ the observed yields in the hadronic signal region, and the predicted yields obtained from
¹⁹⁵⁶ the control samples. In addition to checking whether the predictions are compatible with
¹⁹⁵⁷ a **SM** only hypothesis, the likelihood model is also used to test for the presence of a
¹⁹⁵⁸ variety of signal models.

¹⁹⁵⁹ 4.8.1. Hadronic sample

¹⁹⁶⁰ Let N be the number of bins on H_T , with n^i the number of events observed satisfying
¹⁹⁶¹ all selection requirements in each H_T bin i. The likelihood of the observations can then
¹⁹⁶² be written :

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

¹⁹⁶³ where b^i represents the expected **SM** background

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

¹⁹⁶⁴ and s^i the expected number of signal events from the different **SMS** models interpreted.
¹⁹⁶⁵ Pois refers to the Poisson distribution of these values and is defined as :

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

₁₉₆₆ **4.8.2. H_T evolution model**

- ₁₉₆₇ The hypothesis, that for a process the α_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.
- ₁₉₇₀ The expected QCD background, QCD^i , within a bin i is then modelled as,

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

- ₁₉₇₁ where m_i represent the number of events observed with $\alpha_T \leq 0.55$ in each H_T bin i , and ₁₉₇₂ $\langle H_T \rangle$ represents the mean H_T of each bin. Expressed as functions of just the zeroth bin, ₁₉₇₃ QCD^0 , and k_{QCD} , the QCD expectation is given by

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

₁₉₇₄ **4.8.3. EWK control samples**

- ₁₉₇₅ The **EWK** background estimation within each bin, i , is broken into two components, the ₁₉₇₆ expected yield from $Z \rightarrow \nu\bar{\nu}$ and $t\bar{t}$ -W (plus other residual backgrounds) events. This is ₁₉₇₇ written as, Z_{inv}^i and $t\bar{t}W^i$, and it follows that

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

- ₁₉₇₈ This can be further expressed as

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

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

1979 where f_{Zinv}^i represents the expected yield from $Z \rightarrow \nu\bar{\nu}$ in bin i divided by the expected
1980 **EWK** background EWK^i . This fraction is modelled as a linear component

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

1981 where N again represents the number of H_T bins, and f_{Zinv}^i and f_{Zinv}^{N-1} are float parameters
1982 whose final values are limited between zero and one.

1983 Within each H_T bin there are three background measurements for the different control
1984 samples, n_γ^i , n_μ^i and $n_{\mu\mu}^i$, representing the event yields from the $\gamma +$ jets, $\mu +$ jets and
1985 $\mu\mu +$ jets control samples respectively. Each of these have a corresponding yield in
1986 simulation, MC_γ^i , MC_μ^i and $MC_{\mu\mu}^i$. Within the hadronic signal region there are also
1987 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
1988 used to define

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

1989 where r_p^i represents the inverse of the **TF**'s used to extrapolate the yield of each background
1990 process.

1991 The likelihoods regarding the three measured yields n_γ^i , $n_{\mu\mu}^i$, n_μ^i can then be fully expressed
1992 as

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

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

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

$$(4.31)$$

1993 which contain an additional term s_μ^i , which represents the signal contamination in the
1994 $\mu +$ jets sample. The parameters $\rho_{\gamma Z}^j$, $\rho_{\mu\mu}^j$ and ρ_μ^j represent “correction factors” that

1995 accommodate the systematic uncertainties associated with the control sample based
1996 background constraints.

1997 Each of these equations are used to estimate the maximum likelihood value for relevant
1998 background in the signal region given the observations n_p^i in each of the control samples
1999 (see Section (4.2.3)).

2000 The measurements in each of the control samples and the hadronic signal region, along
2001 with the ratios r_γ^i , $r_{\mu\mu}^i$, and r_μ^i , are all considered simultaneously through the relationships
2002 defined by Equations (4.19),(4.24) and (4.25).

2003 In addition to the Poission product, an additional log-normal term is introduced to
2004 accommodate the systematic uncertainties given by,

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

2005 where $\sigma_{\gamma Z}^j$, $\sigma_{\mu\mu Z}^j$ and $\sigma_{\mu W}^j$ represent the relative systematic uncertainties for the control
2006 sample constraints and Logn is the log-normal distribution [87],

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

2007 Five parameters per control sample are used to span the eight H_T bins, with just one
2008 used for the three H_T bins in the $n_b^{reco} \geq 4$ category. These parameters span the same
2009 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.

2010 Alternatively, in the higher n_b^{reco} categories ($n_b^{reco} = 3$, $n_b^{reco} \geq 4$), the single muon sample
2011 is used to constrain the total EWK background. Therefore the likelihood function is
2012 greatly simplified and is represented by

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

2013 where,

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

2014 **4.8.4. Contributions from signal**

2015 The cross section for each model is represented by x and l represents the total recorded luminosity considered by the analysis in the signal region. Let ϵ_{had}^i and ϵ_μ^i represent the 2016 analysis selection efficiency for that particular signal model in H_T bin i of the hadronic 2017 and $\mu +$ jets control sample respectively. Letting δ represent the relative uncertainty on 2018 the signal yield, assumed to be fully correlated across all bins, and ρ_{sig} the “correction 2019 factor” to the signal yield which accommodates this uncertainty. f represents an unknown 2020 multiplicative factor on the signal cross section, for which an allowed interval is computed. 2021

2022 The expected signal yield s^i is thus given by

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

2023 and signal contamination with the $\mu +$ jets control sample by

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

2024 The systematic uncertainty on the signal is additionally included by the term

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

2025 A discussion of the **SMS** signal models through which the analysis is interpreted can be
2026 found in the following Chapter.

2027 **4.8.5. Total likelihood**

2028 The total likelihood function for a given signal bin $k(n_b^{reco}, n_{jet})$ is given by the product
2029 of the likelihood functions introduced within the previous sections:

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

2030 In categories containing eight H_T bins and utilising the three control samples ($\mu + \text{jets}, \mu\mu$
2031 $+ \text{jets}, \gamma + \text{jets}$), there are 25 nuisance parameters, whilst when just one control sample is
2032 used to estimate the **EWK** background, there are 15 nuisance parameters. Where three
2033 H_T bins are used (the highest n_b^{reco} category), there are 6 nuisance parameters. This
2034 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.

2035 When considering **SUSY** signal models within the likelihood, an additional parameter is
2036 introduced, ρ_{sig} . When multiple categories are fit simultaneously the total likelihood is
2037 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.

²⁰³⁸ Results and Interpretation

²⁰³⁹ Using the statistical framework outlined in the previous chapter, results are compared to
²⁰⁴⁰ a **SM**-only hypothesis (Section (5.1)) and interpreted using various **SMS** models (Section
²⁰⁴¹ (5.2)).

²⁰⁴² 5.1. Standard Model

²⁰⁴³ The **SM** background only hypothesis is tested by removing any signal contributions
²⁰⁴⁴ within the signal and control samples, and the likelihood function is maximised over all
²⁰⁴⁵ parameters using Rootfit [88] and MINUIT [89]. The results of the search consist of the
²⁰⁴⁶ observed yields in the hadronic signal sample, and the $\mu + \text{jets}$, $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$
²⁰⁴⁷ control samples.

²⁰⁴⁸ These observed yields along with the expectations and uncertainties given by the simulta-
²⁰⁴⁹ neous fit for the hadronic signal region are given in Table 5.2. The results obtained from
²⁰⁵⁰ the simultaneous fits, including that of the three control samples, are shown in Figure
²⁰⁵¹ 5.1-5.8, as summarised in Table 5.1.

²⁰⁵² The figures show a comparison between the observed yields and the **SM** expectations
²⁰⁵³ across all H_T bins, for events in both n_{jet} and n_b^{reco} multiplicity categories. In all categories
²⁰⁵⁴ the samples are well described by the **SM** only hypothesis. In particular no significant
²⁰⁵⁵ 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
≥ 4	0	$\mu + \text{jets}, \mu\mu + \text{jets}, \gamma + \text{jets}$	5.4
≥ 4	1	$\mu + \text{jets}, \mu\mu + \text{jets}, \gamma + \text{jets}$	5.5
≥ 4	2	$\mu + \text{jets}$	5.6
≥ 4	3	$\mu + \text{jets}$	5.7
≥ 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- ∞
SM Data	0	≤ 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	≥ 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	≤ 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	≥ 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	≤ 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	≥ 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	≥ 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	≥ 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.

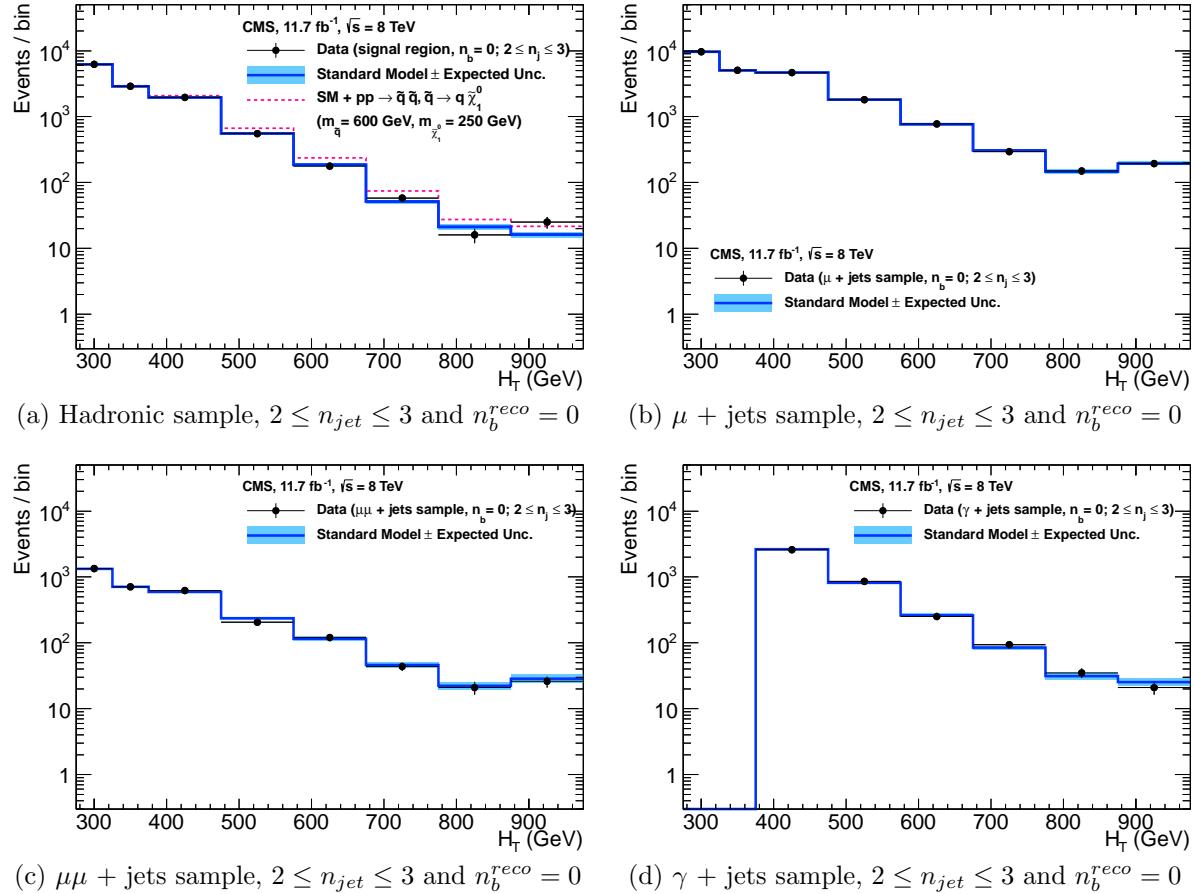


Figure 5.1.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \leq 3$. The observed event yields in data (black dots)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

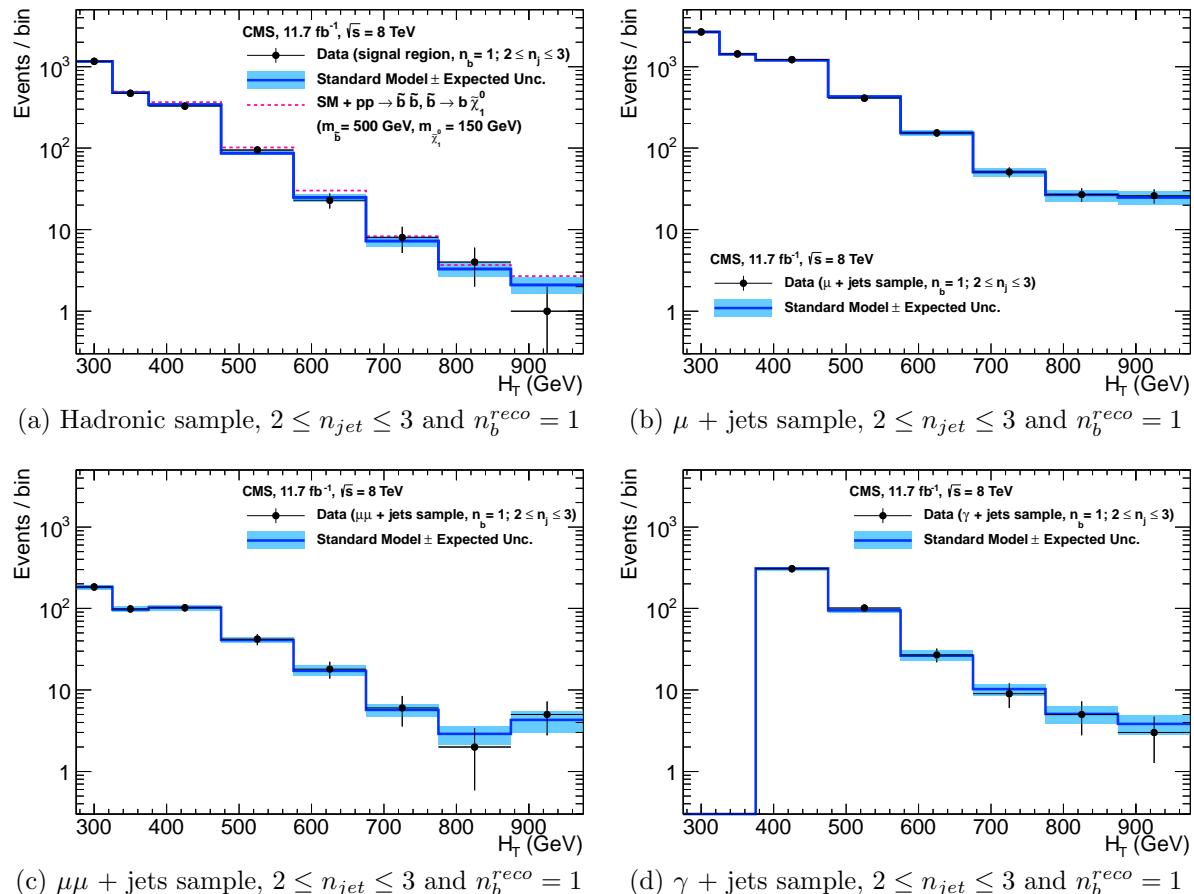


Figure 5.2.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$. The observed event yields in data (black dots)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

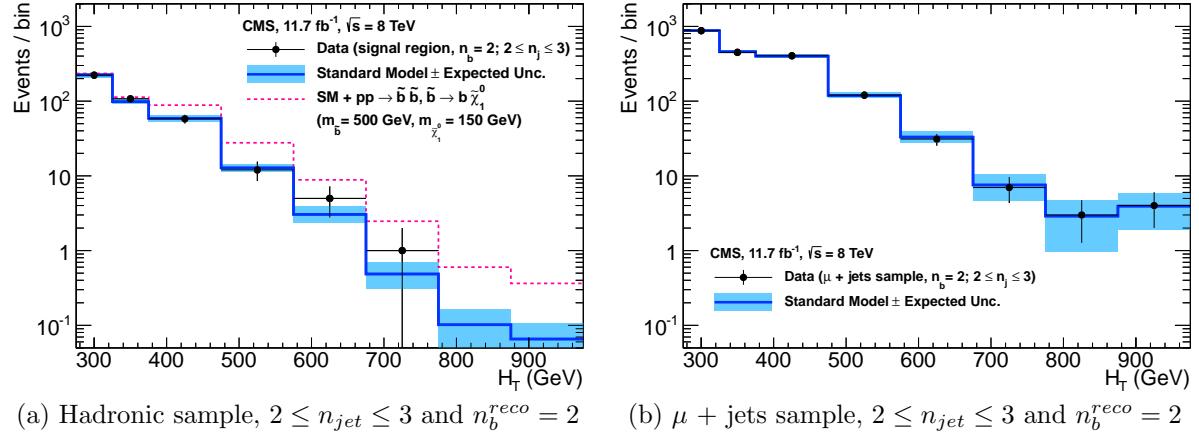


Figure 5.3.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$. The observed event yields in data (black dots)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

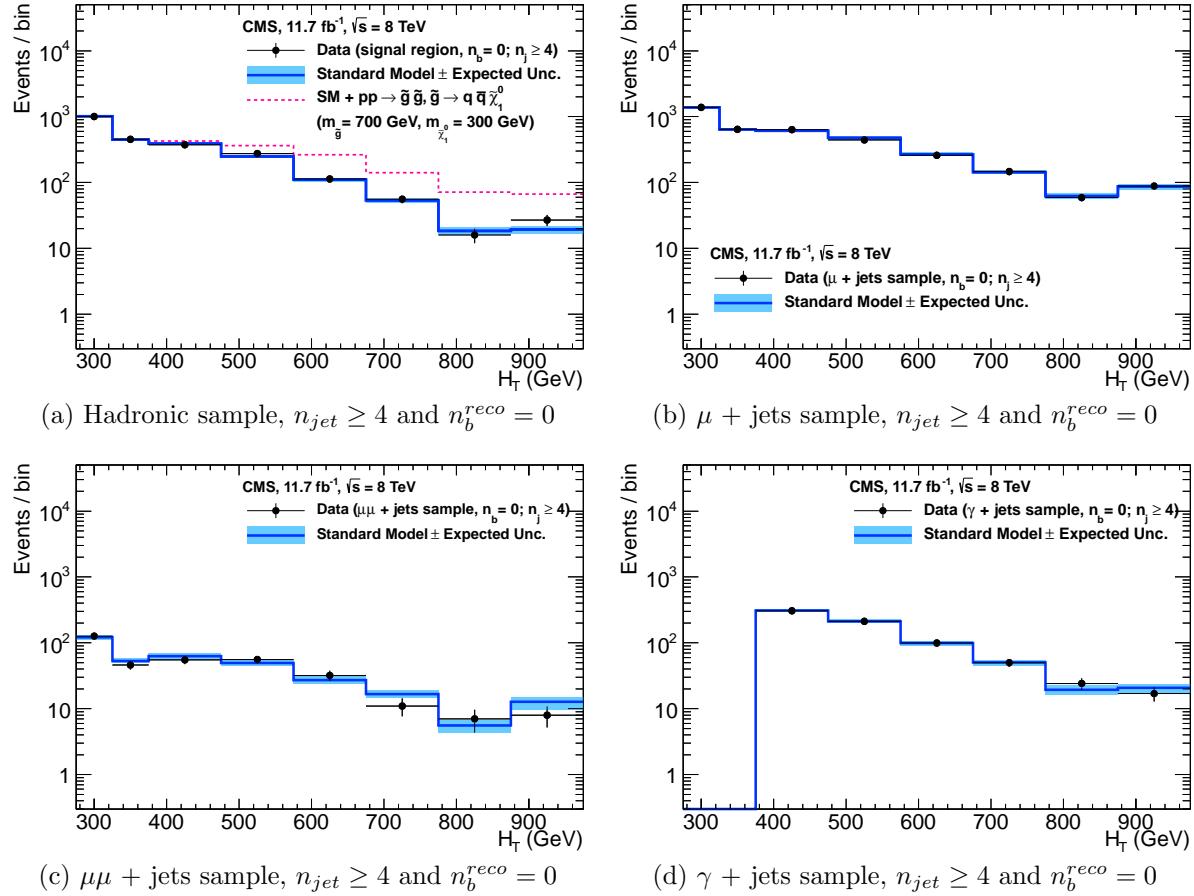


Figure 5.4.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

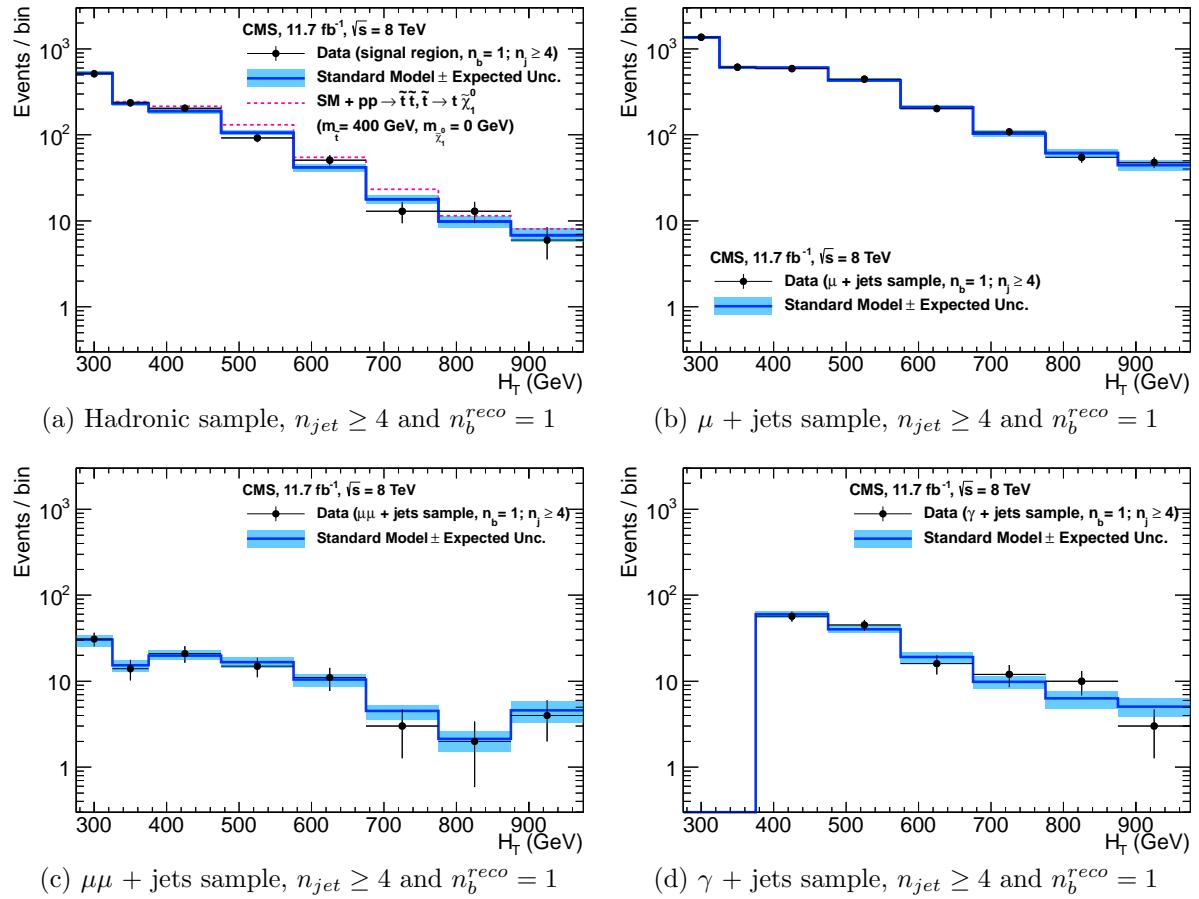


Figure 5.5.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

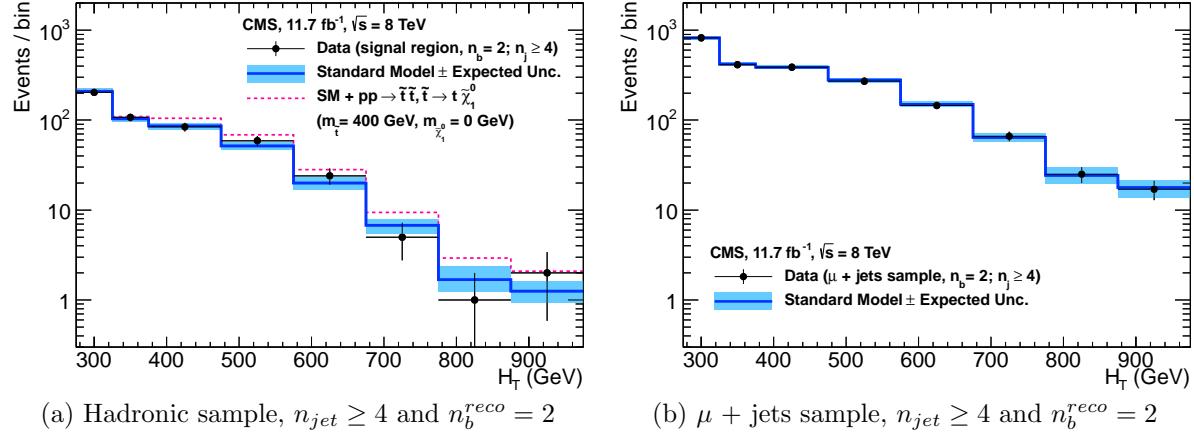


Figure 5.6.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

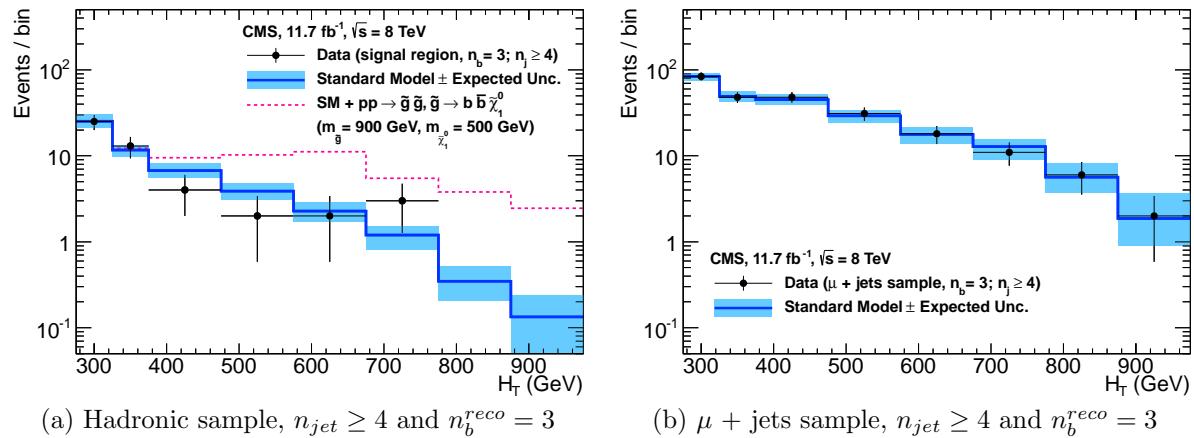


Figure 5.7.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

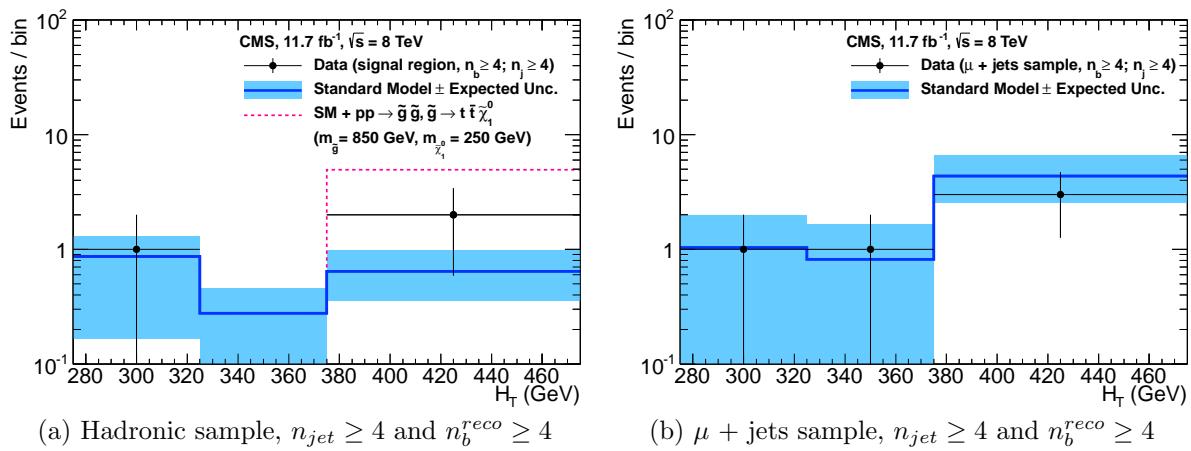


Figure 5.8.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$. The observed event yields in data (black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

2056 **5.2. SUSY**

2057 Limits are set in the parameter space of a set of **SMS** models that characterise both
2058 natural **SUSY** third generation squark production, and compressed spectra where the
2059 mass splitting between the particle and **LSP** is small, leading to soft final state jets.
2060 However as detailed in Section (2.4.1), the individual models are not representative of a
2061 real physical **SUSY** model as only one decay process is considered. Instead these models
2062 represent a way to test for signs of specific signatures indicating new physics.

2063 **5.2.1. The CL_s method**

2064 The CLs method [90][91][92] is used to compute the limits for signal models, with the
2065 one-sided profile likelihood ratio as the test statistic [93].

The test statistic is defined as

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

2066 where

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

2067 represents the profile likelihood ratio, in which $\mu \equiv f$ from Section (4.8.4), is the
2068 parameter characterising the signal strength. $\hat{\mu}$ is defined at the maximum likelihood
2069 value, $\hat{\theta}$ the set of maximum likelihood values of the nuisance parameters and θ_μ the set
2070 of maximum values of the nuisance parameters for a given value of μ .

2071 When $\mu \equiv f = 1$, the signal model is considered at its nominal production cross section.
2072 The distribution of q_μ is built up via the generation of pseudo experiments in order to
2073 obtain two distributions for the background (B) and signal plus background (S+B) cases.
2074 The compatibility of a signal model with observations in data is determined by the
2075 parameter CL_s ,

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

with CL_B and 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 $\text{CL}_s \leq 0.05$ [94].

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	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$	≥ 4	0	5.9(a)	5.10(a)	~ 950	~ 450
T2	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 \bar{q}\tilde{\chi}_1^0$	≤ 3	0	5.9(b)	5.10(b)	~ 775	~ 325
T2bb	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\tilde{\chi}_1^0 \bar{b}\tilde{\chi}_1^0$	≤ 3	1,2	5.9(c)	5.10(c)	~ 600	~ 200
T1tttt	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow t\bar{t}\tilde{\chi}_1^0 t\bar{t}\tilde{\chi}_1^0$	≥ 4	$2,3,\geq 4$	5.9(d)	5.10(d)	~ 975	~ 325
T1bbbb	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\tilde{b}\tilde{\chi}_1^0 b\tilde{b}\tilde{\chi}_1^0$	≥ 4	$2,3,\geq 4$	5.9(e)	5.10(e)	~ 1125	~ 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 1σ 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.

2093 Signal efficiency in the kinematic region defined by $0 < m_{\tilde{g}(\tilde{q})} < 175$ GeV or $m_{\tilde{g}(\tilde{q})} < 300$
2094 GeV is strongly affected by the presence of Initial State Radiation (**ISR**). This region in
2095 which direct (i.e., non-**ISR** induced) production is kinematically forbidden due to the H_T
2096 > 275 GeV requirement, therefore a large percentage of signal acceptance is due to the
2097 effect of **ISR** jets. Given the large associated uncertainties, no interpretation is provided
2098 for this kinematic region.

2099 The estimates on mass limits shown in Table 5.3, are determined conservatively from
2100 the observed exclusion based on the theoretical production cross section, minus 1σ
2101 uncertainty. The most stringent mass limits on pair-produced sparticles are obtained at
2102 low **LSP** masses, while the limits typically weaken for compressed spectra points close to
2103 the diagonal. In particular, for all of the considered **SMS** models, there is an **LSP** mass
2104 beyond which no limit can be set, which can be observed from the figures referenced in the
2105 table.

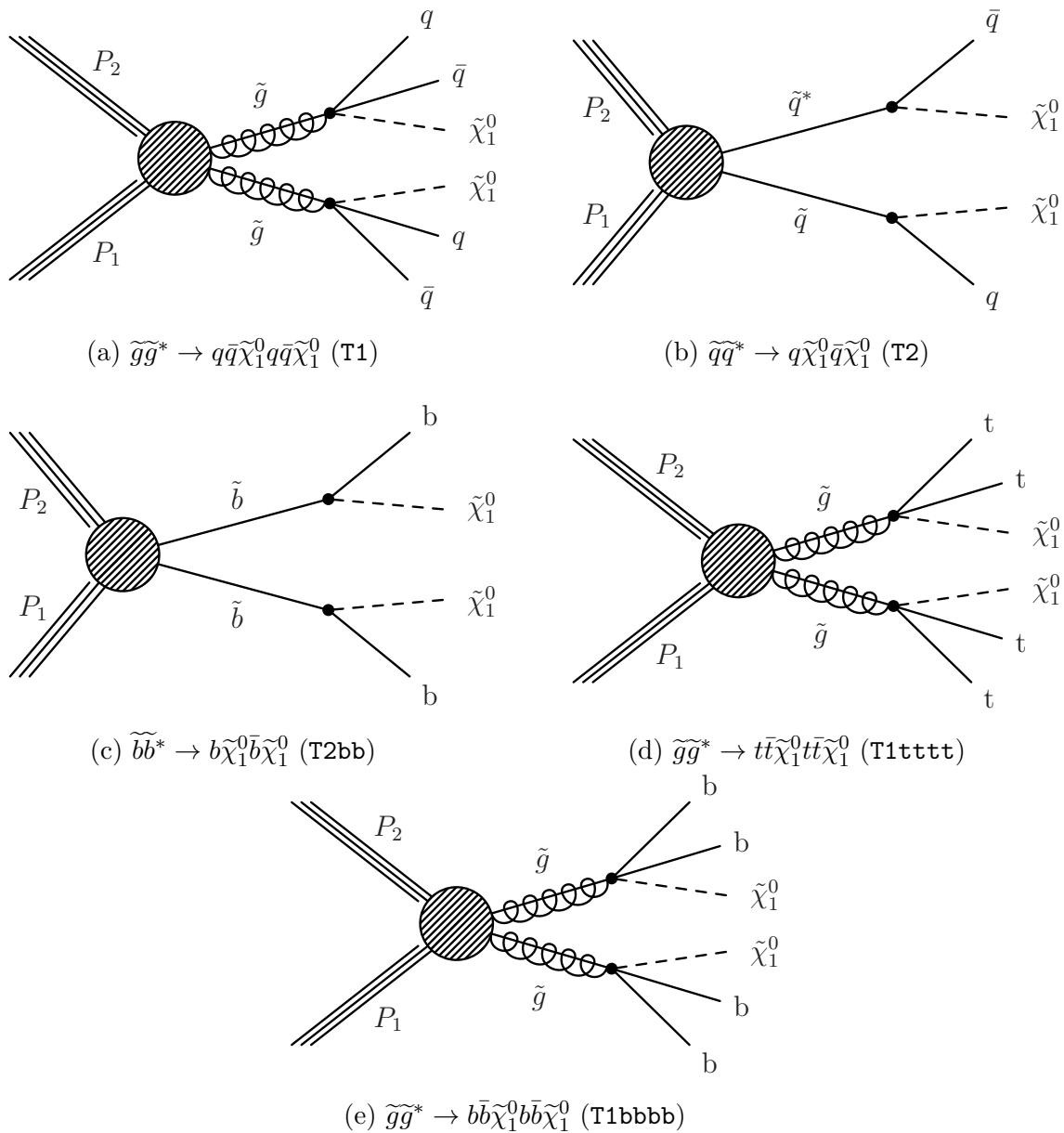


Figure 5.9.: Production and decay modes for the various **SMS** models interpreted within the analysis.

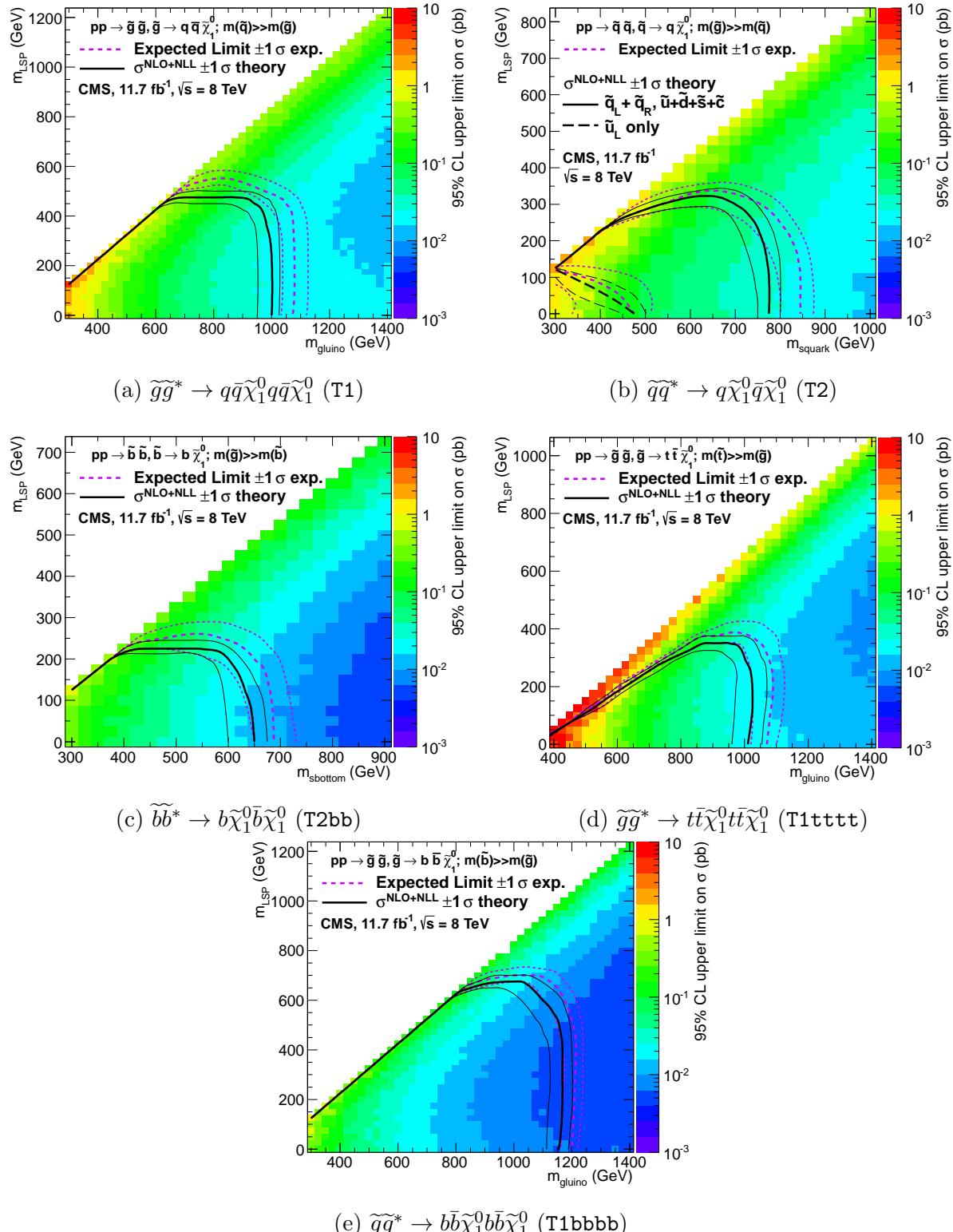


Figure 5.10.: Upper limit of cross section at 95% CL as a function of $m_{\tilde{q}/\tilde{g}}$ and m_{LSP} for various SMS models. The solid thick black line indicates the observed exclusion region assuming NLO and NLL SUSY production cross section. The analysis selection efficiency is measured for 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) 1σ (thin lines) expected exclusion regions.

Chapter 6.

²¹⁰⁶ Searches For Natural SUSY With ²¹⁰⁷ B-tag Templates.

²¹⁰⁸ Within this chapter a complimentary technique is discussed as a means to predict the
²¹⁰⁹ distribution of three and four reconstructed b-quark jets in an event. The recent discovery
²¹¹⁰ of the Higgs boson has made third-generation “Natural SUSY” models attractive, given
²¹¹¹ that light top and bottom squarks are a candidate to stabilise divergent loop corrections
²¹¹² to the Higgs boson mass.

²¹¹³ Using the α_T search as a base, a simple templated fit is employed to estimate the
²¹¹⁴ SM background in higher b-tag multiplicities (3-4) from a region of a low number of
²¹¹⁵ reconstructed b-jets (0-2). As a proof-of-concept, the procedure after being shown to
²¹¹⁶ close in simulation, is applied to the SM enriched $\mu + \text{jets}$ control sample of the α_T
²¹¹⁷ all-hadronic search detailed in Chapter 4. To highlight the relative insensitivity of the
²¹¹⁸ choice of the b-tagging algorithm working points in the effectiveness of the procedure,
²¹¹⁹ results are presented using the CSV tagger (introduced in Section (3.3.2)) for the “Loose”,
²¹²⁰ “Medium” and “Tight” working points.

²¹²¹ 6.1. Concept

²¹²² The dominant SM backgrounds most SUSY searches are typically $t\bar{t} + \text{jets}$, $W + \text{jets}$ and
²¹²³ $Z \rightarrow \nu\bar{\nu} + \text{jets}$. These process are characterised by typically having zero or two underlying
²¹²⁴ b-quarks per event. The first step in this approach is to categorise two templates to be
²¹²⁵ 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.

- 2126 Thus two templates are defined, Z0 and Z2 ($t\bar{t}$ is combined with single top) which
 2127 represent processes which have an underlying b-quark content of zero or two respectively.
- 2128 Both these templates can be generated through the application of the relevant event
 2129 selection and taking the underlying n_b^{reco} distribution directly from simulation. However
 2130 as discussed within Section (4.5), there are large uncertainties for high n_b^{reco} multiplicities
 2131 due to limited MC statistics. This is particularly prominent for the Z0 templates, where
 2132 the number of reconstructed b-tags is driven primarily by the light-quark mis-tagging
 2133 rate. Therefore to improve the statistical precision of the predictions the formula method,
 2134 introduced in Section (4.5.1) is used.
- 2135 The generation of these templates is then dependant upon the jet-flavour content and
 2136 b-tagging rate within the phase space of interest, with the tagging probabilities of a jet
 2137 being a function of the jet p_T , the pseudo-rapidity $|\eta|$, and the jet-flavour. This can be
 2138 observed in Figure 6.1, where the b-tagging / c-quark mis-tagging / light mis-tagging
 2139 efficiency for the three working points of the CSV tagger is shown as a function of jet p_T .
- 2140 Before the templates are generated, the relevant jet p_T and η corrections are applied to
 2141 correct simulation to data, as specified in Section (4.5.3), to then determine the average
 2142 tagging rates per analysis bin.
- 2143 These two templates are then fit to data in the low n_b^{reco} region (0-2). The fit result is
 2144 used, along with the knowledge of the template shapes, to extrapolate an estimate to the
 2145 high n_b^{reco} signal region (3,4), which is then compared to what is observed in data.
- 2146 This method can, in principle, be applied to any analysis where the signal hypothesis
 2147 has a larger underlying b-quark spectra than the SM backgrounds, as it solely relies on
 2148 fitting to the shape of the n_b^{reco} distribution.

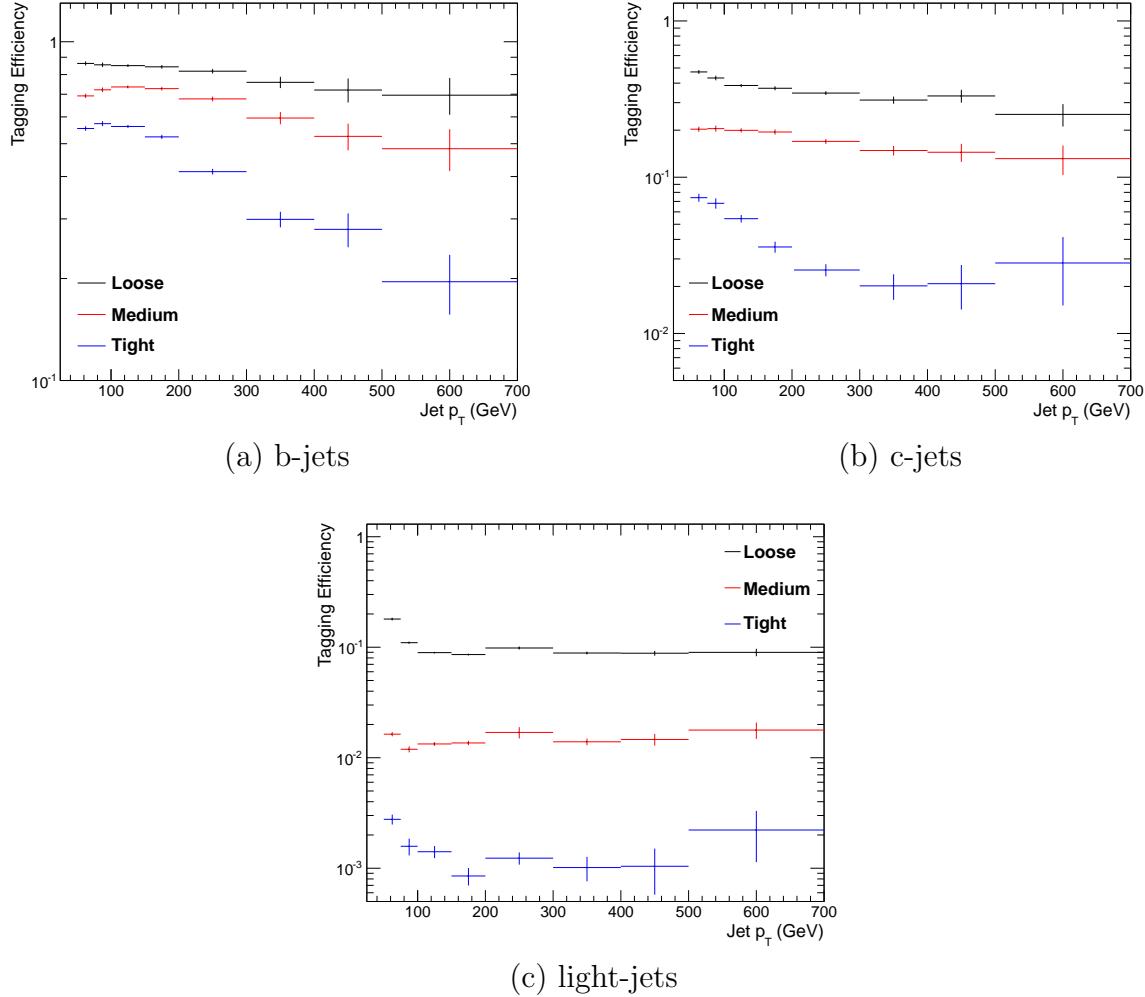


Figure 6.1.: The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging rate (c) as measured in simulation after the α_T analysis, $\mu + \text{jets}$ control sample selection in the region $H_T > 375$.

2149 6.2. Application to the α_T Search

2150 As detailed in the previous chapter, the α_T analysis is a search for **SUSY** particles
 2151 in all-hadronic final states, utilising the kinematic variable α_T to suppress QCD to a
 2152 negligible level. **SM** enriched control samples are used to estimate the background within
 2153 an all-hadronic signal region.

2154 The selection for the $\mu + \text{jets}$ control samples defined in Section (4.2.3) is used to
 2155 demonstrate the template fitting procedure both conceptually in simulation, and also
 2156 when applied in data. This is chosen, as such a selection is dominated by events stemming
 2157 from the **SM** processes with little or no signal contamination from potential new physics..

2158 Neither are contributions from rate **SM** processes with a higher underlying b-quark
2159 content (e.g. $t\bar{t}b\bar{b}$) expected. For these reasons, there is a degree of confidence that the
2160 procedure should close when applied to this phase space.

2161 The analysis presented here is binning in source jet multiplicity bins, of 3,4 and ≥ 5
2162 reconstructed jets per event (di-jet events are not included as there is no contribution
2163 to the high n_b^{reco} region (3,4)) , in order to reduce the kinematic jet p_T dependence.
2164 Furthermore the analysis is split into three H_T regions,

2165 • 275-325 GeV

2166 • 325-375 GeV

2167 • > 375 GeV

2168 contrary to the eight used within the α_T analysis. Templates for both underlying b-quark
2169 content hypotheses are then generated for the nine defined analysis bins.

2170 6.2.1. Proof of principle in simulation

2171 In order to demonstrate that the template procedure produces accurate predictions
2172 within simulation, the simulation samples in the analysis are firstly split into two to allow
2173 for statistically independent fits to be performed.

2174 By combining the relevant ingredients necessary to employ the formula method, n_b^{reco}
2175 templates for $Z = 0$ and $Z= 2$ are generated individually for each n_{jet} and H_T bin using
2176 one half of each simulation sample. A fit of these two templates is then performed in the
2177 low n_b^{reco} (0-2) region, back to the sum of the other halves of each simulation sample in
2178 order to check that the relevant information can be recovered in the n_b^{reco} signal region
2179 (3-4).

2180 The fits are performed independently within each of the defined analysis bins to reduce the
2181 dependence of the shapes of these distributions on simulation. The half of the simulation
2182 sample for which the templates are fitted too, are taken directly from simulation, extending
2183 this procedure to also be a validation of the formula method to accurately estimate
2184 the n_b^{reco} distribution. Additionally as this test is performed in simulation, the relevant
2185 corrections of the b-tagging rates between data and simulation are *not* applied.

2186 Within Figure 6.2, the results of this fitting procedure is shown for each **CSV** working
2187 point. Results are presented for the $n_{jet} \geq 5$ category, using the $\mu +$ jets control sample

2188 selection in the inclusive $H_T > 375$ GeV analysis bin. The grey bands represent the
 2189 statistical uncertainty on the template shapes. Additional fits are shown for other n_{jet}
 2190 category within Appendix D.1.

2191 Furthermore the extrapolated fit predictions within the high n_b^{reco} signal region, are
 2192 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 ± 14.8	387.9 ± 10.2	794.1 ± 14.34
Template $n_b = 3$	820.4 ± 26.7	376.3 ± 11.9	780.1 ± 15.1
Simulation $n_b = 4$	68.2 ± 3.9	27.6 ± 2.7	91.28 ± 4.9
Template $n_b = 4$	72.5 ± 4.7	28.25 ± 2.34	84.4 ± 3.8
Medium working point			
Simulation $n_b = 3$	133.7 ± 5.7	74.5 ± 4.5	164.2 ± 6.4
Template $n_b = 3$	132.8 ± 4.8	74.5 ± 3.9	159.9 ± 5.7
Simulation $n_b = 4$	1.6 ± 0.6	0.6 ± 0.4	3.4 ± 0.9
Template $n_b = 4$	1.8 ± 0.2	1.1 ± 0.2	4.1 ± 0.4
Tight working point			
Simulation $n_b = 3$	26.9 ± 2.6	13.9 ± 1.9	31.8 ± 2.9
Template $n_b = 3$	24.7 ± 1.5	13.8 ± 1.2	28.1 ± 1.5
Simulation $n_b = 4$	0.5 ± 0.4	-	-
Template $n_b = 4$	0.1 ± 0.1	0.1 ± 0.1	0.2 ± 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 fb^{-1} . The uncertainties quoted on the template yields are purely statistical.

2193 The pull distributions for all the fits performed are compatible with a mean of zero and
 2194 standard distributions, see Appendix D.2.

2195 The good overall agreement summarised in the table validates both the formula method
 2196 used to generate the templates as well as the fitting method itself. The application of
 2197 this method to the same selection in data is used to demonstrate necessary control over
 2198 the efficiency and mis-tagging rates.

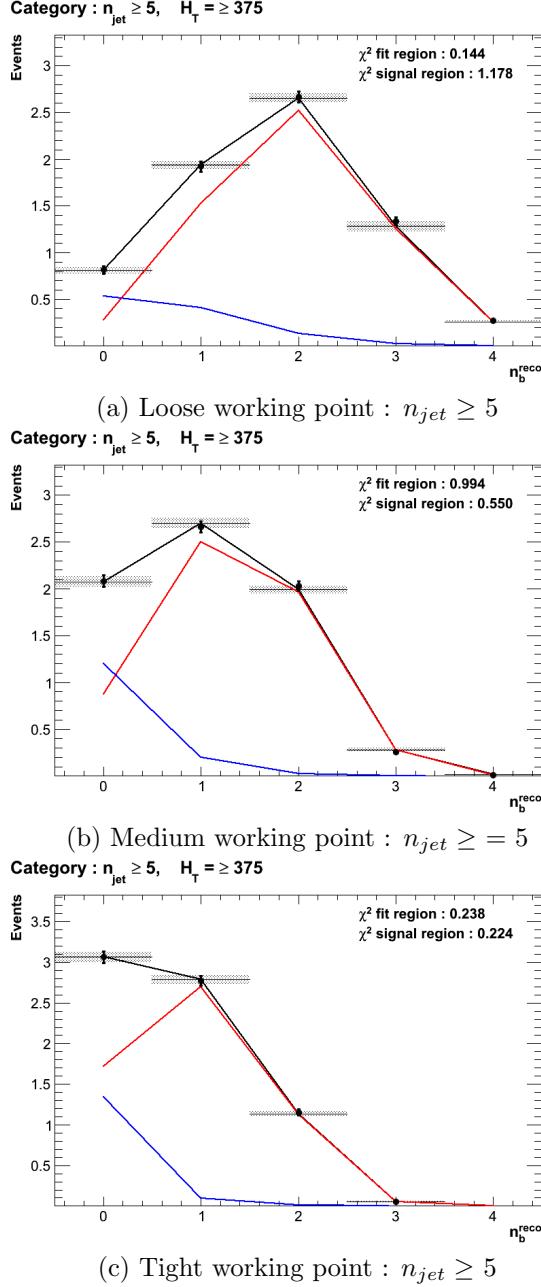


Figure 6.2.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} \geq 5$ category. The blue template represents $Z = 0$, while the red template represents $Z = 2$. Grey bands represent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

²¹⁹⁹ 6.2.2. Results in a data control sample

- ²²⁰⁰ The method above is now applied to the 2012 8 TeV dataset in the $\mu +$ jets control
²²⁰¹ sample, to establish the validity of this method in data. The relevant data to simulation
²²⁰² scale factors are applied to get corrected values of the efficiency and mis-tagging rates
²²⁰³ measured in data [64] [85].
- ²²⁰⁴ Figure 6.3 show the the results of the templates derived from simulation to each of the
²²⁰⁵ three defined H_T bins, in the $n_{jet} \geq 5$ category for the medium working point CSV
²²⁰⁶ tagger (the same working point used within the α_T analysis). Grey bands represent the
²²⁰⁷ statistical uncertainty of the fit combined in quadrature with the systematic uncertainties
²²⁰⁸ of varying the data to simulation scale factors up and down by their measured systematic
²²⁰⁹ uncertainties. Additional fit results for other jet multiplicities are found in Appendix D.3
- ²²¹⁰ The numerical results and extrapolation to the $n_b^{reco} = 3, 4$ bins for all H_T and working
²²¹¹ 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 ± 16.7	372.1 ± 10.1	673.2 ± 14.1
Data $n_b = 4$	81	43	81
Template $n_b = 4$	74.5 ± 2.3	27.6 ± 1.2	71.6 ± 2.6
Medium working point			
Data $n_b = 3$	137	79	152
Template $n_b = 3$	131.2 ± 2.3	65.1 ± 1.7	127.8 ± 2.4
Data $n_b = 4$	1	1	3
Template $n_b = 4$	1.8 ± 0.1	0.9 ± 0.1	3.1 ± 0.1
Tight working point			
Data $n_b = 3$	24	15	25
Template $n_b = 3$	23.0 ± 0.4	10.9 ± 0.3	20.3 ± 0.5
Data $n_b = 4$	0	0	1
Template $n_b = 4$	0.1 ± 0.1	0.1 ± 0.1	0.2 ± 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 fb^{-1} of data at $\sqrt{s} = 8\text{TeV}$. The uncertainties quoted on the template yields are purely statistical.

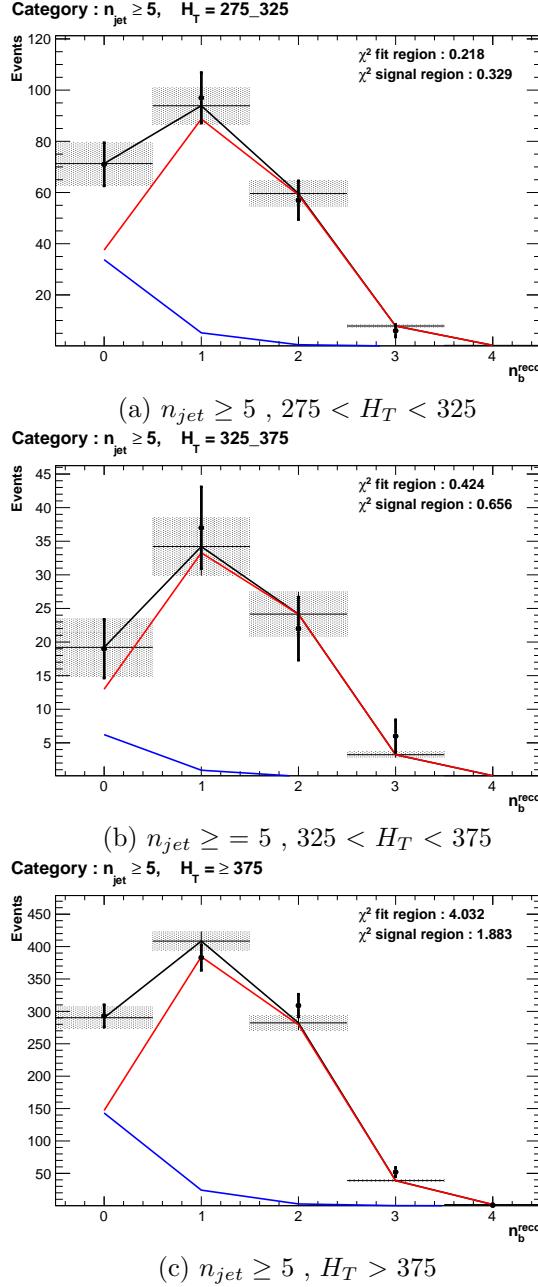


Figure 6.3.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from data, for the $n_{jet} \geq 5$ category and medium CSV working point. The blue template represents $Z = 0$, while the red template represents $Z = 2$. The χ^2 parameter displayed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

2212 The agreement for all working points demonstrates a good control of the b-tagging
 2213 efficiency and mis-tagging rates and gives confidence in the method outlined.

2214 6.2.3. Application to the α_T hadronic search region

2215 As an accompaniment to the background estimation methods outlined by the α_T search.
 2216 The b-tag template method offers a complimentary way of testing the SM only background
 2217 of the hadronic signal region of the search. In the absence presence of a natural SUSY
 2218 signature, containing four underlying b or t quarks, the number of reconstructed three/four
 2219 b-tagged events will be enhanced.
 2220 Figure 6.4 show the the results of the templates derived from simulation to each of
 2221 the three CSV working points, in the $n_{jet} \geq 5$, $H_T > 375$ GeV category. Grey bands
 2222 represent the statistical uncertainty of the fit combined in quadrature with the systematic
 2223 uncertainties of varying the data to simulation scale factors up and down by their
 2224 measured systematic uncertainties. Additional fit results for other jet multiplicities are
 2225 found in Appendix D.4
 2226 The numerical results and extrapolation to the $n_b^{reco} = 3, 4$ bins for all H_T and working
 2227 points is shown in Table 6.4. The excess of data is found and predictions from this
 2228 method are found to be compatible with the α_T maximum likelihood fit results from
 2229 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 ± 9.0	103.4 ± 5.9	124.98 ± 7.4
Data $n_b = 4$	15	9	16
Template $n_b = 4$	15.9 ± 1.2	8.05 ± 0.9	13.1 ± 1.3
Medium working point			
Data $n_b = 3$	28	15	12
Template $n_b = 3$	24.4 ± 0.9	12.7 ± 0.8	19.9 ± 2.4
Data $n_b = 4$	1	0	2
Template $n_b = 4$	0.3 ± 0.2	0.3 ± 0.1	0.5 ± 0.1
Tight working point			
Data $n_b = 3$	5	2	0
Template $n_b = 3$	4.03 ± 0.2	2.4 ± 0.2	3.1 ± 0.2
Data $n_b = 4$	1	0	0
Template $n_b = 4$	0.1 ± 0.1	0.1 ± 0.1	0.0 ± 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 fb^{-1} of data at $\sqrt{s} = 8\text{TeV}$. The uncertainties quoted on the template yields are purely statistical.

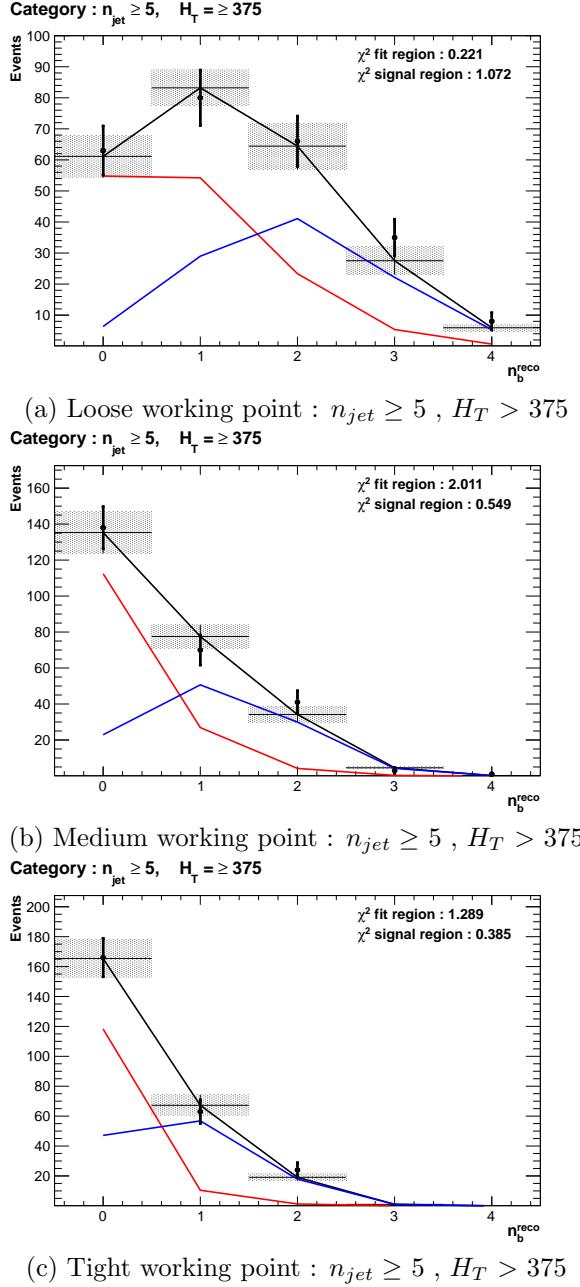


Figure 6.4.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, in the $n_{jet} \geq 5$ and $H_T > 375$ category for all CSV working points. The blue template represents $Z = 0$, while the red template represents $Z = 2$. The χ^2 parameter displayed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

2230 **6.3. Summary**

2231 A **SUSY** signature such as one from gluino-induced third-generation squark production,
2232 would result in a final state with an underlying b-quark content greater than two. In
2233 order to be able to discriminate such signatures from the **SM** background, templates are
2234 generated based on a parameterisation of the number of the **SM** processes, where the
2235 underlying b-quarks per event is typically zero or two. These templates are then fit to
2236 data in a low n_b^{reco} (0-2) control region in order to extrapolate a prediction in a high
2237 n_b^{reco} (3-4) signal region.
2238 The method was demonstrated both in simulation and also in data, using the **SM** enriched
2239 $\mu + \text{jets}$ selection from the α_T search, to prove conceptually and experimentally that the
2240 method works and there is adequate control over the efficiency and mis-tagging rates in
2241 data for all working points of the **CSV** tagger. Additionally this method was also applied
2242 to the α_T analysis signal region where good agreement is observed between data and the
2243 background estimation method of the α_T analysis.

Chapter 7.

²²⁴⁴ Conclusions

²²⁴⁵ A search for supersymmetry is presented based on a data sample of pp collisions collected
²²⁴⁶ at $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 11.7 ± 0.5 fb⁻¹. Final
²²⁴⁷ states with two or more jets and significant \cancel{E}_T , a typical final state topology of R-parity
²²⁴⁸ conserving SUSY models have been analysed. The sum of standard model backgrounds
²²⁴⁹ per bin are estimated from a simultaneous binned likelihood fit to hadronic, $\mu +$ jets,
²²⁵⁰ $\mu\mu +$ jets, and $\gamma +$ jets samples. Systematic errors due to theory, detector effects and
²²⁵¹ analysis choices are quantified through the use of data driven closure tests and accounted
²²⁵² for in the final interpretation.

²²⁵³ No excess of events is observed over the expected SM background. The analysis is
²²⁵⁴ further interpreted in a set of SMS models, with a special emphasis on third generation
²²⁵⁵ squarks and compressed spectra scenarios. In the considered models with gluino pair
²²⁵⁶ production and for small LSP masses, exclusion limits of the gluino mass are in the range
²²⁵⁷ 950-1125 GeV. For SMS with squark pair production, first or second generation squarks
²²⁵⁸ are excluded up to around 775 GeV and bottom squarks are excluded up to 600 GeV,
²²⁵⁹ again for small LSP masses.

²²⁶⁰ A complementary approach using a templated method to estimate the b-tag jet distribu-
²²⁶¹ tion of SM processes, is used to search for gluino induced third generation squark SUSY
²²⁶² production. The α_T analysis is used to demonstrate conceptually and experimentally
²²⁶³ this technique in the $\mu +$ jets control sample. This method is further applied to the
²²⁶⁴ α_T hadronic search region where good agreement is observed between the data and the
²²⁶⁵ background estimation procedure of the α_T analysis.

²²⁶⁶ The performance of the Level-1 trigger for jets and energy sum quantities is also presented.
²²⁶⁷ These studies quantify any change in level-1 performance after the introduction of a 5

₂₂₆₈ GeV jet seed threshold into the jet algorithm configuration. No significant change in
₂₂₆₉ single jet trigger efficiencies is observed and good performance is observed for a range of
₂₂₇₀ level-1 quantities.

₂₂₇₁

Appendix A.

²²⁷² Miscellaneous

²²⁷³ A.1. Noise Filters

²²⁷⁴ For Calo jets the following criteria were applied:

Loose CaloJet Id	
Variable	Definition
$f_{HPD} < 0.98$	Fraction of jet energy contributed from “hottest” HPD , which rejects HCAL noise.
$f_{EM} > 0.01$	Noise from the HCAL 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.

²²⁷⁵ For PF jets the following criteria were applied:

Loose PF jet Id	
Variable	Definition
$\text{nfhJet} < 0.99$	Fraction of jet composed of neutral hadrons. HCAL noise tends to populate high values of neutral hadron fraction.
$\text{nemfJet} < 0.99$	Fraction of jet composed of neutral electromagnetic energy. ECAL noise tends to populate high values of neutral EM fraction.
$\text{nmultiJet} > 1$	Number of constituents that jet is composed from.
$\text{chfJet} > 0$	Fraction of jet composed of charged hadrons.
$\text{cmultiJet} > 0$	Number of charged particles that compose jet.
$\text{cemfJet} < 0.99$	Fraction of jet composed of charged electromagnetic energy.

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

²²⁷⁶ The following noise filters are applied, to remove events with spurious, non-physical jets
²²⁷⁷ or missing transverse energy.

Noise Filters	
Variable	Definition
CSC tight beam halo filter	As proton beams circle the LHC , proton interactions with the residual gas particles or the beam collimators can occur, producing showers of secondary particles which can interact with the CMS detector.
HBHE noise filter with isolated noise rejection	Anomalous noise in the HCAL not due to electronics noise, but rather due to instrumentation issues associated with the HPD 's and Readout Boxes (RBXs).
HCAL laser filter	The HCAL 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	EB and EE have single noisy crystals which are masked in reconstruction. Use the Trigger Primitive (TP) information to assess how much energy was lost in masked cells.
Bad EE Supercrystal filter	Two supercrystals in EE 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.

2278 A.2. Primary Vertices

2279 The pileup per event is defined by the number of 'good' reconstructed primary vertices
2280 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 HLT 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.

2281 L1 Jets

2282 B.1. Jet matching efficiencies

2283 The single jet turn-on curves are derived from events independent of whether the leading
2284 jet in an event is matched to a Level 1 jet using ΔR matching detailed in Section (3.4.3)
2285 or not. These turn-ons are produced from events which are not triggered on jet quantities
2286 and therefore it is not guaranteed that the lead jet of an event will be seeded by a Level
2287 1 jet. Figure B.1 shows the particular matching efficiency of a lead jet to a L1 jet.

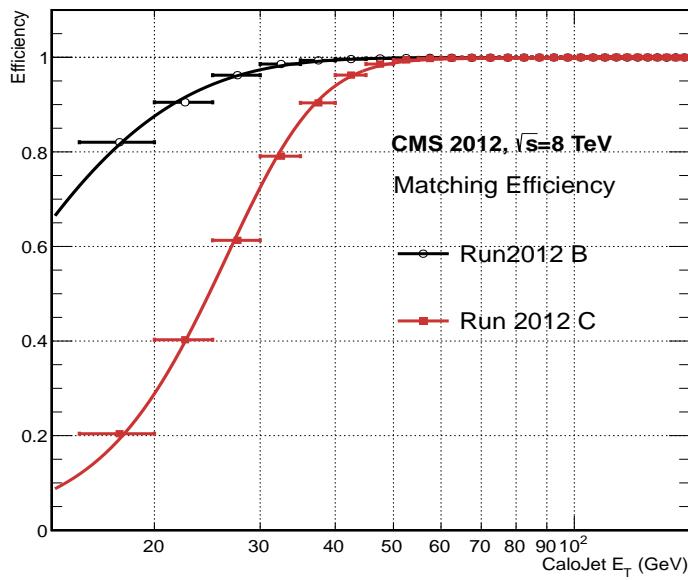


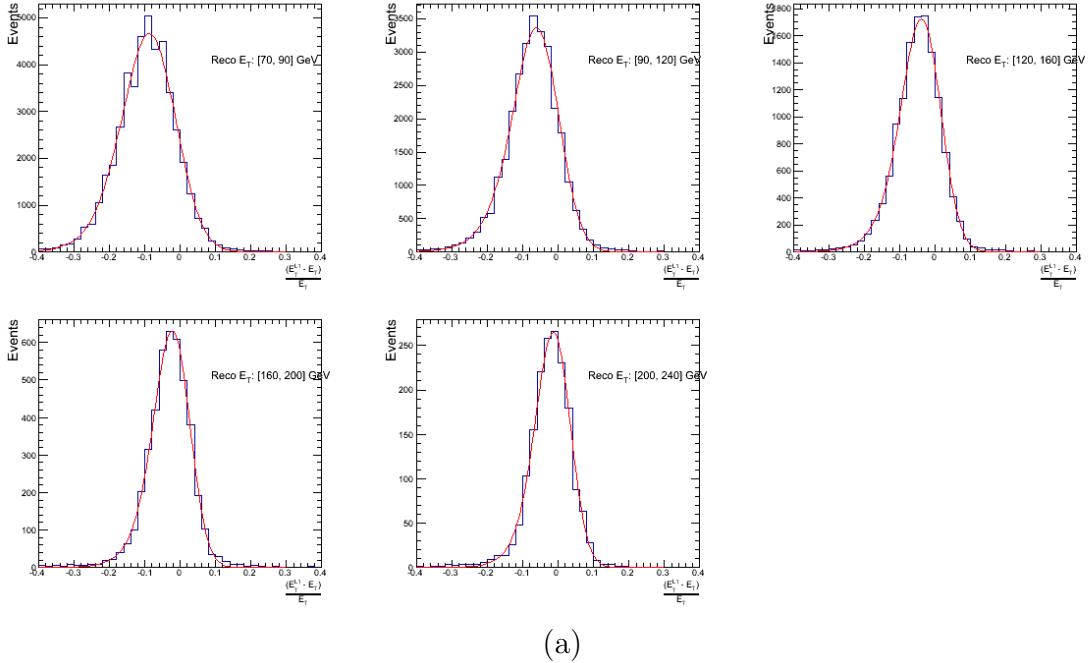
Figure B.1.: Leading jet matching efficiency as a function of the offline CaloJet E_T , measured in an isolated muon triggered dataset in the 2012B and 2012C run periods.

Run Period	μ	σ
2012B	6.62 ± 0.01	0.79 ± 0.03
2012C	19.51 ± 0.03	7.14 ± 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, μ , and resolution, σ , are measured with respect to offline Calo Jet E_T .

- 2288 It can be seen that the turn on is sharper during the 2012B run period. The seed
 2289 threshold requirement of a 5 GeV jet seed in run 2012C results in more events in which
 2290 even the lead offline jet does not have an associated L1 jet. For larger jet E_T thresholds,
 2291 typical of thresholds used in physics analyses, 100% efficiency is observed.
 2292 The matching efficiencies have a μ values of 6.62 GeV and 19.51 GeV for Run 2012B
 2293 and 2012C respectively and is shown in Table B.1.

2294 B.2. Leading Jet Energy Resolution



(a)

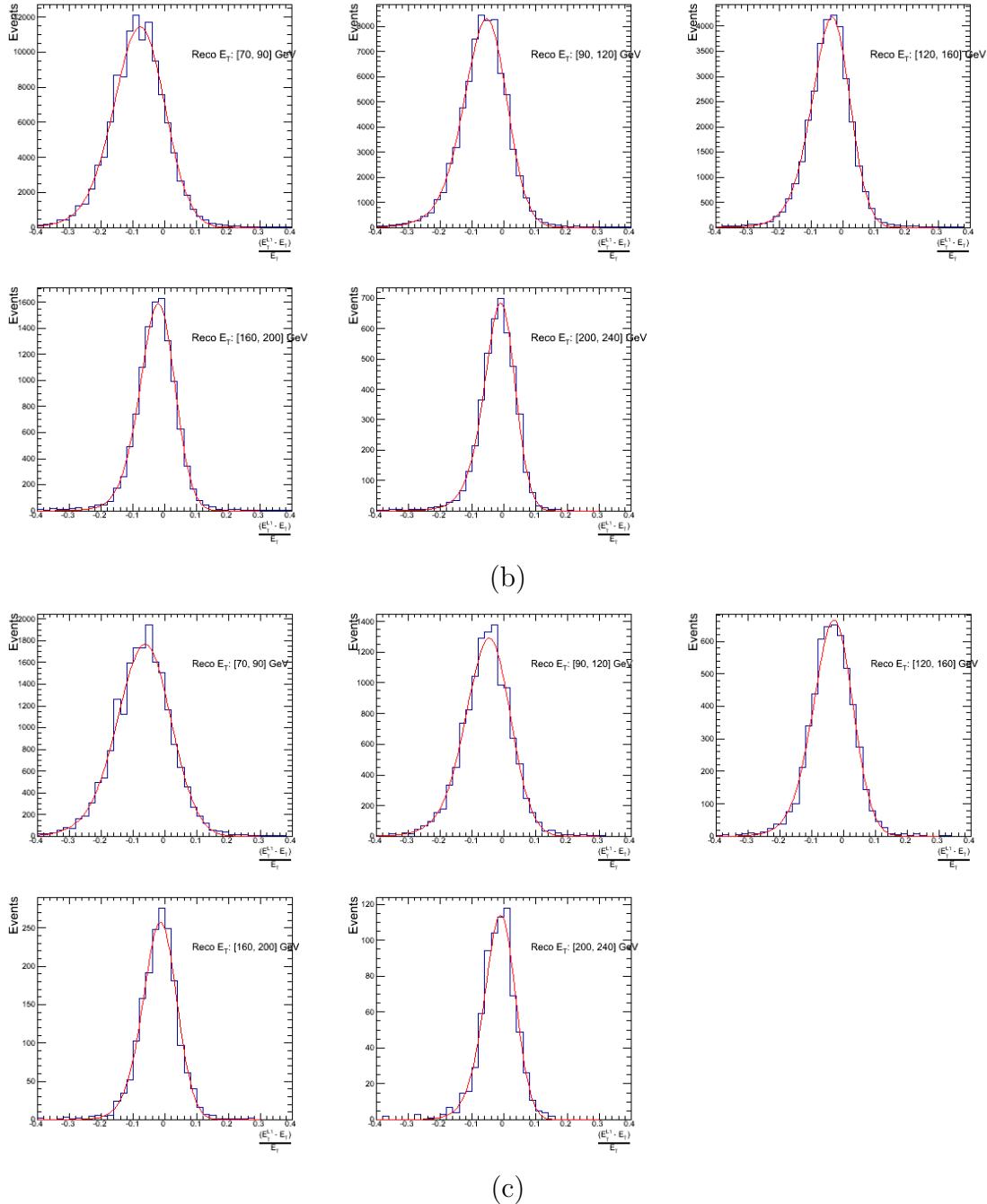
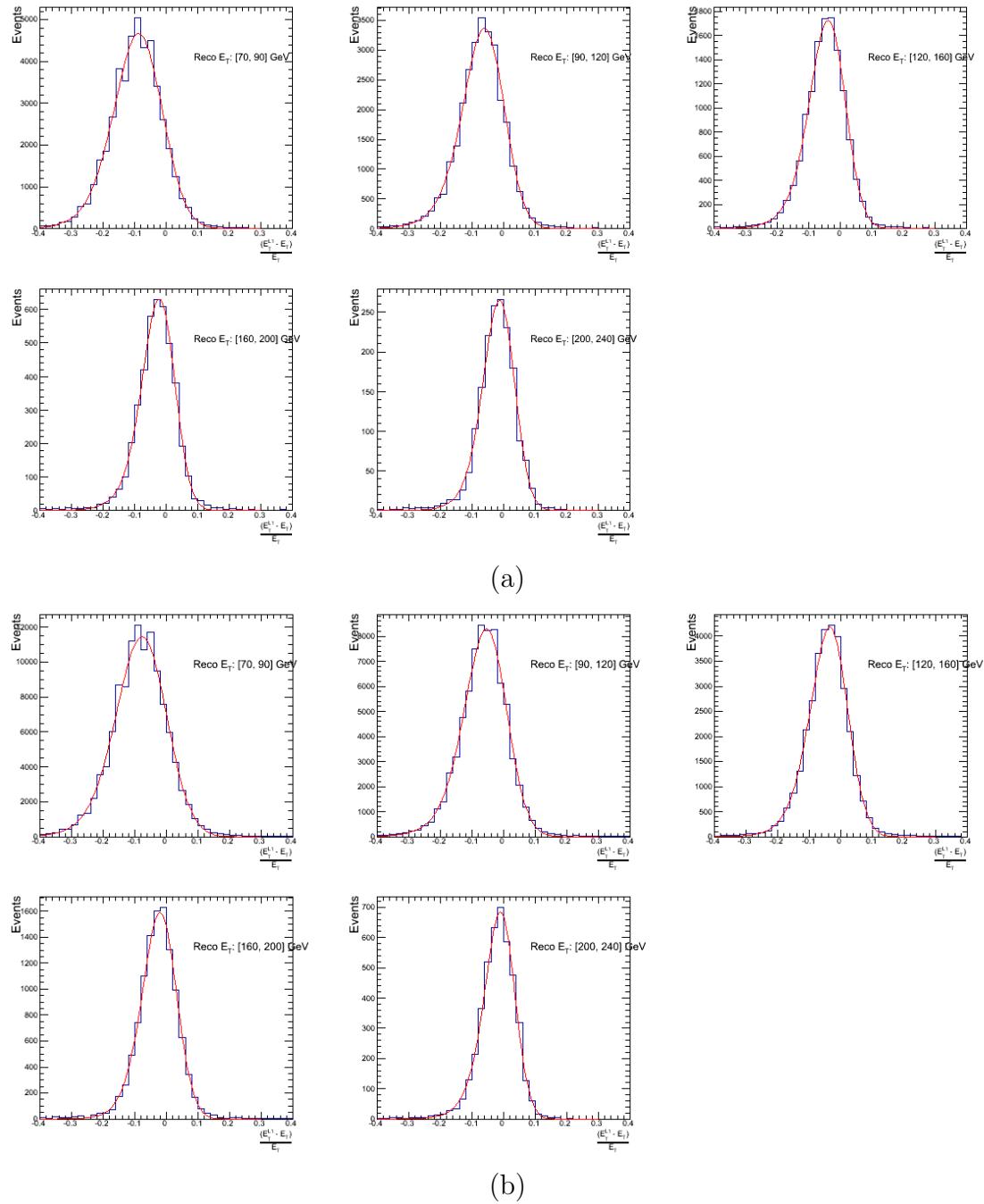
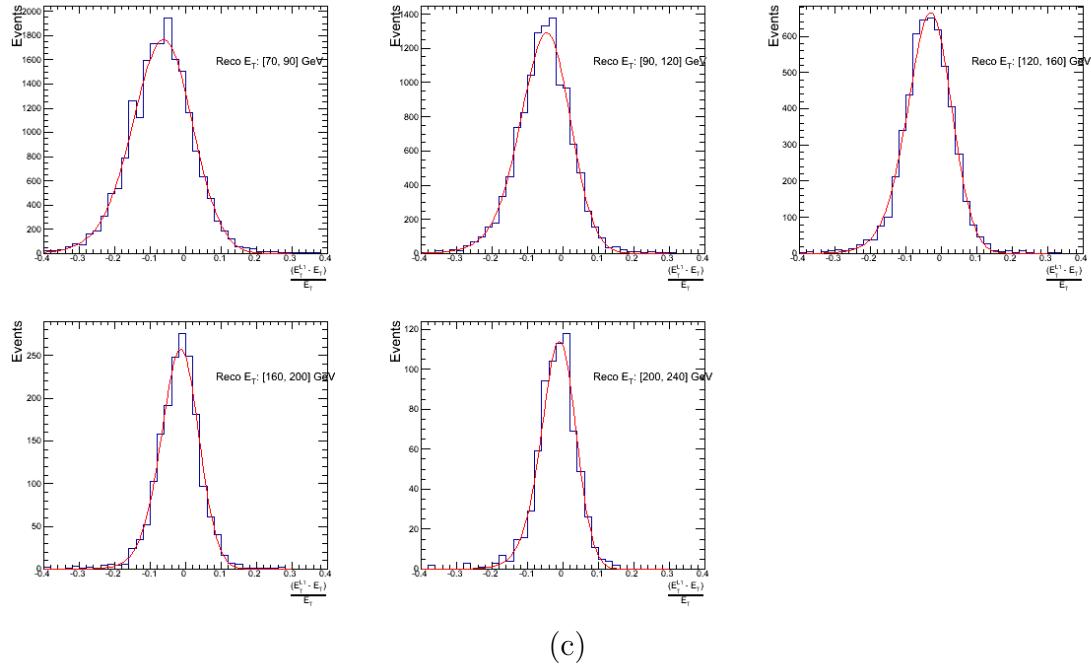


Figure B.2.: Resolution plots of the leading offline jet Calo E_T measured as a function of $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low (a), medium (b) and high (c) pile-up conditions.





(c)

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

2295 B.3. Resolution for Energy Sum Quantities

2296 The following plots show the resolution parameters for the four energy sum quantities as
 2297 a function of the quantity (q) itself. In this case, The mean and RMS of the individual
 2298 $\frac{(L1\ q - Offline\ q)}{Offline\ q}$ distributions, in bins of the quantity q is displayed.

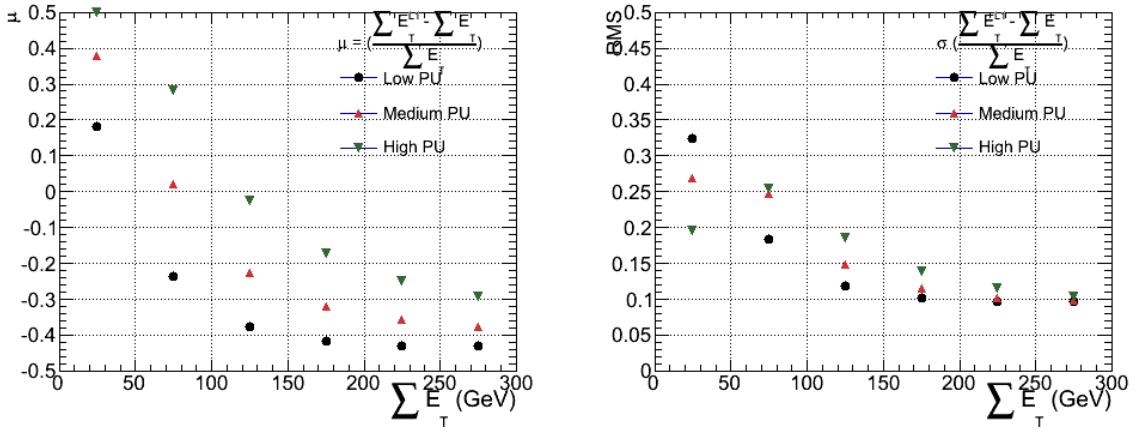


Figure B.4.: $\sum E_T$ resolution parameters in bins of Calo $\sum E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

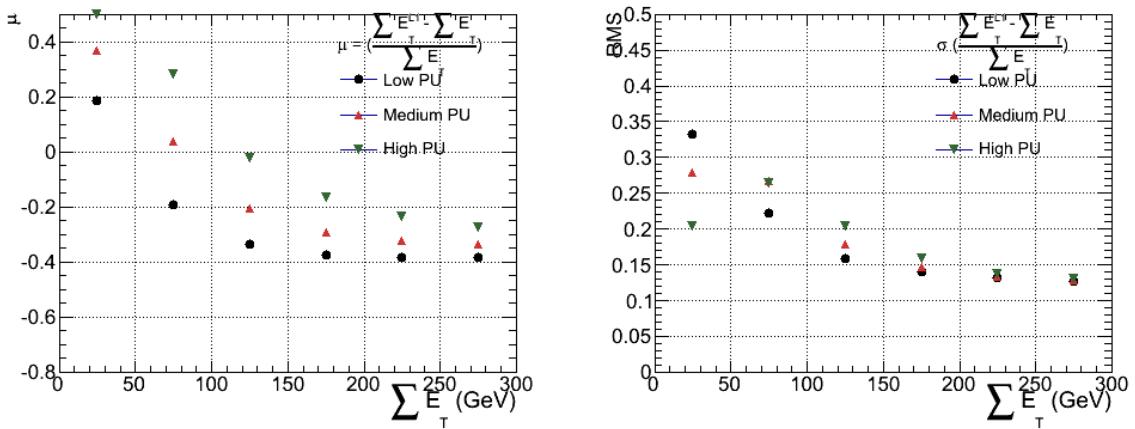


Figure B.5.: $\sum E_T$ resolution parameters in bins of PF $\sum E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

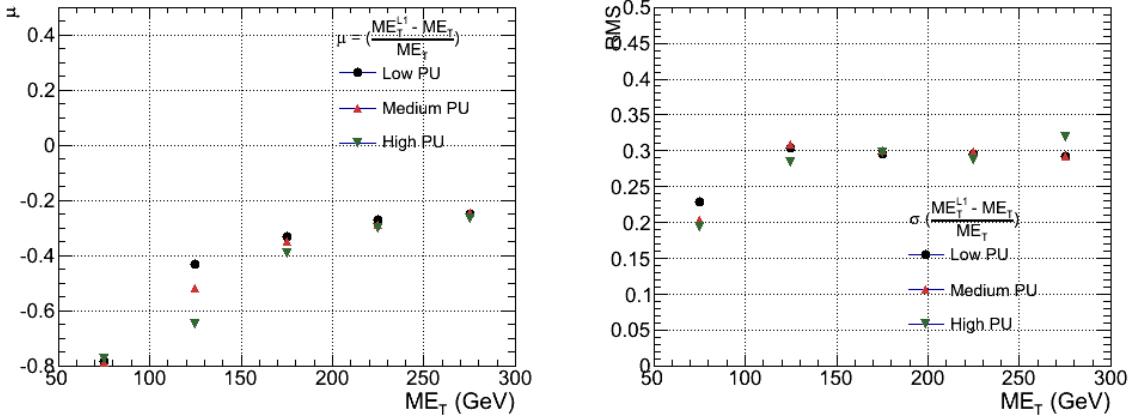


Figure B.6.: \mathcal{E}_T resolution parameters in bins of Calo \mathcal{E}_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

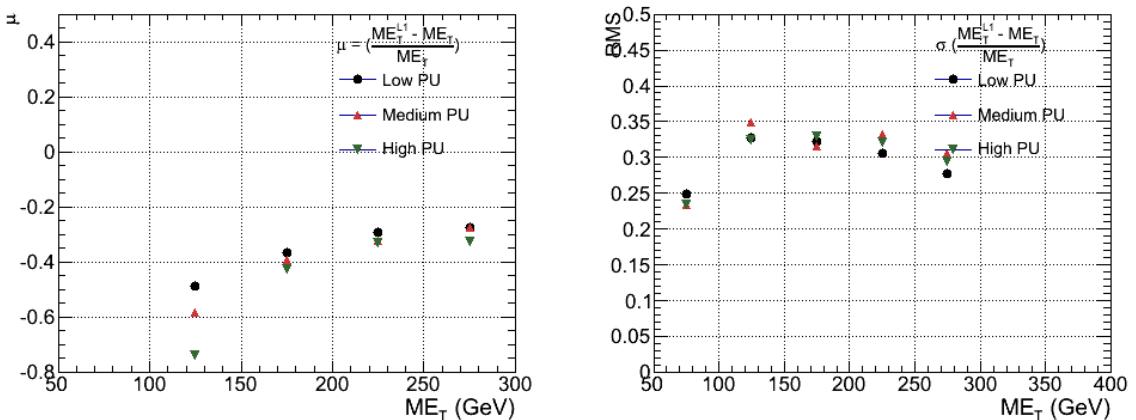


Figure B.7.: \mathcal{E}_T resolution parameters in bins of PF \mathcal{E}_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

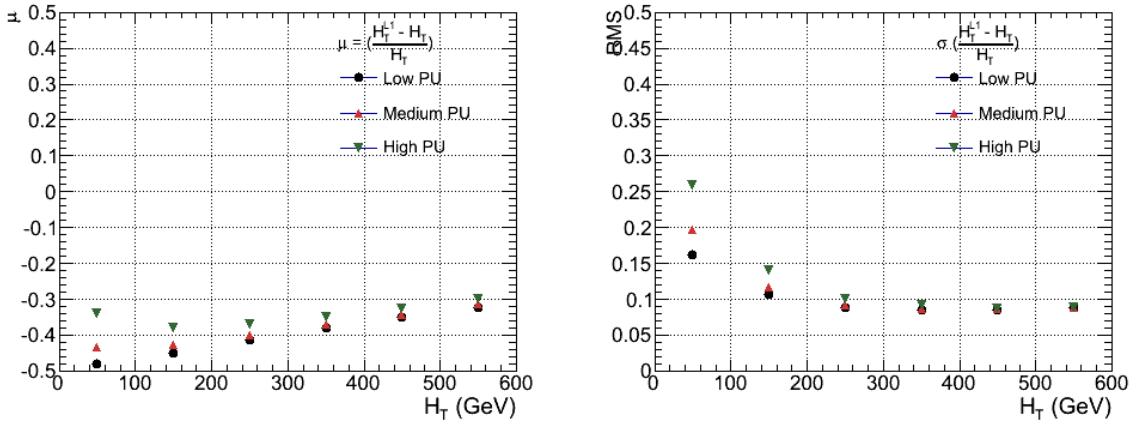


Figure B.8.: H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

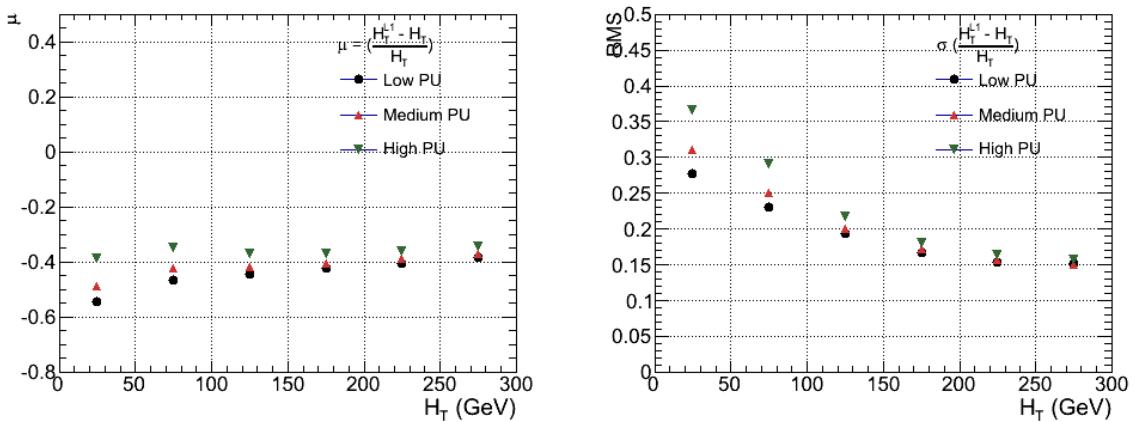


Figure B.9.: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

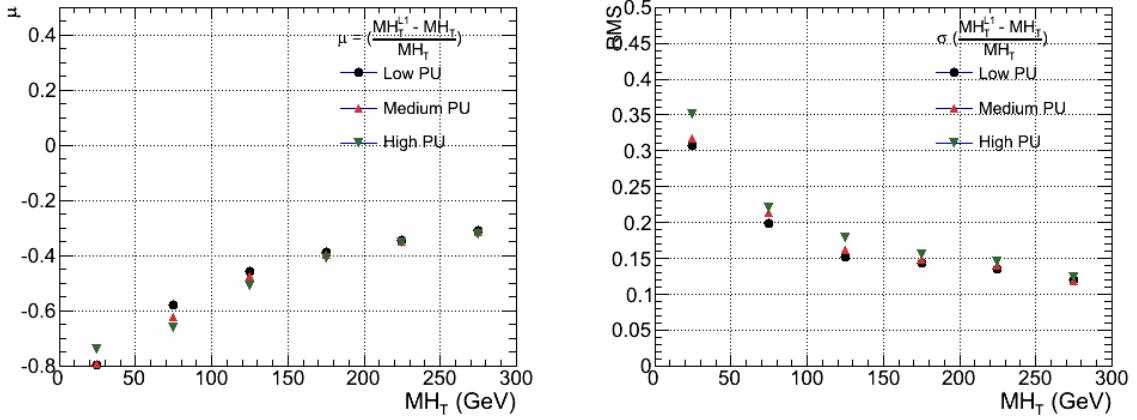


Figure B.10.: H_T resolution parameters in bins of H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

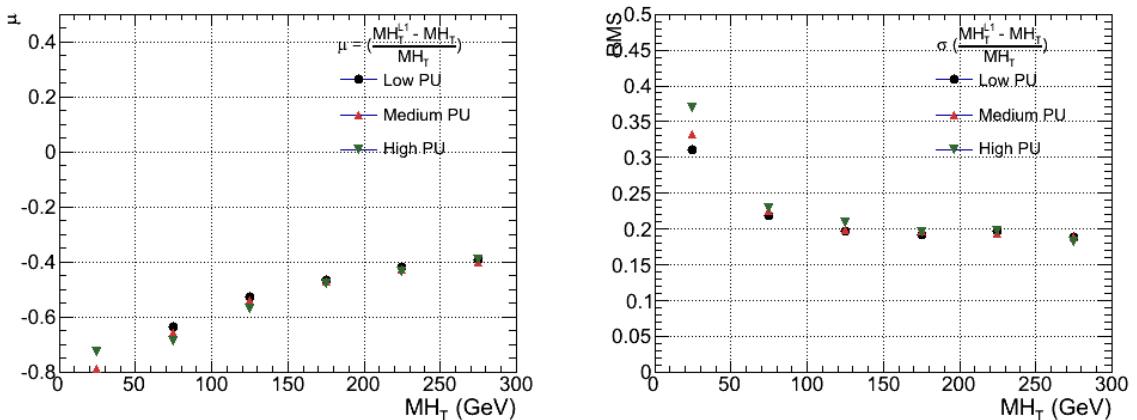


Figure B.11.: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

Appendix C.

²²⁹⁹ Additional material on background ²³⁰⁰ estimation methods

²³⁰¹ C.1. Determination of k_{QCD}

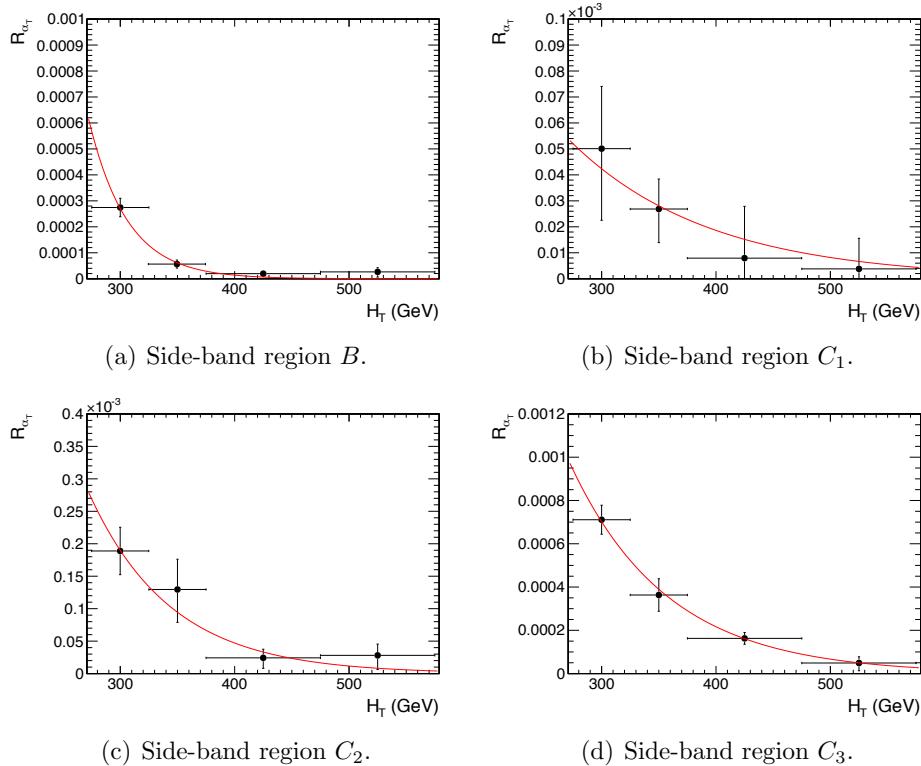


Figure C.1.: $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is conducted between the H_T region $275 < H_T < 575$.

2303 **C.2. Effect of varying background cross sections on**
2304 **closure tests**

2305 Closure tests with cross section variations of +20% and -20% applied to $W + \text{jets}$ and $t\bar{t}$
2306 processes respectively.

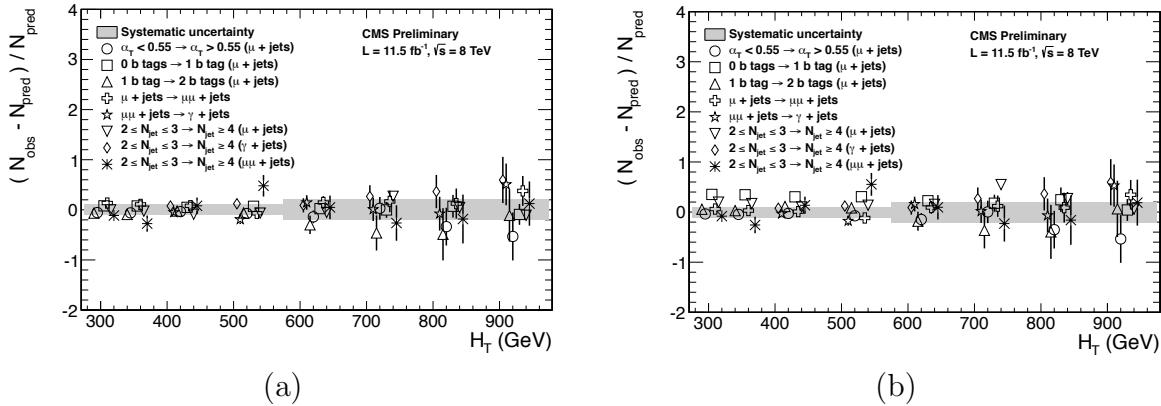


Figure C.2.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins: (a) $2 \leq n_{jet} \leq 3$ and (b) $n_{jet} \geq 4$.

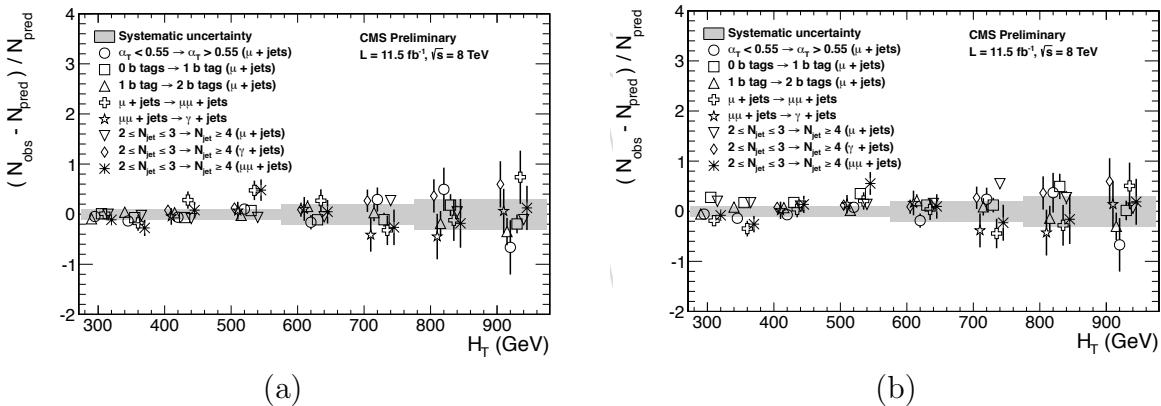


Figure C.3.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins: (a) $2 \leq n_{jet} \leq 3$ and (b) $n_{jet} \geq 4$.

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

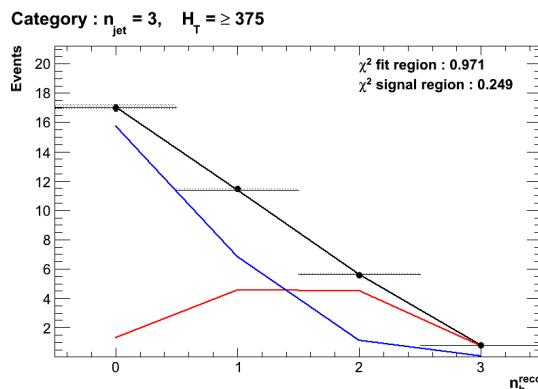
Table C.1.: Translation factors constructed from the $\mu +$ jets control sample and signal selection MC, to predict yields for the $W +$ jets and $t\bar{t}$ back-grounds in the signal region with (a) NNLO cross sections corrected by k-factors determined from a data sideband see Section (4.4), marked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

²³⁰⁷ Additional Material For B-tag ²³⁰⁸ Template Method

²³⁰⁹ D.1. Templates Fits in Simulation

²³¹⁰ Template fits for the three **CSV** working points in the $n_{jet} = 3, H_T > 375$ category :



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

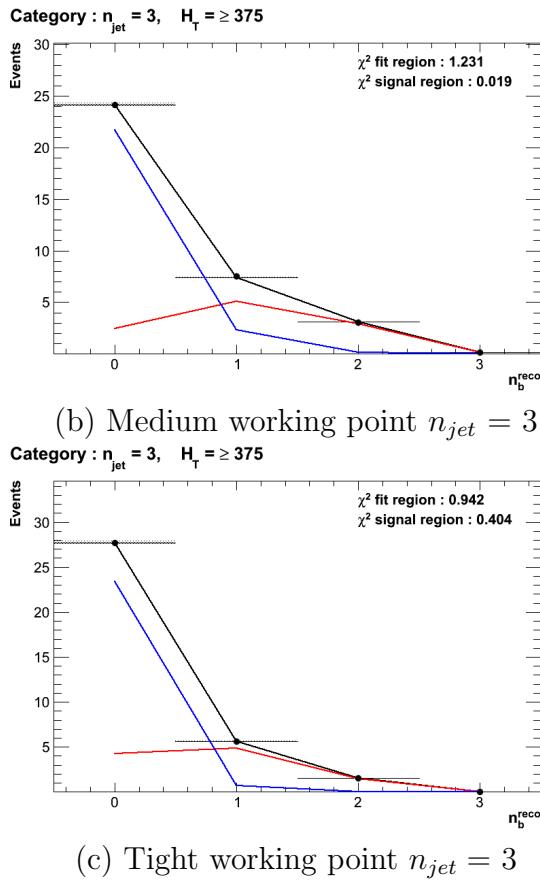
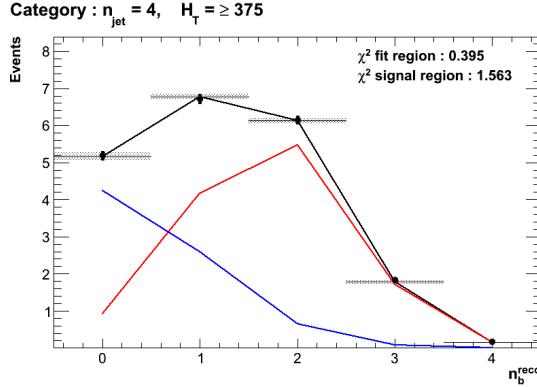
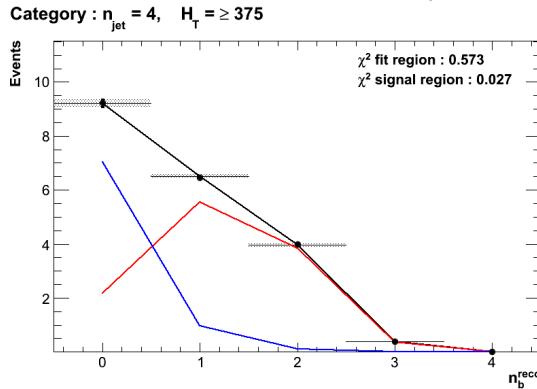


Figure D.1.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 3$ category. The blue template represents $Z = 0$, while the red template represents $Z = 2$. Grey bands represent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

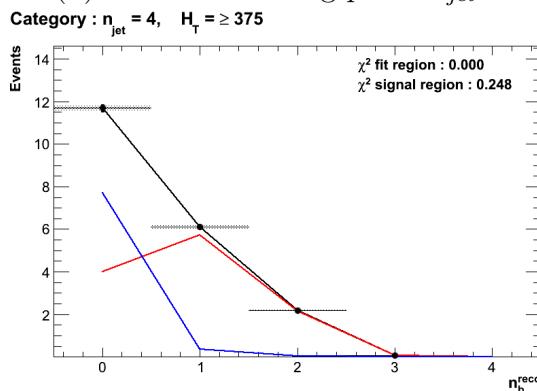
2311 Template fits for the three **CSV** working points in the $n_{jet} = 4, H_T > 375$ category :



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



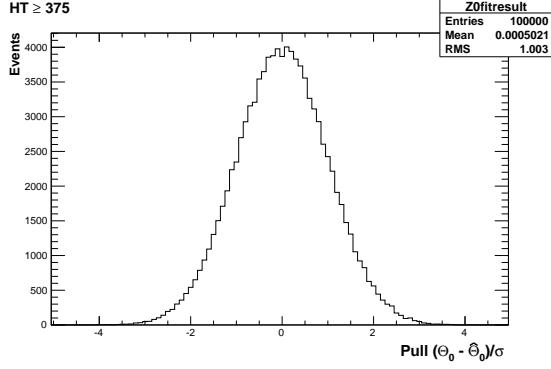
(b) Medium working point $n_{jet} = 4$



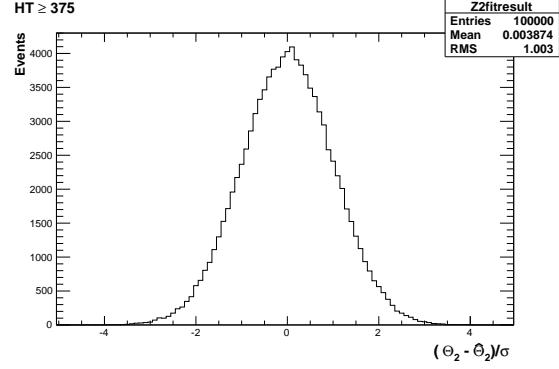
(c) Tight working point $n_{jet} = 4$

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 χ^2 parameter displayed represents the goodness of fit to the low n_b^{reco} (0-2) control region.

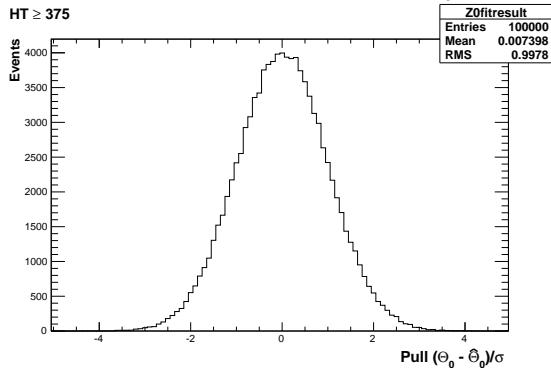
2312 D.2. Pull Distributions for Template Fits



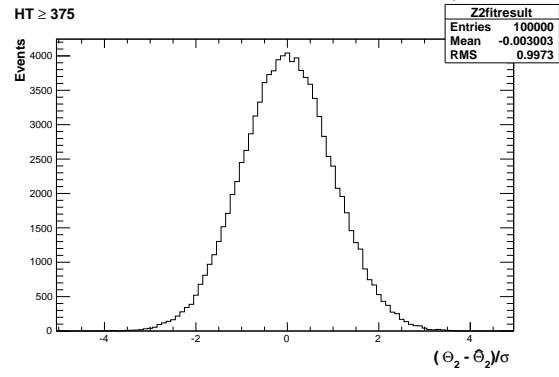
(a) Z0 Template, $H_T > 375$, $n_{jet} = 3$



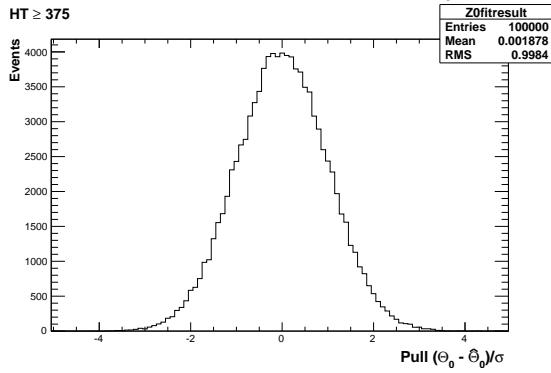
(b) Z2 Template, $H_T > 375$, $n_{jet} = 3$



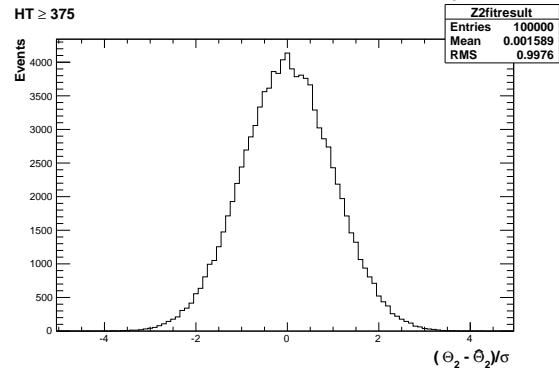
(a) Z0 Template, $H_T > 375$, $n_{jet} = 4$



(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 σ . Distributions are shown for both Z0 and Z2 templates for the medium CSV working point.

₂₃₁₃ **D.3. Templates Fits in Data Control Sample**

₂₃₁₄ Template fits for the three H_T bins, in the $n_{jet} = 3$, medium **CSV** working point:

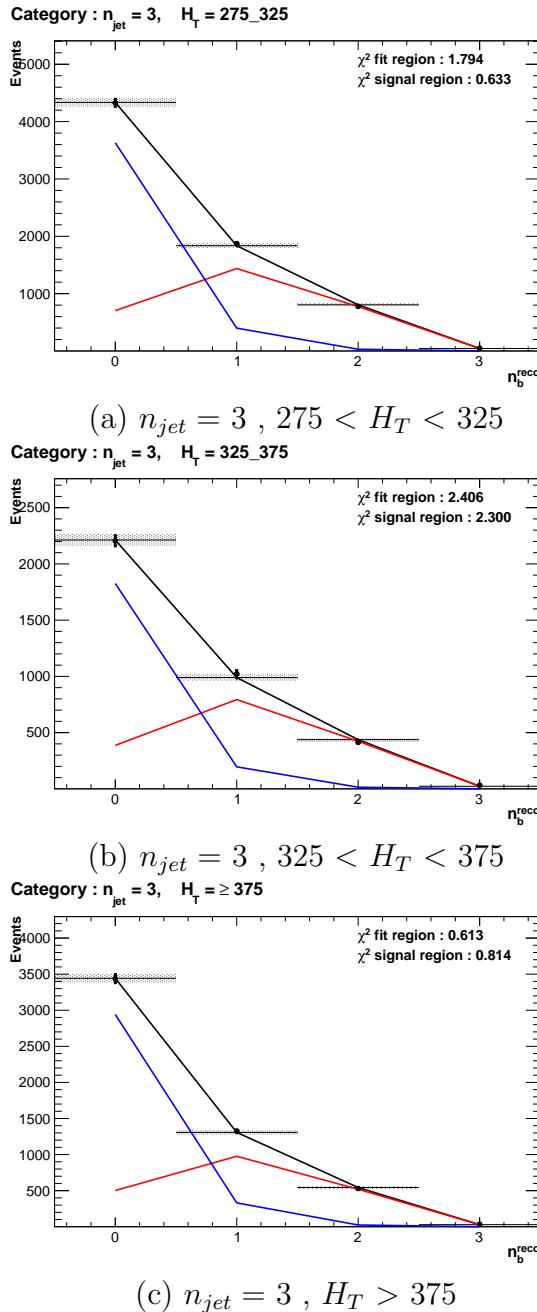


Figure D.4.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} = 3$ category and medium **CSV** working point. The blue template represents $Z = 0$, while the red template represents $Z = 2$. The χ^2 parameter displayed represents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

2315 Template fits for the three H_T bins, in the $n_{jet} = 4$, medium CSV working point:

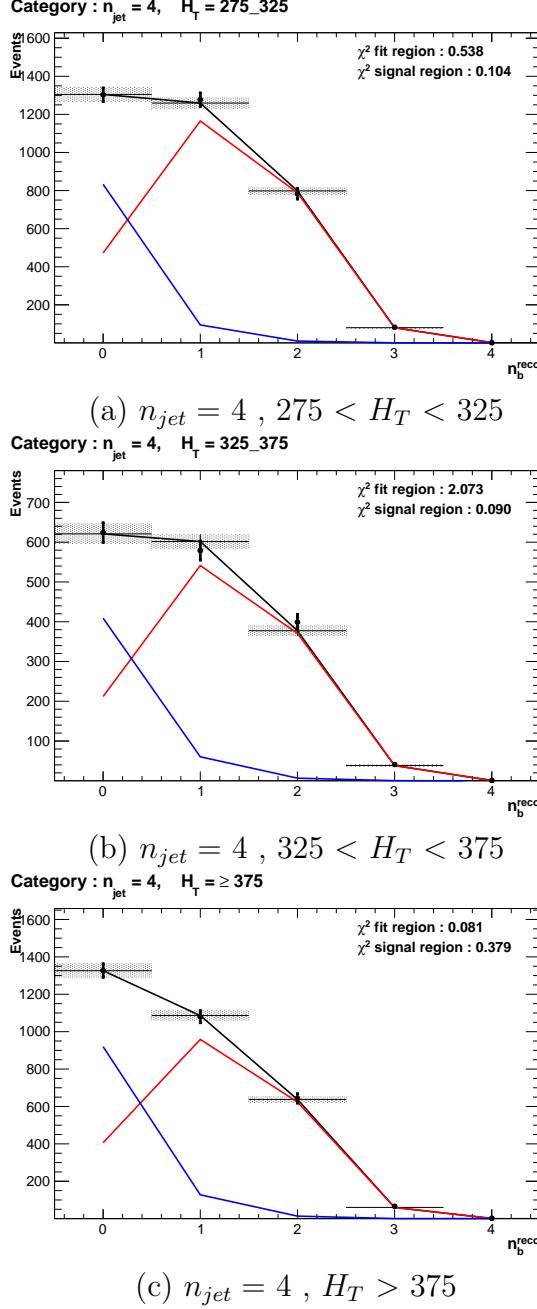
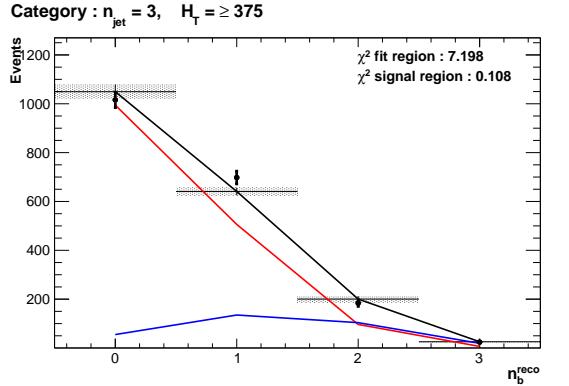


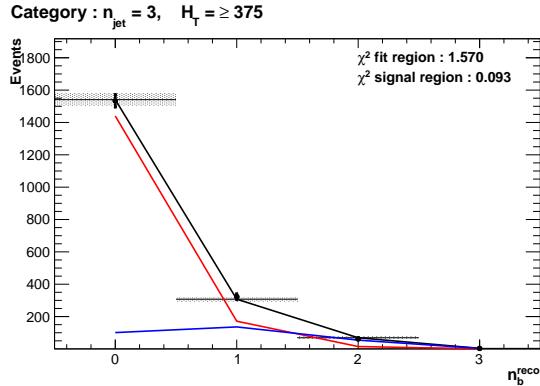
Figure D.5.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} = 4$ category and medium CSV working point. The blue template represents $Z = 0$, while the red template represents $Z = 2$. The χ^2 parameter displayed represents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

₂₃₁₆ **D.4. Templates Fits in Data Signal Region**

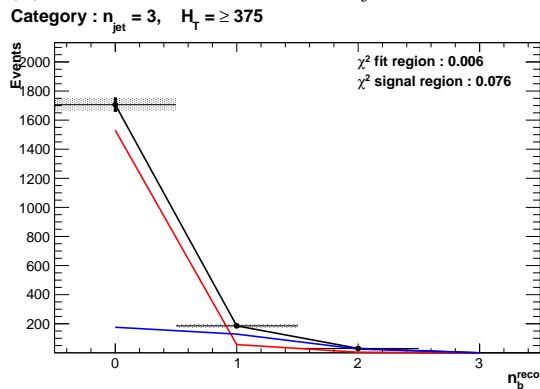
₂₃₁₇ Template fits for the three **CSV** working points, in the $n_{jet} = 3, H_T > 375$ category :



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



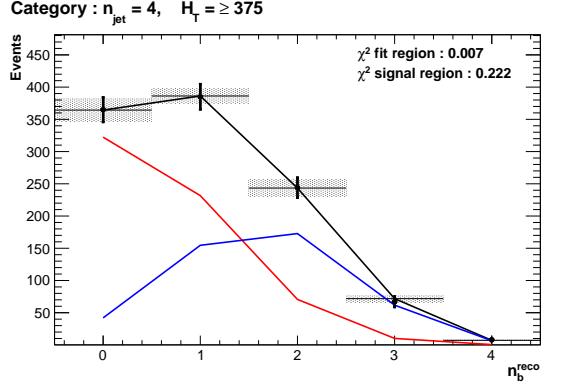
(b) Medium working point : $n_{jet} = 3, H_T > 375$



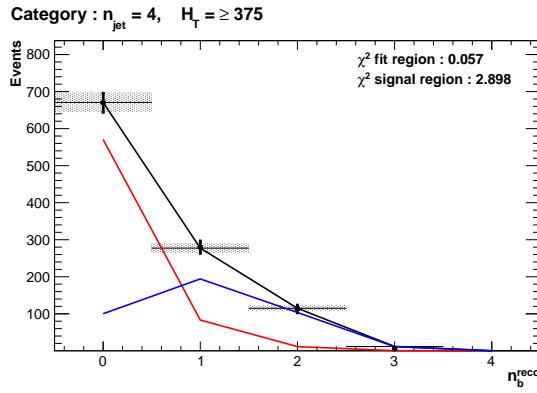
(c) Tight working point : $n_{jet} = 3, H_T > 375$

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 χ^2 parameter displayed represents the goodness of fit to the low n_b^{reco} (0-2) control region.

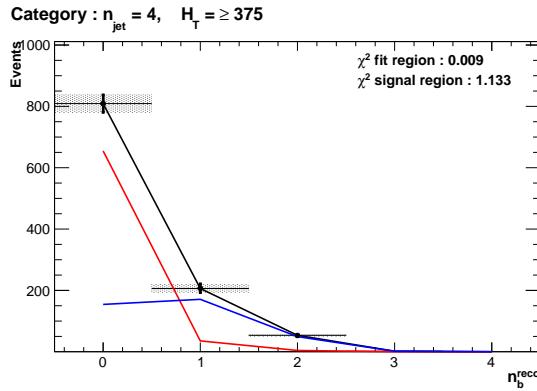
2318 Template fits for the three **CSV** working points, in the $n_{jet} = 4$, $H_T > 375$ category :



(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 χ^2 parameter displayed represents the goodness of fit to the low n_b^{reco} ($0-2$) control region.

2320

Bibliography

- 2321 [1] J. Beringer *et al.*, “Review of Particle Physics (RPP),” *Phys.Rev.*, vol. D86, p. 010001, 2012.
- 2322 [2] G. H. et al., “Nine-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results,” *The Astrophysical Journal Supplement Series*, vol. 208, no. 2, p. 19, 2013.
- 2323 [3] G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys.Lett.*, vol. B716, pp. 1–29, 2012.
- 2324 [4] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys.Lett.*, vol. B716, pp. 30–61, 2012.
- 2325 [5] S. Chatrchyan *et al.*, “Search for supersymmetry in hadronic final states with missing transverse energy using the variables AlphaT and b-quark multiplicity in pp collisions at 8 TeV,” *Eur.Phys.J.*, vol. C73, p. 2568, 2013.
- 2326 [6] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.
- 2327 [7] S. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl.Phys.*, vol. 22, pp. 579–588, 1961.
- 2328 [8] A. Salam, “Weak and Electromagnetic Interactions,” *Conf.Proc.*, vol. C680519, pp. 367–377, 1968.
- 2329 [9] G. Hooft, “Renormalizable lagrangians for massive yang-mills fields,” *Nuclear Physics B*, vol. 35, no. 1, pp. 167 – 188, 1971.
- 2330 [10] F. Hasert *et al.*, “Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment,” *Phys.Lett.*, vol. B46, pp. 138–140, 1973.

- 2345 [11] G. Arnison *et al.*, “Experimental Observation of Lepton Pairs of Invariant Mass
2346 Around 95-GeV/c**2 at the CERN SPS Collider,” *Phys.Lett.*, vol. B126, pp. 398–410,
2347 1983.
- 2348 [12] M. Banner *et al.*, “Observation of Single Isolated Electrons of High Transverse
2349 Momentum in Events with Missing Transverse Energy at the CERN anti-p p
2350 Collider,” *Phys.Lett.*, vol. B122, pp. 476–485, 1983.
- 2351 [13] E. Noether, “Invariante variationsprobleme,” *Nachrichten von der Gesellschaft
2352 der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse*, vol. 1918,
2353 pp. 235–257, 1918.
- 2354 [14] F. Halzen and A. D. Martin, “Quarks and leptons.” Wiley, 1985.
- 2355 [15] *Introduction to Elementary Particles*. Wiley-VCH, 2nd ed., Oct. 2008.
- 2356 [16] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, “Ex-
2357 perimental Test of Parity Conservation in Beta Decay,” *Physical Review*, vol. 105,
2358 pp. 1413–1415, Feb. 1957.
- 2359 [17] P. Higgs, “Broken symmetries, massless particles and gauge fields,” *Physics Letters*,
2360 vol. 12, no. 2, pp. 132 – 133, 1964.
- 2361 [18] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,”
2362 *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.
- 2363 [19] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*,
2364 vol. 13, pp. 508–509, Oct 1964.
- 2365 [20] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global conservation laws and
2366 massless particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.
- 2367 [21] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov
2368 1967.
- 2369 [22] H. Yukawa, “On the interaction of elementary particles. i,” *Progress of Theoretical
2370 Physics Supplement*, vol. 1, pp. 1–10, 1955.
- 2371 [23] Y. e. a. Fukuda, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev.
2372 Lett.*, vol. 81, pp. 1562–1567, Aug 1998.
- 2373 [24] R. Becker-Szendy, C. Bratton, D. Casper, S. Dye, W. Gajewski, *et al.*, “A Search
2374 for muon-neutrino oscillations with the IMB detector,” *Phys.Rev.Lett.*, vol. 69,

- 2375 pp. 1010–1013, 1992.
- 2376 [25] S. P. Martin, “A Supersymmetry primer,” 1997.
- 2377 [26] H. Nilles, *Supersymmetry, Supergravity and Particle Physics*. Physics reports, North-Holland Physics Publ., 1984.
- 2379 [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.
- 2381 [28] E. Witten, “Dynamical Breaking of Supersymmetry,” *Nucl.Phys.*, vol. B188, p. 513, 1981.
- 2383 [29] J. Wess and B. Zumino, “Supergauge transformations in four dimensions,” *Nuclear Physics B*, vol. 70, no. 1, pp. 39 – 50, 1974.
- 2385 [30] H. Muller-Kirsten and A. Wiedemann, *Introduction to Supersymmetry*. World Scientific lecture notes in physics, World Scientific, 2010.
- 2387 [31] I. Aitchison, *Supersymmetry in Particle Physics: An Elementary Introduction*. Cambridge University Press, 2007.
- 2389 [32] K. A. Intriligator and N. Seiberg, “Lectures on Supersymmetry Breaking,” *Class.Quant.Grav.*, vol. 24, pp. S741–S772, 2007.
- 2391 [33] Y. Shadmi, “Supersymmetry breaking,” pp. 147–180, 2006.
- 2392 [34] C. Burgess, P. G. Camara, S. de Alwis, S. Giddings, A. Maharana, *et al.*, “Warped Supersymmetry Breaking,” *JHEP*, vol. 0804, p. 053, 2008.
- 2394 [35] H. Murayama, “Supersymmetry breaking made easy, viable, and generic,” 2007.
- 2395 [36] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events*. Cambridge University Press, 2006.
- 2397 [37] S. P. Martin, “Implications of supersymmetric models with natural r-parity conservation,” 1996.
- 2399 [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.
- 2401 [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.

- 2433 [55] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical
2434 Design Report CMS, Geneva: CERN, 1997.
- 2435 [56] *The CMS muon project: Technical Design Report*. Technical Design Report CMS,
2436 Geneva: CERN, 1997.
- 2437 [57] CMS Collaboration, “The cms physics technical design report, volume 1,”
2438 *CERN/LHCC*, vol. 2006-001, 2006.
- 2439 [58] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,”
2440 *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.
- 2441 [59] “Jet performance in pp collisions at 7 tev,” Tech. Rep. CMS-PAS-JME-10-003,
2442 CERN, Geneva, 2010.
- 2443 [60] X. Janssen, “Underlying event and jet reconstruction in cms,” Tech. Rep. CMS-CR-
2444 2011-012, CERN, Geneva, Jan 2011.
- 2445 [61] T. C. collaboration, “Determination of jet energy calibration and transverse mo-
2446 mentum resolution in cms,” *Journal of Instrumentation*, vol. 6, no. 11, p. P11002,
2447 2011.
- 2448 [62] R. Eusebi and on behalf of the CMS collaboration), “Jet energy corrections and
2449 uncertainties in cms: reducing their impact on physics measurements,” *Journal of
2450 Physics: Conference Series*, vol. 404, no. 1, p. 012014, 2012.
- 2451 [63] “Algorithms for b Jet identification in CMS,” Tech. Rep. CMS-PAS-BTV-09-001,
2452 CERN, 2009. Geneva, Jul 2009.
- 2453 [64] “Performance of b tagging at $\sqrt{s}=8$ tev in multijet, ttbar and boosted topology
2454 events,” no. CMS-PAS-BTV-13-001, 2013.
- 2455 [65] T. C. collaboration, “Identification of b-quark jets with the cms experiment,” *Journal
2456 of Instrumentation*, vol. 8, no. 04, p. P04013, 2013.
- 2457 [66] S. Dasu *et al.*, “CMS. The TriDAS project. Technical design report, vol. 1: The
2458 trigger systems,” 2000.
- 2459 [67] P. Sphicas, “CMS: The TriDAS project. Technical design report, Vol. 2: Data
2460 acquisition and high-level trigger,” 2002.
- 2461 [68] J. B. et al., “Calibration and Performance of the Jets and Energy Sums in the
2462 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] L. Randall and D. Tucker-Smith, “Dijet searches for supersymmetry at the large hadron collider,” *Phys. Rev. Lett.*, vol. 101, p. 221803, Nov 2008.
- [72] “SUSY searches with dijet events,” 2008.
- [73] “Search strategy for exclusive multi-jet events from supersymmetry at CMS,” Tech. Rep. CMS-PAS-SUS-09-001, CERN, 2009. Geneva, Jul 2009.
- [74] “Calorimeter Jet Quality Criteria for the First CMS Collision Data,” Tech. Rep. CMS-PAS-JME-09-008, CERN, 2010. Geneva, Apr 2010.
- [75] 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.
- [76] The CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at 7 TeV,” *Journal of Instrumentation*, vol. 7, p. 2P, Oct. 2012.
- [77] “Search for supersymmetry in events with photons and missing energy,” Tech. Rep. CMS-PAS-SUS-12-018, CERN, Geneva, 2012.
- [78] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas,” *Phys.Lett.*, vol. B659, pp. 119–126, 2008.
- [79] 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.
- [80] “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.
- [81] 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.
- [82] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, “Squark and gluino production at hadron colliders,” *Nucl.Phys.*, vol. B492, pp. 51–103, 1997.
- [83] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, A. Perrotta, and the CMS Collaboration,

- 2492 “The fast simulation of the cms detector at lhc,” *Journal of Physics: Conference*
2493 *Series*, vol. 331, no. 3, p. 032049, 2011.
- 2494 [84] S. Banerjee, M. D. Hildreth, and the CMS Collaboration, “Validation and tuning
2495 of the cms full simulation,” *Journal of Physics: Conference Series*, vol. 331, no. 3,
2496 p. 032015, 2011.
- 2497 [85] C. Collaboration, “Cms btag pog : Cms b-tagging performance database.”
2498 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>, 2013.
- 2499 [86] “CMS Luminosity Based on Pixel Cluster Counting - Summer 2012 Update,” Tech.
2500 Rep. CMS-PAS-LUM-12-001, CERN, Geneva, 2012.
- 2501 [87] R. Cousins, “Probability density functions for positive nuisance parameters.”
2502 <http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf>., 2012.
- 2503 [88] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, “The RooStats project,” 2010.
- 2504 [89] F. James and M. Roos, “Minuit: A System for Function Minimization and Analysis
2505 of the Parameter Errors and Correlations,” *Comput.Phys.Commun.*, vol. 10, pp. 343–
2506 367, 1975.
- 2507 [90] A. L. Read, “Presentation of search results: the cl_s technique,” *Journal of Physics*
2508 *G: Nuclear and Particle Physics*, vol. 28, no. 10, p. 2693, 2002.
- 2509 [91] T. Junk, “Confidence level computation for combining searches with small statistics,”
2510 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators,*
2511 *Spectrometers, Detectors and Associated Equipment*, vol. 434, no. 23, pp. 435 – 443,
2512 1999.
- 2513 [92] A. L. Read, “Modified frequentist analysis of search results (the CL_s method),”
2514 no. CERN-OPEN-2000-205, 2000.
- 2515 [93] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-
2516 based tests of new physics,” *The European Physical Journal C*, vol. 71, no. 2, pp. 1–19,
2517 2011.
- 2518 [94] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-
2519 based tests of new physics,” *European Physical Journal C*, vol. 71, p. 1554, Feb.
2520 2011.