

¹ **Searches for Supersymmetry using the α_T
² variable with the CMS detector at the LHC**

³ Darren Burton

⁴ Imperial College London
⁵ Department of Physics

⁶ A thesis submitted to Imperial College London
⁷ for the degree of Doctor of Philosophy

9 Abstract

10 A search for supersymmetric particles is presented, using the Compact
11 Muon Solenoid detector at the Large Hadron Collider, with a signature
12 of missing energy in events with high p_T jets is presented. The analysis is
13 performed with 11.7 fb^{-1} of data, collected at a center-of-mass energy of
14 8 TeV during the 2012 run period. The dimensionless kinematic variable
15 α_T is used to select events with genuine missing energy signatures, while
16 Standard Model backgrounds in the signal region estimated using data
17 driven control samples. A complementary method to search for natural
18 SUSY signatures with a high number of b-flavoured jets, through the use
19 of a simple template fit is presented. The α_T search is used as a vehicle to
20 demonstrate proof of principle and as a search region for this technique.
21 Additionally the efficiency of the hadronic Level-1 single jet triggers are
22 measured throughout the 2012 run period. Results are presented with
23 a view to comparing L1 jet performance, before and after, a change
24 to the jet seed algorithm implemented during data taking. No excess
25 of events is found over Standard Model expectations in the α_T search.
26 Exclusion limits are set at the 95% confidence level in the parameter
27 space of simplified models, with special emphasis on compressed spectra
28 and natural SUSY scenarios.

29

Declaration

30 I, the author of this thesis, declare that the work presented within this
31 document to be my own. The work presented in Chapters 4, 6, 5 and Section
32 3.4.1, is a result of the author's own work or that of which I have been a
33 major contributor unless explicitly stated otherwise, and is carried out within
34 the context of the Imperial College London and CERN SUSY groups, itself a
35 subsection of the greater CMS collaboration. All figures and studies taken
36 from external sources are referenced appropriately throughout this document.

37

Darren Burton

38

Acknowledgements

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

41 Contents

42	List of Figures	viii
43	List of Tables	xii
44	1. Introduction	2
45	2. A Theoretical Overview	5
46	2.1. The Standard Model	5
47	2.1.1. Gauge Symmetries of the SM	7
48	2.1.2. The Electroweak Sector and Electroweak Symmetry Breaking	9
49	2.2. Motivation for Physics Beyond the Standard Model	13
50	2.3. Supersymmetry Overview	15
51	2.3.1. R-Parity	16
52	2.4. Experimental Signatures of SUSY at the LHC	17
53	2.4.1. Simplified models	18
54	3. The LHC And The CMS Detector	20
55	3.1. The LHC	20
56	3.2. The CMS Detector	23
57	3.2.1. Detector subsystems	23
58	3.2.2. Tracker	24
59	3.2.3. Electromagnetic calorimeter	25
60	3.2.4. Hadronic calorimeter	26
61	3.2.5. Muon systems	28
62	3.3. Event Reconstruction and Object Definition	28
63	3.3.1. Jets	28
64	3.3.2. B-tagging	30
65	3.4. Triggering System	33
66	3.4.1. The level-1 trigger	34

67	3.4.2. The L1 trigger jet algorithm	35
68	3.4.3. Measuring L1 jet trigger efficiencies	37
69	3.4.4. Effects of the L1 jet seed	38
70	3.4.5. Robustness of L1 jet performance against pile-up	40
71	3.4.6. Summary	43
72	4. SUSY Searches In Hadronic Final States	45
73	4.1. An Introduction to the α_T Search	46
74	4.1.1. The α_T variable	48
75	4.2. Search Strategy	50
76	4.2.1. Physics objects	53
77	4.2.2. Event selection	55
78	4.2.3. Control sample definition and background estimation	58
79	4.2.4. Estimating the QCD multi-jet background	65
80	4.3. Trigger Strategy	67
81	4.4. Measuring MC Normalisation Factors via H_T Sidebands	68
82	4.5. Determining MC Simulation Yields with Higher Statistical Precision	69
83	4.5.1. The formula method	70
84	4.5.2. Establishing proof of principle	71
85	4.5.3. Correcting measured efficiencies in simulation to data	71
86	4.6. Systematic Uncertainties on Transfer Factors	72
87	4.6.1. Determining systematic uncertainties from closure tests	75
88	4.7. Statistical Framework	78
89	4.7.1. Hadronic sample	78
90	4.7.2. H_T evolution model	79
91	4.7.3. Electroweak Sector (EWK) control samples	79
92	4.7.4. Contributions from signal	82
93	4.7.5. Total likelihood	83
94	5. Results and Interpretation	85
95	5.1. Statistical Interpretation	85
96	5.2. Interpretation in Simplified Signal Models	85
97	6. Searches For Natural SUSY With B-tag Templates.	86
98	6.1. Concept	86
99	6.2. Application to the α_T Search	88
100	6.2.1. Proof of principle in simulation	89

101	6.2.2. Results in a data control sample	92
102	6.2.3. Application to the α_T hadronic search region	94
103	6.3. Conclusions	94
104	7. Conclusions	95
105	A. Miscellaneous	96
106	A.1. Noise Filters	96
107	A.2. Primary Vertices	98
108	B. L1 Jets	100
109	B.1. Jet matching efficiencies	100
110	B.2. Leading Jet Energy Resolution	101
111	B.3. Resolution for Energy Sum Quantities	104
112	C. Additional material on background estimation methods	109
113	C.1. Determination of k_{QCD}	109
114	C.2. Effect of varying background cross sections on closure tests	110
115	D. Additional Material For B-tag Template Method	111
116	D.1. Templates Fits in Simulation	111
117	D.2. Pull Distributions for Template Fits	114
118	D.3. Templates Fits in Data	114
119	Bibliography	117

¹²⁰ List of Figures

¹²¹ 2.1.	One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion.	¹²² 14
¹²³ 2.2.	Two example simplified model decay chains.	¹²³ 19
¹²⁴ 3.1.	A top down layout of the LHC, with the position of the four main detectors labelled.	¹²⁵ 21
¹²⁶ 3.2.	The total integrated luminosity delivered to and collected by Compact Muon Solenoid (CMS) during the 2012 8 TeV pp runs	¹²⁷ 22
¹²⁸ 3.3.	A pictorial depiction of the CMS detector.	¹²⁸ 24
¹²⁹ 3.4.	Illustration of the CMS Electromagnetic CALorimeter (ECAL).	¹²⁹ 26
¹³⁰ 3.5.	Schematic of the CMS Hadronic CALorimeter (HCAL).	¹³⁰ 27
¹³¹ 3.6.	Combined Secondary Vertex (CSV) algorithm discriminator values in enriched $t\bar{t}$ and inclusive multi jet samples	¹³² 31
¹³³ 3.7.	Data/MC b-tag scale factors derived using the Combined Secondary Vertex Medium Working Point (CSVM) tagger.	¹³⁴ 32
¹³⁵ 3.8.	Data/MC mis-tag scale factors derived using the CSVM tagger.	¹³⁵ 33
¹³⁶ 3.9.	The CMS Level 1 Trigger (L1) Trigger system.	¹³⁶ 34
¹³⁷ 3.10.	Illustration of the Level-1 jet finding algorithm.	¹³⁷ 36
¹³⁸ 3.11.	L1 jet efficiency turn-on curves as a function of the offline CaloJet and PFJet E_T	¹³⁹ 38
¹⁴⁰ 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.	¹⁴¹ 39

142	3.13. Trigger cross section for the L1HTT150 trigger path.	40
143	3.14. L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T	41
144	3.15. L1 jet efficiency turn-on curves as a function of the leading offline E_T Calo	
145	(left) and PF (right) jet, for low, medium and high pile-up conditions.	42
146	3.16. Fit values from an Exponentially Modified Gaussian (EMG) function fitted	
147	to the resolution plots of leading Calo jet E_T measured as a function of	
148	$\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and high pile-up conditions.	43
149	3.17. Fit values from an EMG function fitted to the resolution plots of leading	
150	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	
151	high pile-up conditions.	44
152	4.1. Reconstructed offline H_T for 11.7fb^{-1} of data after a basic pre-selection.	47
153	4.2. The event topologies of background QCD dijet events (right) and a generic	
154	SUSY signature with genuine Z_T (left).	48
155	4.3. The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet	
156	multiplicities after a full analysis selection and shown for $H_T > 375$	50
157	4.4. Pictorial depiction of the analysis strategy employed by the α_T search to	
158	increase sensitivity to a wide spectra of SUSY models.	52
159	4.5. Data/MC comparisons of key variables for the hadronic signal region.	58
160	4.6. Data/MC comparisons of key variables for the $\mu + \text{jets}$ selection.	61
161	4.7. Data/MC comparisons of key variables for the $\mu\mu + \text{jets}$ selection.	63
162	4.8. Data/MC comparisons of key variables for the $\gamma + \text{jets}$ selection.	64
163	4.9. QCD sideband regions, used for determination of k_{QCD}	66
164	4.10. Sets of closure tests overlaid on top of the systematic uncertainty used for	
165	each of the five H_T regions.	77
166	6.1. The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging	
167	rate (c) as measured in simulation after the α_T analysis, $\mu + \text{jets}$ control	
168	sample selection in the region $H_T > 375$	88

169	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.	91
170		
171		
172	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.	93
173		
174		
175	B.1. Leading jet matching efficiency as a function of the offline CaloJet E_T	100
176	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.	102
177		
178	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.	104
179		
180	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.	105
181		
182	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.	105
183		
184	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.	106
185		
186	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.	106
187		
188	B.8. H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile up conditions.	107
189		
190	B.9. H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions.	107
191		
192	B.10. \cancel{H}_T resolution parameters in bins of \cancel{H}_T measured for the defined low, medium and high pile up conditions.	108
193		
194	B.11. \cancel{H}_T resolution parameters in bins of PF \cancel{H}_T measured for the defined low, medium and high pile up conditions.	108
195		
196	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 $275 < H_T < 575$	109
197		

198	C.2. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five H_T regions.	110
199		
200	C.3. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five H_T regions.	110
201		
202	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.	112
203		
204		
205	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.	113
206		
207		
208	D.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 loose CSV working point.	114
209		
210		
211	D.4. The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from data, for the $n_{jet} \geq 5$ category and tight CSV working point.	115
212		
213		

²¹⁴ List of Tables

215	2.1. The fundamental particles of the Standard Model (SM), with spin, charge and mass displayed.	6
216		
217	3.1. Results of a cumulative EMG function fit to the turn-on curves for L1 single jet triggers in 2012 Run Period C.	38
218		
219	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.	40
220		
221	3.3. Results of a cumulative EMG function fit to the turn-on curves for H_T in 2012 run period B and C.	41
222		
223	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, for low,medium and high pile-up conditions.	41
224		
225		
226	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, for low,medium and high pile-up conditions.	42
227		
228		
229	4.1. A summary of the Simplified Model Spectra (SMS) models interpreted in this analysis, involving both direct (D) and gluino-induced (G) production of squarks and their decays.	46
230		
231		
232	4.2. Muon Identification criteria used within the analysis for selection/veto purposes in the muon control/signal selections.	53
233		
234	4.3. Photon Identification criteria used within the analysis for selection/veto purposes in the $\gamma +$ jets control/signal selections.	54
235		
236	4.4. Electron Identification criteria used within the analysis for veto purposes.	55

237	4.5. Jet thresholds used in the three H_T regions of the analysis.	56
238	4.6. Best fit values for the parameters k_{QCD} obtained from sideband regions	
239	B,C ₁ ,C ₂ ,C ₃	67
240	4.7. Measured efficiencies of the H_T and α_T legs of the HT and HT_alphaT	
241	triggers in independent analysis bins.	68
242	4.8. k-factors calculated for different EWK processes.	69
243	4.9. place holder	71
244	4.10. A summary of the results obtained from fits of zeroeth order polynomials	
245	(i.e. a constant) to five sets of closure tests performed in the $2 \leq n_{jet} \leq 3$	
246	bin	74
247	4.11. A summary of the results obtained from fits of zeroeth order polynomials	
248	(i.e. a constant) to five sets of closure tests performed in the $n_{jet} \geq 4$ bin	74
249	4.12. A summary of the results obtained from fits of zeroeth order polynomials	
250	(i.e. a constant) to five sets of closure tests performed in the $2 \leq n_{jet} \leq 3$	
251	bin	75
252	4.13. Calculated systematic uncertainties for the five H_T regions, determined	
253	from the closure tests.	76
254	4.14. The systematic parameters used in H_T bins.	82
255	4.15. Nuisance parameters used within the different hadronic signal bins of the	
256	analysis	84
257	6.1. Summary of the fit predictions in the n_b^{reco} signal region for $n_{jet} = 3, =$	
258	$4, \geq 5$. The fit region is $n_b^{reco} = 0, 1, 2$ and simulation yields are normalised	
259	to an integrated luminosity of 10 fb^{-1}	90
260	6.2. Summary of the fit predictions in the n_b^{reco} signal region of the $\mu +$ jets	
261	control sample, for $n_{jet} = 3, = 4, \geq 5$. The fit region is $n_b^{reco} = 0, 1, 2$	
262	using 11.5 fb^{-1} of data at $\sqrt{s} = 8\text{TeV}$	92
263	A.1. Criteria for a reconstructed jet to pass the loose calorimeter jet id.	96
264	A.2. Criteria for a reconstructed jet to pass the loose PF jet id.	97

265	A.3. Noise filters that are applied to remove spurious and non-physical \cancel{E}_T	98
266	signatures within the CMS detector.	
267	A.4. Criteria for a vertex in an event to be classified as a 'good' reconstructed	
268	primary vertex.	99
269	B.1. Results of a cumulative EMG function fit to the turn-on curves for the	
270	matching efficiency of the leading jet in an event to a Level-1 jet in run	
271	2012C and 2012B data.	101
272	C.1.	110

273 Acronyms

274	ALICE	A Large Ion Collider Experiment
275	ATLAS	A Toroidal LHC ApparatuS
276	APD	Avalanche Photo-Diodes
277	BSM	Beyond Standard Model
278	CERN	European Organization for Nuclear Research
279	CMS	Compact Muon Solenoid
280	CMSSM	Compressed Minimal SuperSymmetric Model
281	CSC	Cathode Stripe Chamber
282	CSV	Combined Secondary Vertex
283	CSVM	Combined Secondary Vertex Medium Working Point
284	DT	Drift Tube
285	ECAL	Electromagnetic CALorimeter
286	EB	Electromagnetic CALorimeter Barrel
287	EE	Electromagnetic CALorimeter Endcap
288	ES	Electromagnetic CALorimeter pre-Shower
289	EMG	Exponentially Modified Gaussian
290	EPJC	European Physical Journal C
291	EWK	Electroweak Sector
292	GCT	Global Calorimeter Trigger
293	GMT	Global MuonTrigger
294	GT	Global Trigger
295	HB	Hadron Barrel
296	HCAL	Hadronic CALorimeter

297	HE	Hadron Endcaps
298	HF	Hadron Forward
299	HLT	Higher Level Trigger
300	HO	Hadron Outer
301	HPD	Hybrid Photo Diode
302	LUT	Look Up Table
303	L1	Level 1 Trigger
304	LHC	Large Hadron Collider
305	LHCb	Large Hadron Collider Beauty
306	LSP	Lightest Supersymmetric Partner
307	NNLO	Next to Next Leading Order
308	POGs	Physics Object Groups
309	PS	Proton Synchrotron
310	QED	Quantum Electro-Dynamics
311	QCD	Quantum Chromo-Dynamics
312	QFT	Quantum Field Theory
313	RBXs	Readout Boxes
314	RPC	Resistive Plate Chamber
315	RCT	Regional Calorimeter Trigger
316	RMT	Regional Muon Trigger
317	SUSY	SUPerSYmmetry
318	SM	Standard Model
319	SMS	Simplified Model Spectra
320	SPS	Super Proton Synchrotron
321	TF	Transfer Factor

³²²	TP	Trigger Primitive
³²³	VEV	Vacuum Expectation Value
³²⁴	VPT	Vacuum Photo-Triodes
³²⁵	WIMP	Weakly Interacting Massive Particle

³²⁶

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

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

³²⁷

Chapter 1.

³²⁸ Introduction

³²⁹ During the 20th century great advances have been made in the human understanding
³³⁰ of the universe, it's origins, it's future and its composition. The Standard Model (**SM**)
³³¹ first formulated in the 1960's is one of the crowning achievements in science's quest to
³³² explain the most fundamental processes and interactions that make up our universe. It
³³³ has provided a highly successful explanation of a wide range of phenomena in Particle
³³⁴ Physics and has stood up to extensive experimental scrutiny [1].

³³⁵ Despite it's successes it is not a complete theory, with significant questions remaining
³³⁶ unanswered. It describes only three of the four known forces with gravity not incorpo-
³³⁷ rated within the framework of the **SM**. Cosmological experiments infer that just $\sim 4\%$
³³⁸ of the observable universe exists as matter, with elusive "Dark Matter" accounting for a
³³⁹ further $\sim 23\%$ [2]. However no particle predicted by the **SM** is able to account for it. At
³⁴⁰ higher energy scales and small distances the (non-)unification of the fundamental forces
³⁴¹ point to problems with the **SM** at least at higher energies not yet probed experimentally.

³⁴² Many theories exist as extensions to the **SM**, which predict a range of observables
³⁴³ that can be detected at the Large Hadron Collider (**LHC**) of which SUperSYmmetry
³⁴⁴ (**SUSY**) is one such example. It predicts a new symmetry of nature in which all current
³⁴⁵ particles in the **SM** would have a corresponding supersymmetric partner. Common to
³⁴⁶ most Supersymmetric theories is a stable, weakly interacting Lightest Supersymmetric
³⁴⁷ Partner (**LSP**), which has the properties of a possible dark matter candidate. The **SM**
³⁴⁸ and the main principles of Supersymmetric theories are outlined in Chapter 2, with
³⁴⁹ emphasis placed on how experimental signatures of **SUSY** may reveal themselves at the
³⁵⁰ **LHC**.

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

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

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

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

380 Additionally a method to search for **SUSY** signatures which are rich in top and bottom
381 flavoured jet final states is disscussd in Chapter 6. A parametrisation of the b-tagging
382 distribution for different Electroweak processes is used to establish templates, which
383 are then used to estimate the expected number of 3 or 4 b-tagged jet events from **SM**

384 processes. The α_T search is used as a cross check for this template method to establish
385 it's functionality.

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

Chapter 2.

³⁹² A Theoretical Overview

³⁹³ Within this chapter, a brief introduction and background to the **SM** is given. Its success
³⁹⁴ as a rigorously tested and widely accepted theory is discussed as are its deficiencies,
³⁹⁵ leading to the argument that this theory is not a complete description of our universe.
³⁹⁶ The motivations for new physics at the TeV scale and in particular Supersymmetric
³⁹⁷ theories are outlined within Section (2.3), with the chapter concluding with how an
³⁹⁸ experimental signature of such theories can be produced and observed at the **LHC**,
³⁹⁹ Section (2.4).

⁴⁰⁰ 2.1. The Standard Model

⁴⁰¹ The **SM** is the name given to the relativistic Quantum Field Theory (**QFT**), where
⁴⁰² particles are represented as excitations of fields, which describe the interactions and
⁴⁰³ properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field
⁴⁰⁴ theory which contains three symmetries: $SU(3)$ for colour charge, $SU(2)$ for weak isospin
⁴⁰⁵ and $U(1)$ relating to weak hyper charge, which require its Lagrangian \mathcal{L}_{SM} to be invariant
⁴⁰⁶ under local gauge transformation.

⁴⁰⁷ Within the **SM** theory, matter is composed of spin $\frac{1}{2}$ fermions, which interact with each
⁴⁰⁸ other via the exchange of spin-1 gauge bosons. A summary of the known fundamental
⁴⁰⁹ fermions and bosons is given in Table 2.1.

⁴¹⁰ Fermions are separated into quarks and leptons of which only quarks interact with
⁴¹¹ the strong nuclear force. Quarks unlike leptons are not seen as free particles in nature,
⁴¹² but rather exist only within baryons, composed of three quarks with an overall integer
⁴¹³ charge, and quark-anti-quark pairs called mesons. Both leptons and quarks are grouped

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].

414 into three generations which have the same properties, but with ascending mass in each
 415 subsequent generation.

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

420 The unification of the electromagnetic and weak-nuclear forces into the current
 421 Electroweak theory yield the weak gauge bosons, W^\pm and Z through the mixing of the
 422 associated gauge fields. The force carriers of this theory were experimentally detected by
 423 the observation of weak neutral current, discovered in 1973 in the Gargamelle bubble

⁴²⁴ chamber located at European Organization for Nuclear Research (**CERN**) [10], with the
⁴²⁵ masses of the weak gauge bosons measured by the UA1 and U2 experiments at the Super
⁴²⁶ Proton Synchrotron (**SPS**) collider in 1983 [11][12].

⁴²⁷ 2.1.1. Gauge Symmetries of the SM

⁴²⁸ Symmetries are of fundamental importance in the description of physical phenomena.
⁴²⁹ Noether's theorem states that for a dynamical system, the consequence of any symmetry
⁴³⁰ is an associated conserved quantity [13]. Invariance under translations, rotations, and
⁴³¹ Lorentz transformations in physical systems lead to conservation of momentum, energy
⁴³² and angular momentum.

⁴³³ In the **SM**, a quantum theory described by Lagrangian formalism, the weak, strong and
⁴³⁴ electromagnetic interactions are described in terms of “gauge theories”. A gauge theory
⁴³⁵ possesses invariance under a set of “local transformations”, which are transformations
⁴³⁶ whose parameters are space-time dependent. The requirement of gauge invariance within
⁴³⁷ the **SM** necessitates the introduction of force-mediating gauge bosons and interactions
⁴³⁸ between fermions and the bosons themselves. Given the nature of the topics covered by
⁴³⁹ this thesis, the formulation of **EWK** within the **SM** Lagrangian is reviewed within this
⁴⁴⁰ section.

⁴⁴¹ The simplest example of the application of the principle of local gauge invariance
⁴⁴² within the **SM** is in Quantum Electro-Dynamics (**QED**), the consequences of which
⁴⁴³ require a massless photon field [14][15].

⁴⁴⁴ Starting from the free Dirac Lagrangian written as

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

⁴⁴⁵ where ψ represents a free non interacting fermionic field, with the matrices $\gamma^\mu, \mu \in$
⁴⁴⁶ $0, 1, 2, 3$ defined by the anti commutator relationship $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} I_4$, with $\eta^{\mu\nu}$
⁴⁴⁷ being the flat space-time metric $(+, -, -, -)$ and I_4 the 4×4 identity matrix.

⁴⁴⁸ Under a local $U(1)$ abelian gauge transformation in which ψ transforms as:

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

⁴⁴⁹ the kinetic term of the Lagrangian does not remain invariant, due to the partial
⁴⁵⁰ derivative interposed between the $\bar{\psi}$ and ψ yielding,

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

⁴⁵¹ To ensure that \mathcal{L} remains invariant, a modified derivative, D_μ , that transforms
⁴⁵² covariantly under phase transformations is introduced. In doing this a vector field A_μ
⁴⁵³ with transformation properties that cancel out the unwanted term in (2.3) must also be
⁴⁵⁴ included,

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

⁴⁵⁵ Invariance of the Lagrangian is then achieved by replacing ∂_μ by D_μ :

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

⁴⁵⁶ An additional interaction term is now present in the Lagrangian, coupling the Dirac
⁴⁵⁷ particle to this vector field, which is interpreted as the photon in QED. To regard this
⁴⁵⁸ new field as the physical photon field, a term corresponding to its kinetic energy must be
⁴⁵⁹ added to the Lagrangian from Equation (2.5). Since this term must also be invariant
⁴⁶⁰ under the conditions of Equation (2.4), it is defined in the form $F_{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

⁴⁶¹ This then leads to the Lagrangian of QED:

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

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

465 2.1.2. The Electroweak Sector and Electroweak Symmetry 466 Breaking

467 The same application of gauge symmetry and the requirement of local gauge invariance
468 can be used to unify QED and the Weak force in the Electroweak Sector (EWK).
469 The nature of EWK interactions is encompassed within a Lagrangian invariant under
470 transformations of the group $SU(2)_L \times U(1)_Y$.

471 The weak interactions from experimental observation [16], are known to violate parity
472 and are therefore not symmetric under interchange of left and right helicity fermions.
473 Thus within the SM the left and right handed parts of these fermion fields are treated
474 separately. A fermion field is then split into two left and right handed chiral components,
475 $\psi = \psi_L + \psi_R$, where $\psi_{L/R} = (1 \pm \gamma^5)\psi$.

476 The $SU(2)_L$ group is the special unitary group of 2×2 matrices U satisfying $UU^\dagger = I$
477 and $\det(U) = 1$. It may be written in the form $U = e^{-i\omega_i T_i}$, with the generators of the
478 group $T_i = \frac{1}{2}\tau_i$ where τ_i , $i \in 1,2,3$ being the 2×2 Pauli matrices

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

479 which form a non Abelian group obeying the commutation relation $[T^a, T^b] \equiv$
480 $if^{abc}T^c \neq 0$. The gauge fields that accompany this group are represented by $\hat{W}_\mu =$
481 $(\hat{W}_\mu^1, \hat{W}_\mu^2, \hat{W}_\mu^3)$ and act only on the left handed component of the fermion field ψ_L .

⁴⁸² One additional generator Y which represents the hypercharge of the particle under
⁴⁸³ consideration is introduced through the $U(1)_Y$ group acting on both components of the
⁴⁸⁴ fermion field, with an associated vector boson field \hat{B}_μ .

⁴⁸⁵ The $SU(2)_L \times U(1)_Y$ transformations of the left and right handed components of ψ
⁴⁸⁶ are summarised by,

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

⁴⁸⁷ where the left handed fermions form isospin doubles χ_L and the right handed fermions
⁴⁸⁸ are isosinglets ψ_R . For the first generation of leptons and quarks this represents

$$\begin{aligned}\chi_L &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_L \\ \psi_R &= e_R, \quad u_R, d_R\end{aligned}\tag{2.9}$$

⁴⁸⁹ Imposing local gauge invariance within \mathcal{L}_{EWK} is once again achieved by modifying
⁴⁹⁰ the covariant derivative

$$D_\mu = \partial_\mu - \frac{ig}{2} \tau^i W_\mu^i - \frac{ig'}{2} Y B_\mu,\tag{2.10}$$

⁴⁹¹ where g and g' are the coupling constant of the $SU(2)_L$ and $U(1)_Y$ groups respectively.
⁴⁹² Taking the example of the first generation of fermions defined in Equation.(2.9), with input
⁴⁹³ hypercharge values of -1 and -2 for χ_L and e_R respectively, would lead to a Lagrangian
⁴⁹⁴ \mathcal{L}_1 of the form,

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

⁴⁹⁵ As in **QED**, these additional gauge fields introduce field strength tensors $B_{\mu\nu}$ and
⁴⁹⁶ $W_{\mu\nu}$,

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

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

⁴⁹⁷ corresponding to the kinetic energy and self coupling of the W_μ fields and the kinetic
⁴⁹⁸ energy term of the B_μ field.

⁴⁹⁹ None of these gauge bosons are physical particles, and instead linear combinations of
⁵⁰⁰ these gauge bosons make up γ and the W and Z bosons, defined as

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

⁵⁰¹ where the mixing angle, $\theta_w = \tan^{-1} \frac{g'}{g}$, relates the coupling of the neutral weak and
⁵⁰² electromagnetic interactions.

⁵⁰³ As in the case of the formulation of the **QED** Lagrangian there remains no mass term
⁵⁰⁴ for the photon. However this is also the case for the W, Z and fermions in the Lagrangian,
⁵⁰⁵ contrary to experimental measurement. Any explicit introduction of mass terms would
⁵⁰⁶ break the symmetry of the Lagrangian and instead mass terms can be introduced through
⁵⁰⁷ spontaneous breaking of the **EWK** symmetry via the Higgs mechanism.

⁵⁰⁸ The Higgs mechanism induces spontaneous symmetry breaking through the introduc-
⁵⁰⁹ tion of a complex scalar SU(2) doublet field ϕ which attains a non-zero Vacuum
⁵¹⁰ Expectation Value (**VEV**) [17][18][19][20].

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

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

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

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

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

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

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

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

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

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

526 where the relations between the physical and electroweak gauge fields from Equation
 527 (2.14) are used. The W^\pm and Z bosons can then be determined to be

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

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

534 2.2. Motivation for Physics Beyond the Standard 535 Model

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

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

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

551 Perhaps one of the most glaring gaps in the predictive power of the SM is that there
 552 exists no candidate to explain the cosmic dark matter observed in galactic structures
 553 through indirect techniques including gravitational lensing and measurement of the
 554 orbital velocity of stars at galactic edges. Any such candidate must be very weakly

555 interacting but must also be stable, owing to the lack of direct detection of the decay
 556 products of such a process. Therefore a stable dark matter candidate, is one of the
 557 main obstacles to address for any Beyond Standard Model (**BSM**) physics model.

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

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

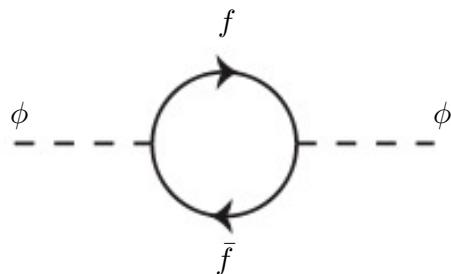


Figure 2.1.: One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion.

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

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

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

569 To recover the mass of the now discovered Higgs boson would require a fine-tuning of
 570 the parameters to cancel out these mass corrections of the Higgs mass to the scale of
 571 30 orders of magnitude. This appears as an unnatural solution to physicists and it is
 572 this hierarchy problem that provides one of the strongest motivations for the theory of
 573 SUperSYmmetry (**SUSY**).

⁵⁷⁴ 2.3. Supersymmetry Overview

⁵⁷⁵ Supersymmetry provides potential solutions to many of the issues raised in the previous
⁵⁷⁶ section. It provides a dark matter candidate, can explain baryogenesis in the early
⁵⁷⁷ universe and also provides an elegant solution to the hierarchy problem [26][27][28][29].
⁵⁷⁸ At its heart it represents a new space-time symmetry that relates fermions and bosons.
⁵⁷⁹ This symmetry converts bosonic states into fermionic states, and vice versa, see Equation
⁵⁸⁰ (2.21) ,

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

⁵⁸¹ where the operator Q is the generator of these transformations. Quantum field theories
⁵⁸² which are invariant under such transformations are called supersymmetric.

⁵⁸³ This symmetry operator therefore acts upon a particles spin altering it by a half integer
⁵⁸⁴ value. The consequences of the introduction of this additional space-time symmetry
⁵⁸⁵ introduce a new rich phenomenology. For example in supersymmetric theories, both
⁵⁸⁶ the left handed $SU(2)$ doublet and right handed singlet of fermions will have a spin-0
⁵⁸⁷ superpartner, containing the same electric charge, weak isospin, and colour as its **SM**
⁵⁸⁸ partner. In the case of the leptons $(\nu_l, l)_L$, they will have two superpartners, a sneutrino
⁵⁸⁹ $\tilde{\nu}_l{}_L$ and a slepton \tilde{l}_L , whilst the singlet l_R also has a superpartner slepton \tilde{l}_R .

⁵⁹⁰ Each particle in a supersymmetric theory is paired together with their superpartners
⁵⁹¹ as a result of these supersymmetric transformations in a so called supermultiplet. These
⁵⁹² superpartners will then consequently also contribute to the corrections to the Higgs mass.
⁵⁹³ Bosonic and fermionic loops contributing to the correction appear with opposite signs,
⁵⁹⁴ and therefore cancellation of these divergent terms will stabilise the Higgs mass, solving
⁵⁹⁵ the hierarchy problem [30][31].

⁵⁹⁶ One of the simplest forms of **SUSY**, is to simply have a set of **SM** supersymmetric
⁵⁹⁷ partners with the same mass and interactions as their counterparts. However the currently
⁵⁹⁸ lack of any experimental evidence for the predicted sparticle spectrum implies **SUSY**
⁵⁹⁹ must be a broken symmetry in which any sparticle masses must be greater than their
⁶⁰⁰ **SM** counterparts.

⁶⁰¹ There exist many techniques which can induce supersymmetric breaking [32][33][34].
⁶⁰² Of particular interest to experimental physicists are those at which the breaking scale

is of an order that is experimentally accessible to the LHC i.e. \sim TeV scale. Whilst there is no requirement for supersymmetric breaking to occur at this energy scale, for supersymmetry to provide a solution to the hierarchy problem, it is necessary for this scale to not differ too drastically from the EWK scale [35][36].

2.3.1. R-Parity

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

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

R-parity is described by the equation

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

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

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

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

630 2.4. Experimental Signatures of SUSY at the LHC

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

- 633 • squark/anti-squark and gluino pairs can be produced via both gluon fusion and
634 quark/anti-quark scattering.
- 635 • a gluino and squark produced together via quark-gluon scattering
- 636 • squark pairs produced via quark-quark scattering

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

640 During the 2011 run period at a $\sqrt{s} = 7$ TeV, particular models were used to
641 benchmark performance and experimental reach of both CMS searches and previous
642 experiments. The Compressed Minimal SuperSymmetric Model (CMSSM) was initially
643 chosen for a number of reasons [38], one of the most compelling being the reduction from
644 up to 105 new parameters that can be introduced by SUSY in addition to the existing
645 19 of the SM, to just 5 free extra free parameters. It was this simplicity, combined with
646 the theory not requiring any fine tuning of particle masses to produce the experimentally
647 verified SM that made it an attractive model to interpret physics results.

648 However recent results from the LHC now strongly disfavour large swathes of CMSSM
649 parameter space [39][40][41]. In the face of such results a more pragmatic model indepen-
650 dent search strategy is now applied across most SUSY searches at the LHC, see Section
651 (2.4.1).

652 As previously stated, a stable LSP that exhibits the properties of a dark matter
653 candidate would be weakly interacting and therefore will not be directly detected in a
654 detector environment. Additionally the cascade decays of supersymmetric particles to
655 the LSP would also result in significant hadronic activity. These signatures can then be
656 characterised through large amounts of hadronic jets (see Section (3.3.1)), leptons and
657 a significant amount of missing energy dependent upon the size of the mass splitting
658 between the LSP and the supersymmetric particle it has decayed from.

659 The SM contains processes which can exhibit a similar event topology to that described
660 above. The largest contribution coming from the general QCD environment of a hadron

661 collider. A multitude of different analytical techniques are used by experimental physicists
662 to reduce or estimate any reducible or irreducible backgrounds, allowing a possible **SUSY**
663 signature to be extracted. The techniques employed within this thesis are described in
664 great detail within Section (4.1).

665 **2.4.1. Simplified models**

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

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

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

- 681 • **T1,T1xxxx**, models represent a simplified version of gluino pair production with
682 each gluino (superpartner to the gluon) undergoing a three-body decay to a quark-
683 antiquark pair and the **LSP** (i.e. $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{LSP}$). The resultant final state from this
684 decay is typically 4 jets + \cancel{E}_T in the absence of initial/final state radiation and
685 detector effects. xxxx denotes models in which the quarks are of a specific flavour,
686 typically t or b quark-antiquarks.
- 687 • **T2,T2xx**, models represent a simplified version of squark anti-squark production
688 with each squark undergoing a two-body decay into a light-flavour quark and **LSP**
689 (i.e. $\tilde{q} \rightarrow q\tilde{\chi}_{LSP}$). This results in final states with less jets than gluino mediated
690 production, typically 2 jets + \cancel{E}_T when again ignoring the effect of initial/final state
691 radiation and detector effects. xx models again represent decays in which both the
692 quark and the squark within the decay is of a specific flavour, typically \tilde{t}/t or \tilde{b}/b .

Models rich in b and t quarks are interpreted within this thesis as they remain of particular interest within “Natural SUSY” scenarios [45][46]. The largest contribution to the quadratic divergence in the Higgs mass parameter comes from a loop of top quarks via the Yukawa coupling. Cancellation of these divergences can be achieved in supersymmetric theories by requiring a light right handed top squark, \tilde{t}_R , and left-handed double $SU(2)_L$ doublet containing top and bottom squarks, $(\tilde{\frac{t}{b}})_L$ [47].

These theories therefore solve the hierarchy problem by predicting light \sim EWK scale third generation sleptons, to be accessible at the LHC. Search strategies involving the requirement of b-tagging (see Section (3.3.2)) are used to give sensitivity to these type of SUSY scenarios and are discussed in greater detail within Chapter 5.

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

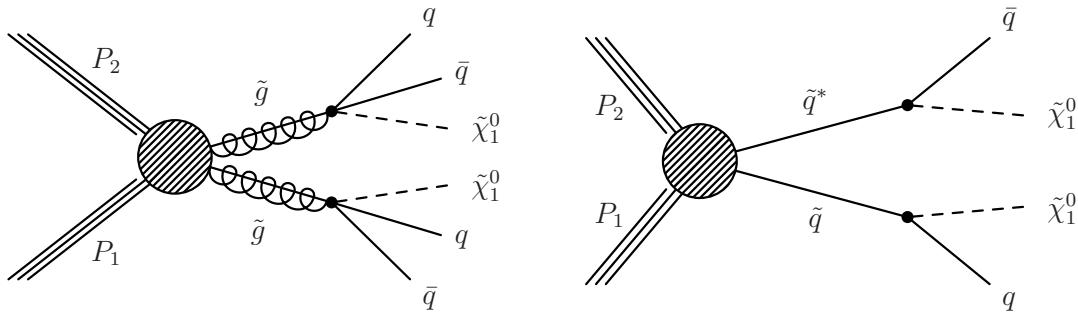


Figure 2.2.: Two example SMS model decays (T1 (left), T2 (right)), which are used in interpretations of physics reach by CMS.

Chapter 3.

⁷⁰⁵ The LHC And The CMS Detector

⁷⁰⁶ Probing the SM for signs of new physics would not be possible without the immensely
⁷⁰⁷ complex electronics and machinery that makes the TeV energy scale accessible for the
⁷⁰⁸ first time. This chapter will describe both the LHC based at European Organization
⁷⁰⁹ for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS) detector, being
⁷¹⁰ the experiment the author is a member of. Section (3.2) serves to introduce an overview
⁷¹¹ of the different components of the CMS detector, with specific components relevant to
⁷¹² the search for supersymmetric particles described in greater detail. Section (3.3) will
⁷¹³ focus on event and object reconstruction again with more emphasis on jet level quantities
⁷¹⁴ which are most relevant to the author's analysis research. Finally Section (3.4) will
⁷¹⁵ cover work performed by the author, as service to the CMS Collaboration, in measuring
⁷¹⁶ the performance of the Global Calorimeter Trigger (GCT) component of the L1 trigger
⁷¹⁷ during the 2012-2013 run period.

⁷¹⁸ 3.1. The LHC

⁷¹⁹ The LHC is a storage ring, accelerator, and collider of circulating beams of protons or
⁷²⁰ ions. Housed in the tunnel dug for the Large Electron-Positron collider (LEP), it is
⁷²¹ approximately 27 km in circumference, 100 m underground, and straddles the border
⁷²² between France and Switzerland outside of Geneva. It is currently the only collider
⁷²³ in operation that is able to study physics at the TeV scale. A double-ring circular
⁷²⁴ synchrotron, it was designed to collide both proton-proton (pp) and heavy ion (PbPb)
⁷²⁵ with a centre of mass energy $\sqrt{s} = 14$ TeV at a final design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$.

⁷²⁶

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 (**LHC-B**) [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.

738

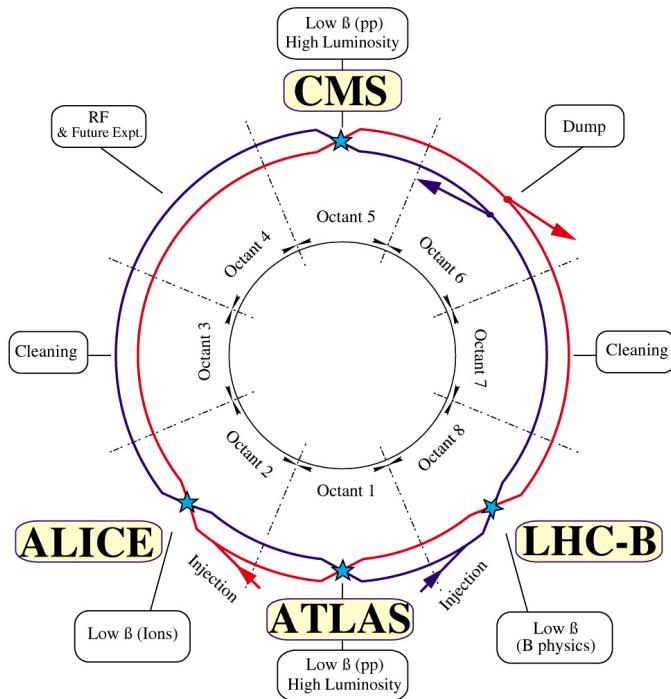


Figure 3.1.: A top down layout of the LHC. [52], with the position of the four main detectors labelled.

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

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

749

750 Colliding the beams produced an instantaneous luminosity of approximately $5 \times$
 751 $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ during the 2012 run. The high number of protons in each bunch increases
 752 the likelihood of multiple interactions with each crossing of the counter-circulating
 753 beams. This leads to isotropic energy depositions within the detectors positioned at these
 754 interaction points, increasing the energy scale of the underlying event. This is known as
 755 pile-up and the counteracting of it's effects are important to the many measurements
 756 performed at the LHC.

757 In the early phase of prolonged operation after the initial shutdown the machine
 758 operated in 2010-2011 at 3.5 TeV per beam, $\sqrt{s} = 7 \text{ TeV}$, delivering 6.13 fb^{-1} of data
 759 [53]. During the 2012-2013 run period, data was collected at an increased $\sqrt{s} = 8 \text{ TeV}$
 760 improving the sensitivity of searches for new physics. Over the whole run period 23.3
 761 fb^{-1} of data was delivered of which 21.8 fb^{-1} was recorded by the CMS detector as shown
 762 in Figure 3.2 [53]. A total of 12 fb^{-1} of 8 TeV certified data was collected by October
 763 2012, and it is this data which forms the basis of the results discussed within this thesis.

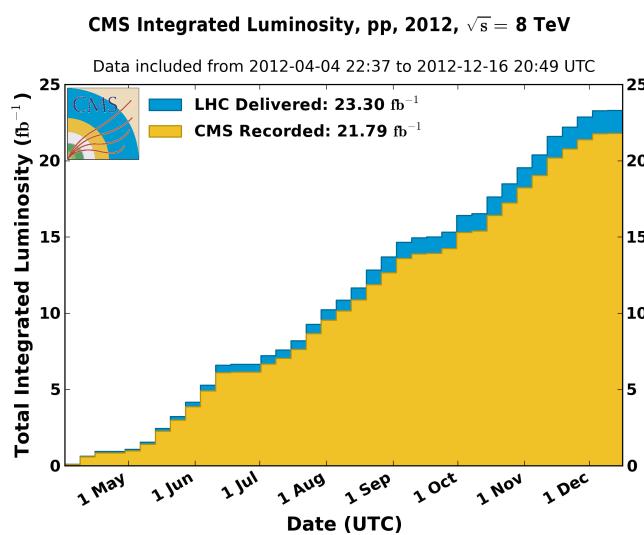


Figure 3.2.: The total integrated luminosity delivered to and collected by CMS during the 2012 8 TeV pp runs.

⁷⁶⁴ 3.2. The CMS Detector

⁷⁶⁵ The Compact Muon Solenoid (**CMS**) detector is one of two general purpose detectors
⁷⁶⁶ at the **LHC** designed to search for new physics. The detector is designed to provide
⁷⁶⁷ efficient identification and measurement of many physics objects including photons,
⁷⁶⁸ electrons, muons, taus, and hadronic showers over wide ranges of transverse momentum
⁷⁶⁹ and direction. Its nearly 4π coverage in solid angle allows for accurate measurement of
⁷⁷⁰ global transverse momentum imbalance. These design factors give **CMS** the ability to
⁷⁷¹ search for direct production of **SUSY** particles at the TeV scale, making the search for
⁷⁷² Supersymmetric particles one of the highest priorities among the wide range of physics
⁷⁷³ programmes at **CMS**.

⁷⁷⁴

⁷⁷⁵ **CMS** uses a right-handed Cartesian coordinate system with the origin at the interaction
⁷⁷⁶ point and the z-axis pointing along the beam axis, the x-axis points radially inwards to
⁷⁷⁷ the centre of the collider ring, with the y-axis points vertically upward. The azimuthal
⁷⁷⁸ angle, ϕ ranging between $[-\pi, \pi]$ is defined in the x-y plane starting from the x-axis. The
⁷⁷⁹ polar angle θ is measured from the z axis. The common convention in particle physics is
⁷⁸⁰ to express an out going particle in terms of ϕ and its pseudorapidity defined as

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

⁷⁸¹ The variable $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is commonly used to define angular distance
⁷⁸² between objects within the detector and additionally energy and momentum is typically
⁷⁸³ measured in the transverse plane perpendicular to the beam line. These values are
⁷⁸⁴ calculated from the x and y components of the object and are denoted as $E_T = E \sin \theta$
⁷⁸⁵ and $p_T = \sqrt{p_x^2 + p_y^2}$.

⁷⁸⁶ 3.2.1. Detector subsystems

⁷⁸⁷ As the range of particles produced in pp collisions interact in different ways with mat-
⁷⁸⁸ ter, **CMS** is divided into subdetector systems, which perform complementary roles to
⁷⁸⁹ identify the identity, mass and momentum of the different physics objects present in
⁷⁹⁰ each event. These detector sub-systems contained within **CMS** are wrapped in layers

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

794

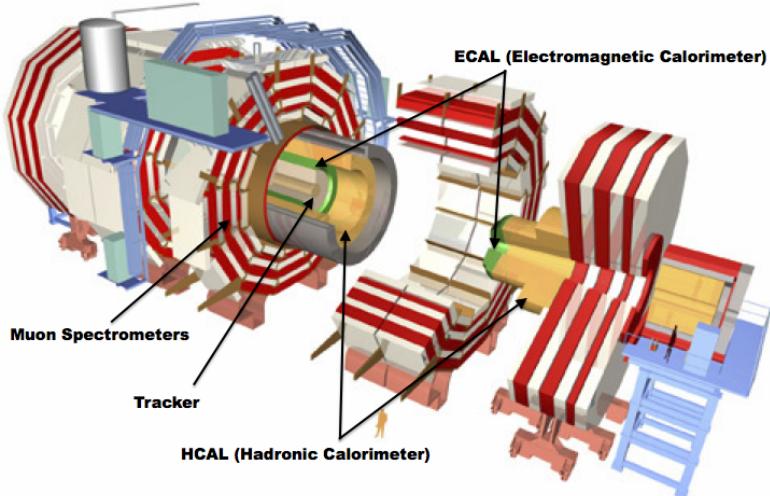


Figure 3.3.: A pictorial depiction of the CMS detector with the main detector subsystems labelled. [54]

795 3.2.2. Tracker

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

805

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

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

811 **3.2.3. Electromagnetic calorimeter**

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

817

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

825

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

834

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

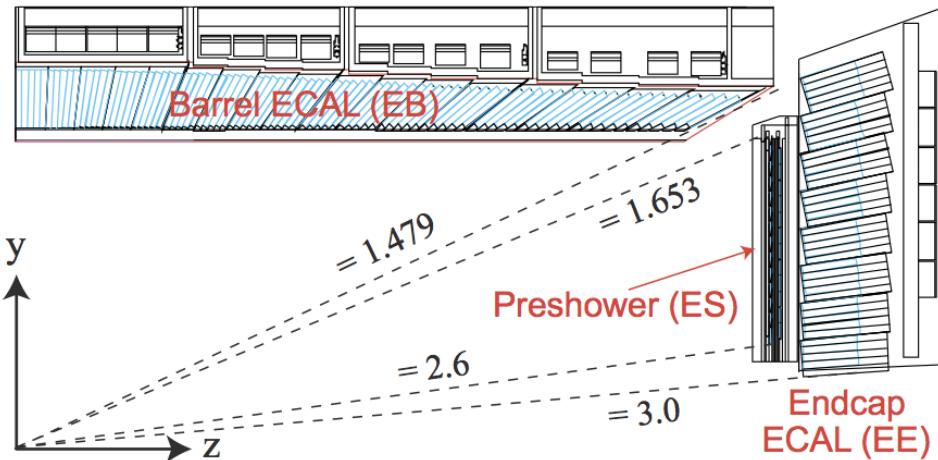


Figure 3.4.: Illustration of the CMS ECAL showing the arrangement of the lead tungstate crystals in the EB and EE. The ES is also shown and is located in front of the EE [55].

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

843 3.2.4. Hadronic calorimeter

844 Beyond the ECAL lies the Hadronic CALorimeter (HCAL) which is responsible for
 845 the accurate measurement of hadronic showers, crucial for analyses involving jets or
 846 missing energy signatures. The HCAL is a sampling calorimeter which consists of al-
 847 ternating layers of brass absorber and plastic scintillator, except in the hadron forward
 848 ($3.0 < |\eta| < 5.0$) region in which steel absorbers and quartz fibre scintillators are used
 849 because of their increased radiation tolerance. Hadron showers are initiated in the
 850 absorber layers inducing scintillation in the plastic scintillator tiles. These scintillation
 851 photons are converted by wavelength shifting fibres for read-out by hybrid photodiodes.
 852

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

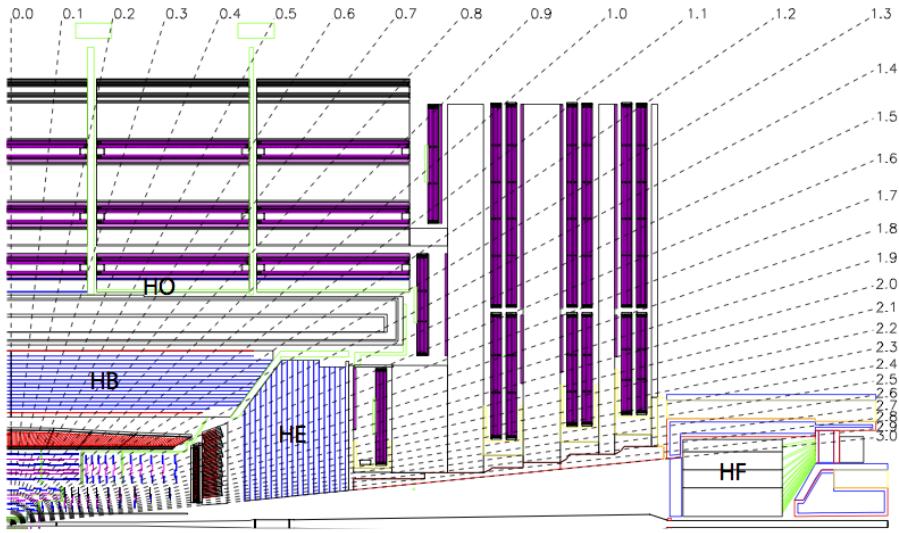


Figure 3.5.: Schematic of the hadron calorimeters in the r - z plane, showing the locations of the **HCAL** components and the **HF**. [50].

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

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

874 **3.2.5. Muon systems**

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

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

885

886 **3.3. Event Reconstruction and Object Definition**

887 The goal of event reconstruction is to take the raw information recorded by the detector
888 and to compute from it higher-level quantities which can be used at an analysis level.
889 These typically correspond to an individual particle’s energy and momenta, or groups of
890 particles which shower in a narrow cone and the overall global energy and momentum
891 balance of the event. The reconstruction of these objects are described in great detail in
892 [57], however covered below are brief descriptions of those which are most relevant to the
893 analysis detailed in Chapter 4.

894 **3.3.1. Jets**

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

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

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

906

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

914 PF jets are formed from combining information from all of the **CMS** subdetectors
915 systems to determine which final state particles are present in the event. Generally,
916 any particle is expected to produce some combination of a track in the silicon tracker,
917 a deposit in the calorimeters, or a track in the muon system. The PF jet momentum
918 and spatial resolutions are greatly improved with respect to calorimeter jets, as the use
919 of the tracking detectors and of the high granularity of **ECAL** allows resolution and
920 measurement of charged hadrons and photons inside a jet, which together constitute \sim
921 85% of the jet energy [61].

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

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

936 **3.3.2. B-tagging**

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

941

942 Many different algorithms developed by CMS select b-quark jets based on variables
943 such as the impact parameters of the charged-particle tracks, the properties of recon-
944 structed decay vertices, and the presence or absence of a lepton, or combinations thereof
945 [63]. One of the most efficient of which is the Combined Secondary Vertex (CSV) which
946 operates based on secondary vertex and track-based lifetime information, benchmarked
947 in ‘Loose’, ‘Medium’ and ‘Tight’ working points, of which the medium point is the tagger
948 used within the α_T search detailed in Section (4.1).

949 Using the CSV tagger, a likelihood-based discriminator distinguishes between jets
950 from b-quarks, and those from charm or light quarks and gluons, which is shown in Figure
951 3.6. The minimum thresholds on the discriminator for each working point correspond to
952 the misidentification probability for light-parton jets of 10%, 1%, and 0.1%, respectively,
953 in jets with an average p_T of about 80 GeV.

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

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

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

962

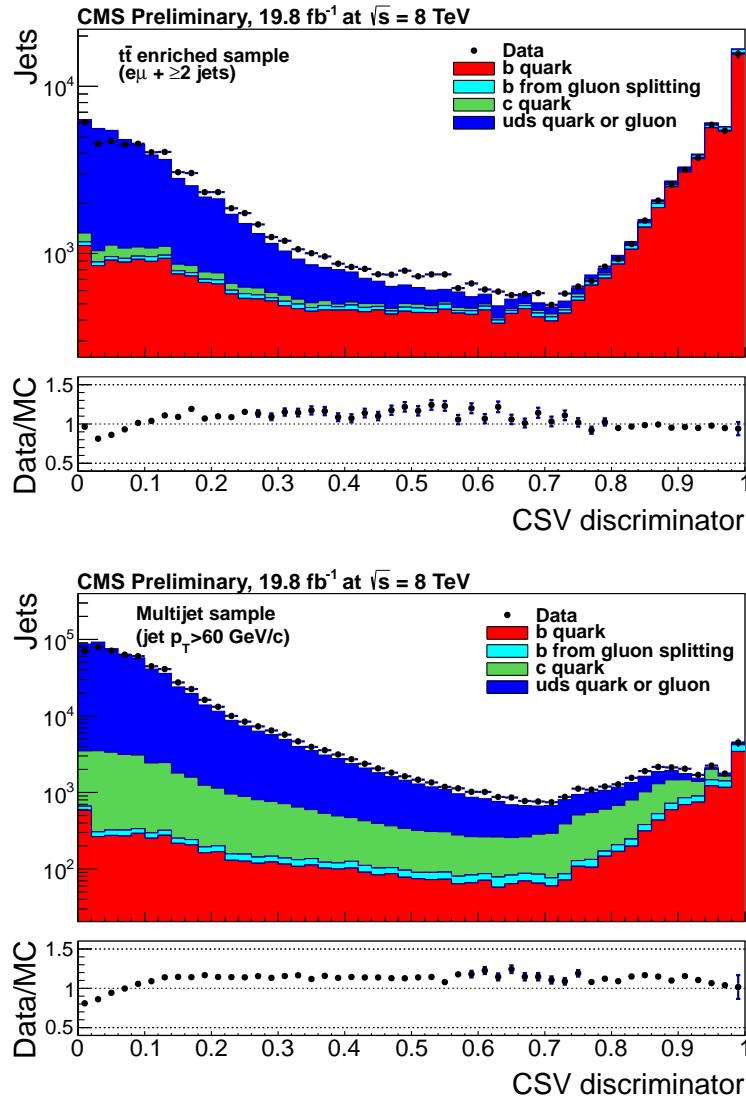


Figure 3.6.: CSV algorithm discriminator values in enriched $t\bar{t}$ (top) and inclusive multi-jet samples (bottom) for b,c and light flavoured jets [64]. Working points are determined from the misidentification probability for light-parton jets to be tagged as a b-jet and are given as 0.244, 0.679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

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

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

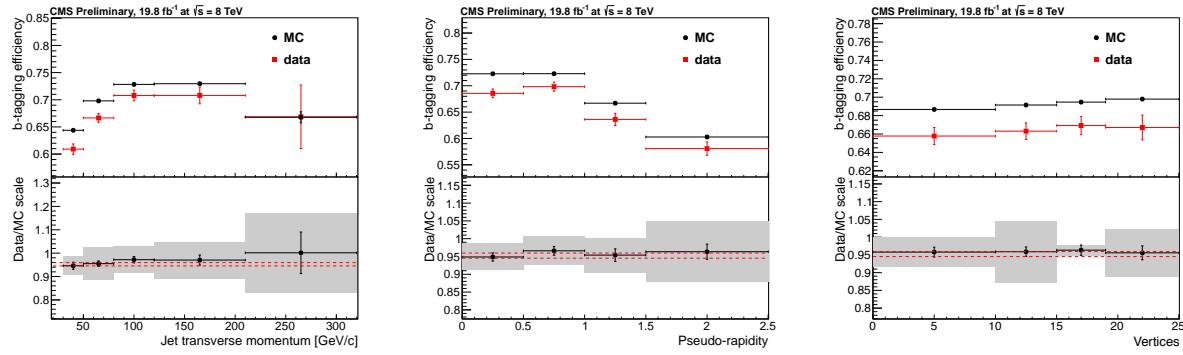


Figure 3.7.: Measured in $t\bar{t} \rightarrow$ di-lepton events using the CSVM tagger: (upper panels) b-tagging efficiencies and (lower panels) data/MC scale factor SF_b as a function of (left) jet p_T , (middle) jet $|\eta|$ and (right) number of primary vertices. In the lower panels, the grey filled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

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

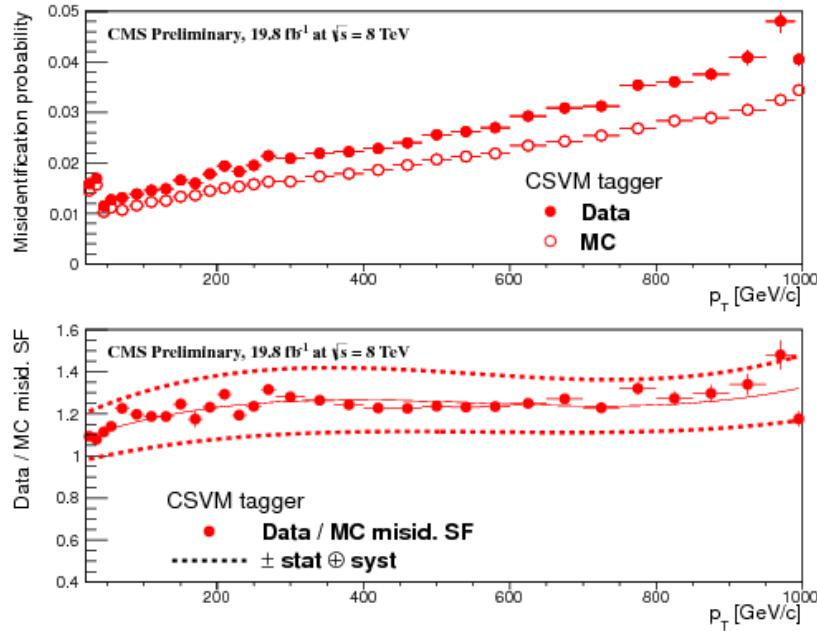


Figure 3.8.: For the **CSV**M tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last p_T bin in each plot includes all jets with $p_T > 1000$ GeV. The solid curve is the result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

981 3.4. Triggering System

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

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

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

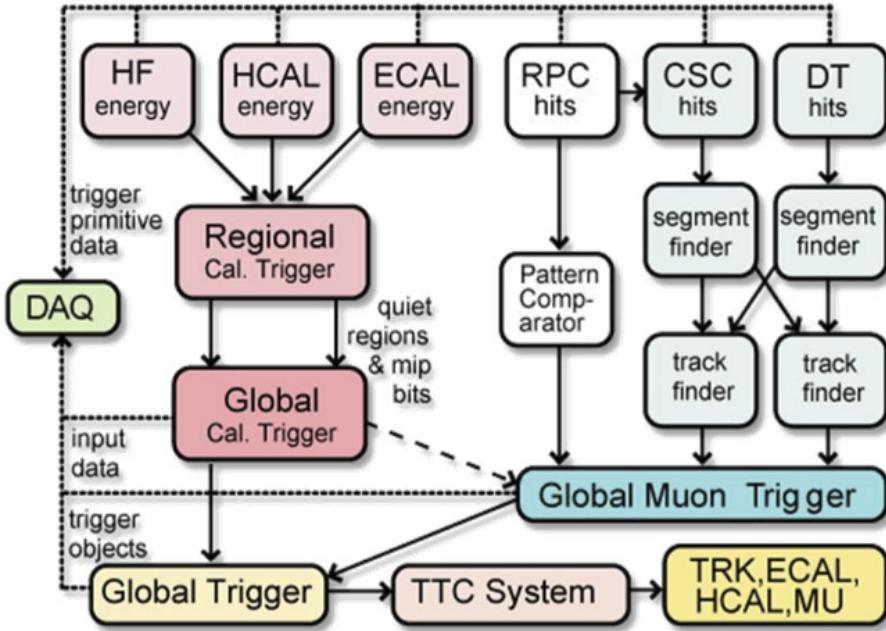


Figure 3.9.: The CMS L1 Trigger system.

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

3.4.1. The level-1 trigger

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

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

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

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

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

3.4.2. The L1 trigger jet algorithm

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

1041 In forming a L1 jet is it required that the central region to be higher than the eight
1042 neighbouring regions $E_{T\text{central}} > E_{T\text{surround}}$. Additionally a minimum threshold of 5 GeV
1043 on $E_{T\text{central}}$ was introduced during the 2012 run period to suppress noise from pile-up,
1044 the effects of which are shown in Section (3.4.4).

1045 The L1 jets are characterised by the E_T , summed over the 3×3 calorimeter regions,
1046 which corresponds to 12×12 trigger towers in barrel and endcap or 3×3 larger **HF**
1047 towers in the **HF**. The ϕ size of the jet window is the same everywhere, whilst the η
1048 binning gets somewhat larger at high η due to calorimeter and trigger tower segmentation.
1049 The jets are labelled by (η, ϕ) indexes of the central calorimeter region.

1050 Jets with $|\eta| > 3.0$ are classified as forward jets, whereas those with $|\eta| < 3.0$ are
1051 classified as central. The four highest energy central, forward and τ jets in the calorimeter
1052 are passed through Look Up Table (**LUT**)'s, which apply a programmable η -dependent
1053 jet energy scale correction. These are then used to make L1 trigger decisions.

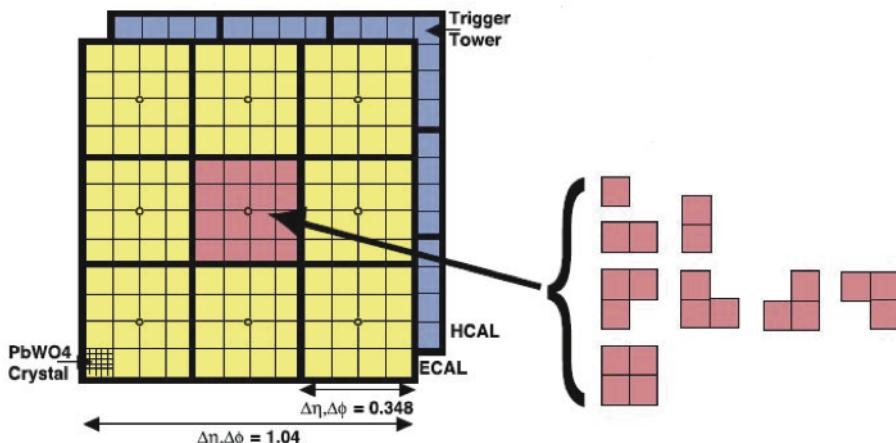


Figure 3.10.: Illustration of the Level-1 jet finding algorithm.

1054 The performance of the L1 jets is evaluated with respect to offline jets, which are
1055 taken from the standard Calo jet and the PF jet reconstruction algorithms of **CMS**.
1056 Jets are corrected for pile-up and detector effects as described in 3.3.1. A moderate
1057 level of noise rejection is applied to the offline jets by selecting jets passing the “loose
1058 identification criteria for both Calo and PF. These criteria are summarised in Appendix
1059 (A).

1060 **3.4.3. Measuring L1 jet trigger efficiencies**

1061 The L1 jet efficiency is defined as the fraction of leading offline jets which were matched
1062 with a L1 tau or central jet above a certain trigger threshold, divided by all the leading
1063 offline jets in the event. This quantity is then plotted as a function of the offline jet E_T ,
1064 η and ϕ .

1065 The efficiency is determined by matching the L1 and reconstructed offline jets spatially
1066 in $\eta - \phi$ space. This is done by calculating the minimum separation in ΔR between the
1067 highest offline reconstructed jet in E_T ($E_T > 10$ GeV, $|\eta| < 3$) and any L1 jet. A jet will
1068 be matched if this value is found to be < 0.5 . Should more than one jet satisfy this, the
1069 jet closest in ΔR is taken as the matched jet. The matching efficiency is close to 100%,
1070 above 30(45) GeV for run 2012B(C) data (see Appendix B.1).

1071 Each efficiency curve is fitted with a function which is the cumulative distribution
1072 function 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:

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

1074 In this functional form, the parameter μ determines the point of 50% of the plateau
1075 efficiency and the σ gives the resolution. This parametrisation is used to benchmark
1076 the efficiency at the plateau, the turn-on points and resolution for each L1 Jet trigger.
1077 The choice of function is purely empirical. Previous studies used the error function
1078 alone, which described the data well at high threshold values but could not describe the
1079 efficiencies well at lower thresholds [68].

1080 The efficiency turn-on curves for various L1 jet thresholds are evaluated as a function
1081 of the offline reconstructed jet E_T for central jets with $|\eta| < 3$. These are measured using
1082 single isolated μ triggers which have high statistics, and are orthogonal and therefore
1083 unbiased to the hadronic triggers under study. The efficiency is calculated with respect to
1084 offline Calo and PF Jets in Figure 3.11. Table 3.1 shows the values of these parameters,
1085 calculated for three example L1 single jet triggers taken from 2012 8 TeV data.

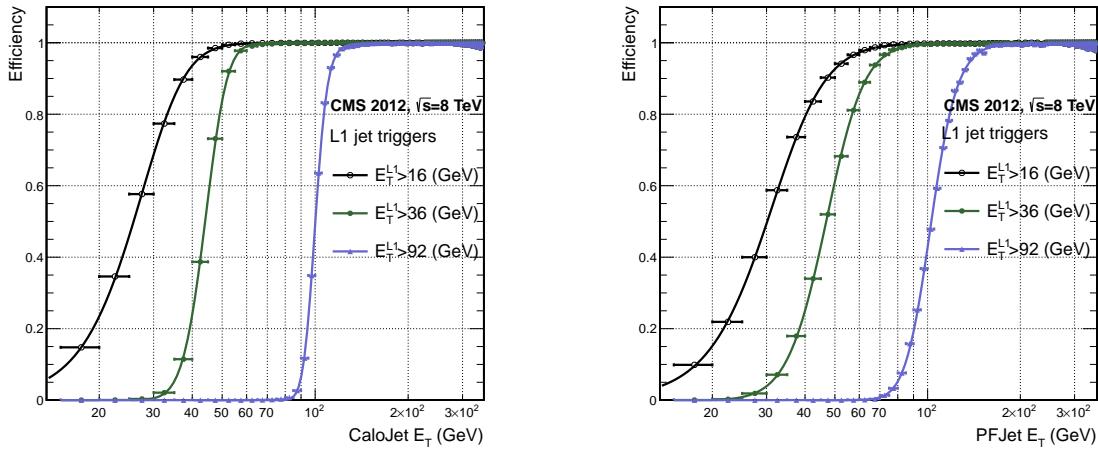


Figure 3.11.: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T (left) and PFJet E_T (right), measured in 2012 Run Period C data and collected with an isolated single μ data sample.

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

Table 3.1.: Results of a cumulative EMG function fit to the turn-on curves for L1 single jet triggers in run 2012 Run Period C, measured in an isolated 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

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

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

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

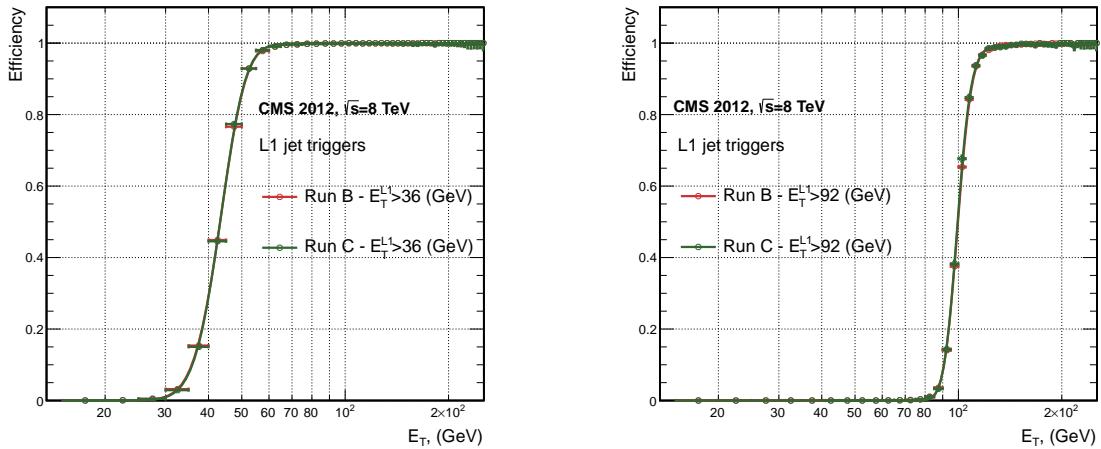


Figure 3.12.: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T , measured for the L1 SingleJet 36 and 92 trigger in 2012 run period B and C collected with an isolated single μ' sample.

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

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

Trigger	2012B		2012C	
	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.

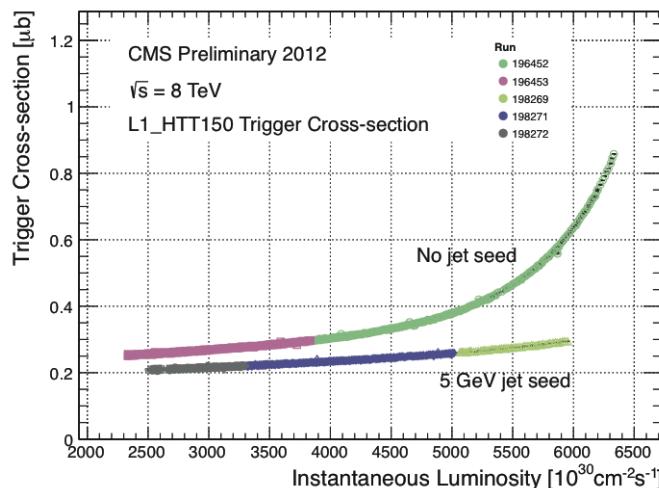


Figure 3.13.: Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1 H_T triggers [70].

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 at CMS in 2012. This is benchmarked relative to Calo and PF jets for the run 2012 C

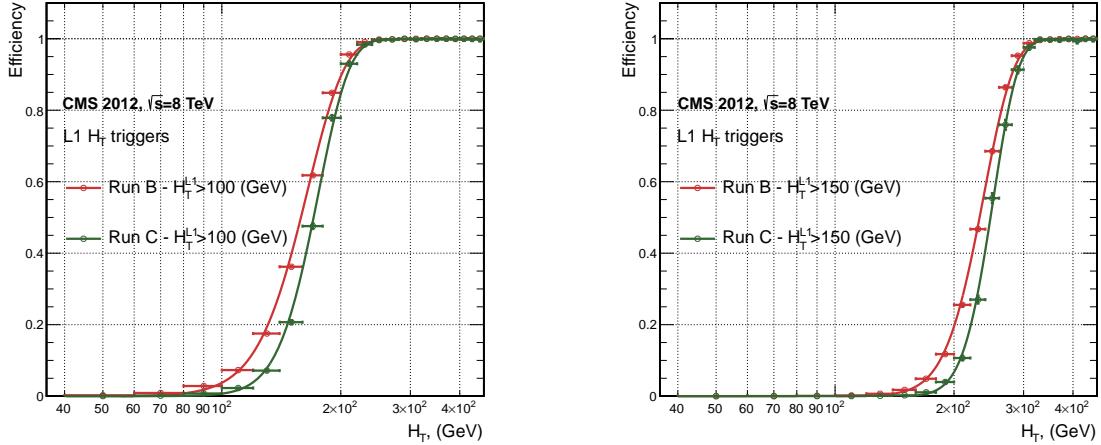


Figure 3.14.: L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T , measured for the L1 H_T 100 and 150 trigger during the run 2012 B and C collected using an isolated single μ triggered sample.

Trigger	2012B		2012C	
	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).

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.

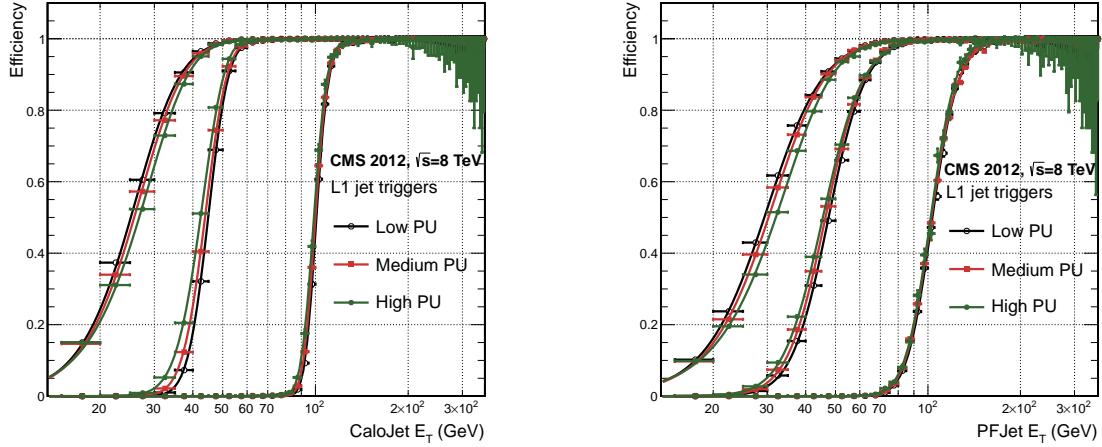


Figure 3.15.: L1 jet efficiency turn-on curves as a function of the leading offline E_T Calo (left) and PF (right) jet, for low, medium and high pile-up conditions.

Vertices	0-10		11-20		> 20	
	μ	σ	μ	σ	μ	σ
L1_SingleJet16	21.1 ± 0.1	7.16 ± 0.05	22.34 ± 0.1	7.9 ± 0.1	24.6 ± 0.2	9.5 ± 0.1
L1_SingleJet36	39.6 ± 0.1	7.4 ± 0.1	38.4 ± 0.1	7.4 ± 0.1	37.1 ± 0.2	7.5 ± 0.1
L1_SingleJet92	91.6 ± 0.3	11.3 ± 0.2	90.4 ± 0.3	11.2 ± 0.1	92.0 ± 0.9	12.1 ± 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.

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

These features are further emphasised when shown as a function of

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

1136 in bins of matched leading offline jet E_T , of which the individual fits can be found in
1137 Appendix (B.2). Each of these distributions are fitted with an EMG function as defined
1138 in Equation (3.3).

1139 The μ , σ and λ values extracted for the low, medium and high pile-up conditions are
1140 shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central value
1141 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
1142 is also observed to improve at higher offline jet E_T .

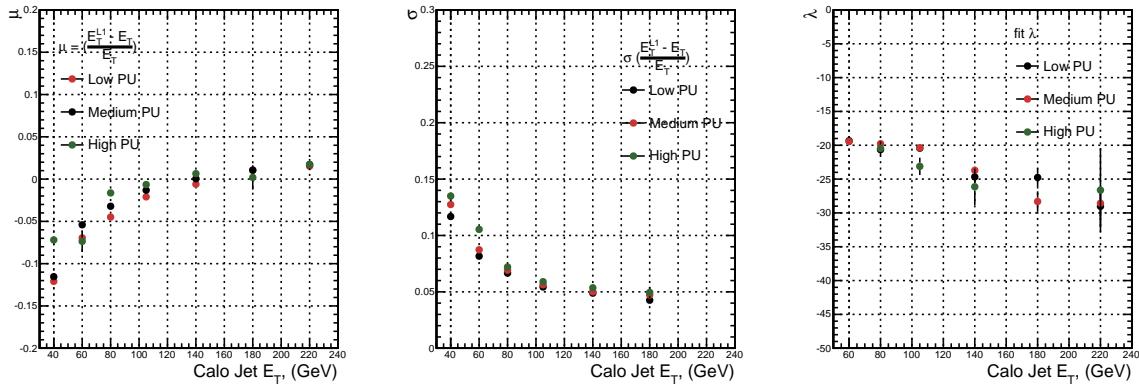


Figure 3.16.: Fit values from an EMG function fitted to the resolution plots of leading Calo jet E_T measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low, medium and high pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian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

1143 The resolution of other L1 energy sum quantities, H_T , \not{E}_T and $\sum E_T$ parameterised
1144 as in Equation (3.4), can be found in Appendix B.3. The same behaviour observed for
1145 the single jet triggers is also found for these quantities, where in the presence of higher
1146 pile-up the μ values are shifted to higher values, with a worsening resolution, σ again
1147 due to the increase in soft pile-up jets and the absence of pile-up subtraction at L1.

1148 3.4.6. Summary

1149 The performance of the CMS Level-1 Trigger has been studied and evaluated for jets and
1150 energy sum quantities using data collected during the 2012 LHC 8TeV run. These studies
1151 include the effect of introduction of a 5 GeV jet seed threshold into the jet algorithm
1152 configuration, the purpose of which is to mitigate the effects of pile-up on the rate of
1153 L1 triggers whilst not adversely affecting the efficiency of these triggers. No significant

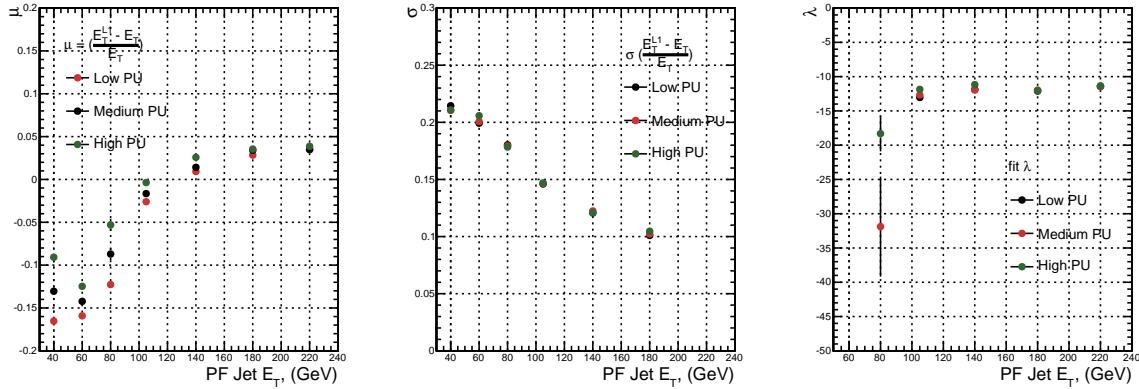


Figure 3.17.: Fit values from an **EMG** function fitted to the resolution plots of leading PF jet E_T measured as a function of $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low and medium pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian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

1154 change in performance is observed with this change and good performance is observed
 1155 for a range of L1 quantities.

Chapter 4.

¹¹⁵⁶ SUSY Searches In Hadronic Final States

¹¹⁵⁸ In this chapter a model independent search for **SUSY** in hadronic final states with \cancel{E}_T
¹¹⁵⁹ using the α_T variable and b-quark multiplicity is introduced and described in detail. The
¹¹⁶⁰ results presented are based on a data sample of pp collisions collected in 2012 at $\sqrt{s} = 8$
¹¹⁶¹ TeV, corresponding to an integrate luminosity of $11.7 \pm 0.5 \text{ fb}^{-1}$.

¹¹⁶² The kinematic variable α_T is motivated as a variable to provide strong rejections
¹¹⁶³ of QCD backgrounds, whilst maintaining sensitivity to a possible **SUSY** signal within
¹¹⁶⁴ Section (4.1). The search and trigger strategy in addition to the event reconstruction
¹¹⁶⁵ and selection are outlined within Sections (4.2-4.3).

¹¹⁶⁶ The method in which the **SM** background is estimated using an analytical technique
¹¹⁶⁷ to improve statistical precision at higher b-tag multiplicities is detailed within Section
¹¹⁶⁸ (4.5), with a discussion on the impact of b-tagging and mis-tagging scale factors between
¹¹⁶⁹ data and MC on any background predictions.

¹¹⁷⁰ A description of the formulation of appropriate systematic uncertainties applied to
¹¹⁷¹ the background predictions to account for theoretical uncertainties and limitations in the
¹¹⁷² simulation modelling of event kinematics and instrumental effects is covered in Section
¹¹⁷³ (4.6).

¹¹⁷⁴ Finally the statistical likelihood model to interpret the observations in the signal and
¹¹⁷⁵ control samples is described in Section (4.7). The experimental reach of the analysis
¹¹⁷⁶ discussed within this thesis is interpreted in two classes of **SMS** models, the topologies
¹¹⁷⁷ of which are detailed in Section (2.4.1). The **SMS** models considered in this analysis

1178 are summarised in Table 4.1. For each model, the **LSP** is assumed to be the lightest
1179 neutralino.

1180 Within the table are also defined reference points, parameterised in terms of parent
1181 gluino/squark and **LSP** sparticle masses, m_{parent} and m_{LSP} , respectively, which are used
1182 within the following chapters to demonstrate potential yields within the signal region
1183 of the search. The masses are chosen to reflect parameter space which is within the
1184 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 (T1bb)	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow b\bar{b}\tilde{\chi}_1^0 b\bar{b}\tilde{\chi}_1^0$	900	500
G3 (T1tt)	$pp \rightarrow \tilde{g}\tilde{g}^* \rightarrow t\bar{t}\tilde{\chi}_1^0 t\bar{t}\tilde{\chi}_1^0$	850	250
D1 (T2)	$pp \rightarrow \tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 q\tilde{\chi}_1^0$	600	250
D2 (T2bb)	$pp \rightarrow \tilde{b}\tilde{b}^* \rightarrow b\tilde{\chi}_1^0 b\tilde{\chi}_1^0$	500	150
D3 (T2tt)	$pp \rightarrow \tilde{t}\tilde{t}^* \rightarrow t\tilde{\chi}_1^0 t\tilde{\chi}_1^0$	400	0

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

1185 4.1. An Introduction to the α_T Search

1186 The experimental signature of **SUSY** signal in the hadronic channel would manifest
1187 itself as a final state containing energetic jets and \cancel{E}_T . The search focuses on topologies
1188 where new heavy supersymmetric, R-parity conserving particles are pair-produced in pp
1189 collisions. These particles decaying to a **LSP** escape the detector undetected, leading to
1190 significant missing energy and missing hadronic transverse energy,

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

1191 defined as the vector sum of the transverse energies of jets selected in an event.
1192 Energetic jets produced in the decay of these supersymmetric particles also can produce
1193 significant visible transverse energy,

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

1194 defined as the scalar sum of the transverse energies of jets selected in an event.

1195 A search within this channel is greatly complicated in a hadron collider environment,
1196 where the overwhelming background comes from inherently balanced multi-jet (“QCD”)
1197 events which are produced with an extremely large cross section as demonstrated within
1198 Figure 4.1. \cancel{E}_T can appear in such events with a substantial mis-measurement of jet
1199 energy or missed objects due to detector miscalibration or noise effects.

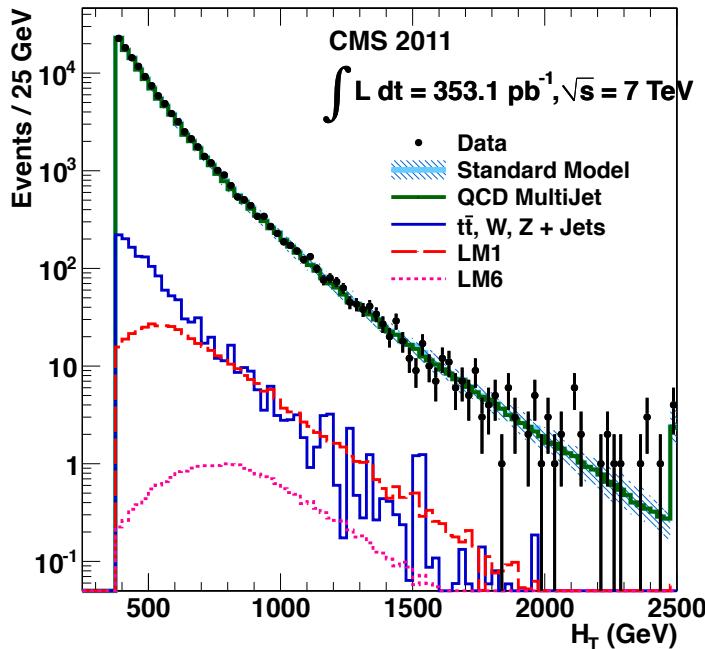


Figure 4.1.: Reconstructed offline H_T for 11.7fb^{-1} of data after a basic pre-selection. Sample is collected from prescaled H_T triggers. Overlaid are expectations from MC simulation of EWK processes as well as a reference signal model (labelled D2 from Table.4.1).

1200 Additional SM background from EWK processes with genuine \cancel{E}_T from escaping
1201 neutrinos comprise the irreducible background within this search and come mainly from:

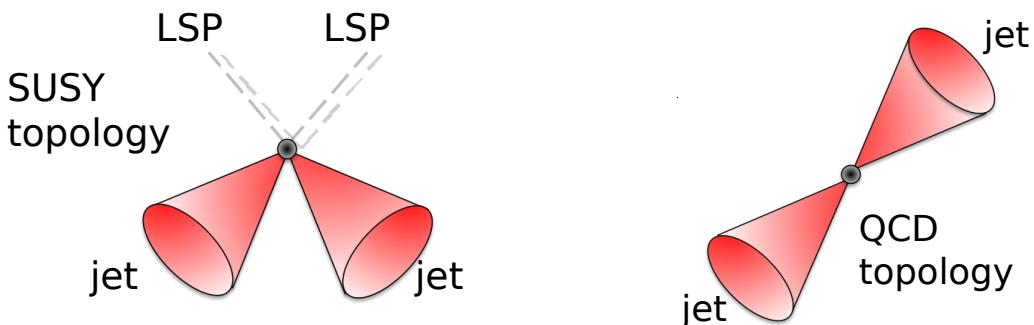
- 1202 • $Z \rightarrow \nu\bar{\nu} + \text{jets}$,

- 1203 • $W \rightarrow l\nu + \text{jets}$ in which a lepton falls outside of detector acceptance, or the lepton
 1204 decays hadronically $\tau \rightarrow \text{had}$,
- 1205 • $t\bar{t}$ with at least one leptonic W decay,
- 1206 • small background contributions from DY, single top and Diboson (WW,ZZ,WZ)
 1207 processes.

1208 The search is designed to have a strong separation between events with genuine and
 1209 “fake” \cancel{E}_T which is achieved primarily through the dimensionless kinematic variable, α_T
 1210 [71][72].

1211 4.1.1. The α_T variable

1212 For a perfectly measured di-jet QCD event, conservation laws dictate that they must be
 1213 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.



1214 **Figure 4.2.:** The event topologies of background QCD di-jet events (right) and a generic SUSY
 1215 signature with genuine \cancel{E}_T (left).

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

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

1216 where E_T^{j2} is the transverse energy of the least energetic of the two jets and M_T
 1217 defined as:

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

1218 A perfectly balanced di-jet event i.e. $E_T^{j_1} = E_T^{j_2}$ would give an $\alpha_T = 0.5$, where as
 1219 events with jets which are not back-to-back, for example in events in which a W or Z
 1220 recoils off a system of jets, α_T can achieve values in excess of 0.5.

1221 α_T can be extended to apply to any arbitrary number of jets, undertaken by modelling
 1222 a system of n jets as a di-jet system, through the formation of two pseudo-jets [73].
 1223 The two pseudo-jets are built by merging the jets present in the event such that the 2
 1224 pseudo-jets are chosen to be as balanced as possible, i.e the $\Delta H_T \equiv |E_T^{pj_1} - E_T^{pj_2}|$ is
 1225 minimised between the two pseudo jets. Using Equation (4.4), α_T can be rewritten as,

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

1226 The distribution of α_T for the two jet categories used within this analysis, 2,3 and
 1227 ≥ 4 jets, is shown in the Figure 4.3, demonstrating the ability of the α_T variable to
 1228 discriminate between multi jet events and EWK processes with genuine \cancel{E}_T in the final
 1229 state.

1230 The α_T requirement used within the search is chosen to be $\alpha_T > 0.55$ to ensure
 1231 that the QCD multijet background is negligible even in the presence of moderate jet
 1232 mis-measurement. There still remains other effects which can cause multijet events to
 1233 artificially have a large α_T value, which are discussed in detail in Section (4.2.2).

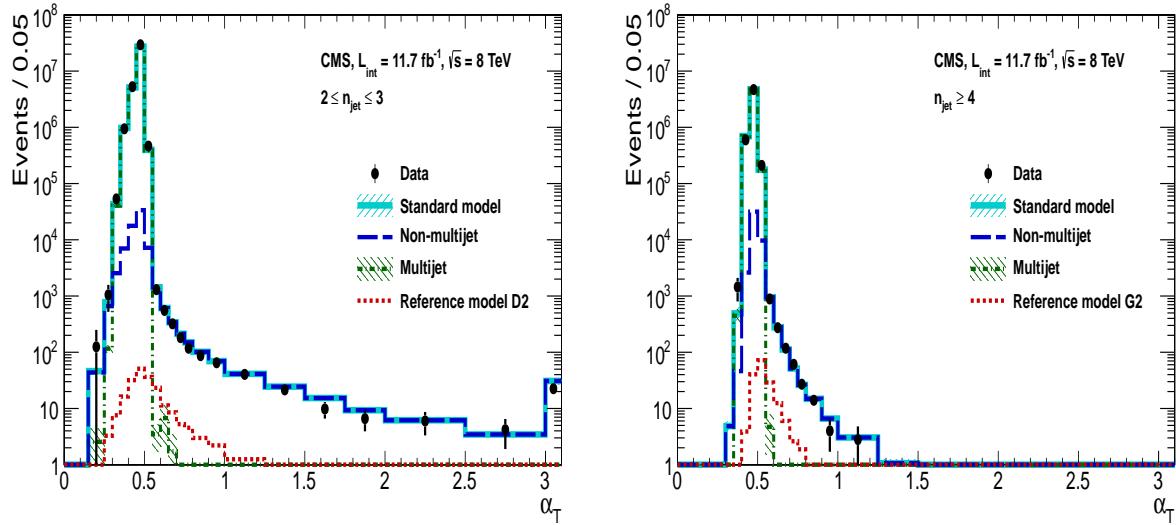


Figure 4.3.: The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet multiplicities after a full analysis selection and shown for $H_T > 375$. Data is collected using both prescaled H_T triggers and dedicated α_T triggers for below and above $\alpha_T = 0.55$ respectively. . Expected yields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

1247 To remain inclusive to a large range of possible **SUSY** models, the signal region is
1248 binned in the following categories to allow for increased sensitivity in the interpretation
1249 of results for different **SUSY** topologies:

1250 **Sensitivity to a range of SUSY mass splittings**

1251 The hadronic signal region is defined by $H_T > 275$, divided into eight bins in H_T .

- 1252 – Two bins of width 50 GeV in the range $275 < H_T < 375$ GeV,
1253 – five bins of width 100 GeV in the range $375 < H_T < 875$ GeV,
1254 – and a final open bin, $H_T > 875$ GeV.

1255 The choice at low H_T is driven primarily by trigger constraints. The mass difference
1256 between the **LSP** and the particle that it decays from is an important factor in the
1257 amount of hadronic activity in the event.

1258 A large mass splitting will lead to hard high p_T jets which contribute to the H_T
1259 sum. From Figure 4.1 it can be seen that the **SM** background falls sharply at high
1260 H_T values, therefore a large number of H_T bins will lead to easier identification
1261 of such signals. Conversely smaller mass splittings lead to softer jet p_T 's which will
1262 subsequently fall into the lower H_T range.

1263 **Sensitivity to production method of SUSY particles**

1264 The production mechanism of any potential **SUSY** signal can lead to different event
1265 topologies. One such way to discriminate between gluino ($g\tilde{g}$ - “high multiplicity”),
1266 and direct squark ($q\tilde{q}$ - “low multiplicity”) induced production of **SUSY** particles is
1267 realised through the number of reconstructed jets in the final state.

1268 The analysis is thus split into two jet categories : 2-3 jets , ≥ 4 jets to give sensitivity
1269 to both of these mechanisms.

1270 **Sensitivity to “Natural SUSY” via tagging jets from b-quarks**

1271 Jets originating from bottom quarks (b-jets) are identified through vertices that
1272 are displaced with respect to the primary interaction. The algorithm used to tag
1273 b-jets is the Combined Secondary Vertex Medium Working Point (**CSV**) tagger,
1274 described within Section (3.3.2). A cut is placed on the discriminator variable of
1275 > 0.679 , leading to a gluon/light-quark mis-tag rate of 1% and a jet p_T dependant
1276 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 (4.7), 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.

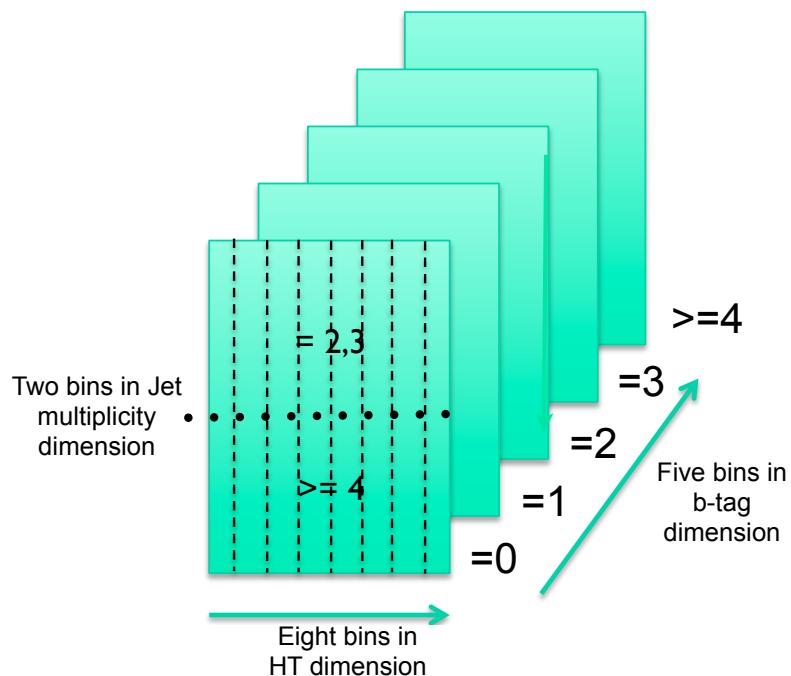


Figure 4.4.: Pictorial depiction of the analysis strategy employed by the α_T search to increase sensitivity to a wide spectra of SUSY models.

1293 **4.2.1. Physics objects**

1294 The physics objects used in the analysis defined below, follow the recommendation of
1295 the various CMS Physics Object Groups ([POGs](#)).

1296 • **Jets**

1297 The jets used in this analysis are CaloJets, reconstructed as described in Section
1298 ([3.3.1](#)) using the anti- k_T jet clustering algorithm.

1299 To ensure the jet object falls within the calorimeter systems a pseudo-rapidity
1300 requirement of $|\eta| < 3$ is applied. Each jet must pass a “loose” identification criteria
1301 to reject jets resulting from unphysical energy, the criteria of which are detailed in
1302 Table A.1 of Appendix A [[74](#)].

1303 • **Muons**

1304 Muons are selected in the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples, and vetoed in
1305 the signal region. The same cut based identification criteria is applied to muons in
1306 both search regions and is summarised in Table 4.2 [[75](#)].

Categories	Criteria
Global Muon	True
PFMuon	True
χ^2	< 10
Muon chamber hits	> 0
Muon station hits	> 1
Transvere impact d_{xy}	< 0.2mm
Longitudinal distance d_z	< 0.5mm
Pixel hits	> 0
Track layer hits	> 5
PF Isolation (DeltaB corrected)	<0.12

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

1307 Additionally muons are required to be within the acceptance of the muon tracking
1308 systems. For the muon control samples, trigger requirements necessitate a $|\eta| <$
1309 2.1 for the selection of muons. In the signal region where muons are vetoed these

1310 conditions are relaxed to $|\eta| < 2.5$ and a minimum threshold of $p_T > 10$ GeV is
1311 required of muon objects.

1312 • **Photons**

1313 Photons are selected within the $\gamma + \text{jets}$ control sample and vetoed in all other
1314 selections. Photons are identified in both cases according to the cut based criteria
1315 listed in Table 4.3 [76].

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.
$\sigma_{i\eta i\eta} < 0.011$	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 [77], 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 $\gamma + \text{jets}$ control/signal selections.

1316 Photon objects are also required to have a minimum momentum of $p_T > 25$ GeV.

1317 • **Electrons**

1318 Electron identification is defined for veto purposes. They are selected according to
1319 the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

1320 Electrons are required to be identified at $|\eta| < 2.5$, with a minimum $p_T > 10$ GeV
1321 threshold to ensure that the electron falls within the tracking system of the detector.

1322 • **Noise and E_T Filters**

1323 A series of Noise filters are applied to veto events which contain spurious non-physical
1324 jets that are not picked up by the jet id, and events which give large unphysical E_T
1325 values. These filters are listed within Table A.3 of Appendix A.

Categories	Barrel	EndCap
$\Delta\eta_{In}$	0.007	0.009
$\Delta\phi_{In}$	0.15	0.10
$\sigma_{in\eta}$	0.01	0.03
H/E	0.12	0.10
d0 (vtx)	0.02	0.02
dZ (vtx)	0.20	0.20
$ (1/E_{ECAL} - 1/p_{track}) $	0.05	0.05
PF Combined isolation/ p_T	0.15	0.15
Vertex fit probability	10^{-6}	10^{-6}

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

1326 4.2.2. Event selection

1327 The selection criteria for events within the analysis are detailed below. A set of common
1328 cuts are applied to both signal (maximise acceptance to a range of **SUSY** signatures), and
1329 control samples (retain similar jet kinematics for background predictions), with additional
1330 selection cuts applied to each control sample to enrich the sample in a particular **EWK**
1331 processes, see Section (4.2.3).

1332 The jets considered in the analysis are required to have a transverse momentum $p_T >$
1333 50 GeV, with a minimum of two jets required in the event. The highest E_T jet is required
1334 to lie within the central tracker acceptance $|\eta| < 2.5$, and the two leading p_T jets must
1335 each have $p_T > 100$ GeV. Any event which has a jet with $p_T > 50$ GeV that either fails
1336 the “loose” identification criteria described in Section(4.2.1) or has $|\eta| > 3.0$, is rejected.
1337 Similarly events in which an electron, muon or photon fails object identification but pass
1338 η and p_T restrictions are identified as an “odd” lepton/photon and the event is vetoed.

1339 At low H_T , the jet threshold requirements applied to be considered as part of the
1340 analysis and enter the H_T sum are scaled downwards. These are scaled down in order to
1341 not restrict phase space, preserving jet multiplicities and background admixture in the
1342 lower H_T bins, as listed in Table 4.5.

1343 Within the signal region to suppress **SM** processes with genuine E_T' from neutrinos,
1344 events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure
1345 multi-jet topology, events are vetoed if an isolated photon is found with $p_T > 25$ GeV.

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.

1346 An α_T requirement of > 0.55 is required to reduce the QCD multi-jet background
1347 to a negligible amount. Finally additional cleaning cuts are applied to protect against
1348 pathological deficiencies such as reconstruction failures or severe energy mis-measurements
1349 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,

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

1350 • Fake \cancel{E}_T and \cancel{H}_T can arise due to significant jet mis-measurements caused by a small
1351 number of non-functioning **ECAL** regions. These regions absorb electromagnetic
1352 showers which are subsequently not added to the jet energy sum. To circumvent
1353 this problem the following procedure is employed : For each jet in the event, the
1354 angular separation

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

1355 is calculated where that jet is itself removed from the event. Here $\Delta\phi^*$ is a measure
1356 of how aligned the \cancel{H}_T of an event is with a jet, a small value is compatible with the
1357 hypothesis of an inherently balanced event in which a jet has been mis-measured.
1358 For every jet in an event with $\Delta\phi^* < 0.5$, if the ΔR distance between the selected jet
1359 and the closest dead **ECAL** region is also < 0.3 , then the event is rejected. Similarly

1360 events are rejected if the jet points within $\Delta R < 0.3$ of the **ECAL** barrel-endcap
1361 gap at $|\eta| = 1.5$.

1362 Some of the key distributions of the data used in this analysis compared to MC
1363 simulation are shown in Figure 4.5. The MC samples are normalised to a luminosity of
1364 11.7 fb^{-1} , with no requirement placed upon the number of b-tagged jets or number of
1365 jets in the events.

1366 The distributions shown are presented for purely illustrative purposes, with the MC
1367 simulation itself not used in absolute term to estimate the yields from background
1368 processes, see Sections (4.2.3,4.5). However it is nevertheless important to demonstrate
1369 that good agreement exists between simulation and observation in data.

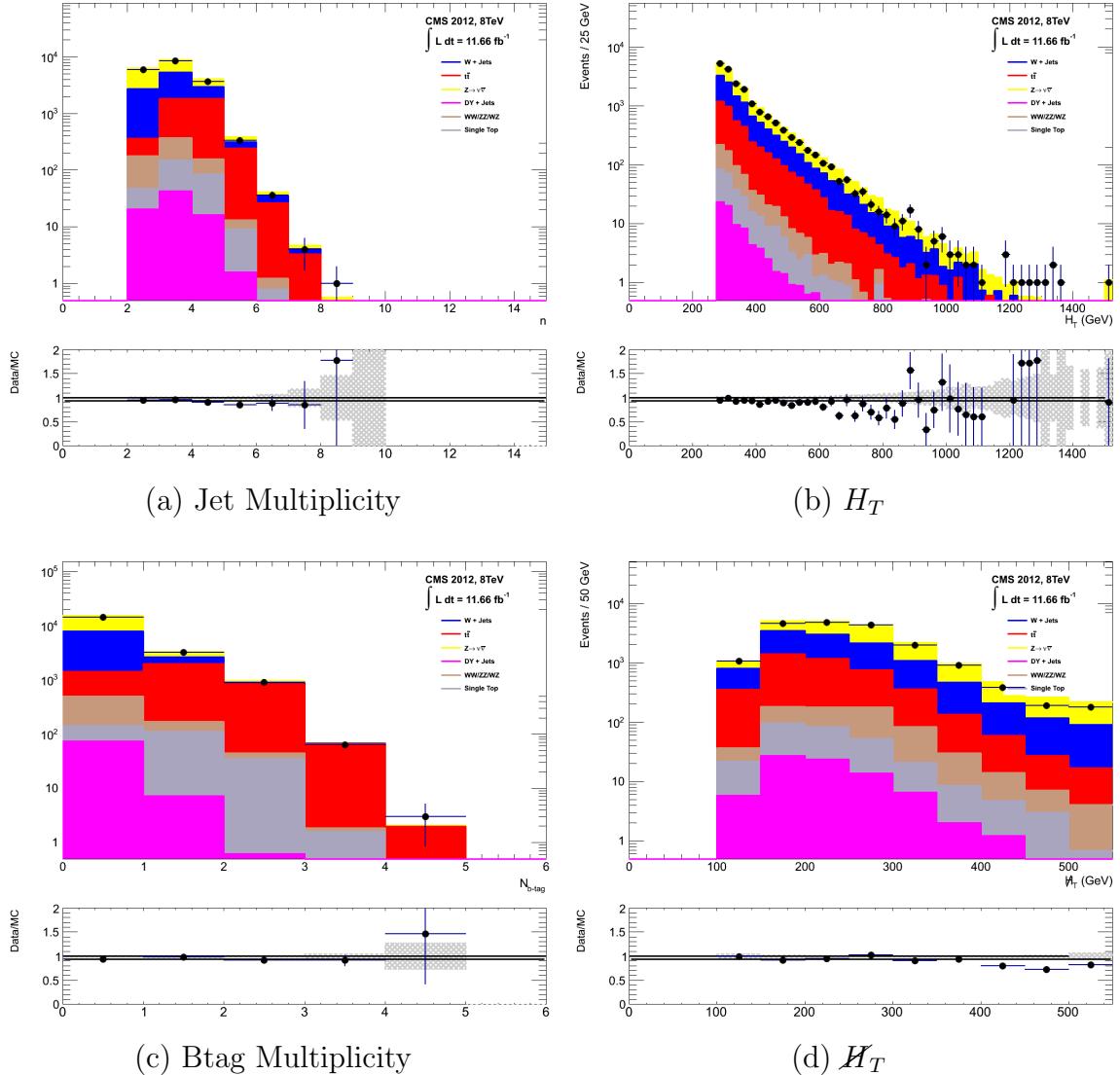


Figure 4.5.: Data/MC comparisons of key variables for the hadronic signal region, following the application of the hadronic selection criteria and the requirements of $H_T > 275$ GeV and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement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

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

1377 All MC samples are normalised to the luminosity of the data samples, 11.7 fb^{-1} .
1378 Through this method, “vanilla” predictions for the **SM** background in the signal region
1379 can be made by considering separately the sum of the prediction from either the $\mu + \text{jets}$
1380 and $\gamma + \text{jets}$ or $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples. However the final background estimation
1381 from which results are interpreted, is calculated via a fitting procedure defined formally
1382 by the likelihood model described in Section (4.7).

1383 The sum of the expected yields from all MC processes, in each control sample enter
1384 the denominator, $N_{\text{MC}}^{\text{control}}$, of the **TF** defined in Eq (4.7). However for the numerator
1385, $N_{\text{MC}}^{\text{signal}}$, only the relevant processes that the control sample is used in estimating a
1386 background for, enter into the **TF**.

1387 For the $\mu + \text{jets}$ sample the simulated MC processes which enter the numerator of
1388 the **TF** are,

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

1389 whilst for both the $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ samples the only MC process used in the
1390 numerator is,

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

1391 The control samples and the **EWK** processes they are specifically tuned to select
1392 are defined below, with distributions of key variables for each of the control samples
1393 shown for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed
1394 upon the number of b-tagged jets or jet multiplicity in the distributions shown. The
1395 MC distributions highlight the background compositions of each control sample, where
1396 in general, good agreement is observed between data and simulation, giving confidence

1397 that the samples are well understood. The contribution from QCD multi-jet events is
1398 expected to be negligible :

1399 **The $\mu +$ jets control sample**

1400 Events from $W +$ jets and $t\bar{t}$ processes enter into the hadronic signal sample due to
1401 unidentified leptons from acceptance or threshold effects and hadronic tau decays.
1402 These leptons originate from the decay of high p_T W bosons.

1403 The control samples specifically identifies $W \rightarrow \mu\bar{\nu}$ decays within the same phase-
1404 space of the signal region, where the muon is subsequently ignored in the calculation
1405 of event level variables, i.e. H_T , \mathcal{H}_T , α_T . All kinematic jet-based cuts are identical
1406 to those applied in the hadronic search region detailed in Section (4.2.2), with the
1407 same H_T , jet multiplicity and b-jet multiplicity binning described above.

- 1408 – Muons originating from W boson decays are selected by requiring one tightly
1409 isolated muon defined in Table 4.2, with a $p_T > 30$ GeV and $|\eta| < 2.1$. Both of
1410 these threshold arise from trigger restrictions.
- 1411 – The transverse mass of the W candidate must satisfy $M_T(\mu, \cancel{E}_T) < 30$ GeV (to
1412 suppress QCD multi-jet events).
- 1413 – Events which contain a jet overlapping with a muon $\Delta R(\mu, \text{jet}) < 0.5$ are vetoed
1414 to remove events from muons produced as part of a jet’s hadronisation process.
- 1415 – Events containing a second muon candidate which has failed id, but passed p_T
1416 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.

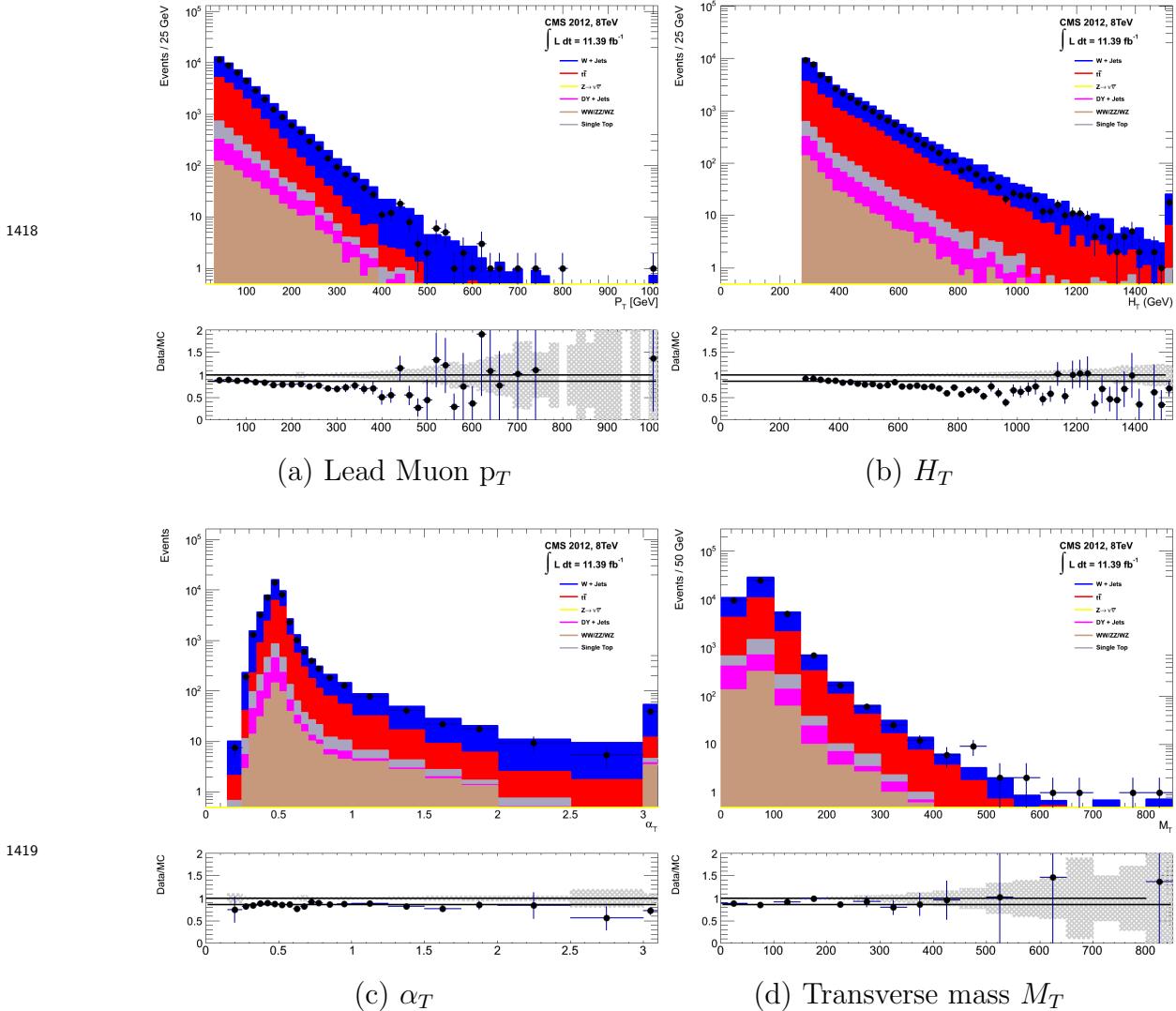


Figure 4.6.: Data/MC comparisons of key variables for the $\mu + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of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

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

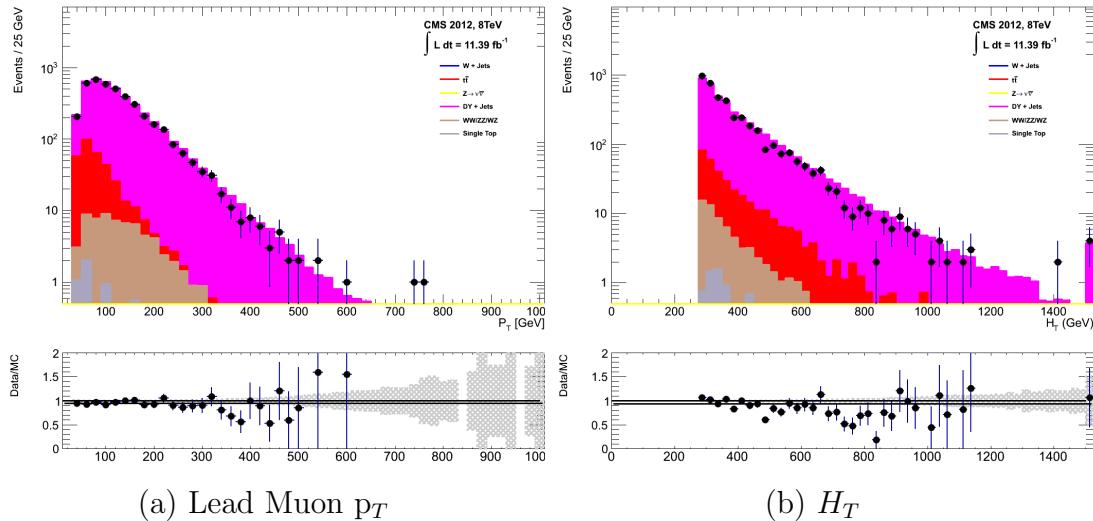
The $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background enters into the signal region from genuine E_T from the escaping neutrinos. This background is estimated using two control samples, the first of which is the $Z \rightarrow \mu\bar{\mu} + \text{jets}$ process, which posses identical kinematic 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.



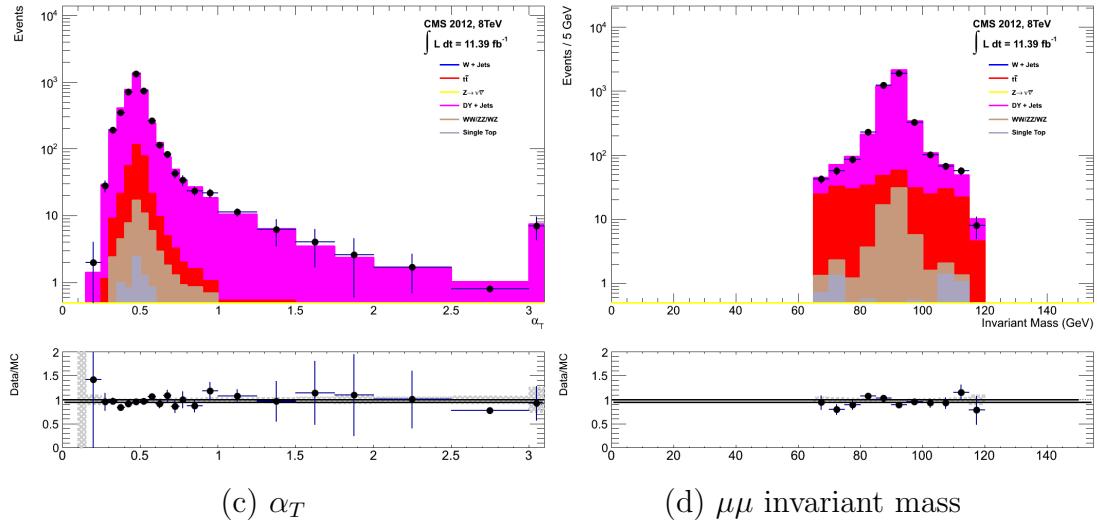


Figure 4.7.: Data/MC comparisons of key variables for the $\mu\mu + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number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

The γ + 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 [78][79]. The photon is ignored for the purpose of the calculation of event level variables, and identical selection cuts to the hadronic signal region are applied.

- Exactly one photon is selected, satisfying identification criteria as detailed in Table 4.3, with a minimum $p_T > 165$ GeV to satisfy trigger thresholds and $|\eta| < 1.45$ to ensure the photon remains in the barrel of the detector.
 - A selection criteria of $\Delta R(\gamma, jet) < 1.0$, between the photon and all jets is applied to ensure the acceptance of only well isolated $\gamma + \text{jets}$ 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).

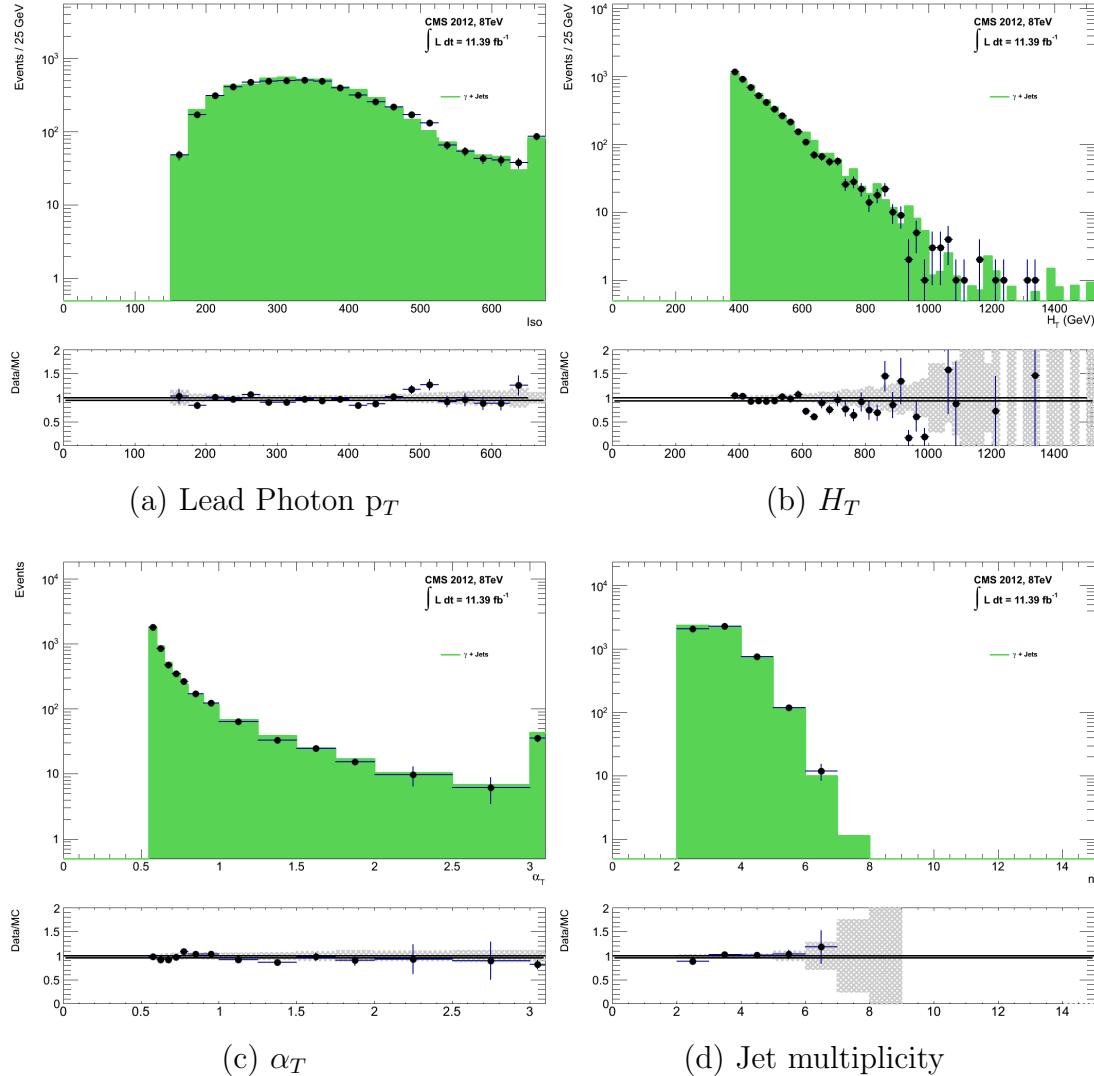


Figure 4.8.: Data/MC comparisons of key variables for the $\gamma + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement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

1466 this way, the acceptance of the two muon control samples can be significantly increased,
1467 which simultaneously improves their predictive power and further reduces the effect of
1468 any potential signal contamination.

1469 The modelling of the α_T variable is probed through a dedicated set of closure tests,
1470 described in Section (4.6), which demonstrate that the different α_T acceptances for the
1471 control and signal samples have no significant systematic bias on the prediction.

1472 4.2.4. Estimating the QCD multi-jet background

1473 A negligible background from QCD multi-jet events within the hadronic signal region
1474 is expected due to the selection requirement, and additional cleaning filters applied.
1475 However a conservative approach is still adopted and the likelihood model, see Section
1476 (4.7.2), is given the freedom to estimate any potential QCD multi-jet contamination.

1477 Any potential contamination can be identified through the variable R_{α_T} , defined as
1478 the ratio of events above and below the α_T threshold value used in the analysis. This is
1479 modelled by a H_T dependant falling exponential function which takes the form,

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

1480 where the parameters A and k_{QCD} are the normalisation and exponential decay
1481 constants respectively.

1482 For QCD event topologies this exponential behaviour is expected as a function of H_T
1483 for several reasons. The improvement of jet energy resolution at higher H_T due to higher
1484 p_T jets leads to a narrower peaked distribution, causing R_{α_T} to fall. Similarly at higher
1485 H_T values > 375 GeV, the jet multiplicity rises slowly with H_T . As shown in Figure 4.3,
1486 at higher jet multiplicities, the result of the combinatorics used in the determination of
1487 α_T , also lead to a narrower α_T distribution.

1488 The value of the decay constant k_{QCD} is constrained via measurements within data
1489 sidebands to the signal region. This is also done to validate the falling exponential
1490 assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet
1491 background and defined as regions where α_T is relaxed or that the R_{miss} cut is inverted.

1492 Figure 4.9 depicts the definition of these data sidebands used to constrain the value of
1493 k_{QCD} .

1494

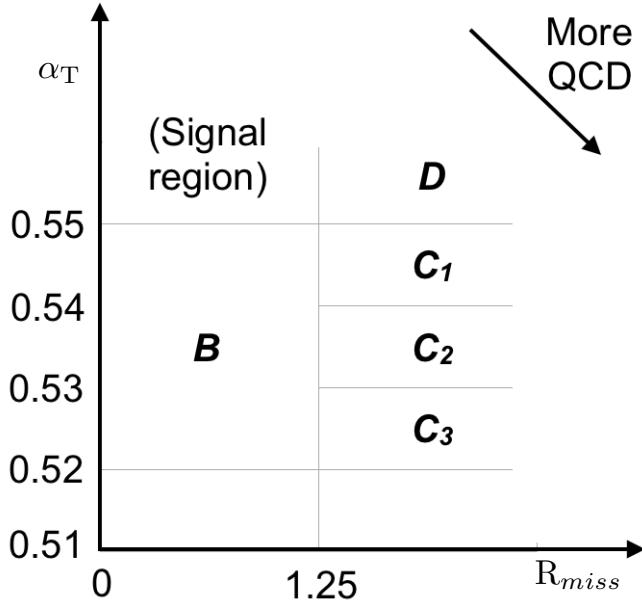


Figure 4.9.: QCD sideband regions, used for determination of k_{QCD} .

1495 The fits to determine the value of k_{QCD} are shown in Appendix (C.1), for which the best
1496 fit value obtained from sideband region B is determined to be $k_{QCD} = 2.96 \pm 0.64 \times 10^{-2}$
1497 GeV^{-1} .

1498 The best fit values of the remaining three C sideband regions are used to estimate
1499 the systematic uncertainty on the central value obtained from sideband region B. The
1500 variation of these measured values is used to determine the error on the determined
1501 central value, and is calculated to be $1.31 \pm 0.26 \times 10^{-2} \text{GeV}^{-1}$. This relative error of \sim
1502 20% gives an estimate of the systematic uncertainty of the measurement to be applied to
1503 k_{QCD} .

1504 Finally the same procedure is performed for sideband region D to establish that the
1505 value of k_{QCD} extracted from a lower α_T slice can be applied to the signal region $\alpha_T >$
1506 0.55. The likelihood fit is performed across all H_T bins within the QCD enriched region
1507 with no constraint applied to k_{QCD} . The resulting best fit value for k_{QCD} shows good
1508 agreement between that and the weighted mean determined from the three C sidebands
1509 regions. This demonstrates that the assumption of using the central value determined
1510 from sideband region B, to provide an unbiased estimator for k_{QCD} in the signal region
1511 ($\alpha_T > 0.55$) is valid.

1512 Table 4.6, summarises the best fit k_{QCD} values determined for each of the sideband
1513 regions to the signal region.

Sideband region	$k_{QCD} (\times 10^{-2} GeV^{-1})$	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.

1514 4.3. Trigger Strategy

1515 A cross trigger based on the quantities H_T and α_T , labelled is used with varying thresholds
1516 across H_T bins to record the events used in the hadronic signal region. The α_T legs of the
1517 HT_alphaT triggers used in the analysis are chosen to fully suppress QCD multi-jet events,
1518 whilst maintaining a sustainable trigger rate. To further maintain an acceptable rate for
1519 these analysis specific triggers, only calorimeter information is used in the reconstruction
1520 of the H_T sum, leading to the necessity for Calo jets to be used within the analysis.

1521 A single object prescaled H_T trigger is used to collect events for the hadronic control
1522 region described above in Section (4.2.4).

1523 The performance of the α_T and H_T triggers used to collect data for the signal and
1524 hadronic control region is measured with respect to a reference sample collected using the
1525 muon system. This allows measurement of both the Level 1 seed and higher level triggers
1526 simultaneously, as the reference sample is collected independent of any jet requirements.

1527 The selection for the trigger efficiency measurement is identical to that described in
1528 Section (4.2.2), with the requirement of exactly one well identified muon with $p_T > 30$
1529 GeV which is subsequently ignored.

1530 The efficiencies measured for the HT_alphaT triggers in bins individual H_T and α_T
1531 legs, is 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.

1532 Data for the control samples of the analysis, detailed in Section (4.2.3), are collected
 1533 using single object photon trigger for the $\gamma +$ jets sample, and a single object muon
 1534 trigger for both the $\mu +$ jets and $\mu\mu +$ jets control samples. The photon trigger is
 1535 measured to be full efficient for the threshold $p_T^{photon} > 150 GeV$, whilst the single muon
 1536 efficiency satisfying $p_T^{muon} > 30 GeV$ is measured to have an efficiency of $(88 \pm 2)\%$ that
 1537 is independent of H_T . In the case of the $\mu\mu +$ jets control sample, the efficiency is
 1538 measured to be $(95 \pm 2)\%$ for the lowest H_T bin, rising to $(98 \pm 2)\%$ for the highest H_T
 1539 bin.

1540 4.4. Measuring MC Normalisation Factors via H_T 1541 Sidebands

1542 The theoretical cross sections of different SM processes at Next to Next Leading Order
 1543 (NNLO) and the number of MC simulated events generated for that particular process,
 1544 is typically used to determine the appropriate normalisation for a MC sample. However
 1545 within the particular high- H_T and high- \cancel{E}_T corners of kinematic phase space probed
 1546 within this search, the theoretical cross sections for various processes are far less well
 1547 understood.

1548 To mitigate the problem of theoretical uncertainties and arbitrary choices of cross
 1549 sections, the normalisation of MC samples used in the analysis are determined through
 1550 the use data sidebands. The sidebands are used to calculate sample specific correct
 1551 factors (k-factors) that are appropriate for the H_T - \cancel{E}_T phase space covered by this analysis.

1552 They are defined within the $\mu +$ jets and $\mu\mu +$ jets control sample, by the region $200 <$
 1553 $H_T < 275$, using the same jet p_T thresholds as the adjacent first analysis bin. Individual

1554 1555 1556 1557 1558 1559 **EWK** processes are isolated within each of these control samples via requirements on jet multiplicity and the requirement on b-tags, summarised in Table 4.8. The purity of the samples are typically $> 90\%$ with any residual contamination corrected for. The resultant k-factor for each process is determined by then taking ratio of the data yield over the MC expectation in the sideband. Subsequently these k-factors are then applied 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.

1560 **4.5. Determining MC Simulation Yields with 1561 Higher Statistical Precision**

1562 Reconstructing events from **EWK** processes with many b-tagged jets (≥ 3), n_b^{reco} , is largely 1563 driven by the mis-tagging of light jets within the event. This is clear when considering 1564 the main **EWK** backgrounds in the analysis, such as $t\bar{t} + \text{jets}$ events, which typically 1565 contain two b-flavoured jets from the decay of the top quarks, whilst W + jets and 1566 Z $\rightarrow \mu\mu + \text{jets}$ events will typically contain no b-flavoured jets.

1567 When the expectation for the number of n_b^{reco} is taken directly from simulation, the 1568 statistical uncertainty at large b-tag multiplicities becomes relatively large. In order to 1569 reduce this uncertainty one approach is to use the information encoded throughout all 1570 events in the simulation sample, to measure each of the four ingredients:

- 1571 1. the b-tagging efficiency in the event selection,
- 1572 2. the charm-tagging efficiency in the event selection
- 1573 3. the mis-tagging rate in the event selection,
- 1574 4. the underlying flavour distribution of the jets in the events,

that determine the n_b^{reco} distribution of the process being measured. This method allows the determination of higher b-tag multiplicities to a higher degree of accuracy reducing the statical uncertainties of the MC which enter into the TF's. For the discussion that follows, these predictions are determined on average (i.e not on an event-by-event basis), and is known as the formula method.

4.5.1. The formula method

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

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

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

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

Using this information the expected number of jets which have been b-tagged can be analytically calculated using the formula :

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

1600 with $N(n_b)$ representing the event yield where n_b jets have been b-tagged, n_b^{tag} ,
 1601 n_c^{tag} and n_q^{tag} represent the number of times that a particular jet flavour results in
 1602 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
 1603 binomial probabilities for that to happen.

1604 This approach ultimately results in a more precise n_b^{reco} distribution prediction as
 1605 information from throughout the entire MC sample is used to estimate the high n_b^{reco}
 1606 bins.

1607 4.5.2. Establishing proof of principle

1608 In order to validate the procedure, the predictions obtained from the formula method
 1609 summarised in Eq (4.11), are compared directly to those obtained directly from simulation.
 1610 These results for the $\mu +$ jets control sample are summarised in Table 4.9, for the 0,1,2
 1611 and 3 n_b^{reco} bins.

Process	Selection	Observation	MC expectation	k-factor
---------	-----------	-------------	----------------	----------

Table 4.9.: place holder

1612 4.5.3. Correcting measured efficiencies in simulation to data

1613 As detailed in Section (3.3.2), it is necessary for certain p_T and η dependant corrections,
 1614 to be applied to both the b-tagging efficiency and mis-tagging rates in order correct
 1615 the efficiencies from simulation to the distributions seen in data. These corrections are
 1616 factored in.

1617 Show plot of before and after correction to btag/mistag rate.

1618 These corrections come with uncertainties..

1619 show plot of effect of scaling correction factor up and down. 2

1620 4.6. Systematic Uncertainties on Transfer Factors

1621 Since the TF's used to establish the background prediction are obtained from simulation,
1622 an appropriate systematic uncertainty is assigned to each factor to account for theoretical
1623 uncertainties [80] and limitations in the simulation modelling of event kinematics and
1624 instrumental effects.

1625 The magnitudes of these systematic uncertainties are established through a set of data
1626 driven method, in which the three independent control samples of the analysis ($\mu + \text{jets}$,
1627 $\mu\mu + \text{jets}$, $\gamma + \text{jets}$) are used to in a series of closure tests. The yields from one of these
1628 control samples, along with the corresponding TF obtained from simulation, are used to
1629 predict the yields in another control sample, using the same method of establishing a
1630 background prediction for the signal region as described in Section (4.2.3).

1631 The level of agreement between the predicted and observed yields is expressed as the
1632 ratio

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

1633 while considering only the statistical uncertainties on N_{pred} , the prediction, and N_{obs} ,
1634 the observation. No systematic uncertainty is assigned to the prediction, and resultantly
1635 the level of closure is defined by the statistical significance of a deviation from the ratio
1636 from zero.

1637 This ratio is measured for each H_T bin in the analysis, allowing these closure tests to
1638 be sensitive to both the presence of any significant biases or any possible H_T dependence
1639 on the level of closure.

1640 Eight sets of closure tests are defined between the three data control samples, con-
1641 ducted independently between the two jet multiplicity ($2 \leq n_{jets} \leq 3$, $n_{jet} \geq 4$) bins.
1642 Each of these tests are specifically chosen to probe each of the different key ingredients
1643 of the simulation modelling that can affect the background prediction.

1644 Each of the different modelling components and the relevant closure tests are described
1645 below :

1646 **α_T modelling**

1647 The modelling of the α_T distribution in genuine Z_T events is probed with the $\mu +$
1648 jets control sample. This test is important to verify the approach of remove the α_T
1649 > 0.55 requirement from the $\mu +$ jets and $\mu\mu +$ jets samples to increase the precision
1650 of the background prediction. The test uses the $\mu +$ jets sample without an α_T cut
1651 to make a prediction into the $\mu +$ jets sample defined with the requirement $\alpha_T >$
1652 0.55 .

1653 **Background admixture**

1654 The sensitivity of the translation factors to the relative admixture of events from
1655 $W +$ jets and $t\bar{t}$ processes is probed by two closure tests. These tests represent
1656 an extremely conservative approach as the admixture of the background remains
1657 similar between the $\mu +$ jets sample and the signal region, contrary to the defined
1658 closure tests which make predictions between two very different admixtures of $W +$
1659 jets and $t\bar{t}$ events.

1660 Within the $\mu +$ jets sample, a W boson enriched sub-sample ($n_b = 0$) is used to
1661 predict yields in a $t\bar{t}$ enriched sub-sample ($n_b = 1$). Similarly the
1662 $t\bar{t}$ enriched sub-sample ($n_b = 1$) is also used to predict yields for a further enriched
1663 $t\bar{t}$ sub-sample ($n_b = 2$).

1664 Similarly a further closure test probes the relative contribution of $Z +$ jets to $W +$ jets
1665 and $t\bar{t}$ events, through the use of the $\mu +$ jets sample to predict yields for the $\mu\mu +$
1666 jets control sample. This closure test, also at some level probes the muon trigger
1667 and reconstruction efficiencies, given that exactly one and two muons are required
1668 by the different selections.

1669 **Consistency between control samples**

1670 An important consistency check between the $\mu\mu +$ jets jets and $\gamma +$ jets, which are
1671 both used in the prediction of the $Z \rightarrow \nu\bar{\nu}$ in the signal region, is measured by using
1672 the $\gamma +$ jets sample to predict yields for the $\mu\mu +$ jets control sample.

1673 **Modelling of jet multiplicity**

1674 The simulation modelling of the jet multiplicity within each control sample is
1675 important due to the exclusive jet multiplicity binning within the analysis. This is
1676 probed via the use of each of the three control samples to independently predict
1677 from the lower jet multiplicity category $2 \leq n_{jet} \leq 3$, to the high jet category ≥ 4 .

1678 For the case of the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples this test this is also a
 1679 further probe of the admixture between $W + \text{jets}/Z + \text{jets}$ and $t\bar{t}$.

1680 To test for the assumption that no H_T dependences exist within the background
 1681 predictions of the analysis, the first five closure tests defined above are taken, with zeroeth
 1682 and first order polynomial fits are applied to each. This is summarised in Table 4.10 and
 1683 Table 4.11 which show the results for both the $2 \leq n_{jet} \leq 3$ and ≥ 4 jet multiplicity bins
 1684 respectively.

Closure test	Symbol	Constant fit		Linear fit	
		Best fit value	p-value	Slope (10^{-4})	p-value
$\alpha_T < 0.55 \rightarrow \alpha_T > 0.55 (\mu + \text{jets})$	Circle	-0.06 ± 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 4.10.: 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 4.11.: 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 $n_{jet} \geq q$ 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.

1685 Table 4.12 shows the same fits applied to the three closure tests that probe the
 1686 modelling between the different n_{jet} bins. The best fit value and its uncertainty is listed
 1687 for each set of closure tests in all three tables, along with the p-value of the constant and
 1688 linear fits applied.

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 4.12.: 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 $2 \leq n_{\text{jet}} \leq 3$ 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.

1689 The best fit value for the constant parameter is indicative of the level of closure,
 1690 averaged across the full range of H_T bins in the analysis, and the p-value an indicator of
 1691 any significant dependence on H_T within the closure tests. The best fit values of all the
 1692 tests are either statistically compatible with zero bias (i.e, less than 2σ from zero) or at
 1693 the level of 10% or less, with the exception of one closure test discussed below.

1694 Within Table 4.12, there exists one test that does not satisfy the above statement,
 1695 which is the $2 \leq n_{\text{jet}} \leq 3 \rightarrow n_{\text{jet}} \geq 4$ test using the $\mu + \text{jets}$ control sample. The low
 1696 p-value can be largely attributed to an outlier in the $675 < H_T < 775$ GeV bin, rather
 1697 than any significant trend in H_T . Removing this single outlier from the constant fit
 1698 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
 1699 0.42. These modified fit results are included within Table 4.12 .

1700 In addition the best fit values for the slope terms of the linear fits in all three tables
 1701 are of the order 10^{-4} , which corresponds to a percent level change per 100 GeV. However
 1702 in all cases, the best fit values are fully compatible with zero (within 1σ) once again with
 1703 the exception detailed above, indicating that the level of closure is H_T independent.

1704 4.6.1. Determining systematic uncertainties from closure tests

1705 Once it has been established that no significant bias or trend has been exist within
 1706 the closure tests, systematic uncertainties are determined. The statistical precision
 1707 of the closure tests is considered a suitable benchmark for determining the systematic
 1708 uncertainties that are assigned to the TF's, which are propagated through to the likelihood
 1709 fit.

1710 The systematic uncertainty band is split into five separate regions of H_T :

1711 1. $275 < H_T < 325\text{GeV}$

1712 2. $325 < H_T < 375\text{GeV}$

1713 3. $375 < H_T < 575\text{GeV}$

1714 4. $575 < H_T < 775\text{GeV}$

1715 5. $H_T > 775\text{ GeV}$

1716 Within each region the square root of the sample variance, σ^2 , is taken over the eight
1717 closure tests to determine the systematic uncertainties to be applied within that region.

1718 Using this procedure the systematic uncertainties for each region are calculated and
1719 are shown in Table 4.13, with the systematic uncertainty to be used in the likelihood
1720 model 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 4.13.: 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.

1721 Figure 4.10 shows the sets of closure tests overlaid on top of grey bands that represent
1722 the H_T dependent systematic uncertainties. These systematic uncertainties are assumed
1723 to fully uncorrelated between the different n_b multiplicity categories and across the five
1724 H_T regions. This can be considered a more conservative approach given that some
1725 correlations between adjacent H_T bins could be expected due to comparable kinematics.

1726 As already referenced. These closure tests represent a conservative estimate of the
1727 systematic uncertainty in making a background prediction for the signal region. This
1728 is due to significant differences in the background composition and event kinematics
1729 between the two sub-samples used in the closure tests. This is contrary to the signal
1730 region prediction where the two sub-samples are both have a comparable background

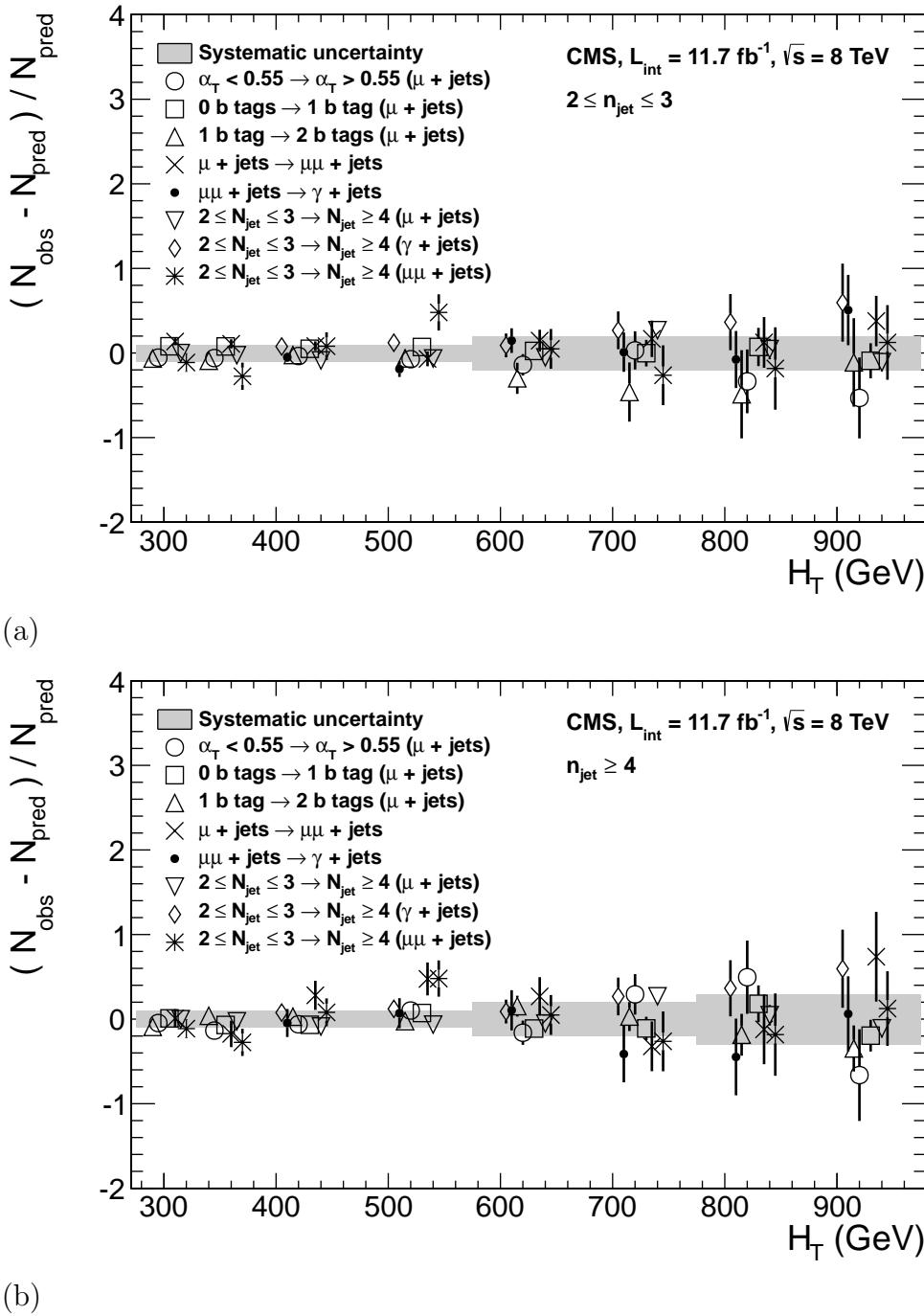


Figure 4.10.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins: (a) $2 \leq n_{\text{jet}} \leq 3$ and (b) $n_{\text{jet}} \geq 4$.

1731 admixture and similar kinematics owing to the fact that the predictions are always made
 1732 using the same $(n_{\text{jet}}, n_b, H_T)$ bin.

1733 This point is emphasised when we examine the sensitivity of the **TF**'s to a change
1734 in the admixture of $W + \text{jets}$ and $t\bar{t}$ with the control and signal samples. This is
1735 accomplished by varying the cross sections of the $W + \text{jets}$ and $t\bar{t}$ by +20% and -20%,
1736 respectively. Figures C.2 and C.3 within Appendix C, show the effect upon the closure
1737 tests for both jet multiplicity categories. Given these variations in cross sections, the
1738 level of closure is found to be significantly worse, with biases as large as $\sim 30\%$, most
1739 apparent in the lowest H_T bins. However the **TF**'s used to extrapolate from control to
1740 signal are seen to change only at the percent level by this large change in cross section,
1741 shown in Table C.1.

1742 Given the robust behaviour of the translation factors with respect to large (and
1743 opposite) variations in the $W + \text{jets}$ and $t\bar{t}$ cross sections, one can assume with confidence
1744 that any bias in the translation factors is adequately (and conservatively) covered by the
1745 systematic uncertainties used in the analysis.

1746 4.7. Statistical Framework

1747 For a given category of events satisfying requirements on both n_{jet} and n_b^{reco} , a likelihood
1748 model of the observations in multiple data samples is used to gauge agreement between
1749 the observed yields in the hadronic signal region, and the predicted yields obtained from
1750 the control samples. In addition to checking whether the predictions are compatible with
1751 a **SM** only hypothesis, the likelihood model is also used to test for the presence of a
1752 variety of signal models.

1753 4.7.1. Hadronic sample

1754 Let N be the number of bins on H_T , with n^i the number of events observed satisfying
1755 all selection requirements in each H_T bin i. The likelihood of the observations can then
1756 be written :

$$L_{had} = \prod_i \text{Pois}(n^i | b^i + s^i), \quad (4.13)$$

1757 where b^i represents the expected **SM** background

$$b^i = EWK_i + QCD_i, \quad (4.14)$$

1758 and s^i the expected number of signal events from the different **SMS** models interpreted.
1759 Pois refers to the Poission distribution of these values and is defined as :

$$\text{Pois}(\chi|\lambda) = \frac{\lambda^\chi \exp^{-\lambda}}{k!}. \quad (4.15)$$

1760 **4.7.2. H_T evolution model**

1761 The hypothesis, that for a process the α_T ratio falls exponentially see Section (4.2.4)
1762 in H_T is defined by Equation (4.10), where k_{QCD} is constrained by measurements in a
1763 signal sideband region.

1764 The expected QCD background, QCD^i , within a bin i is then modelled as,

$$QCD^i = m^i A_{QCD} e^{-k_{QCD} \langle H_T \rangle}, \quad (4.16)$$

1765 where m_i represent the number of events observed with $\alpha_T \leq 0.55$ in each H_T bin i ,
1766 and $\langle H_T \rangle$ represents the mean H_T of each bin. Expressed as functions of just the zeroth
1767 bin, QCD^0 , and k_{QCD} , the QCD expectation is given by

$$QCD^i = QCD^0 \left(\frac{m^i}{m^0} \right) e^{-k_{QCD} (\langle H_T \rangle^i - \langle H_T \rangle^0)}. \quad (4.17)$$

1768 **4.7.3. **EWK** control samples**

1769 The **EWK** background estimation within each bin, i , is broken into two components, the
1770 expected yield from $Z \rightarrow \nu\bar{\nu}$ and $t\bar{t}$ -W (plus other residual backgrounds) events. This is
1771 written as, Z_{inv}^i and $t\bar{t}W^i$, and it follows that

$$EWK^i = Z_{inv}^i + t\bar{t}W^i. \quad (4.18)$$

1772 This can be further expressed as

$$Z_{inv}^i \equiv f_{Zinv}^i \times EWK^i, \quad (4.19)$$

$$t\bar{t}W^i \equiv (1 - f_{Zinv}^i) \times EWK^i, \quad (4.20)$$

1773 where f_{Zinv}^i represents the expected yield from $Z \rightarrow \nu\bar{\nu}$ in bin i divided by the expected
1774 **EWK** background EWK^i . This fraction is modelled as a linear component

$$f_{Zinv}^i = f_{Zinv}^0 + \frac{\langle H_T \rangle^i - \langle H_T \rangle^0}{\langle H_T \rangle^{N-1} - \langle H_T \rangle^0} (f_{Zinv}^{N-1} - f_{Zinv}^i), \quad (4.21)$$

1775 where N again represents the number of H_T bins, and f_{Zinv}^i and f_{Zinv}^{N-1} are float
1776 parameters whose final values are limited between zero and one.

1777 Within each H_T bin there are three background measurements for the different control
1778 samples, n_γ^i , n_μ^i and $n_{\mu\mu}^i$, representing the event yields from the $\gamma +$ jets, $\mu +$ jets and
1779 $\mu\mu +$ jets control samples respectively. Each of these have a corresponding yield in
1780 simulation, MC_γ^i , MC_μ^i and $MC_{\mu\mu}^i$. Within the hadronic signal region there are also
1781 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
1782 used to define

$$r_\gamma^i = \frac{MC_\gamma^i}{MC_{Zinv}^i}; \quad r_{\mu\mu}^i = \frac{MC_{\mu\mu}^i}{MC_{Zinv}^i}; \quad r_\mu^i = \frac{MC_\mu^i}{MC_{t\bar{t}+W}^i}, \quad (4.22)$$

1783 where r_p^i represents the inverse of the **TF**'s used to extrapolate the yield of each background
1784 process.

1785 The likelihoods regarding the three measured yields n_γ^i , $n_{\mu\mu}^i$, n_μ^i can then be fully
1786 expressed as

$$L_\gamma = \prod_i \text{Pois}(n_\gamma^i | \rho_{\gamma Z}^j \cdot r_\gamma^i \cdot Z_{inv}^i), \quad (4.23)$$

$$L_{\mu\mu} = \prod_i \text{Pois}(n_{\mu\mu}^i | \rho_{\mu\mu Z}^j \cdot r_{\mu\mu}^i \cdot Z_{inv}^i), \quad (4.24)$$

$$L_\mu = \prod_i \text{Pois}(n_\mu^i | \rho_{\mu Y}^j \cdot r_\mu^i \cdot Y^i + s_\mu^i), \quad (4.25)$$

(4.26)

which contain an additional term s_μ^i , which represents the signal contamination in the $\mu +$ jets sample. The parameters $\rho_{\gamma Z}^j$, $\rho_{\mu\mu}^j$ and ρ_μ^j represent “correction factors” that accommodate the systematic uncertainties associated with the control sample based background constraints.

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

The measurements in each of the control samples and the hadronic signal region, along with the ratios r_γ^i , $r_{\mu\mu}^i$, and r_μ^i , are all considered simultaneously through the relationships defined by Equations (4.14),(4.19) and (4.20).

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

$$L_{EWK \ syst} = \prod_j \text{Logn}(1.0 | \rho_{\mu W}^j, \sigma_{\mu W}^j) \times \text{Logn}(1.0 | \rho_{\mu\mu Z}^j, \sigma_{\mu\mu Z}^j) \times \text{Logn}(1.0 | \rho_{\gamma Z}^j, \sigma_{\gamma Z}^j), \quad (4.27)$$

where $\sigma_{\gamma Z}^j$, $\sigma_{\mu\mu Z}^j$ and $\sigma_{\mu W}^j$ represent the relative systematic uncertainties for the control sample constraints and Logn is the log-normal distribution [81],

$$\text{Logn}(x | \mu, \sigma_{rel}) = \frac{1}{x\sqrt{2\pi}\ln k} \exp\left(\frac{\ln^2(\frac{x}{\mu})}{2\ln^2 k}\right); \quad k = 1 + \sigma_{rel}. \quad (4.28)$$

1801 Five parameters per control sample are used to span the eight H_T bins, with just one
 1802 used for the three H_T bins in the $n_b^{reco} \geq 4$ category. These parameters span the same
 1803 H_T ranges described in Section (4.6) and is shown in Table 4.14.

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

Table 4.14.: The systematic parameters used in H_T bins. Left: categories with eight bins; right: category with three bins.

1804 Alternatively, in the higher n_b^{reco} categories ($n_b^{reco} = 3, n_b^{reco} \geq 4$), the single muon
 1805 sample is used to constrain the total EWK background. Therefore the likelihood function
 1806 is greatly simplified and is represented by

$$L'_\mu = \prod_i \text{Pois}(n_\mu^i | \rho_{\mu Y}^j \cdot r'_\mu \cdot EWK^i + s_\mu^i), \quad (4.29)$$

(4.30)

1807 where,

$$r'_\mu = \frac{MC_\mu^i}{MC_{tot}^i}. \quad (4.31)$$

1808 4.7.4. Contributions from signal

1809 The cross section for each model is represented by x and l represents the total recorded
 1810 luminosity considered by the analysis in the signal region. Let ϵ_{had}^i and ϵ_μ^i represent the
 1811 analysis selection efficiency for that particular signal model in H_T bin i of the hadronic
 1812 and $\mu +$ jets control sample respectively. Letting δ represent the relative uncertainty on
 1813 the signal yield, assumed to be fully correlated across all bins, and ρ_{sig} the “correction
 1814 factor” to the signal yield which accommodates this uncertainty. f represents an unknown
 1815 multiplicative factor on the signal cross section, for which an allowed interval is computed.

1816 The expected signal yield s^i is thus given by

$$s^i \equiv f \rho_{sig} x l \epsilon_{had}^i \quad (4.32)$$

1817 and signal contamination with the $\mu +$ jets control sample by

$$s_\mu^i \equiv f \rho_{sig} x l \epsilon_\mu^i \quad (4.33)$$

1818 The systematic uncertainty on the signal is additionally included by the term

$$L_{sig} = \text{Logn}(1.0 | \rho_{sig}, \delta). \quad (4.34)$$

1819 A discussion of the **SMS** signal models through which the analysis is interpreted,
1820 including a discussion on the signal efficiencies and systematics can be found in the
1821 following Chapter.

1822 4.7.5. Total likelihood

1823 The total likelihood function for a given signal bin $k(n_b^{reco}, n_{jet})$ is given by the product
1824 of the likelihood functions introduced within the previous sections:

$$L^k = L_{had}^k \times L_\mu^k \times L_\gamma^k \times L_{\mu\mu}^k \times L_{EWKsyst}^k \times L_{QCD}^k. \quad (4.35)$$

1825 In categories containing eight H_T bins and utilising the three control samples ($\mu +$
1826 jets, $\mu\mu +$ jets, $\gamma +$ jets), there are 25 nuisance parameters, whilst when just one control
1827 sample is used to estimate the **EWK** background, there are 15 nuisance parameters.
1828 Where three H_T bins are used (the highest n_b^{reco} category), there are 6 nuisance parameters.
1829 This information is summarised within Table 4.15.

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 4.15.: 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.

When considering SUSY signal models within the likelihood, an additional parameter is introduced, ρ_{sig} . When multiple categories are fit simultaneously the total likelihood is then represented by

$$L = L_{sig} \times \prod_k L_{had}^k \times L_\mu^k \times L_\gamma^k \times L_{\mu\mu}^k \times L_{EWKsyst}^k \times L_{QCD}^k. \quad (4.36)$$

Chapter 5.

¹⁸³³ Results and Interpretation

¹⁸³⁴ Results at 12fb 8TeV

¹⁸³⁵ 5.1. Statistical Interpretation

¹⁸³⁶ Likelihood stuff

¹⁸³⁷ 5.2. Interpretation in Simplified Signal Models

¹⁸³⁸ Result interpretation

Chapter 6.

¹⁸³⁹ Searches For Natural SUSY With ¹⁸⁴⁰ B-tag Templates.

¹⁸⁴¹ Within this chapter a complimentary technique is discussed as a means to predict the
¹⁸⁴² distribution of three and four reconstructed b-quark jets in an event. The recent discovery
¹⁸⁴³ of the Higgs boson has made third-generation “Natural SUSY” models attractive, given
¹⁸⁴⁴ that light top and bottom squarks are a candidate to stabilise divergent loop corrections
¹⁸⁴⁵ to the Higgs boson mass.

¹⁸⁴⁶ Using the α_T search as a base, a simple templated fit is employed to estimate the
¹⁸⁴⁷ SM background in higher b-tag multiplicities (3-4) from a region of a low number of
¹⁸⁴⁸ reconstructed b-jets (0-2). As a proof-of-concept, the procedure after being shown to
¹⁸⁴⁹ close in simulation, is applied to the SM enriched $\mu + \text{jets}$ control sample of the α_T
¹⁸⁵⁰ all-hadronic search detailed in Chapter 4. To highlight the relative insensitivity of the
¹⁸⁵¹ choice of the b-tagging algorithm working points in the effectiveness of the procedure,
¹⁸⁵² results are presented using the CSV tagger (introduced in Section (3.3.2)) for the “Loose”,
¹⁸⁵³ “Medium” and “Tight” working points.

¹⁸⁵⁴ 6.1. Concept

¹⁸⁵⁵ The dominant SM backgrounds most SUSY searches are typically $t\bar{t} + \text{jets}$, $W + \text{jets}$ and
¹⁸⁵⁶ $Z \rightarrow \nu\bar{\nu} + \text{jets}$. These process are characterised by typically having zero or two underlying
¹⁸⁵⁷ b-quarks per event. The first step in this approach is to categorise two templates to be
¹⁸⁵⁸ fitted to the low n_b^{reco} multiplicity in terms of these underlying b-quark event topologies :

1859 Z0 - W + jets, Z $\rightarrow \nu\bar{\nu}$ + jets, DY + jets

1860 Z2 - $t\bar{t}$, single top

1861 where Z0 and Z2 represent processes which have an underlying b-quark content of
1862 zero or two respectively.

1863 Both these templates can be generated through the application of the relevant event
1864 selection and taking the underlying n_b^{reco} distribution directly from simulation. However
1865 as discussed within Section (4.5), there are large uncertainties for high n_b^{reco} multiplicities
1866 due to limited MC statistics. This is particularly prominent for the Z0 templates, where
1867 the number of reconstructed b-tags is driven primarily by the light-quark mis-tagging
1868 rate. Therefore to improve the statistical precision of the predictions the formula method,
1869 introduced in Section (4.5.1) is used.

1870 The generation of these templates is then dependant upon the jet-flavour content and
1871 b-tagging rate within the phase space of interest, with the tagging probabilities of a jet
1872 being a function of the jet p_T , the pseudo-rapidity $|\eta|$, and the jet-flavour. This can be
1873 observed in Figure 6.1, where the b-tagging / c-quark mis-tagging / light mis-tagging
1874 efficiency for the three working points of the CSV tagger is shown as a function of jet p_T .

1875 Before the templates are generated, the relevant jet p_T and η corrections are applied
1876 to correct simulation to data, as specified in Section (4.5.3), to then determine the average
1877 tagging rates per analysis bin.

1878 These two templates are then fit to data in the low n_b^{reco} region (0-2). The fit result
1879 is used, along with the knowledge of the template shapes, to extrapolate an estimate to
1880 the high n_b^{reco} signal region (3,4), which is then compared to what is observed in data.

1881 This method can, in principle, be applied to any analysis where the signal hypothesis
1882 has a larger underlying b-quark spectra than the SM backgrounds, as it solely relies on
1883 fitting to the shape of the n_b^{reco} distribution.

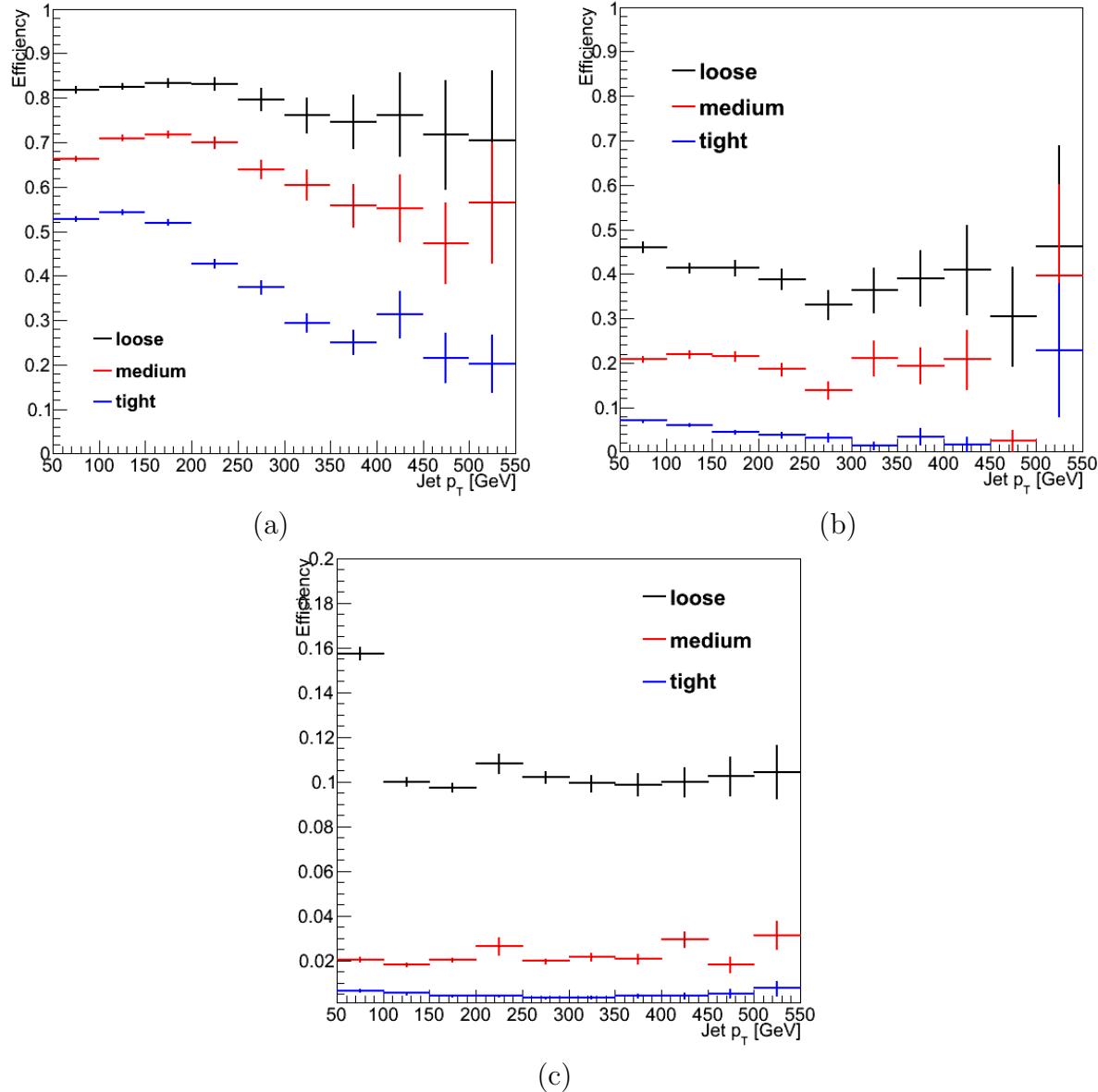


Figure 6.1.: The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging rate (c) as measured in simulation after the α_T analysis, $\mu + \text{jets}$ control sample selection in the region $H_T > 375$.

6.2. Application to the α_T Search

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

The selection for the $\mu +$ jets control samples defined in Section (4.2.3) is used to demonstrate the template fitting procedure both conceptually in simulation, and also when applied in data. This is chosen, as such a selection is dominated by events stemming from the SM processes with little or no signal contamination from potential new physics.. Neither are contributions from rate SM processes with a higher underlying b-quark content (e.g. $t\bar{t}b\bar{b}$) expected. For these reasons, there is a degree of confidence that the procedure should close when applied to this phase space.

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

- 275-325 GeV
- 325-375 GeV
- > 375 GeV

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

6.2.1. Proof of principle in simulation

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

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

The fits are performed independently within each of the defined analysis bins to reduce the dependence of the shapes of these distributions on simulation. The half of the simulation sample for which the templates are fitted too, are taken directly from simulation, extending this procedure to also be a validation of the formula method to

1919 accurately estimate the n_b^{reco} distribution. Additionally as this test is performed in
 1920 simulation, the relevant corrections of the b-tagging rates between data and simulation
 1921 are *not* applied.

1922 Within Figure 6.2, the results of this fitting procedure is shown for each CSV working
 1923 point. Results are presented for the $n_{jet} \geq 5$ category, using the $\mu +$ jets control sample
 1924 selection in the inclusive $H_T > 375$ GeV analysis bin. The grey bands represent the
 1925 statistical uncertainty on the template shapes. Additional fits are shown for other n_{jet}
 1926 category within Appendix D.1.

1927 Furthermore the extrapolated fit predictions within the high n_b^{reco} signal region, are
 1928 summarised for all H_T bins and working points in Table 6.1.

H_T	275-325	325-375	>375
Loose working point			
Simulation $n_b = 3$	344.0 ± 6.8	158.8 ± 4.5	324.9 ± 6.5
Template $n_b = 3$	347.5 ± 11.6	162.6 ± 4.7	322.9 ± 6.9
Simulation $n_b = 4$	29.8 ± 1.9	11.1 ± 1.1	40.2 ± 2.4
Template $n_b = 4$	32.6 ± 2.0	13.0 ± 1.0	37.0 ± 1.8
Medium working point			
Simulation $n_b = 3$	58.2 ± 2.87	33.3 ± 2.1	72.1 ± 3.1
Template $n_b = 3$	60.1 ± 1.9	32.1 ± 1.5	70.8 ± 2.3
Simulation $n_b = 4$	1.0 ± 0.4	0.3 ± 0.2	1.5 ± 0.4
Template $n_b = 4$	1.2 ± 0.1	0.4 ± 0.1	2.2 ± 0.2
Tight working point			
Simulation $n_b = 3$	58.2 ± 2.87	33.3 ± 2.1	72.1 ± 3.1
Template $n_b = 3$	60.1 ± 1.9	32.1 ± 1.5	70.8 ± 2.3
Simulation $n_b = 4$	1.0 ± 0.4	0.3 ± 0.2	1.5 ± 0.4
Template $n_b = 4$	1.2 ± 0.1	0.4 ± 0.1	2.2 ± 0.2

Table 6.1.: 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.

1929 The pull distributions for all the fits performed are compatible with a mean of zero
 1930 and standard distributions, see Appendix D.2.

1931 The good overall agreement summarised in the table validates both the formula method
 1932 used to generate the templates as well as the fitting method itself. The application of
 1933 this method to the same selection in data is used to demonstrate necessary control over
 1934 the efficiency and mis-tagging rates.

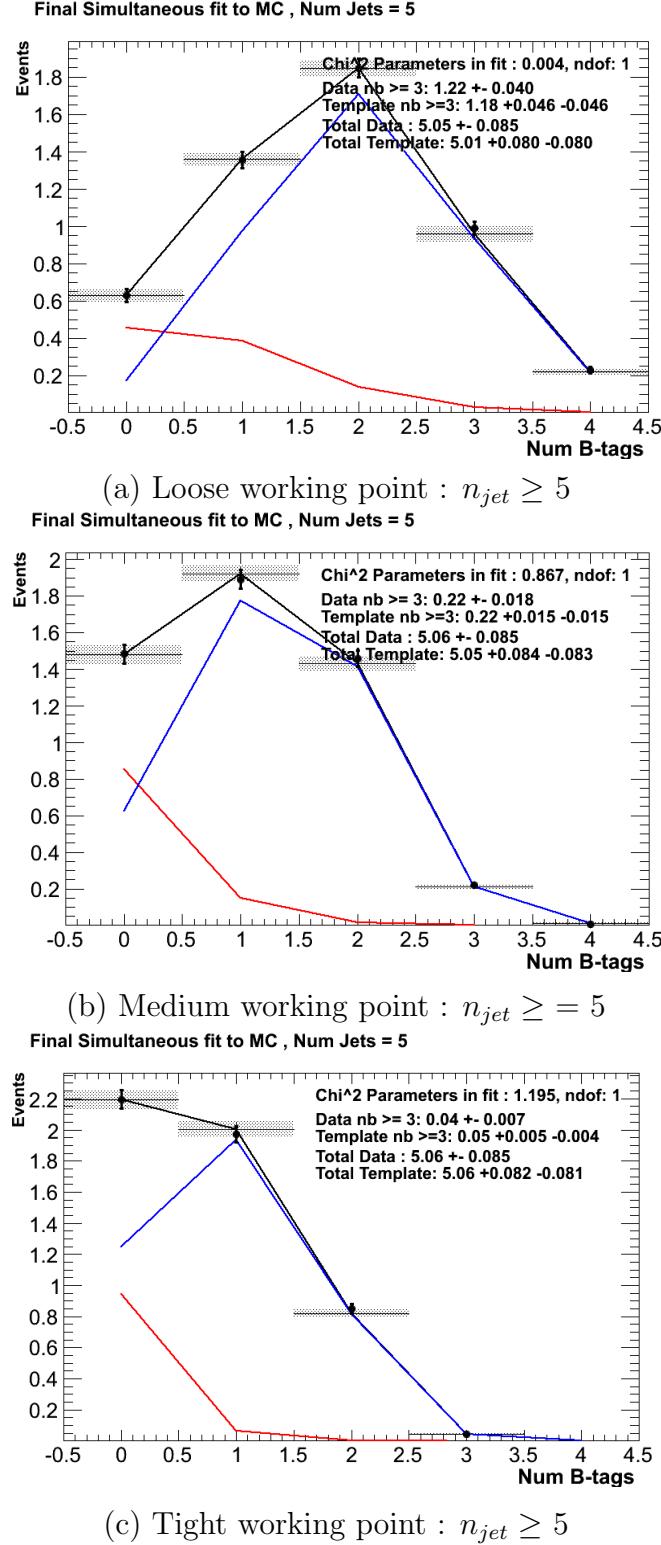


Figure 6.2.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} \geq 5$ category. The red template represents $Z = 0$, while the blue template represents $Z = 2$. Grey bands represent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

 1935 **6.2.2. Results in a data control sample**

1936 The method above is now applied to the 2012 8 TeV dataset in the $\mu +$ jets control
 1937 sample, to establish the validity of this method in data. The relevant data to simulation
 1938 scale factors are applied to get corrected values of the efficiency and mis-tagging rates
 1939 measured in data [64] [82].

1940 Figure 6.3 show the the results of the templates derived from simulation to each of
 1941 the three defined H_T bins, in the $n_{jet} \geq 5$ category for the medium working point CSV
 1942 tagger (the same working point used within the α_T analysis). Grey bands represent the
 1943 statistical uncertainty of the fit combined in quadrature with the systematic uncertainties
 1944 of varying the data to simulation scale factors up and down by their measured systematic
 1945 uncertainties. Additional fit results for the other working points are found in Appendix
 1946 D.3

1947 The numerical results and extrapolation to the $n_b^{reco} = 3, 4$ bins for all H_T and working
 1948 points is shown in Table 6.2.

H_T	275-325	325-375	>375
Loose working point			
Data $n_b = 3$	717	338	618
Template $n_b = 3$	782.6 ± 16.8	340.6 ± 10.2	601.9 ± 14.2
Data $n_b = 4$	68	39	68
Template $n_b = 4$	75.0 ± 2.7	27.6 ± 1.3	71.6 ± 2.6
Medium working point			
Data $n_b = 3$	124	73	137
Template $n_b = 3$	124.3 ± 2.3	62.0 ± 1.7	121.9 ± 2.5
Data $n_b = 4$	1	1	3
Template $n_b = 4$	2.6 ± 0.1	1.3 ± 0.1	4.0 ± 0.1
Tight working point			
Data $n_b = 3$	21	13	23
Template $n_b = 3$	26.7 ± 0.5	11.7 ± 0.3	21.9 ± 0.5
Data $n_b = 4$	0	0	0
Template $n_b = 4$	0.23 ± 0.07	0.09 ± 0.04	0.29 ± 0.09

Table 6.2.: 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 11.5 fb^{-1} of data at $\sqrt{s} = 8\text{TeV}$. The uncertainties quoted on the template yields are purely statistical.

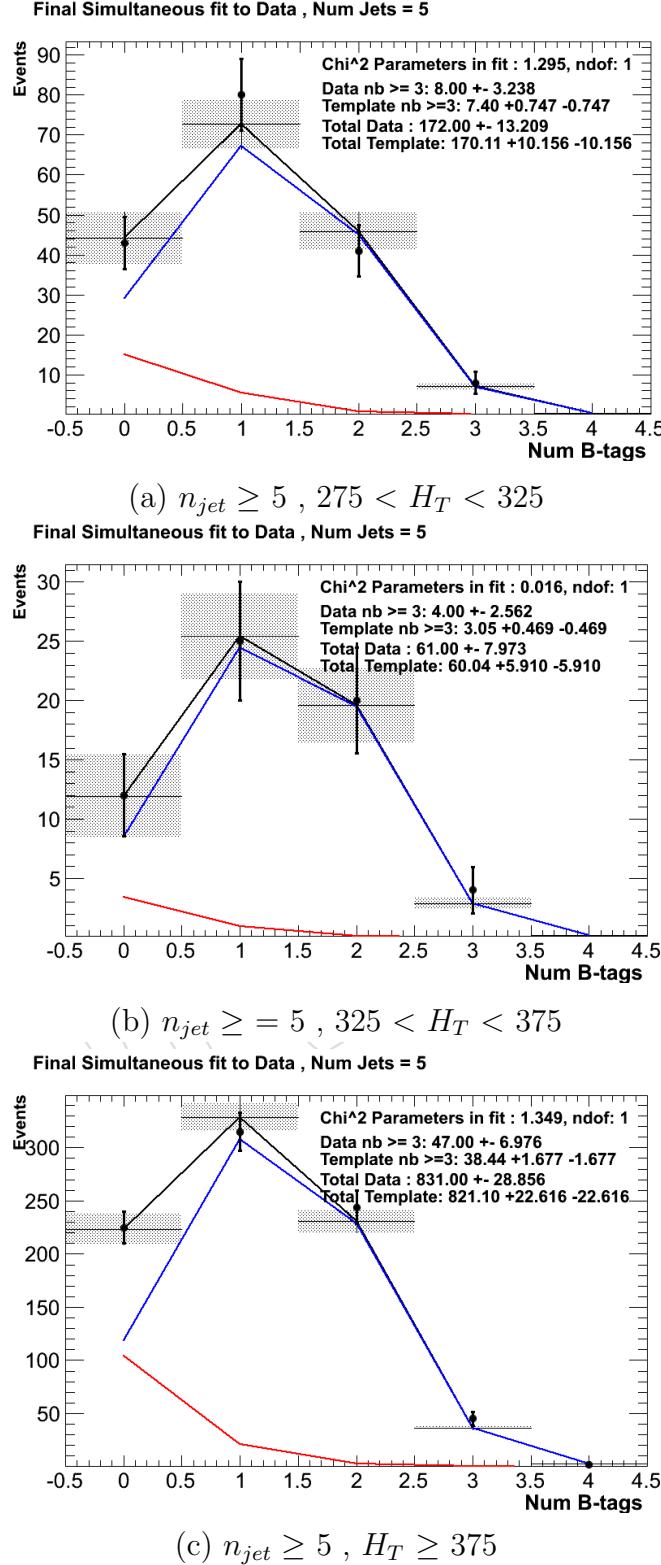


Figure 6.3.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from data, for the $n_{jet} \geq 5$ category and medium CSV working point. The red template represents $Z = 0$, while the blue template represents $Z = 2$. The χ^2 parameter displayed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

1949 The agreement for all working points demonstrates a good control of the b-tagging
1950 efficiency and mis-tagging rates and gives confidence in the method outlined.

1951 6.2.3. Application to the α_T hadronic search region

1952 As an accompaniment to the background estimation methods outlined by the α_T search.
1953 The b-tag template method offers a complimentary way of estimated the background
1954 within the hadronic signal region of the search.

1955 6.3. Conclusions

1956 A **SUSY** signature such as one from gluino-induced third-generation squark production,
1957 would result in a final state with an underlying b-quark content greater than two. In
1958 order to be able to discriminate such signatures from the **SM** background, templates are
1959 generated based on a parameterisation of the number of the **SM** processes, where the
1960 underlying b-quarks per event is typically zero or two. These templates are then fit to
1961 data in a low n_b^{reco} (0-2) control region in order to extrapolate a prediction in a high
1962 n_b^{reco} (3-4) signal region.

1963 The method was demonstrated both in simulation and also in data, using the **SM**
1964 enriched $\mu +$ jets selection from the α_T search, to prove conceptually and experimentally
1965 that the method works and there is adequate control over the efficiency and mis-tagging
1966 rates in data for all working points of the **CSV** tagger. Additionally this method was
1967 also applied to the α_T analysis signal region where good agreement is observed between
1968 data and the background estimation method of the α_T analysis.

Chapter 7.

¹⁹⁶⁹ Conclusions

¹⁹⁷⁰ Conclusions here

¹⁹⁷¹

Appendix A.

₁₉₇₂ Miscellaneous

₁₉₇₃ A.1. Noise Filters

₁₉₇₄ 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.

₁₉₇₅ 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.

1976 The following noise filters are applied, to remove events with spurious, non-physical
1977 jets 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 E_T' spike.
ECAL Laser correction filter	A laser calibration multiplicative factor is applied to correct for transparency loss in each crystal during irradiation. A small number of crystals receive unphysically large values of this correction and become very energetic, resulting in E_T' .

Table A.3.: Noise filters that are applied to remove spurious and non-physical E_T' signatures within the **CMS** detector.

¹⁹⁷⁸ A.2. Primary Vertices

¹⁹⁷⁹ The pileup per event is defined by the number of 'good' reconstructed primary vertices
¹⁹⁸⁰ 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.

¹⁹⁸¹ L1 Jets

¹⁹⁸² B.1. Jet matching efficiencies

¹⁹⁸³ The single jet turn-on curves are derived from events independent of whether the leading
¹⁹⁸⁴ jet in an event is matched to a Level 1 jet using ΔR matching detailed in Section (3.4.3)
¹⁹⁸⁵ or not. These turn-ons are produced from events which are not triggered on jet quantities
¹⁹⁸⁶ and therefore it is not guaranteed that the lead jet of an event will be seeded by a Level
¹⁹⁸⁷ 1 jet. Figure B.1 shows the particular matching efficiency of a lead jet to a L1 jet.

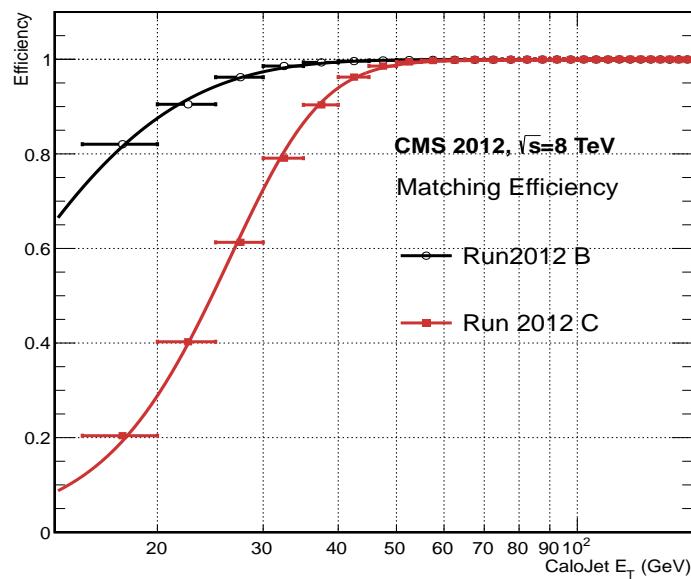


Figure B.1.: Leading jet matching efficiency as a function of the offline CaloJet E_T , measured in an isolated muon triggered dataset in the 2012B and 2012C run periods.

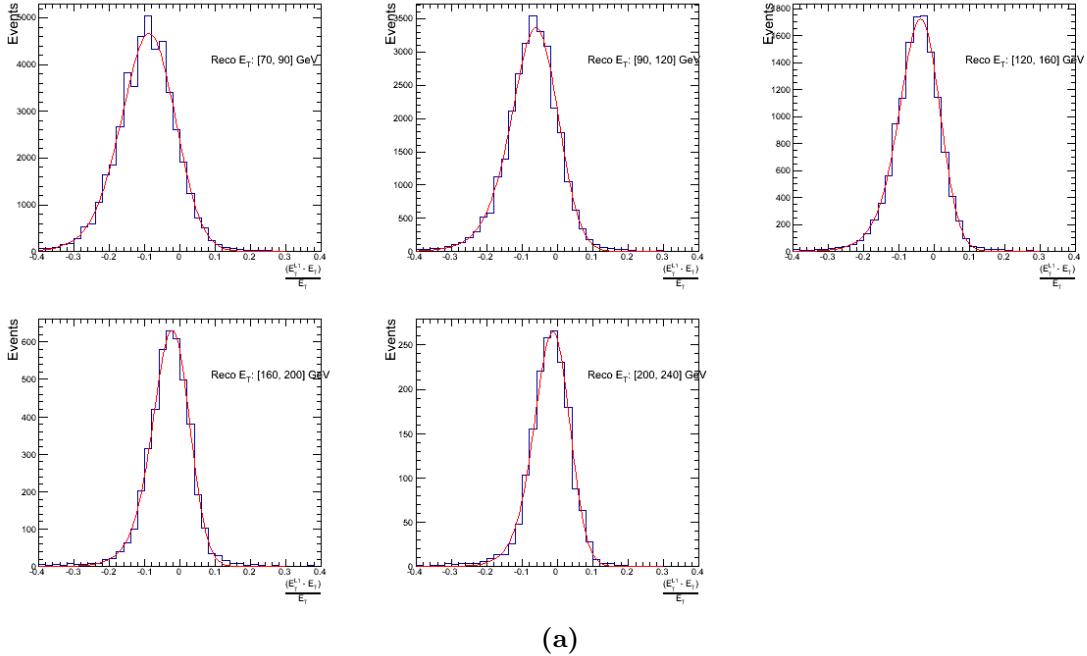
Run Period	μ	σ
2012B	6.62 ± 0.01	0.79 ± 0.03
2012C	19.51 ± 0.03	7.14 ± 0.02

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 .

It can be seen that the turn on is sharper during the 2012B run period. The seed threshold requirement of a 5 GeV jet seed in run 2012C results in more events in which even the lead offline jet does not have an associated L1 jet. For larger jet E_T thresholds, typical of thresholds used in physics analyses, 100% efficiency is observed.

The matching efficiencies have a μ values of 6.62 GeV and 19.51 GeV for Run 2012B and 2012C respectively and is shown in Table B.1.

B.2. Leading Jet Energy Resolution



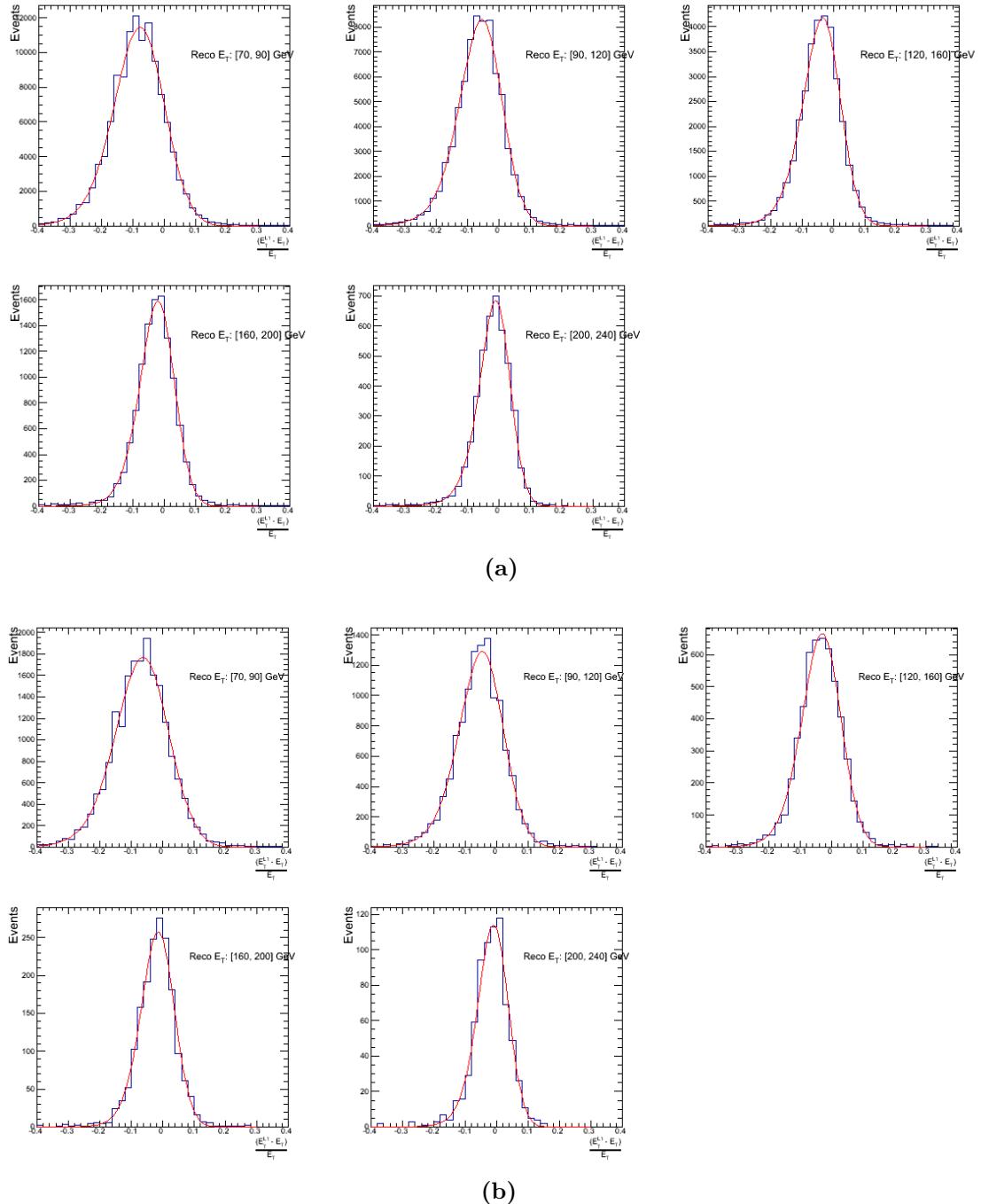
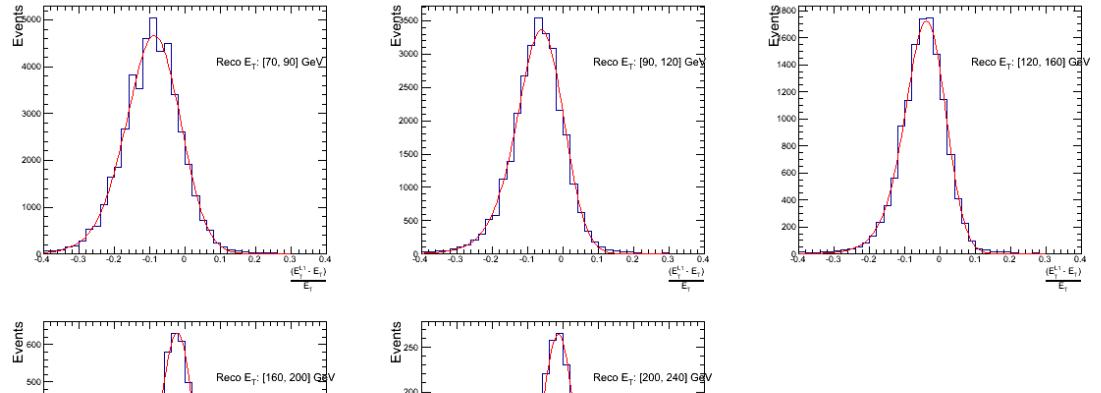
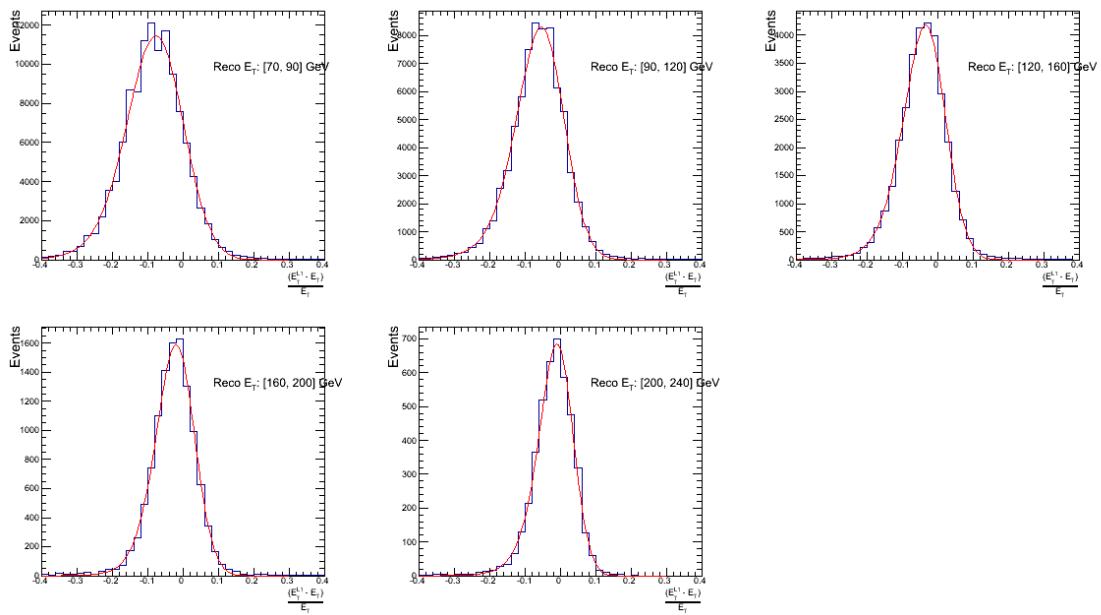


Figure B.2.: Resolution plots of the leading offline jet Calo E_T measured as a function of $\frac{(\text{L1 } E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low (a), medium (b) and high (c) pile-up conditions.



(a)



(b)

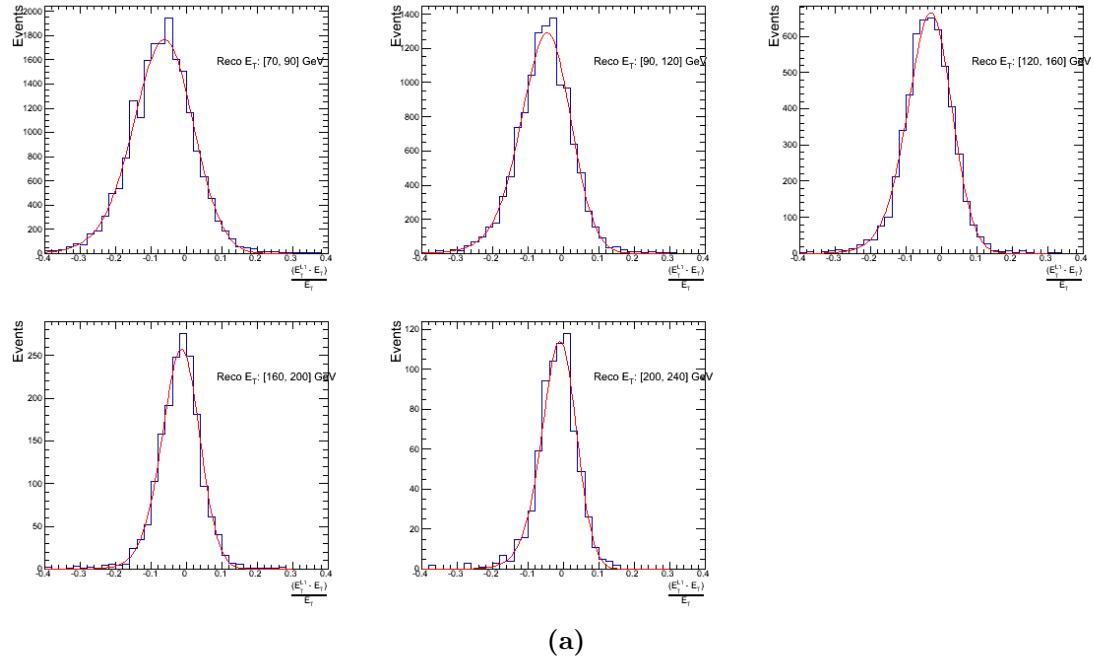


Figure B.3.: Resolution plots of the leading offline jet PF E_T measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low (a), medium (b) and high (c) pile-up conditions.

1995 B.3. Resolution for Energy Sum Quantities

1996 The following plots show the resolution parameters for the four energy sum quantities as
 1997 a function of the quantity (q) itself. In this case, The mean and RMS of the individual
 1998 $\frac{(L1 q - \text{Offline } q)}{\text{Offline } q}$ distributions, in bins of the quantity q is displayed.

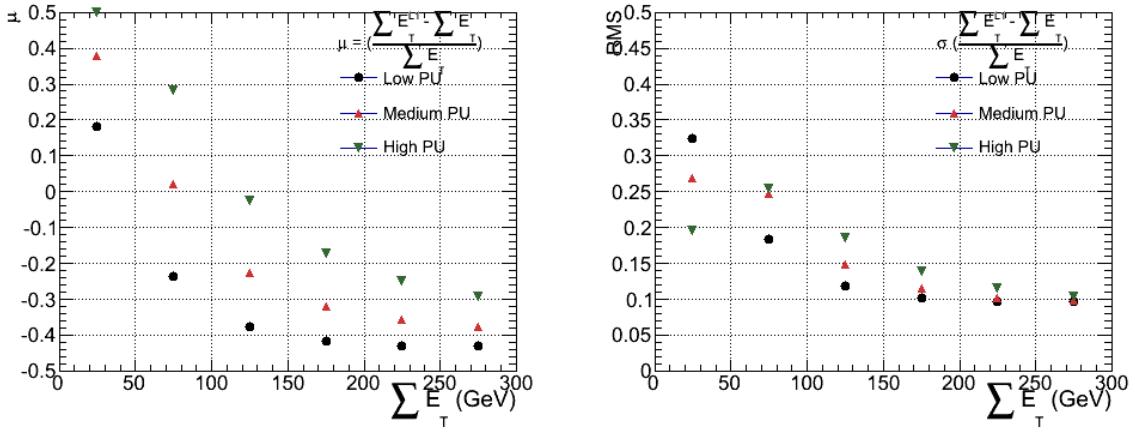


Figure B.4.: $\sum E_T$ resolution parameters in bins of Calo $\sum E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

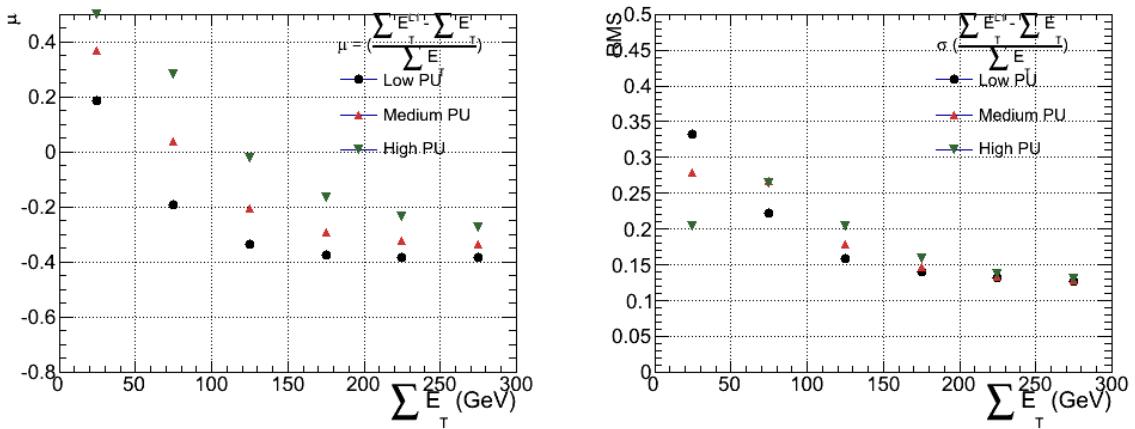


Figure B.5.: $\sum E_T$ resolution parameters in bins of PF $\sum E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

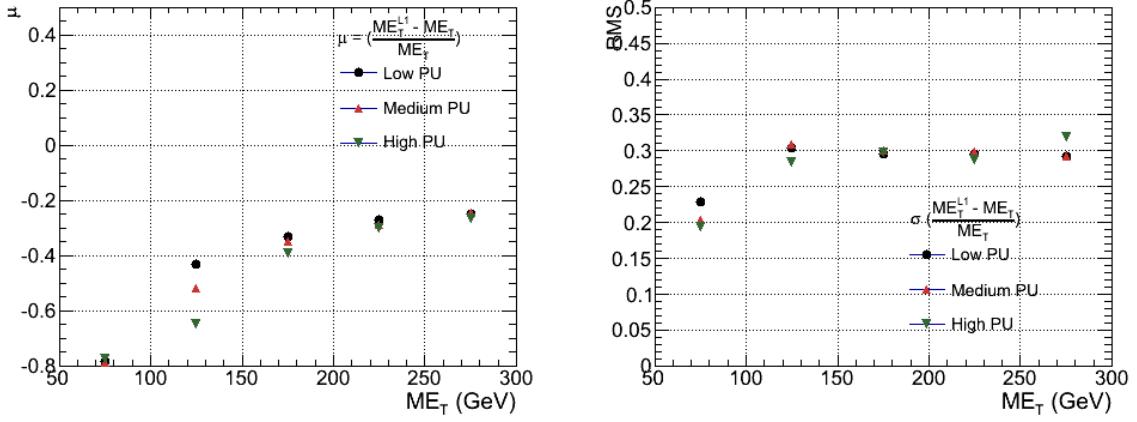


Figure B.6.: \mathcal{E}_T resolution parameters in bins of Calo \mathcal{E}_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

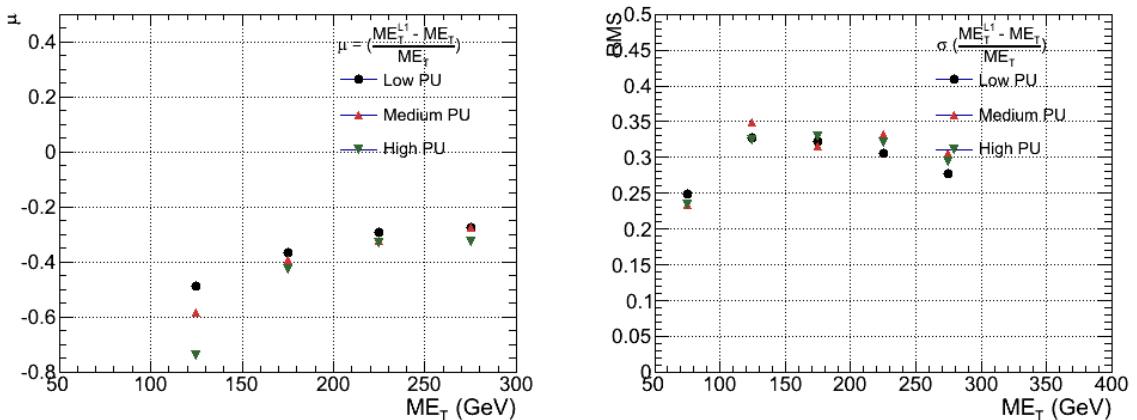


Figure B.7.: \mathcal{E}_T resolution parameters in bins of PF \mathcal{E}_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

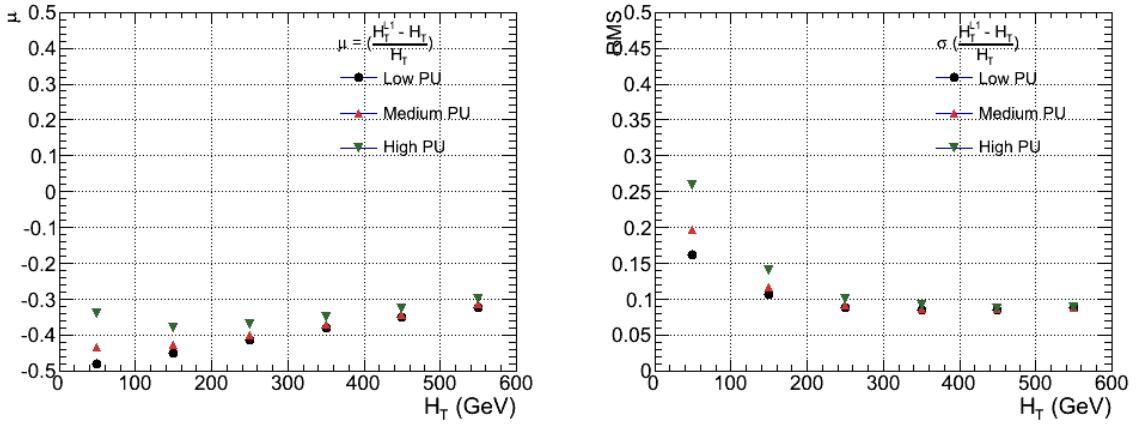


Figure B.8.: H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

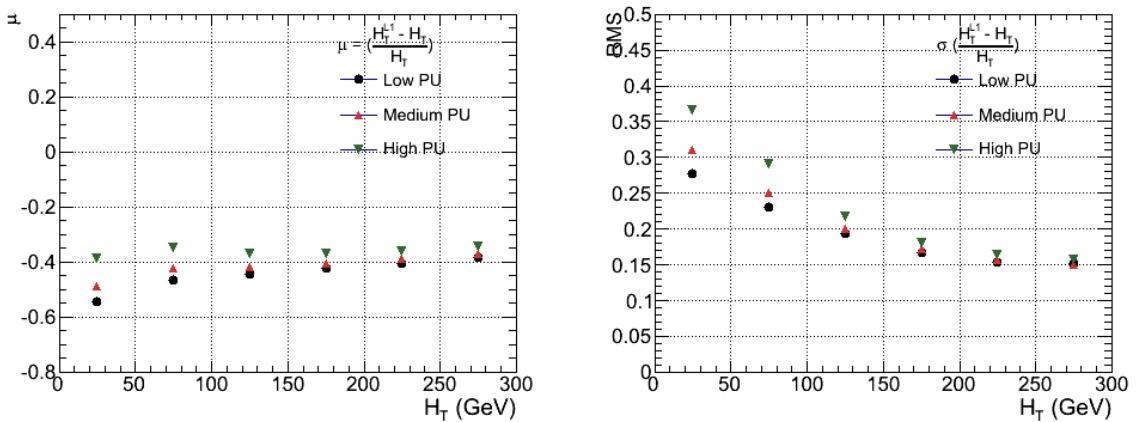


Figure B.9.: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

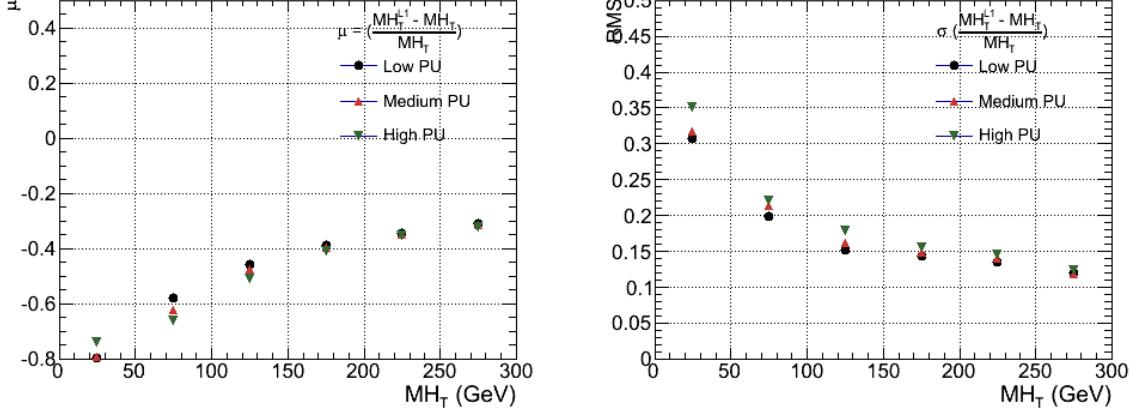


Figure B.10.: H_T resolution parameters in bins of H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

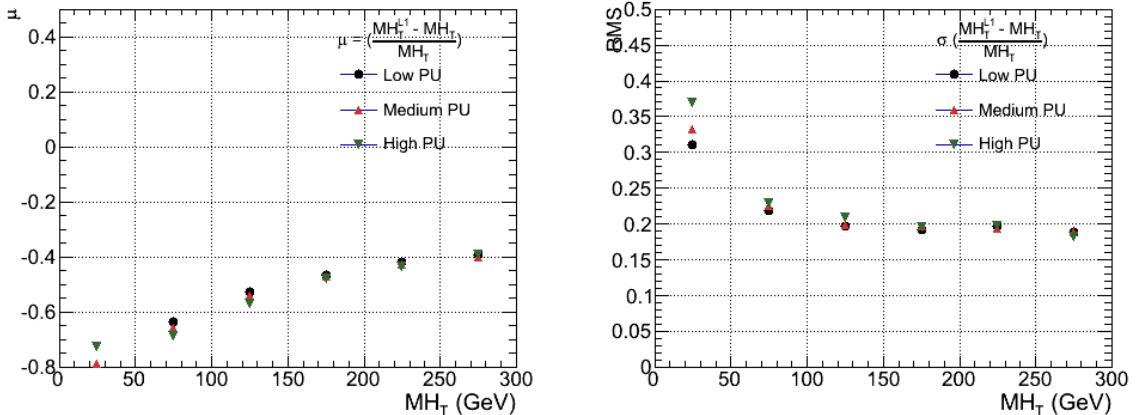


Figure B.11.: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

Appendix C.

Additional material on background estimation methods

C.1. Determination of k_{QCD}

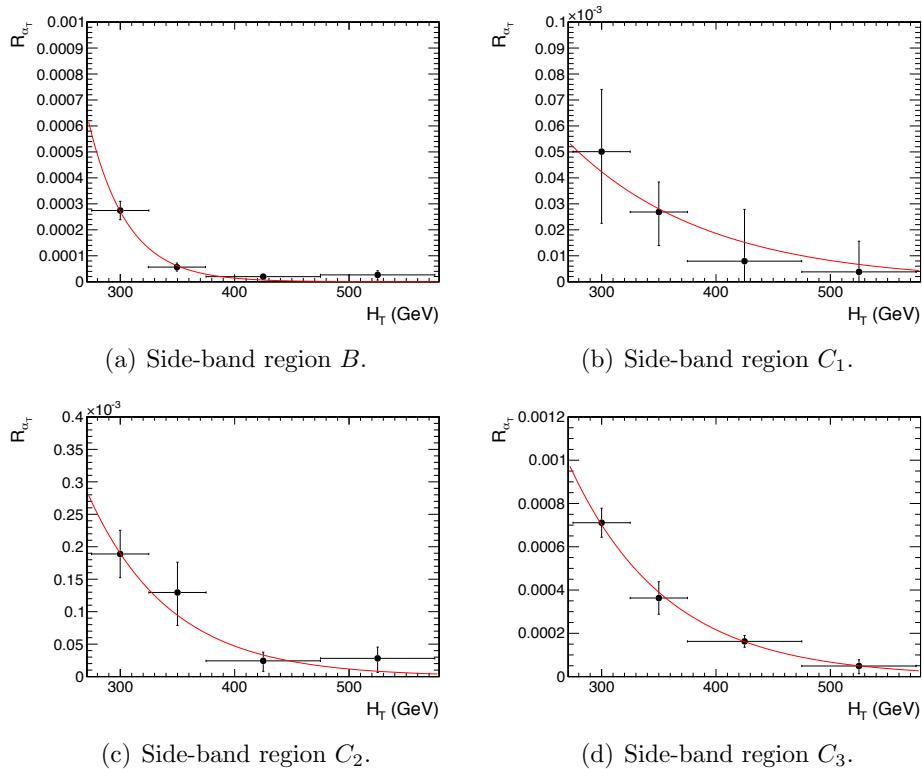


Figure C.1.: $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is conducted between the H_T region $275 < H_T < 575$.

2003 **C.2. Effect of varying background cross sections on**
2004 **closure tests**

2005 Closure tests with cross section variations of +20% and -20% applied to $W + \text{jets}$ and $t\bar{t}$
2006 processes respectively.

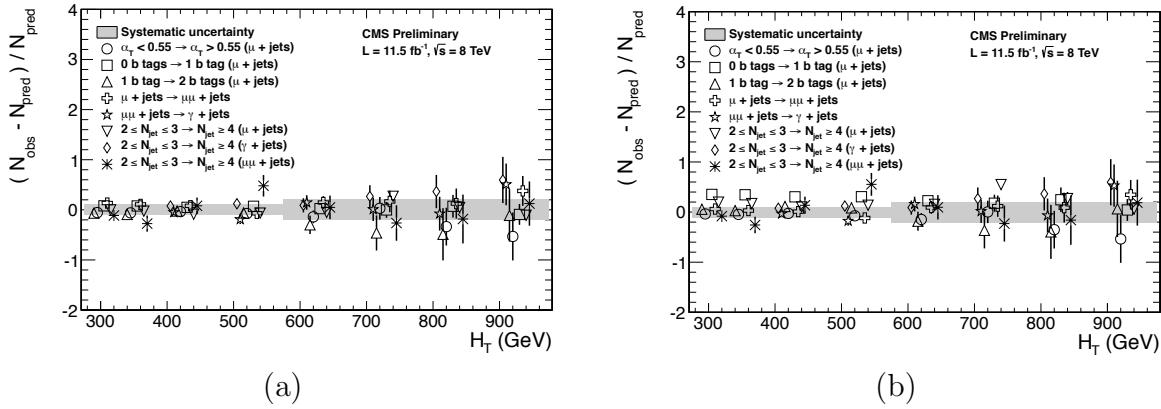


Figure C.2.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins: (a) $2 \leq n_{jet} \leq 3$ and (b) $n_{jet} \geq 4$.

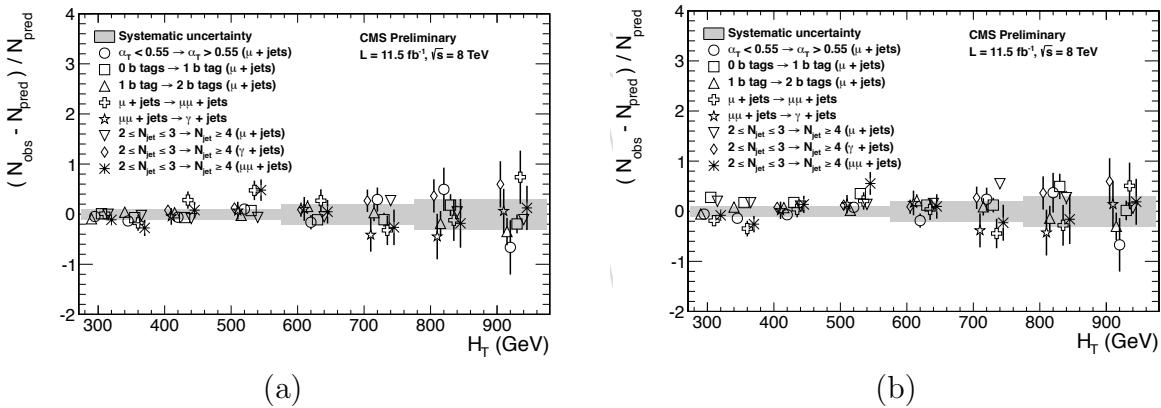


Figure C.3.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins: (a) $2 \leq n_{jet} \leq 3$ and (b) $n_{jet} \geq 4$.

	H_T (GeV)
--	-------------

Table C.1.

Appendix D.

₂₀₀₇ Additional Material For B-tag ₂₀₀₈ Template Method

₂₀₀₉ D.1. Templates Fits in Simulation

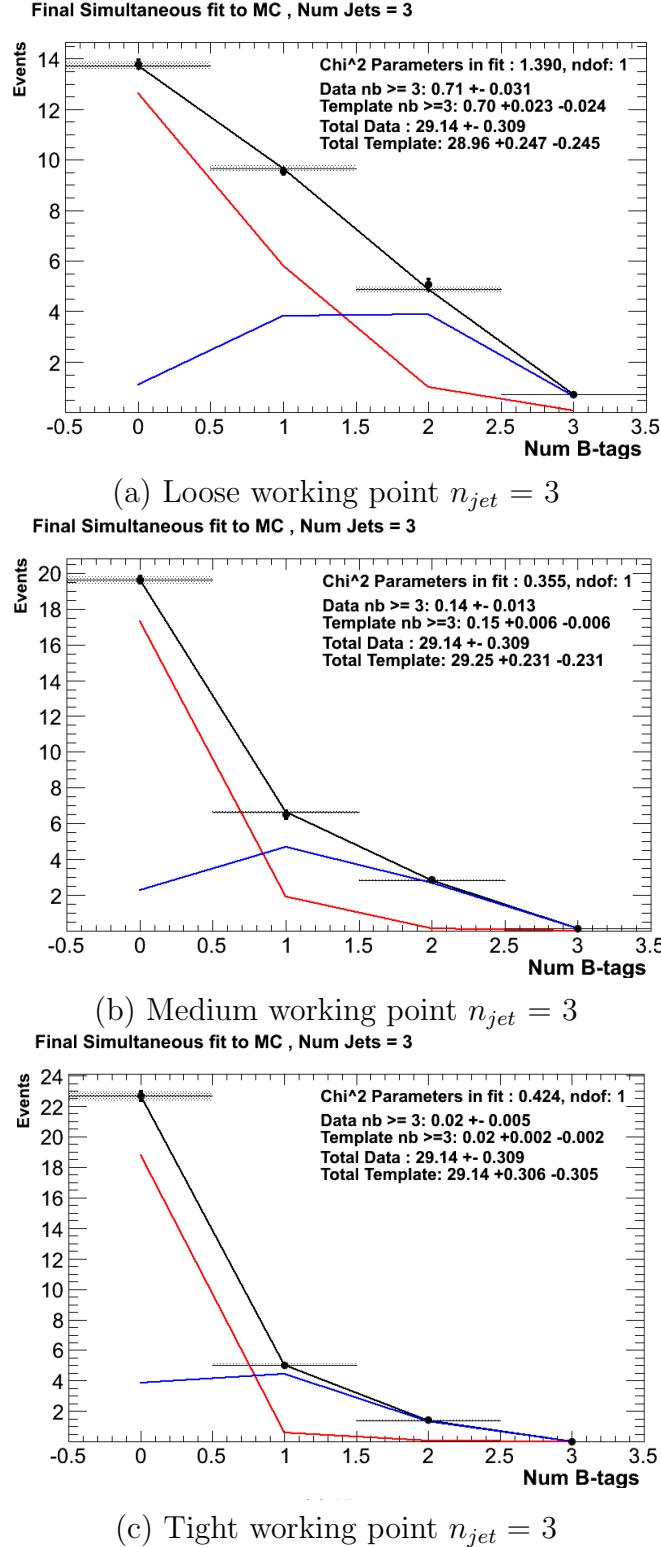


Figure D.1.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 3$ category. The red template represents $Z = 0$, while the blue template represents $Z = 2$. Grey bands represent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

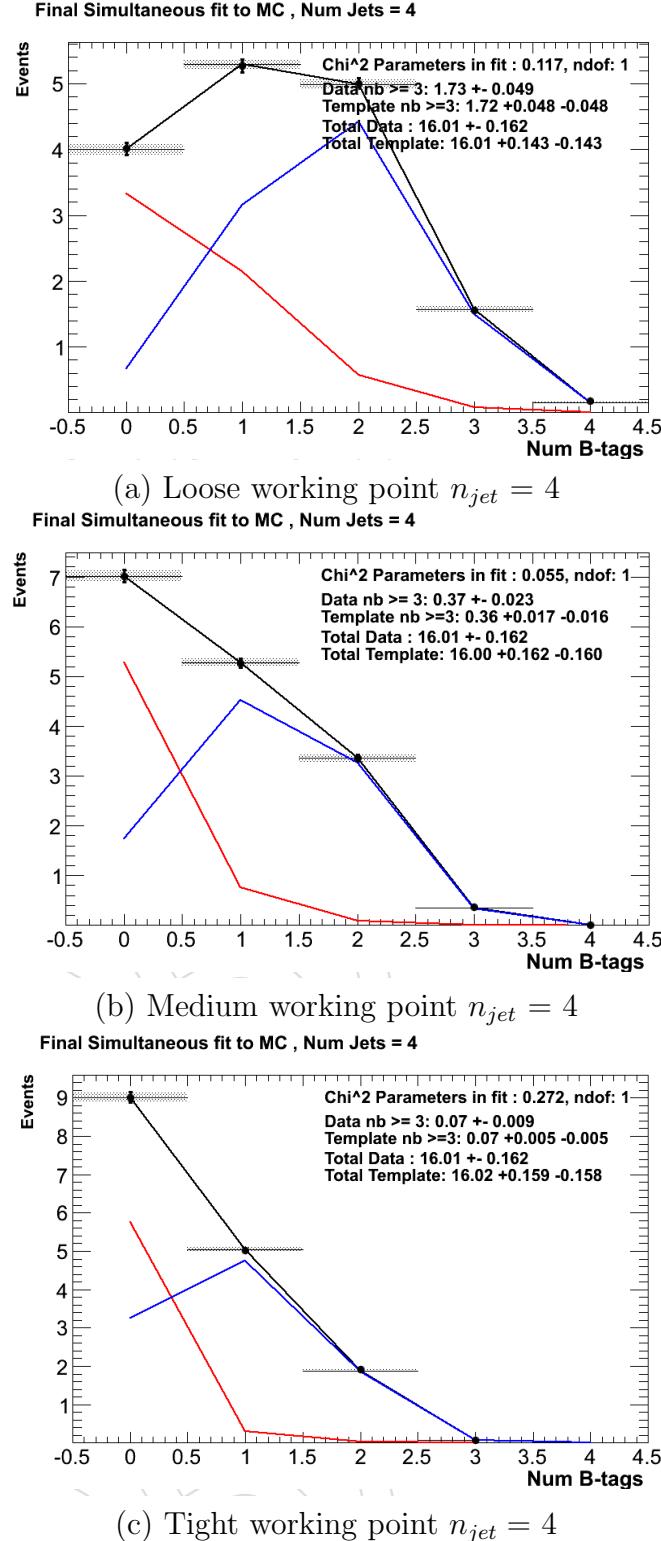


Figure D.2.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 4$ category. The red template represents $Z = 0$, while the blue template represents $Z = 2$. Grey bands represent 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.

2010 **D.2. Pull Distributions for Template Fits**

2011 **D.3. Templates Fits in Data**

2012 Template fits for the loose **CSV** working point :

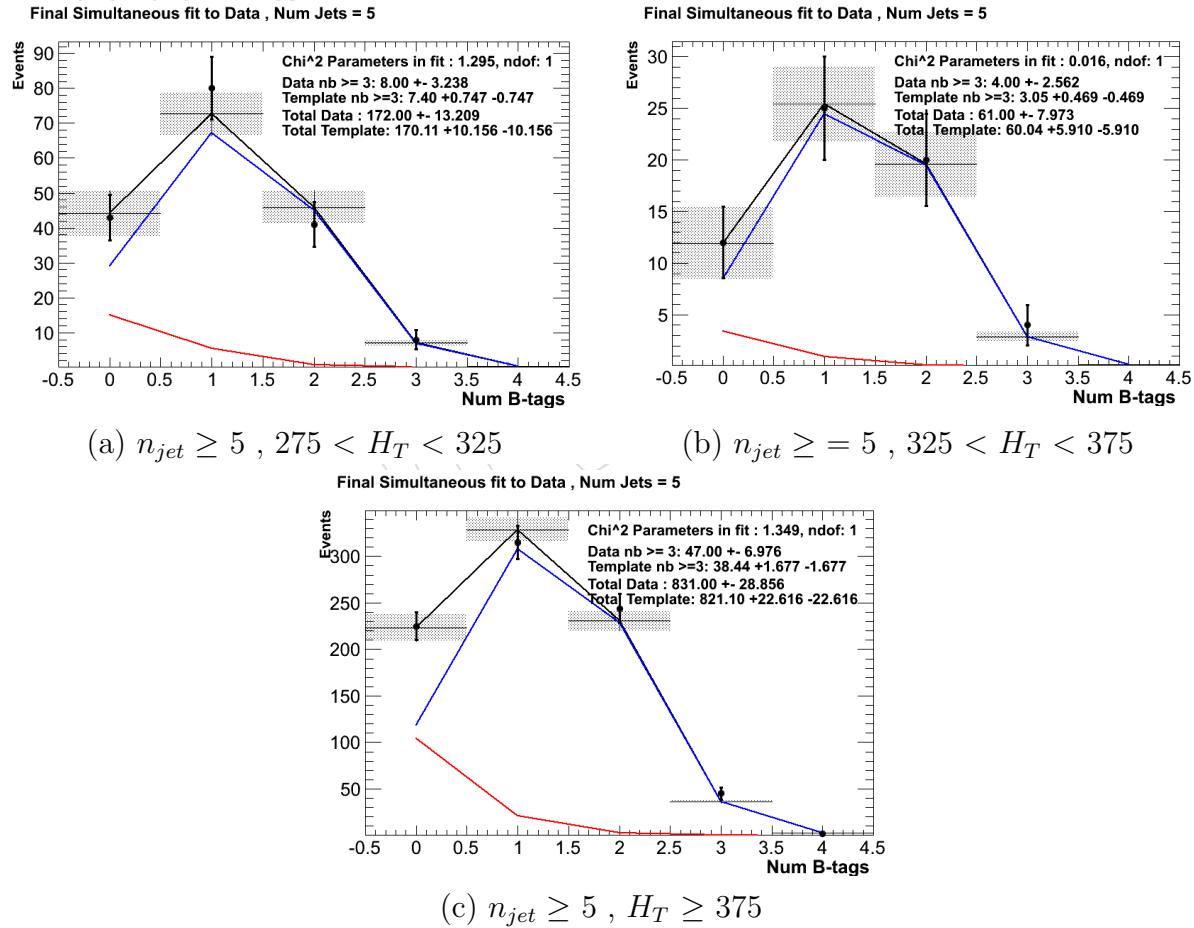


Figure D.3.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} \geq 5$ category and loose **CSV** working point. The red template represents $Z = 0$, while the blue template represents $Z = 2$. The χ^2 parameter displayed represents the goodness of fit to the low n_b^{reco} (0-2) control region.

2013 Template fits for the tight **CSV** working point :

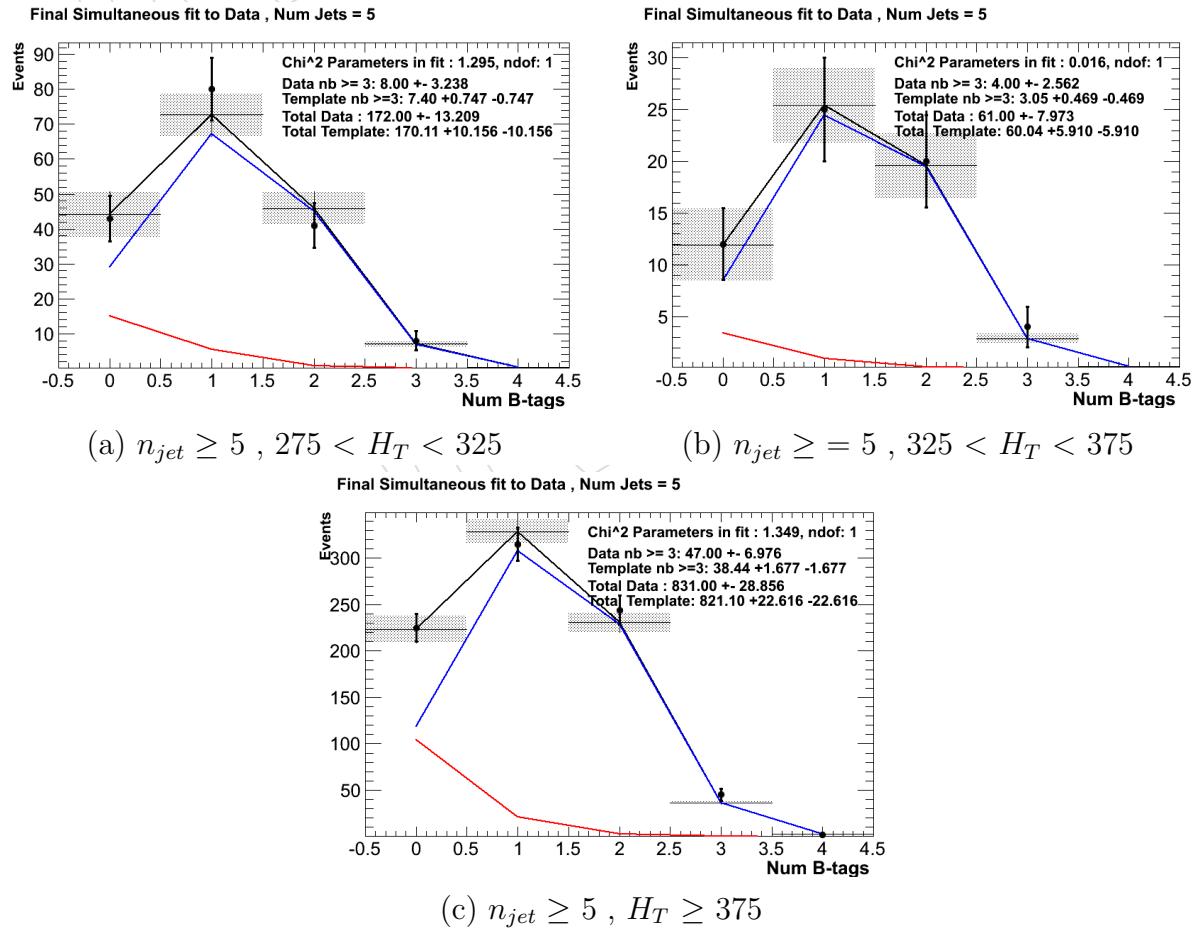


Figure D.4.: The results of fitting the $Z = 0$ and $Z = 2$ templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} \geq 5$ category and tight CSV working point. The red template represents $Z = 0$, while the blue template represents $Z = 2$. The χ^2 parameter displayed represents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

2014

²⁰¹⁵ Bibliography

- ²⁰¹⁶ [1] J. Beringer *et al.*, “Review of Particle Physics (RPP),” *Phys.Rev.*, vol. D86, p. 010001, 2012.
- ²⁰¹⁸ [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.
- ²⁰²¹ [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.
- ²⁰²⁴ [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.
- ²⁰²⁶ [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.
- ²⁰²⁹ [6] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.
- ²⁰³¹ [7] S. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl.Phys.*, vol. 22, pp. 579–588, 1961.
- ²⁰³³ [8] A. Salam, “Weak and Electromagnetic Interactions,” *Conf.Proc.*, vol. C680519, pp. 367–377, 1968.
- ²⁰³⁵ [9] G. Hooft, “Renormalizable lagrangians for massive yang-mills fields,” *Nuclear Physics B*, vol. 35, no. 1, pp. 167 – 188, 1971.
- ²⁰³⁷ [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.

- 2040 [11] G. Arnison *et al.*, “Experimental Observation of Lepton Pairs of Invariant Mass
2041 Around 95-GeV/c**2 at the CERN SPS Collider,” *Phys.Lett.*, vol. B126, pp. 398–410,
2042 1983.
- 2043 [12] M. Banner *et al.*, “Observation of Single Isolated Electrons of High Transverse
2044 Momentum in Events with Missing Transverse Energy at the CERN anti-p p
2045 Collider,” *Phys.Lett.*, vol. B122, pp. 476–485, 1983.
- 2046 [13] E. Noether, “Invariante variationsprobleme,” *Nachrichten von der Gesellschaft
2047 der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse*, vol. 1918,
2048 pp. 235–257, 1918.
- 2049 [14] F. Halzen and A. D. Martin, “Quarks and leptons.” Wiley, 1985.
- 2050 [15] *Introduction to Elementary Particles*. Wiley-VCH, 2nd ed., Oct. 2008.
- 2051 [16] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, “Ex-
2052 perimental Test of Parity Conservation in Beta Decay,” *Physical Review*, vol. 105,
2053 pp. 1413–1415, Feb. 1957.
- 2054 [17] P. Higgs, “Broken symmetries, massless particles and gauge fields,” *Physics Letters*,
2055 vol. 12, no. 2, pp. 132 – 133, 1964.
- 2056 [18] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,”
2057 *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.
- 2058 [19] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*,
2059 vol. 13, pp. 508–509, Oct 1964.
- 2060 [20] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global conservation laws and
2061 massless particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.
- 2062 [21] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov
2063 1967.
- 2064 [22] H. Yukawa, “On the interaction of elementary particles. i,” *Progress of Theoretical
2065 Physics Supplement*, vol. 1, pp. 1–10, 1955.
- 2066 [23] Y. e. a. Fukuda, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev.
2067 Lett.*, vol. 81, pp. 1562–1567, Aug 1998.
- 2068 [24] R. Becker-Szendy, C. Bratton, D. Casper, S. Dye, W. Gajewski, *et al.*, “A Search
2069 for muon-neutrino oscillations with the IMB detector,” *Phys.Rev.Lett.*, vol. 69,

- 2070 pp. 1010–1013, 1992.
- 2071 [25] S. P. Martin, “A Supersymmetry primer,” 1997.
- 2072 [26] H. Nilles, *Supersymmetry, Supergravity and Particle Physics*. Physics reports, North-Holland Physics Publ., 1984.
- 2074 [27] H. E. Haber and G. L. Kane, “The Search for Supersymmetry: Probing Physics
2075 Beyond the Standard Model,” *Phys.Rept.*, vol. 117, pp. 75–263, 1985.
- 2076 [28] E. Witten, “Dynamical Breaking of Supersymmetry,” *Nucl.Phys.*, vol. B188, p. 513,
2077 1981.
- 2078 [29] J. Wess and B. Zumino, “Supergauge transformations in four dimensions,” *Nuclear
2079 Physics B*, vol. 70, no. 1, pp. 39 – 50, 1974.
- 2080 [30] H. Muller-Kirsten and A. Wiedemann, *Introduction to Supersymmetry*. World
2081 Scientific lecture notes in physics, World Scientific, 2010.
- 2082 [31] I. Aitchison, *Supersymmetry in Particle Physics: An Elementary Introduction*.
2083 Cambridge University Press, 2007.
- 2084 [32] K. A. Intriligator and N. Seiberg, “Lectures on Supersymmetry Breaking,”
2085 *Class.Quant.Grav.*, vol. 24, pp. S741–S772, 2007.
- 2086 [33] Y. Shadmi, “Supersymmetry breaking,” pp. 147–180, 2006.
- 2087 [34] C. Burgess, P. G. Camara, S. de Alwis, S. Giddings, A. Maharana, *et al.*, “Warped
2088 Supersymmetry Breaking,” *JHEP*, vol. 0804, p. 053, 2008.
- 2089 [35] H. Murayama, “Supersymmetry breaking made easy, viable, and generic,” 2007.
- 2090 [36] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering
2091 Events*. Cambridge University Press, 2006.
- 2092 [37] S. P. Martin, “Implications of supersymmetric models with natural r-parity conser-
2093 vation,” 1996.
- 2094 [38] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, “Study of constrained
2095 minimal supersymmetry,” *Phys.Rev.*, vol. D49, pp. 6173–6210, 1994.
- 2096 [39] C. Stlege, G. Bertone, D. Cerdeno, M. Fornasa, R. Ruiz de Austri, *et al.*, “Updated
2097 global fits of the cmSSM including the latest LHC SUSY and Higgs searches and
2098 XENON100 data,” *JCAP*, vol. 1203, p. 030, 2012.

- 2099 [40] M. Citron, J. Ellis, F. Luo, J. Marrouche, K. Olive, *et al.*, “The End of the CMSSM
2100 Coannihilation Strip is Nigh,” *Phys.Rev.*, vol. D87, p. 036012, 2013.
- 2101 [41] D. Ghosh, M. Guchait, S. Raychaudhuri, and D. Sengupta, “How Constrained is
2102 the cMSSM?,” *Phys.Rev.*, vol. D86, p. 055007, 2012.
- 2103 [42] D. Alves *et al.*, “Simplified Models for LHC New Physics Searches,” *J.Phys.*, vol. G39,
2104 p. 105005, 2012.
- 2105 [43] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization
2106 of New Physics at the LHC,” *Phys.Rev.*, vol. D79, p. 075020, 2009.
- 2107 [44] S. Chatrchyan *et al.*, “Interpretation of Searches for Supersymmetry with simplified
2108 Models,” *Phys.Rev.*, vol. D88, p. 052017, 2013.
- 2109 [45] J. Hisano, K. Kurosawa, and Y. Nomura, “Natural effective supersymmetry,”
2110 *Nucl.Phys.*, vol. B584, pp. 3–45, 2000.
- 2111 [46] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures,” *JHEP*,
2112 vol. 1209, p. 035, 2012.
- 2113 [47] B. Allanach and B. Gripaios, “Hide and Seek With Natural Supersymmetry at the
2114 LHC,” *JHEP*, vol. 1205, p. 062, 2012.
- 2115 [48] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3,
2116 p. S08002, 2008.
- 2117 [49] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,”
2118 *JINST*, vol. 3, 2008.
- 2119 [50] R. Adolphi *et al.*, “The cms experiment at the cern lhc,” *JINST*, vol. 0803, p. S08004,
2120 2008.
- 2121 [51] A. A. Alves *et al.*, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005,
2122 2008.
- 2123 [52] J.-L. Caron, “Lhc layout.. schema general du lhc..” Sep 1997.
- 2124 [53] C. Collaboration, “Cms luminosity - public results.”
2125 twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults., 2011.
- 2126 [54] CERN, “Cms compact muon solenoid..” <http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg>
2127 Feb 2010.

- 2128 [55] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical
2129 Design Report CMS, Geneva: CERN, 1997.
- 2130 [56] *The CMS muon project: Technical Design Report*. Technical Design Report CMS,
2131 Geneva: CERN, 1997.
- 2132 [57] CMS Collaboration, “The cms physics technical design report, volume 1,”
2133 *CERN/LHCC*, vol. 2006-001, 2006.
- 2134 [58] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,”
2135 *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.
- 2136 [59] “Jet performance in pp collisions at 7 tev,” Tech. Rep. CMS-PAS-JME-10-003,
2137 CERN, Geneva, 2010.
- 2138 [60] X. Janssen, “Underlying event and jet reconstruction in cms,” Tech. Rep. CMS-CR-
2139 2011-012, CERN, Geneva, Jan 2011.
- 2140 [61] T. C. collaboration, “Determination of jet energy calibration and transverse mo-
2141 mentum resolution in cms,” *Journal of Instrumentation*, vol. 6, no. 11, p. P11002,
2142 2011.
- 2143 [62] R. Eusebi and on behalf of the CMS collaboration), “Jet energy corrections and
2144 uncertainties in cms: reducing their impact on physics measurements,” *Journal of
2145 Physics: Conference Series*, vol. 404, no. 1, p. 012014, 2012.
- 2146 [63] “Algorithms for b Jet identification in CMS,” Tech. Rep. CMS-PAS-BTV-09-001,
2147 CERN, 2009. Geneva, Jul 2009.
- 2148 [64] “Performance of b tagging at $\sqrt{s}=8$ tev in multijet, ttbar and boosted topology
2149 events,” no. CMS-PAS-BTV-13-001, 2013.
- 2150 [65] T. C. collaboration, “Identification of b-quark jets with the cms experiment,” *Journal
2151 of Instrumentation*, vol. 8, no. 04, p. P04013, 2013.
- 2152 [66] S. Dasu *et al.*, “CMS. The TriDAS project. Technical design report, vol. 1: The
2153 trigger systems,” 2000.
- 2154 [67] P. Sphicas, “CMS: The TriDAS project. Technical design report, Vol. 2: Data
2155 acquisition and high-level trigger,” 2002.
- 2156 [68] J. B. et al., “Calibration and Performance of the Jets and Energy Sums in the
2157 Level-1 Trigger ,” no. CMS IN 2013/006 (2013), 2013.

- 2158 [69] B. et al., “Study of Level-1 Trigger Jet Performance in High Pile-up Running
2159 Conditions,” no. CMS IN 2013/007 (2013), 2013.
- 2160 [70] J. J. Brooke, “Performance of the cms level-1 trigger,” Tech. Rep. CMS-CR-2012-322.
2161 CERN-CMS-CR-2012-322, CERN, Geneva, Nov 2012.
- 2162 [71] L. Randall and D. Tucker-Smith, “Dijet searches for supersymmetry at the large
2163 hadron collider,” *Phys. Rev. Lett.*, vol. 101, p. 221803, Nov 2008.
- 2164 [72] “SUSY searches with dijet events,” 2008.
- 2165 [73] “Search strategy for exclusive multi-jet events from supersymmetry at CMS,” Tech.
2166 Rep. CMS-PAS-SUS-09-001, CERN, 2009. Geneva, Jul 2009.
- 2167 [74] “Calorimeter Jet Quality Criteria for the First CMS Collision Data,” Tech. Rep.
2168 CMS-PAS-JME-09-008, CERN, 2010. Geneva, Apr 2010.
- 2169 [75] T. C. collaboration, “Performance of cms muon reconstruction in pp collision events
2170 at $\sqrt{s} = 7$ tev,” *Journal of Instrumentation*, vol. 7, no. 10, p. P10002, 2012.
- 2171 [76] “Search for supersymmetry in events with photons and missing energy,” Tech. Rep.
2172 CMS-PAS-SUS-12-018, CERN, Geneva, 2012.
- 2173 [77] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas,” *Phys.Lett.*,
2174 vol. B659, pp. 119–126, 2008.
- 2175 [78] Z. Bern, G. Diana, L. J. Dixon, F. Febres Cordero, S. Hoche, H. Ita, D. A. Kosower,
2176 D. Maitre, and K. J. Ozeren, “Driving missing data at next-to-leading order,” *Phys.
2177 Rev. D*, vol. 84, p. 114002, Dec 2011.
- 2178 [79] “Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus
2179 Jets Search,” Tech. Rep. CMS-PAS-SUS-08-002, CERN, 2009. Geneva, Jan 2009.
- 2180 [80] Z. Bern, G. Diana, L. Dixon, F. Febres Cordero, S. Hoche, *et al.*, “Driving Missing
2181 Data at Next-to-Leading Order,” *Phys.Rev.*, vol. D84, p. 114002, 2011.
- 2182 [81] R. Cousins, “Probability density functions for positive nuisance parameters.”
2183 <http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf>, 2012.
- 2184 [82] C. Collaboration, “Cms btag pog : Cms b-tagging performance database.”
2185 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>, 2013.