

¹ **Searches for Supersymmetric Signatures in
2 all Hadronic Final States with the α_T
3 Variable.**

⁴ Darren Burton

⁵ Imperial College London
⁶ Department of Physics

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

Abstract

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

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

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

38

Declaration

39

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

46

Darren Burton

47

Acknowledgements

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

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

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

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

67 **Contents**

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

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

| | | |
|-----|---|-----|
| 128 | 5.2.2. Interpretation in Simplified Signal Models | 107 |
| 129 | 6. SUSY Searches with B-tag Templates | 111 |
| 130 | 6.1. Concept | 112 |
| 131 | 6.1.1. Fitting Procedure | 112 |
| 132 | 6.2. Application to the α_T Search | 116 |
| 133 | 6.2.1. Proof of Principle in Simulation | 117 |
| 134 | 6.2.2. Results in a Data Control Sample | 120 |
| 135 | 6.2.3. Application to the α_T Hadronic Search Region | 122 |
| 136 | 6.3. Summary | 125 |
| 137 | 7. Conclusions | 126 |
| 138 | A. Miscellaneous | 128 |
| 139 | A.1. Jet Identification Criteria | 128 |
| 140 | A.2. Primary Vertices | 129 |
| 141 | B. L1 Jets | 130 |
| 142 | B.1. Jet Matching Efficiencies | 130 |
| 143 | B.2. Leading Jet Energy Resolution | 131 |
| 144 | B.3. Resolution for Energy Sum Quantities | 134 |
| 145 | C. Additional Material on Background Estimation Methods | 136 |
| 146 | C.1. Determination of k_{QCD} | 136 |
| 147 | C.2. Effect of Varying Background Cross-sections on Closure Tests | 137 |
| 148 | D. Additional Material For B-tag Template Method | 139 |
| 149 | D.1. Templates Fits in Simulation | 139 |
| 150 | D.2. Pull Distributions for Template Fits | 142 |
| 151 | D.3. Templates Fits in Data Control Sample | 143 |
| 152 | D.4. Templates Fits in Data Signal Region | 145 |
| 153 | Bibliography | 148 |

¹⁵⁴ List of Figures

| | | |
|----------------------|---|----|
| ¹⁵⁵ 2.1. | One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion. | 14 |
| ¹⁵⁶ 2.2. | Two example simplified model decay chains. | 19 |
| ¹⁵⁸ 3.1. | A top-down layout of the LHC, with the position of the four main detectors labelled. | 21 |
| ¹⁶⁰ 3.2. | The total integrated luminosity delivered to and collected by Compact Muon Solenoid (CMS) during the 2012 8 TeV pp runs | 22 |
| ¹⁶² 3.3. | A pictorial depiction of the CMS detector. | 24 |
| ¹⁶³ 3.4. | Illustration of the CMS Electromagnetic CALorimeter (ECAL). | 26 |
| ¹⁶⁴ 3.5. | Schematic of the CMS Hadronic CALorimeter (HCAL). | 27 |
| ¹⁶⁵ 3.6. | Combined Secondary Vertex (CSV) algorithm discriminator values in enriched ttbar and inclusive multi-jet samples | 31 |
| ¹⁶⁷ 3.7. | Data/MC b-tag scale factors derived using the Combined Secondary Vertex Medium Working Point (CSVM) tagger. | 32 |
| ¹⁶⁹ 3.8. | Data/MC mis-tag scale factors derived using the CSVM tagger. | 33 |
| ¹⁷⁰ 3.9. | An overview of the different components of the CMS L1 trigger system . | 34 |
| ¹⁷¹ 3.10. | Illustration of the dimensions of the Level-1 jet finder window. | 37 |
| ¹⁷² 3.11. | L1 jet efficiency turn-on curves as a function of the offline CaloJet and PFJet E_T | 39 |
| ¹⁷⁴ 3.12. | L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T for the 2012 run period B and C. | 41 |

| | | |
|-----|--|----|
| 176 | 3.13. L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T | 41 |
| 177 | 3.14. Trigger cross section for the L1HTT150 trigger path. | 43 |
| 178 | 3.15. L1 jet efficiency turn-on curves as a function of the leading offline E_T Calo | |
| 179 | (left) and PF (right) jet, for low, medium and high pile-up conditions. | 44 |
| 180 | 3.16. Fit values from an Exponentially Modified Gaussian (EMG) function fitted | |
| 181 | to the resolution plots of leading Calo jet E_T measured as a function of | |
| 182 | $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low, medium and high pile-up conditions. | 45 |
| 183 | 3.17. Fit values from an EMG function fitted to the resolution plots of leading | |
| 184 | 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 | |
| 185 | high pile-up conditions. | 46 |
| 186 | 4.1. Reconstructed offline H_T distribution in the hadronic signal selection | |
| 187 | (detailed in the following section), from 11.7fb^{-1} of data, in which no α_T | |
| 188 | requirement was made. | 49 |
| 189 | 4.2. The event topologies of background QCD dijet events (right) and a generic | |
| 190 | SUSY signature with genuine Z_T (left). | 50 |
| 191 | 4.3. The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet | |
| 192 | multiplicities after a full analysis selection and $H_T > 375$ requirement. | 52 |
| 193 | 4.4. Pictorial depiction of the analysis strategy employed by the α_T search to | |
| 194 | increase sensitivity to a wide spectra of SUSY models. | 55 |
| 195 | 4.5. Data/MC comparisons of key variables for the hadronic signal region. | 61 |
| 196 | 4.6. Data/MC comparisons of key variables for the $\mu +$ jets selection. | 64 |
| 197 | 4.7. Data/MC comparisons of key variables for the $\mu\mu +$ jets selection. | 66 |
| 198 | 4.8. Data/MC comparisons of key variables for the $\gamma +$ jets selection. | 67 |
| 199 | 4.9. QCD sideband regions, used for determination of k_{QCD} | 69 |
| 200 | 4.10. Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets determined | |
| 201 | from all jets within each H_T category. | 76 |
| 202 | 4.11. Sets of closure tests overlaid on top of the systematic uncertainty used for | |
| 203 | each of the five H_T regions. | 83 |

| | | |
|-----|--|-----|
| 204 | 4.12. Signal efficiencies fo the Simplified Model Spectra (SMS) models (a) T1 205 and (b) T2. | 85 |
| 206 | 5.1. Comparison of the observed yields and Standard Model (SM) expectations 207 given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ 208 jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and 209 $n_{jet} \leq 3$ | 99 |
| 210 | 5.2. Comparison of the observed yields and SM expectations given by the 211 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 212 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$ | 100 |
| 213 | 5.3. Comparison of the observed yields and SM expectations given by the 214 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 215 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$ | 101 |
| 216 | 5.4. Comparison of the observed yields and SM expectations given by the 217 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 218 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$ | 102 |
| 219 | 5.5. Comparison of the observed yields and SM expectations given by the 220 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 221 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$ | 103 |
| 222 | 5.6. Comparison of the observed yields and SM expectations given by the 223 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 224 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$ | 104 |
| 225 | 5.7. Comparison of the observed yields and SM expectations given by the 226 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 227 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$ | 104 |
| 228 | 5.8. Comparison of the observed yields and SM expectations given by the 229 simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ 230 jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$ | 105 |
| 231 | 5.9. Production and decay modes for the various SMS models interpreted 232 within the analysis. | 109 |
| 233 | 5.10. Upper limit of cross section at 95% CL as a function of $m_{\tilde{q}/\tilde{g}}$ and m_{LSP} 234 for various SMS models. | 110 |

| | | |
|-----|---|-----|
| 235 | 6.1. The b-quark (a), c-quark (b), and light-quark (c) tagging efficiency as a function of jet p_T , measured in simulation after the application of the α_T analysis $\mu + \text{jets}$ control sample selection, in the region $H_T > 375$ | 114 |
| 238 | 6.2. An example of a template fit with the defined Z0 (blue) and Z2 (red) templates to data within the low n_b^{reco} control region (left). | 115 |
| 240 | 6.3. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{\text{reco}} = 0\text{-}2$ control region to yields from simulation in the $\mu + \text{jets}$ control sample for the $H_T > 375 \text{ GeV}$, $n_{\text{jet}} \geq 5$ category for all CSV working points. | 119 |
| 243 | 6.4. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{\text{reco}} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSV medium working point, with $n_{\text{jet}} \geq 5$ in each H_T category. | 121 |
| 246 | 6.5. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{\text{reco}} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{\text{jet}} \geq 5$ and $H_T > 375$ category for all CSV working points. | 124 |
| 249 | B.1. Leading jet matching efficiency as a function of the offline CaloJet E_T | 130 |
| 250 | 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 (a) low, (b) medium, and (c) high pile-up conditions. | 132 |
| 252 | 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 (a) low, (b) medium, and (c) high pile-up conditions. | 134 |
| 254 | B.4. H_T resolution parameters in bins of Calo H_T measured in the defined low, medium and high pile up conditions. | 134 |
| 256 | B.5. H_T resolution parameters in bins of PF H_T measured in the defined low, medium and high pile up conditions. | 135 |
| 258 | B.6. \mathcal{H}_T resolution parameters in bins of Calo \mathcal{H}_T measured in the defined low, medium and high pile up conditions. | 135 |
| 260 | B.7. \mathcal{H}_T resolution parameters in bins of PF \mathcal{H}_T measured in the defined low, medium and high pile up conditions. | 135 |
| 262 | C.1. $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is conducted between the H_T region $275 < H_T < 575$ | 136 |

| | | |
|-----|---|-----|
| 264 | C.2. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five H_T regions in the $2 \leq n_{jet} \leq 3$ jet multiplicity category for nominal and varied cross-sections. | 137 |
| 267 | C.3. Sets of closure tests overlaid on top of the systematic uncertainty used for each of the five H_T regions in the $n_{jet} \geq 4$ jet multiplicity category for nominal and varied cross-sections. | 137 |
| 270 | D.1. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to yields from simulation in the $\mu + \text{jets}$ control sample for the $H_T > 375$ GeV, $n_{jet} = 3$ category. | 140 |
| 273 | D.2. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to yields from simulation in the $\mu + \text{jets}$ control sample for the $H_T > 375$ GeV, $n_{jet} = 4$ category. | 141 |
| 276 | D.3. Pull distributions of the normalisation parameter of each template, $\frac{(\hat{\theta} - \theta)}{\sigma}$. Distributions are constructed from 10^4 pseudo-experiments generated by a gaussian distribution with width σ , centred on the nominal template value of each point within the low n_b^{reco} control region. | 142 |
| 280 | D.4. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSVM working point, with $n_{jet} = 3$ in each H_T category. | 143 |
| 283 | D.5. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSV medium working point, with $n_{jet} = 4$ in each H_T category. | 144 |
| 286 | D.6. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{jet} = 3$ and $H_T > 375$ category for all CSV working points. | 145 |
| 289 | D.7. Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{jet} = 4$ and $H_T > 375$ category for all CSV working points. | 146 |

List of Tables

| | | |
|-----|---|----|
| 293 | 2.1. The fundamental particles of the SM, with spin, charge and mass displayed. | 6 |
| 294 | 3.1. Results of a cumulative EMG function fit to the turn-on curves for L1 | |
| 295 | single jet triggers in 2012 Run Period C. | 40 |
| 296 | 3.2. Results of a cumulative EMG function fit to the turn-on curves for L1 | |
| 297 | single jet triggers in the 2012 run period B and C. | 42 |
| 298 | 3.3. Results of a cumulative EMG function fit to the turn-on curves for H_T in | |
| 299 | 2012 run period B and C. | 42 |
| 300 | 3.4. Results of a cumulative EMG function fit to the efficiency turn-on curves | |
| 301 | for L1 single jet triggers in the 2012 run period C, for low,medium and | |
| 302 | high pile-up conditions. | 44 |
| 303 | 3.5. Results of a cumulative EMG function fit to the efficiency turn-on curves | |
| 304 | for Level-1 single jet triggers in the 2012 run period C, for low,medium | |
| 305 | and high pile-up conditions. | 45 |
| 306 | 4.1. A summary of the SMS models interpreted in this analysis, involving both | |
| 307 | direct (D) and gluino-induced (G) production of squarks and their decays. | 48 |
| 308 | 4.2. Muon identification criteria used within the analysis for selection/veto | |
| 309 | purposes in the muon control/signal selections. | 56 |
| 310 | 4.3. Photon identification criteria used within the analysis for selection/veto | |
| 311 | purposes in the $\gamma +$ jets control/signal selections. | 57 |
| 312 | 4.4. Electron identification criteria used within the analysis for veto purposes. | 57 |
| 313 | 4.5. Noise filters that are applied to remove spurious and non-physical E_T | |
| 314 | signatures within the CMS detector. | 58 |

| | | |
|-----|--|----|
| 315 | 4.6. Jet thresholds used in the three H_T regions of the analysis. | 59 |
| 316 | 4.7. Best fit values for the parameters k_{QCD} obtained from sideband regions | |
| 317 | B,C ₁ ,C ₂ ,C ₃ | 70 |
| 318 | 4.8. Measured efficiencies of the H_T and α_T legs of the HT and HT_alphaT triggers in independent analysis bins. | 71 |
| 320 | 4.9. k-factors calculated for different EWK processes. | 72 |
| 321 | 4.10. Comparing yields in simulation within the $\mu + \text{jets}$ selection determined | |
| 322 | from the formula method described in Equation (4.11), and that taken | |
| 323 | directly from simulation. | 75 |
| 324 | 4.11. The absolute change in the Transfer Factor (TF)s used to predict the | |
| 325 | entire signal region SM background, using the $\mu + \text{jets}$ control sample | |
| 326 | when the systematic uncertainties of the data to simulation scale factors | |
| 327 | are varied by $\pm 1\sigma$ | 77 |
| 328 | 4.12. A summary of the results obtained from zeroeth order polynomial (i.e. | |
| 329 | a constant) and linear fits to five sets of closure tests performed in the | |
| 330 | $2 \geq n_{\text{jet}} \geq 3$ category. | 80 |
| 331 | 4.13. A summary of the results obtained from zeroeth order polynomial (i.e. | |
| 332 | a constant) and linear fits to five sets of closure tests performed in the | |
| 333 | $n_{\text{jet}} \geq 4$ category. | 81 |
| 334 | 4.14. A summary of the results obtained from zeroeth order polynomial (i.e. a | |
| 335 | constant) and linear fits to three sets of closure tests performed between | |
| 336 | the $2 \leq n_{\text{jet}} \leq 3$ and $n_{\text{jet}} \geq 4$ categories. | 81 |
| 337 | 4.15. Calculated systematic uncertainties for the five H_T regions, determined | |
| 338 | from the closure tests. | 82 |
| 339 | 4.16. Estimates of systematic uncertainties on the signal efficiency (%) for | |
| 340 | various SMS models when considering points in the region near to the | |
| 341 | diagonal | 90 |
| 342 | 4.17. Estimates of systematic uncertainties on the signal efficiency (%) for | |
| 343 | various SMS models when considering points in the region near to the | |
| 344 | diagonal | 90 |

| | | |
|-----|---|-----|
| 345 | 4.18. The systematic parameters used in H_T bins. | 94 |
| 346 | 4.19. Nuisance parameters used within the different hadronic signal bins of the | |
| 347 | analysis | 96 |
| 348 | 5.1. Summary of control samples used by each fit results, and the Figures in | |
| 349 | which they are displayed. | 98 |
| 350 | 5.2. Comparison of the measured yields in each H_T , n_{jet} and n_b^{reco} jet multiplicity bins for the hadronic sample with the SM expectations and combined | |
| 351 | statistical and systematic uncertainties given by the simultaneous fit. . . | 98 |
| 352 | | |
| 353 | 5.3. A table representing the SMS models interpreted within the analysis. . . | 108 |
| 354 | 6.1. Typical underlying b-quark content of different SM processes which are | |
| 355 | common to many SUSY searches. | 112 |
| 356 | 6.2. Summary of the fit predictions in the n_b^{reco} signal region after combination | |
| 357 | of the $n_{jet} = 3, = 4, \geq 5$ categories compared against yields taken directly | |
| 358 | from simulation. The predictions are extrapolated from a $n_b^{reco} = 0, 1,$ | |
| 359 | 2 control region and simulation yields are normalised to an integrated | |
| 360 | luminosity of 10 fb^{-1} | 119 |
| 361 | 6.3. Summary of the fit predictions in the n_b^{reco} signal region of the $\mu + \text{jets}$ | |
| 362 | control sample, after combination of the $n_{jet} = 3, = 4, \geq 5$ categories.. The | |
| 363 | predictions are extrapolated from a $n_b^{reco} = 0, 1, 2$ control region using | |
| 364 | 11.4 fb^{-1} of $\sqrt{s} = 8\text{TeV}$ data. | 122 |
| 365 | 6.4. Summary of the fit predictions in the n_b^{reco} signal region of the α_T hadronic | |
| 366 | signal selection, after combination of the $n_{jet} = 3, = 4, \geq 5$ categories. The | |
| 367 | predictions are extrapolated from a $n_b^{reco} = 0, 1, 2$ control region using | |
| 368 | 11.7 fb^{-1} of $\sqrt{s} = 8\text{TeV}$ data. | 124 |
| 369 | A.1. Criteria for a reconstructed jet to pass the loose calorimeter jet id. | 128 |
| 370 | A.2. Criteria for a reconstructed jet to pass the loose PF jet id. | 129 |
| 371 | A.3. Criteria for a vertex in an event to be classified as a 'good' reconstructed | |
| 372 | primary vertex. | 129 |

| | | |
|-----|---|-----|
| 373 | B.1. Results of a cumulative EMG function fit to the turn-on curves for the | |
| 374 | matching efficiency of the leading jet in an event to a Level-1 jet in run | |
| 375 | 2012C and 2012B data. | 131 |
| 376 | C.1. Translation factors constructed from the $\mu + \text{jets}$ control sample and signal | |
| 377 | selection MC, to predict yields for the $W + \text{jets}$ and $t\bar{t}$ back-grounds in | |
| 378 | the signal region. | 138 |
| 379 | | |

Chapter 1.

³⁸⁰ Introduction

³⁸¹ During the 20th century, great advances were made in the human understanding of
³⁸² the universe, its origins, its future and its composition. The Standard Model (SM)
³⁸³ first formulated in the 1960s is one of the crowning achievements in science’s quest to
³⁸⁴ explain the most fundamental processes and interactions that make up our universe. It
³⁸⁵ has provided a highly successful explanation for a wide range of phenomena in Particle
³⁸⁶ Physics and has stood up to extensive experimental scrutiny [1].

³⁸⁷ Despite its success it is not a complete theory, with significant questions remaining
³⁸⁸ unanswered. It describes only three of the four known forces with gravity not incorporated
³⁸⁹ within the framework of the SM. Cosmological experiments infer that just $\sim 5\%$ of the
³⁹⁰ observable universe exists as matter, with elusive “Dark Matter” accounting for a further
³⁹¹ $\sim 27\%$ [2]. However no particle predicted by the SM is able to account for it. At higher
³⁹² energy scales, the (non-)unification of the fundamental forces point to problems with the
³⁹³ SM at least at higher energies not yet probed experimentally.

³⁹⁴ Many theories exist as extensions to the SM, predicting a range of observables that can
³⁹⁵ be detected at the Large Hadron Collider (LHC) of which SUperSYmmetry (SUSY) is
³⁹⁶ one such example. It predicts a new symmetry of nature in which all current particles
³⁹⁷ in the SM would have a corresponding supersymmetric partner. Common to most
³⁹⁸ Supersymmetric theories is a stable, weakly interacting Lightest Supersymmetric Partner
³⁹⁹ (LSP), which has the properties of a possible dark matter candidate. The SM and the
⁴⁰⁰ main principles of Supersymmetric theories are outlined in Chapter 2, with emphasis
⁴⁰¹ placed on how experimental signatures of SUSY may reveal themselves in proton collisions
⁴⁰² at the LHC.

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

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

415 The performance of the CMS Level-1 single jet and energy sum triggers is also bench-
416 marked within this chapter. The Level-1 trigger is the first line of the CMS trigger
417 system and is of paramount importance to the collection of physics events. A change in
418 the jet clustering algorithm, via the introduction of a jet seed threshold, was introduced
419 approximately half way through the data taking period. The aim of this change, was
420 to reduce the rate at which collisions not of interest to physics analysis were recorded,
421 whilst avoiding impact to the overall performance of the triggers.

422 Chapter 4, contains a description of the search for direct evidence of the production of
423 supersymmetric particles at the LHC. The main basis of the search centres around the
424 kinematic dimensionless α_T variable; which provides a strong rejection of backgrounds
425 with fake missing transverse energy signatures, whilst maintaining good sensitivity to
426 a variety of SUSY topologies. The author’s work (as an integral part of the analysis
427 group) is documented in detail, and has culminated in numerous publications over the
428 past two years, the latest results having been published in the European Physical Journal
429 C (EPJC) [5].

430 The author in particular has played a major role in the extension of the α_T analysis into
431 additional b-tagged jet (jets identified as originating from a b-quark) and jet multiplicity
432 dimensions, increasing the sensitivity of the analysis to a range of SUSY topologies.
433 Additionally, the author has worked extensively on increasing the statistical precision
434 of the data driven electroweak predictions through analytical techniques. This included
435 work on developing the derivation of data driven systematic uncertainties through the
436 establishment of closure tests within the control samples of the analysis.

- ⁴³⁷ The compatibility of the data collected for the α_T search with a SM only hypothesis is
⁴³⁸ documented in Chapter 5. In the absence of an observed excess, interpretations of the
⁴³⁹ data within the framework of a variety of Simplified Model Spectra (SMS), describing an
⁴⁴⁰ array of possible SUSY event topologies are made.
- ⁴⁴¹ Finally, a method to search for gluino mediated SUSY signatures rich in top and bottom
⁴⁴² flavoured jet final states, is introduced in Chapter 6. These particular SUSY topologies
⁴⁴³ are increasingly of interest to physicists in light of the discovery of the Higgs boson. A
⁴⁴⁴ parametrisation of the b-tagged jet distribution for different electroweak processes is
⁴⁴⁵ used to establish template shapes, which are then fitted at low b-tagged jet multiplicity,
⁴⁴⁶ to extrapolate an expected SM background of 3 and 4 b-tagged jet events within an
⁴⁴⁷ event sample. The α_T control and hadronic signal event selections are used to validate
⁴⁴⁸ the functionality of this template method in both data and simulation. Background
⁴⁴⁹ predictions within the hadronic signal region are compared to those presented in Chapter 5,
⁴⁵⁰ with the intention of serving as a independent crosscheck of the estimated SM backgrounds
⁴⁵¹ from the α_T search.

⁴⁵² Natural units are used throughout this thesis in which $\hbar = c = 1$.

Chapter 2.

⁴⁵³ A Theoretical Overview

⁴⁵⁴ Within this chapter, a brief introduction and background to the SM is given. Its success
⁴⁵⁵ as a rigorously tested and widely accepted theory is discussed as are its deficiencies which
⁴⁵⁶ lead to the argument that this theory is not a complete description of our universe. The
⁴⁵⁷ motivations for new physics at the TeV scale and in particular Supersymmetric theories
⁴⁵⁸ are outlined within Section (2.3). The chapter concludes with how an experimental
⁴⁵⁹ signature of such theories can be produced and observed at the LHC in Section (2.4).

⁴⁶⁰ 2.1. The Standard Model

⁴⁶¹ The SM is the name given to the relativistic Quantum Field Theory (QFT), where
⁴⁶² particles are represented as excitations of fields, which describe the interactions and
⁴⁶³ properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field
⁴⁶⁴ theory which contains three symmetries: $SU(3)$ for colour charge; $SU(2)$ for weak isospin
⁴⁶⁵ and; $U(1)$ relating to weak hyper charge, which requires its Lagrangian \mathcal{L}_{SM} to be
⁴⁶⁶ invariant under local gauge transformation.

⁴⁶⁷ Within the SM theory, matter is composed of spin $\frac{1}{2}$ fermions that interact with each
⁴⁶⁸ other via the exchange of spin-1 gauge bosons. A summary of the known fundamental
⁴⁶⁹ fermions and bosons is given in Table 2.1.

⁴⁷⁰ Fermions are separated into quarks and leptons of which only quarks interact with the
⁴⁷¹ strong nuclear force. Quarks unlike leptons are not seen as free particles in nature, but
⁴⁷² rather exist only within baryons, which are composed of three quarks with an overall
⁴⁷³ integer charge, and quark-anti-quark pairs called mesons. Both leptons and quarks are

| 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].

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

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

The unification of the electromagnetic and weak-nuclear forces into the current Electroweak theory yield the weak gauge bosons W^\pm and Z through the mixing of the associated gauge fields. The force carriers of this theory were experimentally detected by the observation of the weak neutral current. This was first discovered in 1973 by the Gargamelle bubble chamber located at the European Organisation for Nuclear Research (CERN) [10]. The masses of the weak gauge bosons were measured by the UA1 and U2 experiments at the Super Proton Synchrotron (SPS) collider in 1983 [11][12].

⁴⁸⁷ 2.1.1. Gauge Symmetries of the SM

- ⁴⁸⁸ Symmetries are of fundamental importance in the description of physical phenomena.
⁴⁸⁹ Noether's theorem states that for a dynamical system, the consequence of any symmetry
⁴⁹⁰ is an associated conserved quantity [13]. Invariance under translations, rotations, and
⁴⁹¹ Lorentz transformations in physical systems lead to the conservation of momentum,
⁴⁹² energy and angular momentum.
- ⁴⁹³ In the SM, a quantum theory described by Lagrangian formalism, the weak, strong and
⁴⁹⁴ electromagnetic interactions are described in terms of “gauge theories”. A gauge theory
⁴⁹⁵ possesses invariance under a set of “local transformations”, which are transformations
⁴⁹⁶ whose parameters are space-time dependent. The requirement of gauge invariance within
⁴⁹⁷ the SM necessitates the introduction of force-mediating gauge bosons, and interactions
⁴⁹⁸ between fermions and the bosons themselves. Given the nature of the topics covered by
⁴⁹⁹ this thesis, the formulation of Electroweak Sector (EWK) within the SM Lagrangian is
⁵⁰⁰ reviewed within this section.
- ⁵⁰¹ The simplest example of the application of the principle of local gauge invariance within
⁵⁰² the SM is in Quantum Electro-Dynamics (QED), the consequences of which require a
⁵⁰³ massless photon field [14][15].
- ⁵⁰⁴ The free Dirac Lagrangian can be first written as

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

- ⁵⁰⁵ where ψ represents a free non interacting fermionic field, with the matrices γ^μ , $\mu \in 0, 1, 2, 3$
⁵⁰⁶ defined by the anti commutator relationship $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} I_4$, with $\eta^{\mu\nu}$ being the
⁵⁰⁷ flat space-time metric $(+, -, -, -)$, and I_4 the 4×4 identity matrix.
- ⁵⁰⁸ Under a local U(1) abelian gauge transformation, in which ψ transforms as

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

- ⁵⁰⁹ the kinetic term of the Lagrangian will not remain invariant, due to the partial derivative
⁵¹⁰ interposed between the $\bar{\psi}$ and ψ yielding

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

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

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

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

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

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

This then leads to the Lagrangian of QED,

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

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

525 2.1.2. The Electroweak Sector and Electroweak Symmetry 526 Breaking

- 527 The same application of gauge symmetry and the requirement of local gauge invariance
528 can be used to unify QED and the Weak force in the Electroweak Sector (EWK).
529 The nature of EWK interactions is encompassed within a Lagrangian invariant under
530 transformations of the group $SU(2)_L \times U(1)_Y$.
531 The weak interactions from experimental observation [16] are known to violate parity
532 and are therefore not symmetric under interchange of left and right helicity fermions.
533 Thus, within the SM the left- and right-handed parts of these fermion fields are treated
534 separately. A fermion field is then split into two left- and right-handed chiral components,
535 $\psi = \psi_L + \psi_R$, where $\psi_{L/R} = (1 \pm \gamma^5)\psi$.
536 The $SU(2)_L$ group is the special unitary group of 2×2 matrices, U , satisfying $UU^\dagger = I$
537 and $\det(U) = 1$. It may be written in the form $U = e^{-i\omega_i T_i}$, with the generators of the
538 group written as $T_i = \frac{1}{2}\tau_i$ where τ_i , $i \in 1,2,3$ are the 2×2 Pauli matrices

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

- 539 The generators of the group form a non Abelian group obeying the commutation relation
540 $[T^a, T^b] \equiv if^{abc}T^c \neq 0$. The gauge fields that accompany this group are represented by
541 $\hat{W}_\mu = (\hat{W}_\mu^1, \hat{W}_\mu^2, \hat{W}_\mu^3)$ and act only on the left handed component of the fermion field ψ_L .
542 One additional generator, Y , which represents the hypercharge of the particle under
543 consideration is introduced through the $U(1)_Y$ group acting on both components of the
544 fermion field, with an associated vector boson field \hat{B}_μ .
545 The $SU(2)_L \times U(1)_Y$ transformations of the left- and right-handed components of ψ are
546 summarised by,

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

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

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

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

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

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

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

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

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

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

556 corresponding to the kinetic energy and self coupling of the W_μ fields and the kinetic
 557 energy term of the B_μ field.

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

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

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

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

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

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

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

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

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

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

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

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

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

579 where the fluctuations from the vacuum ϕ_0 are parametrized in terms of four real fields,
580 $\theta_1, \theta_2, \theta_3$ and $h(x)$.

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

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

584 where the relations between the physical and electroweak gauge fields from Equation
585 (2.14) are used. The W^\pm and Z boson masses can then be determined to be

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

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

592 **2.2. Motivation for Physics beyond the Standard
593 Model**

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

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

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

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

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

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

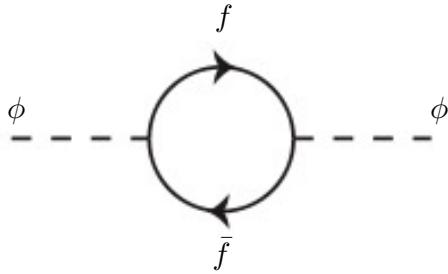


Figure 2.1: One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion.

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

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

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

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

632 2.3. Supersymmetry Overview

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

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

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

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

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

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

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

⁶⁶⁴ 2.3.1. R-Parity

⁶⁶⁵ Supersymmetric theories can also present a solution to the dark matter problem. These
⁶⁶⁶ theories contain a stable Lightest Supersymmetric Partner (LSP), which match the
⁶⁶⁷ criteria of a Weakly Interacting Massive Particle (WIMP) required by cosmological
⁶⁶⁸ observation when R-parity is conserved.

⁶⁶⁹ Baryon (B) and Lepton (L) number conservation is forbidden in the SM by renormal-
⁶⁷⁰ isability requirements. The violation of Baryon or Lepton number results in a proton
⁶⁷¹ lifetime much shorter than those set by experimental limits [37]. Another symmetry
⁶⁷² called R-parity is then often introduced to SUSY theories to maintain baryon and lepton
⁶⁷³ conservation.

⁶⁷⁴ R-parity is described by the equation

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

⁶⁷⁵ where s represents the spin of the particles. $B = \pm \frac{1}{3}$ for quarks/antiquarks and $B = 0$
⁶⁷⁶ for all others, $L = \pm 1$ for leptons/antileptons, $L = 0$ for all others.

⁶⁷⁷ R-parity ensures the stability of the proton in SUSY models, and also has other conse-
⁶⁷⁸ quences for the production and decay of supersymmetric particles. In particle colliders
⁶⁷⁹ supersymmetric particles can then only be pair produced. Similarly the decay of any
⁶⁸⁰ produced supersymmetric particle is restricted to a SM particle and a lighter supersym-
⁶⁸¹ metric particle, as allowed by conservation laws. A further implication of R-parity is
⁶⁸² that once a supersymmetric particle has decayed to the LSP it remains stable, unable to
⁶⁸³ decay into a SM particle.

⁶⁸⁴ A LSP will not interact in a detector at a particle collider, leaving behind a missing
⁶⁸⁵ energy, \cancel{E}_T , signature. The assumption of R-parity and its consequences are used to
⁶⁸⁶ determine the physical motivation and search strategies for SUSY at the LHC.

⁶⁸⁷ 2.4. Experimental Signatures of SUSY at the LHC

⁶⁸⁸ Should strongly interacting sparticles be within the experimental reach of the LHC, then
⁶⁸⁹ it is expected that they can be produced in a variety of ways:

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

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

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

705 However, recent results from the LHC now strongly disfavour large swathes of CMSSM
706 parameter space [39][40][41]. In the face of such results a more pragmatic model indepen-
707 dent search strategy is now applied across most SUSY searches at the LHC, see Section
708 (2.4.1).

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

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

722 2.4.1. Simplified Models

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

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

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

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

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

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

⁷⁵⁵ supersymmetric theories by requiring a light right-handed top squark, \tilde{t}_R , and left-handed
⁷⁵⁶ double $SU(2)_L$ doublet containing top and bottom squarks, $(\tilde{t}, \tilde{b})_L$ [47].

⁷⁵⁷ These theories therefore solve the hierarchy problem by predicting light \sim EWK scale third
⁷⁵⁸ generation sleptons, accessible at the LHC. Search strategies involving the requirement
⁷⁵⁹ of b-tagging (see Section (3.3.2)) are used to give sensitivity to these type of SUSY
⁷⁶⁰ scenarios and are discussed in greater detail within Chapter 4.

⁷⁶¹ Two example decay chains are shown in Figure 2.2; the pair production of gluinos (T1)
⁷⁶² and the pair production of squarks (T2) decaying into SM particles and LSPs.

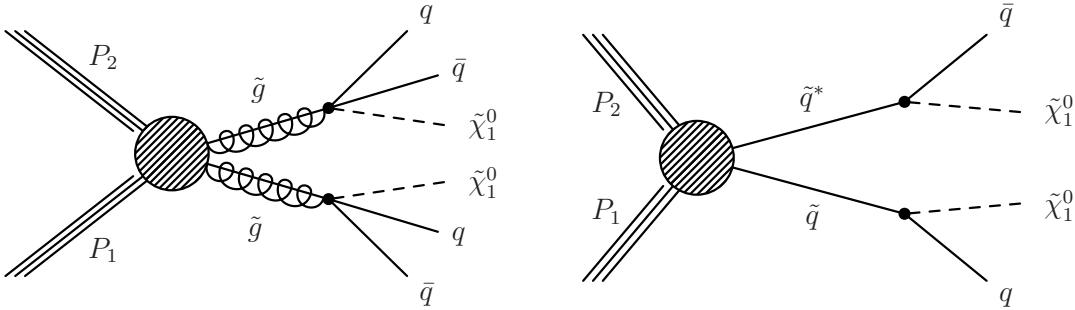


Figure 2.2: Two example SMS model decays (T1 (left), T2 (right)), which are used in interpretations of physics reach by CMS.

Chapter 3.

⁷⁶³ The LHC and the CMS Detector

⁷⁶⁴ Probing the SM for signs of new physics would not be possible without the immensely
⁷⁶⁵ complex electronics and machinery that has made the TeV energy scale accessible to
⁷⁶⁶ physicists for the first time. This chapter will introduce both the LHC based at European
⁷⁶⁷ Organisation for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS)
⁷⁶⁸ detector (of which the author is a member). Section (3.2) serves to present an overview of
⁷⁶⁹ the different components of the CMS detector, with specific components relevant to the
⁷⁷⁰ search for supersymmetric particles described in greater detail. Section (3.3) will focus
⁷⁷¹ on event and object reconstruction, again, with more emphasis on jet level quantities
⁷⁷² which are most relevant to the author's analysis research. Finally, Section (3.4) will
⁷⁷³ describe and detail the service work for the CMS Collaboration performed by the author,
⁷⁷⁴ in measuring the performance of L1 single jet and energy sum triggers in the Global
⁷⁷⁵ Calorimeter Trigger (GCT) during the 2012-2013 run period.

⁷⁷⁶ 3.1. The LHC

⁷⁷⁷ The LHC is a storage ring, accelerator, and collider of circulating beams of protons
⁷⁷⁸ or ions. Housed in the tunnel dug for the Large Electron-Positron Collidior (LEP), it
⁷⁷⁹ is approximately 27km in circumference, 100m underground, and straddles the border
⁷⁸⁰ between France and Switzerland, outside of Geneva. It is currently the only collider
⁷⁸¹ in operation that is able to study physics at the TeV scale. A double-ring circular
⁷⁸² synchrotron, it was designed to collide both proton-proton (pp) and heavy ion (PbPb)
⁷⁸³ with a centre of mass energy $\sqrt{s} = 14$ TeV at a final design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$.

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

796

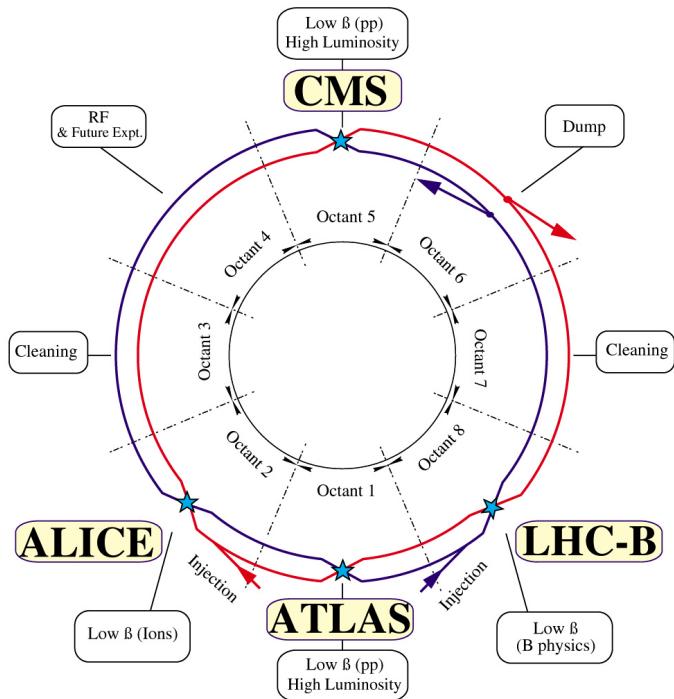


Figure 3.1: A top down layout of the LHC. [52], with the position of the four main detectors labelled.

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 localised into less than 1 ns in the direction of motion). Before collision, the beams are ramped to

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

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

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

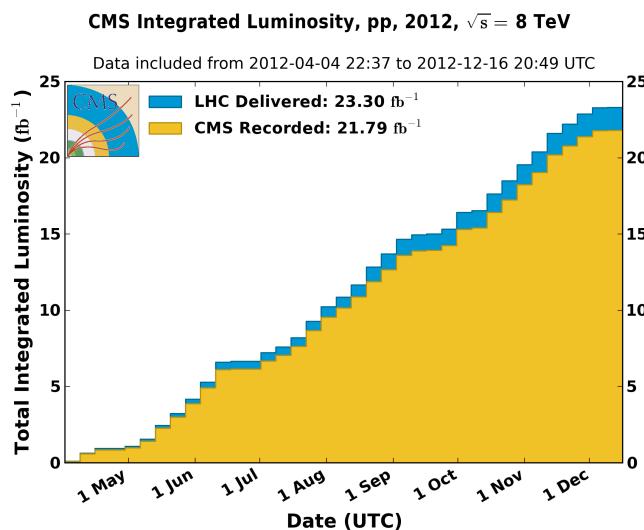


Figure 3.2: The total integrated luminosity delivered to and collected by CMS during the 2012 8 TeV pp runs.

⁸²¹ 3.2. The CMS Detector

⁸²² The Compact Muon Solenoid (CMS) detector is one of two general purpose detectors
⁸²³ at the LHC designed to search for new physics. The detector is designed to provide
⁸²⁴ efficient identification and measurement of many physics objects including photons,
⁸²⁵ electrons, muons, taus, and hadronic showers over wide ranges of transverse momentum
⁸²⁶ and direction. Its nearly 4π coverage in solid angle allows for accurate measurement of
⁸²⁷ global transverse momentum imbalance. These design factors give CMS the ability to
⁸²⁸ search for direct production of SUSY particles at the TeV scale, making the search for
⁸²⁹ Supersymmetric particles one of the highest priorities among the wide range of physics
⁸³⁰ programmes at CMS.

⁸³¹ CMS uses a right-handed Cartesian coordinate system with the origin at the interaction
⁸³² point and the z-axis pointing along the beam axis. The x-axis points radially inwards to
⁸³³ the centre of the collider ring, with the y-axis pointing vertically upward. The azimuthal
⁸³⁴ angle ϕ , ranging between $[-\pi, \pi]$, is defined in the x-y plane starting from the x-axis. The
⁸³⁵ polar angle θ is measured from the z axis. The common convention in particle physics is
⁸³⁶ to express an out-going particle in terms of ϕ and its pseudorapidity defined as

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

⁸³⁷ The variable $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is commonly used to define angular distance between
⁸³⁸ objects within the detector. Additionally, energy and momentum is typically measured in
⁸³⁹ the transverse plane perpendicular to the beam line. These values are calculated from the
⁸⁴⁰ x and y components of the object and are denoted as $E_T = E \sin \theta$ and $p_T = \sqrt{p_x^2 + p_y^2}$.

⁸⁴¹ 3.2.1. Detector Subsystems

⁸⁴² As the range of particles produced from pp collisions interact in different ways with
⁸⁴³ matter, CMS is divided into sub-detector systems, which perform complementary roles
⁸⁴⁴ to identify the identity, the mass, and the momentum of different physics objects present
⁸⁴⁵ in each event. These detector sub-systems contained within CMS are wrapped in layers
⁸⁴⁶ around a central 13m long 4 T super conducting solenoid, as shown in Figure 3.3. With
⁸⁴⁷ the endcaps closed, CMS is a cylinder of length 22m, diameter 15m, and mass 12.5

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

849

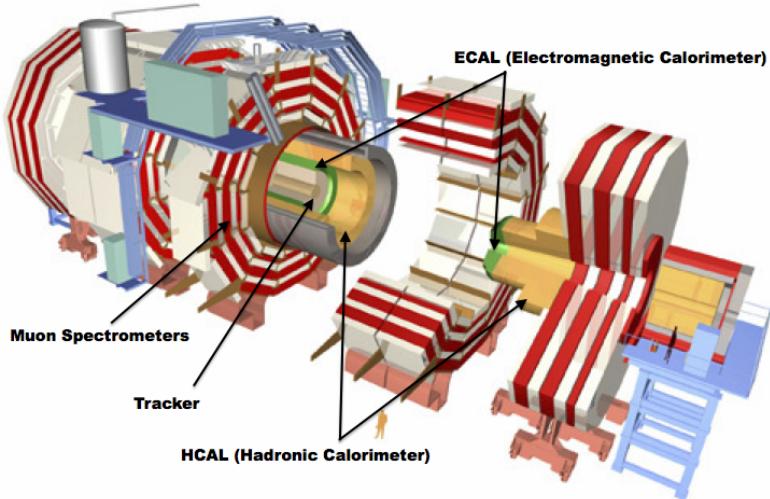


Figure 3.3: A pictorial depiction of the CMS detector with the main detector subsystems labelled [54].

850 3.2.2. Tracker

851 The inner-most sub-detector of the barrel is the multi-layer silicon tracker, formed of a
 852 pixel detector component encased by layers of silicon strip detectors. The pixel detector
 853 consists of three layers of silicon pixel sensors providing measurements of the momentum,
 854 position coordinates of the charged particles as they pass, and the location of primary
 855 and secondary vertices between 4 cm and 10 cm transverse to the beam. Outside the
 856 pixel detector, ten cylindrical layers of silicon strip detectors extend the tracking system
 857 out to a radius of 1.20m from the beam line. The tracking system provides efficient
 858 and precise determination of the charges, momenta, and impact parameters of charged
 859 particles, with the geometry of the tracker extending to cover a rapidity range up to $|\eta| <$
 860 2.5.

861

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

⁸⁶⁷ **3.2.3. Electromagnetic Calorimeter**

⁸⁶⁸ Immediately outside of the tracker, but still within the magnet core, sits the Electromagnetic CALorimeter (ECAL). Covering a pseudorapidity up to $|\eta| < 3$ and comprising
⁸⁶⁹ of over 75×10^3 PbWO₄ (lead tungstate) crystals that scintillate as particles deposit
⁸⁷⁰ energy, the ECAL provides high resolution measurements of the electromagnetic showers
⁸⁷¹ from photons and electrons in the detector.

⁸⁷³

⁸⁷⁴ Lead tungstate is used because of its short radiation length ($X_0 \sim 0.9$ cm) and small
⁸⁷⁵ Molieré radius (~ 2.1 cm) leading to high granularity and resolution. Its fast scintillation
⁸⁷⁶ time (~ 25 ns) reduces the effects of pile-up, which occurs when energy from previous
⁸⁷⁷ collisions are still being read out, and its radiation hardness gives it longevity. The
⁸⁷⁸ crystals are arranged in modules which surround the beam line in a non-projective
⁸⁷⁹ geometry, angled at 3° , with respect to the interaction point to minimise the risk of
⁸⁸⁰ particles escaping down the cracks between the crystals.

⁸⁸¹

⁸⁸² The ECAL is primarily composed of two sections, the Electromagnetic CALorimeter
⁸⁸³ Barrel (EB) which extends in pseudo-rapidity to $|\eta| < 1.479$ with a crystal front cross sec-
⁸⁸⁴ tion of 22×22 mm and a length of 230 mm corresponding to 25.8 radiation lengths. The
⁸⁸⁵ Electromagnetic CALorimeter Endcap (EE) covers a rapidity range of $1.479 < |\eta| < 3.0$,
⁸⁸⁶ which consists of two identical detectors on either side of the EB. A lead-silicon sampling
⁸⁸⁷ ‘pre-shower’ detector Electromagnetic CALorimeter pre-Shower (ES) is placed before
⁸⁸⁸ the endcaps to aid in the identification of neutral pions. Their arrangement is shown in
⁸⁸⁹ Figure 3.4.

⁸⁹⁰

⁸⁹¹ Scintillation photons from the lead tungstate crystals are instrumented with Avalanche
⁸⁹² Photo-Diodes (APD) and Vacuum Photo-Triodes (VPT), located in the EB and EE
⁸⁹³ respectively. They convert the scintillating light into an electric signal which is conse-
⁸⁹⁴ quently used to determine the amount of energy deposited within the crystal. These
⁸⁹⁵ instruments are chosen for their resistance under operation to the strong magnetic field
⁸⁹⁶ of CMS. The scintillation of the ECAL crystals, as well as the response of the APDs,
⁸⁹⁷ vary as a function of temperature; and so cooling systems continually maintain an overall
⁸⁹⁸ constant ECAL temperature $\pm 0.05^\circ C$.

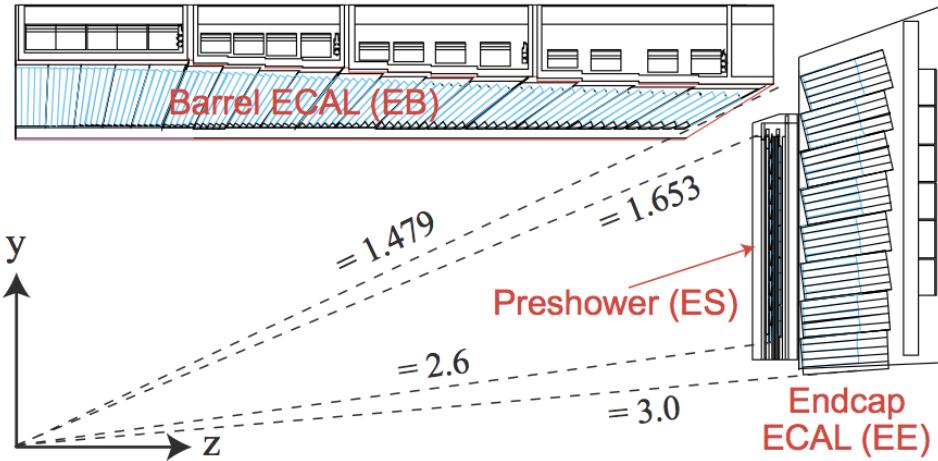


Figure 3.4: Illustration of the CMS ECAL showing the arrangement of the lead tungstate crystals in the EB and EE. The ES is also shown and is located in front of the EE [55].

3.2.4. Hadronic Calorimeter

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

908

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

913

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

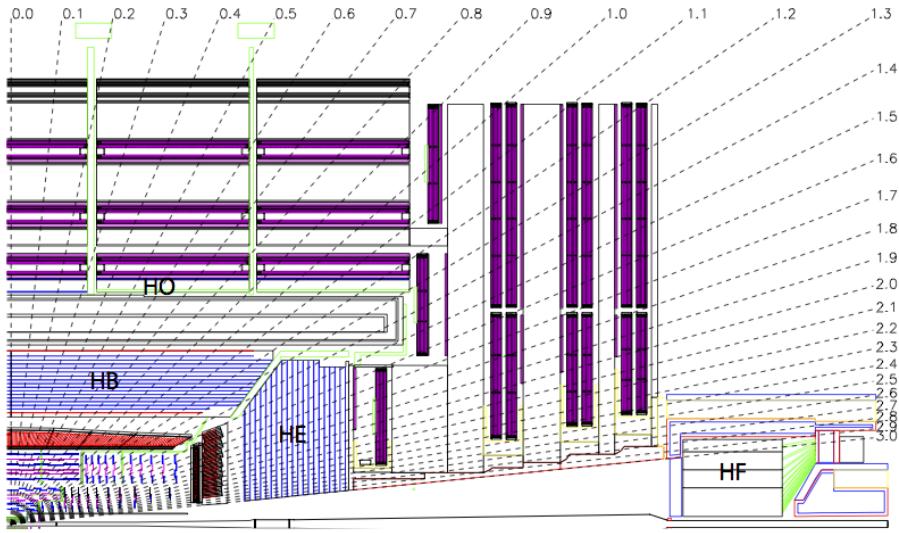


Figure 3.5: Schematic of the hadron calorimeters in the r-z plane, showing the locations of the HCAL components and the HF. [50].

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

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

930 3.2.5. Muon Systems

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

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

941

942 **3.3. Event Reconstruction and Object Definition**

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

950 **3.3.1. Jets**

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

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

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

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

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

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

991 3.3.2. B-tagging

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

996

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

1000 One of the most efficient of which is the Combined Secondary Vertex (CSV) algorithm
 1001 [63]. This operates based on secondary vertex and track-based lifetime information,
 1002 benchmarked in ‘Loose’, ‘Medium’ and ‘Tight’ working points, of which the medium
 1003 point is the tagger used within the α_T search presented in Section (4.1). All figures
 1004 within this sub-section, demonstrating the performance of this b-tagging algorithm are
 1005 taken from [64].

1006 Within the CSV tagger, a likelihood-based discriminator distinguishes between jets from
 1007 b-quarks, and those from charm or light quarks and gluons, shown in Figure 3.6. The
 1008 minimum thresholds on the discriminator for each working point correspond to the
 1009 mis-identification probability for light-parton jets of 10%, 1%, and 0.1%, respectively, in
 1010 jets with an average p_T of about 80 GeV.

1011 The b-tagging performance is evaluated to measure the b-jet tagging efficiency ϵ_b , and the
 1012 misidentification probability of charm ϵ_c and light-parton jets ϵ_s . The tagging efficiencies
 1013 for each of these three jet flavours are compared between data and MC simulation, from
 1014 which a series of p_T and $|\eta|$ dependant jet corrections are determined,

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

1015 These are collectively named ‘Btag Scale Factors’ and allow MC simulation to accurately
 1016 reflect the running conditions and performance of the tagging algorithm in data. A good
 1017 understanding of the tagging efficiency for each of the jet flavours is essential in order to

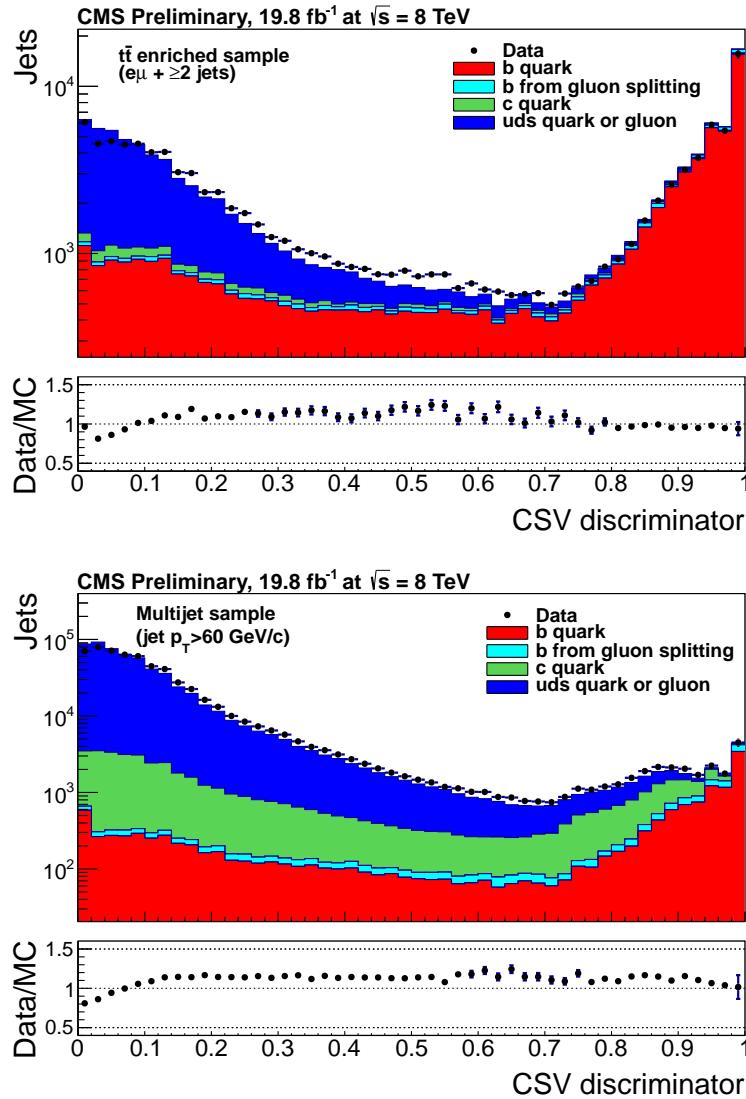


Figure 3.6: CSV algorithm discriminator values in enriched $t\bar{t}$ (top) and inclusive multi-jet samples (bottom) for b, c and light flavoured jets. The discriminator value used for each working points are determined from the misidentification probability for light-parton jets to be tagged as a b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

1018 minimise systematic uncertainties in physics analyses that employ b -tagging.

1019

1020 The b -tagging efficiency is measured in data using several methods applied to multi-jet
 1021 events, primarily based on a sample of jets enriched in heavy flavour content. One method
 1022 requires the collection of events with a poorly isolated muon within a cone $\Delta R < 0.4$
 1023 around the jet axis. Due to the semi-leptonic branching fraction of b hadrons being

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

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

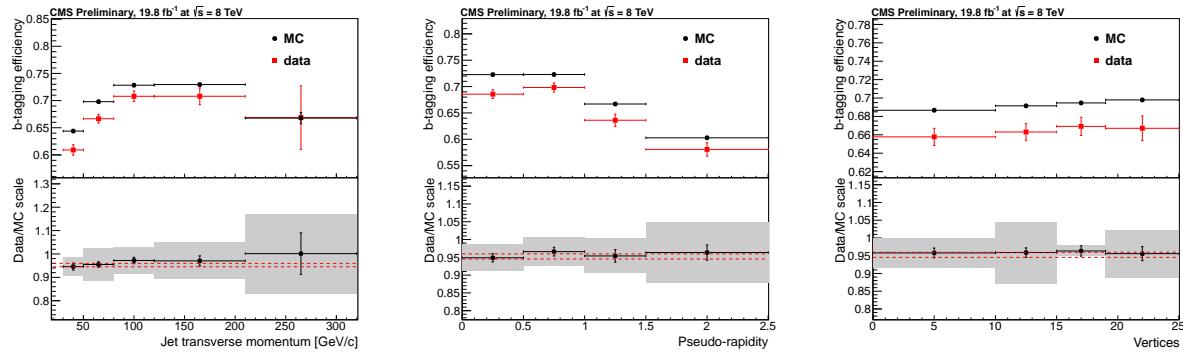


Figure 3.7: Measured in $t\bar{t} \rightarrow$ di-lepton events using the CSV tagger: (upper panels) b-tagging efficiencies and (lower panels) data/MC scale factor SF_b as a function of (left) jet p_T , (middle) jet $|\eta|$ and (right) number of primary vertices. In the lower panels, the grey filled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

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

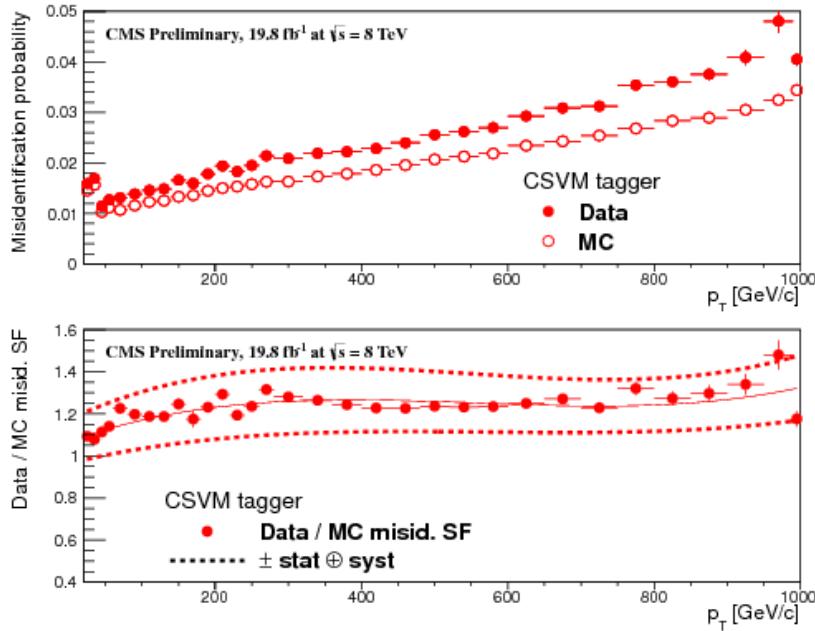


Figure 3.8: For the CSVM tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last p_T bin in each plot includes all jets with $p_T > 1000$ GeV. The solid curve is the result of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

1040 3.4. Triggering System

1041 With bunch crossings separated by just 50 ns, the rate at which data from all collisions
 1042 would have to be written out and processed would be unfeasible. A two-tiered triggering
 1043 system is applied at CMS in order to cope with the high collision rate of protons. The
 1044 CMS trigger is designed to use limited information from each event to determine whether
 1045 to record the event, reducing the rate of data taking to manageable levels whilst ensuring
 1046 a high efficiency of interesting physics object events are selected.

1047 The Level 1 Trigger (L1) is a pipelined, dead-timeless system based on custom-built
 1048 electronics [66], and is a combination of several sub systems which is shown pictorially
 1049 in Figure 3.9. The L1 system is covered in more detail within the following section,
 1050 along with a description of the service work undertaken by the author to benchmark the
 1051 performance of the L1 calorimeter trigger during the 2012 8 TeV run period.

1052 The Higher Level Trigger (HLT) is a large farm of commercial computers [67]. The HLT
 1053 processes events with software reconstruction algorithms that are more detailed, giving
 1054 performance more similar to the reconstruction used offline. The HLT reduces the event

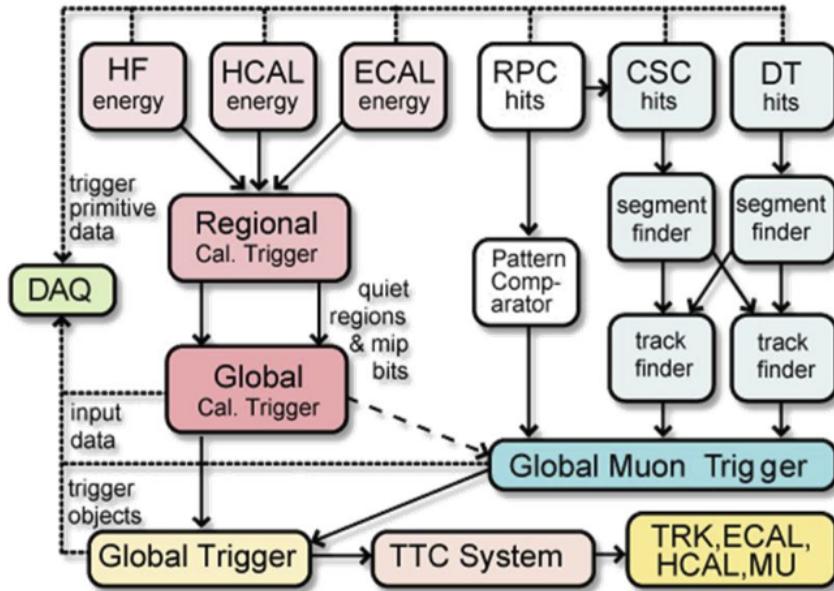


Figure 3.9: An overview of the different components of the CMS L1 trigger system, showing the global calorimeter, muon triggers, and the global trigger.

1055 rate written to disk by a factor of ~ 500 ($\sim 200\text{Hz}$). The recorded events are transferred
 1056 from CMS to the CERN computing centre, where event reconstruction is performed, and
 1057 then distributed to CMS computing sites around the globe for storage and analysis.

1058 3.4.1. The Level-1 Trigger

1059 The L1 trigger reduces the rate of events collected from 20 MHz to $\sim 100\text{ kHz}$ using
 1060 information from just the calorimeters and muon chambers, but not the tracker. This
 1061 is due to requirement that data from each and every bunch crossing be analysed with
 1062 no dead time, drastically reducing time available to process and reconstruct objects in
 1063 making a trigger decision. This facilitates the need for a pipelined processing architecture,
 1064 and so a tree system of triggers is used to decide whether to pass on an event to the HLT
 1065 for further reconstruction.

1066 Calorimeter and muon event information is processed separately by the Regional Calorime-
 1067 ter Trigger (RCT) and Regional Muon Trigger (RMT) systems respectively.

1068 Within the RCT, energy deposits from trigger towers in the ECAL and HCAL calorimeters
1069 are summed into coarser calorimeter regions and sent to the Global Calorimeter Trigger
1070 (GCT) for jet clustering.

1071 Given that electron and photon are much narrower objects than jets, the RCT is used
1072 to identify these candidates but makes no attempt to distinguish between them at this
1073 stage given the lack of tracking information. They are first identified by ensuring the
1074 energy deposits within the central trigger tower and its surrounding cells are above a
1075 certain programmable threshold. To ensure the object is not a hadron, the ratio of HCAL
1076 to ECAL in the central tower is calculated and checked to be below 5%. Additional
1077 algorithms are employed to ascertain whether the e/γ object is isolated/non-isolated.

1078 In the L1 GCT, coarse measurements of the energy deposited in the electromagnetic and
1079 hadronic calorimeters are combined, and by using sophisticated algorithms the following
1080 tasks are performed:

- 1081 • isolated and non-isolated electromagnetic objects are sorted (e and γ), with the four
1082 highest ranked (equivalent to highest transverse energy E_T) objects of each type
1083 passed onto the Global Trigger (GT),
- 1084 • energy sums from the calorimeters supplied by the RCT are used in performing
1085 jet clustering (described in the following section). The clustered jets are then
1086 sub-divided into categories depending on their pseudo-rapidity and the result of
1087 τ identification, being classified as either central, forward, or tau (τ). After being
1088 sorted by rank, the four highest of each category are passed to the GT for use in
1089 trigger decisions,
- 1090 • total transverse energy (E_T), the scalar sum of the energy deposits measured by
1091 L1, and missing transverse energy (\cancel{E}_T), defined as the negative vector sum of the
1092 transverse energy deposits measured at L1 are calculated,
- 1093 • total transverse jet energy (H_T), the scalar sum of the energy of all L1 clustered jet
1094 objects, and missing transverse jet energy (\cancel{H}_T), defined as the negative vector sum
1095 of the energy from L1 clustered jet objects are calculated and passed to the GT.

1096 In addition, quantities suitable for triggering minimum bias events, forward physics and
1097 beam background events are determined. Relevant muon isolation information is also
1098 passed on to the Global MuonTrigger (GMT) to be used in decisions involving the muon
1099 triggers, where it is combined with information from across the three muon sub-systems.

1100 The resultant final accept/reject decision at L1 is then performed by the GT, based on
1101 the objects received from the GCT and GMT ($e/\gamma, \mu$, jets, $E_T, \cancel{E}_T, H_T, \cancel{H}_T$).
1102 The L1 trigger is therefore of upmost importance to the functioning of the detector.
1103 Without a high-performing, efficient trigger and a good understanding of its performance
1104 at ever increasing instantaneous luminosities, the data collected would be useless. Whilst
1105 it would be possible to maintain trigger efficiency by increasing the triggering thresholds
1106 for different jet or energy sum quantities, this is far from ideal. This could result in the
1107 failure to be sensitive to a wide range of new physics signatures, including many types of
1108 compressed spectra SUSY models where the mass splitting between squarks/gluinos and
1109 the LSP is small.
1110 One such method introduced to help maintain low triggering thresholds, was via the
1111 introduction of a jet seed threshold into the L1 jet clustering algorithm. Observations of
1112 how the L1 trigger performance is affected by both the jet seed threshold, and changing
1113 LHC running conditions over the 2012 run period is presented in the following Sections
1114 (3.4.2 - 3.4.6).

1115 3.4.2. The L1 Trigger Jet Algorithm

1116 The L1 jet algorithm clusters jets using the transverse energy sums computed by the
1117 calorimeter trigger regions. Each region consists of a 4×4 trigger tower window which
1118 within the CMS barrel spans a region of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in pseudorapidity-
1119 azimuth.
1120 A L1 jet is defined by a 3×3 window of calorimeter regions, as shown in Figure 3.10.
1121 This corresponds to 12×12 trigger towers in barrel and endcap, or 3×3 larger HF towers
1122 in the HF. The ϕ size of the jet window is the same everywhere, whilst the η binning
1123 gets somewhat larger at high η due to calorimeter and trigger tower segmentation. The
1124 jets are labelled by the (η, ϕ) indices of the central calorimeter region.
1125 A jet candidate is identified if the sum of the transverse ECAL and HCAL energies of a
1126 calorimeter region is larger than all of its 8 neighbouring regions $E_{T \text{ central}} > E_{T \text{ surround}}$.
1127 The jet is then centred at this central calorimeter region.
1128 During the 2012 run period, a minimum threshold of 5 GeV was imposed on the central
1129 seeding region to suppress noise from non-collimated pile-up jets. This threshold is
1130 applied on the raw, uncorrected energy of the calorimeter regions and affects all clustered

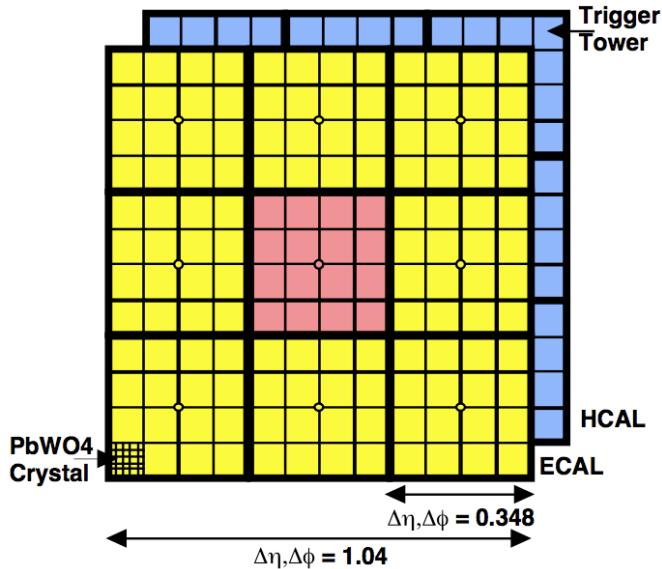


Figure 3.10: Illustration of the dimensions of the Level-1 jet finder window. Each cell represents a trigger tower, which is the sum of the transverse energy contributions from both calorimeter systems.

1131 L1 jets. The effect of such a change to the jet algorithm on the triggering performance
 1132 of L1 quantities is shown in Section (3.4.4).

1133 To form the jet candidates the GCT utilises a pre-clustering algorithm, which employs 18
 1134 jet-finders that operate simultaneously over the whole detector. Each jet-finder spans an
 1135 area of 11 calorimeter regions in η (half the detector) and two in ϕ (40°). Jets are initially
 1136 created in 2×3 mini-clusters in order to reduce the total amount of data duplicated and
 1137 shared between the jet-finders. Information is only shared with the two ϕ strips of the
 1138 neighbouring jet-finders when these mini-clusters jets are found, and is used to form a
 1139 clustered 3×3 L1 jet object. Additional information on this clustering algorithm can be
 1140 found at [68].

1141 Within the $|\eta| < 3$ region, the GCT also determines whether a jet is to be classified as
 1142 a τ or central jet. The hadronic decay modes of the τ typically contain one or three
 1143 isolated pions, thus leading to more collimated energy deposits with fewer constituents
 1144 than non- τ jets. Therefore, for a jet candidate to be classed as a τ jet, up to a maximum
 1145 of one of the eight calorimeter regions neighbouring the central jet seed is permitted to
 1146 have a transverse energy, E_T , above some programmable isolation threshold.

1147 Jets found between $3.0 < |\eta| < 5.0$ are classified as forward jets, whereas those with
 1148 $|\eta| < 3.0$ are classified as either a central or τ -jet. The four clustered jets with the highest

1149 transverse energy in each category (central, forward and τ -jet) are further passed through
1150 Look Up Table (LUT)s, which apply a programmable η -dependent jet energy scale
1151 correction. Finally these jet objects are passed to the GT to make L1 trigger decisions.

1152 The performance of L1 jets within the following sections are evaluated with respect to
1153 offline jets, which are taken from the standard Calo jet and the PF jet reconstruction
1154 algorithms of CMS. These reconstructed offline jets are corrected for pile-up and detector
1155 effects as described in Section (3.3.1). A moderate level of noise rejection is applied to
1156 the offline jets by selecting jets passing the “loose” identification criteria for both Calo
1157 and PF. These jet criteria are listed in Appendix (A.1).

1158 3.4.3. Measuring L1 Single-Jet Trigger Efficiencies

1159 The efficiency of a L1 single-jet trigger at an offline reconstructed jet E_T is defined as; the
1160 fraction of events in a sample containing at least a single reconstructed offline jet, where
1161 the leading offline jet is matched with a L1 central or τ jet that also has a measured L1
1162 energy above the trigger threshold being benchmarked.

1163 A match is determined by comparing the L1 and reconstructed offline jets spatially in
1164 $\eta - \phi$ space. The ΔR separation between the highest offline reconstructed jet ($E_T > 10$
1165 GeV and $|\eta| < 3$) and each L1 jet in the event is calculated. A match is made to the L1
1166 jet with the minimum ΔR to the reconstructed jet on the condition that it also satisfies
1167 $\Delta R < 0.5$.

1168 The matching efficiency for this procedure is found to be close to 100% above an offline
1169 jet threshold of 30(45) GeV for the run 2012B(C) data taking period (see Appendix B.1).

1170 Each efficiency curve is fitted with a function which is the cumulative distribution function
1171 of an Exponentially Modified Gaussian (EMG) distribution:

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

where erfc is the complementary error function defined as:

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

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

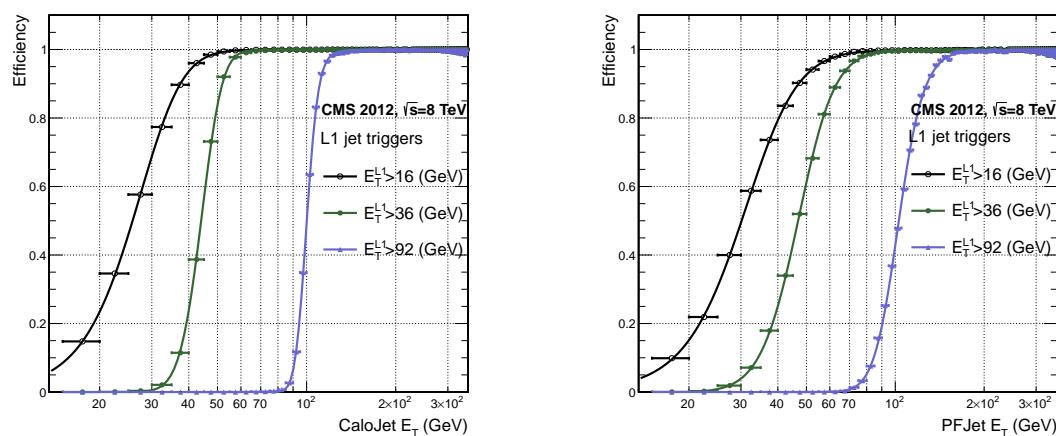


Figure 3.11: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T (left) and PFJet E_T (right), measured in 2012 Run Period C data and collected with an isolated single μ data sample.

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

| Trigger | Calo | | PF | |
|----------------|------------------|-----------------|------------------|------------------|
| | 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).

¹¹⁹⁴ 3.4.4. Effects of the L1 Jet Seed

¹¹⁹⁵ Between run period B and C of the 2012 data taking period, a jet seed threshold
¹¹⁹⁶ was introduced into the L1 jet clustering algorithm. There was previously no direct
¹¹⁹⁷ requirement made on the energy deposited in the central region.

¹¹⁹⁸ The introduction of a jet seed threshold required that the central region have an uncor-
¹¹⁹⁹ rected energy deposit of $E_T \geq 5$ GeV. This value was motivated by studies of the effect
¹²⁰⁰ that different jet seed thresholds had upon the trigger cross-sections and efficiencies of
¹²⁰¹ various H_T , single jet and multi-jet triggers. It was found that the 5 GeV threshold gave
¹²⁰² large reductions in trigger cross-sections particularly in the case of multi-jet and H_T
¹²⁰³ triggers, whilst having a small impact on the measured efficiencies of these triggers [70].

¹²⁰⁴ Its main purpose was to counteract the effects of high pile up running conditions which
¹²⁰⁵ create a large number of soft non-collimated jets, that are then added to the jets from
¹²⁰⁶ the primary interaction or other soft jets from other secondary interactions [71]. This in
¹²⁰⁷ turn causes a large increase in trigger rate, due to the increase in the likelihood that the
¹²⁰⁸ event causes the L1 trigger to fire.

¹²⁰⁹ The effect of the introduction of this jet seed threshold between these two run periods is
¹²¹⁰ benchmarked through a comparison of the efficiency of the L1 jet triggers with respect
¹²¹¹ to offline Calo jets and is shown in Figure 3.12.

¹²¹² The L1 H_T trigger efficiency is also benchmarked at two values, which is shown in Figure
¹²¹³ 3.13. The L1 H_T sum is compared against the offline H_T constructed from Calo jets
¹²¹⁴ with $E_T \geq 40$ GeV. This requirement is imposed to account for the relative difference
¹²¹⁵ between uncorrected jet energy deposits within the GCT used to calculate the L1 H_T
¹²¹⁶ sum, and those same deposits after full object reconstruction has occurred.

1217 To negate any effects from different pile-up conditions in the run periods, the efficiencies
 1218 are measured in events which contain between 15 and 20 primary vertices, as defined in
 1219 Appendix (A.2).

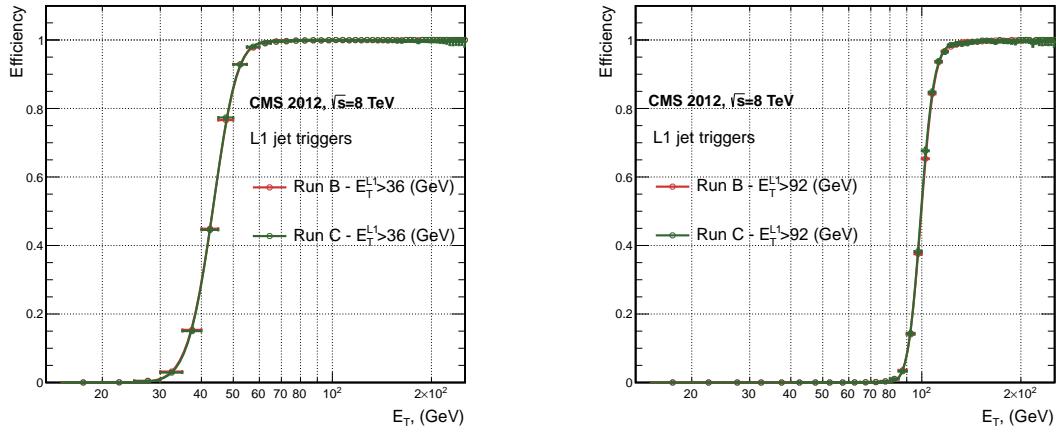


Figure 3.12: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T , measured for the L1 SingleJet 36 and 92 trigger in 2012 run period B and C collected with an isolated single μ sample.

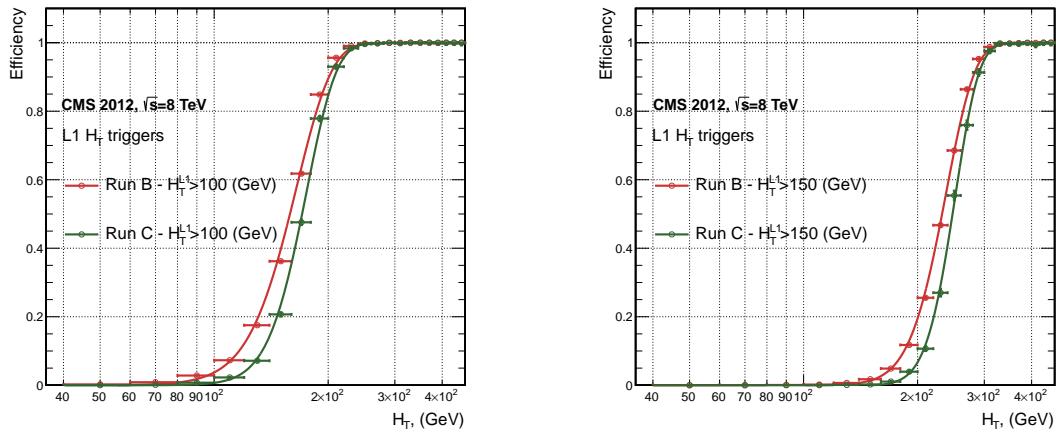


Figure 3.13: L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T , measured for the L1 H_T 100 and 150 trigger during the run 2012 B and C, collected using an isolated single μ triggered sample.

1220 It can be seen that the performance of the $E_T > 36, 92$ single jet triggers are almost
 1221 identical, with the jet seed having no measurable effect on these triggers as shown in
 1222 Table 3.2.

1223 In the case of the H_T triggers, without the jet seed threshold a large increase in the
 1224 trigger cross-section during high luminosity collisions will occur. The low energy threshold

| 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).

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

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

| Trigger | 2012B | | 2012C | |
|-----------|------------------|-----------------|------------------|-----------------|
| | 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).

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

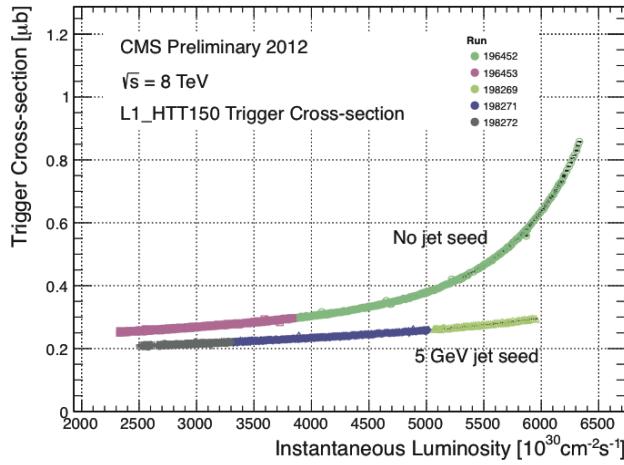


Figure 3.14: Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1 H_T triggers [72].

1238 It can be seen that this slight degradation in the offline value at which these H_T triggers
 1239 become fully efficient due to the jet seed threshold, can be justified from the large
 1240 reduction in the trigger cross-section rate. Any inefficiencies can then if necessary be
 1241 compensated through a reduction in the H_T trigger threshold of the L1 seed.

1242 3.4.5. Robustness of L1 Jet Performance against Pile-up

1243 The performance of the L1 single jet triggers is evaluated in different pile-up conditions
 1244 to determine any dependence on pile-up. Three different pile-up categories of 0-10,
 1245 10-20 and >20 vertices are defined, reflecting the low, medium and high pile-up running
 1246 conditions at CMS in 2012.

1247 The L1 triggers are benchmarked relative to Calo and PF jets in the run period where
 1248 the jet seed threshold *is* applied, for the L1 single jet thresholds of 16, 36 and 92 GeV,
 1249 shown in Figure 3.15. The results of fitting an EMG function to these efficiency turn-on
 1250 curves are given in Table 3.4 and Table 3.5 for Calo and PF jets respectively.

1251 No significant drop in efficiency is observed in the presence of a high number of primary
 1252 vertices. The increase in hadronic activity in higher pile-up conditions, combined with
 1253 the absence of pile-up subtraction for L1 jets, results in the expected observation of
 1254 a decrease in the μ value of the efficiency turn-ons as a function of pile-up. Similarly,
 1255 the resolution, σ , of the turn-ons are found to worsen at a higher number of primary

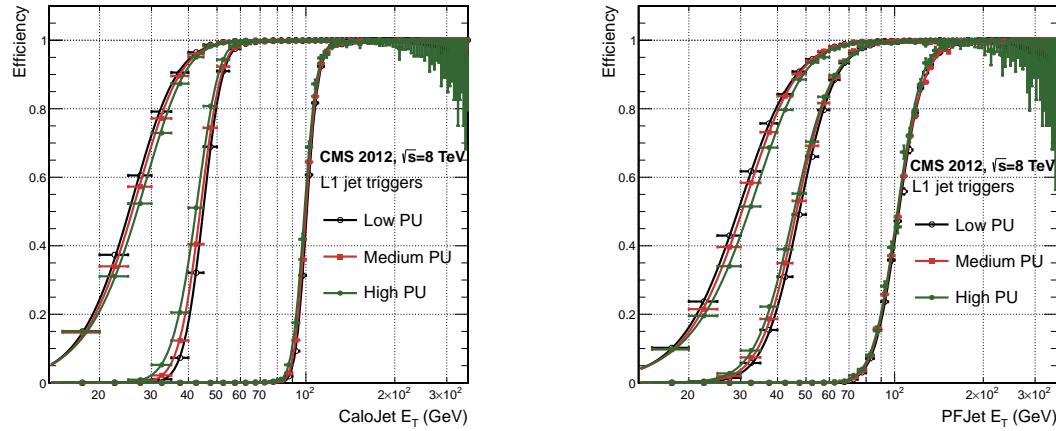


Figure 3.15: L1 jet efficiency turn-on curves as a function of the leading offline E_T Calo (left) and PF (right) jet, for low, medium and high pile-up conditions.

| Vertices | 0-10 | | 11-20 | | > 20 | |
|----------------|----------------|---------------|----------------|---------------|----------------|---------------|
| | μ | σ | μ | σ | μ | σ |
| L1_SingleJet16 | 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.

1256 vertices due to the increasing size of the pile-up corrections being applied to the offline
1257 reconstructed jets.

1258 These features are further emphasised when shown as a function of

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

1259 in bins of matched leading offline jet E_T . The results of these individual fits categorised
1260 as a function of matched leading offline jet E_T can be found in Appendix (B.2), where
1261 each of the distributions are fitted with an EMG function as defined in Equation (3.3).

1262 The μ , σ and λ values extracted for the low, medium and high pile-up conditions are
1263 shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central value

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

Table 3.5: Results of a cumulative EMG function fit to the efficiency turn-on curves for Level-1 single jet triggers in the 2012 run period C, measured from isolated 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.

1264 of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ is observed as expected, to increase as a function of jet E_T , whilst the
 1265 resolution also improves as a function of increasing offline jet E_T for all pile-up categories.

1266 When comparisons are made between the individual pile-up scenarios, it can be seen that
 1267 in the presence of higher pile-up, μ is seen to shift to larger values and a poorer resolution,
 1268 σ , observed. This is particularly evident at low lead jet transverse energy values. These
 1269 differences between the different pile-up scenarios, can once again be attributed to an
 1270 increasing number of soft pile-up jets that add to the transverse energy of the lead jet
 1271 from the primary interaction within each successive pile-up category. However, when
 1272 comparisons of the trigger performance at larger lead jet transverse energy values (> 100
 1273 GeV) are made, similar performance is observed between the separate pile-up categories.

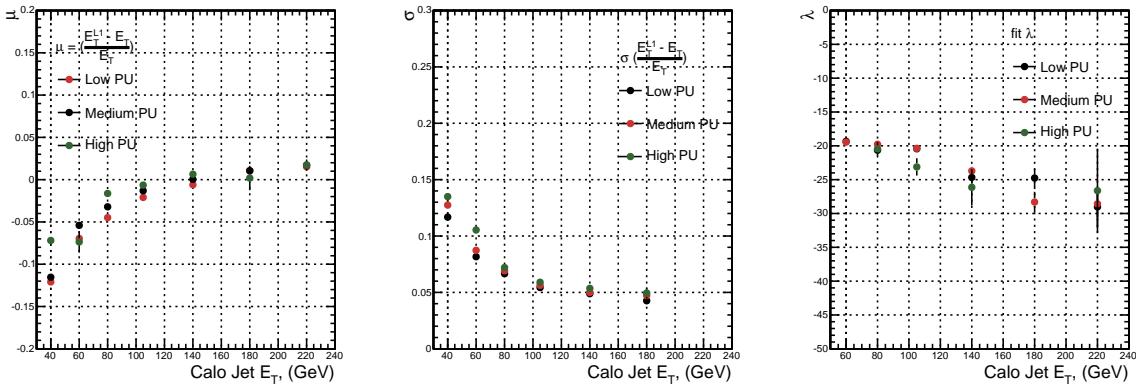


Figure 3.16: Fit values from an EMG function fitted to the resolution plots of leading Calo jet E_T measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for low, medium and high pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian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

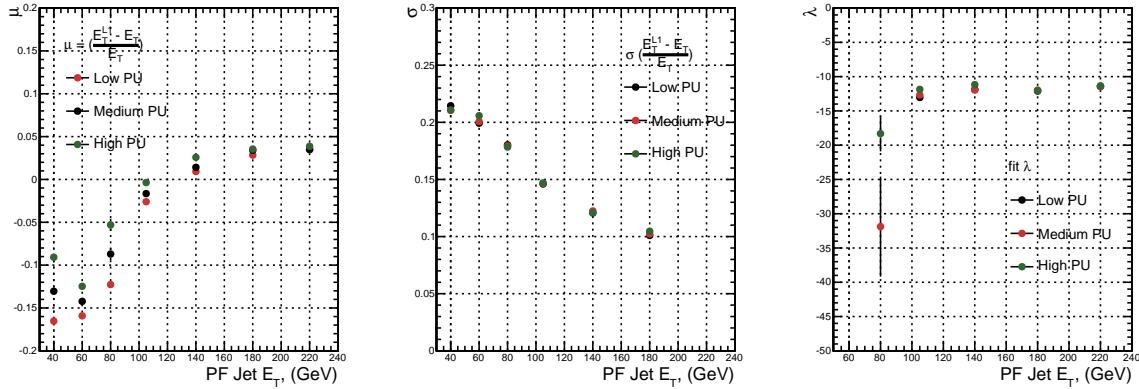


Figure 3.17: Fit values from an EMG function fitted to the resolution plots of leading PF jet E_T measured as a function of $\frac{(L1\ E_T - \text{Offline}\ E_T)}{\text{Offline}\ E_T}$ for low and medium pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian, as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

1274 The resolution of the L1 jet based energy sum quantities, H_T and H_T parameterised as
 1275 in Equation (3.4), can be found in Appendix (B.3).

1276 3.4.6. Summary

1277 The performance of the CMS Level-1 Trigger has been studied and evaluated for jets
 1278 and jet energy sum quantities using data collected during the 2012 LHC 8 TeV run.
 1279 These studies include the effect of the introduction of a 5 GeV jet seed threshold into
 1280 the jet clustering algorithm. The purpose of this change was to mitigate the increase
 1281 in L1 trigger cross-sections, due to larger isotropic energy deposits from an increased
 1282 number of secondary interactions, whilst not adversely affecting the efficiency of these
 1283 triggers. Measurements are made for a range of L1 jet quantities and thresholds, where no
 1284 significant change is observed in the measured efficiencies that would indicate a noticeable
 1285 effect on the overall triggering performance of the detector.

Chapter 4.

¹²⁸⁶ SUSY Searches in Hadronic Final States

¹²⁸⁸ In this chapter a model independent search for SUSY, in hadronic final states with \cancel{E}_T
¹²⁸⁹ using the α_T variable is introduced and described in detail. The results presented are
¹²⁹⁰ based on a data sample of pp collisions collected in 2012 at $\sqrt{s} = 8$ TeV, corresponding
¹²⁹¹ to an integrated luminosity of $11.7 \pm 0.5 \text{ fb}^{-1}$ [5].

¹²⁹² The kinematic variable α_T is motivated as a variable to provide strong rejection of the
¹²⁹³ overwhelming QCD multi-jet background, which is prevalent to jets + \cancel{E}_T final states at
¹²⁹⁴ the LHC. This is achieved whilst maintaining sensitivity to a range of possible SUSY
¹²⁹⁵ signals and is described in Section (4.1). The search and trigger strategy in addition to
¹²⁹⁶ the event reconstruction and selection are outlined within Sections (4.2 - 4.3).

¹²⁹⁷ The method in which the SM background is estimated using data driven control samples
¹²⁹⁸ and an analytical technique to improve statistical precision at higher b-tagged jet
¹²⁹⁹ multiplicities is detailed within Section (4.5). Included in this section is a discussion on
¹³⁰⁰ the impact of b-tagging and mis-tagging scale factors between data and simulation on
¹³⁰¹ any background predictions. Improved precision in estimating background yields at large
¹³⁰² number of b-tagged jets, is important in the context of sensitivity to third generation
¹³⁰³ SUSY models, first outlined in Section (2.4.1).

¹³⁰⁴ A description of the formulation of appropriate systematic uncertainties to be applied
¹³⁰⁵ to the background predictions to account for theoretical uncertainties, limitations in
¹³⁰⁶ the modelling of event kinematics and instrumental effects is covered in Section (4.6).
¹³⁰⁷ Similarly the systematic determination for the SMS signal samples used to interpret the
¹³⁰⁸ physics reach of the analysis are examined in Section (4.7).

1309 Finally the statistical likelihood model to test the compatibility of the data with a SM
1310 only hypothesis, and to interpret the observations within the context of SMS models is
1311 described in Section (4.8). The experimental reach of the analysis discussed within this
1312 thesis is interpreted in two classes of SMS models, both introduced in Section (2.4.1).
1313 The SMS models considered in this analysis are summarised in Table 4.1. For each model,
1314 the LSP is assumed to be the lightest neutralino.

1315 Within the table are also defined reference points, parameterised in terms of parent
1316 gluino/squark and LSP sparticle masses, m_{parent} and m_{LSP} , respectively. These are
1317 used within the following two chapters to demonstrate potential signal yields within the
1318 hadronic search region of the analysis. The masses of each signal topology are chosen to
1319 reflect parameter space which is within the expected sensitivity reach of the search.

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

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.

1320 4.1. An Introduction to the α_T Search

1321 A proton-proton collision resulting in the production and decay of supersymmetric
1322 particles, would manifest as a final state containing energetic jets and \cancel{E}_T in the hadronic
1323 channel. The search focuses on topologies where new heavy supersymmetric, R-parity
1324 conserving particles are pair-produced in pp collisions. These particles decaying to a
1325 LSP escape the detector undetected, leading to significant missing energy and missing
1326 hadronic transverse energy,

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

₁₃₂₇ This is defined as the vector sum of the transverse energies of jets selected in an event.
₁₃₂₈ Energetic jets produced in the decay of these supersymmetric particles also can produce
₁₃₂₉ significant visible transverse energy,

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

₁₃₃₀ defined as the scalar sum of the transverse energies of jets selected in an event.

₁₃₃₁ A search within this channel is greatly complicated in a hadron collider environment;
₁₃₃₂ where the overwhelming background comes from inherently balanced multi-jet (“QCD”)
₁₃₃₃ events, which are produced with an extremely large cross-section as demonstrated within
₁₃₃₄ Figure 4.1. \cancel{E}_T can appear in such events due to a substantial detector mis-measurement,
₁₃₃₅ stochastic fluctuations of jet energy, or missed objects due to detector mis-calibration or
₁₃₃₆ noise effects.

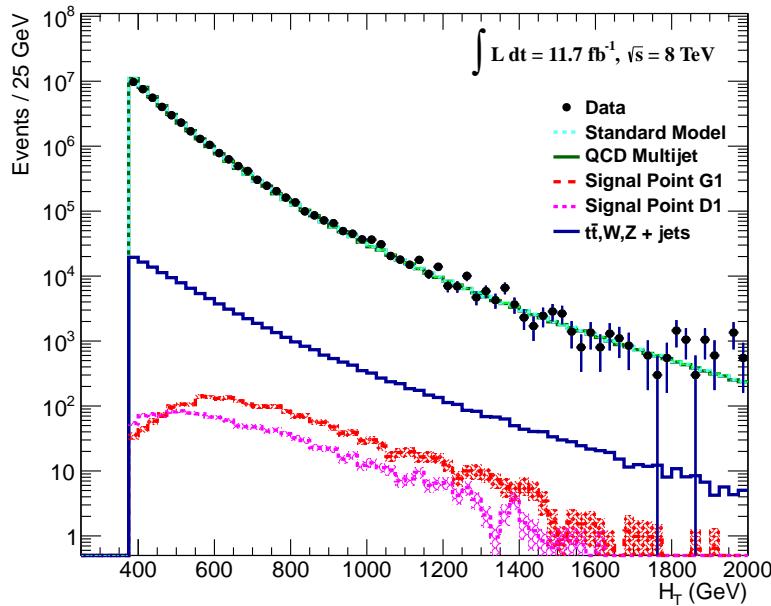


Figure 4.1: Reconstructed offline H_T distribution in the hadronic signal selection (detailed in the following section), from 11.7fb^{-1} of data, in which no α_T requirement was made. The sample is collected from prescaled H_T triggers. Overlaid are expectations from simulation of EWK processes as well as two reference signal models (labelled G1 and D1 from Table 4.1).

1337 Additional SM background from EWK processes with genuine \cancel{E}_T from escaping neutrinos
1338 comprise the irreducible background within this search and come mainly from:

- 1339 • $Z \rightarrow \nu\bar{\nu} + \text{jets}$,
- 1340 • $W \rightarrow l\nu + \text{jets}$ in which a lepton falls outside of detector acceptance, is not
1341 reconstructed, is mis-identified, or the lepton decays hadronically $\tau \rightarrow \text{had}$,
- 1342 • $t\bar{t}$ with at least one leptonically decaying W, which is missed in the detector as
1343 detailed above,
- 1344 • small background contributions from DY, single top and Diboson (WW,ZZ,WZ)
1345 processes.

1346 The search is designed to have a strong separation between events with genuine and
1347 “fake” \cancel{E}_T which is achieved primarily through the dimensionless kinematic variable, α_T
1348 [73][74].

1349 4.1.1. The α_T Variable

1350 For a perfectly measured di-jet QCD event, conservation laws dictate that both jets must
1351 be of equal magnitude and produced in opposite directions. However, in the case of di-jet
1352 events with genuine \cancel{E}_T (as detailed above), no such requirement is made of the two jets,
1353 as depicted in Figure 4.2.

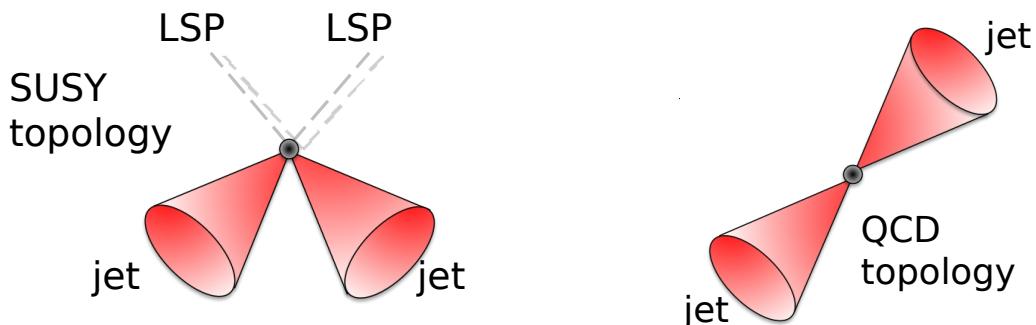


Figure 4.2: The event topologies of background QCD dijet events (right) and a generic SUSY signature with genuine \cancel{E}_T (left).

1353

1354 Exploiting this feature leads to the formulation of α_T (first inspired by [75]) in di-jet
1355 systems defined as,

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

1356 where $E_T^{j_2}$ is the transverse energy of the least energetic of the two jets and M_T defined
1357 as:

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

1358 A perfectly balanced di-jet event i.e. $E_T^{j_1} = E_T^{j_2}$ would yield an α_T value of 0.5. In
1359 processes where a W or Z recoils off a system of jets, these jets will not necessarily be
1360 perfectly balanced and α_T can then achieve values in excess of 0.5. Most importantly,
1361 balanced multi-jet events in which jets *are* mis-measured, will generally result in an α_T
1362 value of less than 0.5, thus giving the α_T variable discriminating power between these
1363 processes.

1364 α_T can be further extended to apply to any arbitrary number of jets. This is undertaken
1365 by modelling a system of n jets as a di-jet system, through the formation of two pseudo-
1366 jets [76]. The two pseudo-jets are built by merging the jets present, and are chosen to
1367 be as balanced as possible, i.e the $\Delta H_T \equiv |E_T^{pj_1} - E_T^{pj_2}|$ is minimised between the two
1368 pseudo jets. Using Equation (4.4), α_T can be rewritten as,

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

1369 The distribution of α_T for the two jet multiplicity categories used within this analysis,
1370 $2 \leq n_{jet} \leq 3$ and $n_{jet} \geq 4$ jets, is shown in the Figure 4.3. It can be seen that the
1371 distributions peak at an α_T value of 0.5, before falling away sharply and being free
1372 of a simulated multi-jet background at larger α_T values. These distributions serve to
1373 demonstrate the ability of the α_T variable to discriminate between multi-jet events and
1374 EWK processes with genuine \cancel{E}_T in the final state.

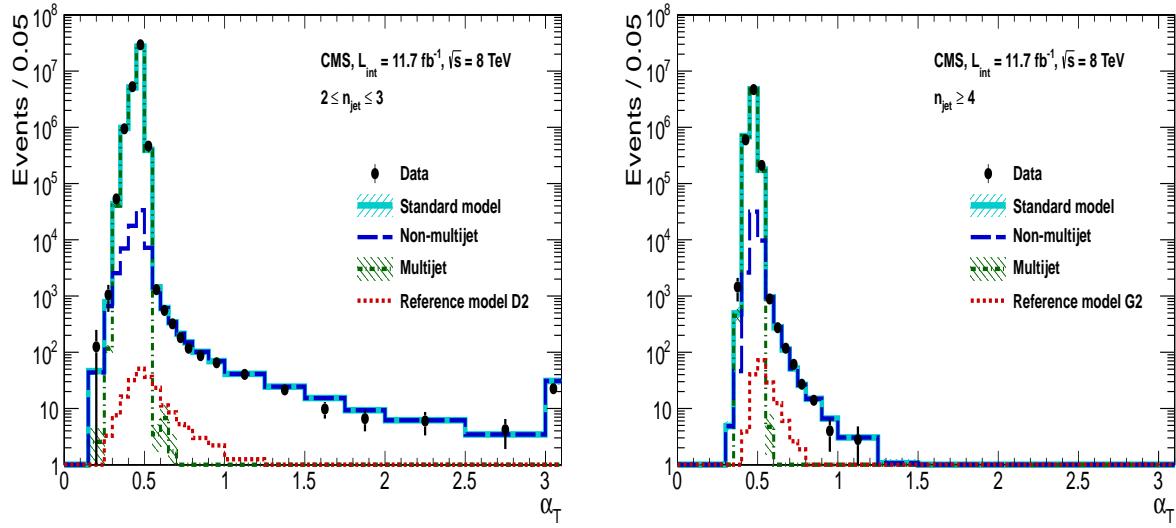


Figure 4.3: The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet multiplicities after a full analysis selection and $H_T > 375$ requirement. Data is collected using both prescaled H_T triggers and dedicated α_T triggers for below and above $\alpha_T = 0.55$ respectively. Expected yields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

1375 The α_T requirement used within the search is chosen to be $\alpha_T > 0.55$ to ensure that
 1376 the QCD multi-jet background is negligible even in the presence of moderate jet mis-
 1377 measurement. There still remain other effects which can cause multi-jet events to
 1378 artificially have a large α_T value, methods to combat them are discussed in detail in
 1379 Section (4.2.2).

1380 4.2. Search Strategy

1381 The aim of the analysis presented in this thesis is to identify an excess of events in data
 1382 over the SM background expectation in multi-jet final states and significant \cancel{E}_T . The
 1383 essential suppression of the dominant multi-jet background for such a search is addressed
 1384 by the α_T variable, described in the previous section. For estimation of the remaining
 1385 EWK backgrounds, three independent data control samples are used to predict the
 1386 different processes that compose the background :

- 1387 • $\mu + \text{jets}$ control sample to determine $W + \text{jets}$, $t\bar{t}$ and single top backgrounds,
- 1388 • $\gamma + \text{jets}$ control sample to determine the irreducible $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background,

1389 • $\mu\mu + \text{jets}$ control sample to also determine the irreducible $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background.

1390 These control samples are chosen to be rich in specific EWK processes, free of QCD
1391 multi-jet events and to also be kinematically similar to the hadronic signal region that
1392 they are estimating the backgrounds of, see Section (4.2.3). The redundancy of using the
1393 $\gamma + \text{jets}$ and $\mu\mu + \text{jets}$ sample to predict the same background within the signal region,
1394 brings an opportunity to reliably crosscheck and validate the background estimation
1395 method, and is utilised in both the determination of background estimation systematics
1396 (Section(4.6)) and in the maximum likelihood fit (Section(4.8)).

1397 To remain inclusive to a large range of possible SUSY models, the signal region is split into
1398 the following categories to allow for increased sensitivity to different SUSY topologies:

1399 **Sensitivity to a range of SUSY mass splittings**

1400 The hadronic signal region is defined by $H_T > 275$, divided into eight bins in H_T .

- 1401 – Two bins of width 50 GeV in the range $275 < H_T < 375$ GeV,
1402 – five bins of width 100 GeV in the range $375 < H_T < 875$ GeV,
1403 – and a final open bin, $H_T > 875$ GeV.

1404 The choice of the lowest H_T bin in the analysis is driven primarily by trigger
1405 constraints. The mass difference between the LSP and the particle that it decays
1406 from is an important factor in the amount of hadronic activity in the event.

1407 A large mass splitting will lead to hard high p_T jets which contribute to the H_T sum.
1408 From Figure 4.1 it can be seen that the SM background falls sharply at high H_T
1409 values, therefore many H_T categories will lead to easier identification of such signals.
1410 Conversely, smaller mass splittings lead to softer jet p_T 's which will subsequently
1411 fall into the lower H_T range.

1412 **Sensitivity to production method of SUSY particles**

1413 The production mechanism of any potential SUSY signal can lead to different event
1414 topologies. One such way to discriminate between gluino ($g\tilde{g}$ - “high multiplicity”),
1415 and direct squark ($q\tilde{q}$ - “low multiplicity”) induced production of SUSY particles is
1416 realised through the number of reconstructed jets in the final state.

1417 The analysis is thus split into two jet categories: $2 \leq n_{\text{jet}} \leq 3$ jets, $n_{\text{jet}} \geq 4$ jets to
1418 give sensitivity to both of these mechanisms.

1419 Sensitivity to “Natural SUSY” via tagging jets from b-quarks

1420 Jets originating from the hadronisation of bottom quarks (b-jets) are identified
1421 through vertices that are displaced with respect to the primary interaction. The
1422 algorithm used in the analysis to identify b-jets is the Combined Secondary Vertex
1423 Medium Working Point (CSVM) tagger, described within Section (3.3.2). A cut
1424 is placed on the discriminator variable of > 0.679 , leading to a gluon/light-quark
1425 tagging efficiency of $\sim 1\%$, a c-quark tagging efficiency of $\sim 20\%$ and a jet p_T
1426 dependant b-tagging efficiency of 60-70% [77].

1427 Natural SUSY models would be characterised through final-state signatures rich
1428 in bottom quarks. A search relying on methods to identify jets originating from
1429 bottom quarks through b-tagging, will significantly improve the sensitivity to this
1430 class of signature. This gain in sensitivity stems from a vast reduction in the vector
1431 boson + jet backgrounds (W, Z) at higher b-tag jet multiplicities, which typically
1432 have no b-flavoured quarks in their decays.

1433 Therefore, events are categorised according to the number of offline reconstructed
1434 b-tagged jets, n_b^{reco} , identified within each event. The following five categories are
1435 used; $n_b^{\text{reco}} = 0, = 1, = 2, = 3$ and ≥ 4 . In the $n_b^{\text{reco}} \geq 4$ category, due to a limited
1436 number of expected signal and background events just three H_T bins are employed:
1437 275-325 GeV, 325-375 GeV, ≥ 375 GeV.

1438 This characterisation is identically mirrored in all control samples, with the infor-
1439 mation from all samples and b-tag categories used simultaneously in the likelihood
1440 model, see Section (4.8).

1441 The combination of the H_T , jet multiplicity and b-tag categorisation of the signal region
1442 as described above, results in 67 different bins in which the analysis is interpreted in. A
1443 visualisation of the analysis categorisation is depicted in Figure 4.4.

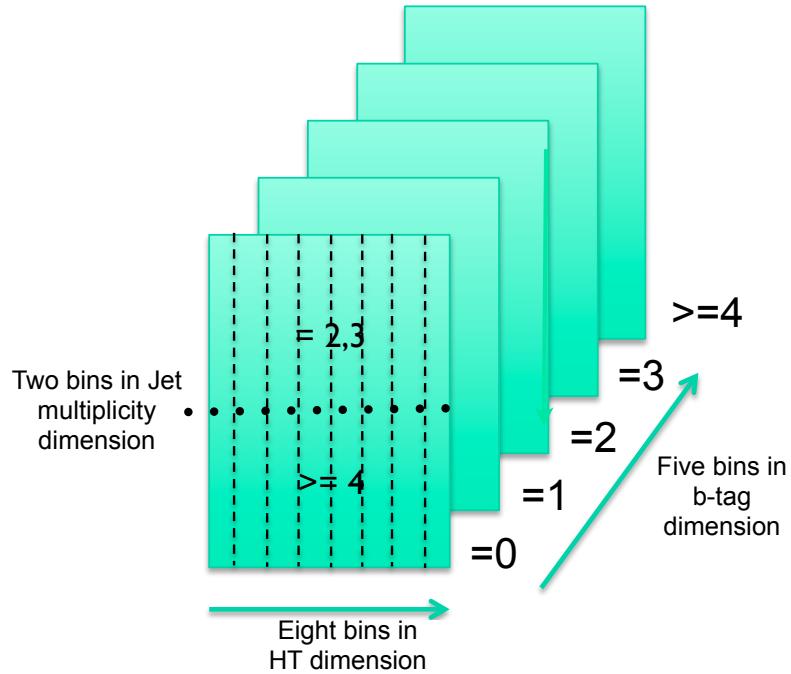


Figure 4.4: Pictorial depiction of the analysis strategy employed by the α_T search to increase sensitivity to a wide spectra of SUSY models.

¹⁴⁴⁴ 4.2.1. Physics Objects

¹⁴⁴⁵ The physics objects used in the analysis are defined below, and follow the recommendation
¹⁴⁴⁶ of the various CMS Physics Object Groups (POGs).

¹⁴⁴⁷ • Jets

¹⁴⁴⁸ The jets used in this analysis are CaloJets, reconstructed as described in Section
¹⁴⁴⁹ (3.3.1) using the anti- k_T jet clustering algorithm.

¹⁴⁵⁰ To ensure the jet object falls within the calorimeter systems a pseudo-rapidity
¹⁴⁵¹ requirement of $|\eta| < 3$ is applied. Each jet must pass a “loose” identification criteria
¹⁴⁵² to reject jets resulting from unphysical energy, the criteria of which are detailed in
¹⁴⁵³ Table A.1 [78].

¹⁴⁵⁴ • Muons

¹⁴⁵⁵ Muons are selected in the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ control samples, and vetoed in
¹⁴⁵⁶ the signal region. The same cut based identification criteria is applied to muons in
¹⁴⁵⁷ both search regions and is summarised in Table 4.2 [79].

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

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

1458 Additionally muons are required to be within the acceptance of the muon tracking
 1459 systems. For the muon control samples, trigger requirements necessitate a $|\eta| <$
 1460 2.1 for the selection of muons. In the signal region where muons are vetoed, these
 1461 conditions are relaxed to $|\eta| < 2.5$ and a minimum threshold of $p_T > 10 \text{ GeV}$ is
 1462 required of muon objects.

1463 **• Photons**

1464 Photons are selected within the $\gamma + \text{jets}$ control sample and vetoed in all other
 1465 selections. Photons are identified in both cases according to the cut based criteria
 1466 listed in Table 4.3 [80].

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

Continued on next page

| | |
|----------------------------|---|
| 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 [81], ECAL, HCAL and tracking isolation sums are used to determine the combined total isolation value. |

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

1467

1468 Photon objects are also required to have a minimum momentum of $p_T > 25$ GeV.

1469 **• Electrons**

1470 Electron identification is defined for veto purposes. They are selected according to
1471 the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

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

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

1472 Electrons are required to be identified at $|\eta| < 2.5$, with a minimum $p_T > 10 \text{ GeV}$
1473 threshold to ensure that the electrons fall within the tracking system of the detector.

1474 • **Noise and \cancel{E}_T Filters**

1475 A series of noise filters are applied to veto events which contain spurious non-physical
1476 jets that are not picked up by the jet id, and events which give large unphysical \cancel{E}_T
1477 values. These filters are listed within Table 4.5.

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

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

1478 **4.2.2. Event Selection**

1479 The selection criteria for events within the analysis are detailed below. A set of common
1480 cuts are applied to both signal (maximise acceptance to a range of SUSY signatures), and
1481 control samples (retain similar jet kinematics for background predictions), with additional
1482 selection cuts applied to each control sample to enrich the sample in a particular EWK
1483 processes, see Section (4.2.3).

- 1484 The jets considered in the analysis are required to have a transverse momentum $p_T > 50$
1485 GeV, with a minimum of two jets required in the event. The highest E_T jet is required
1486 to lie within the central tracker acceptance $|\eta| < 2.5$, and the two leading p_T jets must
1487 each have $p_T > 100$ GeV. Any event which has a jet with $p_T > 50$ GeV that either fails
1488 the “loose” identification criteria described in Section(4.2.1) or has $|\eta| > 3.0$, is rejected.
1489 Similarly events in which an electron, muon or photon fails object identification but
1490 passes η and p_T restrictions, are identified as an “odd” lepton/photon and the event is
1491 vetoed.
- 1492 At low H_T , the jet p_T threshold requirements required to be considered as part of the
1493 analysis and enter the H_T sum are scaled downwards. These are scaled down in order
1494 to extend phase space at low H_T , preserving similar jet multiplicities and background
1495 admixture seen at higher H_T , as listed in Table 4.6.

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

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

- 1496 Within the signal region, to suppress SM processes with genuine \cancel{E}_T from neutrinos,
1497 events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure
1498 multi-jet topology, events are vetoed if an isolated photon is found with $p_T > 25$ GeV.
1499 An α_T requirement of > 0.55 is required to reduce the QCD multi-jet background to
1500 a negligible amount. Finally, additional cleaning cuts are applied to protect against
1501 pathological deficiencies such as reconstruction failures or severe energy mis-measurements
1502 due to detector inefficiencies:

- Significant H_T can arise in events with no real \cancel{E}_T due to multiple jets falling below the p_T threshold for selecting jets. This in turn leads to events which can then incorrectly pass the 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,

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

- 1503 • Fake \cancel{E}_T and \cancel{H}_T can arise due to significant jet mis-measurements caused by a small
 1504 number of non-functioning ECAL regions. These regions absorb electromagnetic
 1505 showers which are subsequently not added to the jet energy sum. To circumvent
 1506 this problem the following procedure is employed: For each jet in the event, the
 1507 angular separation

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

1508 is calculated where that jet is itself removed from the event. Here $\Delta\phi^*$ is a measure
 1509 of how aligned the \cancel{H}_T of an event is with a jet. A small value (i.e. the \cancel{H}_T vector
 1510 lies along the jet axis) is indicative of an inherently balanced event in which a jet has
 1511 been mis-measured. For every jet in an event with $\Delta\phi^* < 0.5$, if the ΔR distance
 1512 between the selected jet and the closest dead ECAL region is also < 0.3 , then the
 1513 event is rejected. Similarly events are rejected if the jet points within $\Delta R < 0.3$ of
 1514 the ECAL barrel-endcap gap at $|\eta| = 1.5$.

1515 Some of the key distributions of the analysis are compared to simulated SM processes,
 1516 shown in Figure 4.5. The simulated samples are normalised to a luminosity of 11.7 fb^{-1} ,
 1517 with no requirement placed upon the number of b-tagged jets or number of jets in the
 1518 distributions shown. In the case of this inclusive selection, the dominant backgrounds
 1519 in the signal regions are, $Z \rightarrow \nu\bar{\nu}$ and $W + \text{jet}$ processes, with a smaller $t\bar{t}$ background
 1520 accompanied by other residual backgrounds.

1521 The distributions shown are presented for purely illustrative purposes, with the simulation
 1522 not used in absolute terms for the estimation of background processes within the signal
 1523 region, see Sections (4.2.3, 4.5). However it is nevertheless important to demonstrate
 1524 that good agreement exists between the modelling of key variables in simulation and
 1525 data.

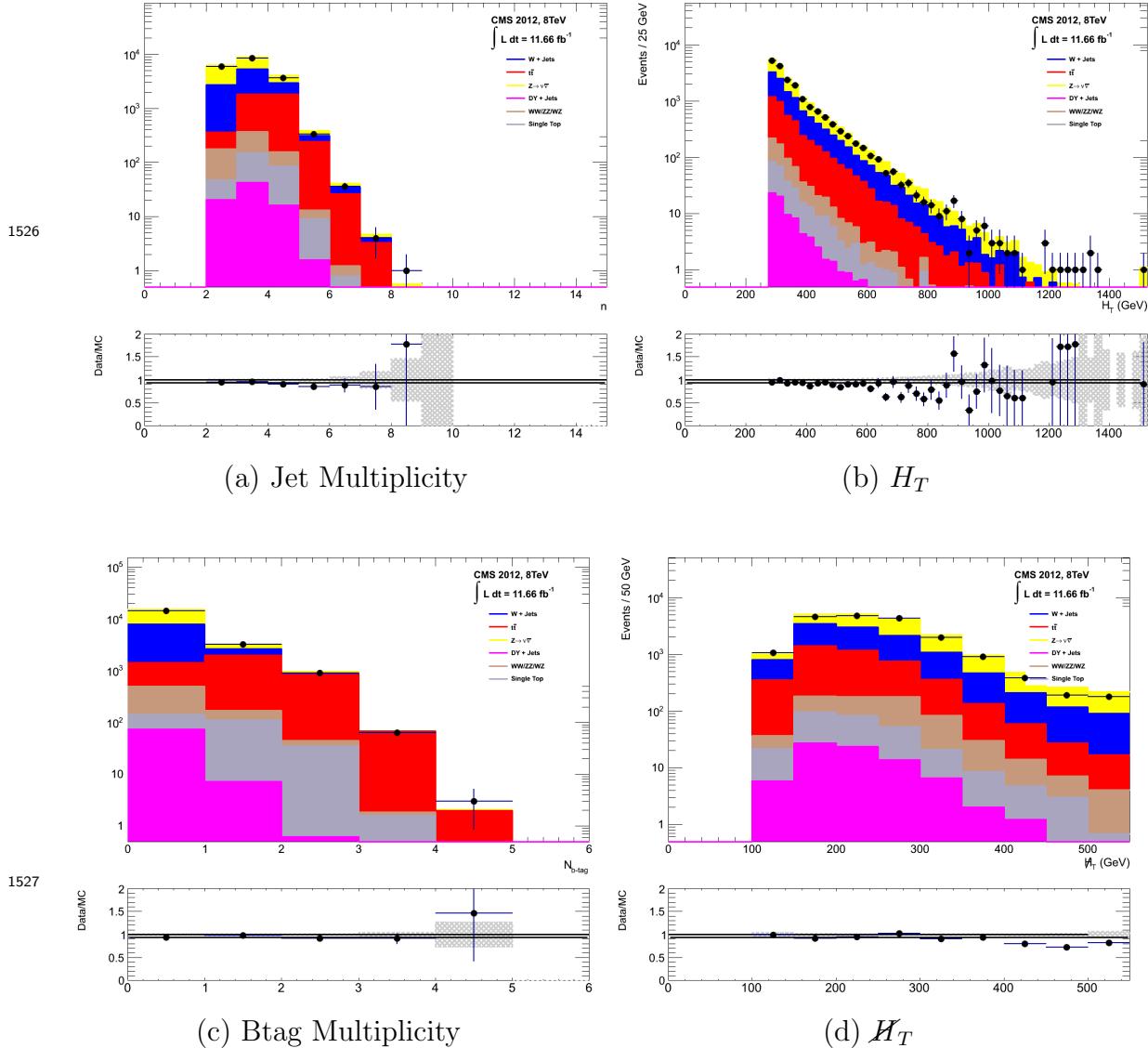


Figure 4.5: Data/MC comparisons of key variables for the hadronic signal region, following the application of the hadronic selection criteria and the requirements of $H_T > 275$ GeV and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement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

1532 in data for a control sample, $N_{\text{obs}}^{\text{control}}$, into a background prediction, $N_{\text{pred}}^{\text{signal}}$, via Equation
1533 (4.7),

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

1534 All simulation samples are normalised to the luminosity of the data samples of the
1535 relevant selection they are being applied to. Through this method, “vanilla” predictions
1536 for the SM background in the signal region can be made by considering separately the
1537 sum of the prediction from the combination of either the $\mu + \text{jets}$ and $\gamma + \text{jets}$, or $\mu +$
1538 jets and $\mu\mu + \text{jets}$ samples.

1539 It must be noted that the final background estimation from which results are interpreted,
1540 is calculated via a fitting procedure defined formally by the likelihood model described
1541 in Section (4.8).

1542 The sum of the expected yields from all simulated processes listed in Section (4.1), enter
1543 the denominator, $N_{\text{MC}}^{\text{control}}$, of the TF defined in Equation (4.7) for each control sample.
1544 However, only the specific processes being estimated by the control sample enter the
1545 numerator, $N_{\text{MC}}^{\text{signal}}$.

1546 For the $\mu + \text{jets}$ sample the processes entering the numerator are,

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

1547 whilst for both the $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ samples the only simulated processes used in
1548 the numerator is,

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

1549 The control samples and the EWK processes they are specifically tuned to select are
1550 defined below, with distributions of key variables for each of the control samples shown
1551 for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed upon the

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

The $\mu +$ jets control sample

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

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

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

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

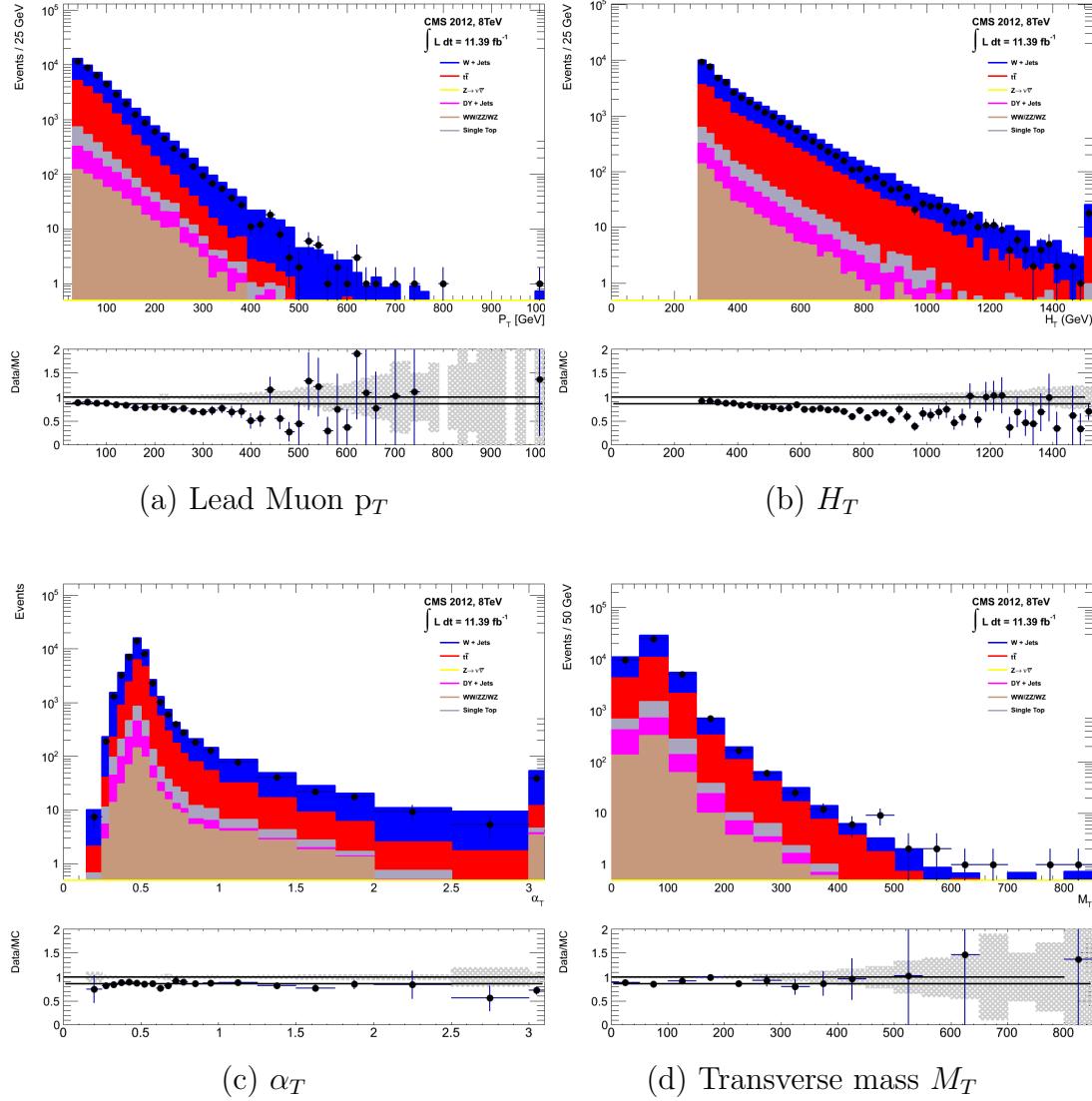


Figure 4.6: Data/MC comparisons of key variables for the $\mu + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of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

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

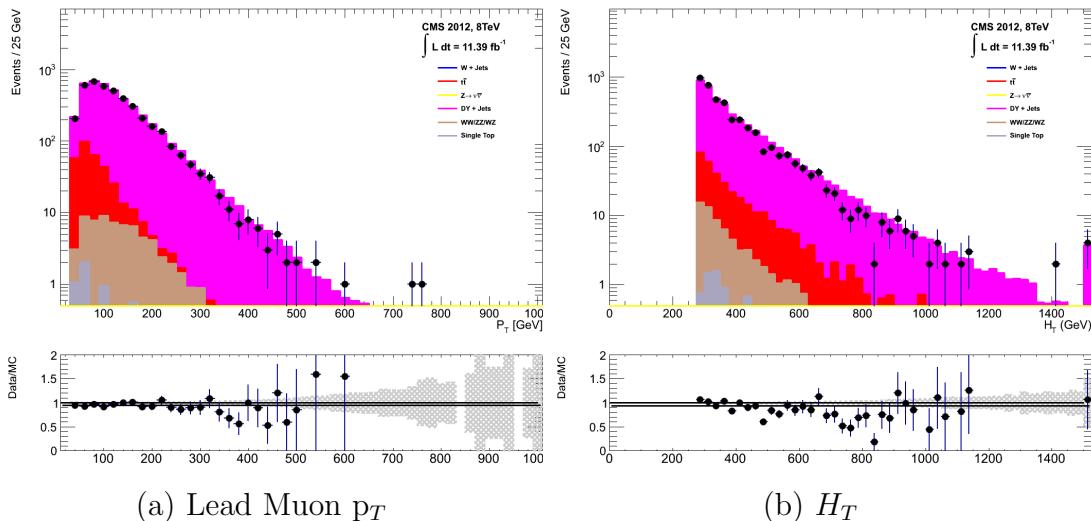
An irreducible $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background enters into the signal region from genuine E_T from the escaping neutrinos. This background is estimated using two control samples, the first of which is the $Z \rightarrow \mu\bar{\mu} + \text{jets}$ process, which posses identical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

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

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

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

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



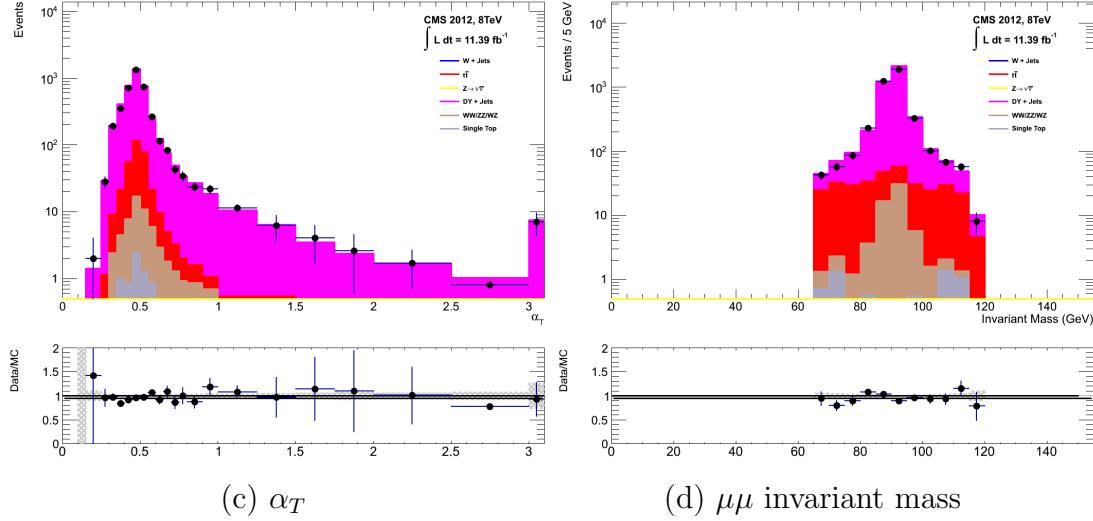


Figure 4.7: Data/MC comparisons of key variables for the $\mu\mu + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number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

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

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

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

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

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

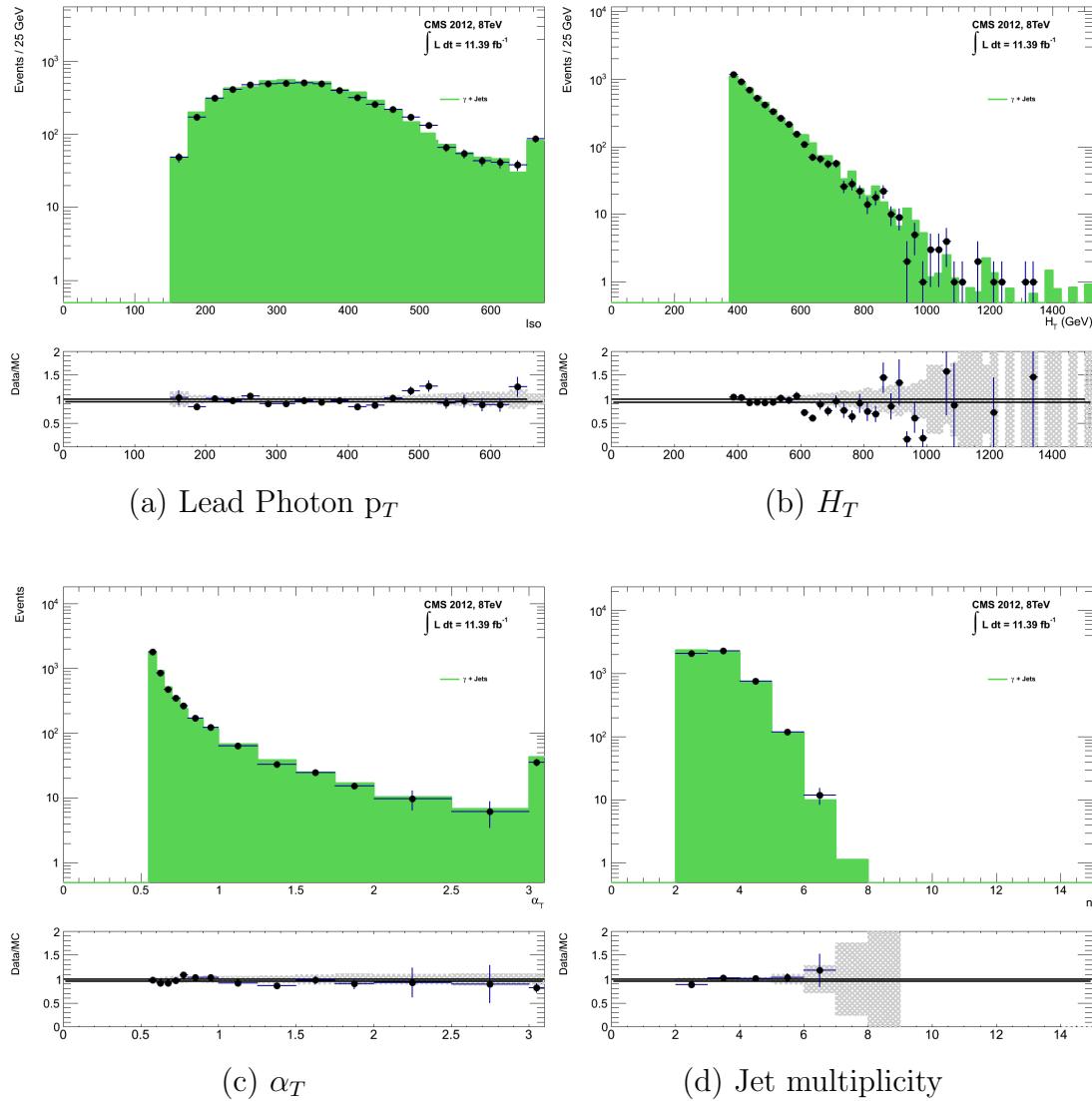


Figure 4.8: Data/MC comparisons of key variables for the $\gamma + \text{jets}$ selection, following the application of selection criteria and the requirements that $H_T > 375 \text{ GeV}$ and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement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

1625 The selection criteria of the three control samples are defined to ensure background
1626 composition and event kinematics mirror closely the signal region. This is done in order
1627 to minimise the reliance on simulation to model correctly the backgrounds and event
1628 kinematics in the control and signal samples.

1629 However in the case of the $\mu + \text{jets}$ and $\mu\mu + \text{jets}$ samples, the α_T requirement is relaxed
1630 in the selection criteria of these samples. This is made possible as contamination from
1631 QCD multi-jet events is suppressed to a negligible level by the other kinematic selection
1632 criteria within the two control samples, selecting pure EWK processes. Thus in this way,
1633 the acceptance of the two muon control samples can be significantly increased, which
1634 simultaneously improves their statistical and predictive power and also dilutes the effect
1635 of any potential signal contamination.

1636 The modelling of the α_T variable is probed through a dedicated set of closure tests,
1637 described in Section (4.6), which demonstrate that the different α_T acceptances for
1638 the control and signal samples have no significant systematic bias on the background
1639 predictions.

1640 4.2.4. Estimating the QCD Multi-jet Background

1641 A negligible background from QCD multi-jet events within the hadronic signal region is
1642 expected due to a combination of selection requirements, and additional applied cleaning
1643 filters. However a conservative approach is still adopted and the likelihood model, see
1644 Section (4.8.2), is given the freedom to accommodate any potential QCD multi-jet
1645 contamination.

1646 Any potential contamination can be identified through the variable R_{α_T} , defined as the
1647 ratio of events above and below the α_T threshold value used in the analysis. This is
1648 modelled by a H_T dependant falling exponential function which takes the form,

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

1649 where the parameters A_{QCD} and k_{QCD} are the normalisation and exponential decay
1650 constants, respectively.

1651 For QCD multi-jet event topologies, this exponential behaviour as a function of H_T is
1652 expected for several reasons. The improvement of jet energy resolution at higher H_T
1653 due to higher p_T jets leads to a narrower peaked α_T distribution, causing R_{α_T} to fall.
1654 Similarly at higher H_T values > 375 GeV, the jet multiplicity rises slowly with H_T . As
1655 shown in Figure 4.3, at higher jet multiplicities the result of the combinatorics used in

1656 the determination of α_T lead to more conservative α_T values, also resulting in a narrower
1657 distribution.

1658 The value of the decay constant k_{QCD} is constrained via measurements within data
1659 sidebands to the signal region. This is also done to validate the falling exponential
1660 assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet
1661 background and defined as regions where either α_T is relaxed or that the R_{miss} cut is
1662 inverted. Figure 4.9 depicts the definition of these data sidebands used to constrain the
1663 value of k_{QCD} .

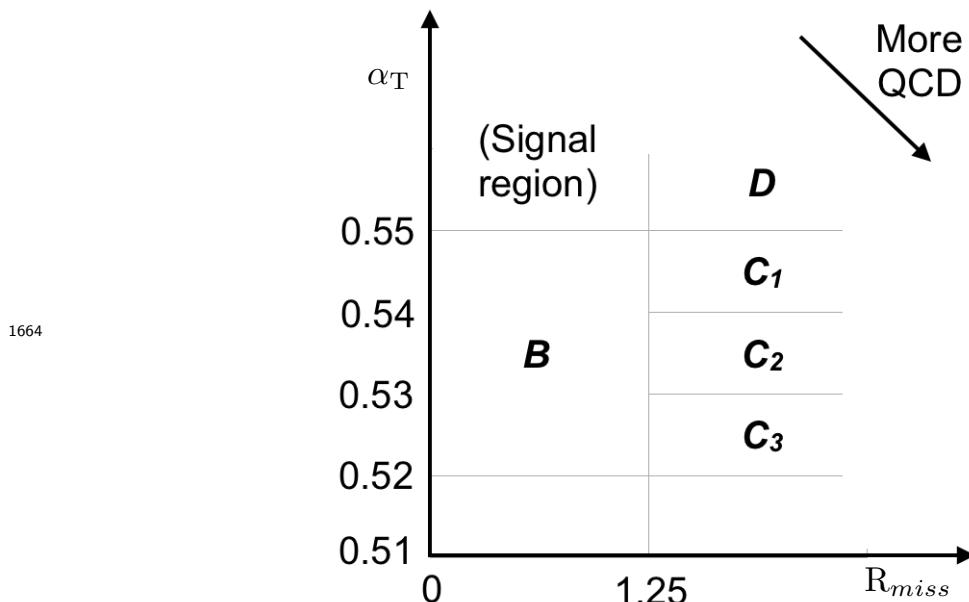


Figure 4.9: QCD sideband regions, used for determination of k_{QCD} .

1664 The fit results used to determine the value of k_{QCD} are shown in Appendix (C.1), for
1665 which the best fit parameter value obtained from sideband region B is determined to be
1666 $k_{\text{QCD}} = 2.96 \pm 0.64 \times 10^{-2} \text{ GeV}^{-1}$.

1667 The best fit values of the remaining three C sideband regions are used to estimate
1668 the systematic uncertainty on the central value obtained from sideband region B. The
1669 variation of these measured values is used to determine the error on the determined
1670 central value, and is calculated to be $1.31 \pm 0.26 \times 10^{-2} \text{ GeV}^{-1}$. This relative error of \sim
1671 20% gives an estimate of the systematic uncertainty of the measurement to be applied to
1672 k_{QCD} .

1673 Finally the same procedure is performed for sideband region D as an independent
1674 crosscheck, to establish that the value of k_{QCD} extracted from a lower α_T slice, can be

1676 applied to the signal region $\alpha_T > 0.55$. The likelihood fit is performed across all H_T
1677 bins within the QCD enriched region with no constraint applied to k_{QCD} . The resulting
1678 best fit value for k_{QCD} shows good agreement between that and the weighted mean,
1679 determined from the three C sideband regions. This demonstrates that the assumption
1680 of using the central value determined from sideband region B, to provide an unbiased
1681 estimator for k_{QCD} in the signal region ($\alpha_T > 0.55$) is valid.

1682 Table 4.7 summarises the best fit k_{QCD} values determined for each of the sideband regions
1683 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 4.7: 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.

1684 4.3. Trigger Strategy

1685 A cross trigger based on the H_T and α_T values of an event, is used with varying thresholds
1686 across H_T bins to record the events used in the hadronic signal region. The α_T legs of
1687 the HT_alphaT triggers used in the analysis, are chosen to suppress QCD multi-jet events
1688 and control trigger rate, whilst maintaining signal acceptance. To maintain an acceptable
1689 rate for these analysis triggers, only calorimeter information is used in the reconstruction
1690 of the H_T sum, leading to the necessity for Calo jets to be used within the analysis.

1691 A single object prescaled HT trigger is used to collect events for the hadronic control
1692 region, described above in Section (4.2.4).

1693 The performance of the α_T and H_T triggers used to collect data for the signal and
1694 hadronic control region is measured with respect to a reference sample collected using
1695 the muon system. This allows measurement of both the Level 1 seed and higher level

1696 triggers simultaneously, as the reference sample is collected independently of any jet
1697 requirements.

1698 The selection for the trigger efficiency measurement is identical to that described in
1699 Section (4.2.2), with the requirement of exactly one well identified muon with $p_T > 30$
1700 GeV. This muon is then subsequently ignored.

1701 The efficiencies measured for the HT_alphaT triggers in each individual H_T and α_T leg,
1702 is summarised in Table 4.8 for each H_T category of the analysis.

| 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 4.8: 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.

1703 Data for the control samples of the analysis, detailed in Section (4.2.3), are collected
1704 using a single object photon trigger for the $\gamma +$ jets sample, and a single object muon
1705 trigger for both the $\mu +$ jets and $\mu\mu +$ jets control samples.

1706 The photon trigger is measured to be fully efficient for the threshold $p_T^{\text{photon}} > 150$ GeV,
1707 whilst the single muon efficiency satisfying $p_T^{\text{muon}} > 30$ GeV is measured to have an
1708 efficiency of $(88 \pm 2)\%$ that is independent of H_T . In the case of the $\mu\mu +$ jets control
1709 sample, the efficiency is measured to be $(95 \pm 2)\%$ for the lowest H_T bin, rising (due to
1710 the average p_T of the second muon in the event increasing at larger H_T) to $(98 \pm 2)\%$ in
1711 the highest H_T category.

1712 4.4. Measuring Standard Model Process 1713 Normalisation Factors via H_T Sidebands

1714 The theoretical cross-sections of different SM processes at Next to Next Leading Order
1715 (NNLO) and the number of available simulated events generated for a particular process,
1716 is typically used to determine the appropriate normalisation for a simulation sample.
1717 However within the particular high- H_T and high- \mathcal{E}_T corners of kinematic phase space

1718 probed within this search, the theoretical cross sections for various processes are far less
1719 well understood.

1720 To mitigate the problem of theoretical uncertainties and arbitrary choices of cross
1721 sections, the normalisation of the simulation samples are determined through the use of
1722 data sidebands. The sidebands are used to calculate sample specific correction factors
1723 (k-factors), that are appropriate for the H_T - \cancel{E}_T phase space covered by this analysis.

1724 They are defined within the $\mu +$ jets and $\mu\mu +$ jets control sample, by the region $200 <$
1725 $H_T < 275$, using the same jet p_T thresholds as the adjacent first analysis bin. Individual
1726 EWK processes are isolated within each of these control samples via requirements on jet
1727 multiplicity and the requirement on b-tag multiplicity, summarised in Table 4.9. The
1728 purity of the samples are typically $> 90\%$ with any residual contamination subtracted
1729 prior to determination of the correction factors. The resultant k-factor for each process
1730 is determined by then taking ratio of the data yield over the expectation from simulation
1731 in the sideband. Subsequently these k-factors are then applied to the processes within
1732 the phase space of the analysis.

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

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

1733 It is worth pointing out that these correction factors have a negligible effect when
1734 providing a background estimation for the signal region. The TFs used in the analysis
1735 are found to be unaffected by application of these k-factors due to the similarity in the
1736 background composition of the control and signal regions. However when systematic
1737 uncertainties are determined in Section (4.6), the closure tests performed are sensitive
1738 to these corrections when extrapolations between different n_b^{reco} and n_{jet} categories are
1739 performed.

¹⁷⁴⁰ 4.5. Determining Monte Carlo Simulation Yields ¹⁷⁴¹ with Higher Statistical Precision

¹⁷⁴² Reconstructing events from EWK processes with many b-tagged jets (≥ 3), n_b^{reco} , is largely
¹⁷⁴³ driven by the mis-tagging of light jets within the event. This is clear when considering the
¹⁷⁴⁴ main EWK backgrounds in the analysis, such as $t\bar{t} + \text{jets}$ events, which typically contain
¹⁷⁴⁵ two underlying b-quarks in the final state from the decay of the top quarks. Conversely
¹⁷⁴⁶ $W + \text{jets}$ and $Z \rightarrow \mu\mu + \text{jets}$ events will typically contain no underlying b-quarks in its
¹⁷⁴⁷ final state.

¹⁷⁴⁸ When the expectation for the number of n_b^{reco} jets is taken directly from simulation,
¹⁷⁴⁹ the statistical uncertainty at large reconstructed b-tagged jet multiplicities becomes
¹⁷⁵⁰ relatively large. One approach to reduce this uncertainty is to use the information
¹⁷⁵¹ encoded throughout all events in the simulation sample, to measure each of the following
¹⁷⁵² four ingredients:

- ¹⁷⁵³ 1. the averaged b-tagging efficiency in the event selection,
- ¹⁷⁵⁴ 2. the averaged charm-tagging efficiency in the event selection,
- ¹⁷⁵⁵ 3. the averaged mis-tagging efficiency in the event selection,
- ¹⁷⁵⁶ 4. the underlying flavour distribution of the jets in the event sample.

¹⁷⁵⁷ Together they can be used to determine the n_b^{reco} distribution of the process being
¹⁷⁵⁸ measured. This method allows the determination of higher b-tagged jet multiplicities to
¹⁷⁵⁹ a higher degree of accuracy, reducing the statistical uncertainties of the simulation yields
¹⁷⁶⁰ which enter into the TF's. For the discussion that follows, this approach will be known
¹⁷⁶¹ as the formula method.

¹⁷⁶² 4.5.1. The Formula Method

¹⁷⁶³ The assigning of jet flavours to reconstruction level jets in simulation is achieved via an
¹⁷⁶⁴ algorithmic method defined as:

- ¹⁷⁶⁵ • attempt to find the parton that most likely determines the properties of the jet and
¹⁷⁶⁶ assign that flavour as the true flavour,

- 1767 • “final state” partons (after showering, radiation) are analysed (also within $\Delta R <$
 1768 0.3 of reconstructed jet cone),
 1769 • if there is a b/c flavoured parton within jet cone: label jet as a b/c flavoured jet,
 1770 • otherwise: assign flavour of the hardest (highest p_T) parton within the jet cone.

1771 This process is employed within each individual simulation sample and independently for
 1772 each H_T - n_{jet} category in the analysis.

1773 Let $N(n_b^{\text{gen}}, n_c^{\text{gen}}, n_q^{\text{gen}})$ represent the 3-dimensional underlying jet flavour distribution in
 1774 simulation, with b underlying b-quarks, c underlying c-quarks and q underlying light
 1775 quarks which are matched to reconstructed jets as detailed above. Light quarks defined
 1776 as those which originate from a u , d , s , g and τ jets, which having similar mis-tagging
 1777 rates are grouped together.

1778 The n_b^{reco} distribution within each H_T - n_{jet} category of the analysis can be constructed
 1779 for each process in turn in an analytical way using the formula:

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

1780 with $N(n)$ representing the yield of n b-tagged jet events of a simulated process as
 1781 calculated by the formula method.

1782 The variables $P(n_b^{\text{tag}}, n_b^{\text{gen}}, \epsilon)$, $P(n_c^{\text{tag}}, n_c^{\text{gen}}, \beta)$ and $P(n_q^{\text{tag}}, n_q^{\text{gen}}, m)$ correspond to the
 1783 binomial probabilities for the tagging of a jet flavour to occur, based on its measured
 1784 tagging efficiency (ϵ , β , m). These efficiencies are measured individually for each analysis
 1785 category from simulation, using all simulated process events passing selection criteria.
 1786 Thus the tagging efficiencies used within the above formula, represent the averaged
 1787 tagging efficiency of each jet flavour within the phase space of the analysis category.

1788 Finally, the constraints $n_{b/c/q}^{\text{tag}}$ signify the number of tagged jets of a particular jet flavour,
 1789 of which the sum of the three terms must equal the number of n tagged jets being
 1790 calculated. Similarly each $n_{b/c/q}^{\text{gen}}$ term represents the identified jet flavour of each jet in
 1791 the event, and is required by definition for the sum of the three terms to fall within the
 1792 n_{jet} category being analysed.

₁₇₉₃ This approach ultimately results in a more precise n_b^{reco} distribution prediction, due to
₁₇₉₄ the utilisation of all events in the simulation sample which pass selection in extracting
₁₇₉₅ the overall n_b^{reco} distribution.

₁₇₉₆ **4.5.2. Establishing Proof of Principle**

₁₇₉₇ In order to validate the procedure, the predictions determined from the formula method
₁₇₉₈ summarised in Equation (4.11), are compared directly with those obtained directly from
₁₇₉₉ simulation. Resultantly no simulation to data correction factors are applied.

₁₈₀₀ This sanity check for the $\mu + \text{jets}$ control sample is presented in Table 4.10, for all n_b^{reco}
₁₈₀₁ and H_T categories with no requirement placed upon the jet multiplicity of the events.

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

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

1802 It can be seen as expected, that there is good consistency between the results determined
 1803 via the formula method and ‘raw’ simulation yields. Similarly the power of this approach
 1804 can be seen in the reduction of this statistical error in the prediction across all H_T and
 1805 n_b^{reco} categories. In particular the statistical uncertainty is reduced by several factors in
 1806 the highest $n_b^{reco} \geq 4$ category.

1807 4.5.3. Correcting Measured Efficiencies in Simulation to Data

1808 As detailed in Section (3.3.2), it is necessary for certain p_T and η dependant corrections,
 1809 to be applied to both the b-tagging efficiency and mis-tagging rates in order to correct
 1810 the efficiencies from simulation to the efficiencies measured in data. These correction
 1811 factors are considered when determining the simulation yields for each selection, which
 1812 are used to construct the TFs of the analysis. The magnitude of this correction are
 1813 measured individually for each H_T category and are shown in Figure 4.10.

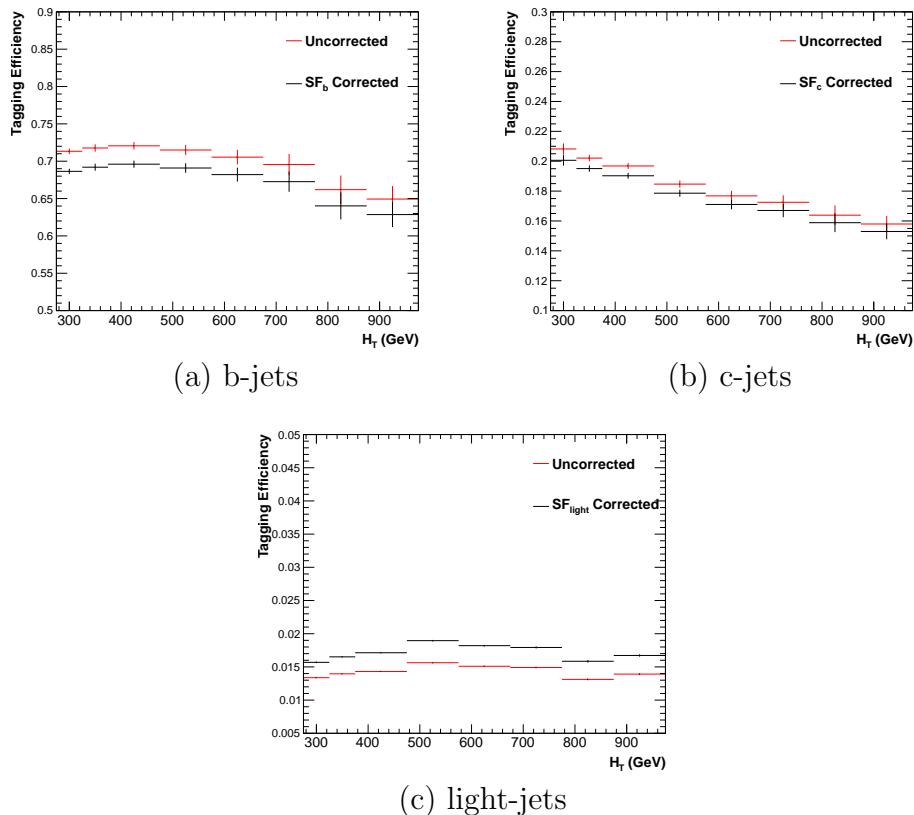


Figure 4.10: Tagging efficiencies of (a) b-jets, (b) c-jets, and (c) light-jets as a function all jets within each H_T category. Efficiencies measured directly from simulation (black) and with data to simulation $SF_{b,c,light}$ correction factors (red) are applied.

1814 Each of the correction factors for the b, c and light flavoured jets come with an associated
 1815 systematic uncertainty. The uncertainties across different jet p_T and η categories, are
 1816 considered as fully correlated. When computing the magnitude of the effect of this
 1817 systematic uncertainty on the TFs of the analysis, the measured tagging efficiencies for
 1818 each jet flavour are scaled up/down simultaneously within each H_T and n_{jet} category by
 1819 the systematic uncertainty of the $SF_{b, c, \text{light}}$ scale factors.

1820 Varying the scale factor corrections by their systematic uncertainty will change the
 1821 absolute yields within each n_b^{reco} bin of all selections. However, ultimately it is the change
 1822 in the TFs which influences the final background prediction from each of the control
 1823 samples. The magnitude of the absolute change in each TF, constructed from when the
 1824 $\mu + \text{jets}$ control sample is used to predict the entire hadronic signal region background,
 1825 is shown in Table 4.11.

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

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

1826 It can be seen that the TFs are found to be relatively insensitive to the systematic
 1827 uncertainty of the b-tag scale factors (showing typically less than $\sim 2\%$ change). This
 1828 can be accounted for by the similar composition of the signal and control sample
 1829 backgrounds, such that any change in the underlying n_b^{reco} distribution will be reflected
 1830 in both signal and control regions and cancel out in the TF.

1831 Any overall systematic effect on the overall background prediction of the analysis from
1832 these b-tag scale factor uncertainties is incorporated within the data driven systematics
1833 introduced in the following section.

1834 4.6. Systematic Uncertainties on Transfer Factors

1835 Since the TFs used to establish the background prediction are obtained from simulation,
1836 an appropriate systematic uncertainty is assigned to account for theoretical uncertainties
1837 [83] and limitations in the simulation modelling of event kinematics and instrumental
1838 effects.

1839 The magnitudes of these systematic uncertainties are established through a data driven
1840 method, in which the three independent control samples of the analysis ($\mu + \text{jets}$, $\mu\mu$
1841 + jets, $\gamma + \text{jets}$) are used to in a series of closure tests. The yields from one of these
1842 control samples, along with the corresponding TF obtained from simulation, are used to
1843 predict the expected yields in another control sample. This procedure therefore utilises
1844 the same method used in determining a background prediction for the signal region as
1845 already established in Section (4.2.3).

1846 The level of agreement between the predicted and observed yields is expressed as the
1847 ratio

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

1848 while considering only the statistical uncertainties on the prediction, N_{pred} , and the
1849 observation, N_{obs} . No systematic uncertainty is assigned to the prediction, and resultantly
1850 the level of closure is defined by the statistical significance of a deviation from the ratio
1851 from zero.

1852 This ratio is measured for each H_T category in the analysis, allowing these closure tests to
1853 be sensitive to both the presence of any significant biases or any possible H_T dependence
1854 to the level of closure.

1855 Eight sets of closure tests are defined between the three data control samples, conducted
1856 independently between the two jet multiplicity ($2 \leq n_{\text{jet}} \leq 3$, $n_{\text{jet}} \geq 4$) categories. Each

1857 of these tests are specifically chosen to probe each of the different key ingredients of the
1858 simulation modelling that can affect the background prediction.

1859 Each of the different modelling components and the relevant closure tests are described
1860 below:

1861 **α_T modelling**

1862 The modelling of the α_T distribution in genuine E_T events is probed with the μ
1863 + jets control sample. This test is important to verify the approach of removing
1864 the $\alpha_T > 0.55$ requirement from the $\mu +$ jets and $\mu\mu +$ jets samples to increase
1865 the precision of the background prediction. The test uses the $\mu +$ jets sample
1866 without an α_T cut to make a prediction into the $\mu +$ jets sample defined with the
1867 requirement $\alpha_T > 0.55$.

1868 **Background admixture**

1869 The sensitivity of the translation factors to the relative admixture of events from
1870 $W +$ jets and $t\bar{t}$ processes is probed by two closure tests.

1871 Within the $\mu +$ jets sample, a W boson enriched sub-sample ($n_b = 0$) is used
1872 to predict yields in a $t\bar{t}$ enriched sub-sample ($n_b = 1$). Similarly, the $t\bar{t}$ enriched
1873 sub-sample ($n_b = 1$) is also used to predict yields for a further enriched $t\bar{t}$ sub-sample
1874 ($n_b = 2$), further probing the modelling of the n_b^{reco} distribution.

1875 Similarly a further closure test probes the relative contribution of $Z +$ jets to $W +$ jets
1876 and $t\bar{t}$ events, through the use of the $\mu +$ jets sample to predict yields for the $\mu\mu +$
1877 jets control sample. This closure test, also at some level probes the muon trigger
1878 and reconstruction efficiencies, given that exactly one or two muons are required by
1879 the different selections.

1880 These tests represent an extremely conservative approach as the admixture of the
1881 two backgrounds remains similar when a prediction is made between the control
1882 samples and the signal region. This is contrary to the closure tests defined above
1883 which make predictions between two very different admixtures of $W +$ jets and $t\bar{t}$
1884 events.

1885 **Consistency check between $Z \rightarrow \nu\bar{\nu}$ predictions**

1886 This is an important consistency check between the $\mu\mu +$ jets and $\gamma +$ jets, which
1887 are both used in the prediction of the $Z \rightarrow \nu\bar{\nu}$ in the signal region. This is conducted

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

Modelling of jet multiplicity

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

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

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

| Closure test | Symbol | Constant fit | | Linear fit | |
|---|----------|------------------|---------|---------------------|---------|
| | | Best fit value | p-value | Slope (10^{-4}) | p-value |
| $\alpha_T < 0.55 \rightarrow \alpha_T > 0.55 (\mu + \text{jets})$ | Circle | -0.06 ± 0.02 | 0.93 | -1.3 ± 2.2 | 0.91 |
| 0 b-jets \rightarrow 1 b-jet ($\mu + \text{jets}$) | Square | 0.07 ± 0.02 | 0.98 | -1.6 ± 1.6 | 1.00 |
| 1 b-jets \rightarrow 2 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 4.12: A summary of the results obtained from zeroeth order polynomial (i.e. a constant) and linear fits to five sets of closure tests performed in the $2 \leq n_{\text{jet}} \leq 3$ category. The two columns show the best fit value for the slope obtained when performing a constant (left) and linear (right) fit and the p-value for that fit.

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

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

| Closure test | Symbol | Constant fit | | Linear fit | |
|--|----------|------------------|---------|---------------------|---------|
| | | Best fit value | p-value | Slope (10^{-4}) | p-value |
| $\alpha_T < 0.55 \rightarrow \alpha_T > 0.55 (\mu + \text{jets})$ | Circle | -0.05 ± 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 4.13: A summary of the results obtained from zeroeth order polynomial (i.e. a constant) and linear fits to five sets of closure tests performed in the $n_{\text{jet}} \geq 4$ category. The two columns show the best fit value for the slope obtained when performing a constant (left) and linear (right) fit and the p-value for that fit.

| Closure test | Symbol | Constant fit | | Linear fit | |
|---------------------------------------|-------------------|------------------|---------|---------------------|---------|
| | | Best fit value | p-value | Slope (10^{-4}) | p-value |
| $\mu + \text{jets}$ | Inverted triangle | -0.03 ± 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 4.14: A summary of the results obtained from zeroeth order polynomial (i.e. a constant) and linear fits to three sets of closure tests performed between the $2 \leq n_{\text{jet}} \leq 3$ and $n_{\text{jet}} \geq 4$ categories. The two columns show the best fit value for the slope obtained when performing a constant (left) and linear (right) fit and the p-value for that fit.

¹⁹¹² statistically compatible with zero bias (i.e. less than 2σ from zero) or at the level of 10%

¹⁹¹³ or less, with the exception of one closure test discussed below.

¹⁹¹⁴ Within Table 4.14, there exists one test that does not satisfy the above statement, which
¹⁹¹⁵ is the $2 \leq n_{\text{jet}} \leq 3 \rightarrow n_{\text{jet}} \geq 4$ test using the $\mu + \text{jets}$ control sample. The low p-value
¹⁹¹⁶ can be largely attributed to an outlier between $675 < H_T < 775$ GeV, rather than any
¹⁹¹⁷ significant trend in H_T . Removing this single outlier from the constant fit performed,
¹⁹¹⁸ gives a best fit value of -0.04 ± 0.01 , $\chi^2 / \text{d.o.f} = 6.07/6$. and a p-value of 0.42. These
¹⁹¹⁹ modified fit results are also included in Table 4.14.

¹⁹²⁰ Additionally, it is found that the best fit values for the slope terms of the linear fits in
¹⁹²¹ all three tables are of the order 10^{-4} , which corresponds to a percent level change per
¹⁹²² 100 GeV. However in all cases, the best fit values are fully compatible with zero (within
¹⁹²³ 1σ) once again with the exception detailed above, indicating that the level of closure is
¹⁹²⁴ indeed H_T independent.

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

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

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

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

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

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

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

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

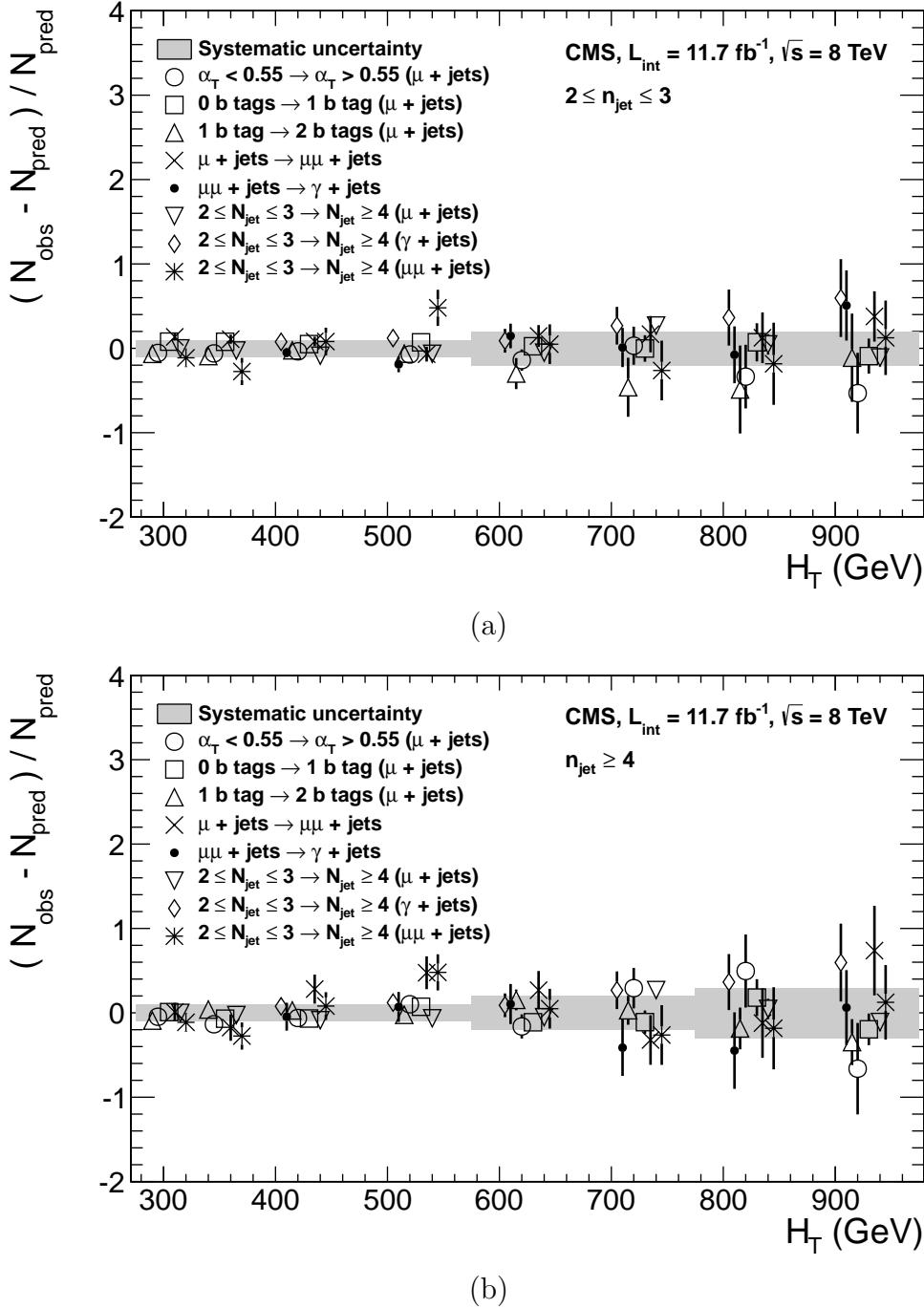


Figure 4.11: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity categories: (a) $2 \leq n_{\text{jet}} \leq 3$ and (b) $n_{\text{jet}} \geq 4$.

¹⁹⁴⁹ This point is emphasised when we examine the sensitivity of the TFs to a change in the
¹⁹⁵⁰ admixture of $W + \text{jets}$ and $t\bar{t}$ with the control and signal samples. This is accomplished
¹⁹⁵¹ by varying the cross-sections of the $W + \text{jets}$ and $t\bar{t}$ by +20% and -20%, respectively.

1952 Figures C.2 and C.3 within Appendix C, show the effect upon the closure tests for both
1953 jet multiplicity categories. Given these variations in cross-sections, the level of closure is
1954 found to be significantly worse, with biases as large as $\sim 30\%$, most apparent in the
1955 lowest H_T bins. However, the TFs used to extrapolate from control to signal are seen to
1956 change only at the percent level by this large change in cross-section, shown in Table C.1.

1957 Given the robust behaviour of the translation factors with respect to large (and opposite)
1958 variations in the $W + \text{jets}$ and $t\bar{t}$ cross-sections, one can assume with confidence that
1959 any bias in the translation factors is adequately (and conservatively) covered by the
1960 systematic uncertainties used in the analysis.

1961 4.7. Simplified Models, Efficiencies and Systematic 1962 Uncertainties

1963 The results of the analysis are interpreted using various SMS signal models, which as
1964 already introduced in Section (2.4.1) offer a natural starting point for quantifying and
1965 characterising SUSY signals, and a means to identify the boundaries of search sensitivity
1966 for different mass splittings, kinematic ranges, and final states.

1967 Each model is parameterised in a two dimensional parameter space, $(m_{\tilde{q}/\tilde{g}}, m_{\text{LSP}})$, from
1968 which upper limits on the production cross-sections of the various SMS models can be
1969 set.

1970 Each signal sample is generated at Leading Order (LO) with Pythia [85], and cross-
1971 sections calculated for Next to Leading Order (NLO) and Next to Leading Logarithmic
1972 Order (NLL) [86], with events simulated using the **Fastsim** framework. This framework
1973 represents a simplified simulation of the CMS detector, but allows for faster production
1974 of various signal topologies with different mass parameters.

1975 A series of correction factors are applied to account for differences between **Fastsim** [87]
1976 and **Fullsim** [88] simulation, which can affect the resultant n_b^{reco} distribution and which
1977 are detailed in Section (4.7.2).

¹⁹⁷⁸ **4.7.1. Signal Efficiency**

¹⁹⁷⁹ The analysis selection efficiency, ϵ , is measured for each mass point of the interpreted
¹⁹⁸⁰ model. This serves as a measure of the sensitivity of the signal selection for that particular
¹⁹⁸¹ sparticle, LSP mass and final state topology. The signal yield is then given by

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

¹⁹⁸² where σ represents the model's cross-section and \mathcal{L} the luminosity. An upper limit on σ
¹⁹⁸³ taken from theory can then allow for the setting of limits in terms of the particle mass.

¹⁹⁸⁴ Figure 4.12 shows the expected signal efficiency of the signal selection for the T1 and
¹⁹⁸⁵ T2 SMS models interpreted in this analysis. The efficiency maps are produced with the
¹⁹⁸⁶ requirement $H_T > 275$ GeV (i.e. no H_T categorisation) and requirements on n_{jet} and
¹⁹⁸⁷ n_b^{reco} are the most sensitive to the model in question.

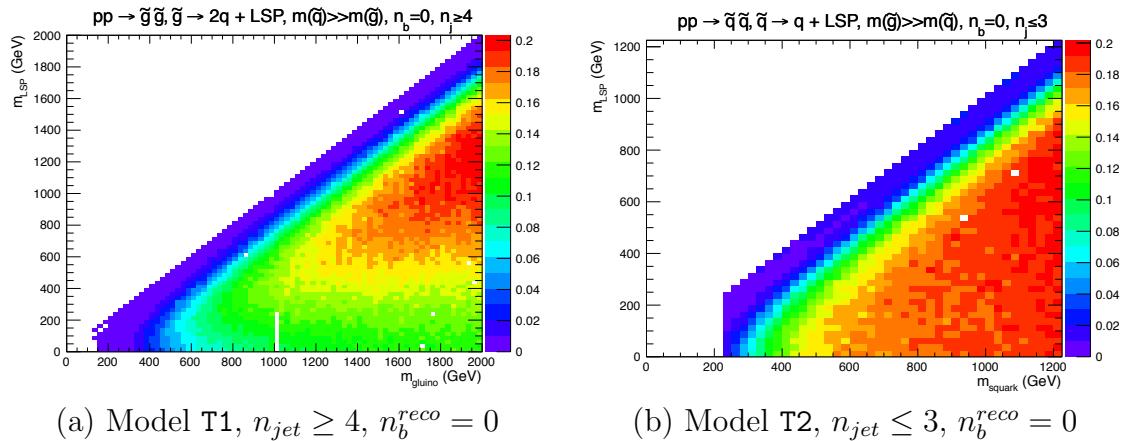


Figure 4.12: Signal efficiencies for the SMS models (a) T1 ($\tilde{g}\tilde{g}^* \rightarrow q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$) and (b) T2 ($\tilde{q}\tilde{q}^* \rightarrow q\tilde{\chi}_1^0 q\tilde{\chi}_1^0$) when requiring $n_{jet} \geq 4$ and ≤ 3 respectively, and $n_b^{reco} =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

¹⁹⁸⁸ The same procedure is conducted in the analysis control samples. It is found in the $\mu +$
¹⁹⁸⁹ jets control samples, that the signal-to-background ratios for the expected signal yields in
¹⁹⁹⁰ each of the SMS models are many times smaller than in the hadronic signal region. The
¹⁹⁹¹ relative contamination for the $\mu\mu +$ jets sample is smaller still due to the requirement of
¹⁹⁹² a second muon. The relative contamination for the $\gamma +$ jets sample is expected to be
¹⁹⁹³ zero for the models under consideration. These small, relative levels of contamination
¹⁹⁹⁴ are accounted for in the fitting procedure, as described in Section (4.8.4).

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

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

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

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

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

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

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

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

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

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

2022 For each jet, the weight of the event is re-weighted according to whether the jet fires the
2023 tagger. In the instance that the jet *is* tagged, the event weight will be modified by,

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

2024 and in the case that the jet does *not* fire the tagger,

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

2025 This procedure is applied to all events that pass the selection criteria, thus correcting
2026 the **FastSim** n_b^{reco} distribution to data.

2027 4.7.3. Experimental Uncertainties

2028 The systematic uncertainty on the expected signal acceptance \times analysis efficiency is
2029 determined independently for the each SMS model considered. These systematics stem
2030 from uncertainties on the parton distribution functions, the luminosity measurement, jet
2031 energy scale, b-tag scale factor measurements and the efficiencies of various selection
2032 criteria used in the signal selection, including the H_T / E_T , dead ECAL cleaning filter
2033 and lepton / photon event vetoes.

2034 Rather than trying to estimate the level of systematic that is applicable point-by-point
2035 in a model space, general behaviours are considered; and instead constant systematics
2036 are estimated in two regions of the SMS models parameter space.

2037 These two regions are defined as, near (small mass splittings) and far (large mass
2038 splittings) from the mass degenerate diagonal, where the far region is bounded by the
2039 condition

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

2040 The total systematics in each region are evaluated in the following ways:

2041 **Jet energy scale:** The relative change in the signal efficiency is gauged by varying
2042 the energy of all jets in an event up or down according to a p_T and η dependent jet
2043 energy scale uncertainty. Within the two systematic regions, the resulting systematic
2044 uncertainties for each SMS model are determined by taking the value of the 68th
2045 percentile for the distributions of the relative change in the signal efficiency.

2046 **Luminosity measurement:** The uncertainty on the measurement of the luminosity
2047 collected propagates through to an uncertainty on the signal event yield when
2048 considering any new physics model, which is currently 4.4% [89].

2049 **Parton density function :** Each signal sample is produced using the CTEQ6L1
2050 parton density function. The effect on the signal acceptance when re-weighting to
2051 the central value of three different parton distribution functions, CT10, MSTW08
2052 and NNPDF2.1 are examined [90]. It is found that the change of the signal efficiency
2053 in different SMS models, due to the alternate PDF sets are typically a few percent,
2054 and approaches 10% at higher squark/gluon and LSP masses.

2055 **$\mathcal{H}_T/\mathcal{E}_T$ cleaning filter:** The ratio of the efficiencies of the cleaning filter are
2056 compared in simulation and data after application of the $\mu +$ jets control sample
2057 selection. No α_T requirement or further event cleaning filters are applied. The
2058 ratio of the efficiencies observed in data and simulation for a cut value of $\mathcal{H}_T/\mathcal{E}_T <$
2059 1.25 and the two jet multiplicity categories, $2 \leq n_{jet} \leq 3$ and $n_{jet} \geq 4$ are $1.028 \pm$
2060 0.007 and 1.038 ± 0.015 respectively. These deviations are taken to represent the
2061 systematic uncertainty on the simulation modelling of this variable.

2062 **Dead ECAL cleaning filter:** The ratio of the efficiencies observed in data and
2063 simulation for this filter in the two jet multiplicity categories, $2 \leq n_{jet} \leq 3$ and
2064 $n_{jet} \geq 4$, are 0.961 ± 0.008 and 0.961 ± 0.009 , respectively. These deviations

2065 from unity are taken to represent the systematic uncertainties in the modelling in
2066 simulation of this filter.

2067 **Lepton and photon vetoes:** The uncertainty on the efficiency of the lepton and
2068 photon vetoes is determined by considering truth information. The efficiency of
2069 the vetoes is measured after applying relevant object filters with identical logic,
2070 but based on truth instead of reconstructed objects. Where the efficiency is found
2071 to not be 100%, it is taken to represent the fraction of signal events that are
2072 incorrectly vetoed. This deviation is taken directly as the systematic uncertainty on
2073 the efficiency. The systematic uncertainty is only non-zero for models which contain
2074 third-generation quarks in the final state, where the uncertainties are at the order
2075 of 1% level.

2076 **B-tag scale factor uncertainties:** The relative change in the signal efficiency
2077 is observed when relevant flavour, p_T and η dependant b-tag correction factors,
2078 are varied up or down by their systematic uncertainty. Within the two systematic
2079 regions, the resulting systematic uncertainties for each SMS model are determined
2080 by taking the value of the 68th percentile for the distributions of the relative change
2081 in the signal efficiency, over all mass points.

2082 Tables 4.16 and 4.17 summarise all the aforementioned systematic uncertainties on the
2083 signal efficiencies for each individual SMS model interpreted in the analysis. In the case
2084 of the T1tttt model, in which pair produced gluinos decay to $t\bar{t}$ pairs and the LSP,
2085 the near region of SMS space is not considered, and so no systematic uncertainties are
2086 included.

2087 In both of the defined regions it is found that the systematic uncertainties are relatively
2088 flat justifying the approach taken. The systematic uncertainties applied to the region
2089 near to the diagonal fall in the range 13-15%; similarly, for the region far from the
2090 diagonal the determined uncertainties are in the range of 12-23%. These uncertainties
2091 are all propagated through to the limit calculation.

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

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

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

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

2092 4.8. Statistical Interpretation

2093 For a given category of events satisfying requirements on both n_{jet} and n_b^{reco} , a likelihood
 2094 model of the observations in multiple data samples is used to gauge agreement between
 2095 the observed yields in the hadronic signal region, and the predicted yields obtained from
 2096 the control samples. In addition to checking whether the predictions are compatible
 2097 with a SM only hypothesis, the likelihood model is also used to test for the presence
 2098 of a variety of signal models. The statistical framework outlined within this section is
 2099 presented in greater detail within [91].

2100 4.8.1. Hadronic Sample

2101 Let N be the number of bins in H_T , with n^i the number of events observed satisfying all
 2102 selection requirements in each H_T bin i. The likelihood of the observations can then be
 2103 written:

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

2104 where b^i represents the expected SM background

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

2105 and s^i the expected number of signal events (see Section (4.8.4)) from the different SMS
2106 models interpreted. Pois refers to the Poisson distribution of these values and is defined
2107 as:

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

2108 **4.8.2. H_T Evolution Model**

2109 The hypothesis, that for a process the α_T ratio falls exponentially (see Section (4.2.4))
2110 in H_T is defined by Equation (4.10), where k_{QCD} is constrained by measurements in a
2111 signal sideband region.

2112 The expected QCD background, QCD^i , within a bin i is then modelled as,

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

2113 where m_i represent the number of events observed with $\alpha_T \leq 0.55$ in each H_T bin i , and
2114 $\langle H_T \rangle$ represent the mean H_T of each bin. Expressed as functions of just the zeroth bin,
2115 QCD^0 , and k_{QCD} , the QCD expectation is given by

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

2116 **4.8.3. EWK Control Samples**

2117 The EWK background estimation within each bin, i , is broken into two components, the
2118 expected yield from $Z \rightarrow \nu\bar{\nu}$ and $t\bar{t}$ -W (plus other residual backgrounds) events. This is
2119 written as, Z_{inv}^i and $t\bar{t}W^i$, and it follows that

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

2120 This can be further expressed as

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

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

2121 where f_{Zinv}^i represents the expected yield from $Z \rightarrow \nu\bar{\nu}$ in bin i divided by the expected
2122 EWK background EWK^i . This fraction is modelled as a linear component

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

2123 where N again represents the number of H_T bins, and f_{Zinv}^0 and f_{Zinv}^{N-1} are float parameters
2124 whose final values are limited between zero and one.

2125 Within each H_T bin there are three background measurements for the different control
2126 samples, n_γ^i , n_μ^i and $n_{\mu\mu}^i$, representing the event yields from the $\gamma +$ jets, $\mu +$ jets and
2127 $\mu\mu +$ jets control samples respectively. Each of these have a corresponding yield in
2128 simulation, MC_γ^i , MC_μ^i and $MC_{\mu\mu}^i$. Within the hadronic signal region there are also
2129 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
2130 used to define

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

2131 where r_p^i represents the inverse of the TFs used to extrapolate the yield of each background
2132 process.

2133 The likelihoods regarding the three measured yields n_γ^i , $n_{\mu\mu}^i$, n_μ^i can then be fully expressed
2134 as

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

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

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

$$(4.31)$$

2135 which contain an additional term s_μ^i , which represents the signal contamination in the
2136 $\mu + \text{jets}$ sample. The parameters $\rho_{\gamma Z}^j$, $\rho_{\mu\mu}^j$ and ρ_μ^j represent “correction factors” that
2137 accommodate the data driven systematic uncertainties derived from the control samples
2138 in Section (4.12).

2139 Each of these equations are used to estimate the maximum likelihood value for relevant
2140 background in the signal region given the observations n_p^i in each of the control samples
2141 (see Section (4.2.3)).

2142 The measurements in each of the control samples and the hadronic signal region, along
2143 with the ratios r_γ^i , $r_{\mu\mu}^i$, and r_μ^i , are all considered simultaneously through the relationships
2144 defined by Equations (4.19),(4.24) and (4.25).

2145 In addition to the Poisson product, an additional log-normal term is introduced to
2146 accommodate the systematic uncertainties given by,

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

2147 The parameters ρ^j , ρ^j and ρ^j represent the already introduced “correction factors”
2148 that accommodate the systematic uncertainties, while the quantities $\sigma_{\gamma Z}^j$, $\sigma_{\mu\mu Z}^j$ and

2149 $\sigma_{\mu W}^j$ represent the relative systematic uncertainties for the respective control sample
2150 constraints. Logn represents the log-normal distribution [92],

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

2151 Five parameters per control sample are used to span the eight H_T categories, with just
2152 one used for the three H_T in the $n_b^{reco} \geq 4$ category. These parameters span the same
2153 H_T ranges described in Section (4.6) and is shown in Table 4.18.

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

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

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

2154 Alternatively, in the higher n_b^{reco} categories ($n_b^{reco} = 2$ and above), only the single muon
2155 sample is used to constrain the total EWK background. This is due to a lack of statistics
2156 in the $\mu\mu +$ jets and $\gamma +$ jets at these n_b^{reco} multiplicities. Therefore the likelihood
2157 functions for the control samples are reduced and simply represented by

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

2158 where

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

2159 4.8.4. Contributions from Signal

2160 The cross-section for each model is represented by x , while l represents the total recorded
2161 luminosity considered by the analysis in the signal region. Let ϵ_{had}^i and ϵ_μ^i represent the
2162 analysis selection efficiency for that particular signal model in H_T bin i of the hadronic
2163 and $\mu +$ jets control sample respectively. Letting δ represent the relative uncertainty on

- 2164 the signal yield, assumed to be fully correlated across all bins, and ρ_{sig} the “correction
2165 factor” to the signal yield which accommodates this uncertainty. f represents an unknown
2166 multiplicative factor on the signal cross section, for which an allowed interval is computed.
2167 The expected signal yield s^i is thus given by

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

- 2168 and signal contamination with the $\mu +$ jets control sample by

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

- 2169 The systematic uncertainty on the signal is additionally incorporated by the term

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

- 2170 A discussion of the SMS signal models through which the analysis is interpreted can be
2171 found in the following chapter.

2172 4.8.5. Total Likelihood

- 2173 The total likelihood function for a given signal category $k(n_b^{\text{reco}}, n_{jet})$ is then given by
2174 the product of the likelihood functions introduced within the previous sections:

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

- 2175 In categories containing eight H_T bins and utilising the three control samples ($\mu +$ jets,
2176 $\mu\mu +$ jets, $\gamma +$ jets), there are 25 nuisance parameters. When just one control sample

2177 is used to estimate the EWK background, this is reduced to 15 nuisance parameters.

2178 In the $n_b^{\text{reco}} \geq 4$ category where only three H_T bins are used, there are just 6 nuisance

2179 parameters. This information is summarised within Table 4.19.

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

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

2180 When considering SUSY signal models within the likelihood, the additional L_{sig} term
2181 is included and therefore when multiple categories are fitted simultaneously the total
2182 likelihood is then represented by

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

Chapter 5.

²¹⁸³ Results and Interpretation

²¹⁸⁴ Using the statistical framework outlined in the previous chapter, results are shown for
²¹⁸⁵ the compatibility of the collected data with a SM-only hypothesis in Section (5.1). The
²¹⁸⁶ data is further interpreted within the context of various SMS models within Section (5.2).

²¹⁸⁷ 5.1. Compatibility with the Standard Model Hypothesis

²¹⁸⁹ The SM background only hypothesis is tested by removing any signal contributions
²¹⁹⁰ within the signal and control samples, and the likelihood function defined in Equation
²¹⁹¹ (4.39) maximised over all parameters using Rootfit [93] and MINUIT [94]. The results of
²¹⁹² the search consist of the observed yields in the hadronic signal sample, and the $\mu + \text{jets}$,
²¹⁹³ $\mu\mu + \text{jets}$ and $\gamma + \text{jets}$ control samples.

²¹⁹⁴ These observed yields along with the expectations and uncertainties given by the simula-
²¹⁹⁵ taneous fit for the hadronic signal region are displayed in Table 5.2. The results obtained
²¹⁹⁶ from the simultaneous fits, including that of the three control samples, are shown in
²¹⁹⁷ Figure 5.1-5.8, and which is summarised in Table 5.1.

²¹⁹⁸ The figures show a comparison between the observed yields and the SM expectations as
²¹⁹⁹ given by the fit across all H_T bins, and in all n_{jet} and n_b^{reco} multiplicity categories. In all
²²⁰⁰ categories the samples are well described by the SM only hypothesis. In particular no
²²⁰¹ significant excess is observed above SM expectation within the hadronic signal region.

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

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

| Cat | n_b^{reco} | n_{jet} | H_T bin (GeV) | | | | | | | |
|------------|--------------|-----------|-----------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| | | | 275-325 | 325-375 | 375-475 | 474-575 | 575-675 | 675-775 | 775-875 | 875- ∞ |
| SM Data | 0 | ≤ 3 | 6235^{+100}_{-67} 6232 | 2900^{+60}_{-54} 2904 | 1955^{+34}_{-39} 1965 | 558^{+14}_{-15} 552 | 186^{+11}_{-10} 177 | $51.3^{+3.4}_{-3.8}$ 58 | $21.2^{+2.3}_{-2.2}$ 16 | $16.1^{+1.7}_{-1.7}$ 25 |
| | | ≥ 4 | 1010^{+34}_{-24} 1009 | 447^{+19}_{-16} 452 | 390^{+19}_{-15} 375 | 250^{+12}_{-11} 274 | 111^{+9}_{-7} 113 | $53.3^{+4.3}_{-4.3}$ 56 | $18.5^{+2.4}_{-2.4}$ 16 | $19.4^{+2.5}_{-2.7}$ 27 |
| SM Data | 1 | ≤ 3 | 1162^{+37}_{-29} 1164 | 481^{+18}_{-19} 473 | 341^{+15}_{-16} 329 | $86.7^{+4.2}_{-5.6}$ 95 | $24.8^{+2.8}_{-2.7}$ 23 | $7.2^{+1.1}_{-1.0}$ 8 | $3.3^{+0.7}_{-0.7}$ 4 | $2.1^{+0.5}_{-0.5}$ 1 |
| | | ≥ 4 | 521^{+25}_{-17} 515 | 232^{+15}_{-12} 236 | 188^{+12}_{-11} 204 | 106^{+6}_{-6} 92 | $42.1^{+4.1}_{-4.4}$ 51 | $17.9^{+2.2}_{-2.0}$ 13 | $9.8^{+1.5}_{-1.4}$ 13 | $6.8^{+1.2}_{-1.1}$ 6 |
| SM Data | 2 | ≤ 3 | 224^{+15}_{-14} 222 | $98.2^{+8.4}_{-6.4}$ 107 | $59.0^{+5.2}_{-6.0}$ 58 | $12.8^{+1.6}_{-1.6}$ 12 | $3.0^{+0.9}_{-0.7}$ 5 | $0.5^{+0.2}_{-0.2}$ 1 | $0.1^{+0.1}_{-0.1}$ 0 | $0.1^{+0.1}_{-0.1}$ 0 |
| | | ≥ 4 | 208^{+17}_{-9} 204 | 103^{+9}_{-7} 107 | $85.9^{+7.2}_{-6.9}$ 84 | $51.7^{+4.6}_{-4.7}$ 59 | $19.9^{+3.4}_{-3.0}$ 24 | $6.8^{+1.2}_{-1.3}$ 5 | $1.7^{+0.7}_{-0.4}$ 1 | $1.3^{+0.4}_{-0.3}$ 2 |
| SM Data | 3 | ≤ 3 | $8.6^{+2.8}_{-0.8}$ 8 | $4.6^{+1.0}_{-0.9}$ 3 | $2.7^{+0.7}_{-0.7}$ 2 | $0.3^{+0.2}_{-0.1}$ 0 | $0.0^{+0.0}_{-0.0}$ 1 | $0.0^{+0.0}_{-0.0}$ 0 | $0.0^{+0.0}_{-0.0}$ 0 | $0.0^{+0.0}_{-0.0}$ 0 |
| | | ≥ 4 | $25.3^{+5.0}_{-4.2}$ 25 | $11.7^{+1.7}_{-1.8}$ 13 | $6.7^{+1.4}_{-1.2}$ 4 | $3.9^{+0.8}_{-0.8}$ 2 | $2.3^{+0.6}_{-0.6}$ 2 | $1.2^{+0.3}_{-0.4}$ 3 | $0.3^{+0.2}_{-0.1}$ 0 | $0.1^{+0.1}_{-0.1}$ 0 |
| SM Data | 4 | ≥ 4 | $0.9^{+0.4}_{-0.7}$ 1 | $0.3^{+0.2}_{-0.2}$ 0 | | | | $0.6^{+0.3}_{-0.3}$ 2 | | |

Table 5.2: Comparison of the measured yields in each H_T , n_{jet} and n_b^{reco} jet multiplicity bins for the hadronic sample with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit. Note that the $n_b^{reco} = 3$, $n_{jet} \leq 3$ category is not used in any interpretations within this section but is included for completeness.

²²⁰² Given the lack of an excess in data hinting at a possible supersymmetric signature within
²²⁰³ the data, interpretations are made on the production masses and cross-section of a range
²²⁰⁴ of SUSY decay topologies within the following section.

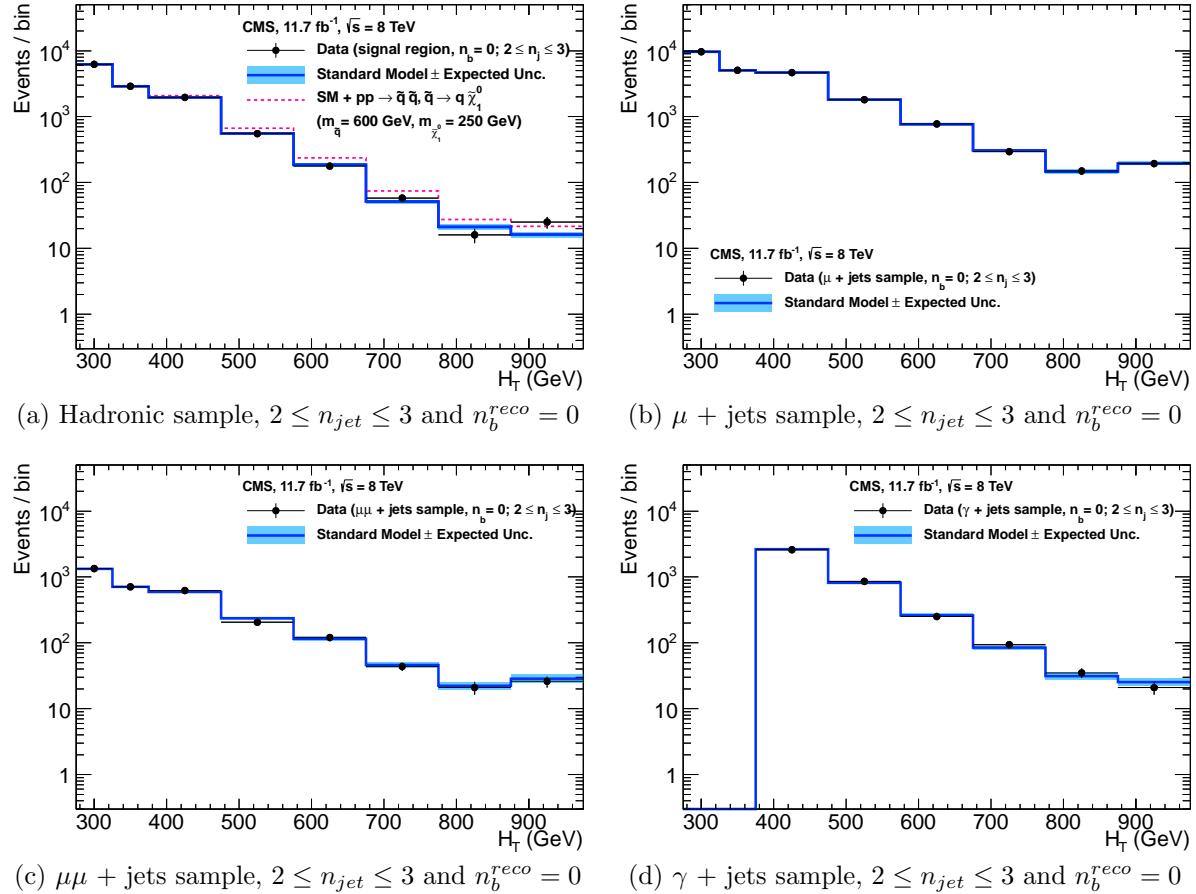


Figure 5.1: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \leq 3$. The observed event yields in data (black dots)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

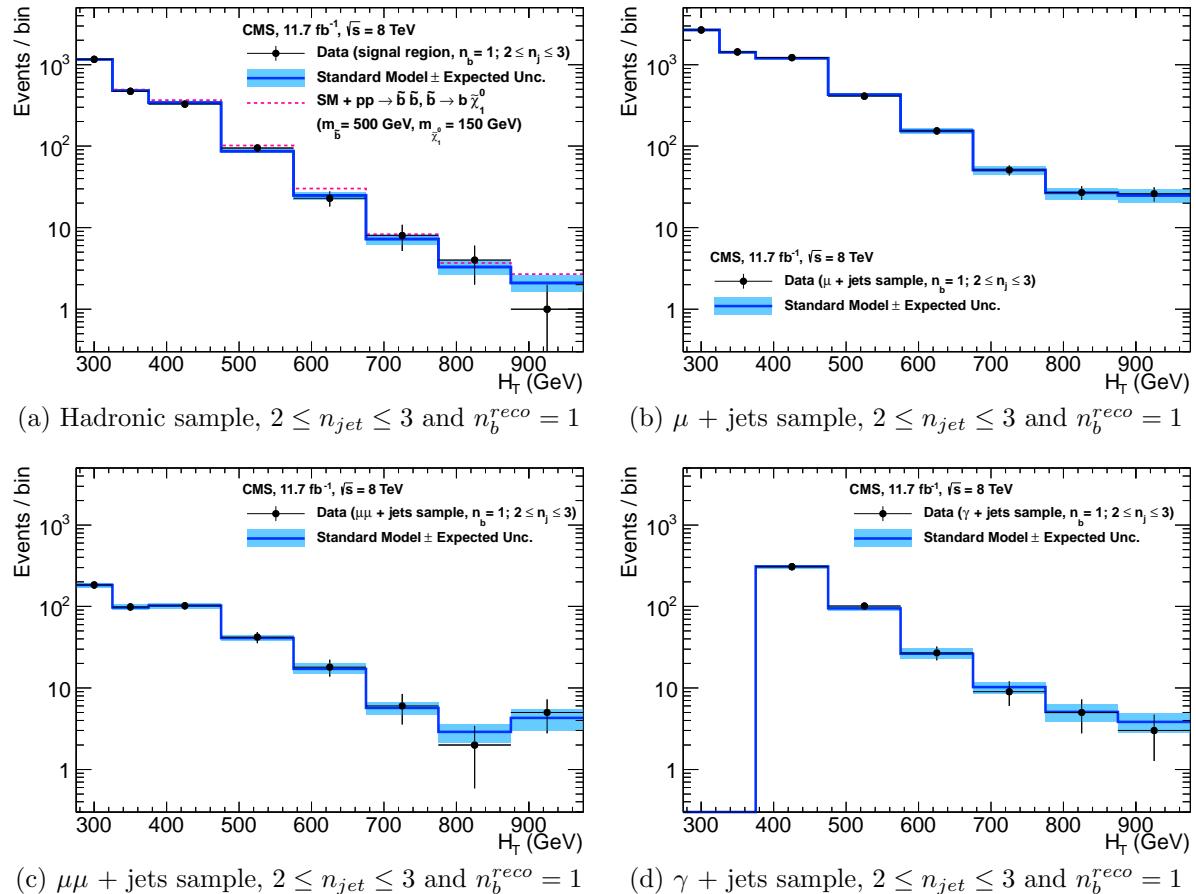


Figure 5.2: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$. The observed event yields in data (black dots)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

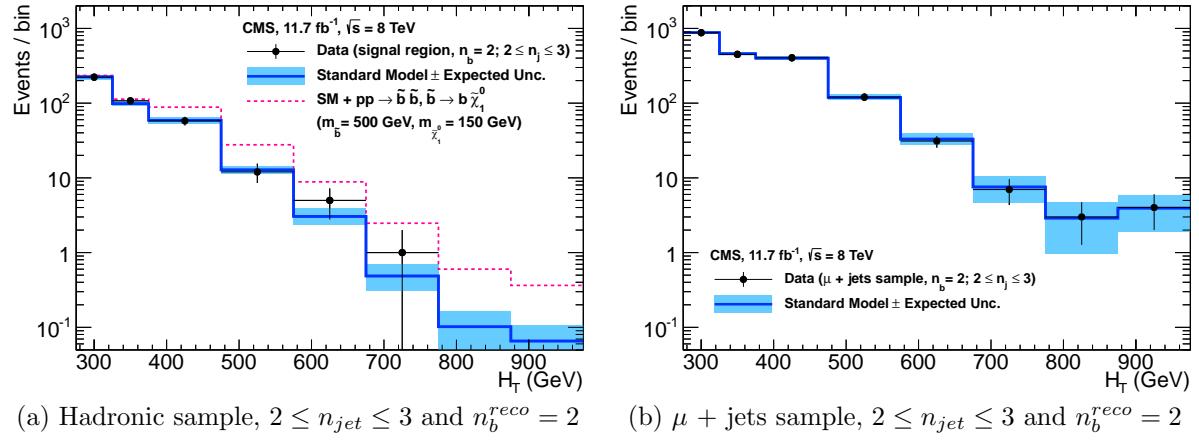


Figure 5.3: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$. The observed event yields in data (black dots)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

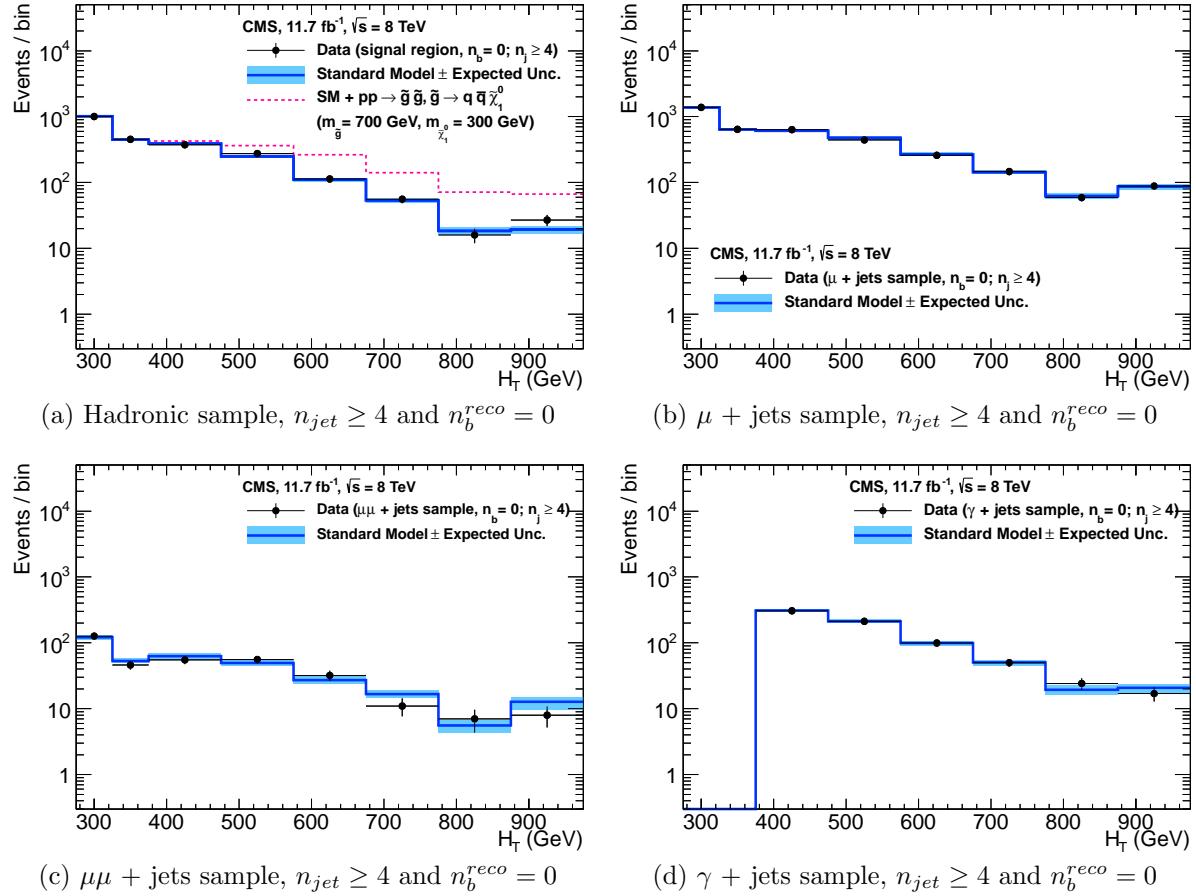


Figure 5.4: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

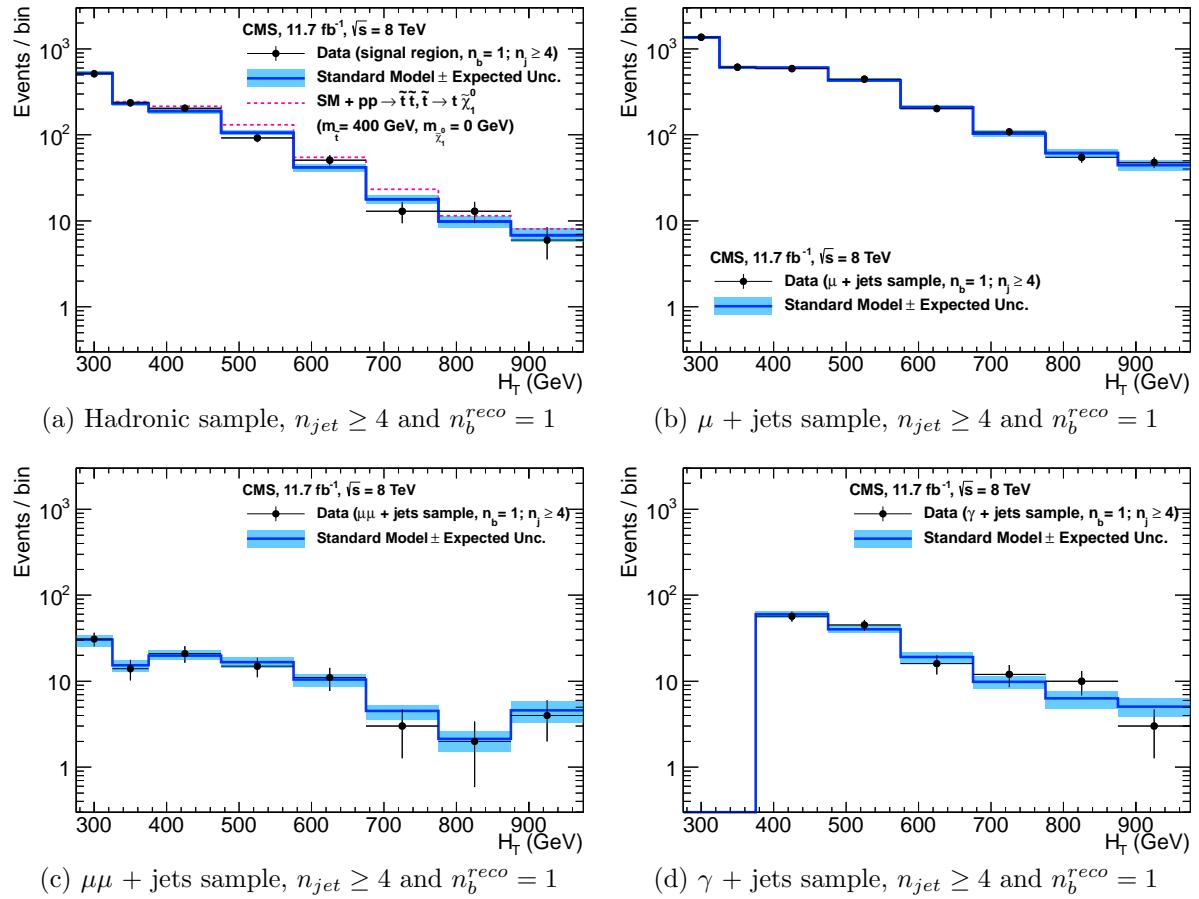


Figure 5.5: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

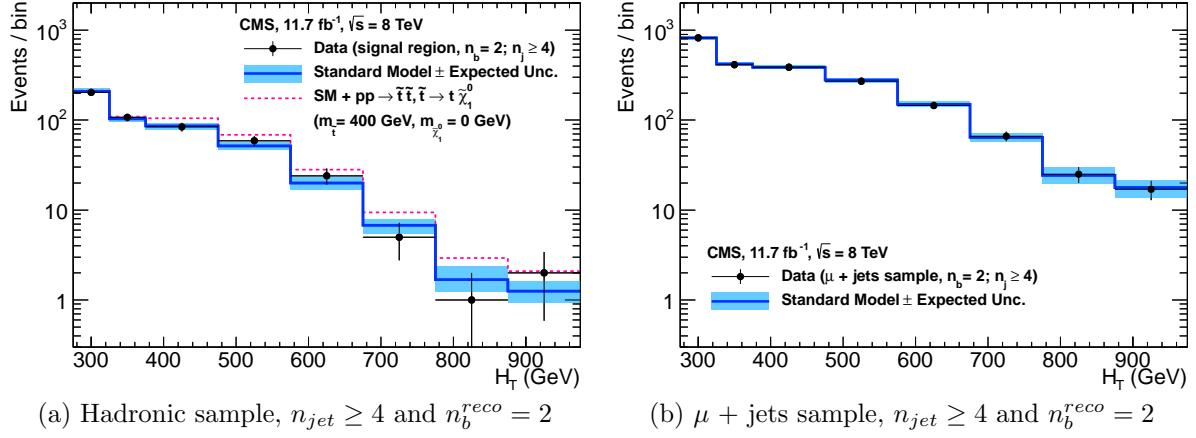


Figure 5.6: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

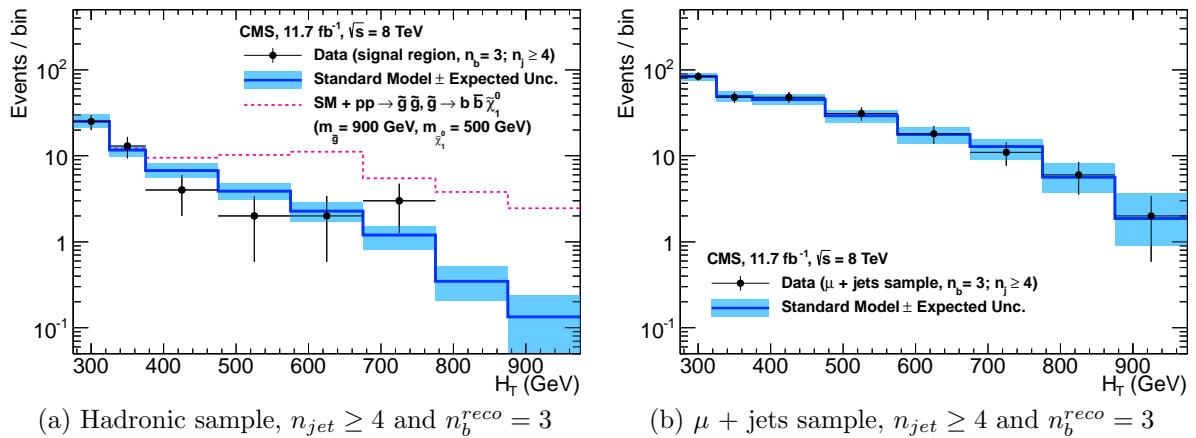


Figure 5.7: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$. The observed event yields in data (black dots)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

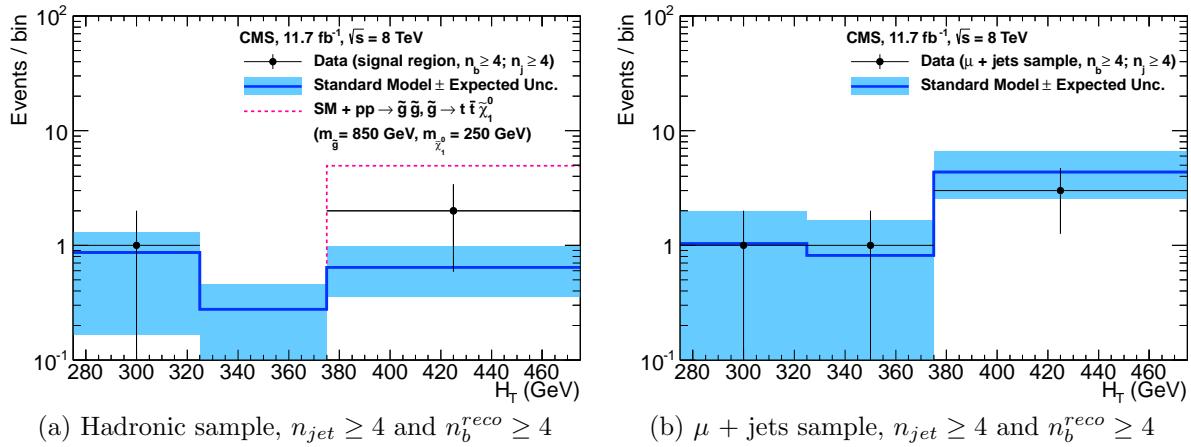


Figure 5.8: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) $\mu +$ jets, (c) $\mu\mu +$ jets and (d) $\gamma +$ jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$. The observed event yields in data (black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

2205 **5.2. SUSY**

2206 Limits are set on sparticle and LSP masses in the parameter space of a set of SMS models
2207 that characterise supersymmetric final states resulting from; direct third generation or
2208 light squark production, and gluino induced production of supersymmetric particles.
2209 However as detailed in Section (2.4.1), the individual models are not representative of a
2210 real physical SUSY model as only one decay process is considered. Instead these models
2211 represent a way to test for signs of specific signatures indicating new physics.

2212 **5.2.1. The CL_s Method**

2213 The CLs method [95][96][97] is used to compute the limits for signal models, with the
2214 one-sided profile likelihood ratio as the test statistic [98].

The test statistic is defined as

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

2215 where

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

2216 represents the profile likelihood ratio, in which $\mu \equiv f$ from Section (4.8.4), is the
2217 parameter characterising the signal strength. $\hat{\mu}$ is defined as the maximum likelihood
2218 value, $\hat{\theta}$ the set of maximum likelihood values of the nuisance parameters and θ_μ the set
2219 of maximum values of the nuisance parameters for a given value of μ .

2220 When $\mu \equiv f = 1$, the signal model is considered at its nominal production cross section.
2221 The distribution of q_μ is built up via the generation of pseudo experiments in order to
2222 obtain two distributions for the background (B) and signal plus background (S+B) cases.

2223 The compatibility of a signal model with observations in data is determined by the
2224 parameter CL_s ,

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

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

5.2.2. Interpretation in Simplified Signal Models

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

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

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

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

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

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

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

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

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

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

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

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

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

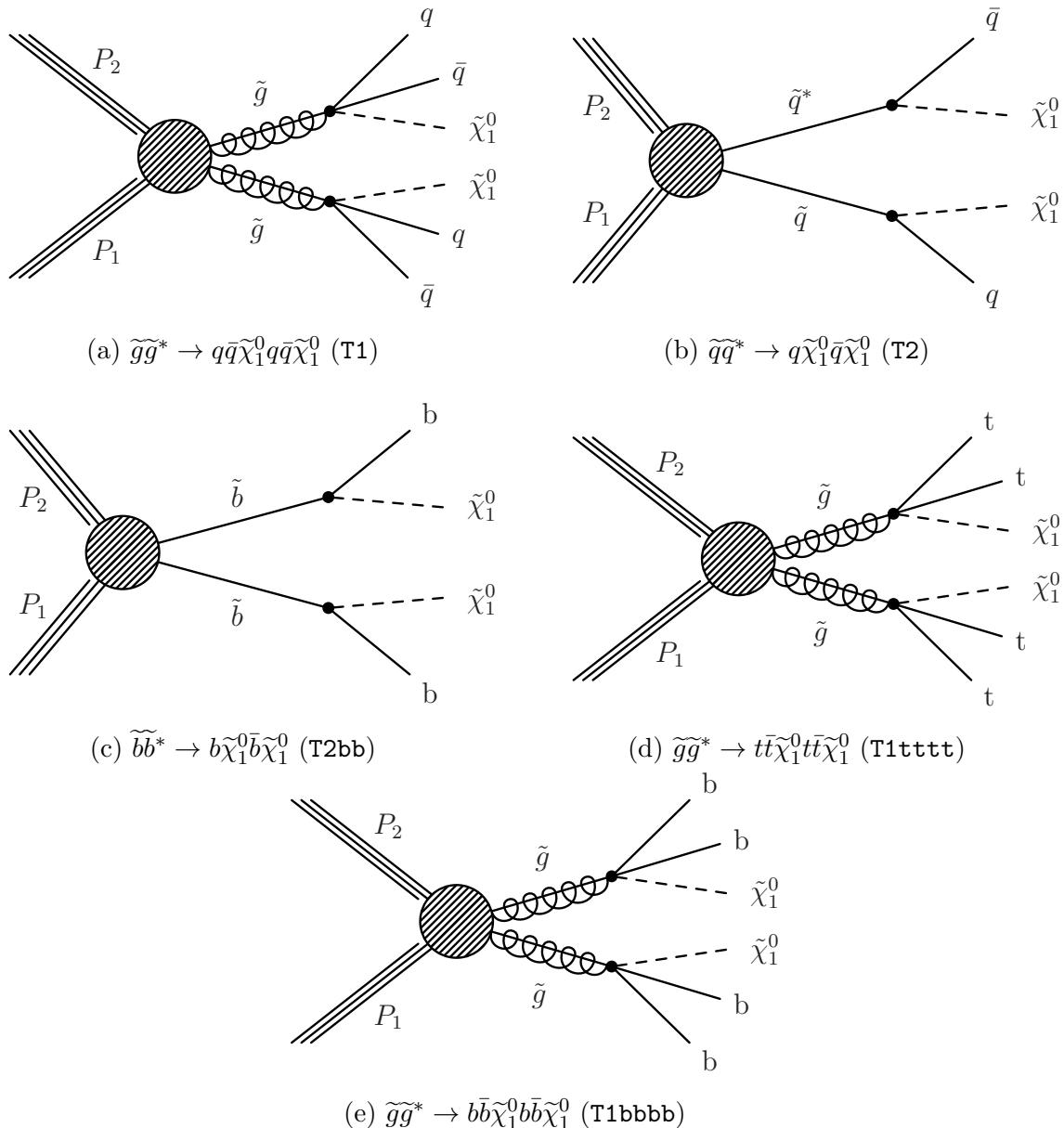


Figure 5.9: Production and decay modes for the various SMS models interpreted within the analysis.

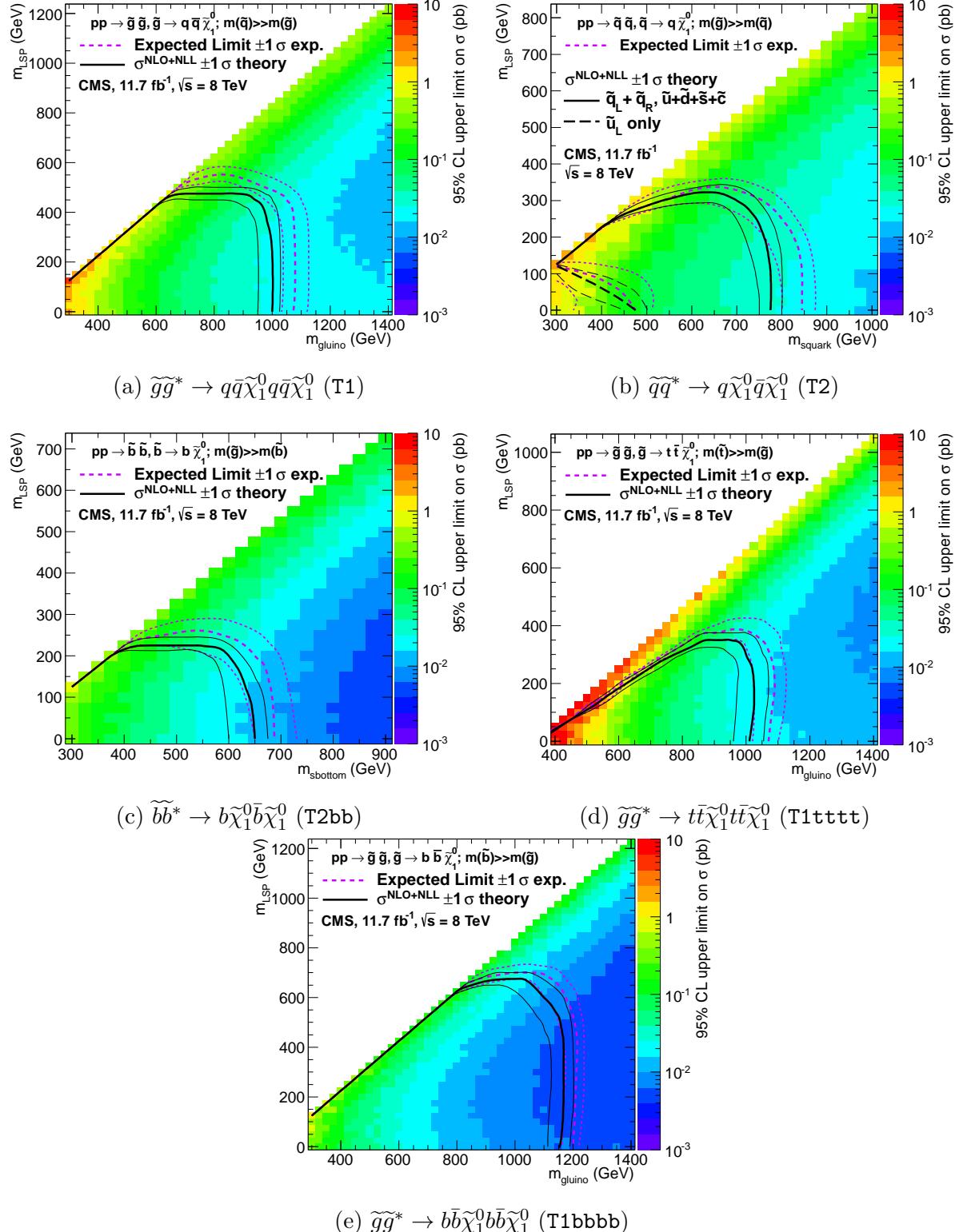


Figure 5.10: Upper limit of cross section at 95% CL as a function of $m_{\tilde{q}/\tilde{g}}$ and m_{LSP} for various SMS models. The solid thick black line indicates the observed exclusion region assuming NLO and NLL SUSY production cross section. The analysis selection efficiency is measured for each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

Chapter 6.

²²⁷⁹ SUSY Searches with B-tag

²²⁸⁰ Templates

²²⁸¹ Within this chapter a complementary technique is discussed as a means to predict the
²²⁸² distribution of three and four reconstructed b-tagged ($n_b^{\text{reco}} = 3, 4$), jets in an event
²²⁸³ sample. The recent discovery of the Higgs boson has made “Natural SUSY” models
²²⁸⁴ attractive, given that light top and bottom squarks are a candidate to stabilise divergent
²²⁸⁵ loop corrections to the Higgs boson mass. A light gluino which subsequently decays to
²²⁸⁶ third generation sparticle pairs, will give rise to many events with a large number of final
²²⁸⁷ state b-tagged jets.

²²⁸⁸ The method described within this chapter is used to estimate the SM background at
²²⁸⁹ high b-tagged jet multiplicities (3-4), from a templated fit conducted in a low b-tagged
²²⁹⁰ jet (0-2) control region of an event sample. This approach can hypothetically be applied
²²⁹¹ to generic supersymmetric searches, to gain sensitivity to signals which contain a higher
²²⁹² number of b-tagged jets than the search’s dominant SM backgrounds.

²²⁹³ As a proof-of-concept, the procedure is applied to the SM enriched $\mu + \text{jets}$ control
²²⁹⁴ sample of the α_T search detailed in Chapter 4, and validated in both data and simulation.
²²⁹⁵ This method is then further utilised to provide an independent crosscheck of the SM
²²⁹⁶ background estimations determined by the α_T search within its hadronic signal region at
²²⁹⁷ high b-tagged jet multiplicities.

²²⁹⁸ To highlight the relative insensitivity of this method to the choice of b-tagging algorithm
²²⁹⁹ working point, results are presented using the CSV tagger (introduced in Section (3.3.2))
²³⁰⁰ for the “Loose”, “Medium” and “Tight” working points.

2301 6.1. Concept

2302 The dominant SM backgrounds of most SUSY searches are typically $t\bar{t} + \text{jets}$, $W + \text{jets}$,
 2303 $Z \rightarrow \nu\bar{\nu} + \text{jets}$ or other rare processes (e.g. Diboson, $t\bar{t}W + \text{jets}$ production in the case of
 2304 hadronic searches) with neutrinos in the final state. These processes are characterised by
 2305 typically having zero or two underlying b-quarks per event as shown in Table 6.1. This
 2306 ultimately means that the resultant shape of the n_b^{reco} distribution for these two types of
 2307 event topologies will differ significantly due to the varying tagging probabilities of the
 2308 different jet flavours present in the final state of these processes.

2309 Similarly, SMS models comprising the gluino-mediated production of third generation
 2310 squarks, such at the T1tttt and T1bbbb models described in the previous chapter, will
 2311 contain four underlying b-quarks in its decay. Therefore the resultant shape of the
 2312 n_b^{reco} distribution from such a signal will be further skewed towards a higher number of
 2313 b-tagged jets. As SM processes with a similarly large number of underlying b-quarks
 2314 are rare, a signal indicative of natural SUSY can potentially be easily identified, via an
 2315 observed excess of $n_b^{\text{reco}} = 3, \geq 4$ events with respect to the expected yields from SM
 2316 processes.

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

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

2317 Within a supersymmetric or indeed any search for new physics, the compatibility of the
 2318 n_b^{reco} distribution in data with SM expectations can be tested, via the shape parameterisation
 2319 of the SM background n_b^{reco} distribution, grouped in terms of these two most
 2320 common underlying b-quark topologies.

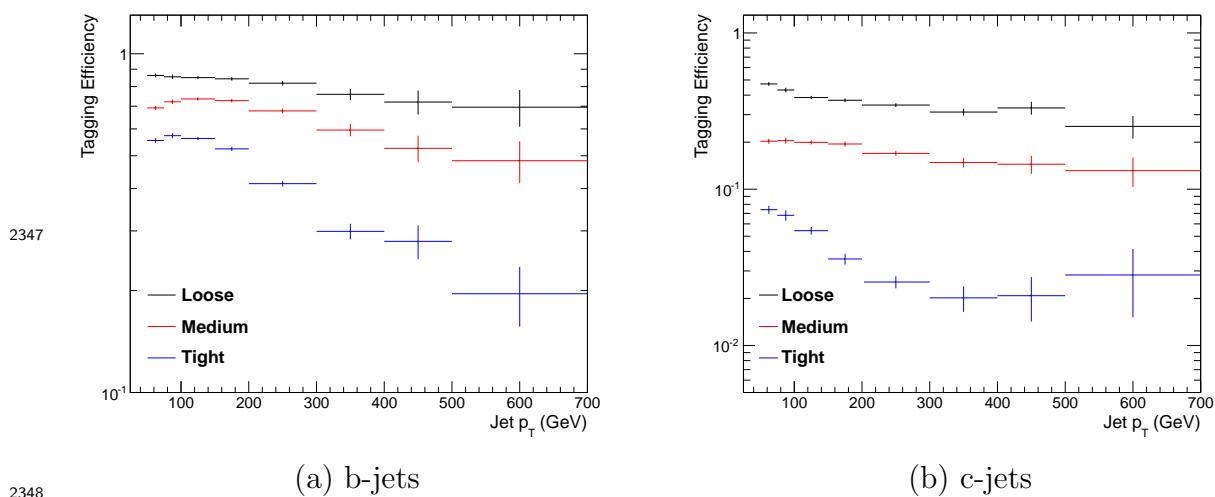
2321 6.1.1. Fitting Procedure

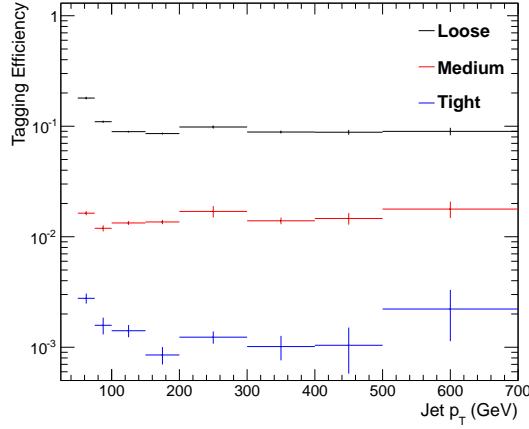
2322 Two templates, representing processes which have an underlying b-quark content of
 2323 zero or two are defined as Z0 and Z2 respectively (single top processes are a negligible

2324 background, $< 1\%$, within the α_T search to which this method is applied in the following
2325 section, and are thus incorporated within the Z2 template). SM background estimates at
2326 high n_b^{reco} multiplicities can then be extrapolated from the fitting of these two template
2327 shapes in a low n_b^{reco} control region (0-2) under the assumption of negligible signal
2328 contamination.

2329 The simplest way to determine the shapes of the n_b^{reco} distributions for both templates
2330 would be, after the application of the relevant event selection, to take the n_b^{reco} distribution
2331 as given directly from simulation. However as discussed within Section (4.5), there are
2332 large statistical uncertainties in simulation at high n_b^{reco} multiplicities (which is the region
2333 in which we wish to use the templates to estimate the SM backgrounds). This statistical
2334 uncertainty is particularly pronounced for processes incorporated within the Z0 templates,
2335 where events with a large number reconstructed b-tagged jets stem largely from the
2336 mis-tagging of all the light-flavoured jets in the final state. Therefore to improve the
2337 statistical precision of the final background prediction at high b-tagged jet multiplicities,
2338 the formula method first introduced in Section (4.5.1) is utilised to generate the template
2339 shapes.

2340 The template shapes of each analysis category (H_T and n_{jet} in the case of the α_T analysis)
2341 are dependant upon the jet-flavour content and their tagging efficiencies within the phase
2342 space of interest, with the tagging efficiency of a jet being a function of the jet p_T , the
2343 pseudo-rapidity $|\eta|$, and jet-flavour. This can be seen in Figure 6.1, where the tagging
2344 efficiency of jets identified as stemming from the hadronisation of a b, c or light quark
2345 from truth information in simulation, is shown for the three working points of the CSV
2346 tagger as a function of jet p_T .





(c) light-jets

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

2349 Therefore, before the template shapes are generated via the formula method, the jet
 2350 p_T and η averaged tagging efficiencies of each jet flavour are determined within each
 2351 individual analysis category. Additionally, the relevant jet p_T and η corrections are
 2352 applied to correct the measured b-tagging rate in simulation to that of data, as specified
 2353 in Section (4.5.3). These corrections propagate through to the average determined tagging
 2354 efficiency for each jet flavour, consequently affecting the final Z0 and Z2 template shape
 2355 of the n_b^{reco} distribution, determined within each analysis category.

2356 Using the truth-level flavour information of each of the defined Z0 and Z2 templates and
 2357 the measured tagging efficiencies of each jet flavour, the template shapes are constructed
 2358 from simulation via the formula method. These two shapes are then fitted to data in a
 2359 low n_b^{reco} control region (0-2), by allowing the normalisation constants θ_{Z0} and θ_{Z2} of the
 2360 two templates to float. The fits are performed independently within each of the defined
 2361 analysis category to remove any dependence on the modelling of jet multiplicity between
 2362 simulation and data. Best fit values of θ_{Z0} and θ_{Z2} are used, along with the fixed shape
 2363 of each template, to extrapolate a SM background estimation within the high n_b^{reco} signal
 2364 region (3,4) as shown in Figure 6.2.

2365 In deriving the uncertainty on the background prediction the following statistical uncer-
 2366 tainties are considered;

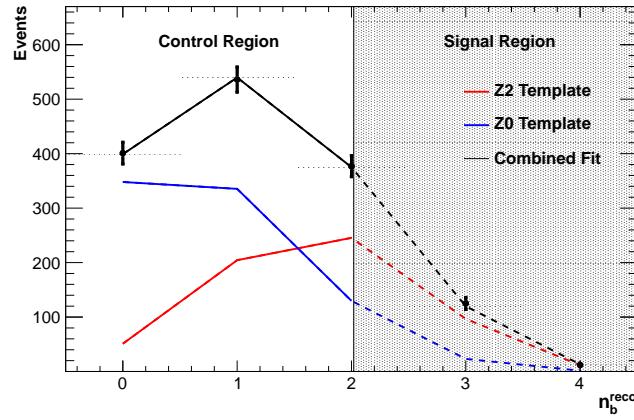


Figure 6.2: An example of a template fit with the defined Z0 (blue) and Z2 (red) templates to data within the low n_b^{reco} control region (left). The shape of the two templates are fixed but the normalisations θ_{Z0} and θ_{Z2} are allowed to vary. The best fit values are then applied to extrapolate a combined background prediction from the shaded signal region (right), represented by the dashed black line. Statistical and systemic uncertainties are not shown within this figure.

2367 **Fit uncertainty:** The statistical uncertainty on the normalisation factors θ_{Z0} and
 2368 θ_{Z2} as determined by the fit to data.

2369 **Measured tagging efficiency uncertainty:** The uncertainty of the template
 2370 shapes due to the uncertainty on the measured average tagging efficiencies of each
 2371 jet flavour from simulation. This uncertainty is propagated through to the template
 2372 prediction for each n_b^{reco} multiplicity by profiling the distribution of the θ_{Z0} and θ_{Z2}
 2373 best-fit values from multiple pseudo-experiments.

2374 For each pseudo-experiment, a Z0 and Z2 template shape is generated and fitted to
 2375 data. The tagging efficiencies of each jet flavour used by the formula method to
 2376 generate the template shape, are determined from a Gaussian distribution centred
 2377 on the nominal measured efficiency with a width equal to its measured statistical
 2378 uncertainty. The uncertainties on the nominal θ_{Z0} and θ_{Z2} normalisation factors
 2379 are then determined from the value of the 68th percentile in the best fit θ_{Z0} and
 2380 θ_{Z2} distributions constructed from all of the pseudo-experiments.

2381 **Formula method statistical error:** The statistical uncertainties of the two
 2382 templates Z0 and Z2 at each n_b^{reco} multiplicity are propagated through to the overall
 2383 uncertainty. This is due to the finite amount of simulated events used in the formula
 2384 method to generate the template shapes.

B-tag scale factor systematic error: When this procedure is applied to data, an additional systematic error is also incorporated into the template uncertainty. This takes into account the uncertainty in correcting the tagging efficiencies measured in simulation to data as first shown in Figures 3.7 and 3.8. The systematic uncertainty for each template is determined by varying these simulation to data scale factors ($SF_{b, c, \text{light}}$), up and down by their systematic uncertainties. These scale factor uncertainties are conservatively taken as fully correlated across all jet flavours [77]. The resultant relative difference due to these variations in the template shape at each n_b^{reco} multiplicity of the template, is taken as the systematic uncertainty on the nominal best fit template value.

All statistical and systematic errors are added in quadrature to determine an overall template fit uncertainty at each n_b^{reco} multiplicity in the control and signal regions. These are represented in all figures by a shaded grey band.

Any large excess in data is an indication that the n_b^{reco} distribution is not adequately described by the SM backgrounds encapsulated by the templates. This could mean there are additional SM backgrounds that fall within the selection of the analysis that need to be considered, or that there is signal present within the data. This method relies solely on fitting to the shape of the n_b^{reco} distribution, and can in principle, be applied to any analysis where the signal hypothesis has a larger underlying b-quark spectra than the SM backgrounds.

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

6.2. Application to the α_T Search

As detailed in the previous chapter, the α_T analysis is a search for supersymmetric particles in all hadronic final states, utilising the kinematic variable α_T to suppress QCD

2416 to a negligible level. SM enriched control samples are used to estimate the background
2417 within a hadronic signal region.

2418 The selection for the $\mu +$ jets control samples defined in Section (4.2.3) is used to
2419 demonstrate the template fitting procedure both conceptually in simulation, and also
2420 when applied in data. This is chosen, as such a selection is dominated by events stemming
2421 from the SM processes with little or no signal contamination from potential new physics.
2422 Contributions from rare SM processes with a higher underlying b-quark content (e.g.
2423 $t\bar{t}b\bar{b}$) are also found to be negligible from studies in simulation. For these reasons, there
2424 is a degree of confidence that the procedure should adequately describe the observations
2425 in data when extrapolated to the signal region.

2426 As a departure from the α_T search strategy described in the previous section, events are
2427 categorised according to jet multiplicity categories of 3, 4 and ≥ 5 reconstructed jets
2428 per event (di-jet events are not included as there is no contribution to the high n_b^{reco}
2429 region (3,4)), in order to reduce the kinematic range of the jet p_T 's within each category.
2430 Furthermore the analysis is split into just three H_T regions, for the purpose of increasing
2431 statistics within the control region,

2432 • 275-325 GeV

2433 • 325-375 GeV

2434 • > 375 GeV

2435 contrary to the eight used within the α_T analysis. Templates for both underlying b-quark
2436 content hypotheses are then generated for the nine defined event categories.

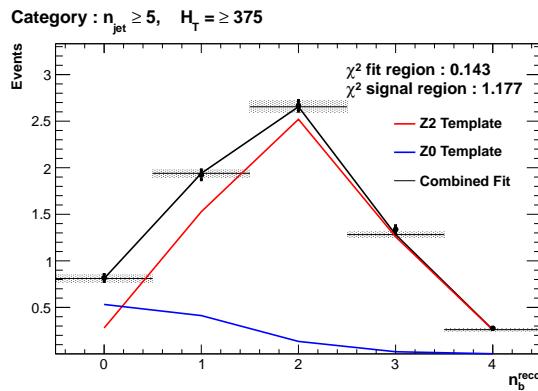
2437 6.2.1. Proof of Principle in Simulation

2438 This template procedure must be first demonstrated to work within simulated events free
2439 from any potential signal contamination before it can be applied to data. By combining
2440 the relevant ingredients necessary to employ the formula method, n_b^{reco} shape templates
2441 are generated individually for each n_{jet} and H_T category using one half of the available
2442 simulated events for each SM process. In this case, as the template shapes are being
2443 fitted to simulation, it is *not* necessary to apply the relevant corrections of the b-tagging
2444 rates between data and simulation.

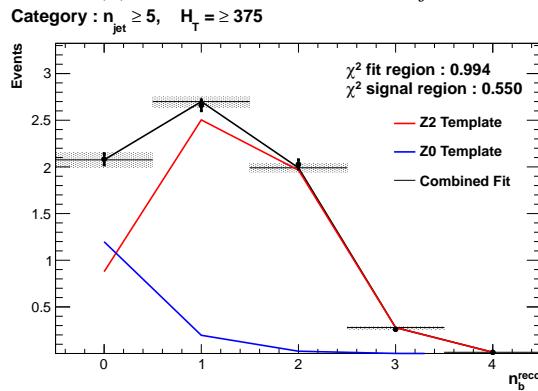
2445 The other half of simulated events is utilised to provide a statistically independent sample
 2446 from which the n_b^{reco} distribution is taken directly. The two generated templates are
 2447 then fit within the low n_b^{reco} (0-2) control region to this pseudo-data, from which a signal
 2448 region prediction is then extrapolated from the template best fit values.

2449 The aim of this procedure is to ensure that the template fit can accurately extrapolate the
 2450 n_b^{reco} distribution within the defined signal region from two independent but kinematically
 2451 identical samples. Furthermore, as the pseudo-data of the n_b^{reco} distribution is taken
 2452 directly from simulation, observation of good closure for both the initial fit of the two
 2453 templates within the control region and after extrapolation to the signal region will serve
 2454 as a validation of the formula method in recovering the original n_b^{reco} distribution itself.

2455 Results are presented in Figure 6.3 for each CSV working point in the $n_{\text{jet}} \geq 5$ category,
 2456 using the $\mu + \text{jets}$ control sample selection and the inclusive $H_T > 375$ GeV analysis bin.
 2457 Additional fit results for other n_{jet} categories which show a similar level of closure can
 2458 be found within Appendix D.1. The grey bands represent the statistical uncertainty of
 2459 the template prediction at each n_b^{reco} multiplicity derived from adding in quadrature the
 2460 statistical uncertainties introduced in the previous section.



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



(b) Medium working point : $n_{\text{jet}} \geq 5$

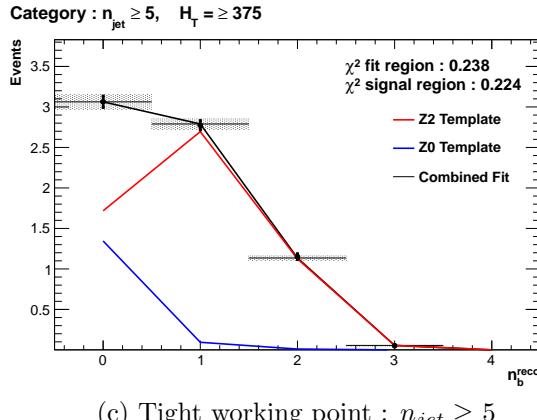
(c) Tight working point : $n_{jet} \geq 5$

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

2462 The extrapolated fit predictions summed over all n_{jet} multiplicities within the high n_b^{reco}
 2463 signal region, are summarised for all H_T bins and working points in Table 6.2.

| H_T | | 275-325 | 325-375 | >375 |
|----------------------|-----------|------------------|------------------|------------------|
| Loose working point | | | | |
| Simulation | $n_b = 3$ | 786.4 ± 14.7 | 392.7 ± 10.3 | 802.2 ± 14.4 |
| Template | | 789.6 ± 27.5 | 375.6 ± 16.6 | 770.1 ± 22.9 |
| Simulation | $n_b = 4$ | 67.4 ± 3.9 | 28.2 ± 2.7 | 93.7 ± 4.9 |
| Template | | 64.5 ± 5.9 | 26.4 ± 3.3 | 82.3 ± 5.8 |
| Medium working point | | | | |
| Simulation | $n_b = 3$ | 134.2 ± 5.8 | 74.4 ± 4.5 | 161.9 ± 6.3 |
| Template | | 129.9 ± 6.6 | 68.3 ± 4.8 | 159.9 ± 7.7 |
| Simulation | $n_b = 4$ | 1.5 ± 0.4 | 0.6 ± 0.4 | 3.1 ± 0.6 |
| Template | | 1.7 ± 0.3 | 0.9 ± 0.3 | 3.9 ± 0.6 |
| Tight working point | | | | |
| Simulation | $n_b = 3$ | 28.1 ± 2.7 | 13.9 ± 1.9 | 29.2 ± 2.7 |
| Template | | 25.9 ± 2.0 | 12.2 ± 1.5 | 28.3 ± 2.4 |
| Simulation | $n_b = 4$ | 0.5 ± 0.4 | - | 0.2 ± 0.2 |
| Template | | 0.1 ± 0.1 | 0.1 ± 0.1 | 0.2 ± 0.1 |

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

2464 The pull distributions for all the fits performed can be found in Appendix D.2, and are
2465 compatible with a mean of zero and standard deviation of one, showing no obvious bias
2466 to the fitting procedure. Each of the fits performed show good compatibility between
2467 the template shapes and data from simulation within the defined control region, with
2468 additional good overall agreement also observed for extrapolation to the signal region as
2469 shown in Table 6.2. This validates both the formula method used in the generation of
2470 the template shapes as well as the method of predicting the SM background in the high
2471 n_b^{reco} signal region.
2472 The application of this method to the same selection in a data control sample is now
2473 used to demonstrate necessary control over the efficiency and mis-tagging rates when
2474 b-tagging scale factors are applied, and to test the assumption of no signal contamination
2475 with the $\mu + \text{jets}$ control sample.

2476 6.2.2. Results in a Data Control Sample

2477 The procedure is now applied to the 2012 8 TeV dataset in the $\mu + \text{jets}$ control sample, to
2478 establish the validity of this method in data. The relevant data to simulation b-tagging
2479 scale factors are applied to produce corrected values of the efficiency and mis-tagging
2480 rates within each analysis category [77].
2481 Figure 6.4 shows the results of the templates derived from simulation to each of the three
2482 defined H_T bins, in the $n_{\text{jet}} \geq 5$ category for the medium working point CSV tagger (the
2483 same working point used within the α_T analysis). Grey bands represent the previously
2484 detailed statistical uncertainty of the fit combined in quadrature with the systematic
2485 uncertainties of varying up and down the simulation to data scale factors by their b-tag
2486 scale factor systematic uncertainties. Additional fit results for other jet multiplicities are
2487 found in Appendix D.3.

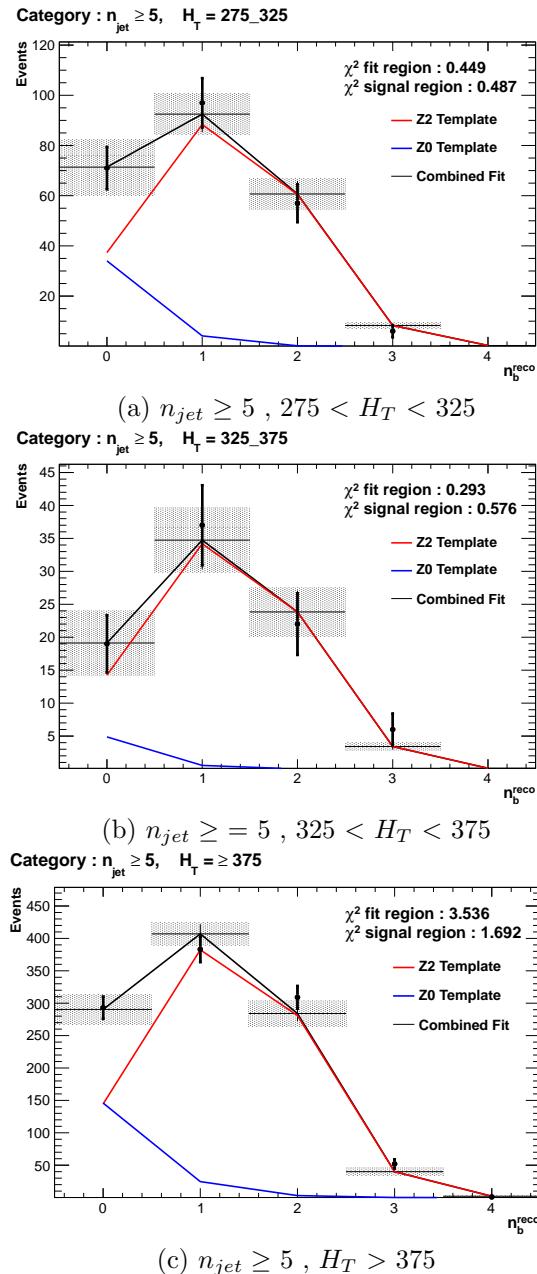


Figure 6.4: Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the $\mu + \text{jets}$ control sample, for the CSV medium working point, with $n_{jet} \geq 5$ in each H_T category. Data is represented by the black circles with the blue, red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

2488 The numerical results and extrapolation to the $n_b^{reco} = 3, 4$ bins for all H_T and working
2489 points, is shown in Table 6.3.

| H_T | | 275-325 | 325-375 | >375 |
|----------------------|-----------|------------------|------------------|------------------|
| Loose working point | | | | |
| Data | $n_b = 3$ | 838 | 394 | 717 |
| Template | | 871.8 ± 46.9 | 369.9 ± 23.7 | 678.5 ± 42.5 |
| Data | $n_b = 4$ | 81 | 43 | 81 |
| Template | | 79.4 ± 9.9 | 32.9 ± 4.2 | 74.4 ± 10.0 |
| Medium working point | | | | |
| Data | $n_b = 3$ | 137 | 79 | 152 |
| Template | | 132.6 ± 9.3 | 69.8 ± 5.4 | 133.1 ± 10.8 |
| Data | $n_b = 4$ | 1 | 1 | 3 |
| Template | | 1.9 ± 0.4 | 0.9 ± 0.3 | 3.2 ± 0.6 |
| Tight working point | | | | |
| Data | $n_b = 3$ | 24 | 15 | 25 |
| Template | | 22.3 ± 1.9 | 12.1 ± 1.2 | 20.3 ± 2.4 |
| Data | $n_b = 4$ | 0 | 0 | 1 |
| Template | | 0.1 ± 0.1 | 0.1 ± 0.1 | 0.2 ± 0.1 |

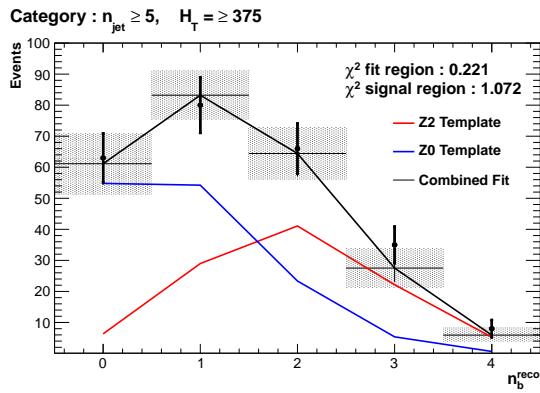
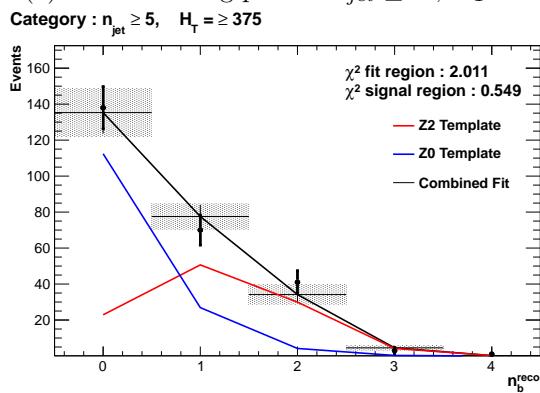
Table 6.3: Summary of the fit predictions in the n_b^{reco} signal region of the $\mu +$ jets control sample, after combination of the $n_{jet} = 3, = 4, \geq 5$ categories. The predictions are extrapolated from a $n_b^{reco} = 0, 1, 2$ control region using 11.4 fb^{-1} of $\sqrt{s} = 8\text{TeV}$ data.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

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

6.2.3. Application to the α_T Hadronic Search Region

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

2505 Figure 6.5 shows the results of the template shapes derived from simulation and
 2506 fitted to data for each of the three CSV working points, in the $n_{jet} \geq 5, H_T > 375$
 2507 GeV category. Grey bands represent the statistical uncertainty of the fit combined in
 2508 quadrature with the systematic uncertainties of varying the simulation to data scale
 2509 factors up and down by their measured systematic uncertainties. Additional fit results
 2510 for other jet multiplicities are found in Appendix D.4.

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

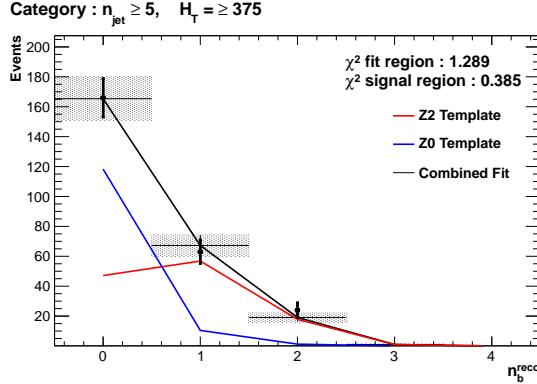
(c) Tight working point : $n_{jet} \geq 5, H_T > 375$

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

| H_T | | 275-325 | 325-375 | >375 |
|----------------------|-----------|----------------------|----------------------|----------------------|
| Loose working point | | | | |
| Data | | 198 | 85 | 126 |
| Template | $n_b = 3$ | 207.1 ± 28.7 | 103.4 ± 12.2 | 124.98 ± 14.4 |
| Data | | 15 | 9 | 16 |
| Template | $n_b = 4$ | 15.9 ± 5.4 | 8.05 ± 2.1 | 13.1 ± 3.2 |
| Medium working point | | | | |
| Data | | 33 | 16 | 14 |
| Template | $n_b = 3$ | 25.4 ± 4.0 | 12.7 ± 2.2 | 19.9 ± 2.9 |
| α_T ML Fit | | $33.9^{+5.7}_{-4.3}$ | $16.3^{+1.9}_{-2.0}$ | $17.5^{+1.4}_{-1.4}$ |
| Data | | 1 | 0 | 2 |
| Template | $n_b = 4$ | 0.3 ± 0.2 | 0.3 ± 0.1 | 0.5 ± 0.2 |
| α_T ML Fit | | $0.9^{+0.4}_{-0.7}$ | $0.3^{+0.2}_{-0.2}$ | $0.6^{+0.3}_{-0.3}$ |
| Tight working point | | | | |
| Data | | 5 | 2 | 0 |
| Template | $n_b = 3$ | 4.03 ± 0.8 | 2.4 ± 0.5 | 3.1 ± 0.6 |
| Data | | 1 | 0 | 0 |
| Template | $n_b = 4$ | 0.1 ± 0.1 | 0.1 ± 0.1 | 0.0 ± 0.1 |

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

²⁵¹³ The numerical results and extrapolation to the $n_b^{reco} = 3, 4$ bins for all H_T and working points are shown in Table 6.4. Also included within the table are total SM background

2515 predictions determined by the maximum likelihood fit in both jet multiplicity categories
2516 of the α_T analysis for the CSV tagger, as introduced in Section (4.8). No excess of data
2517 is found for any of the three CSV working points. Predictions within the signal region
2518 from this method are also found to be statistically compatible with the background
2519 predictions determined by the α_T simultaneous fit (introduced in Table 5.2).

2520 **6.3. Summary**

2521 A SUSY signature such as one from gluino-induced third-generation squark production,
2522 would result in a final state with an underlying b-quark content greater than two. In
2523 order to be able to discriminate such signatures from the SM background, templates are
2524 generated based on a parameterisation of SM processes, where the underlying b-quarks
2525 per event is typically zero or two. These templates are then fit to data in a low n_b^{reco}
2526 (0-2) control region in order to extrapolate a prediction within a high n_b^{reco} (3-4) signal
2527 region. This approach is built upon the assumptions that the defined control region is
2528 almost entirely free of any possible signal contamination from possible signal topologies
2529 with a small number of b quarks in the final state.

2530 The method was demonstrated both in simulation and also in data, using the SM enriched
2531 $\mu + \text{jets}$ selection from the α_T search. This was conducted to prove conceptually and
2532 experimentally that the method is valid and that there is adequate control over the
2533 measurement of the efficiency of each jet flavour for all working points of the CSV tagger.
2534 Additionally this method was further applied to the hadronic signal region of the α_T
2535 analysis, where good agreement is observed between the SM background predictions from
2536 the template method, observations in data and also the background estimation procedure
2537 of the α_T analysis.

Chapter 7.

²⁵³⁸ Conclusions

²⁵³⁹ A search for supersymmetry has been presented based on a data sample of pp collisions

²⁵⁴⁰ collected at $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of $11.7 \pm 0.5 \text{ fb}^{-1}$.

²⁵⁴¹ Final states with two or more jets and significant missing transverse energy, a typical

²⁵⁴² final state topology of R-parity conserving SUSY models have been analysed. The α_T

²⁵⁴³ variable is utilised as the main discriminator between balanced multi-jet backgrounds

²⁵⁴⁴ and those with real missing transverse energy.

²⁵⁴⁵ Within the search presented, Standard Model (SM) backgrounds are estimated from

²⁵⁴⁶ a simultaneous binned likelihood fit to a hadronic signal selection as well as three SM

²⁵⁴⁷ process enriched control samples. The search is split into total transverse hadronic

²⁵⁴⁸ energy (H_T), jets identified as originating for a b-quark (n_b^{reco}), and jet multiplicity

²⁵⁴⁹ (n_{jet}) categories to improve sensitivity to a range of possible supersymmetric final states.

²⁵⁵⁰ Systematic errors due to theory, detector effects and simulation deficiencies are quantified

²⁵⁵¹ through the use of data driven closure tests and accounted for in the final interpretation.

²⁵⁵² Observations in data are found to be compatible with a SM only hypothesis.

²⁵⁵³ In the absence of a signal like excess the analysis is further interpreted in a set of

²⁵⁵⁴ Simplified Model Spectra (SMS) models, representing a set of model independent decay

²⁵⁵⁵ topologies parameterised only by the production process and the masses of their parent

²⁵⁵⁶ sparticle and Lightest Supersymmetric Partner (LSP). In models mediated by gluino pair

²⁵⁵⁷ production and containing a large mass difference between the gluino and LSP, exclusion

²⁵⁵⁸ limits of the gluino mass are set in the range 950-1125 GeV. For SMS models describing

²⁵⁵⁹ direct squark pair production, first or second generation squarks are excluded up to 775

²⁵⁶⁰ GeV, with direct bottom squarks production excluded up to masses of 600 GeV.

2561 In the case of gluino mediated third generation signatures containing many jets originating
2562 from b-quarks in the final state, mass limits are set in the range of 975-1125 GeV for
2563 large mass splittings between the gluino and the LSP. The experimental sensitivity
2564 to these models is attributed to the n_b^{reco} categorisation of the analysis, where the
2565 signal-to-background is enhanced within the phase space of the search at high n_b^{reco} .

2566 Furthermore, a measurement of the performance of the Level-1 trigger for jets and jet
2567 energy sum quantities has also been presented. These studies quantify any change in
2568 Level-1 performance after the introduction of a 5 GeV jet seed threshold into the jet
2569 clustering algorithm. No significant change in single jet trigger efficiencies is observed
2570 and good performance is observed for a range of Level-1 jet energy sum quantities.

2571 This change was introduced to facilitate a reduction in the rate of events triggered by
2572 energy deposits due to soft non-collimated jets from secondary interactions, and which are
2573 not of interest to physics analyses. This was necessary to ensure, that trigger thresholds
2574 can be maintained at low values in the presence of an ever increasing number of bunch
2575 crossings per proton interaction. In the context of SUSY, this is a necessity to keep
2576 CMS sensitive to types of compressed spectra signatures characterised by low transverse
2577 energy jets and small missing transverse energy signatures.

2578 Finally, an approach that uses a template fit method to the n_b^{reco} distribution of SM
2579 processes within a supersymmetric search is introduced and then validated in simulation
2580 and data. The approach can be used to identify any excess in data arising from gluino
2581 mediated third generation supersymmetric signatures. It is utilised within this thesis as
2582 a crosscheck to the α_T background prediction at high b-tagged jet multiplicities. This
2583 method is found to give a SM background estimation that is in good agreement with the
2584 α_T search within the hadronic signal region.

2585 The continued absence of a supersymmetric signal in the α_T search or other analyses
2586 at CMS [101][102][103], puts pressure on the parameter space in which SUSY can reside.
2587 Indeed the smoking gun that many theorists and experimentalists hoped to see at the
2588 LHC has not materialised. Instead identifying a SUSY signal may now only result
2589 from many years of data taking and the incorporation of increasingly advanced analysis
2590 techniques. An unenviable task considering the difficulties of not knowing where SUSY
2591 may reside, but perhaps solace can be taken in remembering that nothing worth having
2592 ever comes easy.

Appendix A.

²⁵⁹⁴ Miscellaneous

²⁵⁹⁵ A.1. Jet Identification Criteria

²⁵⁹⁶ For Calo jets the following identification criteria were applied:

| Loose CaloJet Id | |
|------------------------|---|
| Variable | Definition |
| $f_{HPD} < 0.98$ | Fraction of jet energy contributed from “hottest” 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.

²⁵⁹⁷ For PF jets the following identification criteria were applied:

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

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

2598 **A.2. Primary Vertices**

2599 The pile-up per event is defined by the number of 'good' reconstructed primary vertices
2600 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.3: Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex.

Appendix B.

2601 L1 Jets

2602 B.1. Jet Matching Efficiencies

2603 The single jet turn-on curves are derived from events independent of whether the leading
2604 jet in an event is matched to a Level 1 jet using ΔR matching detailed in Section (3.4.3).
2605 These turn-ons are produced from events which are not triggered on jet quantities and
2606 therefore it is not guaranteed that the lead jet of an event will be seeded by a Level 1 jet.
2607 Figure B.1 shows the particular matching efficiency of a lead jet to a L1 jet.

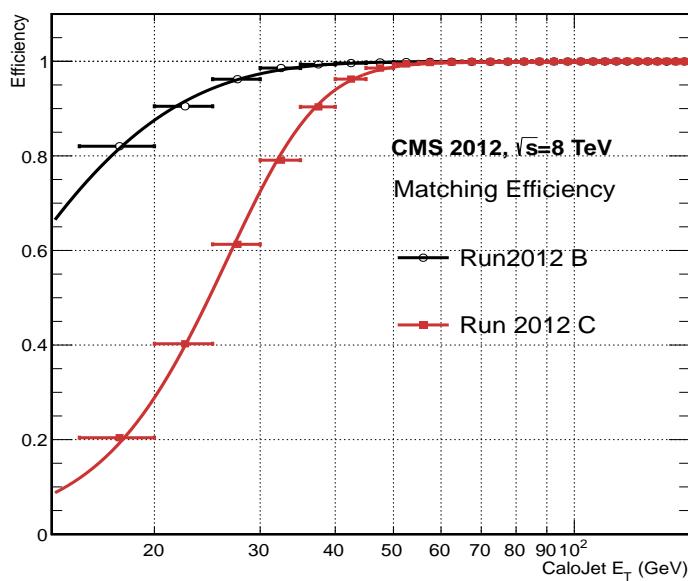


Figure B.1: Leading jet matching efficiency as a function of the offline CaloJet E_T , measured in an isolated muon triggered dataset in the 2012B and 2012C run periods.

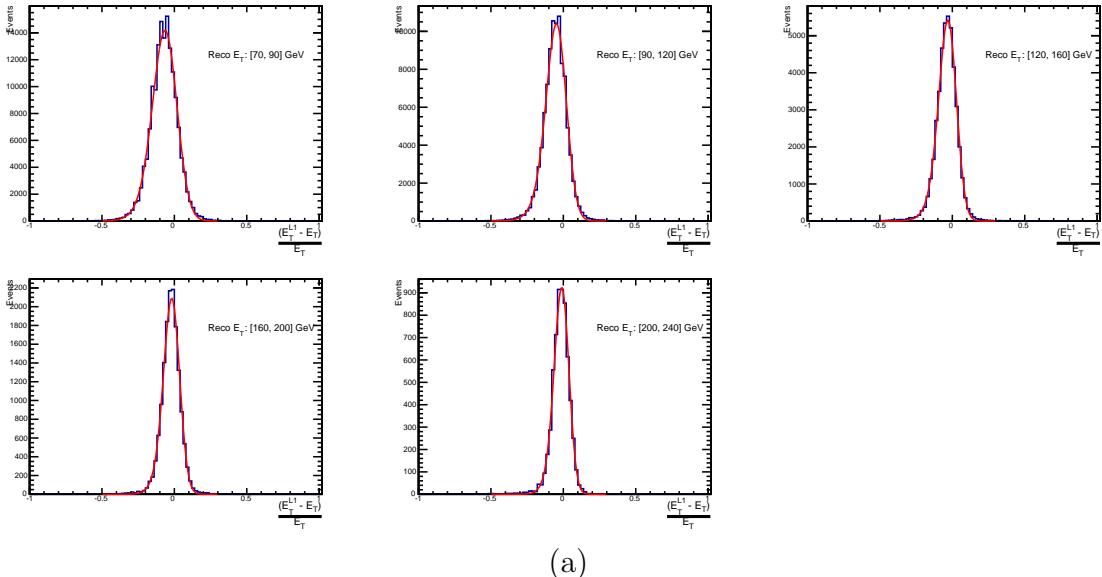
| Run Period | μ | σ |
|------------|------------------|-----------------|
| 2012B | 6.62 ± 0.01 | 0.79 ± 0.03 |
| 2012C | 19.51 ± 0.03 | 7.14 ± 0.02 |

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 .

2608 It can be seen that the turn-on occurs at a lower E_T during the 2012B run period. The
 2609 seed threshold requirement of a 5 GeV jet seed in run 2012C result in more events in
 2610 which the lead offline jet does not have an associated L1 jet. This can be attributed to
 2611 events with soft non-collimated jets in which the energy deposits are not centralised in a
 2612 calorimeter region. However, for larger jet E_T thresholds typical of those used by physics
 2613 analyses, 100% efficiency is observed, and therefore this effect has no impact to overall
 2614 physics performance.

2615 A fit of an EMG function to the matching efficiencies find mean, μ , values of 6.62 GeV
 2616 and 19.51 GeV for Run 2012B and 2012C respectively and is shown in Table B.1.

2617 B.2. Leading Jet Energy Resolution



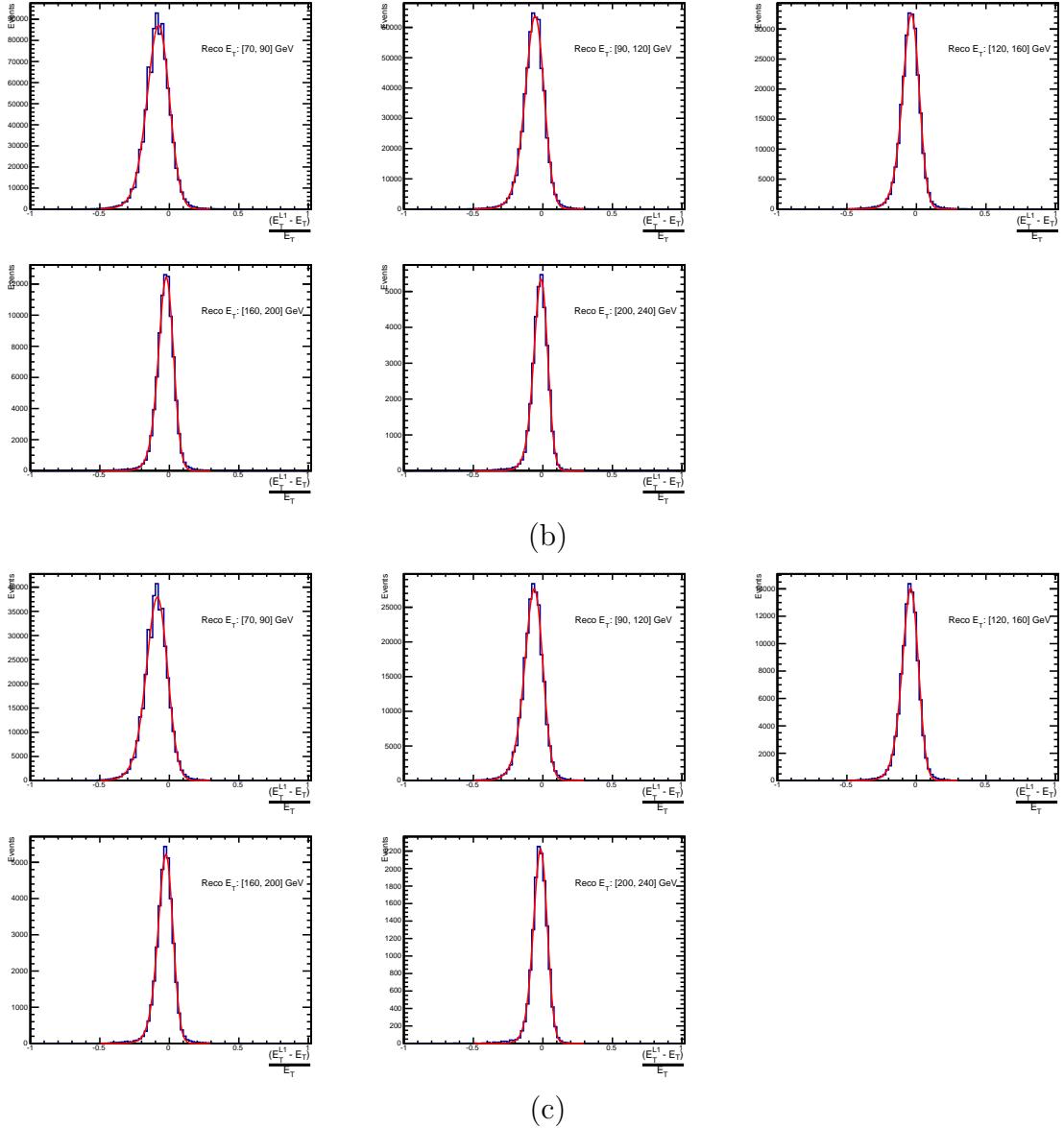
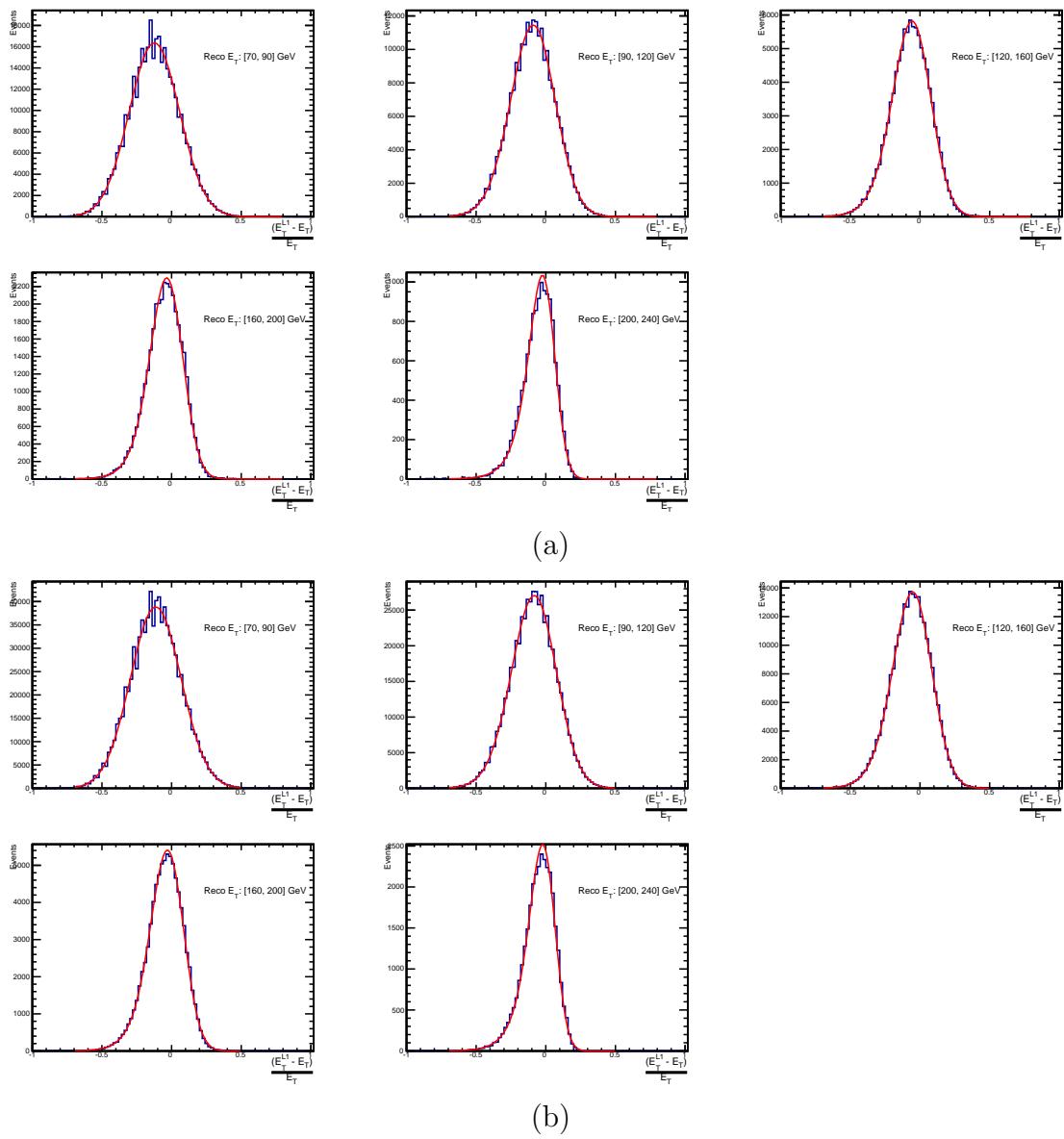
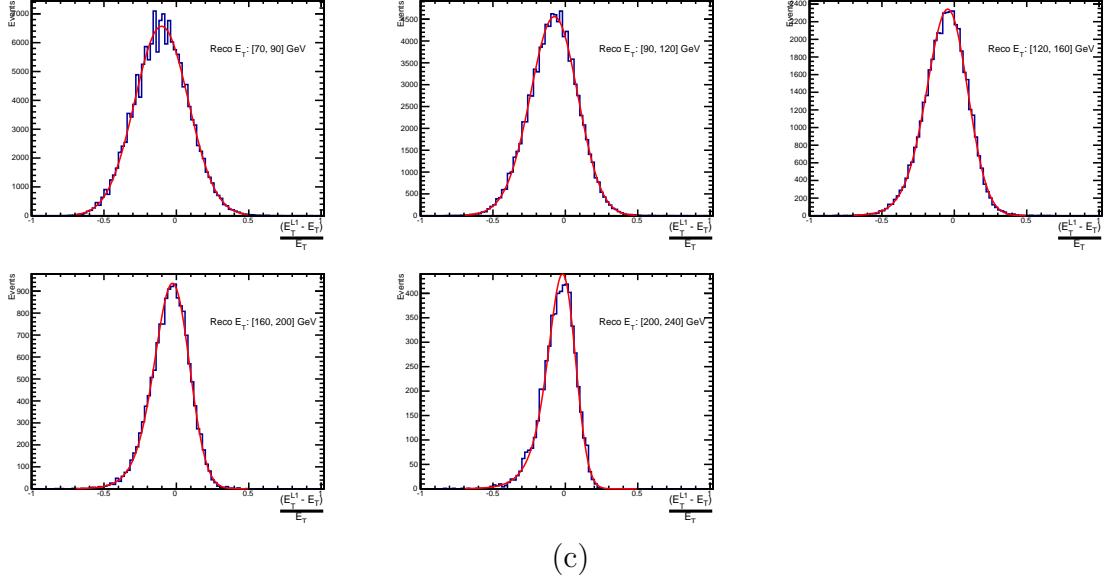


Figure B.2: Resolution plots of the leading offline jet Calo E_T measured as a function of $\frac{(L1 E_T - \text{Offline } E_T)}{\text{Offline } E_T}$ for (a) low, (b) medium, and (c) high pile-up conditions as defined in Section (3.4.4).





(c)

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

2618 B.3. Resolution for Energy Sum Quantities

2619 The following plots show the resolution parameters for energy sum quantities as a function
 2620 of the quantity (q) itself. In this case, the μ , σ and λ fit values to an EMG function
 2621 defined by Equation (3.3) for each of the individual $\frac{(L1 q - \text{Offline } q)}{\text{Offline } q}$ distributions, in bins
 2622 of the quantity (q) is displayed.

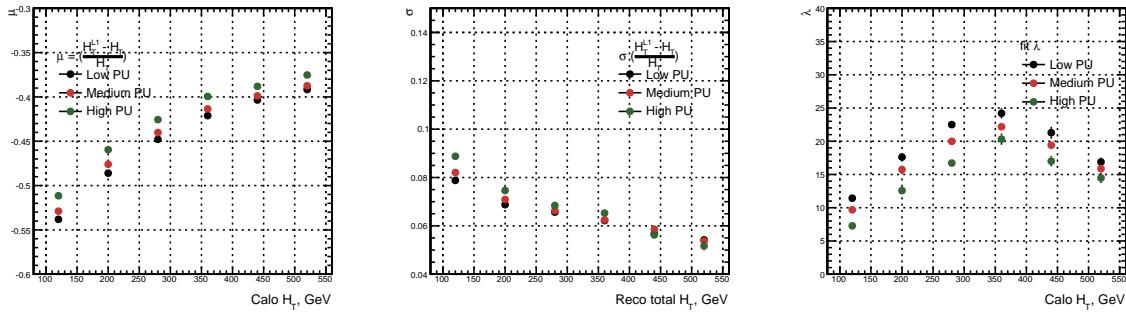


Figure B.4: H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile-up conditions. Shown are the mean μ (left), resolution σ (middle) and λ (right) fit values to an EMG function for the $\frac{(L1 H_T - H_T)}{H_T}$ distributions.

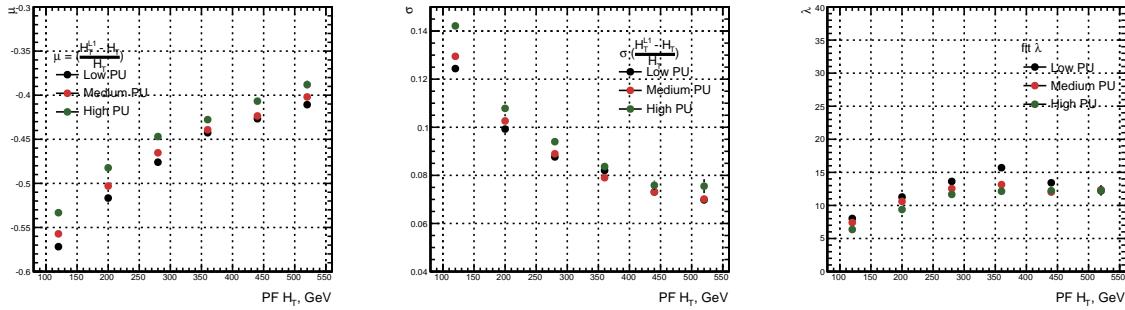


Figure B.5: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile-up conditions. Shown are the mean μ (left), resolution σ (middle) and λ (right) fit values to an EMG function for the $\frac{(L1H_T - H_T)}{H_T}$ distributions.

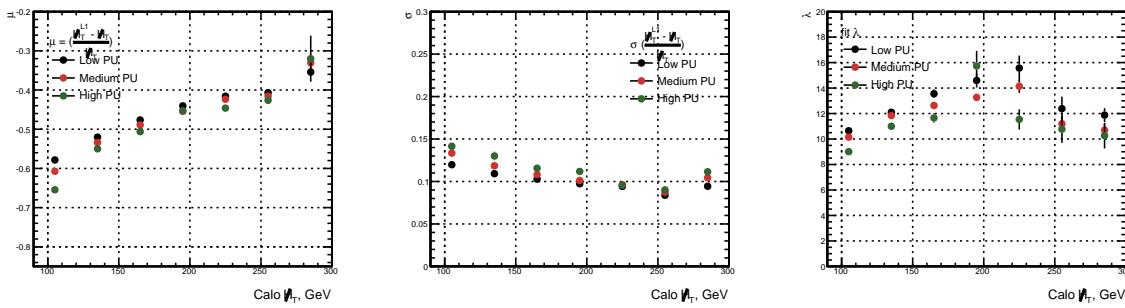


Figure B.6: H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile-up conditions. Shown are the mean μ (left), resolution σ (middle) and λ (right) fit values to an EMG function for the $\frac{(L1H_T - H_T)}{H_T}$ distributions.

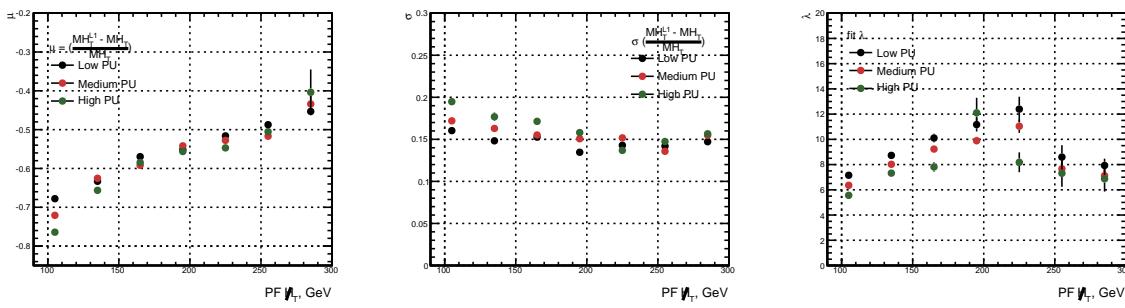


Figure B.7: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile-up conditions. Shown are the mean μ (left), resolution σ (middle) and λ (right) fit values to an EMG function for the $\frac{(L1H_T - H_T)}{H_T}$ distributions.

Appendix C.

²⁶²³ Additional Material on Background ²⁶²⁴ Estimation Methods

²⁶²⁵ C.1. Determination of k_{QCD}

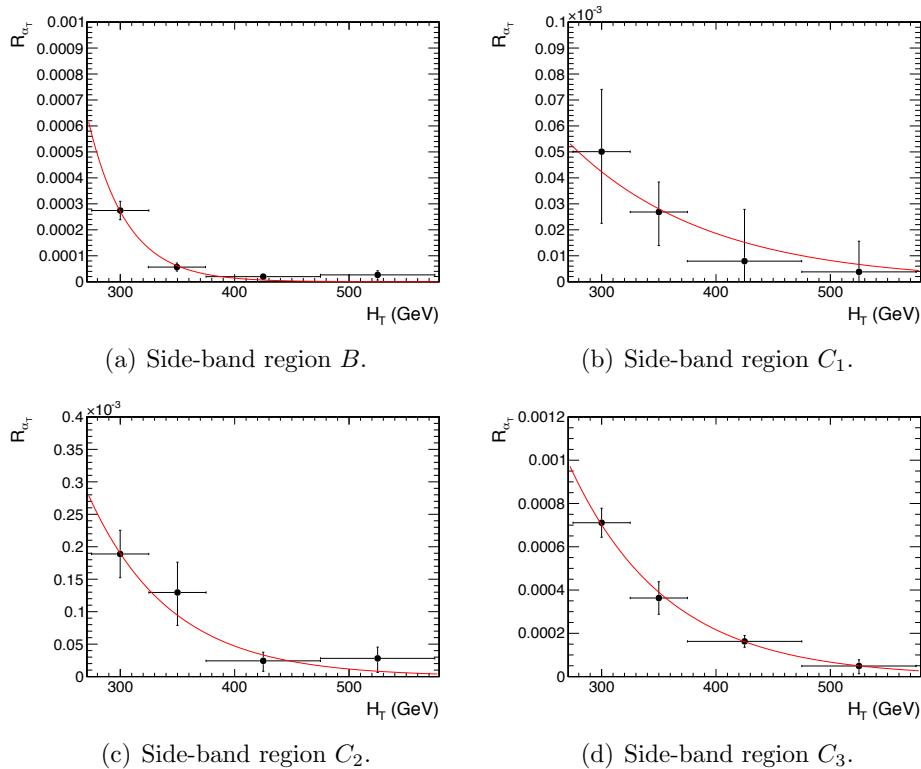


Figure C.1: $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is conducted between the H_T region $275 < H_T < 575$.

2627 **C.2. Effect of Varying Background Cross-sections on**
 2628 **Closure Tests**

2629 Closure tests with cross section variations of +20% and -20% applied to $W + \text{jets}$ and $t\bar{t}$
 2630 processes respectively.

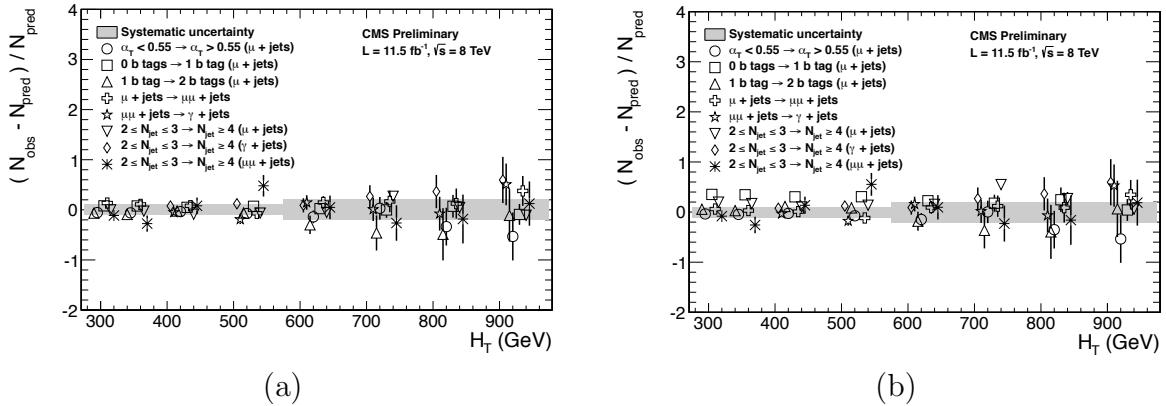


Figure C.2: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) in the $2 \leq n_{jet} \leq 3$ jet multiplicity category for nominal and varied cross-sections; (a) Nominal and (b) Varied $\pm 20\%$.

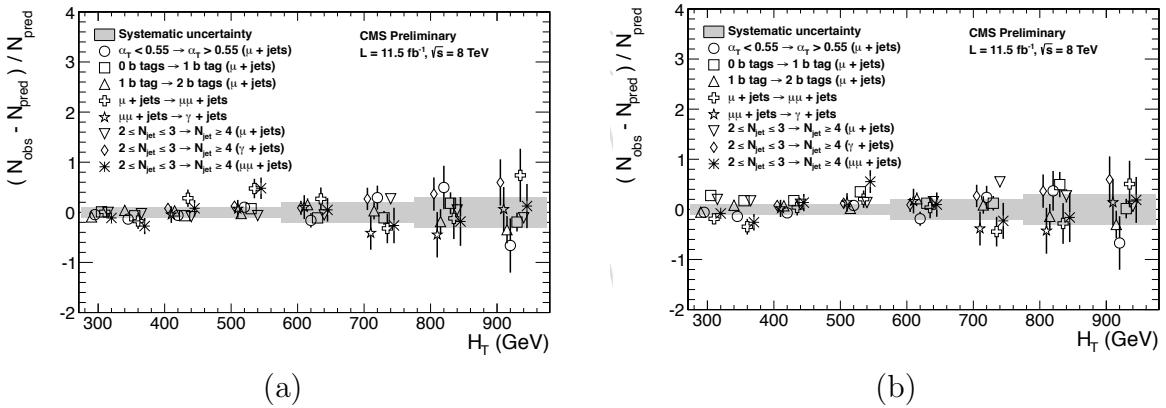


Figure C.3: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) in the $n_{jet} \geq 4$ jet multiplicity category for nominal and varied cross-sections; (a) Nominal (b) Varied $\pm 20\%$.

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

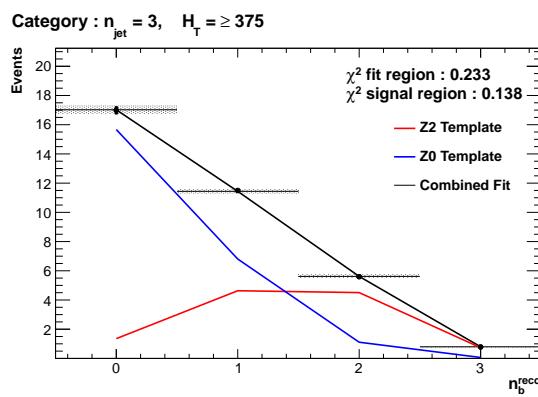
Table C.1: Translation factors constructed from the $\mu +$ jets control sample and signal selection MC, to predict yields for the $W +$ jets and $t\bar{t}$ back-grounds in the signal region with (a) NNLO cross sections corrected by k-factors determined from a data sideband see Section (4.4), marked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

Appendix D.

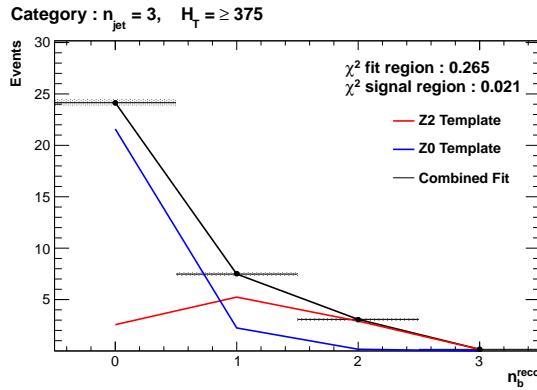
²⁶³¹ Additional Material For B-tag ²⁶³² Template Method

²⁶³³ D.1. Templates Fits in Simulation

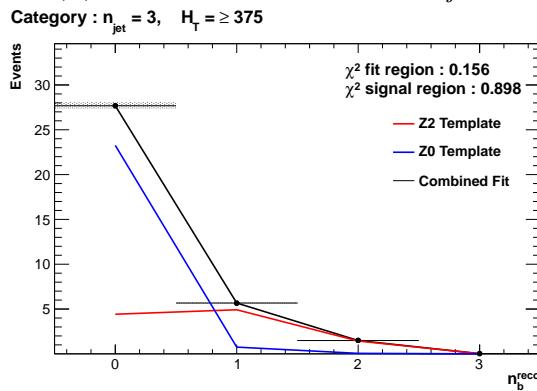
²⁶³⁴ The result of template fits for the three CSV working points in the $n_{jet} = 3, H_T > 375$
²⁶³⁵ category:



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



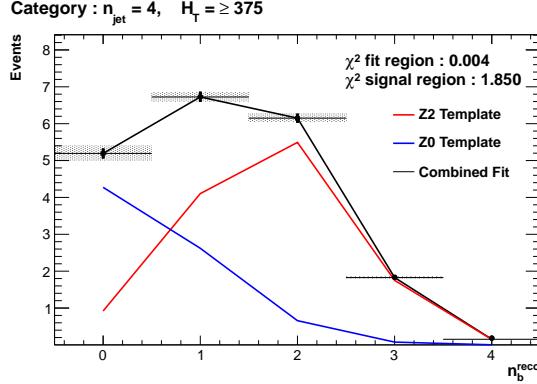
(b) Medium working point $n_{jet} = 3$



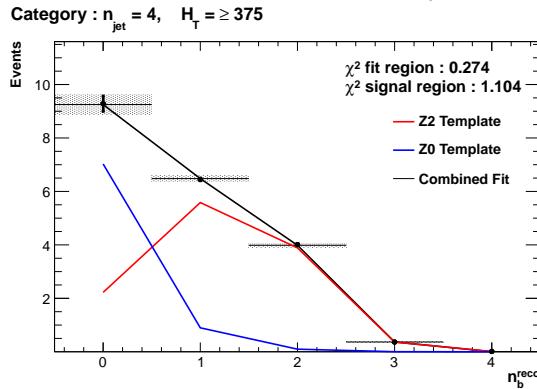
(c) Tight working point $n_{jet} = 3$

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

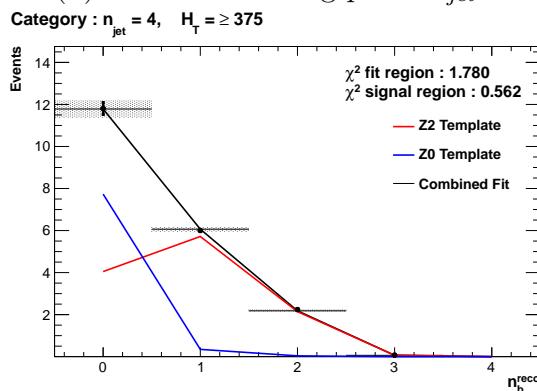
2636 Template fits for the three CSV working points in the $n_{jet} = 4$, $H_T > 375$ category:



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



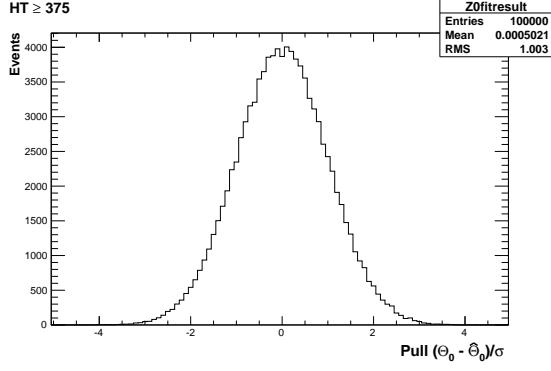
(b) Medium working point $n_{jet} = 4$



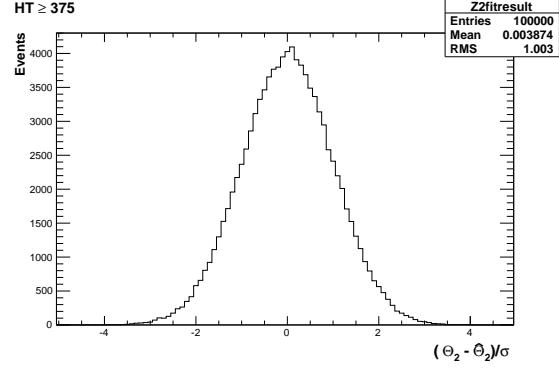
(c) Tight working point $n_{jet} = 4$

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

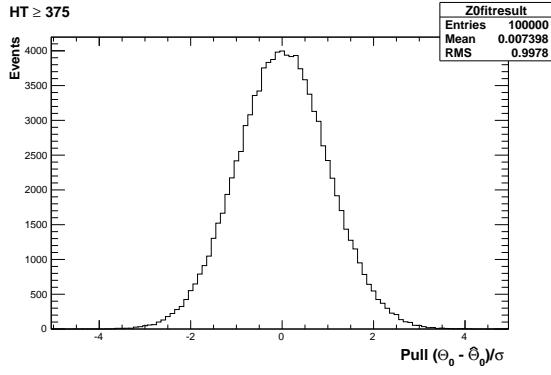
2637 D.2. Pull Distributions for Template Fits



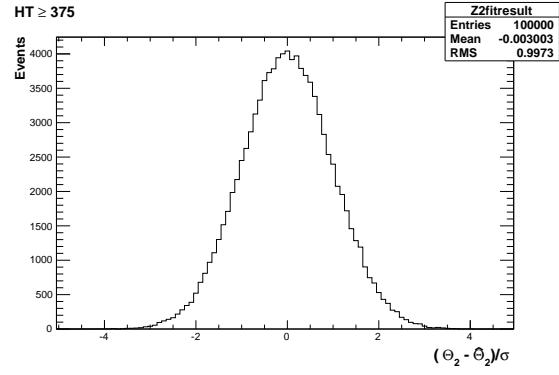
(a) Z0 Template, $H_T > 375$, $n_{jet} = 3$



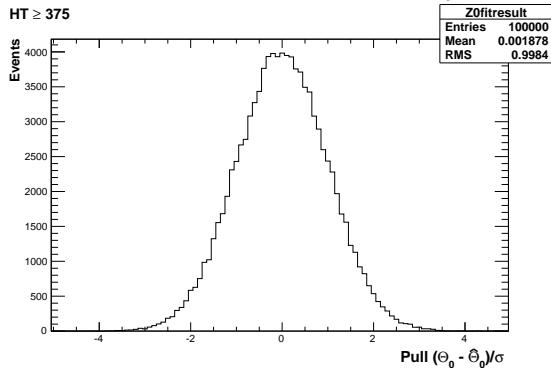
(b) Z2 Template, $H_T > 375$, $n_{jet} = 3$



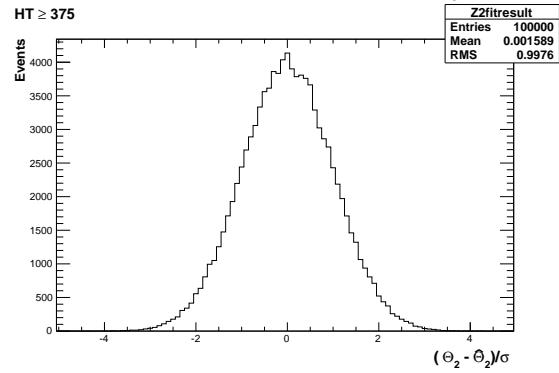
(a) Z0 Template, $H_T > 375$, $n_{jet} = 4$



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



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



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

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

²⁶³⁸ **D.3. Templates Fits in Data Control Sample**

²⁶³⁹ Template fits for the three H_T bins, in the $n_{jet} = 3$, medium CSV working point:

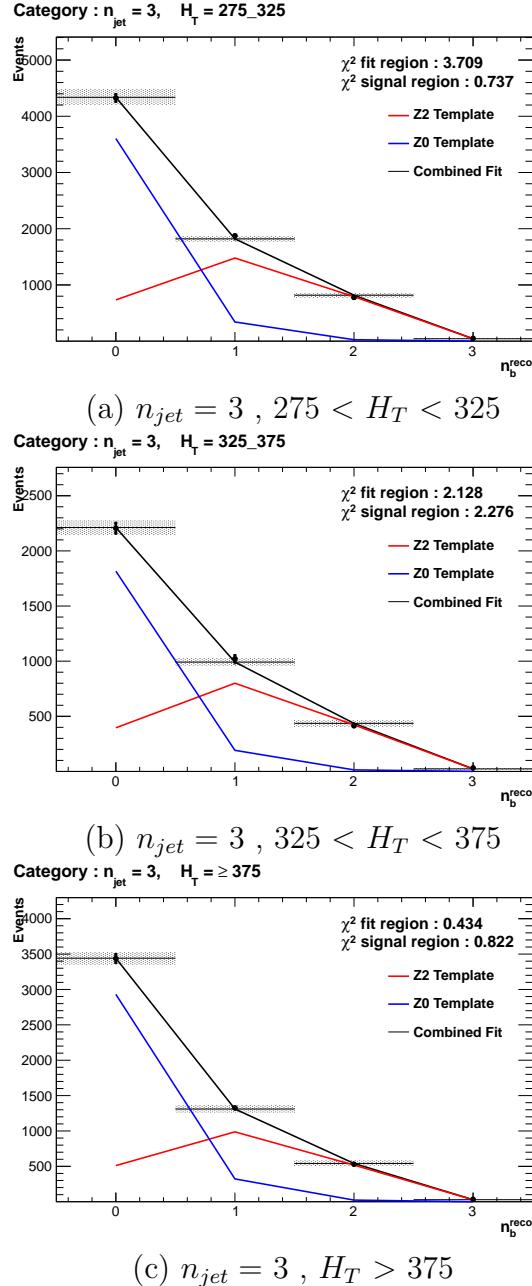


Figure D.4: Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{--}2$ control region to data from the $\mu +$ jets control sample, for the CSVM working point, with $n_{jet} = 3$ in each H_T category. Data is represented by the black circles with the blue, red and black lines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

²⁶⁴⁰ Template fits for the three H_T bins, in the $n_{jet} = 4$, medium CSV working point:

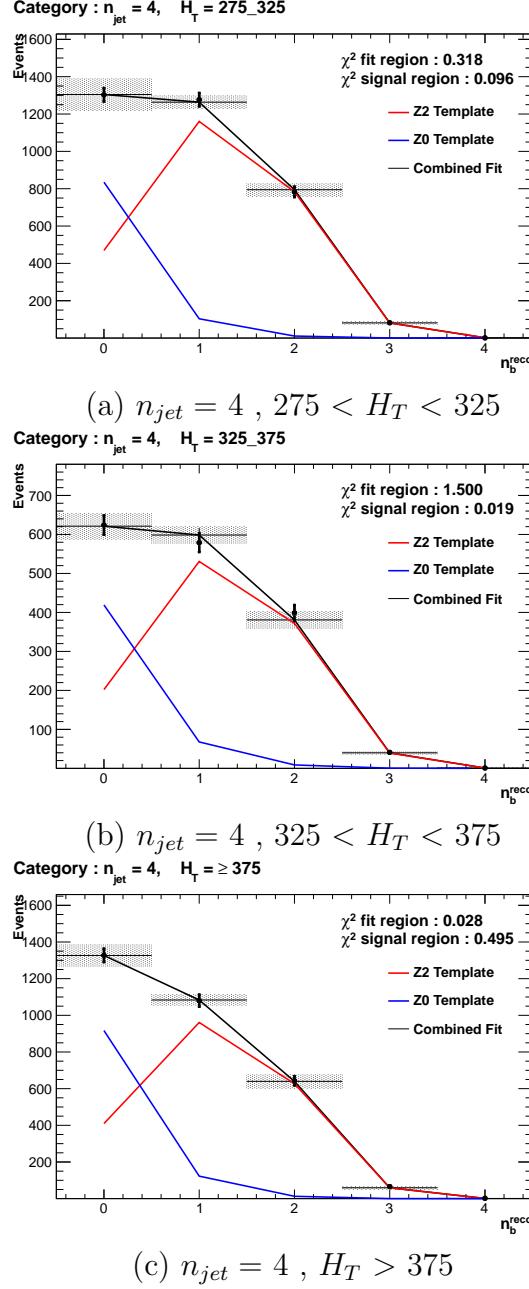


Figure D.5: Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{--}2$ control region to data from the $\mu +$ jets control sample, for the CSVM working point, with $n_{jet} = 4$ in each H_T category. Data is represented by the black circles with the blue, red and black lines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

²⁶⁴¹ **D.4. Templates Fits in Data Signal Region**

²⁶⁴² Template fits for the three CSV working points, in the $n_{jet} = 3, H_T > 375$ category :

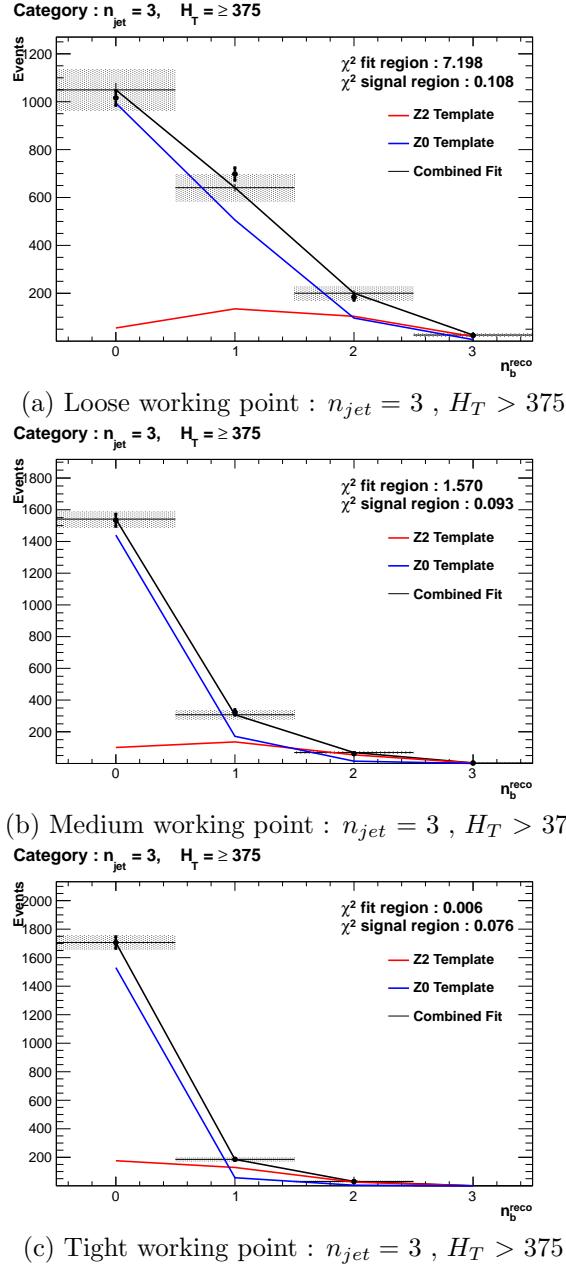


Figure D.6: Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0-2$ control region to data from the hadronic signal selection, in the $n_{jet} = 3$ and $H_T > 375$ category for all CSV working points. Data is represented by the black circles with the blue, red and black lines representing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

2643 Template fits for the three CSV working points, in the $n_{jet} = 4$, $H_T > 375$ category :

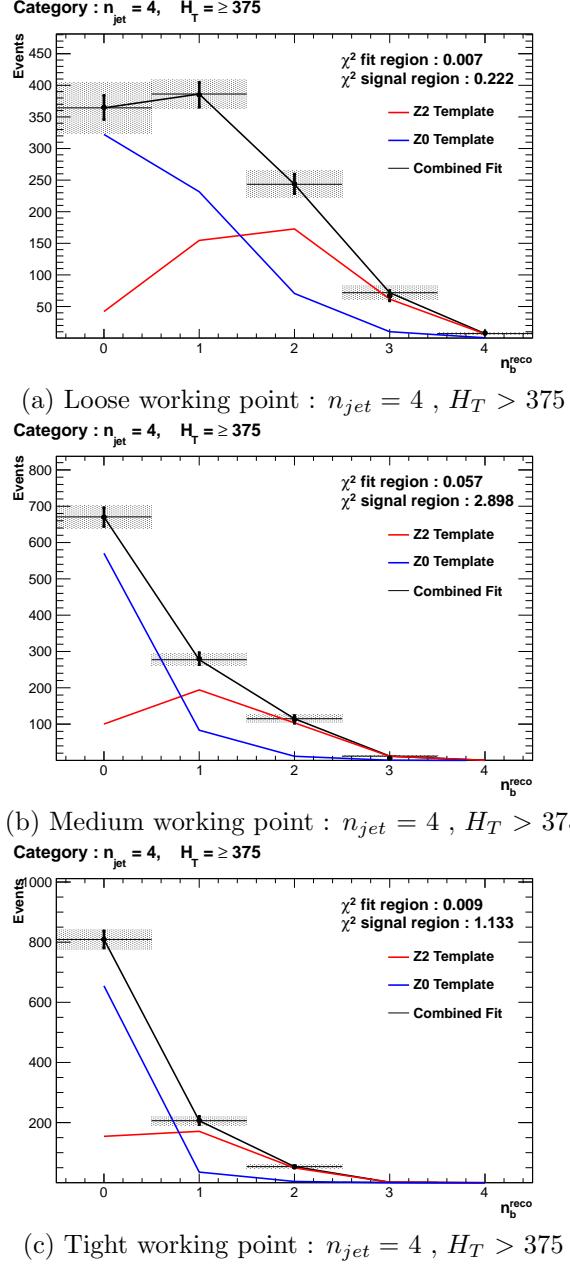


Figure D.7: Results of fitting the $Z = 0$ and $Z = 2$ templates in the $n_b^{reco} = 0\text{-}2$ control region to data from the hadronic signal selection, in the $n_{jet} = 4$ and $H_T > 375$ category for all CSV working points. Data is represented by the black circles with the blue, red and black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

2644

²⁶⁴⁵ Bibliography

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

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

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

- 2729 **D49** (1994) doi:10.1103/PhysRevD.49.6173, arXiv:hep-ph/9312272.
- 2730 [39] C. Strege et al., “Updated global fits of the cMSSM including the latest LHC
2731 SUSY and Higgs searches and XENON100 data”, *JCAP* **1203** (2012)
2732 doi:10.1088/1475-7516/2012/03/030, arXiv:1112.4192.
- 2733 [40] M. Citron et al., “The End of the CMSSM Coannihilation Strip is Nigh”,
2734 *Phys.Rev.* **D87** (2013) doi:10.1103/PhysRevD.87.036012, arXiv:1212.2886.
- 2735 [41] D. Ghosh et al., “How Constrained is the cMSSM?”, *Phys.Rev.* **D86** (2012)
2736 doi:10.1103/PhysRevD.86.055007, arXiv:1205.2283.
- 2737 [42] LHC New Physics Working Group Collaboration, “Simplified Models for LHC
2738 New Physics Searches”, *J.Phys.* **G39** (2012) 105005,
2739 doi:10.1088/0954-3899/39/10/105005, arXiv:1105.2838.
- 2740 [43] J. Alwall et al., “Simplified Models for a First Characterization of New Physics at
2741 the LHC”, *Phys.Rev.* **D79** (2009) 075020, doi:10.1103/PhysRevD.79.075020,
2742 arXiv:0810.3921.
- 2743 [44] CMS Collaboration, “Interpretation of Searches for Supersymmetry with
2744 simplified Models”, *Phys.Rev.* **D88** (2013) 052017,
2745 doi:10.1103/PhysRevD.88.052017, arXiv:1301.2175.
- 2746 [45] J. Hisano et al., “Natural effective supersymmetry”, *Nucl.Phys.* **B584** (2000)
2747 3–45, doi:10.1016/S0550-3213(00)00343-6, arXiv:hep-ph/0002286.
- 2748 [46] M. Papucci et al., “Natural SUSY Endures”, *JHEP* **1209** (2012) 035,
2749 doi:10.1007/JHEP09(2012)035, arXiv:1110.6926.
- 2750 [47] B. Allanach and B. Gripaios, “Hide and Seek With Natural Supersymmetry at the
2751 LHC”, *JHEP* **1205** (2012) 062, doi:10.1007/JHEP05(2012)062,
2752 arXiv:1202.6616.
- 2753 [48] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3**
2754 (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.
- 2755 [49] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron
2756 Collider”, *JINST* **3** (2008) doi:10.1088/1748-0221/3/08/S08003.
- 2757 [50] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **0803**
2758 (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

- 2759 [51] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008)
2760 S08005, doi:10.1088/1748-0221/3/08/S08005.
- 2761 [52] J.-L. Caron, “LHC Layout. Schema general du LHC.”, (Sep, 1997).
- 2762 [53] CMS Collaboration, “CMS Luminosity - Public Results”, , (2011).
2763 http://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.
- 2764 [54] CERN, “CMS Compact Muon Solenoid.”, (Feb, 2010).
2765 http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg.
- 2766 [55] “The CMS Electromagnetic Calorimeter Project: Technical Design Report”.
2767 Technical Design Report CMS. CERN, Geneva, 1997.
- 2768 [56] “The CMS Muon Project: Technical Design Report”. Technical Design Report
2769 CMS. CERN, Geneva, 1997.
- 2770 [57] CMS Collaboration, “The CMS Physics Technical Design Report, Volume 1”,
2771 *CERN/LHCC 2006-001* (2006).
- 2772 [58] M. Cacciari et al., “The anti- k_t jet clustering algorithm”, *Journal of High Energy*
2773 *Physics* **2008** (2008), no. 04, 063.
- 2774 [59] “Jet Performance in pp Collisions at 7 TeV”, CMS-PAS-JME-10-003, CERN,
2775 Geneva, (2010).
- 2776 [60] X. Janssen, “Underlying event and jet reconstruction in CMS”, CMS-CR-2011-012,
2777 CERN, Geneva, (Jan, 2011).
- 2778 [61] CMS Collaboration, “Determination of jet energy calibration and transverse
2779 momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11.,
- 2780 [62] R. Eusebi, “Jet energy corrections and uncertainties in CMS: reducing their impact
2781 on physics measurements”, *Journal of Physics: Conference Series* **404** (2012).
- 2782 [63] CMS Collaboration, “Algorithms for b Jet identification in CMS”,
2783 CMS-PAS-BTV-09-001, CERN, 2009. Geneva, (Jul, 2009).
- 2784 [64] CMS Collaboration, “Performance of b-tagging at $\sqrt{s} = 8$ TeV in multijet, $t\bar{t}$ and
2785 boosted topology events”, CMS-PAS-BTV-13-001, CERN, Geneva, (2013).
- 2786 [65] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”,
2787 *Journal of Instrumentation* **8** (2013), no. 04.,

- 2788 [66] CMS Collaboration, “CMS. The TriDAS Project. Technical design report, Vol. 1:
2789 The trigger systems”, ,
- 2790 [67] CMS Collaboration, “CMS: The TriDAS Project. Technical design report, Vol. 2:
2791 Data acquisition and high-level trigger” ,.
- 2792 [68] G. Iles, Brooke et al., “Revised CMS Global Calorimeter Trigger Functionality &
2793 Algorithms”, ,
- 2794 [69] CMS Collaboration, “ Calibration and Performance of the Jets and Energy Sums
2795 in the Level-1 Trigger”, CMS IN 2013/006 (2013), CERN, Geneva, (2013).
- 2796 [70] B. Shorney-Mathias, “Search for supersymmetry in pp collisions with all-hadronic
2797 final states using the α_T variable with the CMS detector at the LHC”. PhD thesis,
2798 2014.
- 2799 [71] B. Mathias et al., “Study of Level-1 Trigger Jet Performance in High Pile-up
2800 Running Conditions”, CMS-IN-2013-007, CERN, Geneva, (2013).
- 2801 [72] J. J. Brooke, “Performance of the CMS Level-1 Trigger”, CMS-CR-2012-322,
2802 CERN, Geneva, (Nov, 2012).
- 2803 [73] CMS Collaboration, “Search for supersymmetry in final states with missing
2804 transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions
2805 using the variable α_T ”, *JHEP* **1301** (2013) 077,
2806 doi:10.1007/JHEP01(2013)077, arXiv:1210.8115.
- 2807 [74] CMS Collaboration, “SUSY searches with dijet events”, CMS-PAS-SUS-08-005,
2808 (2008).
- 2809 [75] L. Randall and D. Tucker-Smith, “Dijet Searches for Supersymmetry at the Large
2810 Hadron Collider”, *Phys. Rev. Lett.* **101** (Nov, 2008) 221803,
2811 doi:10.1103/PhysRevLett.101.221803.
- 2812 [76] CMS Collaboration, “Search strategy for exclusive multi-jet events from
2813 supersymmetry at CMS”, CMS-PAS-SUS-09-001, CERN, (Jul, 2009).
- 2814 [77] CMS Collaboration, “CMS Btag POG : CMS b-tagging performance database”, ,
2815 (2013). <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>.
- 2816 [78] CMS Collaboration, “Calorimeter Jet Quality Criteria for the First CMS Collision
2817 Data”, CMS-PAS-JME-09-008, CERN, (Apr, 2010).

- 2818 [79] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision
2819 events at 7 TeV”, *Journal of Instrumentation* **7** (October, 2012) 2P,
2820 [doi:10.1088/1748-0221/7/10/P10002](https://doi.org/10.1088/1748-0221/7/10/P10002), [arXiv:1206.4071](https://arxiv.org/abs/1206.4071).
- 2821 [80] CMS Collaboration, “Search for supersymmetry in events with photons and
2822 missing energy”, CMS-PAS-SUS-12-018, CERN, Geneva, (2012).
- 2823 [81] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys.Lett.*
2824 **B659** (2008) [doi:10.1016/j.physletb.2007.09.077](https://doi.org/10.1016/j.physletb.2007.09.077), [arXiv:0707.1378](https://arxiv.org/abs/0707.1378).
- 2825 [82] Z. Bern et al., “Driving missing data at next-to-leading order”, *Phys. Rev. D* **84**
2826 (Dec, 2011) 114002, [doi:10.1103/PhysRevD.84.114002](https://doi.org/10.1103/PhysRevD.84.114002).
- 2827 [83] Z. Bern et al., “Driving Missing Data at Next-to-Leading Order”, *Phys.Rev.*
2828 **D84** (2011) 114002, [doi:10.1103/PhysRevD.84.114002](https://doi.org/10.1103/PhysRevD.84.114002), [arXiv:1106.1423](https://arxiv.org/abs/1106.1423).
- 2829 [84] CMS Collaboration, “Data-Driven Estimation of the Invisible Z Background to the
2830 SUSY MET Plus Jets Search”, CMS-PAS-SUS-08-002, CERN, (Jan, 2009).
- 2831 [85] T. Sjostrand et al., “PYTHIA 6.4 Physics and Manual”, *JHEP* **0605** (2006) 026,
2832 [doi:10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026), [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175).
- 2833 [86] W. Beenakker et al., “Squark and gluino production at hadron colliders”,
2834 *Nucl.Phys.* **B492** (1997) 51–103, [doi:10.1016/S0550-3213\(97\)80027-2](https://doi.org/10.1016/S0550-3213(97)80027-2),
2835 [arXiv:hep-ph/9610490](https://arxiv.org/abs/hep-ph/9610490).
- 2836 [87] S. Abdullin et al., “The Fast Simulation of the CMS Detector at LHC”, *Journal*
2837 *of Physics: Conference Series* **331** (2011), no. 3.,
- 2838 [88] S. Banerjee, M. D. Hildreth, and the CMS Collaboration, “Validation and Tuning
2839 of the CMS Full Simulation”, *Journal of Physics: Conference Series* **331** (2011).
- 2840 [89] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer
2841 2012 Update”, CMS-PAS-LUM-12-001, CERN, Geneva, (2012).
- 2842 [90] M. Botje et al., “The PDF4LHC Working Group Interim Recommendations”,
2843 [arXiv:1101.0538](https://arxiv.org/abs/1101.0538).
- 2844 [91] E. M. Laird, “A Search for Squarks and Gluinos with the CMS Detector”. PhD
2845 thesis, 2012.
- 2846 [92] R. Cousins, “Probability Density Functions for Positive Nuisance Parameters”.
2847 2012. <http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf>.

- 2848 [93] L. Moneta et al., “The RooStats project”, 2010. [arXiv:1009.1003](https://arxiv.org/abs/1009.1003).
- 2849 [94] F. James and M. Roos, “Minuit: A System for Function Minimization and
2850 Analysis of the Parameter Errors and Correlations”, *Comput.Phys.Commun.* **10**
2851 (1975) 343–367, doi:[10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9).
- 2852 [95] A. L. Read, “Presentation of search results: the CL s technique”, *Journal of
2853 Physics G: Nuclear and Particle Physics* **28** (2002).
- 2854 [96] T. Junk, “Confidence level computation for combining searches with small
2855 statistics”, *Nuclear Instruments and Methods in Physics Research Section A* **434**
2856 (1999), no. 23, doi:[http://dx.doi.org/10.1016/S0168-9002\(99\)00498-2](http://dx.doi.org/10.1016/S0168-9002(99)00498-2).
- 2857 [97] A. L. Read, “Modified frequentist analysis of search results (the CL_s method)”,
2858 CERN-OPEN-2000-205, (2000).
- 2859 [98] G. Cowan et al., “Asymptotic formulae for likelihood-based tests of new physics”,
2860 *The European Physical Journal C* **71** (2011)
2861 doi:[10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0).
- 2862 [99] G. Cowan et al., “Asymptotic formulae for likelihood-based tests of new physics”,
2863 *European Physical Journal C* **71** (February, 2011)
2864 doi:[10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0), arXiv:[1007.1727](https://arxiv.org/abs/1007.1727).
- 2865 [100] CMS Collaboration, “Search for supersymmetry in final states with missing
2866 transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions
2867 using the variable α_T ”, *JHEP* **1301** (2013) 077,
2868 doi:[10.1007/JHEP01\(2013\)077](https://doi.org/10.1007/JHEP01(2013)077), arXiv:[1210.8115](https://arxiv.org/abs/1210.8115).
- 2869 [101] CMS Collaboration, “Search for new physics in the multijet and missing transverse
2870 momentum final state in proton-proton collisions at $\sqrt{s} = 8$ TeV”,
2871 arXiv:[1402.4770](https://arxiv.org/abs/1402.4770).
- 2872 [102] CMS Collaboration, “Search for new physics in events with same-sign dileptons
2873 and jets in pp collisions at $\sqrt{s} = 8$ TeV”, *JHEP* **1401** (2014) 163,
2874 doi:[10.1007/JHEP01\(2014\)163](https://doi.org/10.1007/JHEP01(2014)163), arXiv:[1311.6736](https://arxiv.org/abs/1311.6736).
- 2875 [103] CMS Collaboration, “Search for supersymmetry in pp collisions at $\sqrt{s} = 8$ TeV in
2876 events with a single lepton, large jet multiplicity, and multiple b jets”,
2877 arXiv:[1311.4937](https://arxiv.org/abs/1311.4937).

2878 Acronyms

| | | |
|------|--------------|--|
| 2879 | ALICE | A Large Ion Collider Experiment |
| 2880 | ATLAS | A Toroidal LHC ApparatuS |
| 2881 | APD | Avalanche Photo-Diodes |
| 2882 | BSM | Beyond Standard Model |
| 2883 | CERN | European Organisation for Nuclear Research |
| 2884 | CMS | Compact Muon Solenoid |
| 2885 | CMSSM | Compressed Minimal SuperSymmetric Model |
| 2886 | CSC | Cathode Stripe Chamber |
| 2887 | CSV | Combined Secondary Vertex |
| 2888 | CSVM | Combined Secondary Vertex Medium Working Point |
| 2889 | DT | Drift Tube |
| 2890 | ECAL | Electromagnetic CALorimeter |
| 2891 | EB | Electromagnetic CALorimeter Barrel |
| 2892 | EE | Electromagnetic CALorimeter Endcap |
| 2893 | ES | Electromagnetic CALorimeter pre-Shower |
| 2894 | EMG | Exponentially Modified Gaussian |
| 2895 | EPJC | European Physical Journal C |
| 2896 | EWK | Electroweak Sector |
| 2897 | GCT | Global Calorimeter Trigger |
| 2898 | GMT | Global MuonTrigger |
| 2899 | GT | Global Trigger |
| 2900 | HB | Hadron Barrel |
| 2901 | HCAL | Hadronic CALorimeter |

| | | |
|------|-------------|-----------------------------------|
| 2902 | HE | Hadron Endcaps |
| 2903 | HF | Hadron Forward |
| 2904 | HLT | Higher Level Trigger |
| 2905 | HO | Hadron Outer |
| 2906 | HPD | Hybrid Photo Diode |
| 2907 | ISR | Initial State Radiation |
| 2908 | LUT | Look Up Table |
| 2909 | L1 | Level 1 Trigger |
| 2910 | LEP | Large Electron-Positron Collidior |
| 2911 | LHC | Large Hadron Collider |
| 2912 | LHCb | Large Hadron Collider Beauty |
| 2913 | LO | Leading Order |
| 2914 | LSP | Lightest Supersymmetric Partner |
| 2915 | NLL | Next to Leading Logarithmic Order |
| 2916 | NLO | Next to Leading Order |
| 2917 | NNLO | Next to Next Leading Order |
| 2918 | POGs | Physics Object Groups |
| 2919 | PS | Proton Synchrotron |
| 2920 | QED | Quantum Electro-Dynamics |
| 2921 | QCD | Quantum Chromo-Dynamics |
| 2922 | QFT | Quantum Field Theory |
| 2923 | RBXs | Readout Boxes |
| 2924 | RPC | Resistive Plate Chamber |
| 2925 | RCT | Regional Calorimeter Trigger |
| 2926 | RMT | Regional Muon Trigger |

- ₂₉₂₇ **SUSY** SUperSYmmetry
- ₂₉₂₈ **SM** Standard Model
- ₂₉₂₉ **SMS** Simplified Model Spectra
- ₂₉₃₀ **SPS** Super Proton Synchrotron
- ₂₉₃₁ **TF** Transfer Factor
- ₂₉₃₂ **TP** Trigger Primative
- ₂₉₃₃ **VEV** Vacuum Expectation Value
- ₂₉₃₄ **VPT** Vacuum Photo-Triodes
- ₂₉₃₅ **WIMP** Weakly Interacting Massive Particle