

SEARCH FOR W' BOSON WITH BOOSTED AND HADRONIC
TOP-QUARK FINAL STATE IN pp COLLISIONS AT $\sqrt{s} = 13$ TEV
WITH THE ATLAS DETECTOR

By

Kuan-Yu Lin

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics — Doctor of Philosophy

2021

ABSTRACT

SEARCH FOR W' BOSON WITH BOOSTED AND HADRONIC TOP-QUARK FINAL STATE IN PP COLLISIONS AT $\sqrt{S} = 13$ TEV WITH THE ATLAS DETECTOR

By

Kuan-Yu Lin

With the LHC proton-proton collision data at an energy of 13 TeV recorded by the ATLAS detector, this analysis searches for the production of a new particle W' with a yet-to-be-probed mass in the top quark plus bottom quark decay channel. The boosted and hadronically decaying top-quark is reconstructed as a large-radius jet and identified using jet substructures. A secondary-vertex finding algorithm identifies small-radius jets originating from a bottom quark. After the identification stage, the main background events come from the Quantum Chromodynamical production of multi-jet, which is estimated by a data-driven approach. The statistical data analysis searches for a resonant peak in the invariant mass distribution of the system with a jet associated with the top quark and another with the bottom quark. The consistency between data and the background prediction is found in the invariant mass range between 1 TeV to 7 TeV, and an upper limit at the 95% confidence level is set on the production cross-section times branching ratio for a W' boson with right-handed chirality decaying to a top quark and a bottom quark. A lower limit on the W' mass is set at 4.4 TeV accordingly.

ACKNOWLEDGMENTS

First I would like to thank my PhD supervisor, Reinhard, for all the financial and mental support he has offered during my studies. The opportunities to stay at CERN and the Argonne National Laboratories gave me unique experiences of research and living. The numerous time he spent to review my thesis allowed me to strengthen my knowledges and convictions of doing experimental research. The consideration he gave to accommodate my visa situation and prolonged difficulties under COVID was incredible.

All other members in the ATLAS group help me substantially too. Thanks Wade for mentoring me from time to time and his elucidation in physics and statistics. Thanks Hector for all the assistance to my works and friendly discussions within and beyond physics. Thanks Garabed for the fun chat we had together. I hope I can share with you my latest reading on philosophy soon. Thanks Dan for your cheerfulness during all coffee breaks and beer time. Thanks Trisha for instant technical help many time and being a nice person sitting behind me in the office. Thanks Reiner for all the tips regarding traveling and hiking. Thanks Yuri for the companion during my L1Calo general meeting. Although Rui is strictly speaking not a student of MSU, I appreciate to have a like-minded person to work on the same analysis.

I would also like to thanks many other people I met at MSU, CERN, and Argonne that gave me memorable years during my PhD.

At last I want to thank my parents and my sister for helping me to face all kinds of difficulties in life.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
Chapter 1 Introduction	1
1.1 An Introduction to Particle Physics	1
1.2 The Road Forward	3
1.3 The Search for W'	5
1.4 The Structure of This Thesis	7
Chapter 2 From the Standard Model to New Physics: W'	9
2.1 The Standard Model	9
2.2 Beyond the Standard Model with W'	12
Chapter 3 Analysis Strategies	15
Chapter 4 Experimental Apparatus	19
4.1 Large Hadron Collider	19
4.2 ATLAS Detector	24
4.2.1 Inner detectors	26
4.2.2 Calorimeters	28
4.2.3 Muon spectrometers	33
4.2.4 Trigger and data acquisition	34
4.2.5 Level 1 calorimeter trigger	37
Chapter 5 Predicting the W' Signal	39
5.1 NLO Cross-Section Evaluation with ZTOP	39
5.2 Theoretical Uncertainties	44
5.3 Event Generation	47
Chapter 6 The Identification of Top and Bottom Quarks	53
6.1 Jet	53
6.2 Large-radius Jet — Local Cluster Topological Jet	55
6.2.1 Large-radius Jet Trigger	57
6.2.2 DNN top tagging	59
6.3 Small-radius Jet — Particle Flow Jet	63
6.3.1 DL1r b-tagging	65
6.4 Truth-level and Reconstruction-level Observables	67
Chapter 7 Event Selection and Categorization	70

Chapter 8 Background Processes	76
8.1 $t\bar{t}$	76
8.2 Data-driven Backgrounds	79
8.2.1 Top-tagging and b-tagging correlations	81
8.2.2 The systematic uncertainty	85
Chapter 9 Statistical Analysis	91
9.1 Systematic Uncertainties	92
9.1.1 Experimental uncertainties	92
9.1.2 Signal modeling uncertainties	94
9.1.3 Background modeling uncertainties	95
9.1.4 Propagation of uncertainties to the data-driven background	97
9.2 Profile-likelihood Method	97
9.3 The CLs method	99
9.4 Fit validations with pulls and constraints	100
9.5 Nuisance parameter ranking	101
Chapter 10 Statistical Analysis Results	103
10.1 Validating the Data-driven Background Estimation	103
10.2 Test the Signal Region Fit with the Background-only Pseudo-data	106
10.3 Exclusion Upper Limit with the Background-only Pseudo-data	116
10.4 The Signal Region Fit with Observed data	116
10.5 Exclusion Upper Limit with Observed Data	127
10.6 Systematic Uncertainty Rankings for a 4 TeV W'_R	129
10.6.1 Background-only pseudo-data	129
10.6.2 Signal-plus-background pseudo-data	131
10.6.3 Observed data	131
Chapter 11 Conclusion	135
APPENDICES	138
Appendix A The Definition of Electrons and Muons	139
Appendix B Insignificant Background Processes	141
BIBLIOGRAPHY	148

LIST OF TABLES

Table 2.1: Three generations of quark duplets.	12
Table 2.2: Three generations of lepton duplets.	12
Table 3.1: Four regions employed by a data-driven background estimation are shown as four cells on a two-dimensional plane. The horizontal and vertical axes are represented by the tagging decision on the b-candidate jet and the top-candidate jet, respectively.	18

LIST OF FIGURES

<p>Figure 1.1: The Feynman diagram shows a quark (q) from one proton (p^+) annihilates with an antiquark of another kind (\bar{q}') from the other proton to produces a W' boson. The latter then decays to a top quark (t) and an anti-bottom quark (\bar{b}). The vertical axis represents space and the horizontal axis points in time. The diagram is produced by feynMF [21, 22].</p> <p>Figure 2.1: The Feynman diagram shows parton showers from a vector boson's decay. Straight lines with an arrow are quarks. Curly lines are gluons. The vertical axis represents space and the horizontal axis represents time. Diagram cropped from [30].</p> <p>Figure 4.1: The CERN accelerator complex. The injection chain for proton collisions at the LHC is composed of LINAC 2, PS Booster (BOOSTER on the diagram), PS, and SPS. There are four detectors/experiments, ALICE, ATLAS, CMS, and LHCb marked as yellow dots along the LHC ring. [42]</p> <p>Figure 4.2: Increase of integrated luminosities over time of the data delivered by the LHC, recorded by the ATLAS detector, and marked as good for physics analysis. [44]</p> <p>Figure 4.3: The number of interactions per bunch crossing averaged over a luminosity block [45] is shown as integrated-luminosity-weighted distribution on the graph for each year of data taking. The legends tell the overall average for each year. [44]</p> <p>Figure 4.4: The figure points out the main subdetectors of the ATLAS detector. . .</p> <p>Figure 4.5: A simplified representation of a segment of a cross-sections of the ATLAS detector. The toroidal magnets around the muon spectrometer are omitted. The tracking system corresponds to the inner detector (ID). . .</p> <p>Figure 4.6: A 3D view of the barrel region of the inner detectors. [49]</p> <p>Figure 4.7: The ATLAS calorimeter encloses the ID colored in grey and a hard-to-see solenoid magnet around it. The ECAL consists of a liquid-Argon (LAr) barrel part ($\eta < 1.475$) and two LAr end-caps (EMEC, $1.375 < \eta < 3.2$). The tile calorimeter (barrel: $\eta < 1.0$ and extended barrel: $0.8 < \eta < 1.7$) and two LAr hadronic end-caps ($1.5 < \eta < 3.2$, painted in two sheds of copper) comprise the HCAL. At η between 3.1 and 4.9 on each side, there are three components of the LAr forward calorimeter (FCAL). [46]</p>	<p>5</p> <p>10</p> <p>21</p> <p>22</p> <p>23</p> <p>24</p> <p>25</p> <p>27</p> <p>29</p>
---	--

Figure 4.8: Segements of accordion-shaped calorimeters. The silver-colored lead absorbers are sandwiched by copper-clad Kapton electrodes. The electrode receives ionizing signals from the liquid Argon filled in the gap between lead absorbers. [51]	29
Figure 4.9: A wedge of tile hadronic calorimeters consisting of scintillator tiles interleaving with steel absorbers. The coordinate system shown on the bottom left follows the convention introduced at the beginning of this section. The top left and center left diagrams demonstrate how a waveshifter fiber transmitting light signals from scintillator tiles to a photomultiplier tube (PMTs). The top figure shows the PMT's position in the top drawer. The plot is taken from [20] which adapts the original one in [52].	30
Figure 4.10: The pion energy resolution of the hadronic and electromagnetic calorimeters combined is shown as a function of the square root of pion energy inverted. The black (white) dots are measured with test beams in year 1996 (1994), while the white cross is derived from the simulation of hadronic showers. The two fit functions differ in their description of detector noises. [53]	31
Figure 4.11: The p_T resolution of the large-radius jet as a function of the average of the di-jet transverse momentum from the <i>in-situ</i> calibration [54]. The symbol $ \eta_{\text{det}} $ means the absolute pseudorapidity of the jet axis with the detector's center taken as the coordinate's origin. The data (black dashed line) has a light band that sums the statistical and systematic uncertainties in quadrature. The red band is the envelope of <i>in-situ</i> calibration applied to three different QCD di-jet event generators. The yellow dashed line corresponds to the reconstructed jet p_T resolution relative to the truth-leveled jet (Chapter 6) computed from events simulated by the PYTHIA generator.	32
Figure 4.12: The muons detected by the MS are bent by the toroidal magnets in the barrel (green) and end-cap (black). With requirements of three layers of measurements, triggering (RPC and TGC) has a range of $ \eta < 2.4$, limited by the two outer TGC wheels. In comparison, precision-tracking (MDT and CSC) can go futher to 2.7. [56]	33

Figure 4.13: A schematic diagram of the Run-2 ATLAS trigger and data acquisition system. The top-right block shows the detector front-ends (FEs), some of which at the CAL and MS are connected to the corresponding calorimeter trigger and muon trigger at the first-level (L1) trigger block on the top-left block. The entire time consumed by data transmission and logic processing at the L1 is less than $2.5 \mu\text{s}$ before the central trigger flags the L1 Accept command back to the FE at a rate below 100 kHz. The accepted event's data from all detectors are then distributed through the ReadOut System (ROS). The L1 trigger also sends the RoI information to the High-level Trigger (HLT), which accesses the event data through the ROS. As shown on the bottom right block, data passing the HLT's threshold at a rate of about 1 kHz is transferred to permanent storage for offline analysis. (Figure adapted from [57])

35

Figure 4.14: The architecture of the L1Calo during Run-2 (2015-2018). All modules have output to Read-Out Drivers (RODs), allowing diagnostics, calibrations, and performance studies. In particular, RoIs from the CP and the JEP are routed through the CMX to the RODs for later retrieval by the HLT. RODs pack, buffer and send signals to the ROS. [62]

37

Figure 5.1: The leading order Feynman diagram for the signal's hard-scattering matrix element. This diagram's final state is referred as the parton level. Time points from left to right. The spatial dimension is vertical. For aesthetic purposes, the momentum flow is only accurate in the time direction. The charge-conjugated diagram is not shown but included in the event generation. The diagram is produced by `feynMF` [21]

40

Figure 5.2: Feynman diagrams of NLO terms in calculating the cross-section $\sigma(q\bar{q}' \rightarrow W' \rightarrow t\bar{b})$, where q and \bar{q}' represents an up-type quark and a down-type antiquark emerging from protons, respectively. The vertical (horizontal) axis corresponds to the space (time) axis. The charge-conjugated processes are included in the cross-section calculation. Feynman diagrams are made with `TikZ-Feynman` [64].

41

Figure 5.3: The cross-section $\sigma(pp \rightarrow W')$ times branching ratio $B(W' \rightarrow tb)$ with W' 's rest masses in $0.2\sim 5.0$ TeV is evaluated with ZTOP [65–67] at NLO of QCD. Figure (a) is for a left-handed chiral W'_L and (b) is for a right-handed chiral W'_R . The central black curve is the nominal value, including almost invisible statistical errors. The red band stands for the theoretical uncertainties. Interfaced by LHAPDF [68], the PDF set is `PDF4LHC15_nlo_mc_pdfas` documented by [69]. This set combines CT14 [70], MMHT14 [71], and NNPDF3.0 [72] with technologies found in [73, 74].

42

Figure 5.4: The relative theoretical uncertainty of $\sigma(pp \rightarrow W' \rightarrow tb)$ for different $m_{W'}$ and chirality. The uncertainty from shifting m_t up (red) or down (orange) by 1 GeV dominates at low $m_{W'}$. Larger $m_{W'}$ sees increases of uncertainty from both scaling μ ($= \mu_F = \mu_R$) by a factor of 0.5 (blue) or 2.0 (purple) and the one-sided PDF+ α_s uncertainty (green). The latter dominates in the highest region of $m_{W'}$	46
Figure 5.5: Parton-level (without FSR) invariant mass shapes of the top quark and bottom quark decaying from (a) left-handed or (b) right-handed W' with 5 different rest masses are compared by normalizing the area to 1. Each distribution is a cumulation of around 300K events.	48
Figure 5.6: (a) The top-quark transverse momenta and (b) the top-quark to bottom quarks pseudorapidity difference's absolute value for right-handed W' with 5 different rest masses are compared by normalizing the area to 1. Left-handed W' bosons share similar kinematic properties	49
Figure 5.7: (a) The angle between the down-quark's momentum in the top-quark's rest frame and the top-quark's momentum in the W' rest frame and (b) the maximum angular distance ($\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$) between the top-quark and its three decay products in the lab frame for a left-handed (blue) or right-handed (red) W' with a 4 TeV rest mass are compared by normalizing the area to 1. The rightmost bin in (b) is an overflow bin that encompasses everything above 1.	50
Figure 6.1: The probabilities of truth-level large-radius jets having a top quark and various combinations of the top decay products within $\Delta R < 0.75$ from the jet axis are shown as a function of the top quark's transverse momentum. A top quark decays first to a bottom quark (b) and a W -boson, which further decays to two first or second generation quarks (q_1, q_2). [80]	55
Figure 6.2: The efficiencies of three large- R jet triggers for data collected in the year 2015 (black), 2016 (red), and 2017 to 2018 (blue) are shown as a function of the (offline) leading large-radius transverse momentum. The legend lists each trigger's (online) transverse energy thresholds. The efficiency is computed by the coincidence of the targeted trigger and a lower threshold trigger – the (online) leading small-radius jet's $E_T > 260$ GeV. All three trigger efficiencies reach 100% at (offline) leading large-radius jet $p_T > 500$ GeV (drawn as a vertical black line). The shaded area represents statistical uncertainties.	58
Figure 6.3: Feynman diagrams of (a) the top-quark signal and (b) a light-quark background for the top-tagging algorithm.	59

Figure 7.1: The flow chart showing the classification of events that lead to different regions. The top input block's top-candidate jet means the top-tagged large-radius jets.	72
Figure 7.2: For each of the two categories, events are divided into six regions on a two-dimensional plane. The vertical axis represents the top-candidate jet's top-tagging score from low to high upward, while the horizontal axis represents the b-candidate jet's b -tagging score from low to high rightward. SR stands for the signal region, VR for the validation region, TR for the template region, and CR for the control region.	74
Figure 7.3: The absolute pseudorapidity difference between the top-candidate jet and the b-candidate jet for a 4 TeV right-handed W' signal and the two main backgrounds are compared in the signal region with the highest background rejection rate via the top-tagging and b -tagging (Signal Region 1 in Figure 7.2). Both the top-quark-pair ($t\bar{t}$) and QCD multi-jet backgrounds are simulated by MC generators (details in Chapter 8), and their sums are normalized to have unit area.	75
Figure 8.1: The Feynman diagrams of top quark pair production at the leading order of strong coupling. The vertical axis represents space, and the horizontal axis represents time. Feynman diagrams are made with TikZ-Feynman [64].	77
Figure 8.2: The invariant mass distribution of a large-radius top-candidate jet and a small-R b-candidate jet in region SR1 (right) and the neighboring region TR1 (left) (c.f. Figure 7.2). (a) The template region has the data plotted against two channels of $t\bar{t}$ background – the all-hadronic and the non-all-hadronic – shown as stacked histograms. Their statistical uncertainties are summed in quadrature into a single $t\bar{t}$ statistical error. The ratio of the summed $t\bar{t}$ distribution to the data is shown in the lower panel with statistical errors from both data and the $t\bar{t}$ background. (b) In the signal region, the data-driven background (see the next section) is drawn above the two $t\bar{t}$ channels with separate statistical errors. The ratio shown in the lower panel is the division of the $t\bar{t}$ to the total background. Covariances are taken into account in error propagations. On either figure's ratio plot, there is a horizontal reference line showing the ratio of total events.	78
Figure 8.3: The invariant mass M_{tb} distribution of simulated QCD multi-jet background in the signal region SR1 (black dot) is compared with its extrapolation (blue lines) from the template region TR1 and the control regions CR1 and CR2. The vertical axis is the number of events for a bin width of 200 GeV. The lower panel shows the ratio of the signal region histogram to the extrapolated one.	81

Figure 8.4: Examplified LO Feynman diagrams of the QCD multi-jet background at the LHC. The vertical axis represents space, and the horizontal axis represents time.	82
Figure 8.5: Figure (a) shows fractions of different parton flavors associated with the top-candidate/proxy jet as a function of the top-tagging score in nature logs. The top-tagging dependence of non- b -tagged jets are presented in figure (b). The last two bins' lower edges, corresponding to the 80% working point and the 50% working point, are not exactly the same as the label. The invariant mass M_{tb} is between 1.2 and 1.4 TeV for both figures. Light quarks include up, down, and strange quarks as their masses are below the QCD condensation scale. The charm and bottom quark fractions are multiplied by factors for viewing purposes.	83
Figure 8.6: (a) The b -candidate jet's parton flavor fraction as a function of the top-candidate/proxy jet's top-tagging score's natural log in the “1 b -tag in” category. The last two bins' lower edges, corresponding to the 80% working point and the 50% working point, are not the same as the label. The invariant mass M_{tb} is between 3 TeV and 3.4 TeV. Non- b -tagged b -candidate jets are shown for reading the b -tag ratio variations. (b) The b -tagging ratio (blue with left vertical-axis) and top-tagging ratio (red with right vertical-axis) as a function of the b -candidate jet p_T in the same M_{tb} range and category. The top-tagging ratio is between the 50% WP top-tagged region and the upper control region (CRXa, X = numbers), as defined in Figure 8.7. The b -tagging ratio is also calculated in the control region.	85
Figure 8.7: The categorization in the control regions (CRs) defined in Figure 7.2 is refined for computing the data-driven background's systematic uncertainties. The two new thresholds are placed on e^{-4} and e^{-7} of the top-tagging DNN score.	86
Figure 8.8: Validating plots for the data-driven background's systematic uncertainties by the simulated QCD multi-jet background. In each diagram, the “SR” histogram is obtained directly from the top-tagging and b -tagging, and the extrapolation is the result of the data-driven method. The red hatches show the statistical uncertainty of the extrapolation. The error bars on the lower panel only include the statistical uncertainty for the SR distribution. The purple triangular dots show the correlation factor obtained in the control region of each “ b -tag in” category, yielding the cyan-colored systematic uncertainties after symmetrization (same as a reflection) and smoothing procedure.	89

Figure 8.9: The data-driven background’s systematic uncertainties. The extrapolation is the data-driven background derived from data subtracting the $t\bar{t}$ ’s simulated events. Statistical uncertainties are drawn separately for the extrapolation and the SR histogram as in the previous figure. The purple triangular dots show the correlation factor obtained in the control region of each “b-tag in” category, yielding the cyan-colored systematic uncertainties after symmetrization (same as a reflection) and smoothing procedure.	90
Figure 10.1: Data and predicted background M_{tb} distribution in the validation region (top-tagged at the 80% but below the 50% working point and in the “0 b-tag in” category) (a) before and (b) after the profile-likelihood fit. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.	104
Figure 10.2: The pull and constraint plots of the data-driven-background systematic uncertainties in the background-model fit to data in the validation region. The central value (θ_0) is the pre-fit value, which is 0 by construction. For each nuisance parameter, the black dot shows the post-fit value ($\hat{\theta}$) with the post-fit uncertainty as its error bar, which is in general larger than the pre-fit uncertainty of unit size ($\Delta\theta$).	105
Figure 10.3: Predicted background M_{tb} distribution in the signal regions: (a) SR1, (b) SR2, and (c) SR3 before the profile-likelihood fit. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.	107
Figure 10.4: Predicted background M_{tb} distribution in the signal regions: (a) SR1, (b) SR2, and (c) SR3 after the profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.	108
Figure 10.5: The nuisance parameter pulls for jet energy scale uncertainties common to large-radius jets and small-radius jets after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” “EtaInter” means the <i>in situ</i> calibrations across different η regions of the detector.	109

109

110

Figure 10.8: The nuisance parameter pulls for large-radius jet mass scale and resolution uncertainties after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” “Rtrk” is the ratio between jet mass estimated by tracks to that measured by calorimeters. The difference in Rtrk between simulation and data serves as a systematic uncertainty. “COMB” refers to “combined mass,” a mass observable combining the tracking and calorimeter measurements. 110

111

112

112

Figure 10.12: The nuisance parameter pulls for the modeling uncertainties of the top-quark pair background after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The term “shower,” “pdf,” “muR,” “muF,” “ME,” “ISR,” “FSR” refer to the parton shower generator, Parton Distribution Function, renormalization scale, factorization scale, initial state radiation, and final state radiation respectively. These terms are described in section 2.1 and subsection 9.1.3. The region code (Figure 7.2) at the end further specifies the signal region where the $t\bar{t}$ background is to be varied. Nuisance parameters of “DataDriven (DataDriv)” account for the uncertainty from subtracting the $t\bar{t}$ background in the specified signal region’s neighboring template region (the b-candidate jet is not <i>b</i> -tagged rather than tagged) in the data-driven background estimation.	113
Figure 10.13: The nuisance parameter pulls for the 1.7% luminosity uncertainty and the 4% top-quark pair background production cross-section uncertainty after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.”	114
Figure 10.14: The correlation matrix between nuisance parameters after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The matrix is symmetric over the diagonal from the top left to the bottom right corner. Only nuisance parameters with more than 20% of correlation with another are shown. The meaning of individual nuisance parameters can be found in the preceding Figure 10.5 to Figure 10.13.	115
Figure 10.15: The upper limit on the W'_R production cross-section times the decay branching ratio of $W'_R \rightarrow tb$ at the 95% Confidence Level is evaluated with the background-only pseudo-data. The blue dashed curve indicates the median, while the green and yellow bands show the one and two Gaussian quantiles of the limit under statistical plus systematic uncertainties. The Next-to-Leading-Order theoretical cross-section times the branching ratio is drawn as the red dashed curve with an asymmetric theoretical uncertainty band (Chapter 5).	117
Figure 10.16: Data versus background M_{tb} distributions in the signal regions: (a) SR1, (b) SR2, and (c) SR3 before the profile-likelihood fit. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.	118
Figure 10.17: Data versus background M_{tb} distribution in the signal regions: (a) SR1, (b) SR2, and (c) SR3 after the profile-likelihood fit with under the background-only hypothesis. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.	119

Figure 10.18: The nuisance parameter pulls for jet energy scale uncertainties common to large-radius jets and small-radius jets after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.5.	120
Figure 10.19: The nuisance parameter pulls for large-radius jet energy scale and resolution uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.6.	121
Figure 10.20: The nuisance parameter pulls for small-radius jet energy scale and resolution uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.7.	121
Figure 10.21: The nuisance parameter pulls for large-radius jet mass scale and resolution uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.8.	122
Figure 10.22: The nuisance parameter pulls for the top-tagging uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.9.	122
Figure 10.23: The nuisance parameter pulls for the b -tagging uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.10.	123
Figure 10.24: The nuisance parameter pulls for the modeling uncertainties of the data-driven background estimation after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.11.	123
Figure 10.25: The nuisance parameter pulls for the modeling uncertainties of the top-quark pair background after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.12.	124
Figure 10.26: The nuisance parameter pulls for the 1.7% luminosity uncertainty and the 4% top-quark pair background production cross-section uncertainty after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$).	124

125

Figure 10.28: The upper limit on the W'_R production cross-section times the decay branching ratio of $W'_R \rightarrow tb$ at the 95% Confidence Level is evaluated with data. The blue dashed curve indicates the median expected limit, while the green and yellow bands show the one and two Gaussian quantiles under statistical plus systematic uncertainties. The black solid curve shows the observed limit. The Next-to-Leading-Order theoretical cross-section times the branching ratio is drawn as the red dashed curve with an asymmetric theoretical uncertainty band (Chapter 5).

128

Figure 10.29: The nuisance parameter rankings for the 4 TeV W'_R signal with respect to the pseudo-data with “ $\mu = 0$.” Color-filled bands show the fifteen-highest impact on the best-fit μ ($\Delta\mu$) by fixing a particular nuisance parameter to its best-fit value plus-or-minus the post-fit error ($\hat{\theta} \pm \Delta\hat{\theta}$). The unfilled bands with colored boundaries correspond to $\Delta\mu$ when the nuisance parameter is shifted by its pre-fit uncertainty ($\hat{\theta} \pm \Delta\theta$). The scale of $\Delta\mu$ is shown on the top horizontal line. The bottom horizontal scale concerns with the black dots and bars, the best-fit nuisance parameters and their post-fit errors, as in the pull plots. The captions from Figure 10.5 to Figure 10.13 explain the naming convention of the nuisance parameters. 130

Figure 10.30: The nuisance parameter rankings for the 4 TeV W'_R signal with respect to the pseudo-data with “ $\mu = 1$.” Color-filled bands show the fifteen-highest impact on the best-fit μ ($\Delta\mu$) by fixing a particular nuisance parameter to its best-fit value plus-or-minus the post-fit error ($\hat{\theta} \pm \Delta\hat{\theta}$). The unfilled bands with colored boundaries correspond to $\Delta\mu$ when the nuisance parameter is shifted by its pre-fit uncertainty ($\hat{\theta} \pm \Delta\theta$). The scale of $\Delta\mu$ is shown on the top horizontal line. The bottom horizontal scale concerns with the black dots and bars, the best-fit nuisance parameters and their post-fit errors, as in the pull plots. The captions from Figure 10.5 to Figure 10.13 explain the naming convention of the nuisance parameters. 132

Figure 10.31: The nuisance parameter rankings for the 4 TeV W'_R signal with respect to the observed data. Color-filled bands show the fifteen-highest impact on the best-fit μ ($\Delta\mu$) by fixing a particular nuisance parameter to its best-fit value plus-or-minus the post-fit error ($\hat{\theta} \pm \Delta\hat{\theta}$). The unfilled bands with colored boundaries correspond to $\Delta\mu$ when the nuisance parameter is shifted by its pre-fit uncertainty ($\hat{\theta} \pm \Delta\theta$). The scale of $\Delta\mu$ is shown on the top horizontal line. The bottom horizontal scale concerns with the black dots and bars, the best-fit nuisance parameters and their post-fit errors, as in the pull plots. The captions from Figure 10.5 to Figure 10.13 explain the naming convention of the nuisance parameters. 133

Figure 10.32: The nuisance parameter pulls for the modeling uncertainties of the data-driven background estimation after a profile-likelihood fit to data with μ determined by the fit. The meaning of the nuisance parameters are explained in Figure 10.11. 134

Figure B.1: The Feynman diagrams for hard-scatter matrix elements of the single-top processes at the LHC. The vertical axis represents space and the horizontal axis represents time. The diagrams above implicitly include three colors of quarks, eight types of gluons, and the combination rules between colors and gluon types. The q and q' is a duo of up-type and down-type quarks in the first two generation. The four diagrams are labelled by the virtual particle's Mandelstam variables and the final state particles. Feynman diagrams are made with TikZ-Feynman [64]. 142

Figure B.2: The Feynman diagrams for hard-scatter matrix elements of the V +jet processes at the LHC. The vertical axis represents space and the horizontal axis represents time. The diagrams above implicitly include three colors of quarks, eight types of gluons, and the combination rules between colors and gluon types. For the W -boson diagrams, q and q' is a duo of one of the two lighter up-type quark and one down-type quark. For the neutral boson diagrams, q is any quark but the top quark. γ^* means a photon carrying nonzero mass, i.e., being off-shell. Even though the initial state's flavor and color degrees of freedom are reminiscent of the QCD multi-jet events, the small electroweak coupling suppresses the vector boson production relative to the former background. Feynman diagrams are made with TikZ-Feynman [64]. 144

Figure B.3: (a) The invariant mass distribution of a large-radius top-candidate jet and a small-R b-candidate jet in one of the template regions. The template region has the data plotted with two types of V +jets distributions. The W +jets channel, which has a W boson and some high p_T partons in the final state, is stacked on top of the Z/γ^*+ jets channel, where the W boson is replaced with a neutral vector boson. The statistical uncertainties of the two channels are summed in quadrature into a single V +jets statistical error. The ratio of the V +jets to data is shown in the lower panel with the statistical errors of V +jets and data. There is a horizontal line given by the ratio between total area. (b) The b -tagging ratio – the ratio of the signal region to the template regions shown in (a) – is shown as a function of M_{tb} . Since simulated events of V +jets are scarce, the two channels are added together, and the number of bins is reduced. The b -tagging ratio obtained in the upper control region (top-tagging DNN $> e^{-4}$ while below the 80% working point) with $t\bar{t}$ -subtracted data is compared. The hatched area represent the latter’s statistical error, which is minuscule compared with the V +jets. The bottom panels shows the division of the two ratios with the statistical uncertainties propagated. . 146

Chapter 1

Introduction

1.1 An Introduction to Particle Physics

The history of particle physics is an ongoing record of humans uncovering the fundamental principles and constituents of nature. At the beginning of the twentieth century, physicists raised some questions when they obtained inexplicable results from observing the inner structures of atoms [1] and decays of particles originating from cosmic rays or their interactions with the earth's atmosphere [2].

A question concerning the atomic structure at the time was, “how could the electrons in an atom move around the nucleus without falling into it by losing their energy to electromagnetic radiation according to classical electrodynamics?” Meanwhile, it was mysterious that the energy and electric charge conservations were insufficient to predict if a decay process could take place or not. Moreover, some particles had unexplainedly longer mean lifetimes than others.

These puzzles, at first, only had crude solutions rather than self-contained ones. Regarding the atomic structure, one of the initial hypothesis was that electrons in an atom behaved like standing waves whose wavelengths were inversely proportional to their momenta. Regarding the questions of particle decays, physicists assigned “quantum numbers” to each type of particle and hypothesized that null observations or rare occurrences of a decay process

were due to nonconservations of quantum numbers. As an example, a proton cannot decay to an electron and a photon since this process violates both the lepton number and baryon number conservation. Also, Kaons are known to decay infrequently because all of their decay processes do not conserve “the strange number.”

Later, revolutionary theories arrived one by one, providing quantitative explanations that incorporated more phenomena at once. To start with, the wave-like nature of electrons, together with the particle-like nature of light, indicated by the photo-electric effect, for example, was unified in quantum mechanics. Under quantum mechanics, positions and momenta of matter cannot be measured arbitrarily precisely at the same time. Instead, only their distributions from measurements are predicted by the theory.

The generalization of quantum mechanics to field theories, together with the gradual understanding that all fundamental interactions are mediated by gauge bosons, finally led to the birth of the Standard Model (SM). This general theory sheds light on the various structures behind the particle decays. Take the slow decays of Kaons as an example. It turns out that each Kaon consists of a strange quark (the identity of the strange number) and an up quark or a down quark. When a Kaon decays, all of the lighter particles it can turn into do not contain a strange quark. As a result, the strange quark has to turn into an up quark via the weak interaction carried by the W boson. Since the W boson is a massive particle, this interaction has a short effective range, leading to a low likelihood of the Kaon decay.

The SM has brilliantly passed numerous experimental confirmations. The top quark [3, 4] and the bottom quark [5] are revealed at Fermilab in the United States. Across the Atlantic, the European Council for Nuclear Research (CERN) has hosted experiments that discovered other particles predicted by the SM, such as the W and Z bosons found by the UA1 [6,

7] and UA2 [8, 9] collaborations and the Higgs boson by the ATLAS [10] and CMS [11] collaborations in 2012. Since then, the ATLAS and CMS collaborations have collected more data and have done an enormous amount of researches based on them. All of the available results are still consistent with the predictions of the SM.

1.2 The Road Forward

Coming through all the glory days of advancement, could it be said that the tale of particle physics has come to an end? Certainly not! The SM’s foundation, Quantum Field Theory, gives infinite quantum corrections to probability amplitudes when integrating to infinitely high energy [12]. Even though (through a technique called renormalization) this problem can be removed systematically such that the theoretical values of lifetimes and interaction rates of particles become finite [12], this constraint leads to beliefs that Quantum Field Theory cannot be valid at arbitrarily high energy scales. Instead, some other new physical models will be required to explain new phenomena that might appear at high energy scales. Planck scale, a high energy scale at 10^{19} GeV [13], where gravity has similar strength as the other fundamental interactions formulated by the SM, is regarded by many as a definite point for the SM to be inadequate.

Even though the Planck scale foretells the existence of physics beyond the SM, it is around 10^{15} times more than the energy achievable by the Large Hadron Collider (LHC) at 13 TeV ($=1.3 \times 10^4$ GeV) today, leaving no indication of new physics accessible in the foreseeable future. Even though there are other phenomena beyond the SM’s scope, such as neutrino oscillation and dark matter, the new physics behind them can appear at a very high energy scale while being consistent with current experimental results [14]. The only strong

motivation for new physics beyond the SM at the LHC energy is the “hierarchy problem.” It means the energy-dependent quantum corrections to the Higgs mass at several TeV are significantly larger than the experimental value [15]. In other words, new physics appearing at a higher energy scale could not easily contrive to produce the SM Higgs’ mass [14]. But since the hierarchy problem is a conceptual one, it does not entail any experimental discrepancy which requires physics beyond the SM at the LHC.

Despite the lack of indication, experimentalists at the LHC still search for new physics models that solve the SM’s conceptual problems and embed it into a more comprehensive theory. For example, Supersymmetry has been probed by many experimentalists at the LHC since it could alleviate the hierarchy problem, unify all three interactions in the SM at the so-called Grand Unification Theory (GUT) scale of 10^{16} GeV [13] and incorporate candidates of dark matter. Though searches for Supersymmetry resemble those for the electroweak gauge bosons and the Higgs boson predicted by the SM, the former is more difficult due to the abundance of insufficiently constrained parameters in typical Supersymmetric models.

In addition to the direct searches for new physics, “precision measurements,” which include measuring the parameters of the SM and counting events produced at a low rate by the SM processes, have been done at the LHC. Through a laborious effort to minimize uncertainties, they are immensely sensitive to potential inconsistencies between data and the SM. The power of precision measurements is seen from the discovery of CP violation in weak interactions via measuring the frequencies of a rare Kaon decay $K_L \rightarrow \pi\pi$ [16] that happened much earlier than the first theoretical account [17] for it.

“Generic searches” are methods to facilitate the quest for new physics as well. The idea is to find out new physics without discerning the exact origin of it. One example of this is searching for a new type of particle appearing in many new physics models. If

the LHC has enough energy to produce the particle, there will be a peak, referred to as a resonance, in the invariant mass distribution of the particles originating from its decay. Discoveries of new particles by resonances occurred many times in the twentieth century, such as the J/Ψ particle at SLAC [18] and Brookhaven [19]. Since the LHC collides protons at unprecedently high energy, it facilitates the generic search for any particle with a mass much higher than the most massive SM particle, the top quark, weighing 173 GeV [20].

1.3 The Search for W'

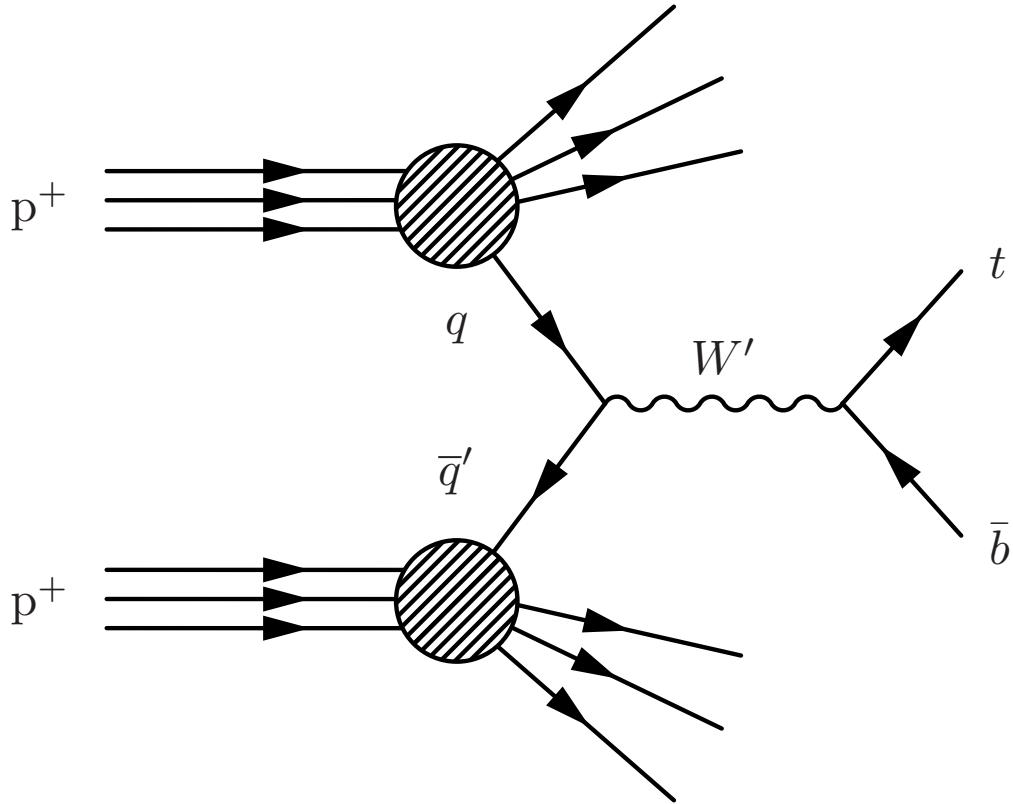


Figure 1.1: The Feynman diagram shows a quark (q) from one proton (p^+) annihilates with an antiquark of another kind (\bar{q}') from the other proton to produces a W' boson. The latter then decays to a top quark (t) and an anti-bottom quark (\bar{b}). The vertical axis represents space and the horizontal axis points in time. The diagram is produced by **feynMF** [21, 22].

One of the massive particle that might be found by resonance searches at the LHC is the

W' boson. It is a one-unit-charged and spin-one particle, which is produced by annihilation between quarks inside the smashed protons at the LHC. This thesis looks for an above-a-TeV-mass W' boson through its decay to a top quark and a bottom quark (Figure 1.1). From a theoretical viewpoint, W' could be one of the particles emerging in composite Higgs models or an extra-dimensional state of the W boson [20]. The composite Higgs models resolve the hierarchy problem by providing a natural cut-off to the quantum corrections to the Higgs masses [13]. Some extra-dimensional theories can reduce the Planck scale such that the quantum corrections are naturally small [13].

Searching for the W' boson via the final state of a top quark and a bottom quark is valuable as there are models where the W' only couples to quarks [23] or even preferentially to the third generation fermions [24]. Furthermore, the final state has unique features helping to reduce SM backgrounds. Top quarks decay predominantly via weak interactions to a W boson and a bottom quark. About two-thirds of the W boson turns further into two quarks [20] – the hadronic mode targeted by this analysis. Since the W' boson mass of interest is about ten or more times the top quark mass, the top quark's decay products constitute a Lorentz-boosted large-radius jet with substructures distinguishable from ordinary jets [25]. The bottom quark is also identifiable as a bottom-quark-containing hadron has a longer lifetime and higher mass than most other types hadrons [20].

Searches for W' boson productions with a t and b quark final state had been done at the Tevatron. The maximum excluded mass ranges of a right-handed W' from CDF [26] and DØ [27] collaborations are 800 and 885 GeV. They made use of data recorded when the Tevatron collided a proton beam with an anti-proton beam at a center-of-mass energy of 1.96 TeV. Recently, having data recorded with proton-proton collisions at a center-of-mass energy of 13 TeV at the LHC, the ATLAS [28] and CMS [29] collaborations pushed the mass

limit further to 3.25 and 3.6 TeV each.

Extending the search for the W' decaying a top quark and a bottom quark, the full Run-2 dataset recorded with the ATLAS detector at a collision energy of 13 TeV is analyzed in this thesis. In each event, the boosted and hadronic decay of the top quark is identified using a large-radius jet. And the bottom quarks coming from the W' and the top quark are selected using small-radius jets. The signal region is established by applying a top-tagging algorithm to the large-radius jet and a b -tagging algorithm to a back-to-back small-radius jet. The presence or not of any small-radius jet fulfilling a threshold of the b -tagging algorithm within the large-radius jet classifies the events into two categories

Despite the two algorithms, the signal regions still have a substantial amount of SM background events with (or resembling) a boosted top jet accompanied by a bottom quark. Some background events have a pair of top quarks in their final states, which can be estimated by Monte Carlo (MC) simulations. While the most significant background, the production of multi-jet events via the QCD strong interaction, is modeled by extrapolation from the control regions, defined by inverting either one or both of the two tagging requirements for each category. Four-momenta of the large-radius jet and the small-radius jet are summed into a W' boson proxy. The resulting mass spectrums in the signal regions are fitted in a profile-likelihood analysis to set exclusion limits on the cross-section for the production of a W' boson with decay to a top quark and a bottom quark.

1.4 The Structure of This Thesis

This thesis details the analysis, which is separated into eleven chapters, excluding the appendices. After this introduction, the SM and some of its extensions containing the W'

boson is summarized in Chapter 2. Chapter 3 will be about the analysis strategy to illuminate the arrangement of the remaining chapters. In Chapter 4, the LHC and the ATLAS detector will be introduced. Afterward, Chapter 5 will describe both the next-to-leading order (NLO) cross-section predictions and the leading order (LO) MC simulation of the W' signal. The conceptual bases of the top- and b -tagging algorithms and their applications are mentioned in Chapter 6. Chapter 7 will then provide each step in the event selections and categorization, followed by the descriptions of the background processes present in the signal regions in Chapter 8. Chapter 9 will discuss the statistical analysis methods and give a summary of systematic uncertainties. The analysis results in Chapter 10 include fitted W' mass distributions and exclusion limits on the W' production cross-section times the decay branching ratio. This thesis will conclude in Chapter 11.

Chapter 2

From the Standard Model to New Physics: W'

2.1 The Standard Model

The SM relates all high-energy and short-distance physics to interactions among a small number of elementary particles. These are eleven spin-one force-carrying gauge bosons, twelve spin-one-half fermions, and a complex-doublet scalar boson, the Higgs boson. The twelve fermions are further divided into six types of quarks and six types of leptons. All elementary particles are quantum fields, i.e., they have quantum properties at each four-dimensional spacetime coordinate.

The SM gauge bosons are exchanged between elementary particles, producing three kinds of fundamental interactions. The eight gluons are responsible for the strong interaction, the photon for the electromagnetic interactions, and the W and Z bosons for the weak interaction. In simplified mathematical notation, a fermion field $\psi(x)$ couples with a gauge boson $A(x)$ by modifying the spacetime derivatives of the former into

$$i\partial\psi(x) - egA(x). \quad (2.1)$$

(Spin, spacetime indices and quantum operator symbols are suppressed.) The gauge symmetry determines the allowed values of the quantum number e , and the coupling constant g controls the interaction strength. If $|g| \ll 1$ – the interaction is weak – quantum amplitudes can be well-approximated by a perturbative series in the coupling constant.

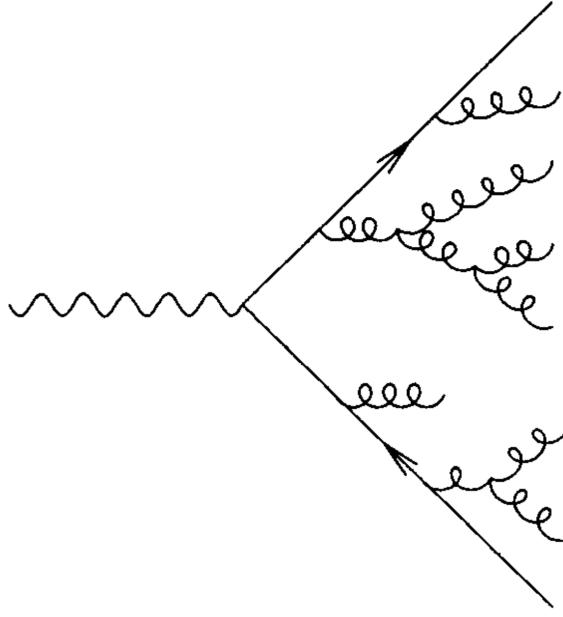


Figure 2.1: The Feynman diagram shows parton showers from a vector boson’s decay. Straight lines with an arrow are quarks. Curly lines are gluons. The vertical axis represents space and the horizontal axis represents time. Diagram cropped from [30].

The strong interaction acts only on quarks and gluons as they carry the “color” quantum numbers. Due to vacuum polarization, the strong interaction’s coupling constant g_s has less strength in shorter distances, or equivalently, at higher energy scales [30]. At an energy scale as low as typical hadron masses around 1 GeV, the strong force that holds quarks and gluons into hadrons is too strong to be considered a small perturbation. But when the energy increases, the strong coupling constant asymptotically approaches zero, whereby quarks and gluons become separate entities exchanging gluons. Therefore, high energy events at the LHC can be viewed as quarks and gluons – or partons – emerging from protons and undergo

“hard-scattering” which is perturbatively calculable (shown as Feynman diagrams).

QCD describes the conversion between partons and hadrons in both directions. For partons inside hadrons such as protons, their momentum fractions relative to the parent hadron follow the probabilistic distribution called Parton Distribution Functions (PDF) whose scale variation is described by parton radiations. In a hard-scattering interaction, initial-state partons emit partons before they interact with partons emerging from a counter-moving hadron. Also, final-state partons lose energy through a cascade of parton emission, called the parton showering (Figure 2.1). After each QCD radiation, the average parton energy decreases, making the g_s strong enough to bound partons back into hadronic states. The top quark is an exception, as it decays more rapidly than hadron formation time-scales.

The electromagnetic (EM) interaction is present between electrically charged particles, which include all of the quarks, half of the leptons, and the W boson. As the photon has zero mass and the EM interaction has weaker dependence on the energy scale than the strong interaction, it has an infinite interaction range.

Every SM particle is subject to the weak interaction. But as its name suggests, the weak interaction does not pose a significant force at lower energy. The reason is that the W and Z bosons are massive, equivalent to very short Compton wavelengths. While being less powerful than the strong and EM interaction, it has many unique behaviors. As an example, the W boson only couples to left-handed fermions, violating the parity conservation. Also, the W boson connects an up-type fermion to another down-type fermion. For this reason, the six quarks (Table 2.1) and six leptons (Table 2.2) are commonly displayed in three generations of duplets.

The last piece in the list, the Higgs boson, plays the role of giving masses to all quarks, charged leptons, the W and Z bosons, and itself. This role is necessary because the gauge

type	generations		
up	u	c	t
down	d	s	b

Table 2.1: Three generations of quark duplets.

type	generations		
up	ν_e	ν_μ	ν_τ
down	e	μ	τ

Table 2.2: Three generations of lepton duplets.

symmetry forbids the W and Z boson to be massive and because the chiral symmetry is contradictive to explicit fermion masses. Even though the physical Higgs boson has been discovered at the LHC, the Higgs field’s self-interaction, the Higgs potential, remains a theoretically intriguing. The Higgs potential has a coefficient with a dimension of energy squared. This coefficient receives quantum corrections proportional to the square of the energy scale. Therefore, if new physics appears at an energy scale much higher than the electroweak scale, such as the Planck scale, the model parameters have to be highly fine-tuned to accommodate such a gigantic scale difference. The dissatisfaction toward the fine-tuning is usually referred to as the “hierarchy problem.” Solving this problem is one of the motivations for many new physics models.

2.2 Beyond the Standard Model with W'

There are many open questions in the universe to be answered by physics “beyond the SM.” These questions surrounds neutrino mass, quantum gravity, matter-antimatter asymmetry, and origins of dark matter. Motivated by the hierarchy problem, many new physics models beyond the SM predict new particles and interactions discoverable at the LHC [13], such as

massive vector bosons, referred to as the W' boson in this thesis. Some examples are given in the following.

The first example is the composite Higgs model. Composite Higgs models (see section 94 of Ref. [20] and references therein) postulate that the Higgs bosons are bound states of fermions. They can suppress the hierarchy problem if the binding energy appears not too far above the electroweak scale. The W' boson can arise at the energy scale that resolves the composite Higgs' constituents, analogous to asymptotic free partons in QCD. But since electroweak precision experiments show no sign of a new strong force below 10 TeV, a minimal composite Higgs model is insufficient [13]. The "Little Higgs" model (see Ref. [31] and references therein) proposes a remedy to the composite Higgs by removing the one-loop order quantum corrections to the coefficient responsible for the hierarchy problem. The cancellation is made possible by enlarging the electroweak symmetry group into which the W' boson fits.

A different kind of solution to the hierarchy problem involves adding extra spatial dimensions to the model. Some extra-dimensional models try to explain the weakness of gravitational interactions by hypothesizing that gravitational potential extends to more spatial dimensions than three other fundamental interactions [13]. Randall-Sundrum models [32] take a different approach where the added fourth spatial dimension is warped. The warped space has one end residing at the Planck scale, the UV brane, and the other at the electroweak scale, the IR brane. The Higgs bosons can be confined to the IR brane such that the underlying parameter, the warp factor, does not have to be fine-tuned. In some variants of the Randall-Sundrum models (see Ref. [33] and references therein), the W boson has excited states localizing in the region between the two branes. These massive W bosons can be interpreted as the W' boson.

W bosons naturally emerge as a consequence of an additional SU(2) symmetry. It is present in the left-right symmetric model [34] or intermediate states of symmetry breaking in some grand unified theories [35]. In the left-right symmetric model, the W' boson has right-handed chirality. If the right-handed neutrino predicted by the model is more massive than the right-handed W' boson, its resonance state will only decay to quarks, motivating hadronic searches for the W' boson. In addition, in models having one SU(2) symmetry for the first two generations of fermions and the other for the third generation before the symmetry breaking (such as in [24, 36]), the top quark and the bottom quark, the target of this analysis, have an enhanced coupling to the W' boson.

Chapter 3

Analysis Strategies

The current chapter has two primary functions. First, it serves as an extended outline, emphasizing the logical connections between later chapters. It also gives the general idea of the data analysis process that involves understanding the experimental apparatus and techniques, invoking physics to interpret data, and providing statistical interpretation for the presence of the W' boson.

Chapter 2 has motivated a search for a W' boson via its production in proton collisions at the LHC and its decay to a top quark and a bottom quark. These two quarks have unique properties allowing them to be differentiated from other quarks. They are observed as jets, reconstructed by the ATLAS detector. The jet four momenta for the top quark and bottom quark will be summed as the W' four-momentum, whose mass is the invariant mass M_{tb} . With only SM physics, the M_{tb} distribution is expected to have a smoothly descending shape. If a W' with a 4 TeV mass is found in data, there should be a peak-like excess around 4 TeV.

Jets are useful observables for partons (quarks and gluons) in high-energy proton-proton collisions (Chapter 6). High energy partons lead to collinear parton showers (Chapter 2). Thus, hadrons formed by showering partons are confined in a cone-shaped region around the original parton's moving direction, giving an image of a jet. This jet of hadrons is observed as energy clusters by the calorimeter (Chapter 4), and the direction and energy of calorimeter

clusters give the jet four-momentum. This momentum is corrected back to the initial parton momentum up to differences due to detector resolutions and pile-up contaminations. Pile-ups are additional proton interactions in the same bunch-crossing as the hard-scattering of interest (Chapter 4). In the later text, a top-quark jet will mean a jet produced by a top quark. The same rule applies to other quarks and gluons.

A jet can be thought of as a cone with a circular base, whose radius characterizes its size (Chapter 6). This analysis defines two jet sizes: 0.4 and 1.0, named small-radius jets and large-radius jets. The small-radius jet is sufficient to enclose the parton shower from the bottom quark decaying from the W' ¹. However, it is challenging to assign one small-radius jet to each of the three partons from the top quark decay, as their showers might overlap due to the top quark’s Lorentz boost. For a W' boson with a mass of 3 TeV (the upper excluded W' mass in the last search), the maximally 1.5 TeV energy received from the W' decay over the top quark mass of less than 0.2 TeV gives a significant Lorentz boost factor γ . Therefore, large-radius jets, broad enough for enclosing hadrons induced by the top quark’s three decay particles, are reconstructed along with small-radius jets.

Reconstructed jets are examined by dedicated tagging algorithms to identify the flavor of the jet-emitting parton. A top-tagging algorithm (top-tagger) seeks large-radius jet attributes from a top quark decay, such as a prominently high mass and the three-prong-like energy density distributions. A b -tagging algorithm (b -tagger) finds patterns of a B hadron decay among tracks reconstructed from the ATLAS inner detector (Chapter 4). Both taggers are augmented by Deep Neural Network techniques capable of discerning non-trivial correlations between multiple input variables. A large (small)-radius jet is considered top (b)-tagged when the tagging algorithm’s output is in a range preferred by a top-quark (bottom-quark)

¹to be distinguished from the bottom quark decaying from the top quark

jet.

With jets as observables, the quest for W' bosons starts with removing proton-collision events (or events in short) due to SM backgrounds while preserving W' signal events. Signal regions (SR) are defined with event selection criteria favoring signal events. For example, the top quark and bottom quark decaying from the W' tend to have higher transverse momenta² (Chapter 5), and so should the corresponding top-quark and bottom-quark jets. Thus, the signal region first requires exactly one top-tagged large-radius jet with high transverse momentum (p_T) – the top-quark jet’s candidate (top-candidate jet). Secondly, on the far side of the top-candidate jet, the highest p_T small-radius has to be b -tagged – the bottom-quark jet’s candidate (b-candidate jet). The full list of selection criteria can be found in Chapter 7. The selections are optimized using the signal (Chapter 5) and background (Chapter 8) models without examining data in the SR to avoid human biases.

The dominant SM backgrounds in the SR are the top-quark pair production ($t\bar{t}$) and the QCD production of multiple non-top-quark jets (QCD multi-jet). Simulated events are adopted to obtain the SR distribution of the $t\bar{t}$ background. In contrast, as the QCD multi-jet events are on the tail of output discrimination distribution for top-tagging and b -tagging, its yield is better extrapolated from the three sideband regions shown in Table 3.1. The M_{tb} distributions of QCD multi-jet backgrounds in these regions are estimated as data minus the $t\bar{t}$ background. If the top-tagging and b -tagging decisions are independent, the data-driven background estimate in region A is given by region B times C over D for every M_{tb} bin (Chapter 8).

The accuracy of the data-driven method relies heavily on controlling the correlation between the two candidate jet’s tagging. In defining the non-top-tagged region, an alternative

²The absolute value of a momentum vector perpendicular to the LHC beamline

Top-candidate jet tagged	C	A
Top-candidate jet not-tagged	D	B
	B-candidate jet not-tagged	B-candidate jet tagged

Table 3.1: Four regions employed by a data-driven background estimation are shown as four cells on a two-dimensional plane. The horizontal and vertical axes are represented by the tagging decision on the b-candidate jet and the top-candidate jet, respectively.

top-candidate jet definition might be sensitive to different quark flavors and gluons, thus biasing the b-candidate jet towards specific parton flavors. An improved approach is to consider any high p_T large-radius jet a top-candidate jet, as long as we can find a back-to-back small-radius jet with high p_T as its b-candidate jet. Furthermore, the sideband regions can be subdivided into blocks of four cells, among which the correlation can serve as an estimate of systematic uncertainties.

There are other systematic uncertainties in measuring jet momenta, correcting top-tagging and b -tagging rates for simulated events, and calculating the hard-scattering amplitudes that shape the signal and background-predicted M_{tb} distribution (Chapter 9). There are more than 100 nuisance parameters (θ) representing these uncertainties. Their joint effects are modeled by a profile-likelihood function $\mathcal{L}(\mu, \theta)$ of the SR's M_{tb} bins for a signal strength μ . As this analysis is a generic search, μ is not limited to the adopted W' model's value but is considered a parameter-of-interest. The nuisance parameters are fitted to data by maximizing \mathcal{L} , while μ is either fixed or floating ($\hat{\mu}$) during the fit. A small ratio of $\mathcal{L}(\mu)$ to $\mathcal{L}(\hat{\mu})$ signifies deviation between data and the μ -signal plus background model. Those μ -values with sufficient inconsistency are excluded. The lowest excluded signal size for each W' mass is plotted as an upper limit on the W' production cross-section times the tb -decay branching fraction as a function of W' mass. A hint for an excess compatible with the signal model will show a significant bump on the limit.

Chapter 4

Experimental Apparatus

This chapter describes the principle function and utilization of the experimental apparatus. The LHC accelerator parameters, the proton injection chain, and the LHC data analyzed in this thesis are described in the first section. The following ATLAS section introduces the instrumental structures and main subsystems for identifying and measuring physical objects, with more details provided for those components crucial to the QCD jet reconstruction.

4.1 Large Hadron Collider

Located 45 m to 170 m underground on the border between France and Switzerland, the Large Hadron Collider (LHC) [37] accelerates two counter-circulating proton beams to 6.5 TeV energy per proton. Along its approximately 27 km circumference, opposing protons collide at four interaction points each of which houses a detector to observe particles coming out of the collision: ALICE [38], ATLAS [39], LHCb [40], and CMS [41]. The total 13 TeV center-of-mass energy provides sufficiently energetic partons to annihilate and produce the targeted TeV-mass W' boson (Figure 1.1).

The proton beams are exerted by a standing electromagnetic wave when they pass through a section of cavities at the LHC ring. Protons are separated into bunches such that the longitudinal electric field is in the phase of maximum acceleration when protons traverse through the field-generating device. Each beam has 2808 proton bunches of several

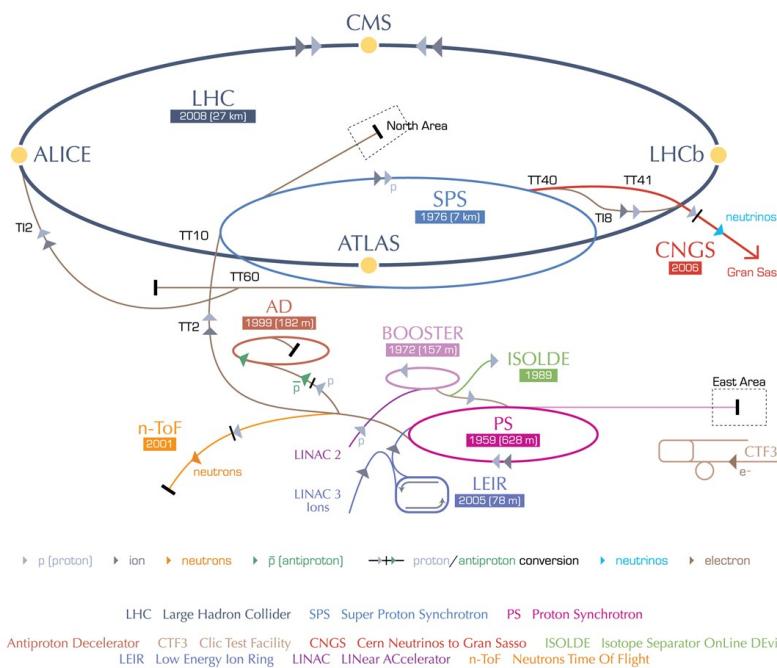
centimeters long separated by few meters, leading to a 24.95 ns duration between collisions [20, 37]. The circular motion of protons along the beam pipes is supported by 1232 dipole superconducting magnets having a peak 8.33 TeV Tesla magnetic field. Liquid helium cools the temperature of the magnet to 1.9 K to maintain its superconducting state. Quadrupole and sextupole magnets are also employed to correct proton trajectories and focus the protons at the interaction points to enhance luminosities. The luminosity measures the area density of protons per unit time.

The luminosity of the LHC is critical to the W' search. The rate of interaction (dN/dt) is the product of instantaneous luminosity ($\mathcal{L}(t)$) times the interaction cross-sections (σ) [20]. Since the predicted production cross-section of a 1 TeV W' is of the order of 10^{-35} cm^2 and decreases further with masses (Figure 5.3a and Figure 5.3b and using the conversion of 1 pico-barn = 10^{-36} cm^{-2}), the $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ or higher instantaneous luminosity of the LHC could yield significant numbers of potential signal events after years of LHC running.

The acceleration of protons starts at the accelerator chain prior to proton injections into the LHC. The initial stage is a linear accelerator, LINAC 2, which brings proton energy to 50 MeV [43]. Then there are three more synchrotrons in succession – Proton Synchrotron (PS) Booster, PS, and Super Proton Synchrotron (SPS) – where the final proton energy is 1.4 GeV, 25 GeV, and 450 GeV each. The ultimate acceleration at the LHC pushes the beam energy from 450 GeV to 6.5 TeV. The accelerator chain is also responsible for allocating proton bunches [37]. The six proton bunches injected from the PS Booster are split into 72 in the PS and are subsequently ejected to the SPS. The SPS is filled by three to four batches of 72-bunch protons before its ejection to the LHC. Each beam transfer requires a change in the field configuration of the kicker, requiring gaps between proton bunches.

This thesis utilizes LHC data recorded between 2015 to 2018 by the ATLAS detector

CERN's accelerator complex



European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

© CERN 2008

Figure 4.1: The CERN accelerator complex. The injection chain for proton collisions at the LHC is composed of LINAC 2, PS Booster (BOOSTER on the diagram), PS, and SPS. There are four detectors/experiments, ALICE, ATLAS, CMS, and LHCb marked as yellow dots along the LHC ring. [42]

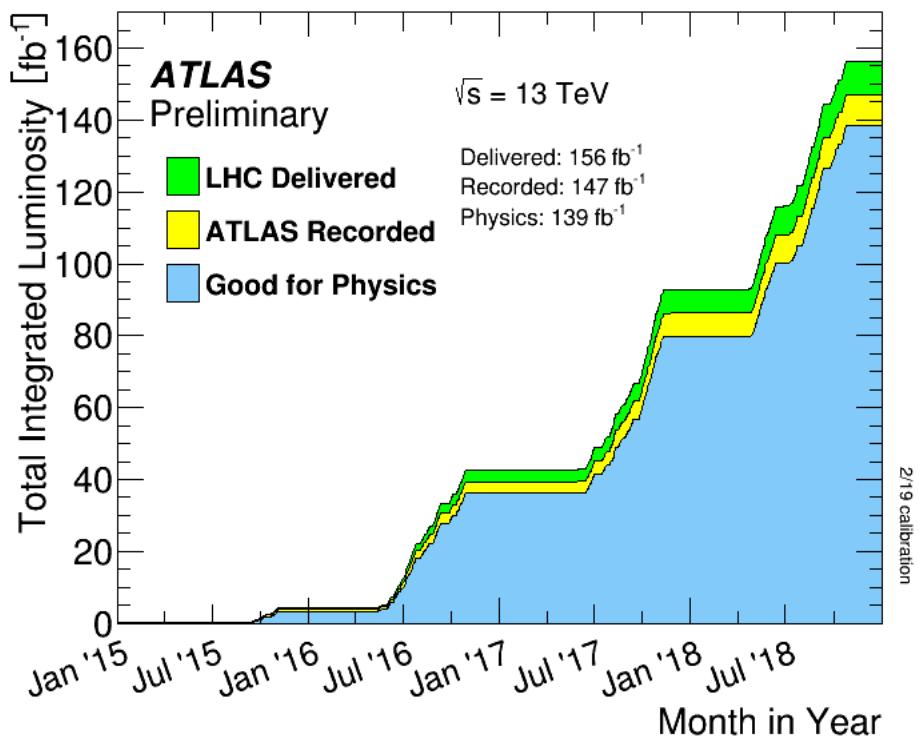


Figure 4.2: Increase of integrated luminosities over time of the data delivered by the LHC, recorded by the ATLAS detector, and marked as good for physics analysis. [44]

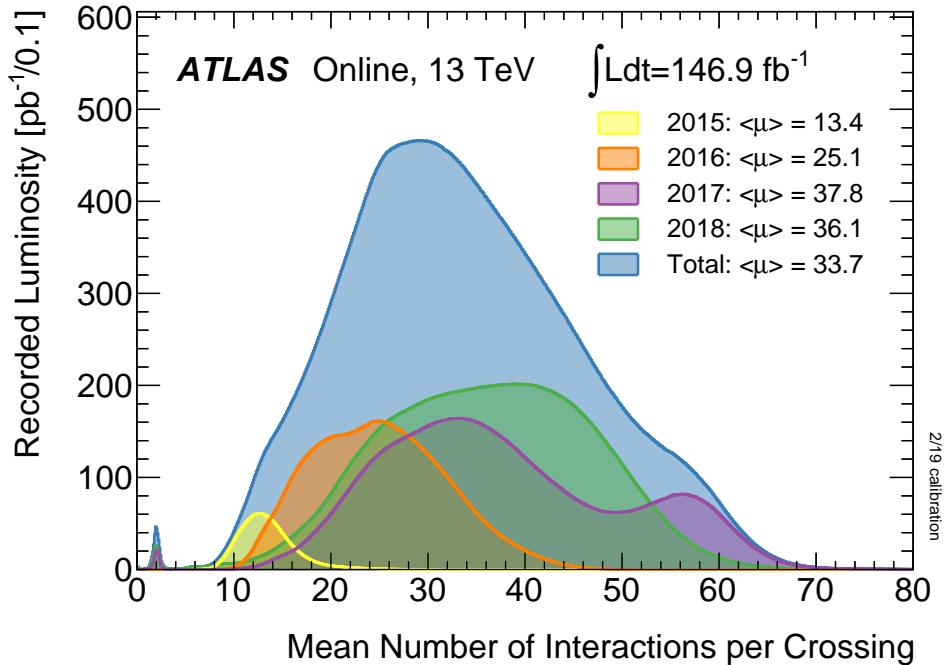


Figure 4.3: The number of interactions per bunch crossing averaged over a luminosity block [45] is shown as integrated-luminosity-weighted distribution on the graph for each year of data taking. The legends tell the overall average for each year. [44]

– the Run 2 period. During this period, the LHC delivered data equivalent to 156 fb^{-1} integrated luminosity ($\int \mathcal{L}(t)dt$), among which 147 fb^{-1} are recorded by the detector. The entire ATLAS system is under conditions suitable for physics analysis in a smaller subset of 139 fb^{-1} (Figure 4.2). In this dataset, the beam energy is always at 6.5 TeV. Meanwhile, the LHC successfully attained higher instantaneous luminosity at the interaction points towards the end of the run [45]. Although higher luminosity helps the search for rare physics, it also elevates the number of proton-proton interactions in the same bunch-crossing (Figure 4.3). Most of them are soft (low energy exchange) QCD processes called pile-up interactions as they overlap with the hard-scattering physics¹. Pile-up yields numerous tracks and abundant energy deposits to the detector, which have to be mitigated in reconstructing final-state

¹On the detector side, these pile-up interactions are called in-time pile-ups. There are also out-of-time pile-ups due to overlapping electronic signals due to the long detector time resolution compared to the collision time window.

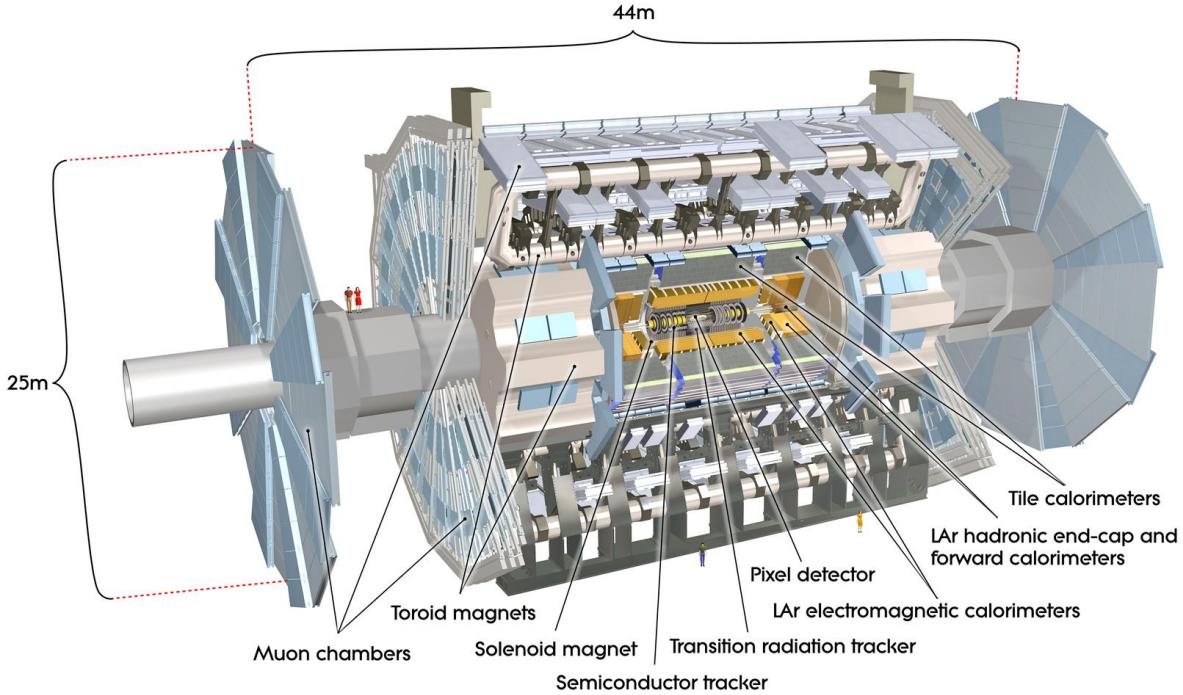


Figure 4.4: The figure points out the main subdetectors of the ATLAS detector.

objects from hard-scattering physics.

4.2 ATLAS Detector

A Toroidal LHC ApparatuS [46], or ATLAS, is a multi-purpose detector designed to measure physical objects such as charged leptons, jets, and missing transverse energy originating in proton or heavy-ion collisions at the LHC. The cylinder-shaped detector is about 25 meters in diameter, 44 meters long, centering at the interaction point, and coaxial with the LHC beam pipe. Except for the two proton beam's traveling directions, the ATLAS detector is covered by layers of subdetectors. They are Muon Spectrometer (MS), Calorimeter (CAL), and Inner Detector (ID) from the outer to the inner layer. The CAL is further subdivided into the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). Four

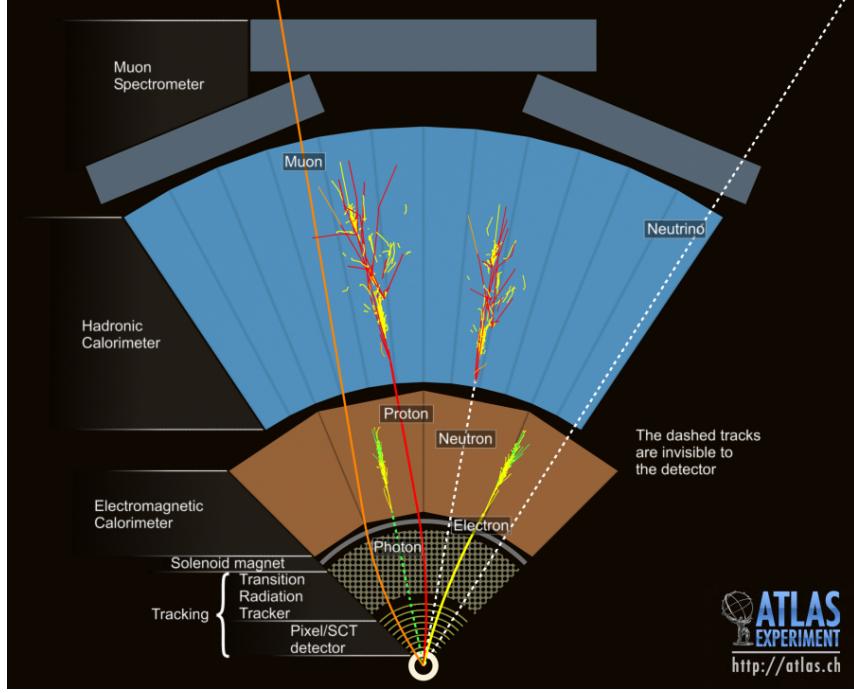


Figure 4.5: A simplified representation of a segment of a cross-sections of the ATLAS detector. The toroidal magnets around the muon spectrometer are omitted. The tracking system corresponds to the inner detector (ID).

superconducting magnets are instrumented: A solenoid immerses the ID in a longitudinal magnetic field, and two end-cap plus one barrel toroids provide a magnetic field for the MS.

Figure 4.4 is a 3D sectional view of the detector.

The ATLAS subdetectors and magnets corroborate to identify particle properties via their interactions with matters, as shown in Figure 4.5. The ID tracks charged particles (muons, electrons, and protons in the figure) by their ionizing hits. The particle momentum is given by the track's curvature under the solenoid's magnetic field. Afterward, the CAL absorbs and measure energy from high energy particle's cascades. The electromagnetic calorimeter contains the electromagnetic showers from electrons and photons. In contrast, hadronic showers are broader and prolonged [47], penetrating deep into the hadronic calorimeters. Muons, with their long lifetime and minimal interactions [48], escape the calorimeter and are detected by the muon spectrometer. The MS drift chambers measure the muon momentum

by tracking their paths under the toroidal magnetic field. Finally, neutrino is intangible to all subdetectors, but a significant imbalance of an event’s transverse energy signals its existence.

The subdetector and measured physical objects are described by a coordinate system that takes the nominal interaction point as its origin. Its z-axis points along the LHC ring counterclockwise as viewed from the top, then the x-axis directs to the center of the LHC ring, and hence its y-axis points upward. The azimuthal angle ϕ is the angle around the z-axis as measured from the x-axis, while the polar angle θ is the angle from the z-axis. Another widely used angular variable is the pseudorapidity $\eta = -\ln \tan(\theta/2)$ (which is equal to the rapidity $y = 1/2 \ln [\tan((E + p_z)/(E - p_z))]$ for massless particles) which is used to define the angular distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Finally, the norm of the momentum p on the x-y plane is termed as the transverse momentum, p_T . Similarly, the energy E times $\cos\theta$ is the transverse energy, E_T .

4.2.1 Inner detectors

The ID reconstructs charged particle tracks using pattern recognition algorithms. These tracks are extrapolated to their closest approaches to the beam pipe, forming primary vertices and secondary vertices. A primary vertex is where a hard-scattering collision or a pile-up interaction occur. A final-state particle emerging from these interactions can travel a distance resolvable by the ID before it decays. Tracks from a long-lived particle decay intersect at a secondary vertex. As mentioned in Chapter 3, the identification of the bottom-quark jet utilizes tracks from the secondary vertex. Since the ID tracks have better angular resolutions than the calorimeter cells, they are also employed to remove pile-up energy and tracks in small-radius jets and to calibrate the large-radius jet substructure and mass (Chapter 6).

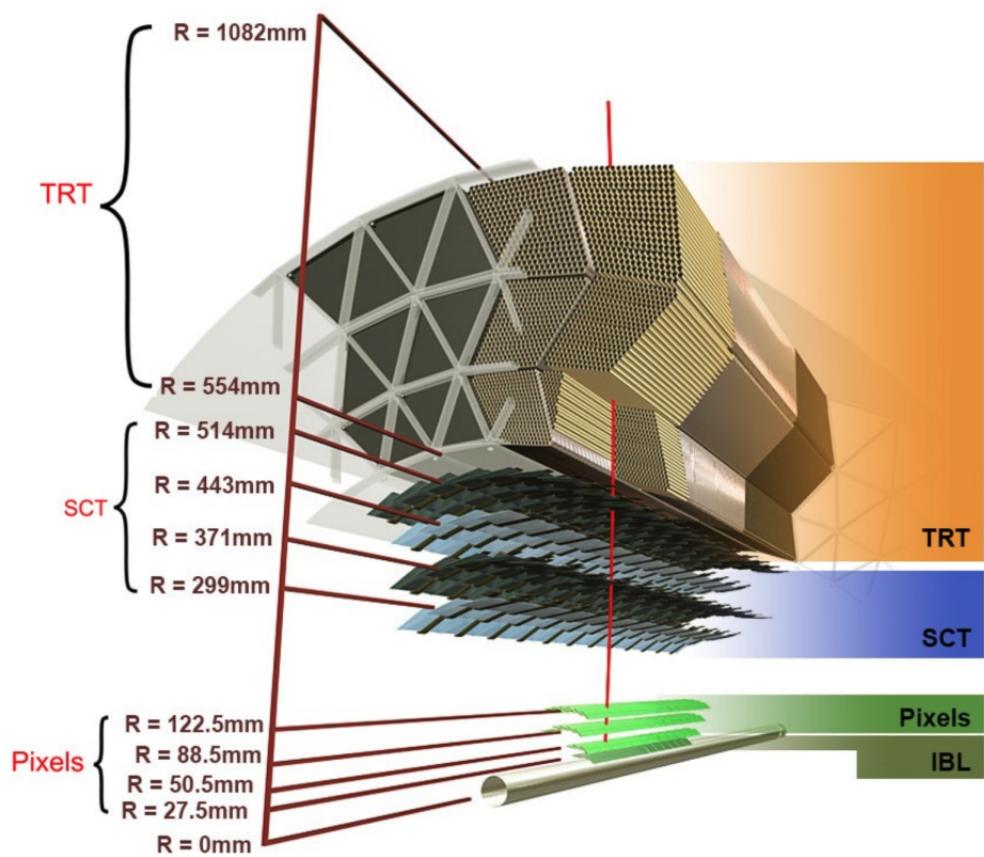


Figure 4.6: A 3D view of the barrel region of the inner detectors. [49]

The ID comprises three subdetectors – Pixel detector, Semiconductor Tracker (SCT), and Transition Radiation Tracker (TRT) – as shown in Figure 4.6. Both Pixel and SCT utilize silicon sensors. The former is closer to the interaction point, and its silicon modules have smaller sizes to achieve the best track reconstruction resolution. The Insertable B-Layer (IBL) is an additional Pixel layer installed in 2014 for coping with higher track multiplicities and risks of Pixel module failures brought by increasing instantaneous luminosities [50]. The TRT is composed of drift tubes. Despite having a lower resolution than the Pixel and SCT, it extends the ID’s depths to recognize long-lived particles. It also contain transition radiator detectors to discriminate electrons from hadrons.

4.2.2 Calorimeters

The calorimeters measure the energy profile of physical objects: electrons, photons, jets. They cause electromagnetic and hadronic showers in the CAL. Active materials in the ATLAS calorimeter sample a fraction of the energy deposits while the rest is absorbed by dense matter. Figure 4.7 presents the geometry of individual calorimeter subsystems. The absorber and sensor layers are interspersed densely, such as in the ECAL (Figure 4.8) and (Figure 4.9), leaving no cracks in the ϕ coverage. The ECAL is closer to the ID, using a high atomic number absorber – lead – to contain electromagnetic radiations due to energetic electrons and photons. In contrast, jets are made of hadrons that lose energy by both electromagnetic showers and nuclear interactions extending deep into the HCAL. Since the ratio of electromagnetic versus hadronic energy deposit of jets has randomness and some portion of hadronic energy is not measurable by the active material [20], hadronic clusters formed by combining ECAL and HCAL measurements have a local weighting scheme to improve the jet energy resolution (Chapter 6).

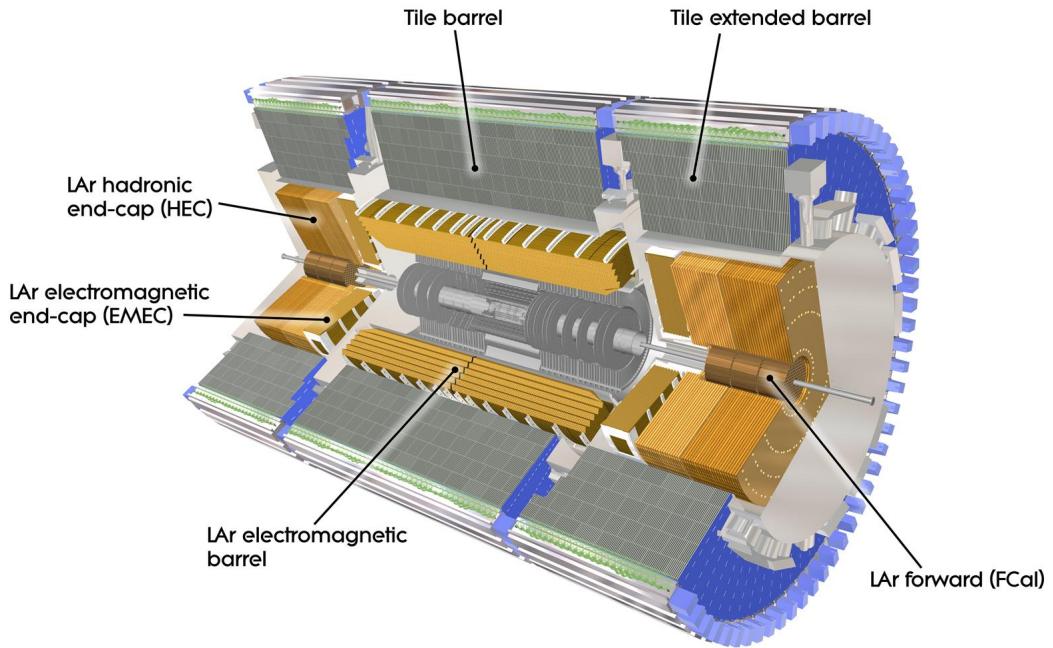


Figure 4.7: The ATLAS calorimeter encloses the ID colored in grey and a hard-to-seen solenoid magnet around it. The ECAL consists of a liquid-Argon (LAr) barrel part ($|\eta| < 1.475$) and two LAr end-caps ($1.375 < |\eta| < 3.2$). The tile calorimeter (barrel: $|\eta| < 1.0$ and extended barrel: $0.8 < |\eta| < 1.7$) and two LAr hadronic end-caps ($1.5 < |\eta| < 3.2$, painted in two sheds of copper) comprise the HCAL. At $|\eta|$ between 3.1 and 4.9 on each side, there are three components of the LAr forward calorimeter (FCAL). [46]

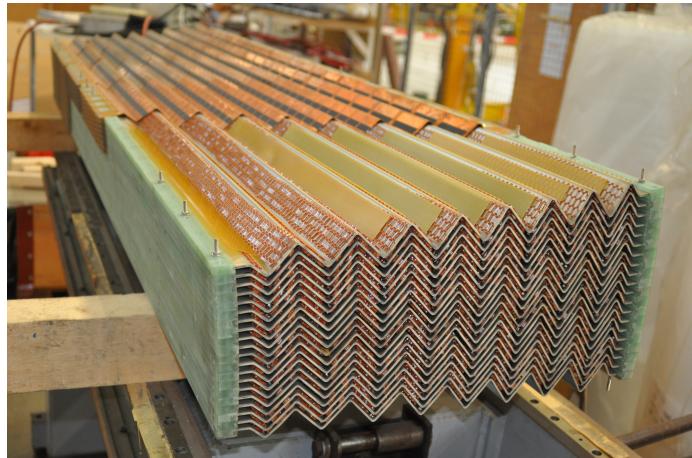


Figure 4.8: Segments of accordion-shaped calorimeters. The silver-colored lead absorbers are sandwiched by copper-clad Kapton electrodes. The electrode receives ionizing signals from the liquid Argon filled in the gap between lead absorbers. [51]

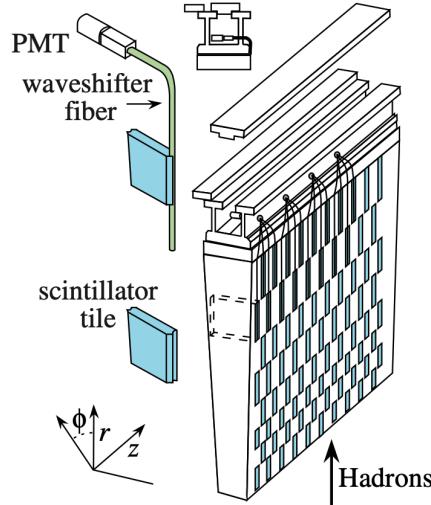


Figure 4.9: A wedge of tile hadronic calorimeters consisting of scintillator tiles interleaving with steel absorbers. The coordinate system shown on the bottom left follows the convention introduced at the beginning of this section. The top left and center left diagrams demonstrate how a waveshifter fiber transmitting light signals from scintillator tiles to a photomultiplier tube (PMTs). The top figure shows the PMT's position in the top drawer. The plot is taken from [20] which adapts the original one in [52].

The radii R of the large-radius and small-radius jets used in this analysis are 1.0 and 0.4, respectively (Chapter 6). Their jet axes are also limited to have $|\eta| < 2.0$ and 2.5 each; therefore, the input clusters are built from cells in the barrel and end-cap calorimeters, excluding the FCAL. The energy resolution, the Gaussian width of the calorimeter response to hadronic energy deposits, is shown in Figure 4.10 in a combined test of the ECAL and Tile HCAL as determined in pion test beams. The curve is dominated by the stochastic term that scales with the inverse square root of the test beam energy. An equivalent quantity, the jet transverse momentum resolution, is calibrated *in-situ* using the p_T of back-to-back jets (More on the jet properties and calibrations see Chapter 6). The p_T resolution of large-radius jets is presented in Figure 4.11 for both data and simulated events in the barrel region. Comparing the value at 300 GeV between Figure 4.10 and Figure 4.11 shows that the energy resolution for a large-radius jet is about 40% more than that for pions in the test

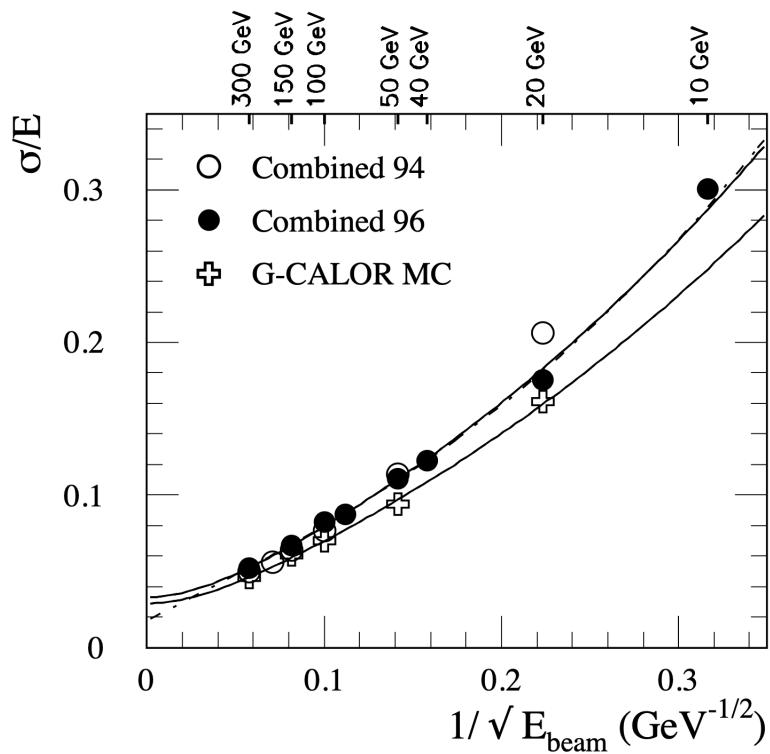


Figure 4.10: The pion energy resolution of the hadronic and electromagnetic calorimeters combined is shown as a function of the square root of pion energy inverse. The black (white) dots are measured with test beams in year 1996 (1994), while the white cross is derived from the simulation of hadronic showers. The two fit functions differ in their description of detector noises. [53]

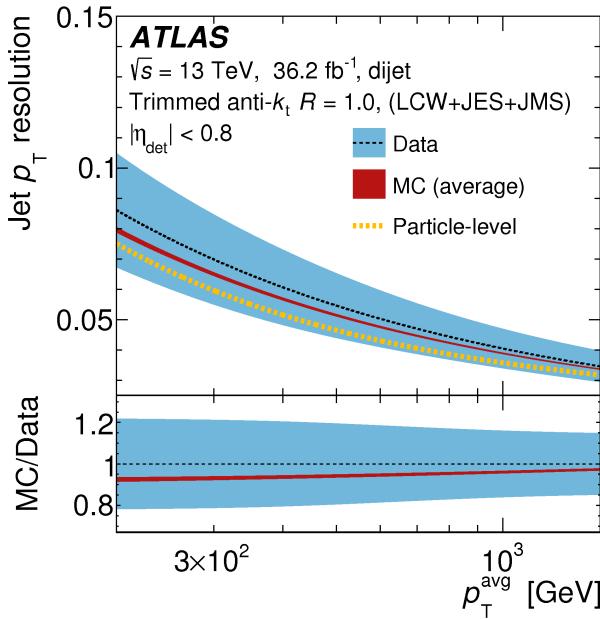


Figure 4.11: The p_T resolution of the large-radius jet as a function of the average of the di-jet transverse momentum from the *in-situ* calibration [54]. The symbol $|\eta_{\text{det}}|$ means the absolute pseudorapidity of the jet axis with the detector's center taken as the coordinate's origin. The data (black dashed line) has a light band that sums the statistical and systematic uncertainties in quadrature. The red band is the envelope of *in-situ* calibration applied to three different QCD di-jet event generators. The yellow dashed line corresponds to the reconstructed jet p_T resolution relative to the truth-leveled jet (Chapter 6) computed from events simulated by the PYTHIA generator.

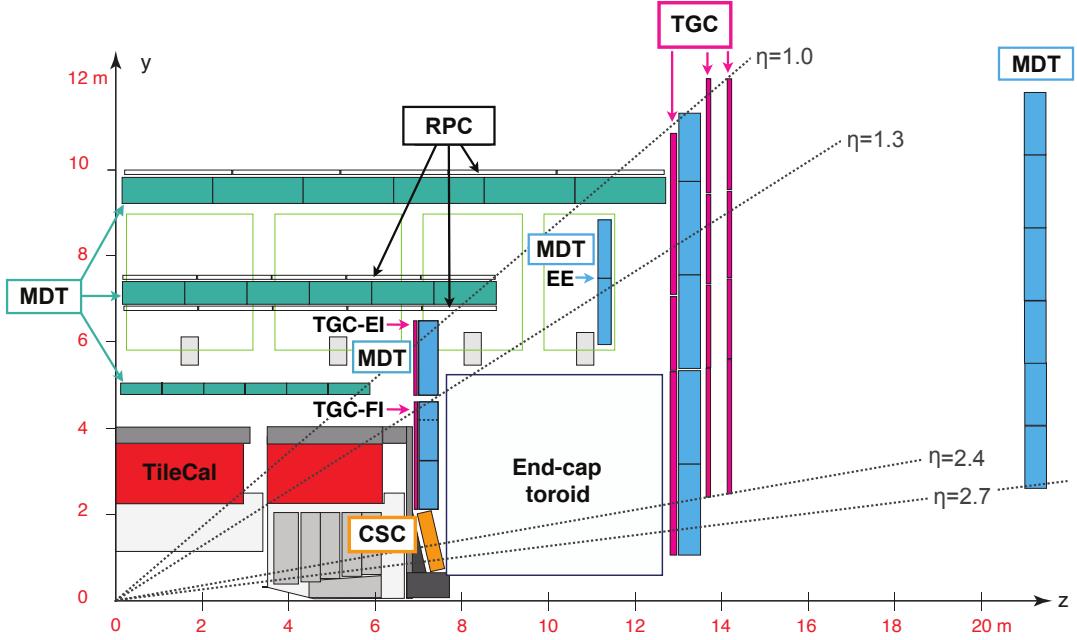


Figure 4.12: The muons detected by the MS are bent by the toroidal magnets in the barrel (green) and end-cap (black). With requirements of three layers of measurements, triggering (RPC and TGC) has a range of $|\eta| < 2.4$, limited by the two outer TGC wheels. In comparison, precision-tracking (MDT and CSC) can go further to 2.7. [56]

beam data. In contrast, the small-radius jet p_T resolution is closer to the latter [55]. The effect of jet energy resolution from the top-quark jet and the bottom-quark jet broadens the reconstructed M_{tb} far beyond the W' width at the parton level (Chapter 5).

4.2.3 Muon spectrometers

The MS identifies, triggers, and measures the momentum of muons with the three surrounding toroidal magnets. Each of them has eight octagonally arranged superconducting coils. Their air-core structure minimizes multiple scattering, which changes a particle's moving direction randomly. The magnetic field pointing in the ϕ -direction bends the muon trajectories on the $r - z$ plane (in a cylindrical coordinate system), changing the η of a muon track. These tracks are reconstructed by ionization hits in three to four layers of gas chambers and

then combined with ID tracks to reduce reconstruction uncertainties.

Figure 4.12 shows the relative position of four subsystems of the MS. There are two precision-tracking chambers: Monitored Drift-tube Tubes (MDT) and Cathode Strip Chambers (CSC). There are three layers of MDT chambers in the barrel region (green) and four layers in the end-cap (blue), with each layer making three to eight measurements in η [46]. In the innermost layer at high $|\eta|$ region, the MDT is replaced with the CSC, having a better resolution to cope with more abundant fake-muons in the forward area. The CSC consists of four chambers equipped with η and ϕ measuring cathode strips. The two trigger chambers are Resistive Plate Chambers (RPC) attaching to the barrel MDT and Thin Gap Chambers (TGC) near the end-cap MDT. They have a shorter signal-processing time than the MDT and CSC, thus suitable for fast triggering. They also provide ϕ coordinates for track reconstruction, complementary to the MDT.

4.2.4 Trigger and data acquisition

Not all events are saved for data analysis at the ATLAS detector since the collision rate of proton bunches of 40.08 MHz far exceeds the ability of data processing and capacity of data storage. Furthermore, compared to the inelastic scattering cross-section of 80 mb at $\sqrt{s} = 13$ TeV [46], events with high energy physical objects are much less likely to appear. For example, the cross-section $\sigma(pp \rightarrow W'_R \rightarrow tb)$ for a 2 TeV W'_R mass is less than 1 pb, 10^{-9} times 1 mb. Since the start of Run-2, the ATLAS detector has been equipped with a two-tier trigger and data acquisition system (TDAQ), which selects events fulfilling signatures of interests efficiently before writing the data to disks.

Figure 4.13 illustrates the trigger system's main modules (on the left), the data readout system (on the right), and the data links between them. The two-tier trigger system consists

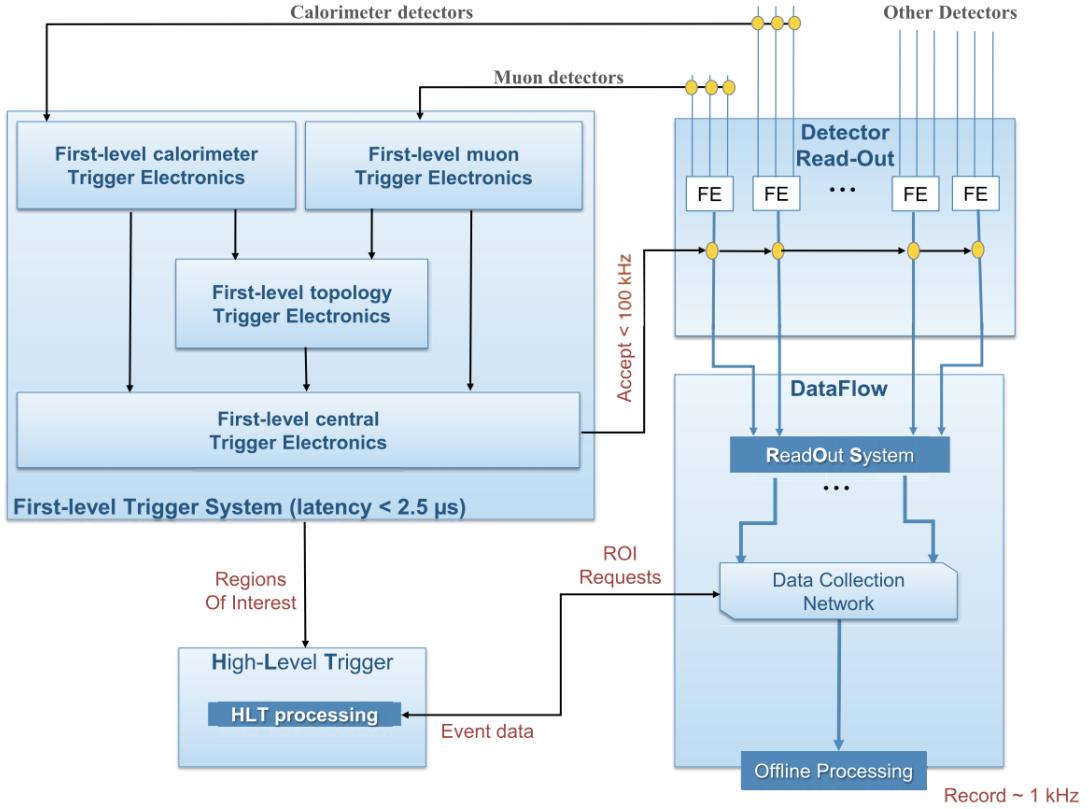


Figure 4.13: A schematic diagram of the Run-2 ATLAS trigger and data acquisition system. The top-right block shows the detector front-ends (FEs), some of which at the CAL and MS are connected to the corresponding calorimeter trigger and muon trigger at the first-level (L1) trigger block on the top-left block. The entire time consumed by data transmission and logic processing at the L1 is less than $2.5 \mu\text{s}$ before the central trigger flags the L1 Accept command back to the FE at a rate below 100 kHz. The accepted event's data from all detectors are then distributed through the ReadOut System (ROS). The L1 trigger also sends the ROI information to the High-level Trigger (HLT), which accesses the event data through the ROS. As shown on the bottom right block, data passing the HLT's threshold at a rate of about 1 kHz is transferred to permanent storage for offline analysis. (Figure adapted from [57])

of the First-level (L1) and the High-level (HLT). By reading only partial information from the CAL and MS, the L1 trigger efficiently rejects events without significant energy deposits. The remaining events are examined by the downstream HLT to execute more time-consuming algorithms and test more stringent trigger criteria than the associated L1 trigger.

The L1 trigger is a hardware-based system with customized electronic chips. It contains units for the Calorimeter (L1Calo), the MS (L1Muon), and topological triggers taking input from both systems (L1Topo). Each L1 trigger system has individual trigger items with different requirements. The detector region identified by a trigger item is called the Region-of-Interests (RoIs), locating physical objects for the HLT. The final L1-Accept decision made at the L1 central trigger processor (L1CTP) combines the output from L1Calo, L1Muon, and L1Topo. The time consumption at the L1 stage is about $2\mu\text{s}$ with 0.5μ for back-up [57]. Since this duration is much longer than the bunch-crossing interval, a pipeline memory installed at each detector's front-end holds output until the L1 decision is made.

One example of L1 trigger is the jet trigger at $|\eta| < 3.2$. The trigger uses a sliding window algorithm to find a localized calorimeter energy exceeding a defined E_{T} level [58]. The ECAL and HCAL energy is summed to form jet elements with a size of 0.2×0.2 in $\Delta\eta \times \Delta\phi$ ($\Delta\eta$ increases to 0.3 at the end-cap region). The jet trigger defines the RoI as 2×2 jet elements with maximal E_{T} among its neighbors. For the trigger decision, a set of energy thresholds is compared to the E_{T} sum in the RoI and the surrounding 3×3 and 4×4 trigger towers.

The HLT's trigger items are software algorithms executed in computing farms. The HLT examines physical objects reconstructed with algorithms simplified from the offline version to meet time constraints. With the full granularity in the detector input and more computational time, the HLT triggers have a sharper turn-on curve of trigger acceptance rate (efficiency) than the L1 [59]. Some HLT trigger items are seeded by the RoI to save

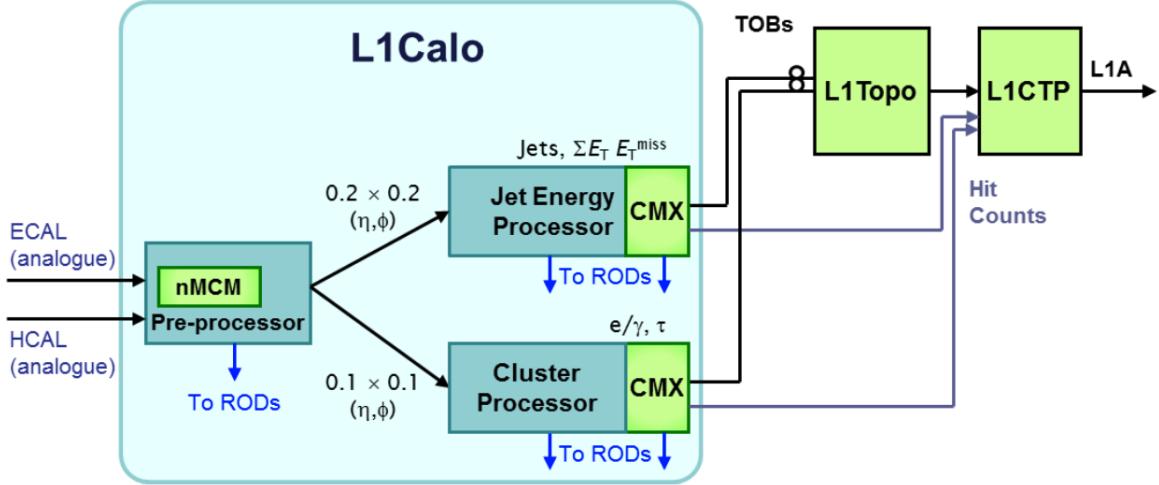


Figure 4.14: The architecture of the L1Calo during Run-2 (2015-2018). All modules have output to Read-Out Drivers (RODs), allowing diagnostics, calibrations, and performance studies. In particular, RoIs from the CP and the JEP are routed through the CMX to the RODs for later retrieval by the HLT. RODs pack, buffer and send signals to the ROS. [62]

time and resources while taking a higher risk of object misidentification. Therefore, a full scan over the entire detector is performed for any HLT algorithm whenever possible.

This analysis adopts an HLT trigger requiring at least one large-radius jet with E_T thresholds ranging from 380 to 460 GeV (Chapter 6). The anti- k_t algorithm is used to reconstruct jets from topological clusters of hadronic energy deposits. The sharpness of turn-on curves is reduced by less sophisticated calibration of the jet energy scale than the offline analysis [60]. Thus more stages of calibrations were ported to the HLT jet reconstruction as Run-2 progressed [61].

4.2.5 Level 1 calorimeter trigger

The Level-1 Calorimeter Trigger (L1Calo) [63], one of the components of the L1 trigger system, is critical to the fast identification of energy deposits in the calorimeters from electrons, photons, and jets. Its input is the trigger tower that sums the HCAL and ECAL energy

into cells of 0.1×0.1 to 0.4×0.4 in $\Delta\eta \times \Delta\phi$, from the central to forward regions. Analog signals of these cells are distributed to L1Calo modules housing the following components, Pre-processor, Cluster Processor (CP), Jet Energy Processor (JEP), and Cluster Merger Module Extended (CMX) (Figure 4.14).

The first component, the Pre-processor, digitizes analog signals at a rate of 80.16 MHz, synchronizes them to compensate for the varying transmission time from each calorimeter cell, and reduce electronic and pile-up noises in the electric signal's baseline (pedestal). As the calorimeter's electric pulses span many bunch-crossing intervals, the Pre-processor employs a filtering algorithm to extract each trigger tower's E_T at the pulse's peak.

The digital signals from the Pre-processor are transmitted to the CP and JEP, respectively. The CP searches for candidate electrons, photons, and τ -leptons. The JEP looks for jets and calculates the sum ($\sum E_T$) and imbalances (E_T^{miss}) of the transverse energy. The L1 trigger algorithms, executed at the CP and JEP modules, define physical objects with various windows of trigger towers, such as a 2×2 core with a surrounding isolation region used in the CP. Note that the base element for jet-finding windows is already a two-by-two combination of trigger towers, as shown in Figure 4.14. The CP and JEP modules write out Trigger OBjects (TOBs) that record the proximity (RoI), energy, and the type of physical objects inspected by them.

The output of CP and JEP modules are merged by the Common Merger module eXtended (CMX) to count the TOBs exceeding a set of trigger thresholds and calculate global E_T variables per event. The CMX then transmits the TOBs to the L1Topo system to compute composite variables like the ΔR between physical objects. A table of local trigger counts and global variables (with the event-wide variable's trigger counts) are sent to the L1CTP for the Level-1 Accept decision (L1A).

Chapter 5

Predicting the W' Signal

This chapter introduces a model that parameterizes a massive vector boson, W' , with which the production cross-section times $W' \rightarrow tb$ branching ratio is evaluated at the next-to-leading order (NLO) in the QCD coupling. The predicted cross-section times branching ratio's theoretical uncertainties are given. Finally, the leading order (LO) matrix element with hadronic top-quark decay (Figure 5.1) is interfaced with parton showers, hadronization, and particle interactions with the detector. Note that the notation $W' \rightarrow tb$ represents both $W'^+ \rightarrow t\bar{b}$ and $W'^- \rightarrow \bar{t}b$, for brevity.

5.1 NLO Cross-Section Evaluation with ZTOP

The phenomenological model of W' , that characterizes vector resonances beyond the SM, is a generalization of the SM W boson. The interaction Lagrangian between SM fermions f and the W' boson is written as

$$\mathcal{L} = \frac{V'_{ij}}{\sqrt{2}} \bar{f}_i \gamma_\mu \left[g'^R \left(\frac{1 + \gamma^5}{2} \right) + g'^L \left(\frac{1 - \gamma^5}{2} \right) \right] W'^\mu f_j + h.c. \quad (5.1)$$

V'_{ij} is a matrix in the flavor space that parametrizes the coupling of the W' boson with an up-type fermion of generation i and a down-type fermion of generation j . The matrices γ_μ and γ^5 are Pauli matrices for fermion's spins. The $(1 + (-)\gamma^5)$ projects the fermion's left

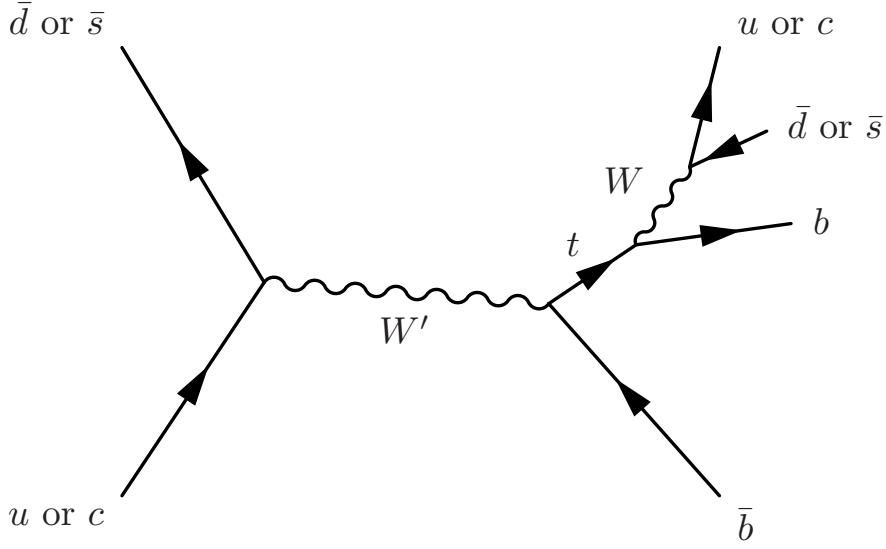
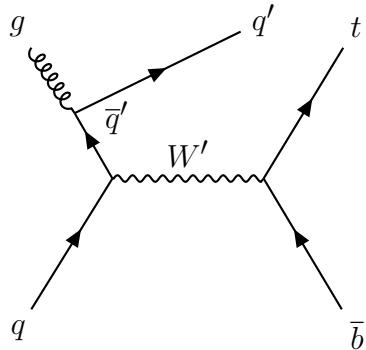


Figure 5.1: The leading order Feynman diagram for the signal’s hard-scattering matrix element. This diagram’s final state is referred as the parton level. Time points from left to right. The spatial dimension is vertical. For aesthetic purposes, the momentum flow is only accurate in the time direction. The charge-conjugated diagram is not shown but included in the event generation. The diagram is produced by **feynMF** [21]

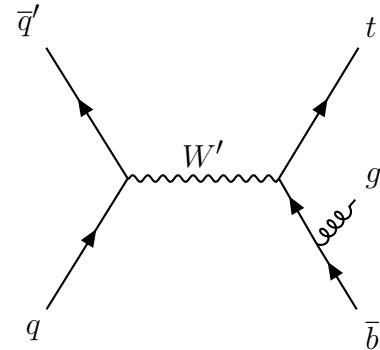
(right) handed component in chirality. The coupling constant for the left-handed chirality is denoted as g'^L and that for the right-handed chirality as g'^R . The symbol *h.c.* stands for the previous term’s Hermitian conjugate.

Although these parameters can be tuned to a range preferred by a specific model, this analysis considers only two scenarios for simplicity – the left-handed W' boson, W'_L and the right-handed W' boson, W'_R . The W'_L only interacts with left-handed fermions by setting $g'^R = 0$ and $g'^L = g_{SM} \approx 0.63$ [20]. The value 0.63 is the SM W boson’s coupling constant. Reversing the coupling constant – $g'^L = 0$ and $g'^R = g_{SM}$ – leads to the W'_R . For both scenarios, the matrix V'_{ij} for quarks also replicates the SM W boson’s CKM matrix. Hence, W' couples sufficiently to the initial state light quarks inside protons and the final state t and b quarks.

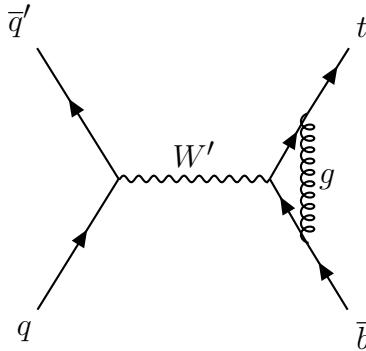
The W' production cross-section depends on the energy scale of the initial state quarks, determined by the parton distribution function (PDF, introduced in Chapter 2). The search



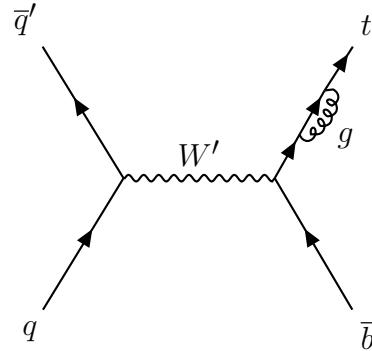
(a) A gluon splits into a quark-antiquark pair



(b) A gluon radiates off the b -quark



(c) One-loop vertex correction



(d) One-loop mass correction to the t -quark

Figure 5.2: Feynman diagrams of NLO terms in calculating the cross-section $\sigma(q\bar{q}' \rightarrow W' \rightarrow t\bar{b})$, where q and \bar{q}' represents an up-type quark and a down-type antiquark emerging from protons, respectively. The vertical (horizontal) axis corresponds to the space (time) axis. The charge-conjugated processes are included in the cross-section calculation. Feynman diagrams are made with TikZ-Feynman [64].

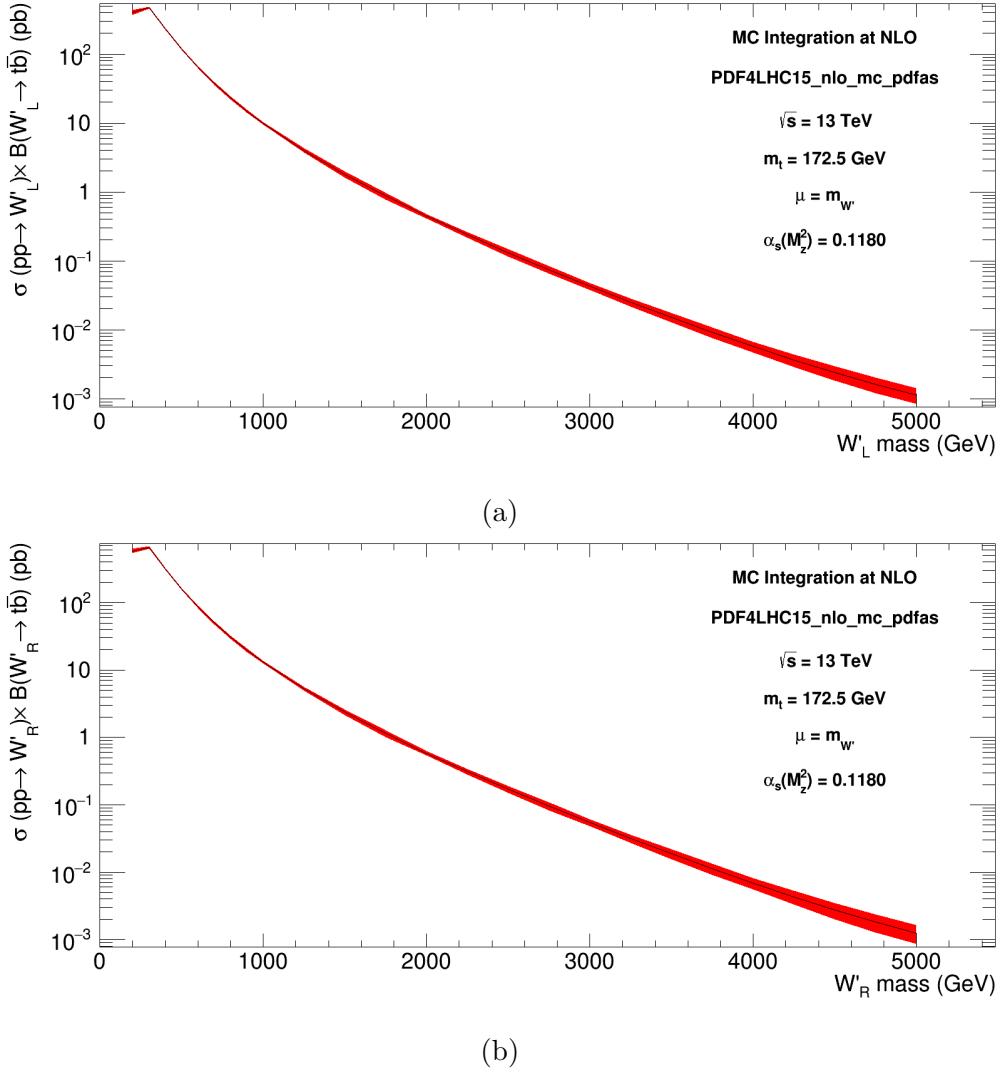


Figure 5.3: The cross-section $\sigma(pp \rightarrow W')$ times branching ratio $B(W' \rightarrow tb)$ with W' 's rest masses in 0.2~5.0 TeV is evaluated with ZTOP [65–67] at NLO of QCD. Figure (a) is for a left-handed chiral W'_L and (b) is for a right-handed chiral W'_R . The central black curve is the nominal value, including almost invisible statistical errors. The red band stands for the theoretical uncertainties. Interfaced by LHAPDF [68], the PDF set is PDF4LHC15_nlo_mc_pdfas documented by [69]. This set combines CT14 [70], MMHT14 [71], and NNPDF3.0 [72] with technologies found in [73, 74].

range of W' mass is limited by trigger threshold to be above 1 TeV, nearly 8% of the proton-proton collisions' center-of-mass energy. Thus, one of the initial state parton is likely to be an up quark or a down quark that carries most of a proton's longitudinal momentum. By contrast, the second parton has to be an antiquark, holding less momentum due to proton structures. At higher orders in QCD, the gluon splitting contributes significantly to the cross-section as partons are dominantly gluons at LHC's energy scale [75] (Figure 5.2a).

A convolution of parton distribution functions and the matrix element, both up to the next-to-leading order (NLO) in QCD (the latter shown in Figure 5.2), yields the parton-level cross-section times branching ratio: $\sigma(pp \rightarrow W' \rightarrow tb)$. Varying the W' mass ($m_{W'}$) from 200 GeV to 5 TeV, the result is plotted for W'_L in Figure 5.3a and W'_R in Figure 5.3b. The W'_L assumes about 1/4 less tb -decay branching ratio than the W'_R as only the former can decay into leptons. The yet-to-discovered right-handed neutrino is hypothesized to be more massive than W'_R , thus forbidding the decay. For both W'_L and W'_R , the kink at 250 GeV is due to the shrinking decay phase spaces when $m_{W'}$ approaches the top quark mass (m_t) of 172.5 GeV. At higher mass points, two effects cause the cross-section time branching ratio's rapid decrease: the cross-section proportionate with the resonance mass scale's inverse squared (Breit-Wigner formula [20]) and the diminishing number of partons to have longitudinal momentum close to the LHC proton beam energy of 6.5 TeV.

Perturbative calculations in QCD require additional parameters – the strong coupling constant α_s , the renormalization scale μ_R , and the factorization scale μ_F . Their nominal values are $\alpha_s(M_Z^2) = 0.1180$, and $\mu_R = \mu_F = m_{W'}$. The scale dependence and the m_t 's uncertainty is the next section's topic. For the all-hadronic W' search, the branching ratio of a top-quark decaying to a bottom quark and two first- or second-generation quarks (Figure 5.1) is 0.665 [20].

5.2 Theoretical Uncertainties

There are four sources of theoretical uncertainties for the NLO cross-section – the PDF, scale, α_s , and top-quark mass (m_t). The first three arise from perturbative expansion to calculate probabilistic amplitudes. The PDF is subject to strong dynamics incalculable by perturbation theory, thus is determined empirically. The factorization scale μ_F governs the separation between strong dynamics and free partons in handling the PDF’s infrared singularities [30]. The renormalization scale μ_R removes ultraviolet singularities while tuning the strong coupling constant, denoted as $\alpha_s(\mu_R^2)$. α_s fitting to data at the scale of the Z boson mass defines $\alpha_s(M_Z^2)$. For better convergence in the perturbative series, μ_R and μ_F are evolved to $m_{W'}$ in this analysis by employing the renormalization group equations [20].

Following the PDF4LHC recommendation [69], the cross-section’s PDF uncertainty is derived from the sample standard deviation of 100 cross-sections, each calculated with a different MC replica in the PDF set `PDF4LHC15_nlo_mc_pdfas`

$$\delta^{\text{pdf}}\sigma = \sqrt{\frac{1}{99} \sum_{k=1}^{100} (\sigma^{(k)} - \langle \sigma \rangle)^2}, \text{ where } \langle \sigma \rangle = \frac{1}{100} \sum_{k=1}^{100} \sigma^{(k)}. \quad (5.2)$$

Note that the cross-sections’ mean $\langle \sigma \rangle$ is not identical to the cross-section computed with the PDF replicas’ mean, $\sigma^{(0)}$, which is the nominal PDF. The PDF set also provides two separate PDFs to account for $\alpha_s(M_Z^2)$ ’s up and down variation by 0.0015. The recommended uncertainty of $\alpha_s(M_Z^2)$ is given by the averaged difference between the cross-section with the two PDFs

$$\delta^{\alpha_s}\sigma = \frac{\sigma(\alpha_s(M_Z^2) = 0.1195) - \sigma(\alpha_s(M_Z^2) = 0.1165)}{2}. \quad (5.3)$$

Then, the above two uncertainties are summed in quadrature, giving a symmetric uncertainty

$$\delta^{\text{pdf}+\alpha_s}\sigma = \sqrt{(\delta^{\text{pdf}})^2 + (\delta^{\alpha_s})^2}. \quad (5.4)$$

Although the PDF and α_s uncertainties could be interpreted as statistical errors of fitting to supplementary measurements, the renormalization and factorization scale uncertainties are educated guesses of the size of missing higher-order terms in the perturbative expansion of cross-section. As the sum of all terms should be scale-independent, this size is conventionally estimated by scaling up or down the renormalization and factorization scales by a factor of two [20]. Here both scales are shifted simultaneously. The scale variation, in general, is asymmetric, so there is an up uncertainty

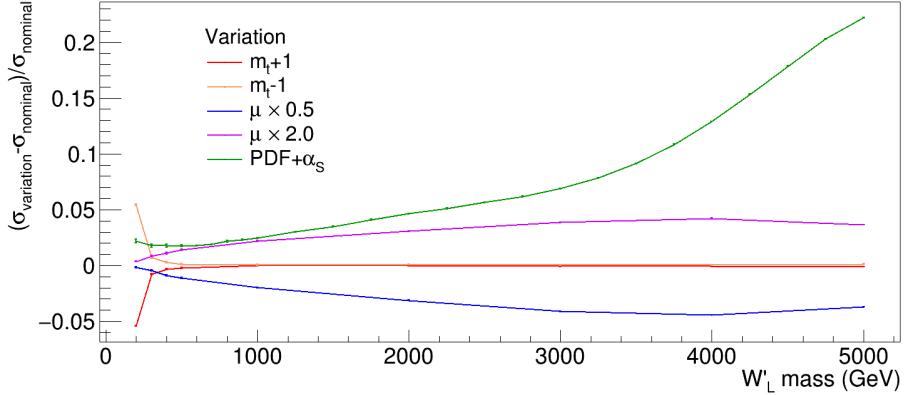
$$\delta^{\mu+}\sigma = \sigma(\mu_R = \mu_F = 2m_{W'}) - \sigma(\mu_R = \mu_F = m_{W'}), \quad (5.5)$$

and a down uncertainty

$$\delta^{\mu-}\sigma = \sigma(\mu_R = \mu_F = m_{W'}) - \sigma(\mu_R = \mu_F = 0.5m_{W'}). \quad (5.6)$$

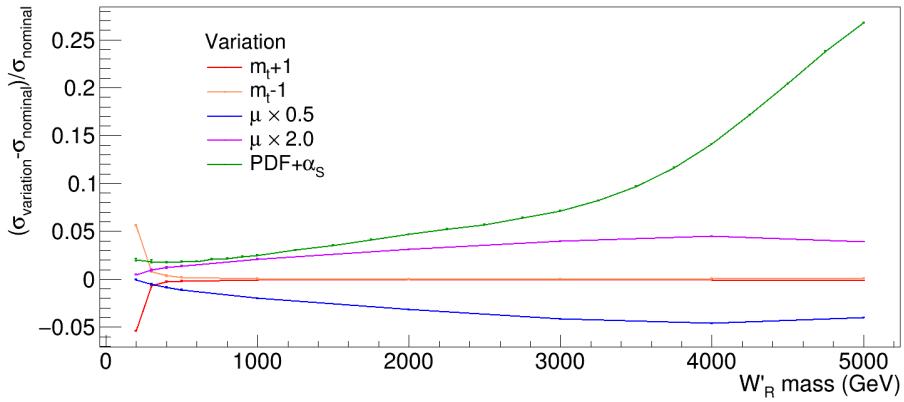
The fourth crucial theoretical uncertainty concerns the top-quark mass, which has a significant impact on the decay width $\Gamma(W' \rightarrow tb)$ when m_t is close to the W' mass. The chosen top quark mass of 172.5 GeV with an error of 1 GeV is consistent with the ATLAS combined result of 172.69 ± 0.25 (stat.) ± 0.41 (syst.) GeV [76]. Recomputing cross-sections with varied m_t provides another pair of asymmetric errors. Since the m_t is anti-correlated

Relative Uncertainties of the Cross-Section of $p p > W'_L \rightarrow tb$ at 13 TeV and NLO



(a)

Relative Uncertainties of the Cross-Section of $p p > W'_R \rightarrow tb$ at 13 TeV and NLO



(b)

Figure 5.4: The relative theoretical uncertainty of $\sigma(pp \rightarrow W' \rightarrow tb)$ for different $m_{W'}$ and chirality. The uncertainty from shifting m_t up (red) or down (orange) by 1 GeV dominates at low $m_{W'}$. Larger $m_{W'}$ sees increases of uncertainty from both scaling μ ($= \mu_F = \mu_R$) by a factor of 0.5 (blue) or 2.0 (purple) and the one-sided PDF+ α_s uncertainty (green). The latter dominates in the highest region of $m_{W'}$.

to the decay width (hence the cross-section too), the up uncertainty for m_t is given by

$$\delta^{m_t+}\sigma = \sigma(m_t = 171.5 \text{ GeV}) - \sigma(m_t = 172.5 \text{ GeV}), \quad (5.7)$$

and the down uncertainty by

$$\delta^{m_t-}\sigma = \sigma(m_t = 172.5 \text{ GeV}) - \sigma(m_t = 173.5 \text{ GeV}). \quad (5.8)$$

Finally, the total theoretical uncertainty is the sum of the PDF+ α_s , scale, and m_t uncertainties in quadrature

$$\delta^\pm\sigma = \sqrt{(\delta^{\text{pdf}+\alpha_s}\sigma)^2 + (\delta^{\mu\pm}\sigma)^2 + (\delta^{m_t\pm}\sigma)^2}, \quad (5.9)$$

The three uncertainties' ratios to the total are plotted for W'_L in Figure 5.4a and for W'_R in Figure 5.4b.

5.3 Event Generation

Even though ZTOP can compute differential cross-sections, such as the transverse momentum distribution of the top quark from the W' 's decay, its prediction cannot be compared with experimental observables straightforwardly. As mentioned in Chapter 2, the quarks and gluons coming out of the hard-scattering undergo parton showers and coalesce into hadrons. What the ATLAS detector measures are these hadrons or their decay products' energy deposits. Recorded events employ algorithms to construct physical objects such as jets from the energy deposit to infer the physics occurring at the hard-scattering. Generating such a

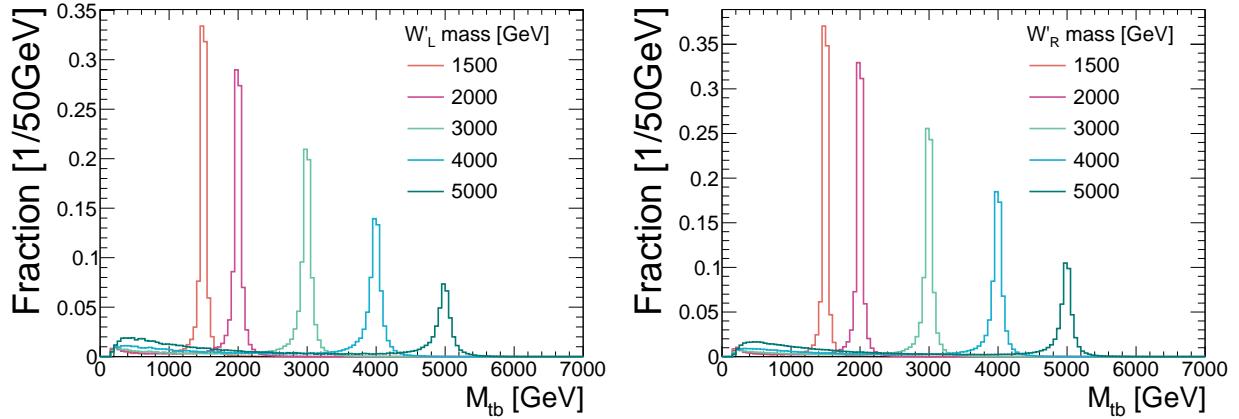


Figure 5.5: Parton-level (without FSR) invariant mass shapes of the top quark and bottom quark decaying from (a) left-handed or (b) right-handed W' with 5 different rest masses are compared by normalizing the area to 1. Each distribution is a cumulation of around 300K events.

complicated evolution requires dedicated programs.

The first stage is the parton-level event generation using `MADGRAPH5_AMC@NLO` v2.2.3 [77]. The program calculates the hard-scattering matrix-element of a W' boson produced in proton-proton collisions with decay to a top quark and a bottom quark. Five W' rest masses are considered for both left- and right-handed W' . Although the matrix element is computed only at the leading order in QCD, the W' production cross-sections, branching ratios, and decay widths are set to their next-to-leading-order values. The program's output contains the final-state t, b quarks' four momenta per event. Their four-momentum-sum's mass is the invariant mass M_{tb} , whose distribution is shown in Figure 5.5 for each mass point and chirality. Common to both W'_L and W'_R , the invariant mass centers at the W' rest mass, corresponding to a resonance. As the rest mass ascends, the peak widens and has a long tail to the low end due to an increasing decay width¹ and decreasing parton densities carrying high momenta. What's uncommon is that the W'_L has a broader resonance than the W'_R with the same mass, reflecting the former's extra leptonic decay channels.

¹roughly proportionate with mass

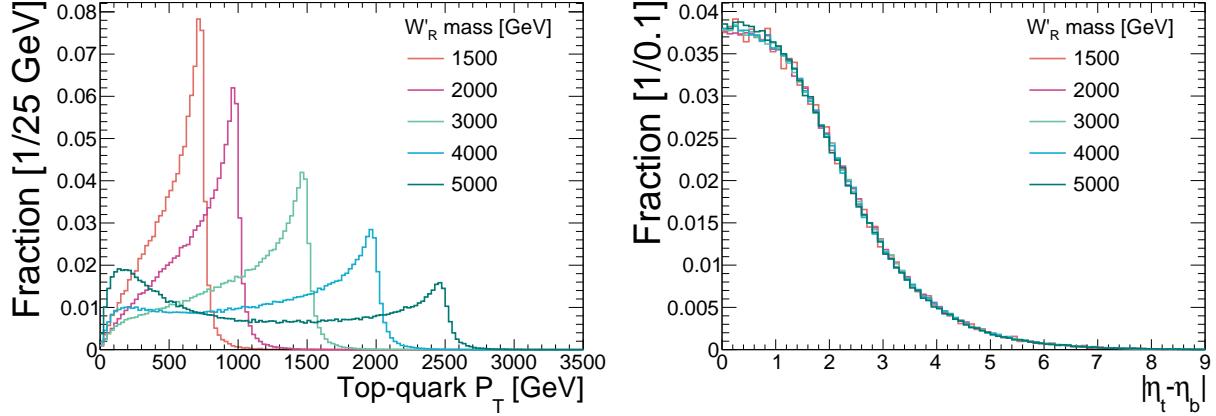


Figure 5.6: (a) The top-quark transverse momenta and (b) the top-quark to bottom quarks pseudorapidity difference's absolute value for right-handed W' with 5 different rest masses are compared by normalizing the area to 1. Left-handed W' bosons share similar kinematic properties

Peaking M_{tb} at the W' rest mass ($m_{W'}$) also influences the top-quark and bottom-quark transverse momenta. Given the initial-state partons' insignificant transverse momenta, the top-quark and bottom-quark are produced back-to-back on the transverse plane. Furthermore, as the top-quark and bottom-quark masses are negligible compared to $m_{W'}$ at the TeV scale, both quarks' transverse momenta share at most one-half the invariant mass by energy conservation. As shown in Figure 5.6(a), the top-quark momentum has a parabola-like increase and then a sharp drop-off near $0.5m_{W'}$. Figure 5.6(b) depicts the top-quark and bottom-quark's pseudorapidity difference, a quantity invariable under the Lorentz boost along the beam pipe. The preference for lower pseudorapidities are uniform among various $m_{W'}$ and is consistent with the rising top-quark transverse momenta.

What distinguishes a left-handed W' from a right-handed one the most is the top-quark spin. A left (right)-handed W' boson only couples to fermions with left (right)-handed chirality. A fermion's chiral eigenstate approaches the helicity eigenstate of the same handedness (or the opposite handedness for an anti-fermion) when its Lorentz boost factor γ is signif-

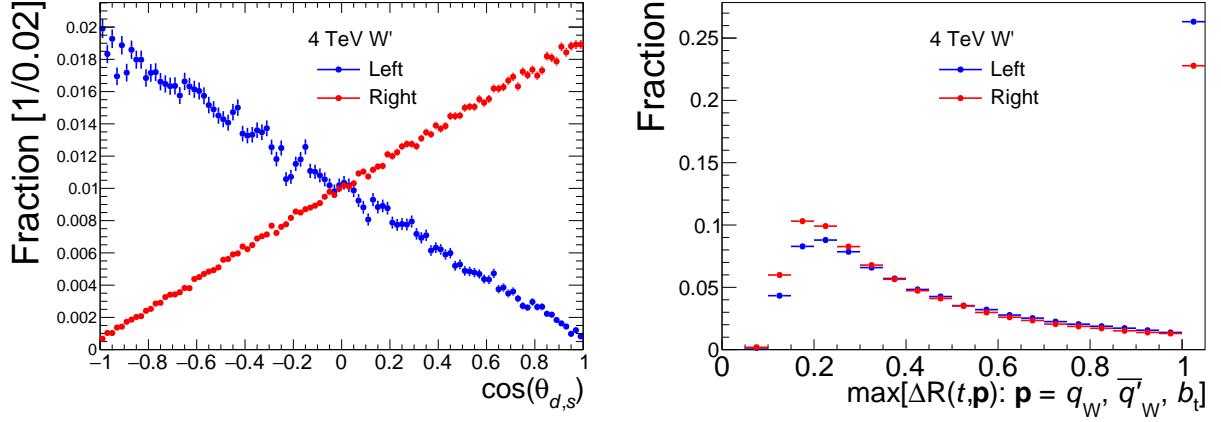


Figure 5.7: (a) The angle between the down-quark’s momentum in the top-quark’s rest frame and the top-quark’s momentum in the W' rest frame and (b) the maximum angular distance ($\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$) between the top-quark and its three decay products in the lab frame for a left-handed (blue) or right-handed (red) W' with a 4 TeV rest mass are compared by normalizing the area to 1. The rightmost bin in (b) is an overflow bin that encompasses everything above 1.

icantly more than 1. In the center-of-mass frame of a W' with a mass of several TeV, the daughter top-quark with around 173 GeV mass assumes γ of order 10. So the top-quark decaying from a left-handed W' mostly has left-handed helicity, meaning that the top-quark spin is anti-parallel to its momentum. Following similar reasoning, a right-handed W' makes the top-quark spin almost always parallel to its momentum.

A top quark decays to a bottom quark and an SM W boson whose subsequent hadron-decay-particles (specific to this analysis) consist of a down-type antiquark (\bar{d} or \bar{s}) and an up-type quark (u or c). In the top quark’s rest frame, the energy released to the outgoing quarks is large enough for them to be considered helicity eigenstate, as the top quark has been in the W' ’s rest frame. Together with the SM W boson’s left-handed-chiral couplings, the bottom quark must have left helicities. If the top-quark is a spin eigenstate toward positive z-direction, conservation of angular momentum yields negative (positive)-z bottom-quark spin for a W boson with left (longitudinal) helicity. The W boson’s momentum thus preferably

points to the negative (positive) z-axis. Further helicity analysis on the W boson's daughter quarks will reveal that the down-type antiquark tends to have a positive z-momentum, i.e., parallel to the top-quark spin [75].

Given that the top-quark momentum (p_t) is anti-parallel (parallel) to its spin in the left (right)-handed W' 's rest frame, the down-type antiquark's momentum in the top quark's rest frame ($p_{d,s}$) is more likely to be away from (toward) p_t . This claim is supported by the distribution of $\cos(\theta_{d,s}) = p_t \cdot p_{d,s} / |p_t||p_{d,s}|$, as shown in Figure 5.7(a). While the down-type antiquark's moving direction relative to the top quark has considerable dependence on the W' chirality, this distinction is much weaker among the other two quarks from the top quark decay [78]. So the down-type antiquark's behavior alone can help understand why the right-handed W' is more likely to have all three decay products within a cone of $|\Delta R| = 1$ around the top quark momentum in the lab frame. Such an effect is reflected by the distinctive maximum angular distance between the top quark and its three daughter quarks shown in Figure 5.7(b). The radius parameter is the same as the large-radius jet for reconstructing the hadronic top-quark decay (c.f. Chapter 6).

The parton-level events given by the **MADGRAPH** simulation only contain the bottom quark directly from the W' decay and the top-quark decay's three quarks. The second program, **PYTHIA** v8.186 [79], adds initial-state and final-state QCD parton showering in a repeated fashion. These showering partons are in the soft-and-collinear limit of QCD, meaning they have small momenta or small angles relative to their parent partons. After each vertex, the energy is taken away by the radiated partons. Once the parton energy reduces to near the hadronization scale, QCD coupling becomes strong enough to bind partons into hadrons. **PYTHIA** describes the hadronization process as the breaking of strings whose potential energy is linear to the separation distance between a quark-antiquark pair. **PYTHIA**

is also a toolbox of other QCD phenomena to emulate the pile-up of proton interactions in the same LHC bunch-crossing.

The hadrons are generally unstable and decay as they traverse through the ATLAS detector. Particles with lifetimes of more than 10 picoseconds (a picosecond = 10^{-12} second) are collectively specified as the particle level or the truth level [80]. These relatively stable particles will leave marks in the tracking system if they carry electric charges, and those who live longer proceed to stimulate electromagnetic and hadronic interactions in the calorimeter. GEANT4 [81, 82] generates these interactions step-by-step. Simulated detector signals will be reconstructed by algorithms coded in Athena [83] to form physical objects like jets, mimicking event reconstructions for data collected during actual proton collisions.

Chapter 6

The Identification of Top and Bottom Quarks

This chapter introduces jet as a physical observable to identify the top and bottom quarks in the W' signal's final state. Large-radius jets that characterize a hadronic-top quark is discussed, emphasizing the associated boosted top-quark DNN algorithm. A parallel description is given for the small-radius jet and its bottom-quark discriminating variable, DL1r.

6.1 Jet

QCD jets, jets of hadronic particles, or simply jets refer to bundles of hadrons shown as isolated energy flows commonly observed in the proton-proton collision at the LHC. Recall in Chapter 2 and Chapter 3, QCD describes jets as repeated soft and collinear parton radiation from final-state partons in a hard-scattering event and subsequent hadron formation. While this phenomenon resembles the electromagnetic shower composed of Bremsstrahlung and photon conversions into electron-positron pairs, hadronic jets' cascades are more common to appear since the QCD coupling is larger than the QED counterpart and gluons can directly couple to gluons [84]. Another difference is due to virtual masses: high-energy partons generally have off-shell massess while electrons need matter nuclei to interact with to radiate

photons without violating four-momentum conservation.

The rulebook of assigning each event’s hadronic energy spreading in the coordinate space into jets is the jet algorithm. For jets being both experimentally and theoretically defined objects, three different levels of jets are defined: parton-levels in perturbative calculations of matrix elements, truth-level after MC simulation of parton showers and hadronization, and reconstruction-level with particle tracks or calorimeter clusters from detector emulation or as recorded from data. Consistent jet configurations in all three levels require jet algorithms to be insensitive to additional soft or collinear radiation from parton radiation or hadronic decays [25].

One group of algorithms meeting the criterion is the sequential recombination algorithm. It clusters jet constituents separated by an angular distance ΔR less than a specified parameter R into jets, starting from the nearest neighbors. For a specific algorithm, the sequence of combining jet constituents can depend on their p_T ’s and different schemes to assign jet four-momenta exist (a comprehensive review is given in [25]). This analysis adopts the direct four-momentum summing scheme (the E -scheme) and two particular algorithms differing in their p_T sorting: the anti- k_t [85] that considers p_T ’s maximum among jet constituents and the k_t [86–88] that takes the minimum. The anti- k_t is used to reconstruct large-radius jets for the top-quark and small-radius jets for a bottom quark, while the k_t is utilized to identify a large-radius jet’s substructures. Note that an anti- k_t jet’s radius refers to the size parameter R , as the algorithm’s prioritization of high p_T particles gives circular boundaries to the higher p_T jets in an event.

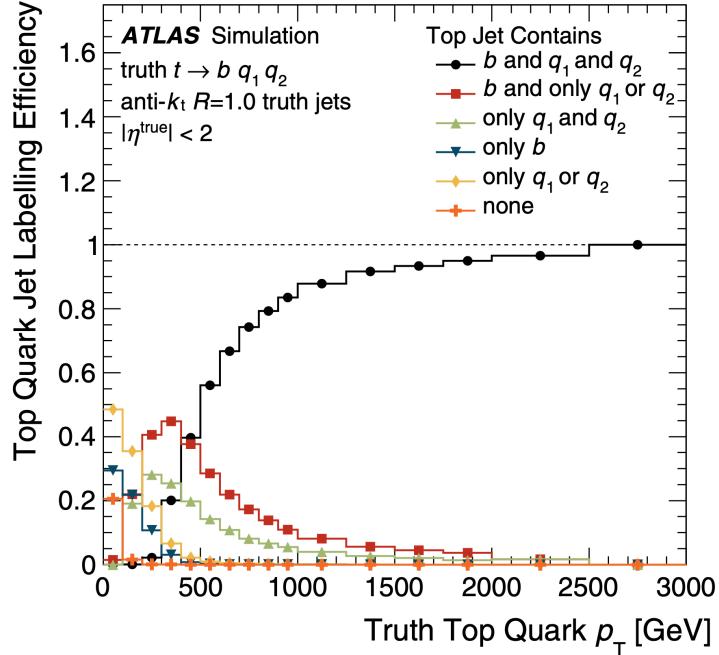


Figure 6.1: The probabilities of truth-level large-radius jets having a top quark and various combinations of the top decay products within $\Delta R < 0.75$ from the jet axis are shown as a function of the top quark’s transverse momentum. A top quark decays first to a bottom quark (b) and a W -boson, which further decays to two first or second generation quarks (q_1, q_2). [80]

6.2 Large-radius Jet — Local Cluster Topological Jet

As pointed out in Chapter 3, when the W' boson with \sqrt{s} times more mass than the daughter top quark, the angles between the decay products of the top quark are shrunk by a Lorenz boost. The Lorentz boost’s effect is illustrated by a simplified example of particle 1 with large transverse momentum p_T decaying into two particles 2 and 3 with negligible masses compared to p_T . Assuming that m_1 is the mass of particle 1 and particle 2 and 3’s

p_T , η , and ϕ variables are labeled with their corresponding subscripts:

$$m_1 = \sqrt{2p_{T_2}p_{T_3}(\cosh(\eta_2)\cosh(\eta_3) - \cos(\phi_2 - \phi_3) - \sinh(\eta_2)\sinh(\eta_3))} \quad (6.1)$$

$$\begin{aligned} &\approx \sqrt{p_{T_2}p_{T_3}((\eta_2 - \eta_3)^2 + (\phi_2 - \phi_3)^2)} \\ &= \sqrt{p_{T_2}p_{T_3}}\Delta R_{23} \end{aligned} \quad (6.2)$$

Therefore, the angular distance between particle 2 and 3, ΔR_{23} , shrinks with their transverse momenta, p_{T_2} and p_{T_3} , that increase with p_T . In this analysis, the large-radius jet is an anti- k_t jet with its size parameter $R=1$, and only large-radius jets with $p_T > 500$ GeV are considered. According to Figure 6.1, 60 percents of the large-radius truth-level jets with a p_T of 500 GeV contain all three quarks from the top quark decay, and the rate increases further with p_T .

Reconstructing a large-radius jet starts with applying anti- k_t algorithms to topological clusters with local hadronic calibrations [89]. These clusters are built from calorimeter cells energetic enough compared with typical electronic and pile-up noises. The clusters' energy has to be corrected to compensate for energy loss for various reasons such as less energy reading from energy deposits by charged pions than by electrons, photons, or neutral pions. The difference in calorimeter's response to electromagnetic versus hadronic showers is corrected by applying weights to hadron-like energy deposits.

After the cell-by-cell corrections to the topological clusters, there could still be energy from pile-up [90], multi-parton scattering, and initial state radiation that contaminate the showering pattern of a top quark decaying to three quarks. This issue is mitigated by applying jet trimming [91]. It begins with sorting the topological clusters in each large- R jet into sub-jets with $R = 0.2$, using the k_t algorithm. The sub-jets with transverse momenta of

less than five percent relative to the original large- R jets are removed. Topological clusters from the remaining sub-jets are recombined into the “trimmed” large- R jet, which is the final reconstructed large-radius jets.

Each reconstructed jet’s energy and mass are calibrated to the truth level [54]. In simulated events, each trimmed reconstructed jet is matched to the closest trimmed truth-level jet on the $\eta - \phi$ plane. The reconstruction-level jet energy’s ratio to the truth-level energy is fitted to a Gaussian distribution whose central value is the energy’s correction factor. The jet mass calibration follows the same method but combines track-based and calorimeter-based measurements to minimize the mass resolution. Additional *in-situ* jet energy and jet mass correction is applied to data to account for uneven detector response in η and residual differences after MC-based correction factors.

6.2.1 Large-radius Jet Trigger

Large-radius jets are reconstructed both online in the High-Level Trigger (HLT) and offline in the analysis stage. During the ATLAS run, events triggered by a Level-1 jet trigger of 100 GeV (Chapter 4) are processed through topological cluster formation and jet reconstructions adapted from the offline analysis software. The large-radius jet trigger is based on anti- k_t $R = 1$ jet with different pile-up removal methods and energy calibration depending on the data-taking year. Since 2017, jet trimming, jet energy and mass correction were implemented in the HLT system [61].

This analysis preserves only data triggered by large-radius jet triggers with E_T thresholds ranging from 380 GeV to 460 GeV as the signal event’s physical signatures. A higher E_T threshold is necessary to contain trigger rates within the data acquisition system’s readout capacity since events with more pile-up interactions meet the same trigger threshold more

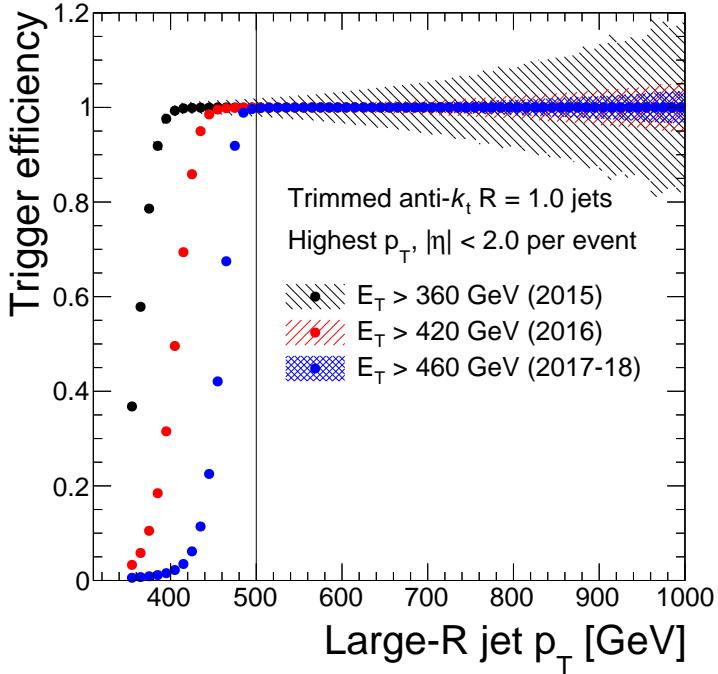


Figure 6.2: The efficiencies of three large- R jet triggers for data collected in the year 2015 (black), 2016 (red), and 2017 to 2018 (blue) are shown as a function of the (offline) leading large-radius transverse momentum. The legend lists each trigger’s (online) transverse energy thresholds. The efficiency is computed by the coincidence of the targeted trigger and a lower threshold trigger – the (online) leading small-radius jet’s $E_T > 260$ GeV. All three trigger efficiencies reach 100% at (offline) leading large-radius jet $p_T > 500$ GeV (drawn as a vertical black line). The shaded area represents statistical uncertainties.

frequently. Figure 4.3 shows that the pile-up increases with years, resulting in more hadronic energy to fulfill the trigger cutoff. To prevent the data stream from overloading, some low energy triggers’ acceptance are considered only every N times. This number N is termed the “pre-scaling factor,” resulting in triggered event losses. The analysis avoids this loss by having an offline selection criterion of large-radius jet’s $p_T > 500$ GeV, thereby ensuring full trigger efficiencies, as shown in Figure 6.2.

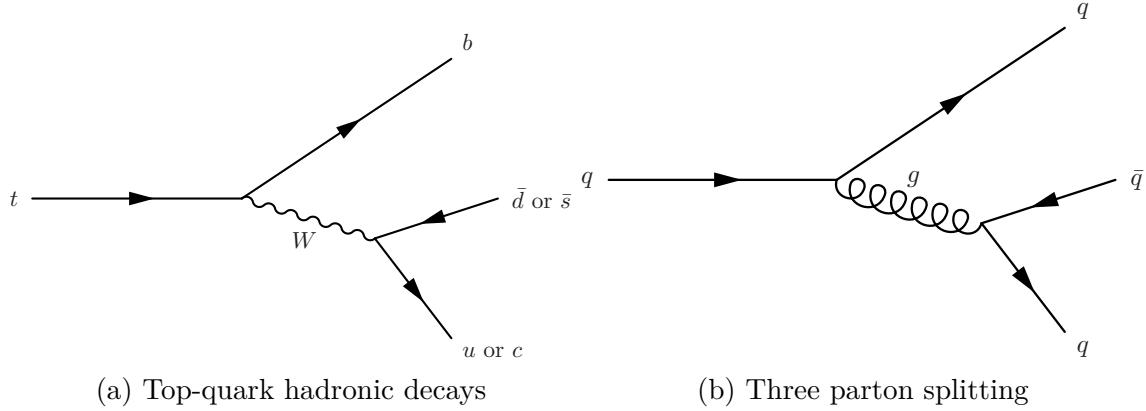


Figure 6.3: Feynman diagrams of (a) the top-quark signal and (b) a light-quark background for the top-tagging algorithm.

6.2.2 DNN top tagging

Top tagging is a technique for finding large- R jets containing a hadronic top-quark decay and reject those containing only QCD showering of quarks and gluons by applying selection to variables that discriminate between the two. The three quarks in a hadronic top quark's decay, shown in Figure 6.3(a), frequently have similar energy and wide opening angles between them. In contrast, three partons produced by QCD radiation, shown in Figure 6.3(b), are more likely to be collinear in momentum and less democratic in energy shares. Thus, many jet substructure variables exploit the distinction between the two kinds of jets' internal energy distributions. One of the variables is the jet mass, calculated from the jet constituents' four-momenta. As the top quark is the heaviest elementary particle, a top-quark large-radius jet's mass is, on average, significantly more than the jet mass originating from parton splitting in light quark jets or gluon jets.

A second jet substructure variable is the splitting scale [93]. Its calculations start with combining hadronic clusters inside a large-radius jet into proto-jets using the k_t algorithm with $R = 1$. The combination sequence is determined by the “closeness” between two (merged or not) constituents of the jet, labeled as i and j , inside a large- R jet. The closeness is defined

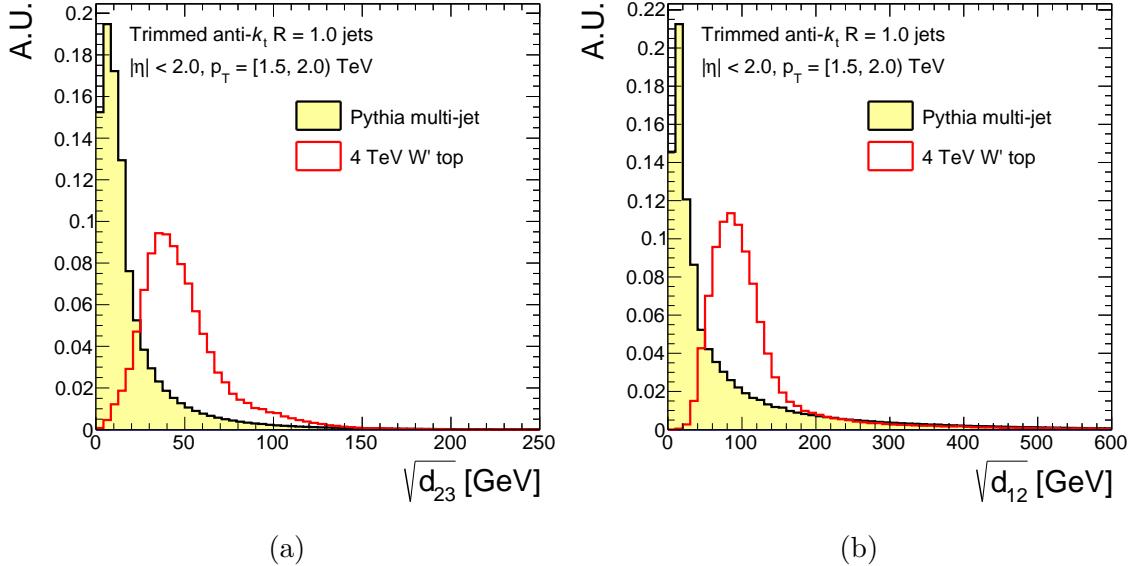


Figure 6.4: The splitting scales $\sqrt{d_{23}}$ and $\sqrt{d_{12}}$ of large- R jets are shown for two cases – top-quark decay from a W' boson with a rest mass of 4 TeV (red line) and a gluon or non-top quark from QCD multi-jet processes simulated by PYTHIA (black line with yellow fill). In the first case, large- R jets are matched to top quarks first by requiring $\Delta R < 0.75$ between the reconstructed jet and its closest truth-level jet, and between the truth-level jet and the top quark. Moreover, there must be at least one B hadron ghost-associated [92] with the truth-level jet. Lastly, the reconstructed jet mass has to surpass 140 GeV. This definition of top-quark jets is the one used in training the DNN top-tagger.

by the splitting scale

$$\sqrt{d_{ij}} = \min(p_{T_i}, p_{T_j}) \frac{\Delta R_{ij}}{R}, \quad (6.3)$$

where the R parameter is the same as the concerned anti- k_t large-radius jet. The k_t algorithm repeatedly combines the pair i and j with the lowest $\sqrt{d_{ij}}$ into a proto-jet until one proto-jet is left, which is just the original jet¹. As soft and collinear parton radiation suppresses the p_T and the ΔR factor in the splitting scale, they are combined into proto-jets corresponding to hard partons initiating the QCD showers. In a top-quark jet, the three last proto-jets are likely to correspond to the three decay quarks depicted in Figure 6.3(b). The splitting scale between these proto-jets can be associated with their mother particle's mass by comparing

¹The beam jet distance d_{iB} in the k_t algorithm for jet-finding is irrelevant since all jet elements clustered by the anti- k_t must have their mutual $\Delta R < R$

Equation 6.3 with Equation 6.2. Thus, the splitting scale for merging three proto-jets into two, $\sqrt{d_{23}}$, peaks around half the W boson mass (Figure 6.4a). Similarly, the splitting scale for clustering two proto-jets into one, $\sqrt{d_{12}}$, has a maximum of around half of the top quark mass (Figure 6.4b).

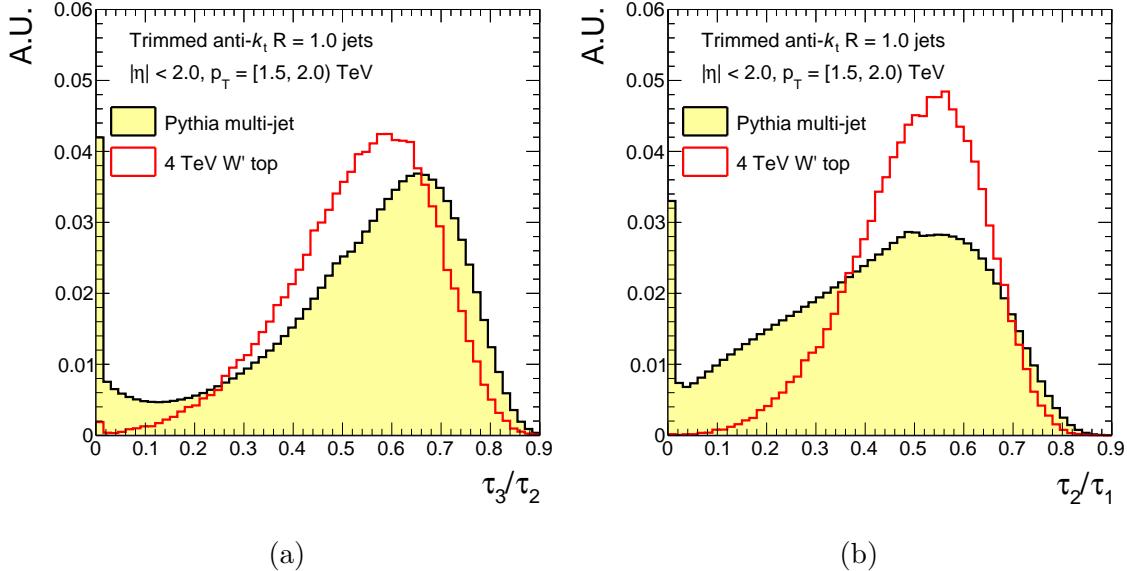


Figure 6.5: The ratios of 3-subjettiness (τ_3) to 2-subjettiness (τ_2) and of 2-subjettiness (τ_2) to 1-subjettiness (τ_1). The definition of top-quark matched large- R jets is written on the caption under the plots for the splitting scale. PYTHIA simulated multijet events' subjettiness ratios have a spike at 0 since large- R jets consisting solely of parton showering could have only one or two constituents. The former causes τ_1 to be 0 and the latter causes τ_2 to be 0.

While the splitting scale measures the mass scale among subjets, the N -subjettiness [94, 95] calculates the energy spread among individual jet constituents – topological clusters in this analysis – relative to the number of subjets (N). Assuming that there are M constituents and ΔR_{ij} is the radius between the constituent i and subjet j , the definition of N -subjettiness adopted in this analysis is

$$\tau_N = \left(\sum_i^M p_{T_i} \min(\Delta R_{i1}, \Delta R_{i2}, \dots, \Delta R_{iN}) \right) \Bigg/ \left(\sum_i^M p_{T_i} R \right), \quad (6.4)$$

where R is the radius of large- R jets. Similar to splitting scales, the sub-jet is found by the k_t algorithm. But to reduce the effect of wide-angle, low energy radiation on deflecting the jet-axis, the “winner-take-all” scheme [96] of combining sub-jet momenta is applied. As the jet constituents’ energy for top-quark jets is expected to be more aligned with three axes than two axes or one axis, the ratio between τ_3 and τ_2 for top-quark jets tends to have a lower value than QCD jets (Figure 6.5a). In contrast, top-quark jets are outnumbered by QCD jets in low ratios of τ_2 to τ_1 when the latter has few or only 1 jet constituent (Figure 6.5b).

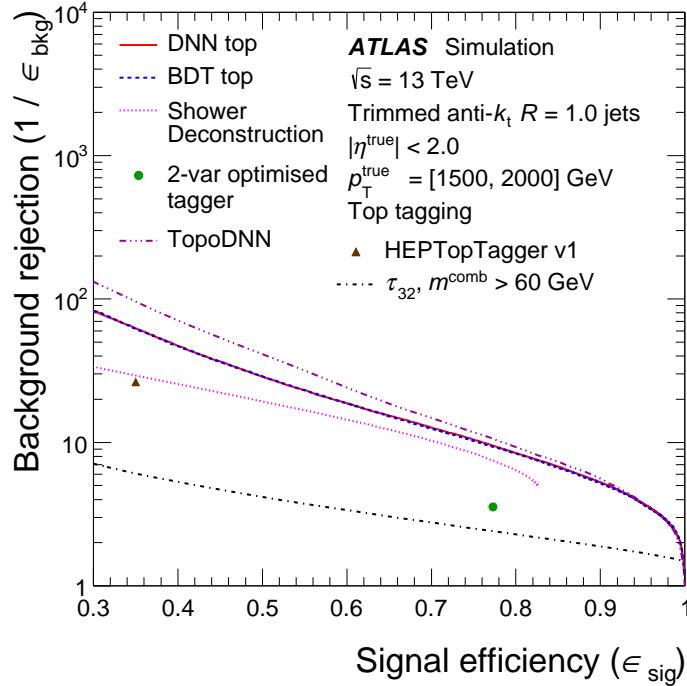


Figure 6.6: The Receiver Operating Characteristic (ROC) curves of different types of top-tagging studied in [80]. The curves on the plot show how many QCD large-radius jets are rejected per one accepted (y-axis), given what fraction of top-quark large- R jet is retained (x-axis) by taking a cutoff(s) on relevant tagging variable(s). The ROC curve of the DNN top-tagger employed in this analysis is shown in solid red, overlapping with the curve of the Boosted Decision Tree (BDT) top-tagger.

Jet substructure variables like jet mass, splitting scale, and N-subjettiness can be combined by multivariate techniques to enhance the discriminatory power of top-tagging. Furthermore, top-tagging is a “classification problem” in machine learning as the jet substructure

variables can be thought of as different “features” of two different classes of large- R jets – top-quark jet and QCD jets [97]. In this analysis, an advanced machine technique called the Deep Neural Network (DNN) [98] takes jet substructure variables as input. A selection cut can then be placed on the DNN’s output for a fixed likelihood of retaining top-quark large- R jets. By varying the selection cut, the power of rejecting QCD large- R jets for a given top-quark jet’s acceptance rate is presented as a ROC curve. The comparison of ROC curves in Figure 6.6 shows an improvement of background rejection power by DNN top-tagging compared to the Shower Deconstruction [99, 100] top-tagging, which is used in a recent search for the W' boson by ATLAS [101]. The top-tagging rates in the MC-simulated events are corrected by p_T dependent scale factors determined in data enriching in top quark jets [80, 102].

6.3 Small-radius Jet — Particle Flow Jet

A small-radius jet is an anti- k_t jet with a size parameter $R = 0.4$, which is employed to reconstruct the kinematics of the bottom quark directly decaying from the W' signal. Furthermore, a bottom-quark containing hadron – B -hadron – has a flight path resolvable by tracking its charged decay particles’ hits in the Inner Detector (ID). Thus, small-radius jets overlapping with large-radius jets are also found for the bottom quark identification. The small-radius jet’s input is Particle Flow objects [103] that consists of not only topological clusters used in the large-radius jet reconstructions but also ID tracks.

In the Particle Flow algorithm, charged hadron’s momenta are reconstructed from tracks while neutral hadrons’ energy is reconstructed from hadronic clusters. To avoid double-counting hadronic energy, calorimeter clusters calibrated to the EM scales are matched to

tracks closed on the angular distance relative to cluster widths. Energy deposited by charged hadrons of the track is removed from calorimeter cells within the clusters using the shower profile of pions. The profile is also used to add two clusters associated with a single charged hadron and determine the consistency between the remanent and neutral hadrons after cell-by-cell energy subtractions from a cluster. Since high-energy pions are more likely to be at the jet core area with dense hadronic activities, it would be more difficult to reconstruct and match tracks with hadronic clusters accurately with increasing pion p_T [103]. Therefore, a p_T -dependent selection criterion is imposed to prevent cluster energy vastly exceeding the expected pion energy deposit from being subtracted [55]. When the track $p_T > 100$ GeV, the associated hadronic cluster is not considered for removal.

One of the advantages of using tracks for charged hadron identification is the Inner Detector's superior energy resolution than the calorimeter in the lower p_T regimes. Since the segmentation of ID subdetectors is substantially less than the calorimeter cell divisions, the angular resolution is improved. Another benefit comes from the pile-up removal. Tracks originating from non-hard-scatter vertices are discarded in jet reconstructions. As the Particle Flow algorithm removes tracks-associated energy deposits, charged hadrons from pile-up interactions are automatically removed.

Remaing pile-up contribution is dealt with by p_T corrections modeled on per-event average jet p_T per jet area [92], number of reconstructed primary vertrices, and interaction counts (μ) [55]. Theres is also a jet-vertex tagger [104] evaluting the compatibility of jet's ghost-associated [92] tracks with pile-up vertices. Only jets with $p_T < 60$ GeV and $|\eta| < 2.4$ are examined jets by the tagger.

Like large- R jets, the energy of small- R jets is also calibrated to corresponding truth-level small- R jets after pile-up corrections. Unlike large- R jets, the energy scale correction

for hadronic clusters in the Particle Flow algorithms has not incorporated the local weighting procedure for hadronic showers. Therefore, the electromagnetic shower energy is the calibration scale, namely the EM scale. Afterward, there is a sequence of corrections based on tracks momentum and energy deposits in specific layers of calorimeters [105] to alleviate jet energy resolution's degradation due to showering differences between quark and gluon jets. There are also data-specific energy corrections derived *in-situ*. The *in-situ* calibration uses prior-calibrated physical objects like electrons decaying from a Z boson in Z +jets events to account for data-MC differences and di-jets events for inter-calibrations between the central and forward detectors.

6.3.1 DL1r b-tagging

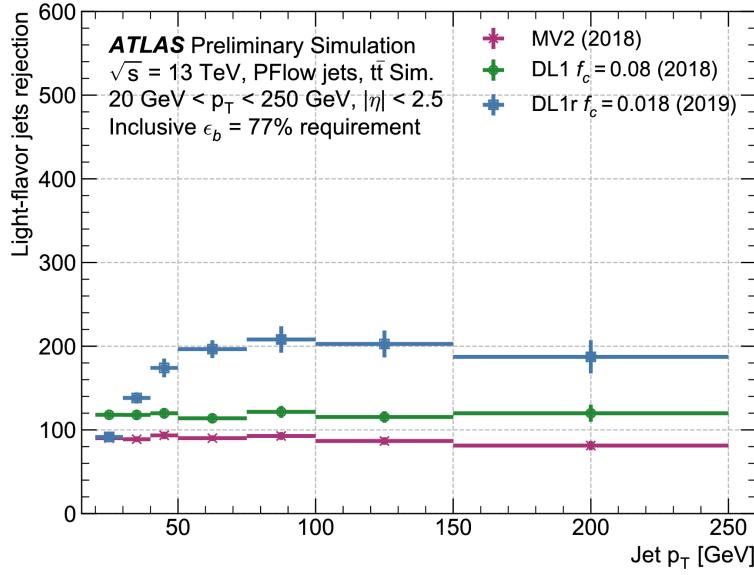


Figure 6.7: Three b -taggers' ability to reject light-quark jets as a function of the p_T of Particle Flow (PFlow) small- R jets are compared. The plot shows one out of how many light-quark small- R jets, a large background to this analysis, are accepted, given a 77% chance of retaining a b -quark jet with $p_T > 25$ GeV. The value f_c means the proportion of c -quark jet included with light quark and gluon jets in the background events. [106]

As stated in Chapter 2, a bottom quark is not an observable object because the strong

force at low energy scales binds it into B -hadrons through QCD fragmentation. Therefore, the two b-quark jets in the W' signal are identified by tracking the B -hadron decay's charged hadrons around a small-radius jets. Each jet considers all tracks within a cone around the jet axis with a radius ΔR of 0.45 at 20 GeV of jet p_T , decreasing gradually to 0.26 at 150 GeV [107]. The decreasing angular distance follows the Lorentz boost as described in the discussion of large-radius jets.

For each event, the spatial coordinates of particle interactions – primary vertices – are reconstructed using ID tracks [108, 109]. The particle's origin is traced by its closest approach – perigee – to the beam axis, following its motion due to the longitudinal magnetic field. The cluster of perigees of two or more tracks with $p_T > 0.5$ GeV allows for vertex fitting, and events without a vertex are rejected. Among all primary vertices, the one with the highest scalar sum of matched tracks' transverse momenta ($\sum p_T$) is deemed where the hard-scattering process occurs, thus given the name “the hard-scatter primary vertex².” In addition to a bottom-quark identification's purposes, the hard-scatter primary vertex is picked as the origin of the coordinate system for hadronic clusters constituting the neutral part of the small-radius jet or all hadronic energy in the large-radius jet.

The b -tagging variables aim to find features differentiating B -hadrons from other long-lived hadrons, photon conservations into leptons, and di-photon decay from neutral pions, or even particle interacting with detector materials. Sets of b -tagging variables evaluate the significance of jet-associated tracks' displacements from the hard-scatter primary axis. More sophisticated variables make use of pairs of tracks approaching each other in spaces. By first rejecting two-track vertices with properties consistent with backgrounds, a fast search for

²In some literature, this is called the primary vertex when secondary vertices of a hadron's or a τ lepton's decay are irrelevant

secondary vertices among numerous tracks in a high particle multiplicities can be achieved. In response to missing tracks, kinematic fitters are also employed to reconstruct the B -hadron and its decay particle's flight paths.

As in the case of top-tagging, these b -tagging variables serve as input to multivariate approaches [107]. Among these are a deep neural network called DL1 (ibid.) and a recurrent neural network (RNN) [110] for finding the sequential relation between the tracks. Figure 6.7 shows that the DL1 tagger incorporating the RNN, called the DL1r algorithm, could reject more light-quark jets than a boosted decision tree, the MV2 algorithm employed in the previous analysis [101]. The probability of the combined DL1r algorithm to tag a PFlow jet is calibrated for different hadron species emanating from jets as described in [107, 111, 112].

6.4 Truth-level and Reconstruction-level Observables

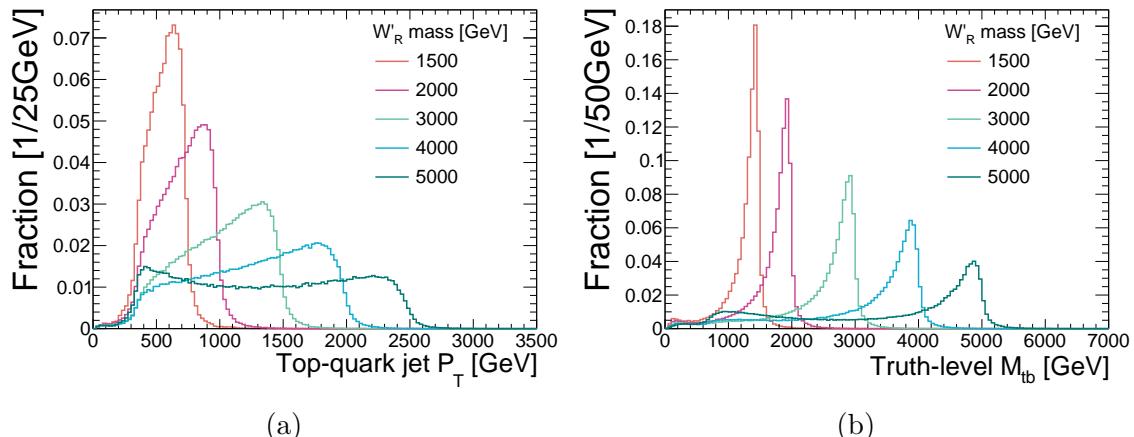


Figure 6.8: The truth-level (a) top-quark jet transverse momentum and (b) invariant mass M_{tb} for right-handed W' bosons with 5 different masses are shown by rescaling all histograms' area to unity.

The W' boson is observed by the $t - b$ invariant mass around its rest mass and the top quark and bottom quark transverse momenta's peaking edge at half of the rest mass. These

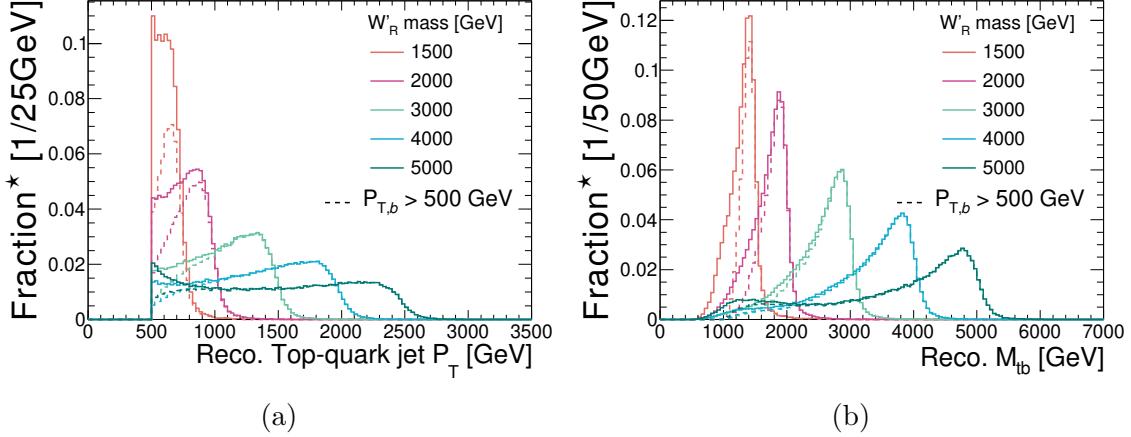


Figure 6.9: The reconstruction-level (a) top-quark jet transverse momentum and (b) invariant mass M_{tb} for right-handed W' bosons with 5 different masses. The solid-line histograms are scaled to have unit area, while the dashed-line histograms are normalized by the solid-line histogram's original area. Their difference lies in that dashed histogram excludes those events having its (W' -coupled) bottom-quark small-radius jet with $p_T \leq 500$ GeV.

distributions have been shown in Figure 5.5 and Figure 5.6 using MC simulated events for 5 different W' masses. Truth-level jets and reconstruction-level jets in the simulation are mapped to the top-quark and bottom quark to examine the impact of parton radiation and jet reconstruction on these signal characteristics.

Clustering hadrons (specified at the end of Chapter 5), truth-level jets shown in Figure 6.8 follow similar kinematic features as the top-quark and bottom-quark at the parton level. The highest p_T truth-level large-radius jet with $\Delta R < 1$ from the top-quark is selected as the truth-level top-quark jet. The bottom-quark jet is found in the same fashion among truth-level jets with $\Delta R = 0.4$. About 25% of W' events cannot have both jets located due to the jet reconstruction requirement of $|\eta| < 2.5$ and other jet p_T constraints related to data processing capacities. The p_T selection causes a hill-like shape around 500 GeV on the M_{tb} distribution.

Compared with Figure 5.6(a), the jet p_T spreads further to the lower end, meaning that the hadronic energy falling outside the jet algorithm's clustering range (R) is on average more

than the gain from initial-state radiation or pile-up. The combined effect from the top-quark jet and the bottom-quark jet broadens the invariant mass distribution that the full-width at half maximum (FWHM) becomes two times than that in Figure 5.5(b) across all W' masses. Though not shown on the graph, the M_{tb} 's width narrows by around one-tenth after using only truth-level top-quark jets containing the three quarks in the top-quark decay in the jet radius $R = 1$. Thus, the enlarged width is more associated with jet reconstruction and parton showers than the R -parameter not being wide enough.

Figure 6.9 shows similar plots with events having reconstructed large-radius jets satisfying a 500 GeV p_T threshold for the trigger requirement, leading to ten percent or less event losses from the truth-level for a 2 TeV or a higher mass W' . Among these high p_T reconstructed jets, the top-quark jet is the closest to the truth-level top-quark jet. The bottom quark jet is not required to be close to the bottom-quark-labeled truth-level jet since the latter might be surrounded by small-radius jets initiated by final-state radiation. Instead, the highest p_T jet with an azimuthal separation $\Delta\phi > 2$ from the reconstructed top-quark jet is chosen. The resulting bottom-quark jet at the reconstruction level is further examined with a restriction of $p_T > 500$ GeV.

From the truth level to the reconstruction level, the large-radius jet p_T shifts to the lower end further and the M_{tb} width doubles again. The p_T requirement on the bottom-quark jet reduces events with reconstructed top-quark jet $p_T < 1$ TeV significantly, narrowing the 1.5 and 2 TeV W' boson's invariant mass. The result shows that such a p_T cutoff can reject backgrounds with smoothly falling p_T distributions while retaining most signal events. The top-quark jet p_T distribution after this cut is more aligned to the truth-level one, suggesting that the reconstruction-level jet finds similar hadronic energy as the truth-level one when the identified top-quark and bottom-quark jets are more symmetric in p_T .

Chapter 7

Event Selection and Categorization

The previous chapter introduced large-radius and small-radius jets as physical observables to reconstruct the top-quark and bottom-quark decaying from the W' boson, as shown in Figure 1.1. In this chapter, the top-tagging and b -tagging algorithms and few to-be-introduced event selections will separate events into the signal region for the statistical analysis in Chapter 9 and the template, control, and validation regions for the data-driven background estimation in Chapter 8. Discussions on the data-driven method is given to motivate a specific usage of the large-radius jet called the top-proxy jet in defining the control region.

Charged lepton veto. Aside from this analysis, there is another search for the W' decaying to a top quark and a bottom quark using the leptonic top-quark decay. Combing the hadronic and leptonic search channel's likelihood functions (Chapter 3 and Chapter 9) enhances the deviation between the signal-plus-background hypothesis and the background-only hypothesis, namely, the signal's statistical significance. If the two channel's events are mutually exclusive, the joint likelihood function is just the product of the two individual likelihood functions. Therefore, events with any reconstructed electron or muon are excluded from this analysis for the statistical independence between the two searches. This charged lepton veto also helps to reject background events. The reconstruction and identification of electrons and muons are described in Appendix A.

Top-candidate jet. From events satisfying the large-radius jet trigger thresholds listed in Chapter 6, large-radius jets with $p_T > 500$ GeV and $|\eta| < 2.0$ are examined with the top-tagging DNN algorithm. If only one large-radius jet is top-tagged, that jet is considered the top-quark jet’s candidate, or say the top-candidate jet. Since the top-tagging rate for the bottom-quark large-radius jet is low, two top-tagged jets are mostly background events with two top quarks in the hard scatter interaction’s final state ($t\bar{t}$, more discussions in Chapter 8). Thus events with two or more top-tagged jets are removed to reject these background events.

Top-proxy jets. Data events without any top-tagged large-radius jets serve are utilized for background estimation. Ideally, all non-top-quark jets have identical top-tagging probabilities, so any large-radius jet satisfying the p_T and η requirements should be considered. Moreover, the number of events with exactly one top-tagged jet is proportional to the number of examined large-radius jets, implying that using all qualified large-radius jets is more suitable than choosing one particular jet. The above consideration motivates the definition of top-proxy jets as large-radius jets with $p_T > 500$ GeV and $|\eta| < 2.0$ when all such jets in the same event fail to be top-tagged. Top-candidate jets and top-proxy jets are on the two sides of the top-tagging axis for the background estimation shown in Table 3.1.

B-candidate jet. By the conservation of momentum, the bottom-quark small-radius jet’s momentum points in the opposite direction to the top-quark large-radius jet on the transverse plane, with deviations due to jet reconstruction and large-angle final-state radiation forming jets. So the bottom-quark jet’s candidate – b-candidate jet – is the highest p_T jet among all small-radius jets with $\Delta\phi > 2$ from the top-candidate jet or a top-proxy jet. The b-candidate jet’s b -tagging score serves as the second axis of the data-driven background’s grid.

Figure 7.1’s flowchart shows the steps in using the top/b-candidate and top-proxy jets

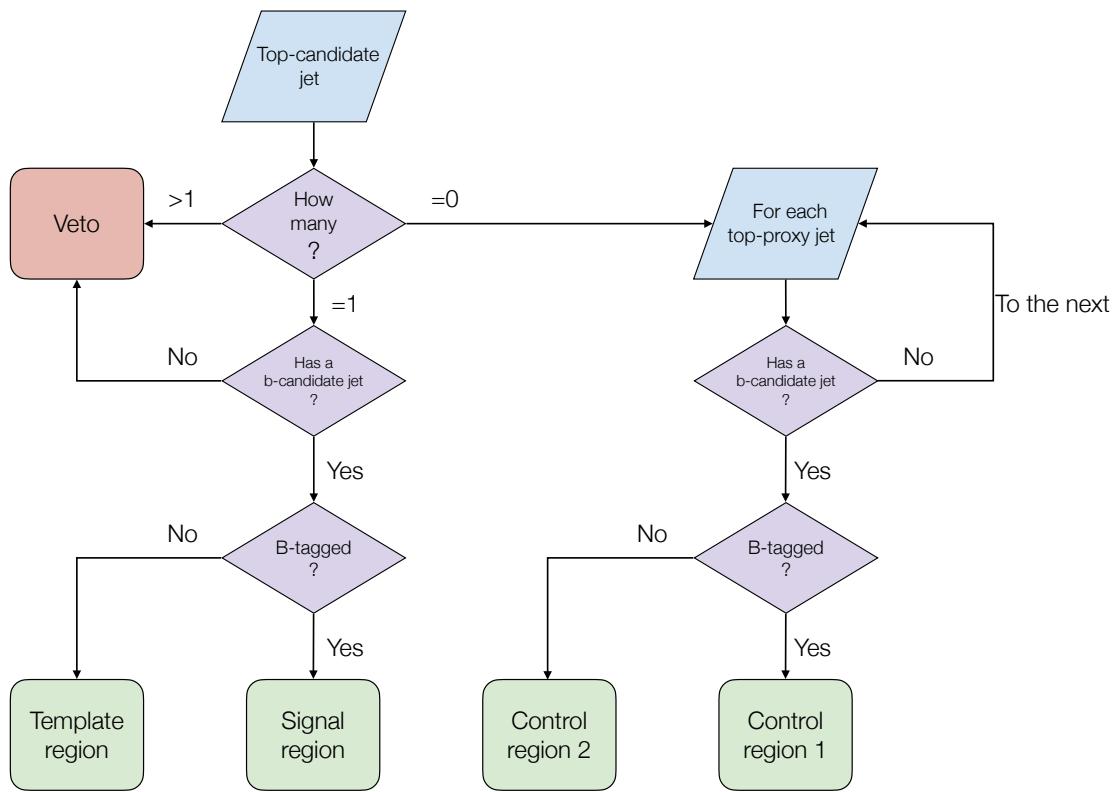


Figure 7.1: The flow chart showing the classification of events that lead to different regions. The top input block's top-candidate jet means the top-tagged large-radius jets.

to form four non-overlapping regions, named after their roles in the data-driven background estimation: the control regions (CR) 1 and 2, whose ratios are to be multiplied by the template region (TR) event counts bin-by-bin, giving the data-driven background estimation in the signal region (SR).

Number of b -tags inside the top-candidate (top-proxy) jets. The top quark’s decay almost always include a bottom-quark [20], thus having b -tagged small-radius jets with $\Delta R < 1$ from a top-candidate jet discriminates the signal from background events. Furthermore, hard-scatter final-state partons from the QCD multi-jet events can have two high p_T bottom quarks or two charm quarks favored by the b -tagging algorithm. Therefore large-radius jets with a bottom/charm quark origin are expected to see a higher b -tagging rate than a gluon/light-quark large-radius jet in the paired b -candidate jets. The change in b -tagging likelihood necessitates estimating backgrounds if b -tagged small-radius jets are present inside top-candidate (top-proxy) jets. These large-radius jets, no matter in the signal, template, and control regions, are separated into the **1 b-tag in category**. The **0 b-tag in category** incorporates the rest.

Top-tagging division. The DNN top-tagging algorithm has two thresholds on the DNN output. The threshold used for the top-candidate jet’s definition has an estimated 80% probability of retaining a top-quark jet [80] and is labeled as the **80 WP**. The higher threshold **50 WP** divides the signal and template regions to reject more backgrounds. After the categorization of b -tagged jets in large-radius jets and the further top-tagging divisions, the four regions from Figure 7.1 are turned into 12 cells shown in Figure 7.2. Except for renumbering, the signal region in the “0 b-tag in” category and situated between the 50 WP and 80 WP of top-tagging is renamed as the validation (VR). It has a negligible portion of signal events, allowing for verifying the accuracy of the data-driven background estimation

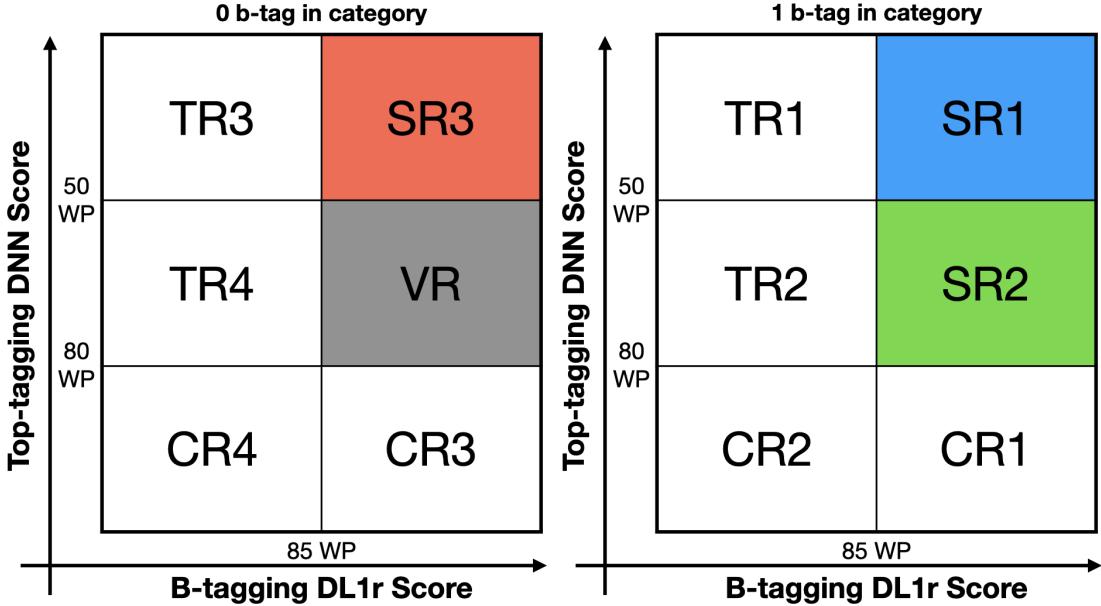


Figure 7.2: For each of the two categories, events are divided into six regions on a two-dimensional plane. The vertical axis represents the top-candidate jet's top-tagging score from low to high upward, while the horizontal axis represents the b-candidate jet's b -tagging score from low to high rightward. SR stands for the signal region, VR for the validation region, TR for the template region, and CR for the control region.

method.

Pseudorapidity requirement. The final selection is to remove events or top-proxy jets if the pseudorapidity difference between the top-candidate or top-proxy jet and its b-candidate jet ($\Delta\eta_{tb}$) is 2.0 or above. The parton-level studies in Figure 5.6(b) and the robustness of the p_T spectrum's overall peaking structure in the reconstruction-level of jets in Figure 6.9 reflect that the W' signal's preference for low $\Delta\eta_{tb}$ due to the s -channel resonance production. In contrast, the dominant background events tend to be more prominent in the t -channel or u -channel than s -channel of the 2-to-2 matrix elements at high energy scales. The comparison between the signal and background is made in Figure 7.3 using a 4 TeV W'_R , which is in the mass range targeted for exclusion in this analysis.

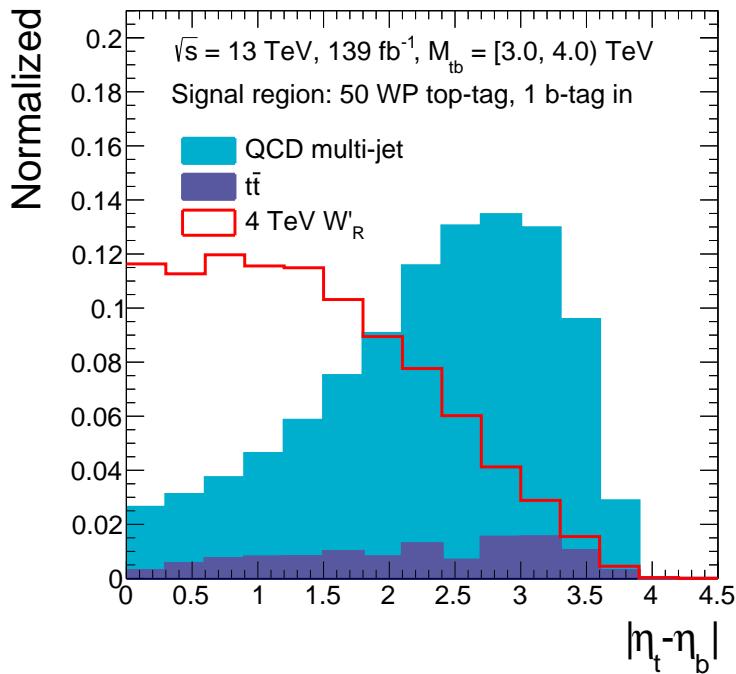


Figure 7.3: The absolute pseudorapidity difference between the top-candidate jet and the b-candidate jet for a 4 TeV right-handed W' signal and the two main backgrounds are compared in the signal region with the highest background rejection rate via the top-tagging and b -tagging (Signal Region 1 in Figure 7.2). Both the top-quark-pair ($t\bar{t}$) and QCD multi-jet backgrounds are simulated by MC generators (details in Chapter 8), and their sums are normalized to have unit area.

Chapter 8

Background Processes

This chapter describes methods to estimate the SM background in the signal regions defined in Chapter 7. One significant background is top quark pair production, whose final-state top quarks and subsequent bottom quark decays resemble the signal event. Precise cross-section calculations and event simulations for this background are introduced. The more dominant background from the QCD multi-jet events is estimated by a data-driven approach. Since this method utilizes the top-tagging and b -tagging ratios, the correlation between them is studied thoroughly.

8.1 $t\bar{t}$

Top quark pair production ($t\bar{t}$) is the most substantial top-quark contributor at the LHC [20]. The hard-scattering event's top quark and anti-top quark are produced by the strong interaction between two gluons or between a quark and its antiquark. Feynman diagrams associated with these two initial states at leading order of α_s are shown in Figure 8.1. Due to the vast gluon populations in protons, the gluon-initiated process leads to vast numbers of $t\bar{t}$ backgrounds in exotic particle searches with top quark final states.

Since the top quark decays must include a W boson, the $t\bar{t}$ event can be classified by the W -boson decay products into the all-hadronic, semileptonic, and di-leptonic channels. Among all $t\bar{t}$ events, the all-hadronic channel constitutes about 46% [84]. Even though the

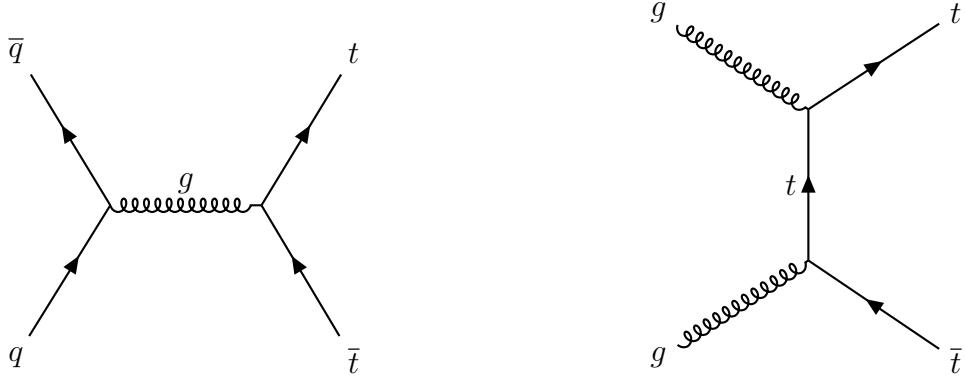


Figure 8.1: The Feynman diagrams of top quark pair production at the leading order of strong coupling. The vertical axis represents space, and the horizontal axis represents time. Feynman diagrams are made with TikZ-Feynman [64].

event selections veto events with top-tagged large-radius jets, the non-all-hadronic contribution is suppressed by the requirement of 0 leptons. Thus the semileptonic and di-leptonic channels will not be shown as separate channels in the following M_{tb} plots.

A significant amount of $t\bar{t}$ background is observed in the signal and template regions by comparing its simulated event numbers with data or the data-derived background. The most prominent regions are displayed in Figure 8.2, where the $t\bar{t}$ event's share is presented in the ratio plot. In the template region, the ratio of the $t\bar{t}$ to data stays around 10% in the low mass region and falls slowly to about 6% in the high mass region. In the signal region, the ratio starts as high as 25% and then stumbles rapidly to almost 6% in the last bin. Since the W' signal is negligible in the template region, the $t\bar{t}$ percentages in the two regions show that the b -tagging ratio is about three times higher for the $t\bar{t}$ background than for other backgrounds at low mass. However, such a boost diminishes as the M_{tb} ascends, reflecting worsening discriminatory power of the b -tagging with higher p_T .

The $t\bar{t}$ background has significant contributions from high-order Feynman diagrams. Ref. [113] found that $t\bar{t}$ plus one or more jets have a higher production cross-section than with no additional jets. In this analysis, the $t\bar{t}$ production cross-section is evaluated to the next-

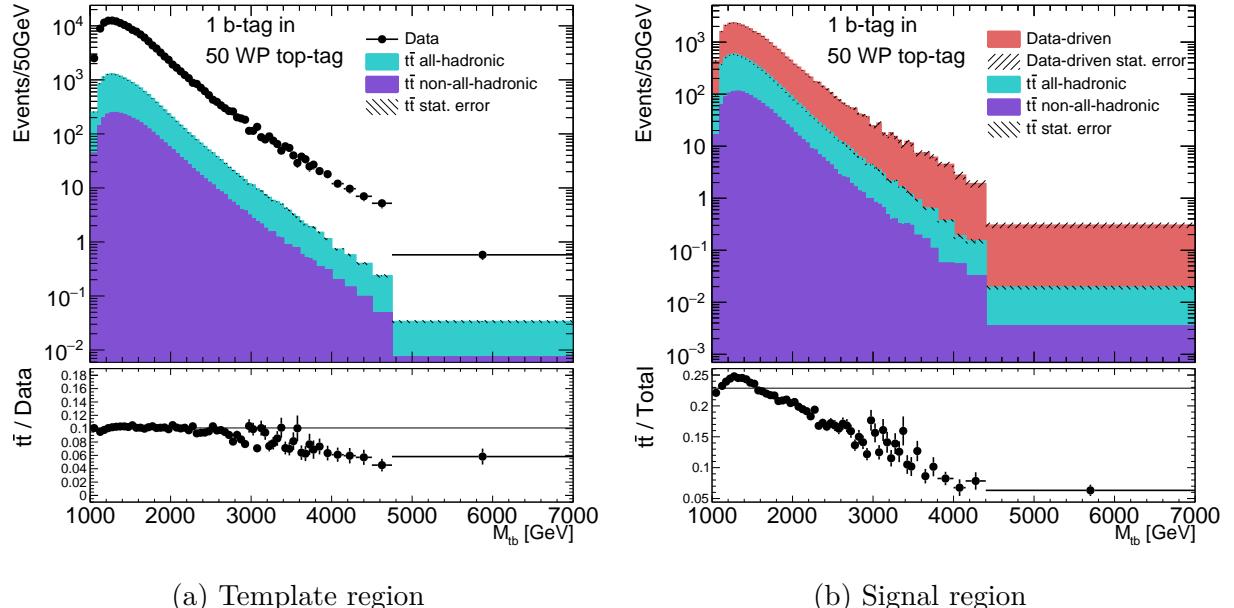


Figure 8.2: The invariant mass distribution of a large-radius top-candidate jet and a small-R b-candidate jet in region SR1 (right) and the neighboring region TR1 (left) (c.f. Figure 7.2). (a) The template region has the data plotted against two channels of $t\bar{t}$ background – the **all-hadronic** and the **non-all-hadronic** – shown as stacked histograms. Their statistical uncertainties are summed in quadrature into a single $t\bar{t}$ statistical error. The ratio of the summed $t\bar{t}$ distribution to the data is shown in the lower panel with statistical errors from both data and the $t\bar{t}$ background. (b) In the signal region, the data-driven background (see the next section) is drawn above the two $t\bar{t}$ channels with separate statistical errors. The ratio shown in the lower panel is the division of the $t\bar{t}$ to the total background. Covariances are taken into account in error propagations. On either figure’s ratio plot, there is a horizontal reference line showing the ratio of total events.

to-next-to-leading order (NNLO) of α_s with next-to-next-to-leading logarithm resummation using TOP++2.0 [114–120]. Theoretical uncertainties (see Chapter 5 for the means of computations) to this order are less than 4% [121].

The $t\bar{t}$ event simulation that enables jet reconstruction and tagging is generated with POWHEG-Box v2 [122–124] interfaced with PYTHIA to perform parton showering and hadronization. POWHEG calculates the next-to-leading order (NLO) quantum amplitude matched to the highest p_T showering parton to avoid double-counting the radiating diagram. It also uses Parton Distribution Functions derived at NLO – NNPDF3.0 [72]. At

this order of α_s , the scale uncertainty of the differential cross-section is below 10% in the top-quark pair invariant mass ($m_{t\bar{t}}$) from 300 GeV to 2 TeV [125], allowing for a precise prediction for the $t\bar{t}$ background in the M_{tb} spectrum.

8.2 Data-driven Backgrounds

Background events in the signal regions other than the $t\bar{t}$ are estimated using data in the template and control regions. After the subtraction of $t\bar{t}$ events, the rest of SM backgrounds are dominantly QCD multi-jet with no top quarks and few bottom/charm quarks in the hard-scatter final state. The data-driven estimation introduced in Chapter 3 is chosen to exploit the lack of correlation between the top-tagging and b -tagging probabilities among these gluon or light quark jets. Other less significant backgrounds like W -boson+charm or W -boson+top, which has a W mass-scale large-radius jet and a heavy flavor quark jet, can induce tagging correlations. Furthermore, the tb -decaying W -production, enhanced by the two taggers, distorts the double-ratio extrapolation. However, these correlating background's presence in data varies data-driven backgrounds within the range of statistical plus systematic uncertainties (detailed in Appendix B). Therefore, all backgrounds except $t\bar{t}$ are accounted for by the data-driven background estimates.

Table 3.1 gives a framework for applying the data-driven estimation defined in Figure 7.2. Since the top-tagging and b -tagging rates of background events vary with jets' kinematic properties, the extraction is performed bin-by-bin over the invariant mass distribution. Thus the data-driven background in the bin no. i in region SR1, $N(i, \text{SR1})$, is the data minus MC-simulated $t\bar{t}$ background of the same bin in the neighboring template region $(D(i, \text{TR1}) - M(i, \text{TR1}))$ times the b -tagging ratio computed in the “1 b -tag in” control

regions (CR1/CR2). The formula for all three signal regions and the validation region are summarized as

$$N(i, \text{SR1}) = C(i, \text{SR1}) \times \frac{(D(i, \text{TR1}) - M(i, \text{TR1}))(D(i, \text{CR1}) - M(i, \text{CR1}))}{D(i, \text{CR2}) - M(i, \text{CR2})}; \quad (8.1)$$

$$N(i, \text{SR2}) = C(i, \text{SR2}) \times \frac{(D(i, \text{TR2}) - M(i, \text{TR2}))(D(i, \text{CR1}) - M(i, \text{CR1}))}{D(i, \text{CR2}) - M(i, \text{CR2})}; \quad (8.2)$$

$$N(i, \text{SR3}) = C(i, \text{SR3}) \times \frac{(D(i, \text{TR3}) - M(i, \text{TR3}))(D(i, \text{CR3}) - M(i, \text{CR3}))}{D(i, \text{CR4}) - M(i, \text{CR4})}; \quad (8.3)$$

$$N(i, \text{VR}) = C(i, \text{VR}) \times \frac{(D(i, \text{TR4}) - M(i, \text{TR4}))(D(i, \text{CR3}) - M(i, \text{CR3}))}{D(i, \text{CR4}) - M(i, \text{CR4})}. \quad (8.4)$$

The correlation factor C characterizes the correlation between the top-tagging and b -tagging algorithms. $C = 1$ is the ideal case where no correlation is present, while a deviation from 1 implies the b -candidate jet's b -tagging rate is sensitive to the top-tagging. The dominant part of the data-driven background is composed of the QCD multi-jet backgrounds whose MC-simulated events are employed to estimate C .

The correlation factor for the SR1 is shown in Figure 8.3 using PYTHIA-generated QCD multi-jet events. The signal region's yield N on the left-hand side of Equation 8.1 is directly given and drawn as the black-dotted histogram on the upper panel. The blue histogram directly applies the data-driven method to the template and control regions' distribution, as the subtraction of $t\bar{t}$ from data is unnecessary. Thus the ratio of the two gives the correlation factor across M_{tb} . The null correlation, namely, the ratio of 1, is observed over the entire spectrum except for the first few bins. While solid evidence for $C = 1$ is also found in the other two signal regions and the validation region, a further investigation into potential sources of correlation provides systematics uncertainties of the data-driven method.

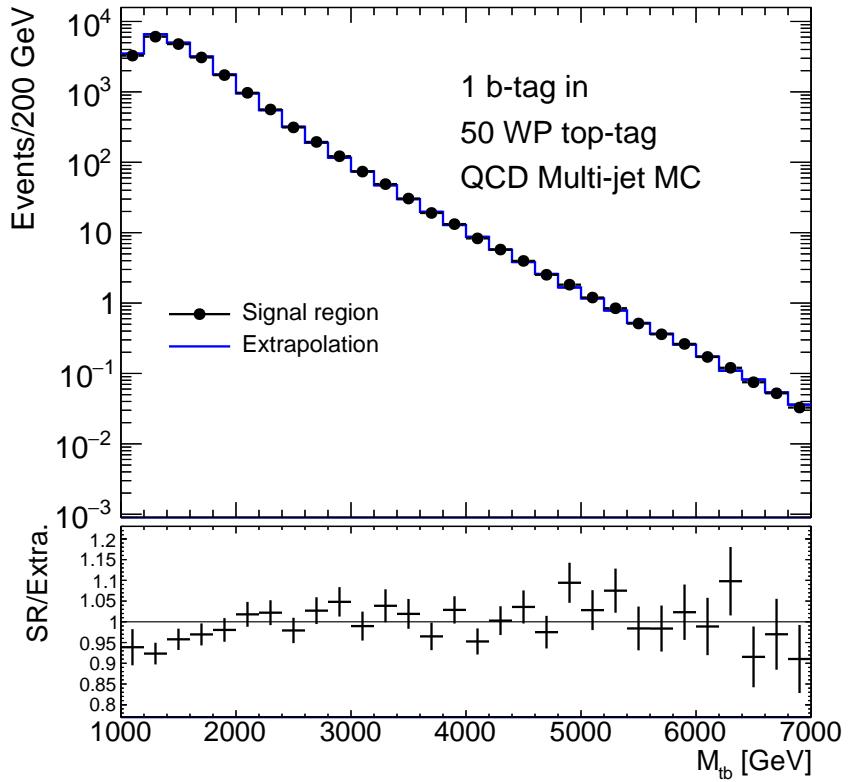


Figure 8.3: The invariant mass M_{tb} distribution of simulated QCD multi-jet background in the signal region SR1 (black dot) is compared with its extrapolation (blue lines) from the template region TR1 and the control regions CR1 and CR2. The vertical axis is the number of events for a bin width of 200 GeV. The lower panel shows the ratio of the signal region histogram to the extrapolated one.

8.2.1 Top-tagging and b-tagging correlations

The correlation between the top-tagging and b -tagging can be caused by the flavor of partons initiating top-candidate and b -candidate jets. At the leading order of QCD, the hard-scatter part of QCD multi-jet production is described by two-to-two Feynman diagrams, selections of which are shown in Figure 8.4. One final-state parton is mapped to the large-radius top-proxy or top-candidate jet by minimizing their ϕ angles, which is within 0.2 radion in more than 90% large-radius jet with $p_T > 500$ GeV. The other outgoing parton is then assigned to the small-radius b -candidate jet. Figure 8.5(a) illustrates the change of the large-radius jet's

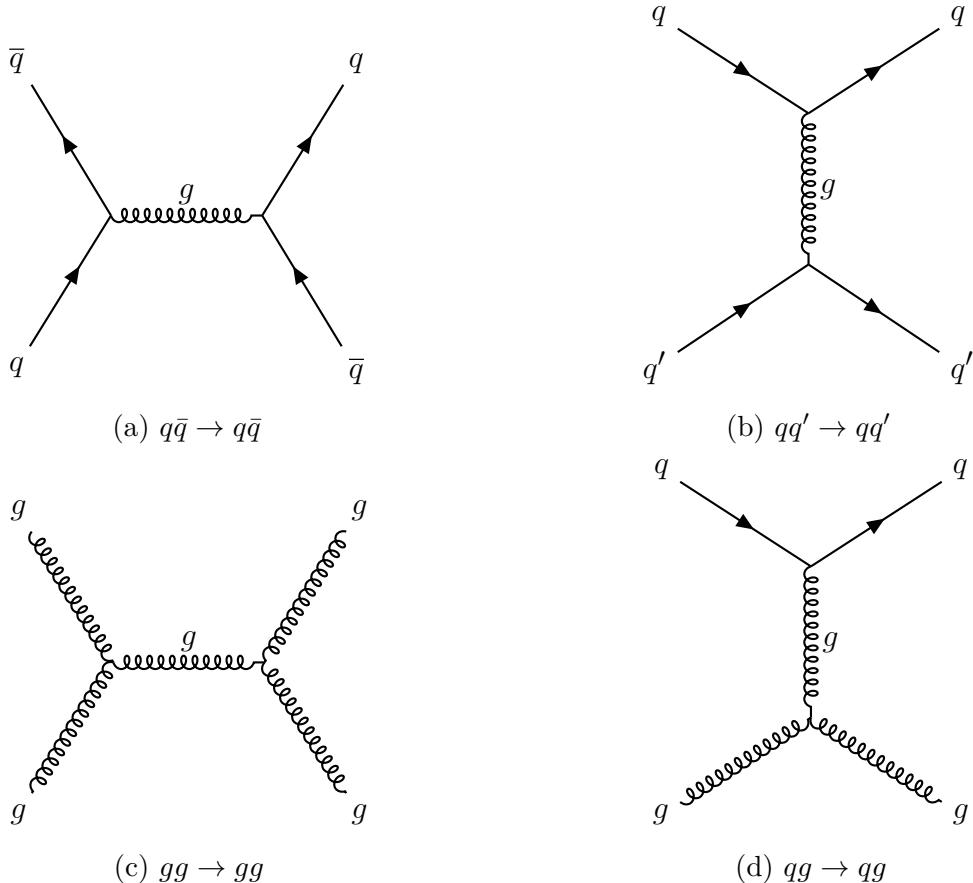


Figure 8.4: Examplified LO Feynman diagrams of the QCD multi-jet background at the LHC. The vertical axis represents space, and the horizontal axis represents time.

associated flavor by the top-tagging DNN score in the natural log scale for the “1 b-tag in” category and M_{tb} ranging from 1.2 to 1.4 TeV. The fraction of gluon-matched large-radius jets increases monotonically with the top-tagging score while suppressing the quark jets’ fractions. This trend is attributed to a gluon’s broader parton-showering [25] than quarks, giving gluon jet a higher probability to resemble the signal top-quark jet. The turning point near the last bin follows from the DNN’s abrupt cut-off at jet mass above 210 GeV, shunning more gluon to quark jets.

The parton types associated with the large-radius jet, when varied by the top-tagging algorithm, impact the b-candidate jet’s parton fraction and, in turn, its b -tagging ratios.

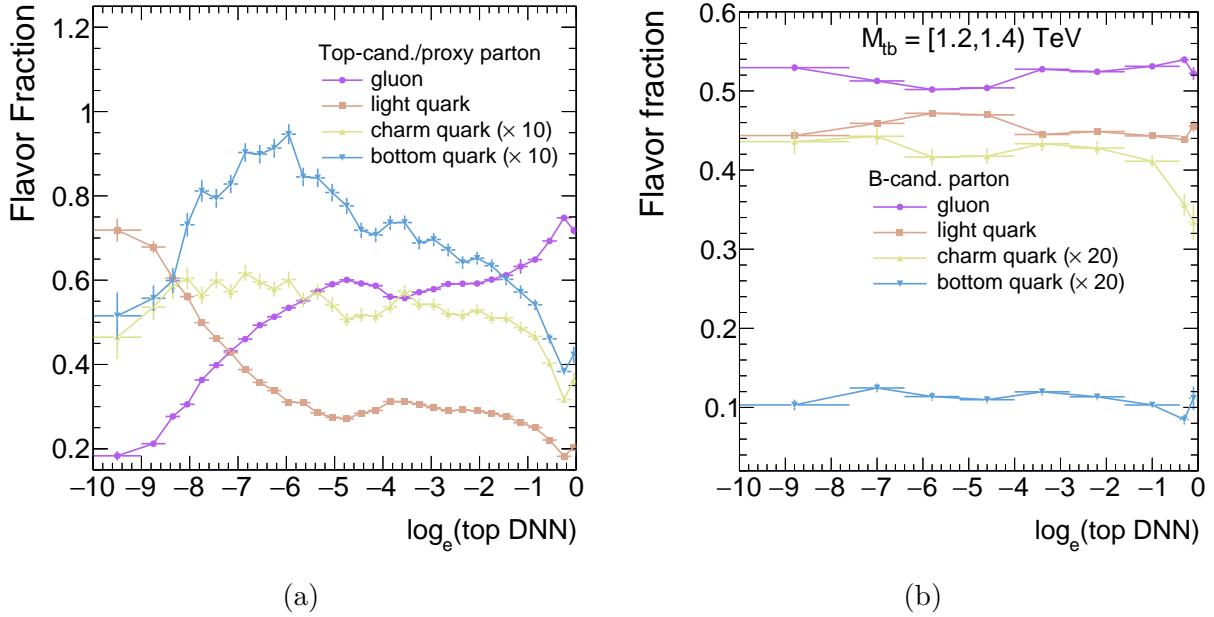


Figure 8.5: Figure (a) shows fractions of different parton flavors associated with the top-candidate/proxy jet as a function of the top-tagging score in nature logs. The top-tagging dependence of non- b -tagged jets are presented in figure (b). The last two bins' lower edges, corresponding to the 80% working point and the 50% working point, are not exactly the same as the label. The invariant mass M_{tb} is between 1.2 and 1.4 TeV for both figures. Light quarks include up, down, and strange quarks as their masses are below the QCD condensation scale. The charm and bottom quark fractions are multiplied by factors for viewing purposes.

Figure 8.5(b) plots the b-candidate jet's parton composition when the b-candidate jet is not b -tagged. While gluon-typed large-radius jets increase in proportion with the top-tagging, the b-candidate jet's light-quark fraction rises faster than its gluon fraction initially, thanks to the qg scattering as shown in Figure 8.4(d). For even higher DNN scores, di-gluon production as in Figure 8.4(c) takes over, leading to the gluon jet's upturn. Since the difference in b -tagging ratios between light-quark jets and gluon jets is around 20%, few percent changes of their portions cause per-mille correlations. In contrast, the fall in the charm-quark jet proportion at a sub-percent level near the top-tagged line (note the multiplying factor) can reduce the overall b -tagging ratio by 2% since the ratio of b -tagged to not- b -tagged charm-quark jets are 5 times higher than that for gluon jets and light-quark jets. Although the

bottom-quark jet’s b -tagging ratio is another 5 times higher, its impact here is less significant due to a relatively stable fraction.

We have seen that the different responses of hard-scatter partons to top-tagging and b -tagging generate correlations. An additional source of correlation comes from transverse momentum conservation which pairs a high p_T top-candidate/top-proxy jet with another high p_T b -candidate jet. Therefore the two taggers’ passing ratios increase coherently if both of them increase with p_T . The tagger’s momentum dependencies are related to poorer background rejections. When a top-quark large-radius jet receives a Lorentz boost, the top quark’s three-prong decay energy spreads across fewer calorimeter clusters, which are less distinguishable from a single-prong jet. In maintaining a similar tagging rate for top-quark jets as in lower p_T ranges, the background large-radius jet’s top-tagging ratios have to go up with p_T . The b -tagger has similar difficulties as more energetic charged particle tracks have less curvature under the inner detector’s (ID) magnetic field, challenging the particle vertex finding. Moreover, relativistic time dilation let high energy particle decay outside the ID’s innermost layer, leaving fewer hits for track reconstructions.

The p_T -induced correlation gets greater with higher M_{tb} as b -candidate jets are more likely to be initiated by light quarks only. A higher mass requires the initial-state partons in the hard-scattering process to carry a higher momentum fraction (x) relative to their parent protons. Since the parton density of light quarks surpasses that of gluons in the high x regime, relatively more light quarks should be produced through the process depicted in Figure 8.4(b). Indeed, Figure 8.6(a) shows that more than 68% of b -candidate jets originate from light quarks when M_{tb} is 3 to 3.4 TeV, thus leaving less room for the top-tagging DNN score to vary flavor compositions than in Figure 8.5(b). Meanwhile, as top-candidate and

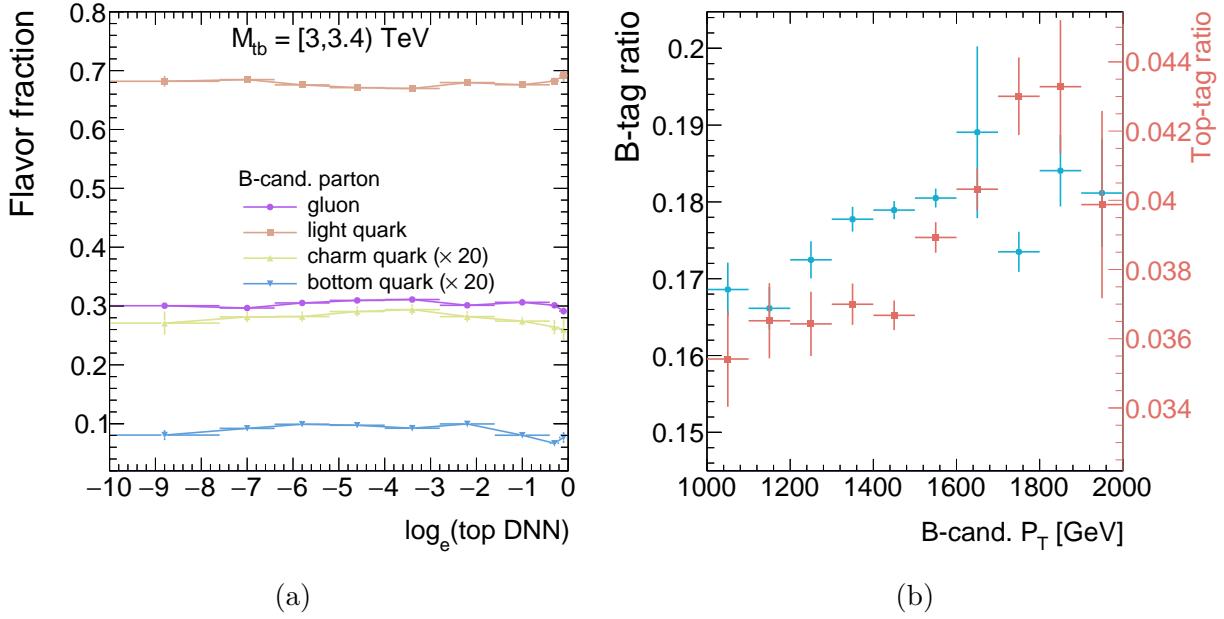


Figure 8.6: (a) The b-candidate jet’s parton flavor fraction as a function of the top-candidate/proxy jet’s top-tagging score’s natural log in the “1 b-tag in” category. The last two bins’ lower edges, corresponding to the 80% working point and the 50% working point, are not the same as the label. The invariant mass M_{tb} is between 3 TeV and 3.4 TeV. Non- b -tagged b-candidate jets are shown for reading the b -tag ratio variations. (b) The b -tagging ratio (blue with left vertical-axis) and top-tagging ratio (red with right vertical-axis) as a function of the b-candidate jet p_T in the same M_{tb} range and category. The top-tagging ratio is between the 50% WP top-tagged region and the upper control region (CRXa, X = numbers), as defined in Figure 8.7. The b -tagging ratio is also calculated in the control region.

b-candidate jets for greater M_{tb} have a broader range of p_T ¹ to generate positive correlation as tagging probabilities of both taggers increase with p_T , as shown in Figure 8.6(b). In conclusion, the correlation induced by jet flavors or p_T can introduce errors to the data-driven estimate at percent levels, consistent with the observation in Figure 8.3.

8.2.2 The systematic uncertainty

The motivation to study potential sources to the correlations between the top-tagging and b -tagging is to estimate systematic uncertainties for the data-driven background. Never-

¹But limited by the constraint of $|\eta_{tb}| < 2$

theless, the systematic uncertainties will be ultimately determined in data rather than in the simulated dataset employed in the preceding study since the latter has non-negligible unknowns, particularly in the high M_{tb} regime. For example, the parton flavor fractions of b-candidate jets are affected by the relative sizes of Parton Distribution Functions (PDFs) of different flavors. PDF uncertainties are substantial for high x , which overlaps with larger M_{tb} . Meanwhile, the top-tagging versus b -tagging correlations due to the p_T dependences of tagging gain uncertainties from insufficient data enriched with top quark jets or bottom quark jets with high p_T .

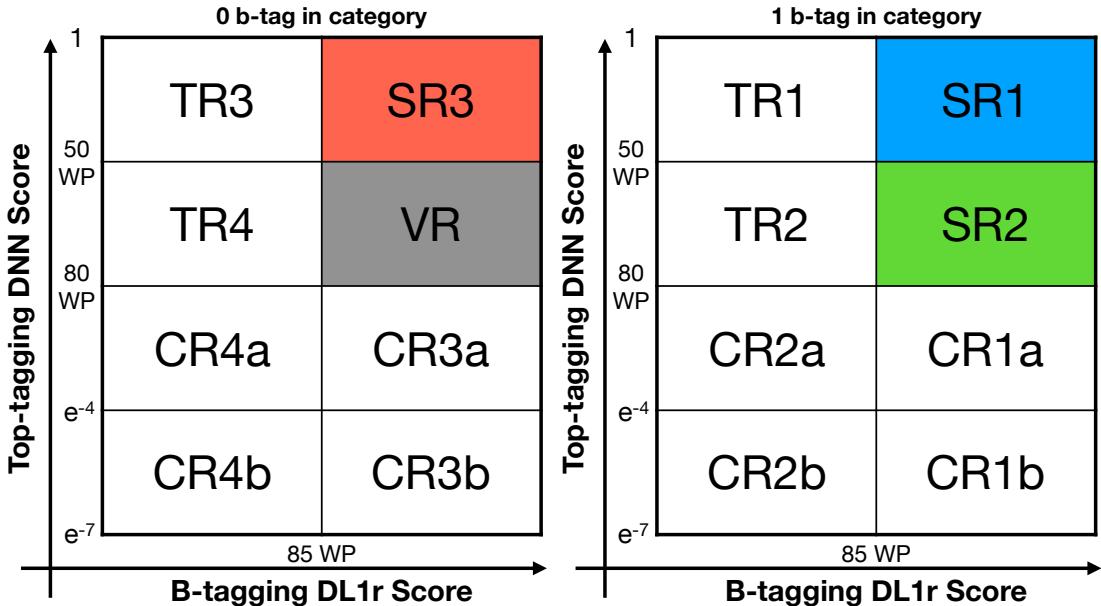


Figure 8.7: The categorization in the control regions (CRs) defined in Figure 7.2 is refined for computing the data-driven background's systematic uncertainties. The two new thresholds are placed on e^{-4} and e^{-7} of the top-tagging DNN score.

Alternatively, the systematic uncertainties of the data-driven background can be estimated using the correlation factor between subdivided control regions in data. The simulated samples are employed to add top-tagging boundaries reflecting the b -tagging variation. Figure 8.5(b) shows that the flavor composition of the b-candidate jet has two turning points

below the top-tagging line (the edge to the second rightest bin) – around e^{-7} and e^{-4} of the top-tagging DNN score. They form two bands with distinct flavor-induced correlation below the top-tagging threshold. Thus the control regions are redrawn as in Figure 8.7. Now, only the control regions closer in DNN scores to the signal regions enter Equation 8.1 through Equation 8.4. The correlation factor between the upper and lower CR are symmetrized and smoothed to serve as the systematic uncertainty. The smoothing procedure alleviate abrupt drops in correlation factors only in few M_{tb} bins. This uncertainty also incorporates correlation due to jet p_T and other smooth b -tagging variations due to the top-tagging scores.

This data-driven uncertainty is first cross-checked with the deviation between the signal region distribution and the data-driven method’s output using events of QCD multi-jet MC, as shown in Figure 8.8. The correlation factor in all regions are well covered by the systematic plus statistical uncertainties. Notably, the downshift in the low M_{tb} range of subfigure (d) (region SR1) is accounted for by the systematic uncertainties. Although the correlation factors from the control regions are larger than 1 (purple inverted triangle), in the opposite direction to the factor for the signal region (black dots). The systematic uncertainties do not outsize the extrapolation’s bin-by-bin statistical uncertainties too much in all signal plus validation regions. The systematic uncertainties remain below 5% in $M_{tb} < 5$ TeV, which is the highest W' mass of interest to set an exclusion limit on the production cross-section times branching ratio.

The estimated data-driven background and its systematic uncertainties are shown in Figure 8.9. The M_{tb} spectra in the control regions are merged more aggressively than in the template regions to reduce statistical fluctuation in estimating the b -tagged ratios and the correlation factors. The systematic uncertainties are close to the extrapolation’s statistical uncertainties in the “0 b-tag in” category, while the former outgrows the latter in the “1

“b-tag in” category. The systematic uncertainties estimated in data are noticeably smaller (greater) than the simulated-event counterpart (cyan lines in Figure 8.8) in low (high) mass regimes. This difference might be partially attributed to higher (lower) event yields of data than the MC simulation in these two ends. However, statistical errors have been controlled by using broader bin widths in calculating the correlation factors. Therefore, we suspect that the correlation is less due to flavor composition than the tagging ratio’s p_T dependence in data.

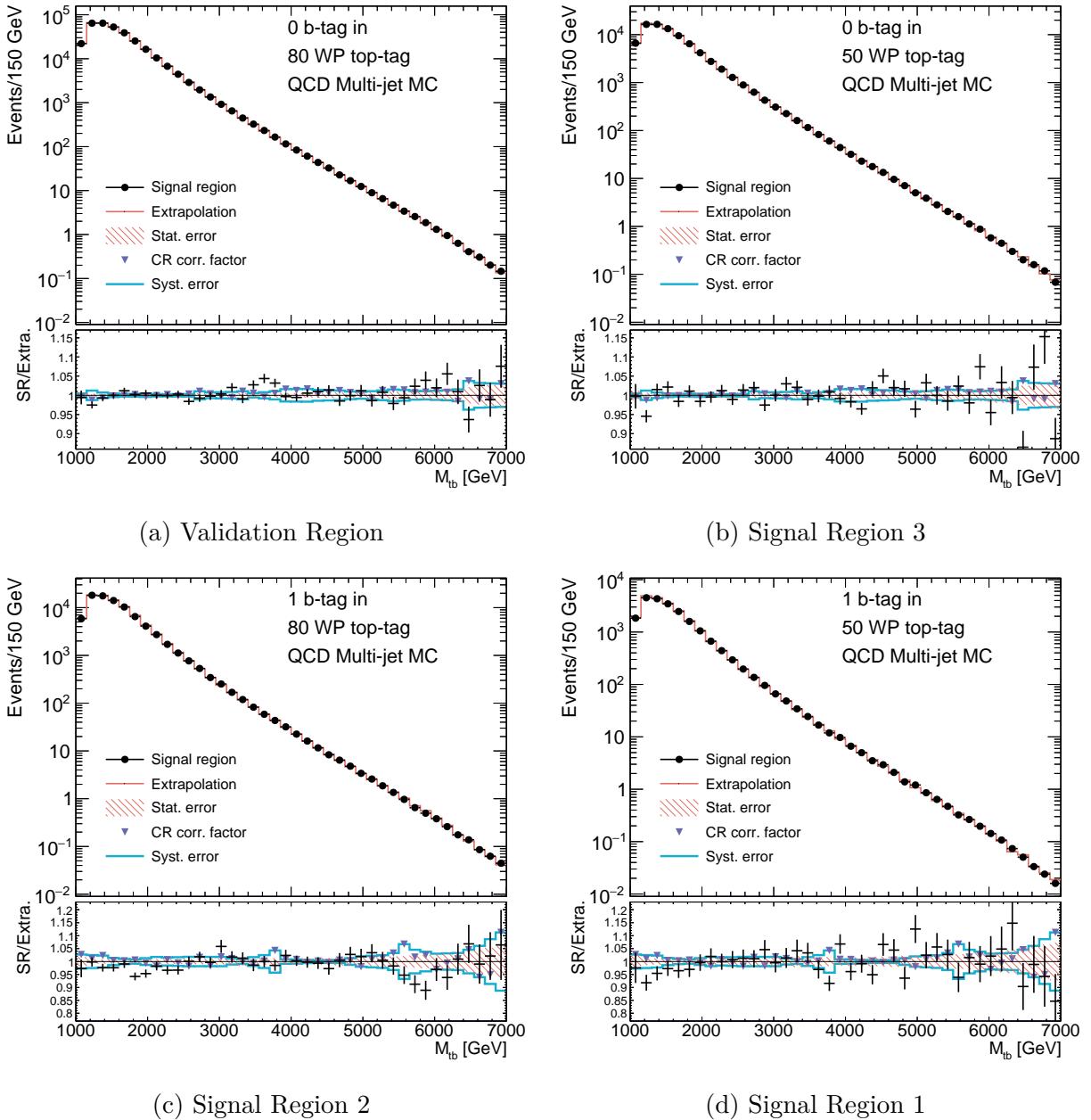


Figure 8.8: Validating plots for the data-driven background’s systematic uncertainties by the simulated QCD multi-jet background. In each diagram, the “SR” histogram is obtained directly from the top-tagging and b -tagging, and the extrapolation is the result of the data-driven method. The red hatches show the statistical uncertainty of the extrapolation. The error bars on the lower panel only include the statistical uncertainty for the SR distribution. The purple triangular dots show the correlation factor obtained in the control region of each “ b -tag in” category, yielding the cyan-colored systematic uncertainties after symmetrization (same as a reflection) and smoothing procedure.

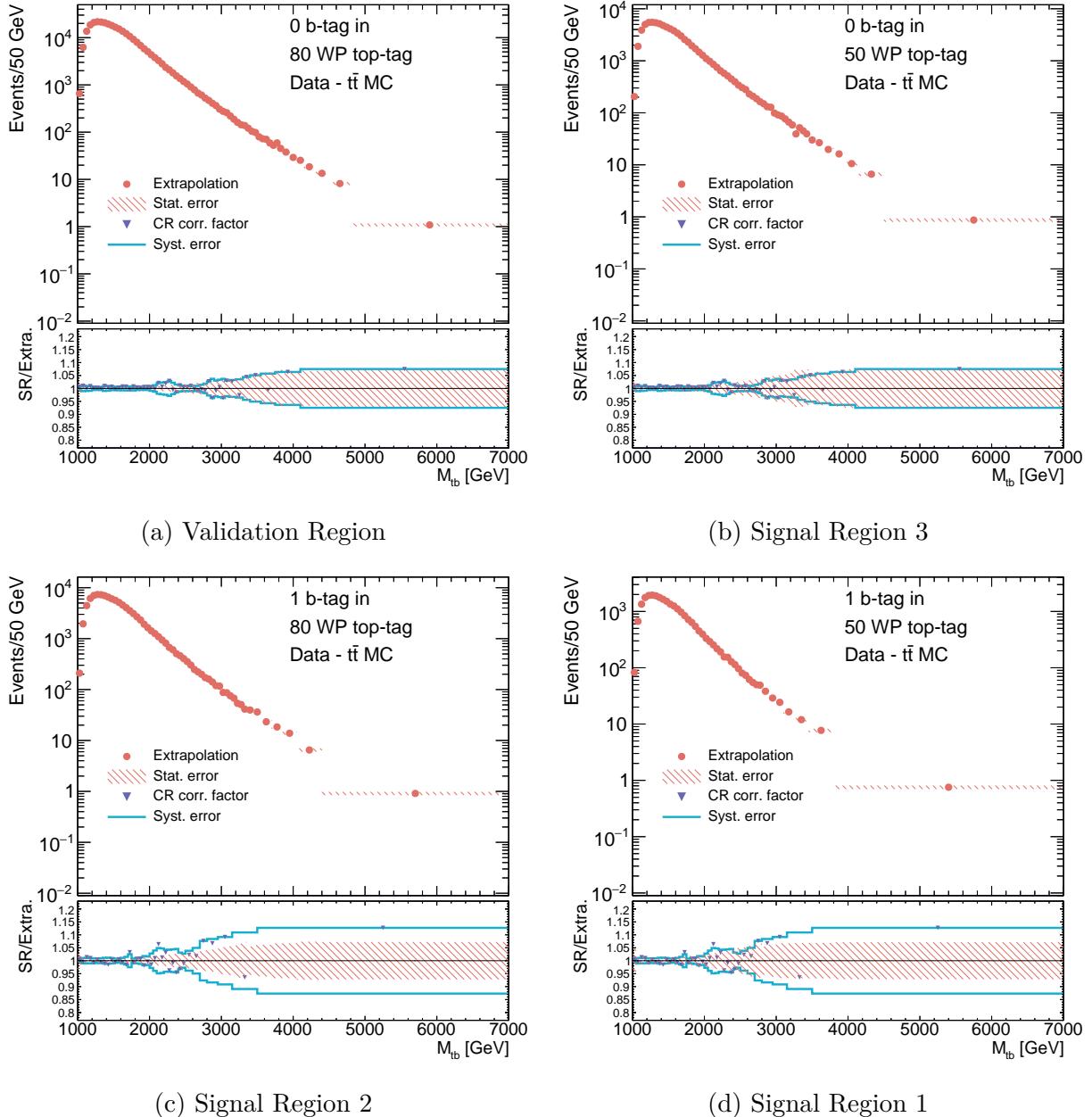


Figure 8.9: The data-driven background's systematic uncertainties. The extrapolation is the data-driven background derived from data subtracting the $t\bar{t}$'s simulated events. Statistical uncertainties are drawn separately for the extrapolation and the SR histogram as in the previous figure. The purple triangular dots show the correlation factor obtained in the control region of each “b-tag in” category, yielding the cyan-colored systematic uncertainties after symmetrization (same as a reflection) and smoothing procedure.

Chapter 9

Statistical Analysis

The Chapter 5 describes the calculation of the NLO cross-section and event generations by MC simulation for the W' signal. Top-quark and bottom-quark jets in data and simulated events are identified by the tagging algorithms introduced in Chapter 6 and selected into different categories and regions as explained in Chapter 7. In Chapter 8, the SM background contribution was estimated, thus completing to model all possible events in the signal regions. The reconstructed top-jet-plus-bottom-jet mass is analyzed in the signal regions to infer the presence of W' bosons in data.

The signal cross-section is quantified by the profile-likelihood method, which fits the signal and background models to data. The likelihood function is parameterized by the signal strength and nuisance parameters, characterizing the systematic uncertainties. The sources of these systematic uncertainties are reviewed in the first part of this chapter. Then, the formula for the profile-likelihood function is followed by the introduction to the CLs method, which is used to set limits on the W' cross-section times the branching ratio. A brief explanation of methods to study the fit qualities in terms of pulls and constraints is also given. A technique to rank the impact of nuisance parameters is summarized at the end of the chapter.

9.1 Systematic Uncertainties

Systematic uncertainty is a broad concept that can be abstracted as, but not limited to, potential errors in the instrument or assumptions in interpreting the measured quantities. Some systematic uncertainties are well-defined statistical uncertainties of physical parameters, such as the calibrated jet energy relative to the reference energy scales. However, in doing the calibration, some assumptions on the physics associating the two energy scales and the event selection necessary to obtain a pure sample could bias the result. Such uncertainties can only be estimated indirectly, such as by testing various physics models or varying the selection thresholds.

In this analysis, the systematic uncertainties are classified either as experimental or modeling. The experimental uncertainties arise mainly from calibrating the jet properties and correcting the top- and b -tagging rates in MC simulation. The modeling uncertainties include PDF and missing higher-order terms in calculating the signal or top-quark-pair background's total cross-sections and simulating jet kinematics. The uncertainty on the correlation between top-tagging and b -tagging ratios for the data-driven background estimation also belongs to the category of modeling uncertainties. But as it will be seen in the following, the calibration processes also make assumptions on the physical models and some modeling systematics involve experimental determination. Thus, the classification is not strictly defined.

9.1.1 Experimental uncertainties

One set of experimental uncertainties originate from calibrating the four-momenta of large-radius and small-radius jet in the simulated events (Chapter 6). Each nuisance parameter

of jet systematics varies each event’s jet energy or jet mass by one standard deviation and indirectly modifies the p_{T} -dependent top-tagging and b -tagging ratios. Generally, there are several systematic components in the category of modeling, calibration sample statistical errors, event selections, or detector effects. Combinations of components across all categories are found by diagonalizing the correlation matrix. This gives a set of uncorrelated systematic uncertainties. Only the combination with significant eigenvalues are considered in modelling the signal and the $t\bar{t}$ events.

Jet Energy Scale (JES) uncertainties shift the mean jet transverse momenta around the calibrated values determined in the *in situ* calibration of large-radius [54] and small-radius [55] jets. There are 19 JES variations exclusive to large-radius jets, 25 to small-radius jets, and 5 common to both jet types. Only two eigen-variations of the large-radius jet JES uncertainties and two common systematics – differences in detector responses to quark versus gluon jets and inter η calibrations – are non-negligible compared to background modeling systematics. The width of reconstructed jet energy’s detector effect underestimated in the MC simulation is accounted for by the **Jet Energy Resolution (JER)** uncertainties. There are 8 components for the small-radius jets and 12 for the large-radius jets, both of which insignificantly shift the M_{tb} of both the W' signal and $t\bar{t}$ backgrounds since the jet p_{T} in this analysis is much higher than the energy ranges of electronic noises, pile-up, and the sampling calorimeter’s statistical fluctuation.

Jet Mass Scale (JMS) corrects the energy distribution within jets after the overall jet energy has been calibrated, thus having independent systematic uncertainties of the JES ones. Large-radius jets [54, 126] have 18 components of JMS uncertainties, but all are too small compared to the jet energy scale to affect the invariant mass M_{tb} . Although energy spread uncertainties inside the large-radius jet can impact top-tagging rates, these effects are

negligible relative to MC modeling uncertainties. Like the JMS uncertainties, two **Jet Mass Resolution (JMR)** uncertainties are insignificant. Applying the above logic to small-radius jets with even a smaller clustering area to vary its jet substructures leads to no consideration of small-radius jet mass uncertainties.

The jet energy and mass uncertainties vary the number of jets satisfying the p_T requirement of event selections and the invariant mass spectrum for both the W' signal and $t\bar{t}$ background. In contrast, the top-tagging and b -tagging uncertainties modify the relative proportions of simulated events among different regions. Besides the propagated jet energy uncertainties, there are 18 top-tagging systematic uncertainties and 21 b -tagging uncertainties. Dominant uncertainties appear in both cases are parton shower modeling and p_T extrapolation uncertainties. The b -tagging extrapolation uncertainties, including the effect of B hadron interactions with the innermost tracking layer, increases with p_T to 20% in tagging probabilities at 2 TeV [127], becoming the most substantial uncertainties in this analysis.

Other experimental uncertainties have minor impacts, such as the luminosity uncertainties of around 1.7% for the $t\bar{t}$ background normalization [45, 128] and the variations in MC simulated pile-up events.

9.1.2 Signal modeling uncertainties

Although the mathematical formalism has no uncertainties, the computation of W' signal production cross-section has uncertainties in practice. As mentioned in section 5.2, the top-quark mass (m_t) and the strong coupling constant (α_s) are free parameters in the SM. The other issue is that the hard-scatter matrix element is evaluated order-by-order in the strong coupling constant. Higher-order terms involve exponentially more Feynman diagrams and

more complicated divergences to be regulated. The renormalization (μ_R) and factorization scales (μ_F) are varied to estimate the size of missing higher-order terms. To make things more challenging, the low energy end of QCD has only phenomenological models. Therefore, independent measurements (such as [129]) are employed to fit the non-perturbative Parton Distribution Function (PDF) describing the incoming parton momenta. Since the W' signal in this analysis is a benchmark model of a yet discovered vector boson, theoretical uncertainties – M_t , α_s , PDF, μ_R , μ_F – are only summed in quadrature for the production cross-section as a reference (Left-handed W' signal in Figure 5.3a and right-handed W' in Figure 5.3b). In other words, they are not included in the set of nuisance parameters, unlike the theoretical uncertainties of the background.

9.1.3 Background modeling uncertainties

While a similar combination is done for the $t\bar{t}$ background's production cross-section at the next-to-next-to-leading order, the resulting 4% error, which constraints the total number of top-quark pair events, constitutes a nuisance parameter in the profile-likelihood fit. Besides normalization, the $t\bar{t}$ background shape uncertainties include the PDF+ α_s and scale variation on the matrix element. The renormalization and factorization scale uncertainties are among the top systematic uncertainties in the entire M_{tb} . In contrast, the PDF uncertainties enlarge with the invariant mass and overtake the scale uncertainties at about 5 TeV.

Two more significant $t\bar{t}$ modeling uncertainties are given by replacing either the matrix element or the parton shower generator. The discrepancy between the default event generator – POWHEG [124] – and its alternative – MADGRAPH [130] – measures the effect brought by matching parton showers to MEs evaluated at the next-to-leading order in QCD [125]. The difference in invariant mass distributions between the two generators increases with M_{tb}

and becomes as large as the scale uncertainties in high mass ranges. The second uncertainty, the switch from the nominal showering program of PYTHIA to HERWIG [131, 132], results in similar shape variation as the ME in the low M_{tb} regime, while the effect subsides gradually. Other less significant shape uncertainties concern the initial state and final state radiations.

All $t\bar{t}$ background modeling uncertainties are decorrelated among the three signal regions. Namely, the same systematic source is characterized by three nuisance parameters in the profile-likelihood fit – one for each region. The modeling uncertainties are educated guesses on the size of missing parts or possible pitfalls in the model assumptions rather than well-defined parameters. The conservative approach allows different sizes of modeling among different regions.

The uncertainties of correlations between the top-tagging and the b -tagging (section 8.2) are modeling uncertainties for the data-driven background. The effect of this uncertainty on the total background is comparable to the $t\bar{t}$ modeling uncertainties as the data-driven background is more copious than the $t\bar{t}$ background, as shown in Figure 8.2. This modeling uncertainty is decorrelated among the three signal regions (and the validation region). The separation is physically motivated by changes in parton flavor compositions due to different top-tagging score ranges of and the application of b -tagging inside the large-radius jet. In each region, the uncertainty is further decorrelated between the high (> 2 TeV) and low (≤ 2 TeV) M_{tb} ranges. The motivation is that correlation in the low mass area is attributed mainly to the jet flavor composition, while the jet kinematics contributes more in the high mass area.

9.1.4 Propagation of uncertainties to the data-driven background

Since the $t\bar{t}$ distribution in the template region is subtracted from data in the data-driven background estimation, the experimental and modeling uncertainties of $t\bar{t}$ from the template region are propagated through the subtraction of the shifted $t\bar{t}$ distributions. The resulting experimental uncertainties on the data-driven background are correlated to those for the $t\bar{t}$ background and the W' signal, i.e., there is only one nuisance parameter per uncertainty. This combination, in turn, reduces experimental uncertainties since the upward variation on the $t\bar{t}$ background reduces the data-driven background in the template region, which is a downward variation. In contrast, the variation due to $t\bar{t}$ modeling uncertainties in each template region is decorrelated to the original $t\bar{t}$ modeling uncertainties in the neighboring signal region. Briefly speaking, the correlation between the physical parameter behind the modeling uncertainties, such as the missing higher-order terms attemptedly covered by the renormalization scale uncertainty, with the b -tagging of the b -candidate jet is unknown a priori. This conservative approach yields significant systematic uncertainties for the data-driven estimate.

9.2 Profile-likelihood Method

The statistical analysis tests the compatibility of data with a hypothesis of signal strength μ , a multiplicative factor to the signal normalization¹ The binned profile-likelihood function defined on signal region M_{tb} histograms is employed to incorporate systematic uncertainties. The number of data events in each bin i of region r ($\nu_{r,i}$) follows a Poisson distribution with

¹The signal production cross-section times the decay branching ratio ($\sigma(pp \rightarrow W' \rightarrow tb \rightarrow q\bar{q}'b\bar{b})$), multiplied by the integrated luminosity ($\int \mathcal{L}(t)dt$). This total number of signal events, times the fraction of events passing the event selection and categorization, gives the signal counts for each signal/validation/template/control regions.

a mean of $\mu s_{r,i} + b_{r,i}$, where $s_{r,i}$ and $b_{r,i}$ stand for the nominal signal and background counts. Furthermore, systematic uncertainties introduced in the last section are parameterized in the list of nuisance parameters $\vec{\theta}$. Each nuisance parameter θ_k , where k runs through the list of systematic uncertainties, is constrained by a Gaussian distribution where one standard deviation (σ_k) from the mean (θ_{k0}) moves the nominal signal or background by the amount corresponding to the systematic uncertainty.

The mathematical formula for the binned likelihood function is then written as

$$\mathcal{L}(\mu; \vec{\theta}) = \prod_{r,i} Pois(\nu_{r,i} | \mu s_{r,i}(\vec{\theta}) + b_{r,i}(\vec{\theta})) \prod_k Gauss(\theta_k | \theta_{k0}, \sigma_k). \quad (9.1)$$

The product $\prod_{r,i}$ runs over histogram bin i of signal region r . Note that $b_{r,i}$ sums the data-driven and $t\bar{t}$ backgrounds. The nuisance parameters as a whole are represented by a vector $\vec{\theta}$. By maximizing the function \mathcal{L} , the signal strength and nuisance parameter preferred by data are estimated by the maximum likelihood method. Moreover, a hypothesis with a fixed signal strength is tested against data by leaving μ fixed to μ_0 in maximizing \mathcal{L} .

The ratio of maximum likelihood between floating μ and fixed μ_0 forms the likelihood ratio to measure compatibility between data and the given hypothesis. By convention, the negative log-likelihood-ratio is defined as

$$q(\mu_0) = -2 \log_e \frac{\max\{\mathcal{L}(\mu; \vec{\theta}) : \mu = \mu_0, \theta_k \in \mathbb{R}\}}{\max\{\mathcal{L}(\mu; \vec{\theta}) : \mu \in (-\infty, \mu_0], \theta_k \in \mathbb{R}\}}. \quad (9.2)$$

The μ in the denominator only moves up to μ_0 since a higher estimated μ is considered consistent with the signal hypothesis. Instead, if μ moves far below μ_0 , the nuisance parameter can no longer draw close the data and prediction, leading to a higher, a more extreme value

of $q(\mu_0)$. The probability of a given signal hypothesis to have at least or more extreme $q(\mu_0)$ is the *p* – *value*. The asymptotic formula of the $q(\mu_0)$ distribution calculated in [133] is then employed to exclude signal strengths with *p* – *values* below 0.05.

9.3 The CLs method

Sometimes, the *p* – *value* could pass the exclusion threshold, even though neither the signal plus background nor the background alone describes the data well. A straightforward example is underestimating the background events in a counting experiment. A countermeasure is the *CLs* method [134] which requires the rejection of the signal plus background hypothesis using a different *p* – *value* definition:

$$p_{s+b} = \frac{p_s}{1 - p_b} < 0.05, \quad (9.3)$$

where p_s denotes the *p* – *value* for the signal-plus-background hypothesis ($\mu > 0$) and p_b for the background-only hypothesis ($\mu = 0$). The denominator $1 - p_b$ measures the sensitivity of the statistical test [135].

The same reference and [133] point out a way by which we use to calculate the expected limit that shows the signal exclusion sensitivity under the background-only hypothesis. Pseudo-data are created by setting all nuisance parameters to their best-fit values given by data and then reset the signal strength to 0². This dataset gives the median $q(\mu_0)$ under the background-only hypothesis while testing a signal-plus-background hypothesis of signal strength of μ_0 . The effect of statistical and systematic errors on the expected limit is

²This is the expected background-only model when data have been unblinded. In the optimization phase of the analysis strategy, the nuisance parameters are all fixed to their nominal values.

computed using the standard deviation of the expected $q(\mu_0)$ distribution. The resulting expected limit and uncertainty bands are compared with the observed limit given by data to evaluate performances.

9.4 Fit validations with pulls and constraints

The quality of fit is crucial to ensure the hypothesis test is sensible. Indicators of fit qualities include *pull* and *constraint*. Pull is defined as the difference in nuisance parameter between the best-fit and nominal value, divided by the pre-fit uncertainty ($(\theta_{\text{fit}} - \theta_{\text{nominal}}) / \sigma_\theta$). A significant pull requires further inspections if the corresponding systematic uncertainty is expected to be inconsequential. A constraint means a significant decrease in the standard deviation of a nuisance parameter after fitting. Substantial shrinking of post-fit errors indicates that the systematic uncertainties are comparable to or larger than the standard deviation of the Poisson terms in Equation 9.1.

Since a systematic uncertainty usually affects more than one region in the fit, one region with the least statistical error could constrain the associated nuisance parameter of another region with a larger error. For example, the “80 WP top-tag, 1 b-tag in” region has the tiniest statistical uncertainty among all signal regions. Despite having fewer expected signal yields than the “50 WP top-tag, 1 b-tag in” region, the fit allows the former to constrain substantial systematic uncertainty like the one for *b*-tagging extrapolation, thereby enhancing the search sensitivity. But the mutual constraint across the two regions does not apply to the top-tagging systematics, as the calibration of the tagger’s 80% working point is not conducted simultaneously with the 50% working point, leaving an undetermined correlation between the nuisance parameters of the same source for the two working points. Thus they

are left uncorrelated, i.e., remain independent parameters in the likelihood function and let the fit calculate their covariance.

9.5 Nuisance parameter ranking

The statistical significance of data under the hypothesis with a signal strength μ is measured by the negative log-likelihood ratio (Equation 9.2), which is the ratio between the minimized profile-likelihoods with a fixed μ and that with an unconstrained μ determined by the fit ($\hat{\mu}$). Therefore, the degree that $\hat{\mu}$ is shifted by the fit in response to the variation of nuisance parameters can tell the relative impacts of the parameters on the minimization of the negative log-likelihood function and in turn, the limit setting.

More precisely, the global best-fit signal strength $\hat{\mu}$ and the simultaneously best-fit $\vec{\theta}$ of N nuisance parameters $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_N$ are given by minimizing the negative log-likelihood

$$-\log \mathcal{L}(\hat{\mu}, \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_N) = \min\{-\log \mathcal{L}(\mu; \vec{\theta}) : \mu, \theta_1, \theta_2, \dots, \theta_N \in \mathbb{R}\}. \quad (9.4)$$

The symbol $\hat{\mu}$ is the maximum likelihood estimator of μ commonplace in literature, such as the Statistics chapter in [20]. In these contexts, the cap notation already implies that it is a variable that (co-)minimizes the negative log-likelihood (a computational trick equivalent to maximizing the likelihood function) for given data. But for more easily denoting the change of μ by fixing one nuisance parameter, here $\hat{\mu}$ is reserved for the μ given by the minimizing operation on the right-hand side of Equation 9.4.

Re-minimizing the negative log-likelihoods by fixing a nuisance parameter to the global best-fit value $\hat{\theta}_i$ plus-or-minus the uncertainty σ_i requires μ to be shifted to $\mu \pm \Delta\mu$, as in

the following formula:

$$-\log \mathcal{L}(\hat{\mu} \pm \Delta\mu; \dots) = \min\{-\log \mathcal{L}(\mu; \vec{\theta}) : \mu, \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_N \in \mathbb{R}; \theta_i = \hat{\theta}_i \pm \sigma_i\}. \quad (9.5)$$

In Chapter 10, nuisance parameters will be ranked by $\Delta\mu$ using the post-fit uncertainty estimated from the global fit represented by Equation 9.4.

Chapter 10

Statistical Analysis Results

This chapter employs the statistical procedures introduced in Chapter 9 to study the compatibility between data and prediction and infer the possible existence of the W' signal. It starts with verifying the robustness of the statistical analysis procedures. The preciseness of the data-driven background estimation is first studied in the signal-insensitive validation region. For the signal regions, the background-only model is used as pseudo-data to provide expected constraints, correlations, and blinded expected limits (definitions can be found in Chapter 9). The background-only fit is repeated but with data instead. The pull plots of the nuisance parameters and correlations are shown for the observed data too. The lack of data excess relative to the background is found, and hence the exclusion limit is set. In the end of this chapter, we analyze the psuedo-data and data with a unconstrained fit, which yields a list of the most impactful systematic uncertainties.

10.1 Validating the Data-driven Background Estimation

The dominant background in the signal regions is estimated by a data-driven method, which is verified by comparing the total background predictions to data in the validation region where the W' signal is negligible (VR shown in Figure 7.2). The top-tagging requirement for

this region is the same as the “SR2” signal region, which has less background rejection than the tagger threshold applied to the other two signal regions. This less stringent top-tagging criterion and the “0 b-tag in” categorization for this region render the number of signal events insignificant relative to the predicted backgrounds even around the W' resonance peak in M_{tb} .

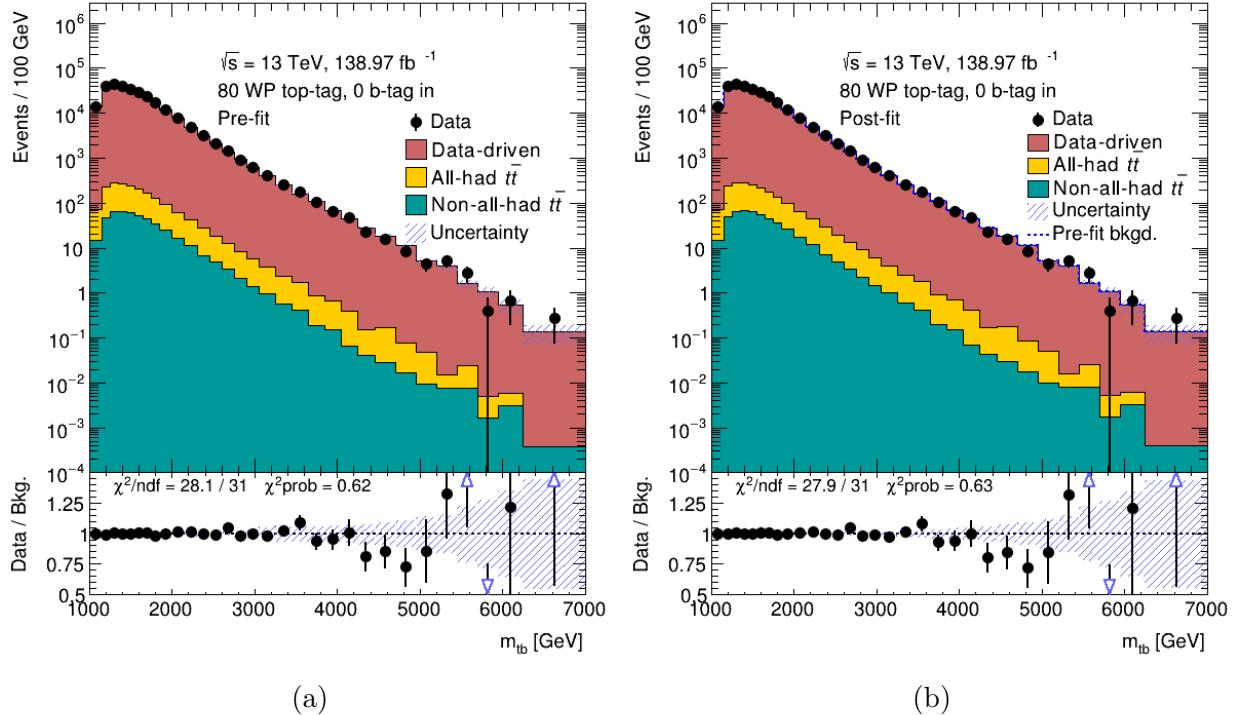


Figure 10.1: Data and predicted background M_{tb} distribution in the validation region (top-tagged at the 80% but below the 50% working point and in the “0 b-tag in” category) (a) before and (b) after the profile-likelihood fit. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.

The comparison between data and the predicted background distribution in the validation region is presented in Figure 10.1. As the data-driven background is two orders of magnitude larger than the $t\bar{t}$ background in the whole mass range, the systematic uncertainties of the $t\bar{t}$ have an insignificant effect on validating the data-driven estimate. The relative difference between the data and background distributions stays within 5% from 1 TeV to 3 TeV of M_{tb} .

and below 10% up to 4 TeV in Figure 10.1(a), where no profile-likelihood fit is performed. No significant discrepancies between the two compared to the statistical plus systematic uncertainties are found over the entire mass range. A pre-fit χ^2 of 28.7 per 31 degrees of freedom supports this observation. A profile-likelihood fit of the background model to data yields the distributions in Figure 10.1(b), shifting the background prediction by a negligible amount from the pre-fit. The χ^2 value decreases only by 0.2 to 27.9 after this background-only fit. The small shift implies that the pre-fit background model parameters are near the best-fit values favored by data.

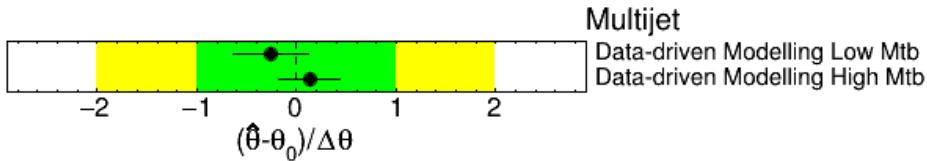


Figure 10.2: The pull and constraint plots of the data-driven-background systematic uncertainties in the background-model fit to data in the validation region. The central value (θ_0) is the pre-fit value, which is 0 by construction. For each nuisance parameter, the black dot shows the post-fit value ($\hat{\theta}$) with the post-fit uncertainty as its error bar, which is in general larger than the pre-fit uncertainty of unit size ($\Delta\theta$).

The nuisance parameter pulls associated with the systematic uncertainties of data-driven estimation are shown in Figure 10.2. Two nuisance parameters are assigned to the low ($M_{tb} < 2$ TeV) and high ($M_{tb} \geq 2$ TeV) mass regimes each. Recall from Chapter 8 that this systematic uncertainty is due to the correlation between top-tagging and b -tagging. Thus, the two nuisance parameters' consistency with their pre-fit value implies that data are compatible with a lack of correlation between the two taggers. The significant constraints are expected as the validation region has sufficient data to constrain systematic errors at a percent level. All the other systematic uncertainties are not shown as they have insignificant pulls and constraints. The definition of pulls and constraints can be found in Chapter 9.

10.2 Test the Signal Region Fit with the Background-only Pseudo-data

Before comparing data and the background predictions in the signal regions, we performed the profile-likelihood fit to the M_{tb} distributions by treating the background predictions as pseudo-data. This dataset is expected to attain maximum likelihood without shifting the nuisance parameters, thus allowing us to study the constraints and correlations between systematic uncertainties. This method helps to find ill-behaved nuisance parameters before conducting the fit to data. The signal region M_{tb} distributions before the profile-likelihood fit with the “ $\mu = 0$ ” pseudo-data are shown in Figure 10.3.

The M_{tb} distributions after the profile-likelihood fit under the constraint of $\mu = 0$ are shown in Figure 10.4. There is no change in the heights at each bin since the pseudo-data is identical to the total backgrounds. The uncertainty bands shrink substantially in the low mass region as the fit constrained the 1σ deviation of the nuisance parameter representing the systematic uncertainties. The pre-fit systematic uncertainty can be thought of as the width of a physical parameter from the complementary measurement. In this sense, the comparison of data and background predictions serves as an independent measurement such that the combined uncertainty, the post-fit value, is naturally less than the pre-fit one. The statistical uncertainties of the background distribution, which dominate the uncertainty band in the high mass region, are less constrained, for the data-driven estimation has considerably lower statistical errors than data by adopting wide bin widths (Figure 8.9). Moreover, the statistical uncertainties are independent between bins, further loosen the constraint.

The pull plots after the fit are shown in Figure 10.5 to Figure 10.13. Most of the nuisance parameters are not constrained under the fit. Notable exceptions are “Top-tag

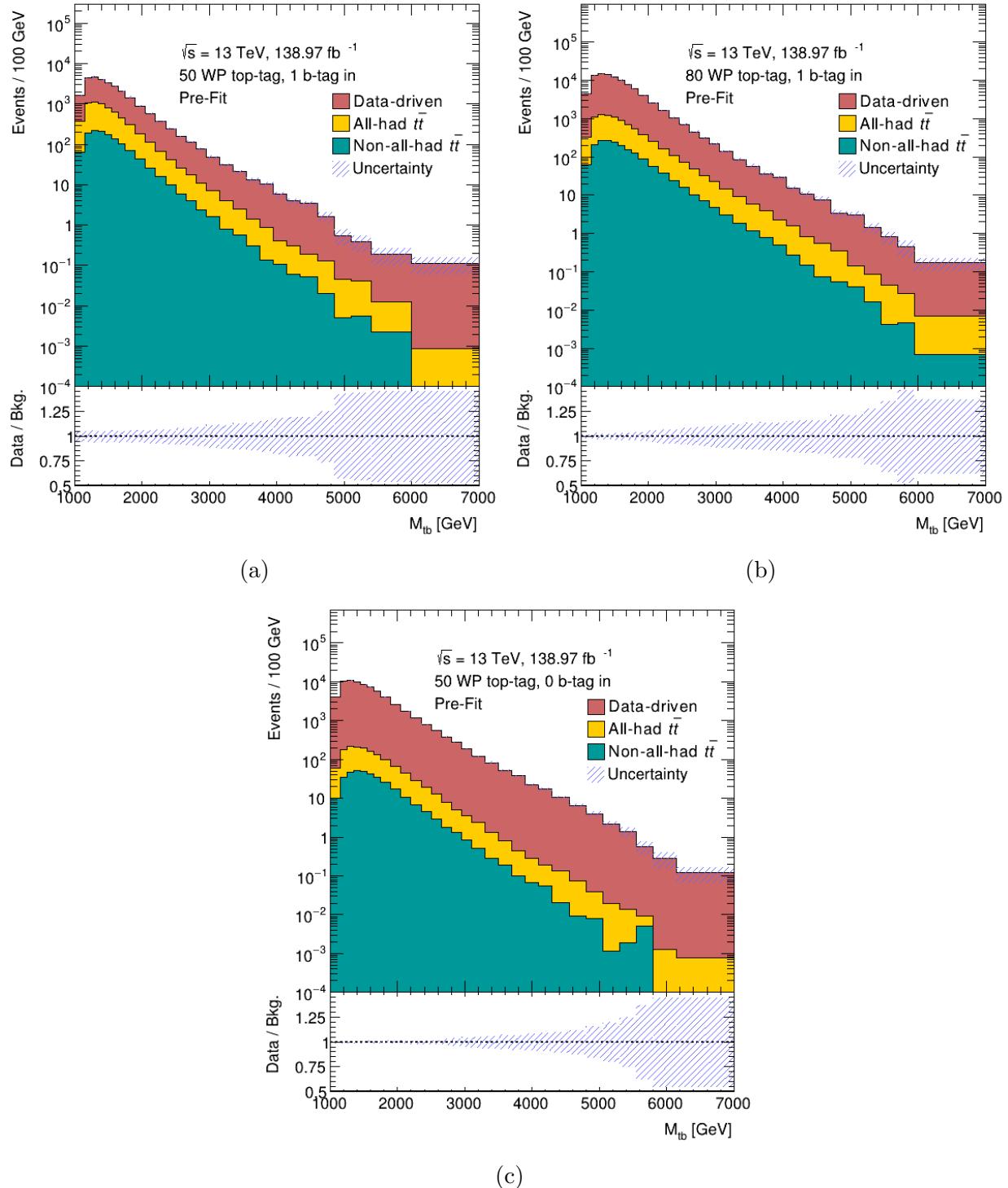


Figure 10.3: Predicted background M_{tb} distribution in the signal regions: (a) SR1, (b) SR2, and (c) SR3 before the profile-likelihood fit. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.

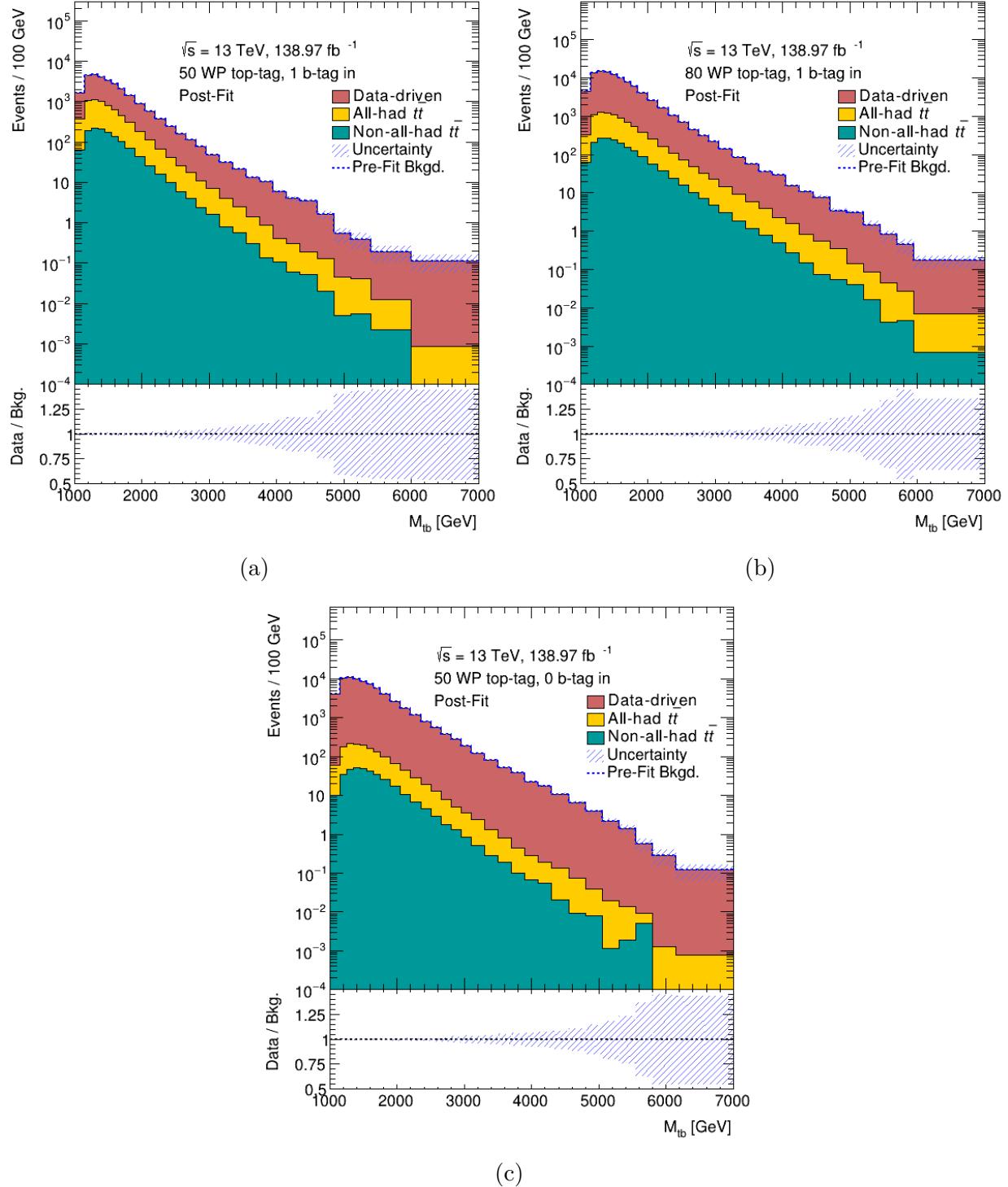


Figure 10.4: Predicted background M_{tb} distribution in the signal regions: (a) SR1, (b) SR2, and (c) SR3 after the profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.

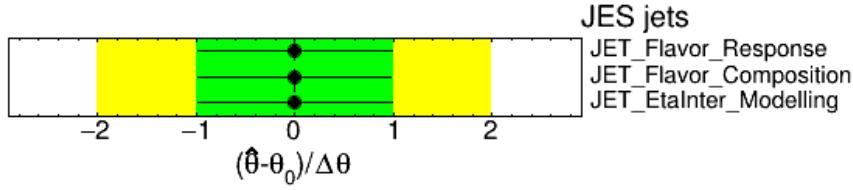


Figure 10.5: The nuisance parameter pulls for jet energy scale uncertainties common to large-radius jets and small-radius jets after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” “EtaInter” means the *in situ* calibrations across different η regions of the detector.

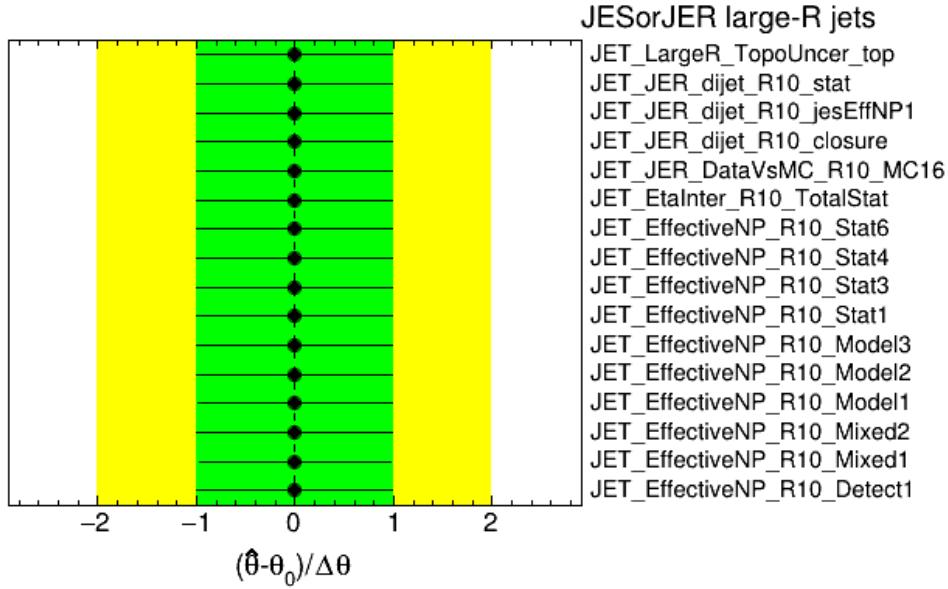


Figure 10.6: The nuisance parameter pulls for large-radius jet energy scale and resolution uncertainties after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” “TopoUncer” refers to jet energy modeling uncertainty for top-quark jets. “JER” denotes jet energy resolution. “jetEffNP1” is the first component in the principle component decomposition for propagating the jet energy scale uncertainty. Each “EffectiveNP” is an eigenvector in the principle component analysis. The suffices, “Stat,” “Model,” “Detect,” and “Mixed” means that the eigenvector combines statistical, modeling, detector, or a mixture of systematic error sources.

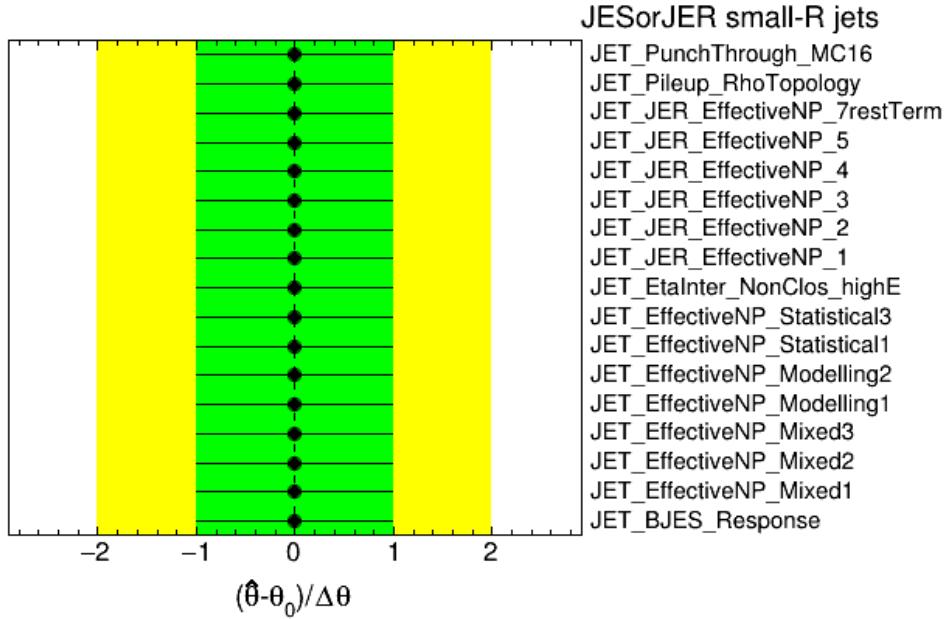


Figure 10.7: The nuisance parameter pulls for small-radius jet energy scale and resolution uncertainties after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” “PunchThrough” means the hadronic shower of a jet is too deep to be contained by the ATLAS calorimeter. “Rho” is the energy profile of an event used to correct for pile-up contributions in the calorimeter measurements. “restTerm” is a quadratic sum of the most insignificant JER components. “NonClos” and “highE” refer to non-closure and high-energy, respectively. “BJES” means the jet energy scale for a bottom-quark jet. Some other terms are defined in the captions of Figure 10.5 and Figure 10.6.

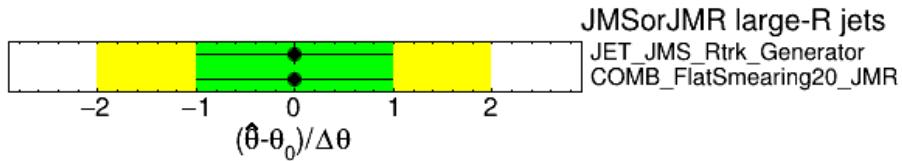


Figure 10.8: The nuisance parameter pulls for large-radius jet mass scale and resolution uncertainties after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” “Rtrk” is the ratio between jet mass estimated by tracks to that measured by calorimeters. The difference in Rtrk between simulation and data serves as a systematic uncertainty. “COMB” refers to “combined mass,” a mass observable combining the tracking and calorimeter measurements.

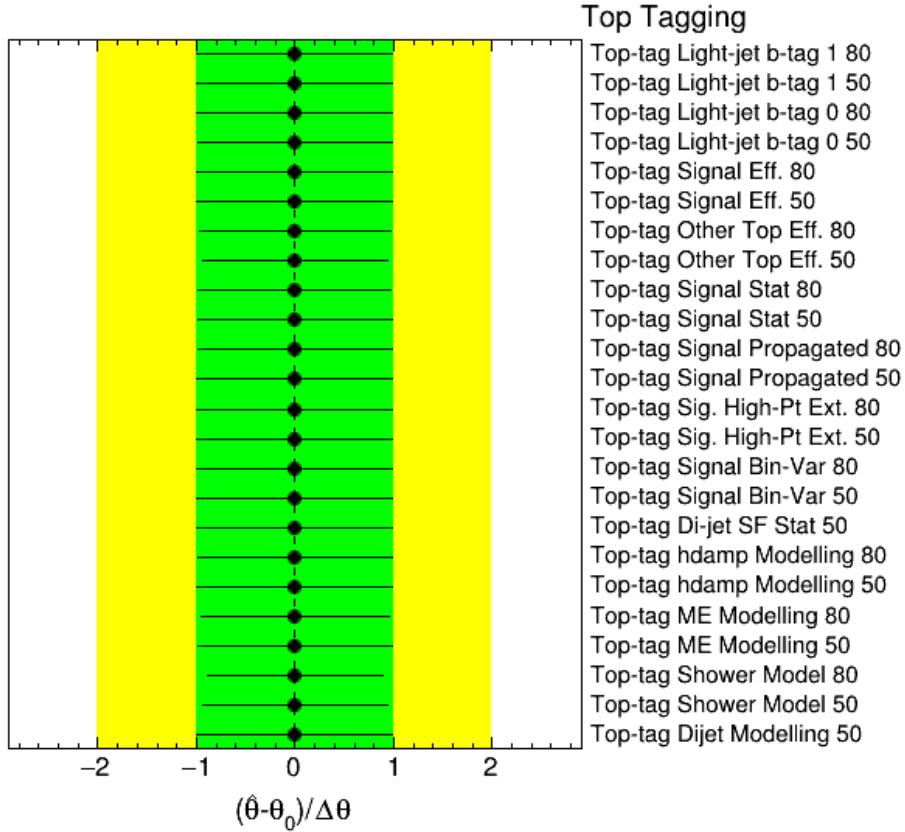


Figure 10.9: The nuisance parameter pulls for the top-tagging uncertainties after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The number 80 and 50 mean the large-radius jet is top-tagged at the 80% working point but not the 50% working point or top-tagged at the 50% working point, respectively. “Light-jet b-tag” are b -tagging nuisance parameters for the light-quark or gluon jet since the top-tagging calibration selects events with b -tagged jets. “Signal Eff.” accounts for the chirality difference between the $t\bar{t}$ events with unpolarized top-quarks and the signal events with polarized top-quarks. “Other Top Eff.” are uncertainty in top-tagging efficiency for top-quark jets with non-optimal reconstructions such as missing top-quark-decay products within a truth-level jet. “Signal Propagated” is a mixture of insignificant jet energy scale uncertainty components than those propagated via the JES uncertainties for well-reconstructed top-quark jets. “High-Pt Ext.” accounts for uncertainties in high transverse momentum beyond the range calibrated with respect to data. “Bin-Var” concerns the p_T -bin choices in fitting the efficiency curve. “Di-jet” means a particular type of QCD multi-jet event with two high p_T jets. “SF” stands for Scale Factor, which scales the tagging probability in simulated events to the calibrated value. “hdamp” is the matching scale for the NNLO event generator, POWHEG [125]. “ME” refers to the matrix element. “Shower” is the parton shower. Alternative event generators of ME and showers are discussed in Chapter 9.

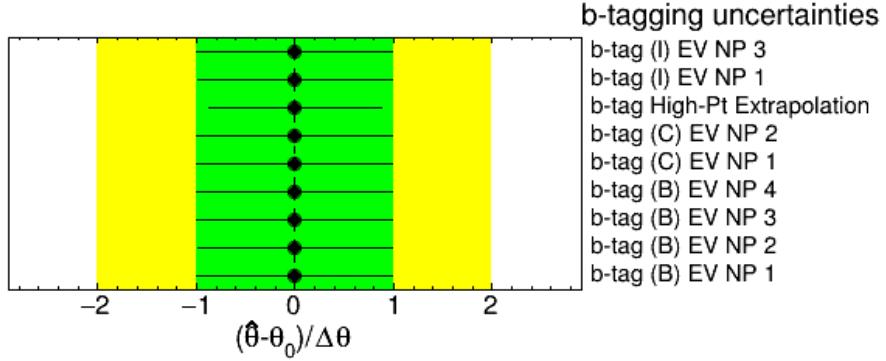


Figure 10.10: The nuisance parameter pulls for the b -tagging uncertainties after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The name “(I) EV NP 3” means the third eigenvector of nuisance parameters for calibrating the b -tagging probability for a light-hadron jet (no heavy hadrons that contains a charm or bottom quark). The same convention applies to a charm-hadron jet by replacing the capitalized letter “I” with “C” and to bottom-hadron jet by replacing it with “B.” “High-Pt Extrapolation” refers to the uncertainties caused by insufficient data to calibrate b -tagging probabilities for small-radius jets with high transverse momenta (for more descriptions see Chapter 9 and Ref. [127]).

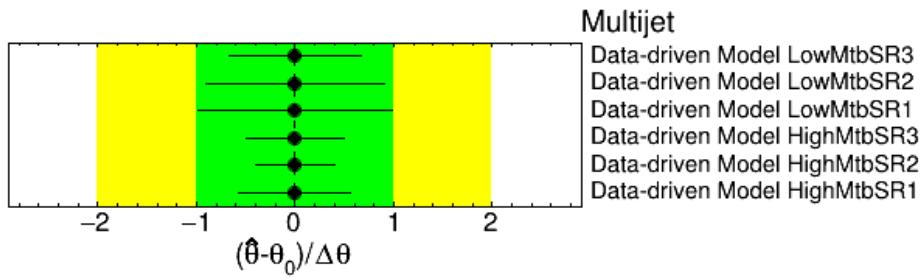


Figure 10.11: The nuisance parameter pulls for the modeling uncertainties of the data-driven background estimation after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The symbols “LowMtb” and “HighMtb” mean that the nuisance parameter varies the data-driven background in invariant mass M_{tb} below or above 2 TeV. The region code (Figure 7.2) at the end further specifies the signal region whose data-driven background is to be varied.

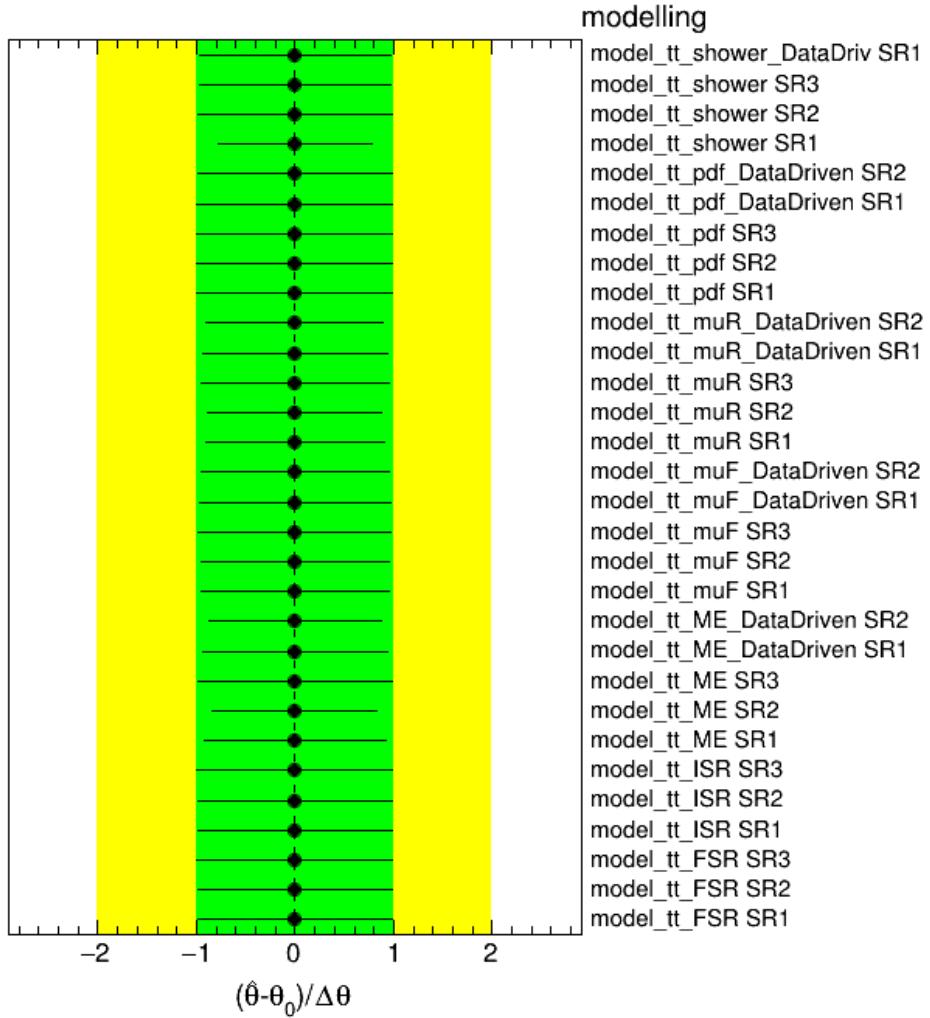


Figure 10.12: The nuisance parameter pulls for the modeling uncertainties of the top-quark pair background after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The term “shower,” “pdf,” “muR,” “muF,” “ME,” “ISR,” “FSR” refer to the parton shower generator, Parton Distribution Function, renormalization scale, factorization scale, initial state radiation, and final state radiation respectively. These terms are described in section 2.1 and subsection 9.1.3. The region code (Figure 7.2) at the end further specifies the signal region where the $t\bar{t}$ background is to be varied. Nuisance parameters of “DataDriven (DataDriv)” account for the uncertainty from subtracting the $t\bar{t}$ background in the specified signal region’s neighboring template region (the b-candidate jet is not b -tagged rather than tagged) in the data-driven background estimation.

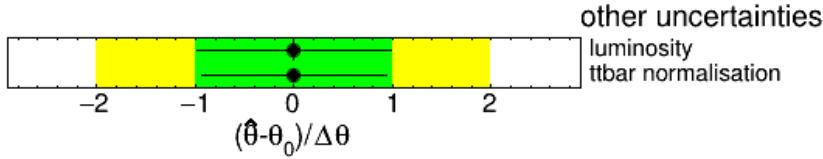


Figure 10.13: The nuisance parameter pulls for the 1.7% luminosity uncertainty and the 4% top-quark pair background production cross-section uncertainty after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.”

Shower Model 80” in Figure 10.9, “b-tag High-Pt Extrapolation” in Figure 10.10, all data-driven systematics in Figure 10.11, and six $t\bar{t}$ modeling uncertainties in Figure 10.12 – “model_tt_shower SR1,” “model_tt_muR_DataDriven SR1,” “model_tt_muR SR1 and SR2,” “model_tt_ME_DataDriven SR2,” and “model_tt_ME SR1.” The nuisance parameters for the data-driven background are even more constrained than those for the $t\bar{t}$ background since the former accounts for 80% or higher of the total background in each region (The region with the highest portion of $t\bar{t}$ background is shown in Figure 8.2(b)). The dominance increase with M_{tb} , leaving fewer degrees of freedom provided by nuisance parameters of the $t\bar{t}$ background modeling. A sudden jump of systematic variation around 2 TeV (Figure 8.9) outweighs the increases in Poisson errors, which scales with the square root of event counts, leading to more significant constraints found in the high M_{tb} components of the systematics.

The correlations between nuisance parameters are obtained from the covariance matrices of the likelihood function evaluated at the best-fit parameters, which are identical to their nominal values in the $\mu = 0$ pseudo-data. Figure 10.14 shows the most significant correlation factors between the nuisance parameters, none of whose absolute value exceeds 30%. It could be counterintuitive at first sight that there are nonzero correlation factors between the nuisance parameters of the data-driven background in different signal and M_{tb} regions. But these isolated nuisance parameters share the degree of freedom with other nuisance

Data-driven Model HighMtbSR1	100.0	8.0	-8.0	-29.0	11.9	-2.3	1.8	-0.7	2.7	1.1	-18.6	1.7
Data-driven Model HighMtbSR2	8.0	100.0	-3.4	-25.7	-1.8	22.3	-21.2	-12.6	1.3	3.5	3.5	0.8
Data-driven Model LowMtbSR3	-8.0	-3.4	100.0	11.0	-11.9	0.5	-0.2	-0.6	28.0	0.0	-7.4	21.8
b-tag High-Pt Extrapolation	-29.0	-25.7	11.0	100.0	4.1	7.7	-6.9	4.7	-4.5	-5.0	-7.1	-2.7
Top-tag Other Top Eff. 50	11.9	-1.8	-11.9	4.1	100.0	0.3	0.3	-0.9	3.2	0.5	24.7	2.4
model_tt_ME_SR2	-2.3	22.3	0.5	7.7	0.3	100.0	29.9	-13.9	-0.2	19.2	-0.9	-0.1
model_tt_ME_DataDriven_SR2	1.8	-21.2	-0.2	-6.9	0.3	29.9	100.0	24.3	0.1	-24.4	-0.6	0.1
model_tt_muR_SR2	-0.7	-12.6	-0.6	4.7	-0.9	-13.9	24.3	100.0	0.1	24.0	1.8	0.1
model_tt_muR_SR3	2.7	1.3	28.0	-4.5	3.2	-0.2	0.1	0.1	100.0	0.1	2.1	-5.6
model_tt_muR_DataDriven_SR2	1.1	3.5	0.0	-5.0	0.5	19.2	-24.4	24.0	0.1	100.0	-1.1	0.0
model_tt_shower_SR1	-18.6	3.5	-7.4	-7.1	24.7	-0.9	-0.6	1.8	2.1	-1.1	100.0	1.5
model_tt_shower_SR3	1.7	0.8	21.8	-2.7	2.4	-0.1	0.1	0.1	-5.6	0.0	1.5	100.0

Data-driven Model HighMtbSR1
 Data-driven Model HighMtbSR2
 Data-driven Model LowMtbSR3
 b-tag High-Pt Extrapolation
 Top-tag Other Top Eff. 50
 model_tt_ME_SR2
 model_tt_ME_DataDriven_SR2
 model_tt_muR_SR2
 model_tt_muR_SR3
 model_tt_muR_DataDriven_SR2
 model_tt_shower_SR1
 model_tt_shower_SR3

Figure 10.14: The correlation matrix between nuisance parameters after a profile-likelihood fit to the pseudo-data with “ $\mu = 0$.” The matrix is symmetric over the diagonal from the top left to the bottom right corner. Only nuisance parameters with more than 20% of correlation with another are shown. The meaning of individual nuisance parameters can be found in the preceding Figure 10.5 to Figure 10.13.

parameters common to all regions, such as the “b-tag High-Pt Extrapolation,” that share some shape similarities with the former. The majority of stronger correlations are associated with the $t\bar{t}$ background modeling systematics of different regions (including those directly applied to the signal regions versus those propagated through the data-driven estimation from the template regions – “DataDriven”). Still, the small correlation justifies their separations a posteriori.

10.3 Exclusion Upper Limit with the Background-only Pseudo-data

Chapter 9 introduces a method of excluding signal cross-sections too large to be compatible with data when the CLs is less than 5%. Figure 10.15 shows this 5% boundary on the cross-section times the branching ratio, the upper limit, by treating the background predictions as data. The limit curve consists of discrete points calculated for five W'_R signals with masses ranges from 1.5 TeV to 5 TeV (Chapter 5). From the crossing point between the theoretical cross-sections and the expected limit, the mass of the right-handed W' is excluded up to about 4.15 TeV at the 95% Confidence Level. This limit represents the expected result to be compared with the limit computed with the actual data in later sections.

10.4 The Signal Region Fit with Observed data

Data is compared with the background predictions before and after the profile-likelihood fit under the background-only hypothesis (μ fixed to 0 in practice), as shown in Figure 10.16 and Figure 10.17, respectively. While the fit reduces the background uncertainties significantly,

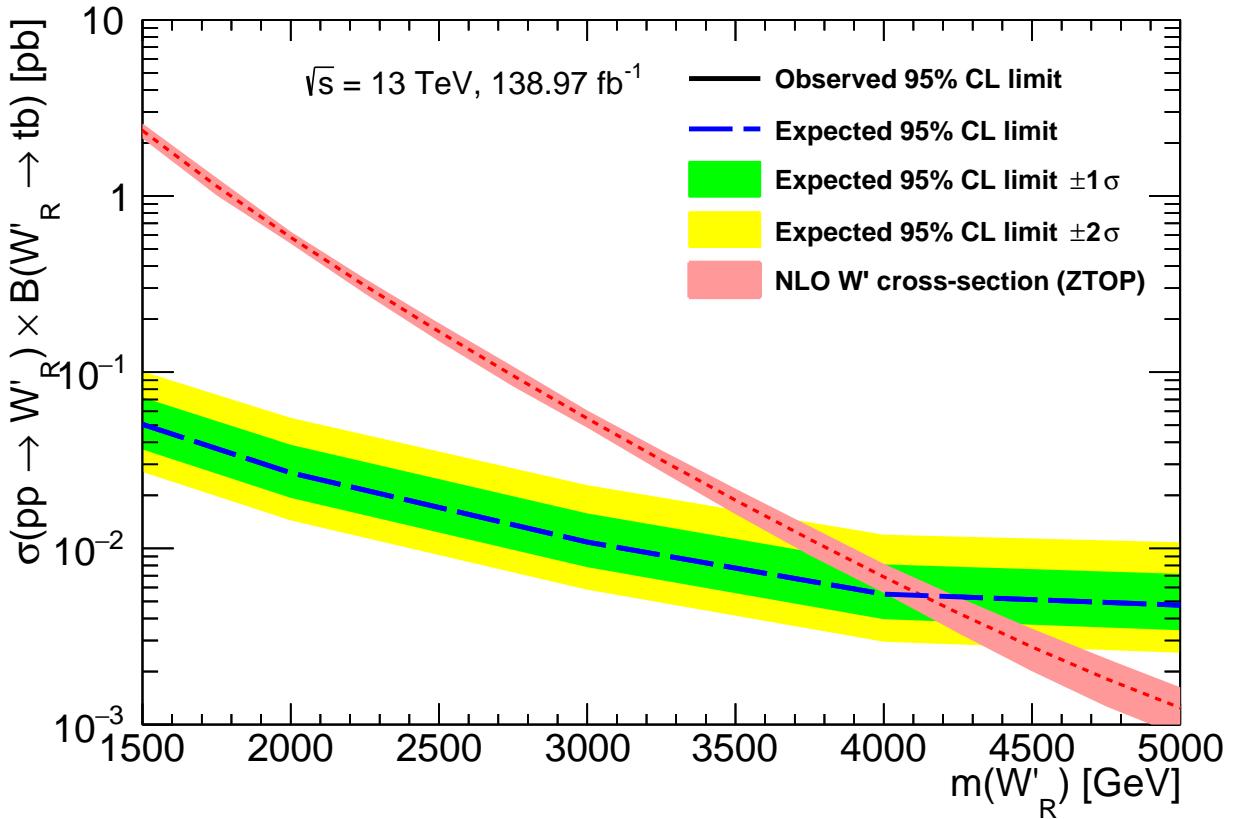


Figure 10.15: The upper limit on the W'_R production cross-section times the decay branching ratio of $W'_R \rightarrow tb$ at the 95% Confidence Level is evaluated with the background-only pseudo-data. The blue dashed curve indicates the median, while the green and yellow bands show the one and two Gaussian quantiles of the limit under statistical plus systematic uncertainties. The Next-to-Leading-Order theoretical cross-section times the branching ratio is drawn as the red dashed curve with an asymmetric theoretical uncertainty band (Chapter 5).

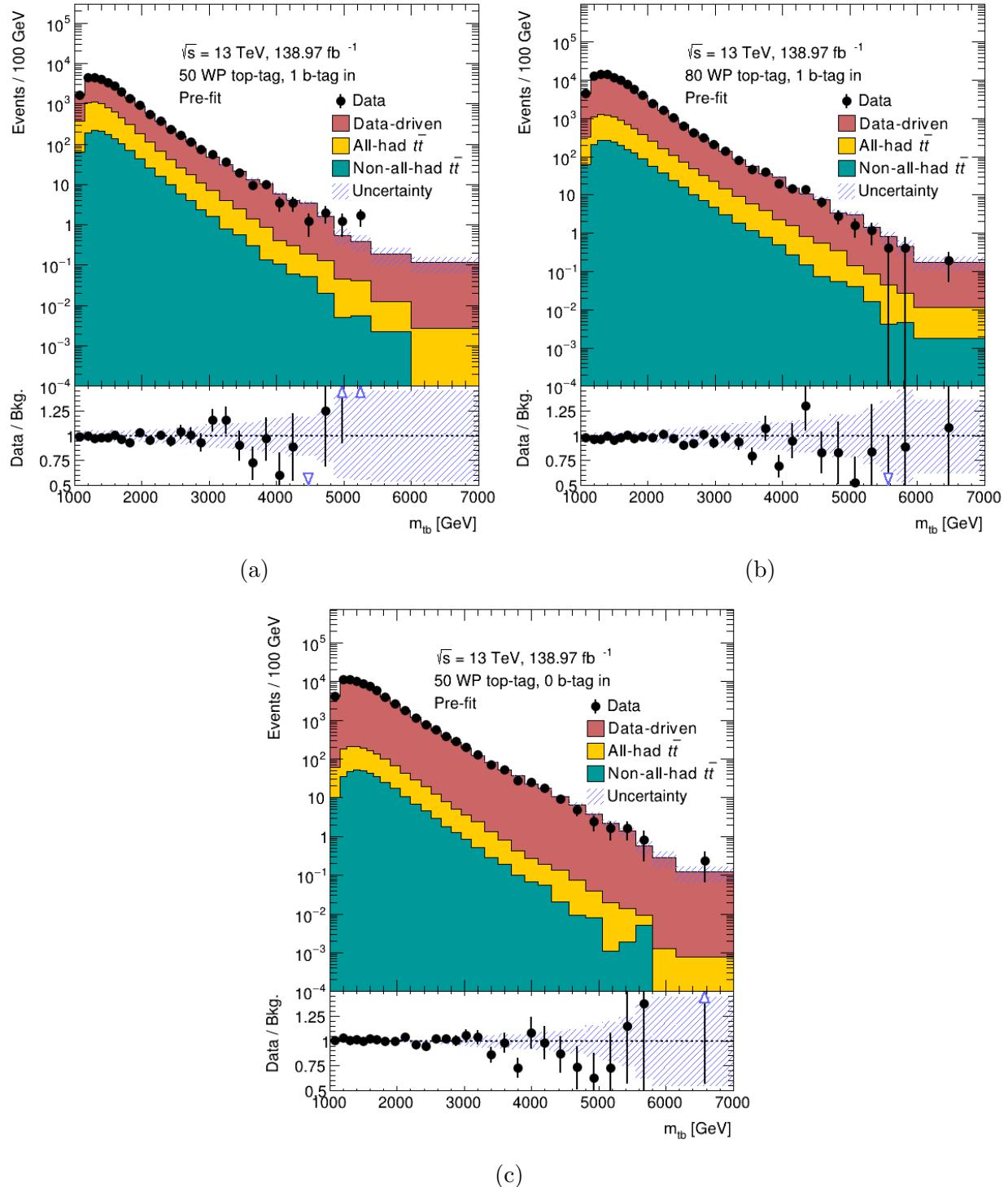


Figure 10.16: Data versus background M_{tb} distributions in the signal regions: (a) SR1, (b) SR2, and (c) SR3 before the profile-likelihood fit. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.

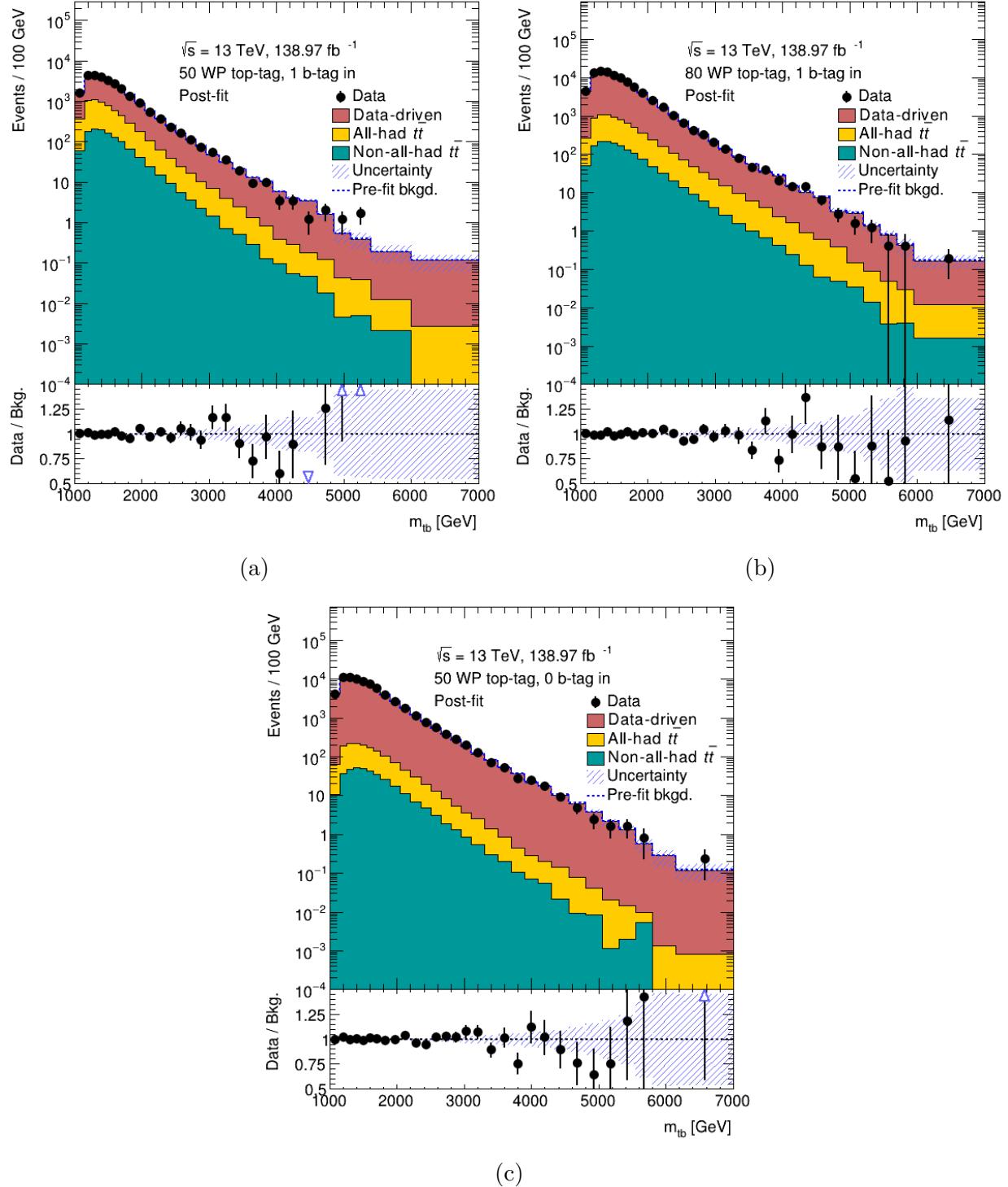


Figure 10.17: Data versus background M_{tb} distribution in the signal regions: (a) SR1, (b) SR2, and (c) SR3 after the profile-likelihood fit with under the background-only hypothesis. The hatched uncertainty bands include the systematic uncertainties and the statistical uncertainties of the background predictions.

the background predictions hardly deviate from the pre-fit values in all three regions. With excellent consistency between data and the nominal background for $M_{tb} < 3$ TeV, the largest shift of bin heights of the background after the fit is observed in the high mass regions in the “80 WP top-tag, 1 b-tag in” (SR2, Figure 10.17(b)) region: 5% down and the “50 WP top-tag, 0 b-tag in” (SR3, Figure 10.17(c)) region: 3% down. Both of them signify a surplus of predicted background events. The only significant excess is seen in one bin around 5 TeV in the “50 WP top-tag, 1 b-tag in” (SR1, Figure 10.17(a)) region, where the total post-fit background = 1.18 ± 0.51 versus 5 counts in data. It corresponds to background rejection significance of just 2.2σ (according to formula (25) in Ref. [136]).

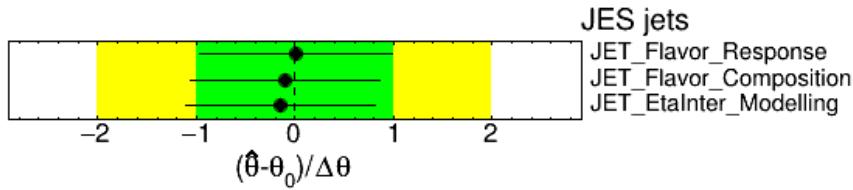


Figure 10.18: The nuisance parameter pulls for jet energy scale uncertainties common to large-radius jets and small-radius jets after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.5.

The pull plots for nuisance parameters after a background-only profile-likelihood fit to data are shown from Figure 10.18 to Figure 10.26. None of the nuisance parameters is pulled beyond one σ of the pre-fit uncertainty. Most of the significant pulls are seen in the nuisance parameters specific to the “80 WP top-tag, 1 b-tag in” (SR2, Figure 10.17(b)) region, such as those bearing the name “80” in their suffices in Figure 10.22 and also those with “SR2” in Figure 10.24 and Figure 10.25. The low data to total background ratios in high M_{tb} of the region and the lowest statistical errors among all signal regions contribute to the shift. A similar effect explains the substantial pull for the data-driven background uncertainty for the high M_{tb} part in the “SR3” region (Figure 10.24). The region’s requirement of 0 b-tagged

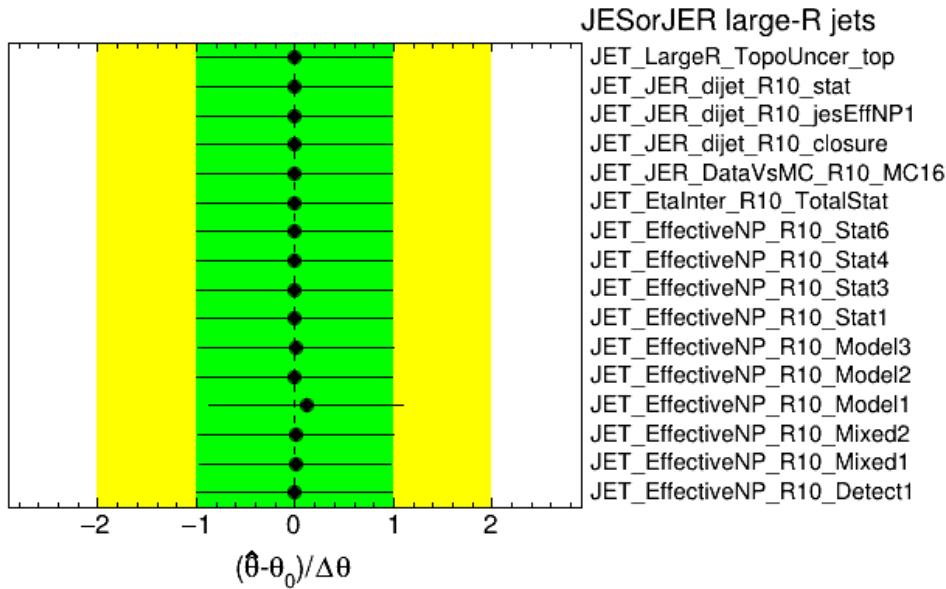


Figure 10.19: The nuisance parameter pulls for large-radius jet energy scale and resolution uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.6.

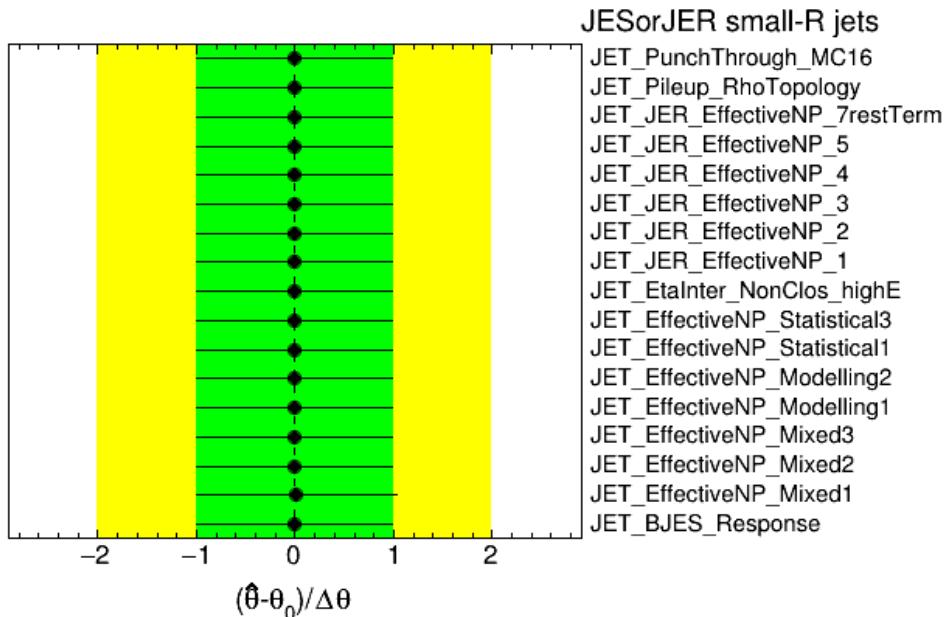


Figure 10.20: The nuisance parameter pulls for small-radius jet energy scale and resolution uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.7.

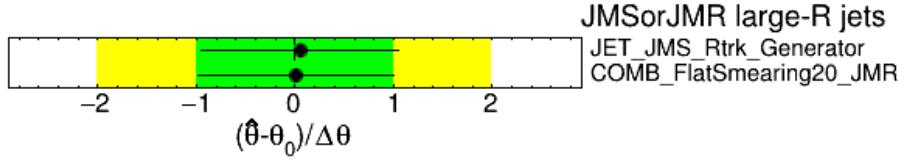


Figure 10.21: The nuisance parameter pulls for large-radius jet mass scale and resolution uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.8.

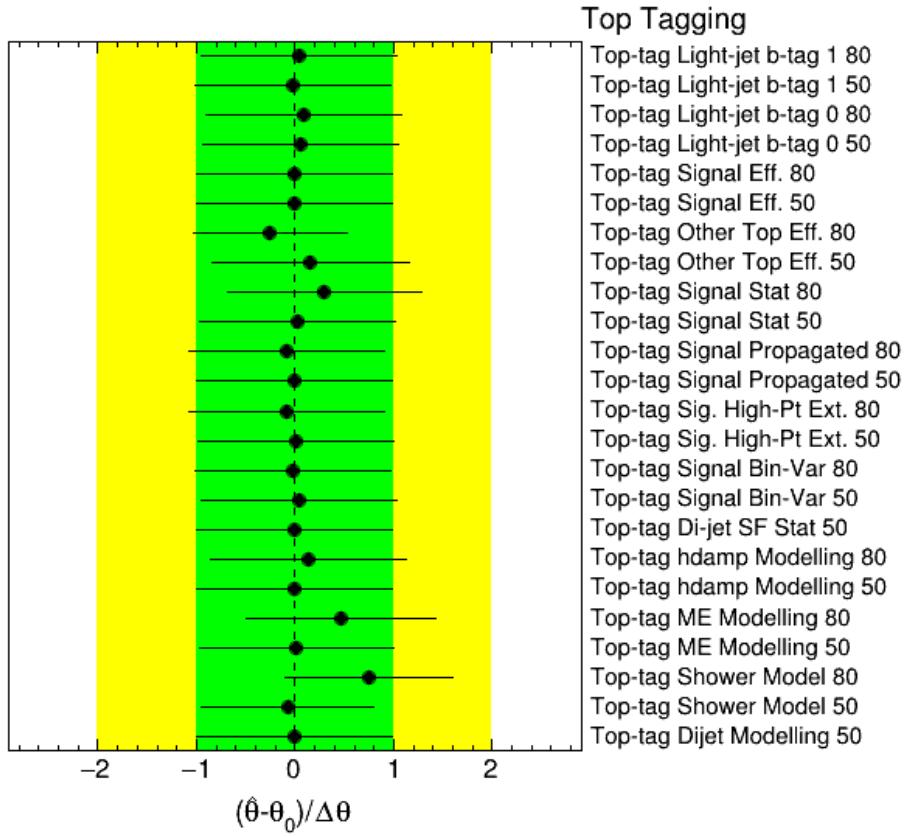


Figure 10.22: The nuisance parameter pulls for the top-tagging uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.9.

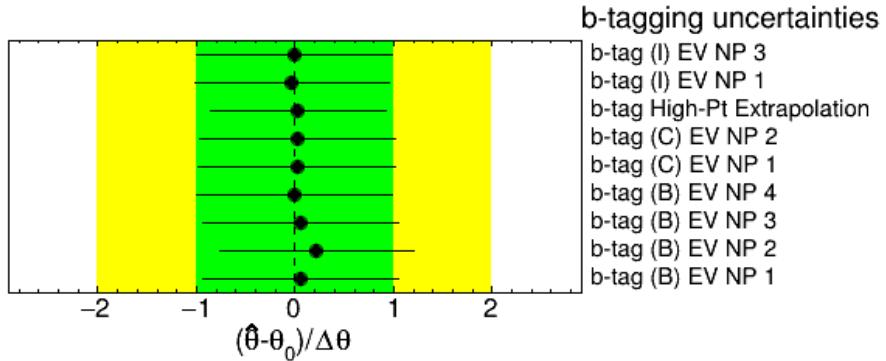


Figure 10.23: The nuisance parameter pulls for the b -tagging uncertainties after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.10.

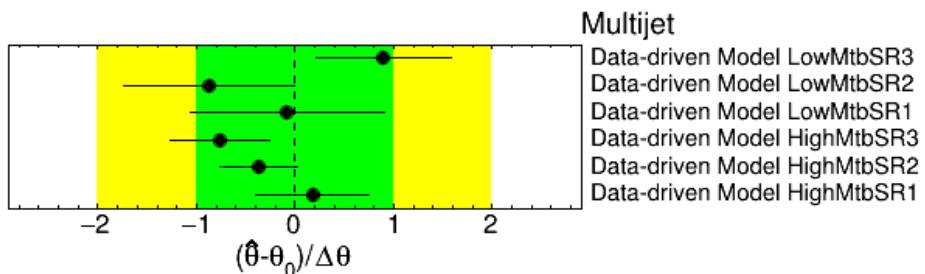


Figure 10.24: The nuisance parameter pulls for the modeling uncertainties of the data-driven background estimation after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.11.

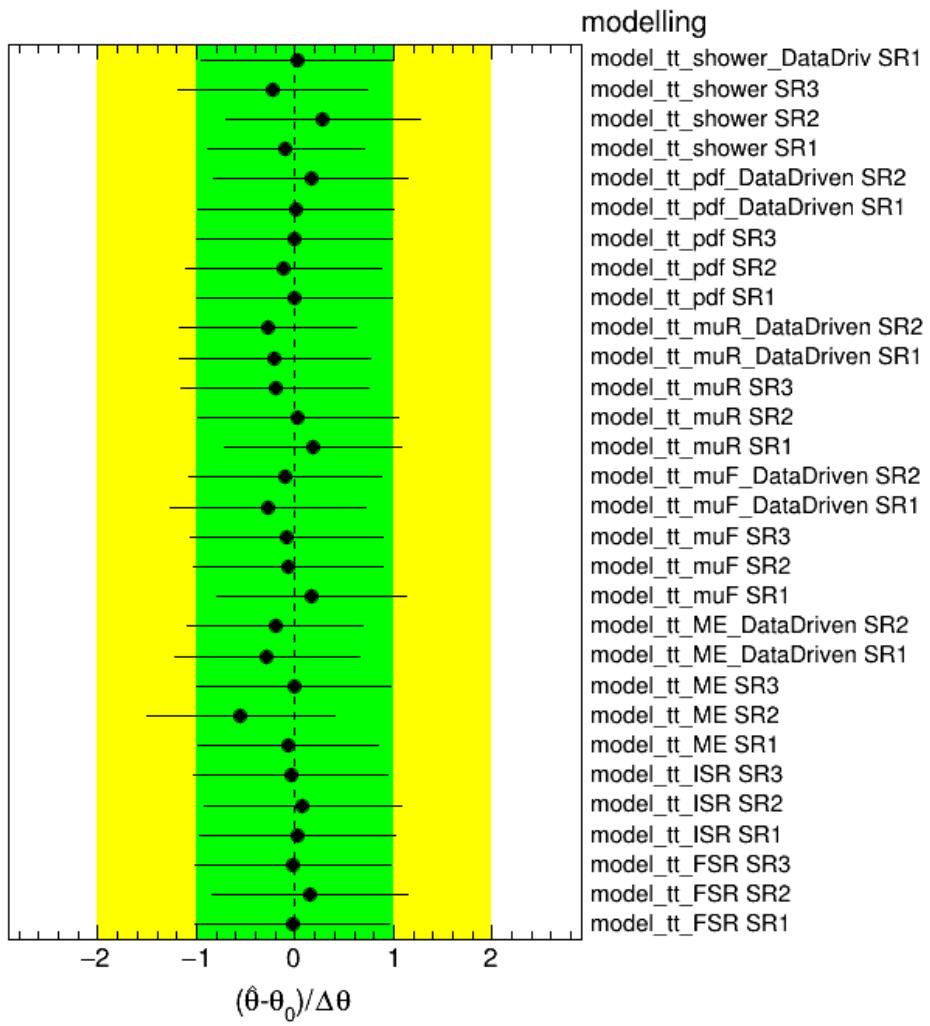


Figure 10.25: The nuisance parameter pulls for the modeling uncertainties of the top-quark pair background after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The meaning of the nuisance parameters are explained in Figure 10.12.

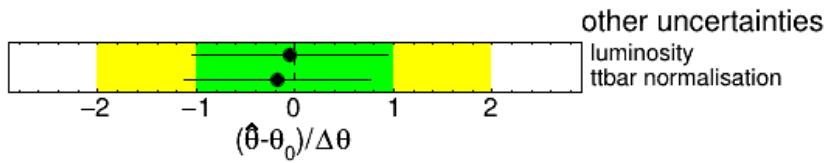


Figure 10.26: The nuisance parameter pulls for the 1.7% luminosity uncertainty and the 4% top-quark pair background production cross-section uncertainty after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$).

	Data-driven Model HighMtbSR1	8.3	-0.9	-9.2	-28.7	16.3	-1.2	2.1	-0.4	3.1	1.3	-15.0	1.7
Data-driven Model HighMtbSR2	8.3	100.0	22.0	-2.9	-27.3	-2.6	22.9	-18.1	-4.6	1.7	9.8	3.0	0.9
Data-driven Model LowMtbSR2	-0.9	22.0	100.0	0.7	1.0	1.4	-7.1	-9.1	10.4	-0.2	-18.0	-1.8	-0.1
Data-driven Model LowMtbSR3	-9.2	-2.9	0.7	100.0	9.1	-17.7	0.4	0.0	-0.3	28.3	0.4	-10.0	22.4
b-tag High-Pt Extrapolation	-28.7	-27.3	1.0	9.1	100.0	6.0	3.9	-8.1	2.8	-5.8	-5.9	-6.4	-3.2
Top-tag Other Top Eff. 50	16.3	-2.6	1.4	-17.7	6.0	100.0	0.4	0.4	-0.8	2.9	0.9	29.8	1.8
model_tt_ME SR2	-1.2	22.9	-7.1	0.4	3.9	0.4	100.0	29.3	-1.5	-0.2	25.7	-0.5	-0.1
model_tt_ME_DataDriven SR2	2.1	-18.1	-9.1	0.0	-8.1	0.4	29.3	100.0	30.3	0.3	-21.3	-0.3	0.1
model_tt_muR SR2	-0.4	-4.6	10.4	-0.3	2.8	-0.8	-1.5	30.3	100.0	-0.1	33.3	1.1	-0.0
model_tt_muR SR3	3.1	1.7	-0.2	28.3	-5.8	2.9	-0.2	0.3	-0.1	100.0	0.1	2.1	-3.9
model_tt_muR_DataDriven SR2	1.3	9.8	-18.0	0.4	-5.9	0.9	25.7	-21.3	33.3	0.1	100.0	-1.0	0.1
model_tt_shower SR1	-15.0	3.0	-1.8	-10.0	-6.4	29.8	-0.5	-0.3	1.1	2.1	-1.0	100.0	1.3
model_tt_shower SR3	1.7	0.9	-0.1	22.4	-3.2	1.8	-0.1	0.1	-0.0	-3.9	0.1	1.3	100.0

Figure 10.27: The correlation matrix between nuisance parameters after a profile-likelihood fit to data under the background-only hypothesis ($\mu = 0$). The matrix is symmetric over the diagonal from the top left to the bottom right corner. Only nuisance parameters with more than 20% of correlation with another are shown. The meaning of individual nuisance parameters can be found in the preceding Figure 10.5 to Figure 10.13.

small-radius jet inside the top-candidate jet suppressed the $t\bar{t}$ background, thus relying on varying the data-driven background to achieve the best fit.

As suggested by a previous section, pulls and constraints of a nuisance parameter provide information on the physics behind the corresponding systematic uncertainty assessed by asantz or calibrations. The data-driven estimation in Chapter 8 assumes no correlation between the top-tagging and b -tagging for the pre-fit background and assigns systematic uncertainties to model the nonzero correlation estimated from the control regions. The flavor composition of partons is found to be a physical source for a decreasing b -tagging ratio with increasing top-tagging level (less background), while the two taggers' high- p_T -preferring tagging probabilities imply an increasing b -tagging ratio (more background). As none of the nuisance parameters show a significant inconsistency with 0 in either positive or negative direction (coinciding with the background's increase and decrease), the two sources do not manifest in either the low or high M_{tb} range. Nonetheless, an intriguing situation is observed in the region “SR3” that the pull directions seen for the two mass ranges misalign with the prior assumption that the b -tagging ratio is suppressed by the differences in flavor composition mainly in low invariant mass and raised by the kinematic effect in the high mass region.

No affirmative physical interpretation can be attained from the pull plots as there is no significant pull relative to the uncertainty. Especially, Figure 10.24 cannot tell which effect is more dominant in data: the negative correlation (below 0 on the plot) caused by the parton flavor compositions of the jet or the positive correlation (above 0 on the plot) caused by the increasing probabilities to either top-tag or b -tag a jet with p_T (Chapter 8). Nevertheless, the differences in nuisance parameter shifts between different signal regions, as seen for the data-driven background estimation uncertainties (Figure 10.24) and for the $t\bar{t}$

background uncertainties (Figure 10.22), suggest that top-tagging and b -tagging algorithms have sensitivities to physics underlying the modeling systematics.

The correlation coefficients between nuisance parameters after the fit to data are shown in Figure 10.27. There is no correlation coefficient exceeding 34% in absolute values and no significant change from the values for the fit to the background-only pseudo-data in Figure 10.14. One noticeable difference is the appearance of entries corresponding to the low M_{tb} correlation systematic uncertainties for Signal Region 2 (named “Data-driven Model LowMtbSR2), with a correlation of 22% with the high M_{tb} counterpart, probably to accommodate pulls from other systematic uncertainties in this region. This correlation coefficient was around 12% in the pseudo-data fit, thus not shown in Figure 10.14.

10.5 Exclusion Upper Limit with Observed Data

The limit on the W'_R production cross-section times the branching ratio of the tb decay channel – excluded at the 95% Confidence Level – is set for five separate W' masses. They are connected as the exclusion curves shown in Figure 10.28. As mentioned in Chapter 9, the unblinded expected limit employs pseudo-data by letting $\mu = 0$ in the signal-plus-background model after a fit to data with the signal strength μ unconstrained. This procedure “fixes” the nuisance parameters to their best-fit values to form the background hypothesis corrected by data while reducing the potential signal by removing the signal shape scaled by μ . The resulting expected limit is consistent with that obtained by the blinded pseudo-data (Figure 10.15), reflecting the insubstantial nuisance parameter pulls in the fit to the observed data. The observed limit is calculated using the observed data without the fixation of μ to 0. Thus, the excess or deficit of events of the observed data as compared with the pseudo-data

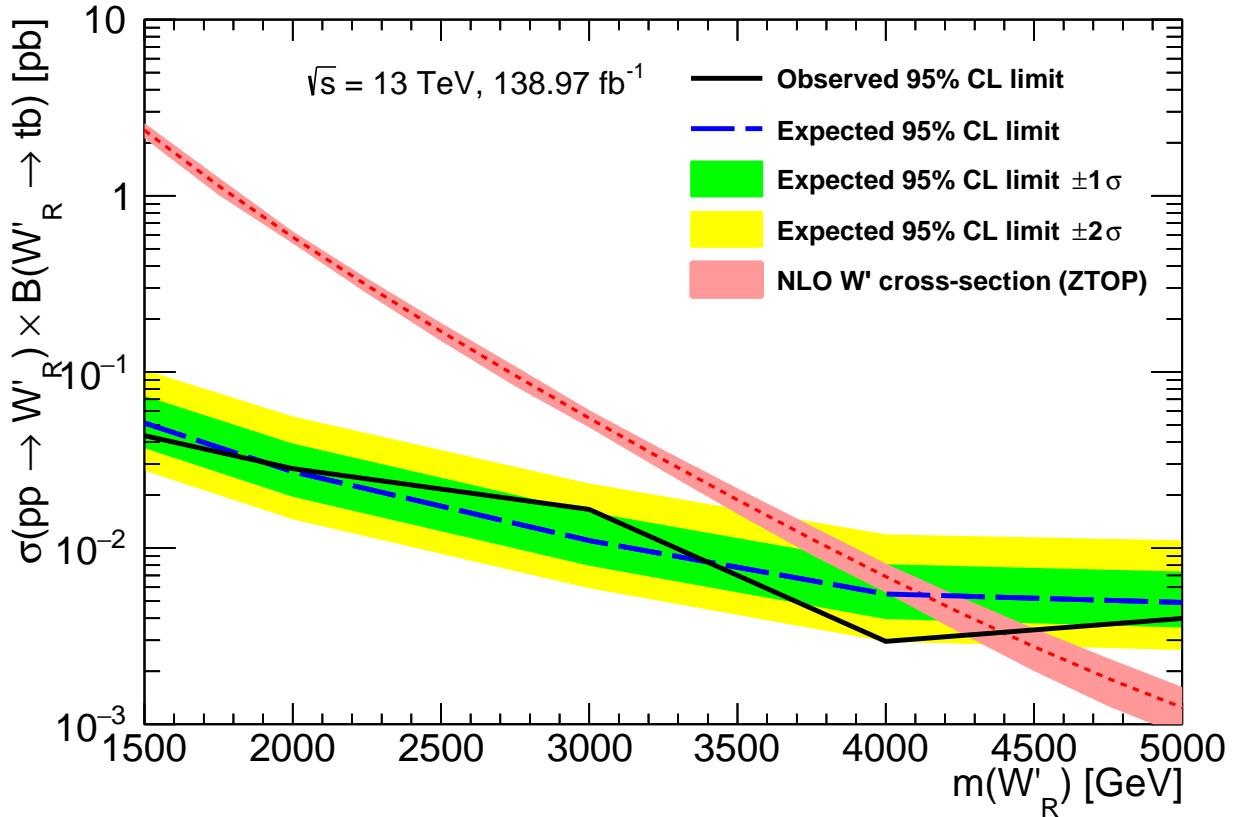


Figure 10.28: The upper limit on the W'_R production cross-section times the decay branching ratio of $W'_R \rightarrow tb$ at the 95% Confidence Level is evaluated with data. The blue dashed curve indicates the median expected limit, while the green and yellow bands show the one and two Gaussian quantiles under statistical plus systematic uncertainties. The black solid curve shows the observed limit. The Next-to-Leading-Order theoretical cross-section times the branching ratio is drawn as the red dashed curve with an asymmetric theoretical uncertainty band (Chapter 5).

causes the upward and downward discrepancies between the two limit curves. The crossing point between the theory curve and the interpolated limit between the 4 and 5 TeV mass gives the observed mass limit at around 4.4 TeV for a right-handed W' .

10.6 Systematic Uncertainty Rankings for a 4 TeV W'_R

Since the 4 TeV right-handed W' is the most massive signal model excluded at the 95% Confidence Level, the most impactful systematic uncertainties in its fit are studied in the following with three different datasets: the pseudo-data composed only of background predictions, the pseudo-data with backgrounds plus a 4 TeV W'_R signal scaled to the blinded expected limit shown in Figure 10.15, and the observed data.

10.6.1 Background-only pseudo-data

The background-only pseudo-data is fitted with μ free-floating. The fit returns $\hat{\mu} = 0.00 \pm 0.46$, in the unit of the blinded expected limit for the 4 TeV W'_R signal (Figure 10.15), and without any nuisance parameter shifted. Figure 10.29 shows the nuisance parameter ranking by repeating the fit while varying one nuisance parameter by its pre-fit or post-fit error at a time (as described in Chapter 9). Only the top three nuisance parameter yield a significant change in $\hat{\mu}$ (μ given by the fit), as compared to the systematic plus statistical uncertainty of μ . These three nuisance parameters characterize the correlation uncertainty in the data-driven background estimation in $M_{tb} > 2$ TeV (Chapter 8) for the indicated signal region. Since they are all highly constrained in the pseudo-data fit (Figure 10.11), it is unsurprising that the distortion of the background distribution incurred by the shift by one σ needs to be balanced by a substantial change in μ to minimize the likelihood function.

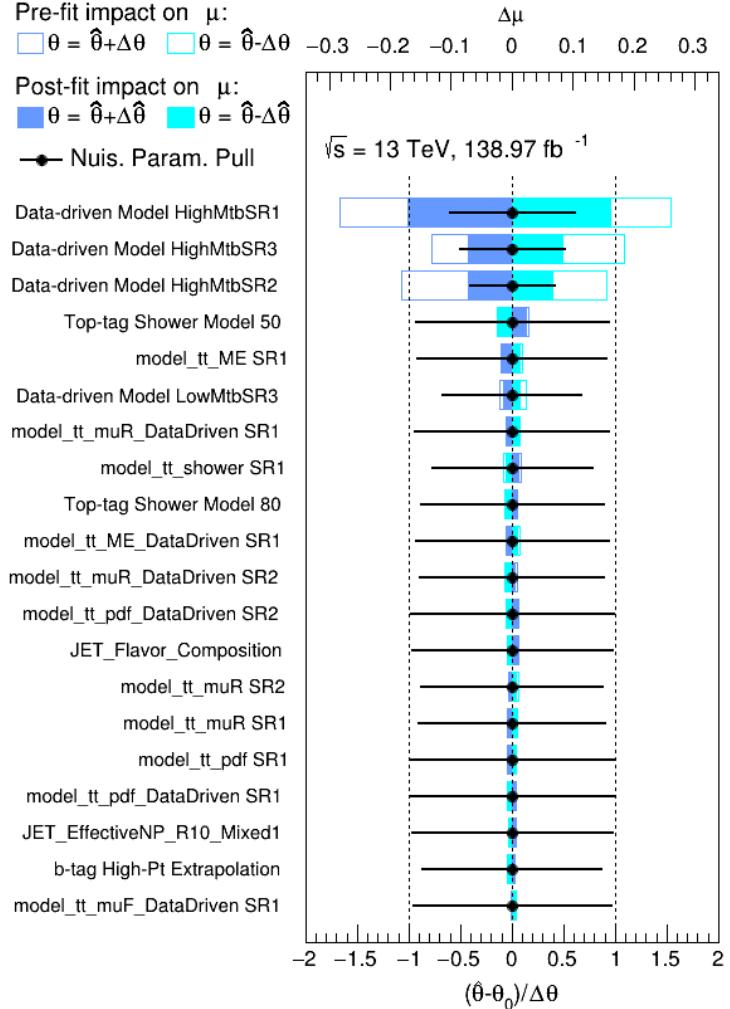


Figure 10.29: The nuisance parameter rankings for the 4 TeV W'_R signal with respect to the pseudo-data with “ $\mu = 0$.” Color-filled bands show the fifteen-highest impact on the best-fit μ ($\Delta\mu$) by fixing a particular nuisance parameter to its best-fit value plus-or-minus the post-fit error ($\hat{\theta} \pm \Delta\hat{\theta}$). The unfilled bands with colored boundaries correspond to $\Delta\mu$ when the nuisance parameter is shifted by its pre-fit uncertainty ($\hat{\theta} \pm \Delta\theta$). The scale of $\Delta\mu$ is shown on the top horizontal line. The bottom horizontal scale concerns with the black dots and bars, the best-fit nuisance parameters and their post-fit errors, as in the pull plots. The captions from Figure 10.5 to Figure 10.13 explain the naming convention of the nuisance parameters.

10.6.2 Signal-plus-background pseudo-data

The background-plus-signal pseudo-data is fitted with the μ free-floating. The fit returns $\hat{\mu} = 1.00 \pm 0.53$, in the unit of the blinded expected limit for the 4 TeV W'_R signal (Figure 10.15), and without any nuisance parameter shifted. Figure 10.30 shows the nuisance parameter ranking by repeating the fit while varying one nuisance parameter by its pre-fit or post-fit error at a time. Only the top three significantly change $\hat{\mu}$ in comparison with the systematic plus statistical uncertainty of μ . Since a significant amount of signal events are injected into the pseudo-data, μ becomes susceptible to the nuisance parameter variations from the signal sample than it does in the fit with the background-only pseudo-data shown in Figure 10.29. This difference explains why the top-ranked nuisance parameter is replaced by the high p_T uncertainty on the b -tagging probability, which only affects the most dominant background indirectly via subtracting the $t\bar{t}$ background in the template region. The emergence of signal could also explain the increased influences on the fit outcome by the top-tagging systematics but fall short for the rise of top-quark pair background modeling and the fall of the data-driven uncertainty of Signal Region 1 and 3. The shape similarities between the varied nuisance parameter and the signal shape could play a role. For example, the variation for Signal Region 1 and 2 stays flat above 3.5 TeV in M_{tb} (Figure 8.9(d) and (c)), while that for Signal Region 3 (Figure 8.9(b)) has a rising structure above 4 TeV that better mimics the signal shape.

10.6.3 Observed data

The observed data is fitted with μ free-floating. The fit returns $\hat{\mu} = -1.67 \pm 0.95$, in the unit of the observed limit for the 4 TeV W'_R signal (Figure 10.28), with nuisance parameters pulled

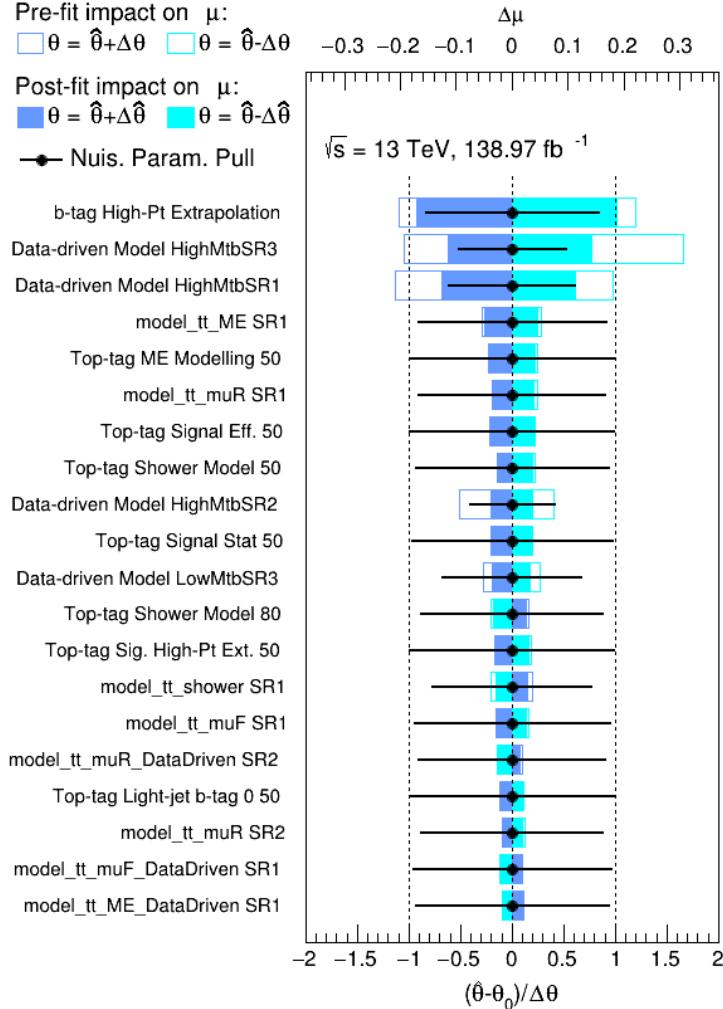


Figure 10.30: The nuisance parameter rankings for the 4 TeV W'_R signal with respect to the pseudo-data with “ $\mu = 1$.” Color-filled bands show the fifteen-highest impact on the best-fit μ ($\Delta\mu$) by fixing a particular nuisance parameter to its best-fit value plus-or-minus the post-fit error ($\hat{\theta} \pm \Delta\hat{\theta}$). The unfilled bands with colored boundaries correspond to $\Delta\mu$ when the nuisance parameter is shifted by its pre-fit uncertainty ($\hat{\theta} \pm \Delta\theta$). The scale of $\Delta\mu$ is shown on the top horizontal line. The bottom horizontal scale concerns with the black dots and bars, the best-fit nuisance parameters and their post-fit errors, as in the pull plots. The captions from Figure 10.5 to Figure 10.13 explain the naming convention of the nuisance parameters.

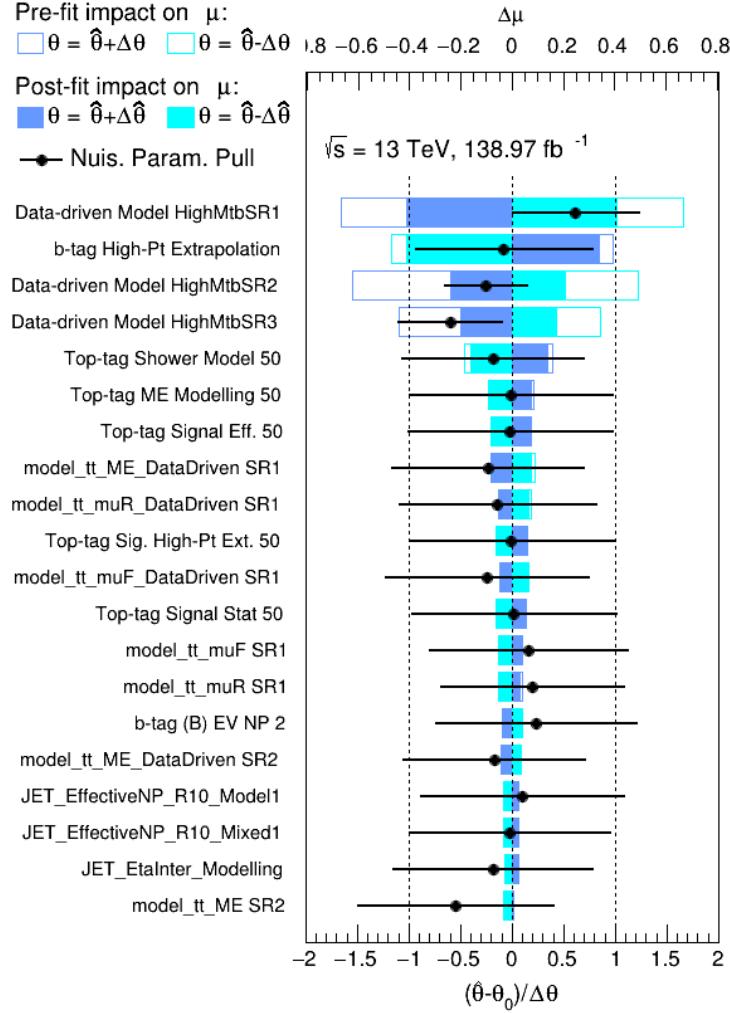


Figure 10.31: The nuisance parameter rankings for the 4 TeV W'_R signal with respect to the observed data. Color-filled bands show the fifteen-highest impact on the best-fit μ ($\Delta\mu$) by fixing a particular nuisance parameter to its best-fit value plus-or-minus the post-fit error ($\hat{\theta} \pm \Delta\hat{\theta}$). The unfilled bands with colored boundaries correspond to $\Delta\mu$ when the nuisance parameter is shifted by its pre-fit uncertainty ($\hat{\theta} \pm \Delta\theta$). The scale of $\Delta\mu$ is shown on the top horizontal line. The bottom horizontal scale concerns with the black dots and bars, the best-fit nuisance parameters and their post-fit errors, as in the pull plots. The captions from Figure 10.5 to Figure 10.13 explain the naming convention of the nuisance parameters.

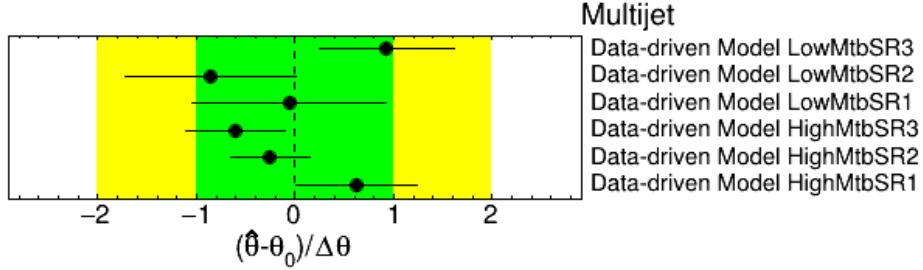


Figure 10.32: The nuisance parameter pulls for the modeling uncertainties of the data-driven background estimation after a profile-likelihood fit to data with μ determined by the fit. The meaning of the nuisance parameters are explained in Figure 10.11.

similarly to those shown from Figure 10.18 to Figure 10.26. The largest difference in pulls is seen for the high M_{tb} data-driven background systematics for the region “50 WP top-tag, 1 b-tag in,” which can be observed by comparing Figure 10.24 with Figure 10.32. This nuisance parameter is ranked at the top in Figure 10.31. Here, only the top five nuisance parameters have a significant $\Delta\mu$, as compared to the uncertainty of μ . The deficit of data in the high M_{tb} regions shown in Figure 10.17 draws $\hat{\mu}$ to a negative value, leading to more $\Delta\mu$ by nuisance parameters affecting the signal sample, as illustrated in the previous analysis on the pseudo-data with signal injection.

Chapter 11

Conclusion

While this text is dedicated to the search for a new particle W' , this thesis also aims to give general views on experimental methods of particle physics at the Large Hadron Collider. Physics concepts scatter across various chapters beyond the theoretical introductions in Chapter 2 are crucial to comprehend the logical chains that link electronic signals to the transient physical phenomena after every proton-proton collision. The particle W' , unlike an atom, cannot be visualized as microscopic humps imaged by an electronic microscope. But its existence can be indirectly seen by combing the jet momenta under the framework of Quantum Chromodynamics (QCD). Thus, whether the existence of W' could be established or not, the utilized apparatus and methodologies deserve appreciation.

On the surface, conducting an analysis, especially a search for new particles beyond the Standard Model, has seemingly been reduced to intricate plays with software. Thanks to the whole ATLAS collaboration's efforts in writing computer programs, mechanical constructions, and calibrations, experimental studies of a physics model can be embarked directly on the selection criterion placed on physical observables that enhance the significance of signal versus background. Even though designing event selection and performing statistical analysis still calls upon physics intuition from time to time, the interpretation of data still inevitably becomes less straightforward as the experimental toolbox advances. Analyzers see the physics information condensed to various figures and might need to reverse-engineer the

analysis tools to obtain the full idea associated with data points and curves.

Without unravelling details of the analysis process, this thesis attempts to add sense to this complicated plot-making machine. The concept of jets is first described in the theoretical discussion of parton showers and hadronization, and later the calorimeter measurement of jet energy is introduced with the machinery. The physical object chapter elaborates on the jet algorithms to connect partons with reconstructed jets. In the following section on large-radius jets, the comparison between the Feynman diagram of hadronic top-quark decay to the three-prong QCD radiation further illustrates jet substructure variables. Finally, the method of data-driven background estimation provides evidence of different parton flavors at play at the high-energy scale through their mutual influences on the top-tagging and b -tagging ratios.

A scrupulous survey of systematic uncertainty is of equal importance to revealing the physics logic in the analysis strategies. If significant systematic errors are not accounted for in the background models, then the fitting can overestimate the signal to fill the gap between backgrounds and data. Particular attention is paid to the significant systematic uncertainty from the data-driven background estimation. Contributors to the correlation between top-tagging and b -tagging have to be located to draw a limit on their effect on the estimation. Studies on parton flavor and the kinematic effect motivate the control region definitions and the assignments of nuisance parameters based on the invariant mass ranges. The resulting (pre-fit) systematic uncertainties turn out to be more or as significant as the data-driven background statistical uncertainties. The associated nuisance parameters are consistent with no correlation between the two taggers in two invariant mass ranges of each region after the fit with data, different from the expectation given by the simulation of QCD multi-jet events. Thus, a direction to reduce this systematic uncertainty might be possible

by understanding the discrepancy better in the future.

The three signal regions finally find consistencies between the background prediction and data. An insignificant but nonnegligible deficit of data around 4 TeV is observed in all three regions. Thus, among all five W' masses between 1.5 and 5 TeV studied, only the 4 TeV right-handed W' has the exclusion limit deviate more than one standard deviation from the expectation of background-only hypothesis. The interpolation between the 4 TeV and 5 TeV lead to the exclusion of the W'_R boson with the Standard Model electroweak couplings up to 4.4 TeV at the 95% Confidence Level.

With the labor spent on analysis strategies documented in this thesis, subsequent analysis can focus on testing more signal models with more sophisticated treatment. For example, the left-handed W' has quantum interferences with the W boson. The interference part on the M_{tb} distribution scales differently from the signal peak. Varying the coupling constant offers another desirable path as it fit the W' into more Beyond-the-Standard-Model scenarios. Enhancement of M_{tb} width of the signal due to a large coupling constant could become as significant as the effect of jet energy and angular resolution.

APPENDICES

Appendix A

The Definition of Electrons and Muons

This appendix lists the properties fulfilled by selected electrons and muons. They both have a minimum transverse momentum (p_T) requirement of 25 GeV.

Electrons are reconstructed by matching the inner-detector-fitted tracks to topological clusters of calorimeter cells, located in $|\eta|$ from 0 to 1.37 and 1.52 to 2.47 [137]. The calorimeter clusters must have more than 50% energy deposit in the electromagnetic calorimeter part (ECAL) to be considered in the electron reconstruction. The empty region in $|\eta|$ is the transition region from the barrel to the end-cap region, where substantial inactive materials intercept the area between the inner tracker and the ECAL.

The reconstructed tracks and clusters are classified by a likelihood function constructed to reject hadrons, leptons from hadron decays, photons, and pair productions from photons. The parameters of the likelihood function are based on physical properties such as depth, width, and energy profile of the shower. This analysis employs the tightest identification criterion where approximately 80% to 85% of electrons are accepted, and 99.7% to 99.9% of hadrons are rejected for p_T from 25 to 80 GeV [138].

Reconstructed muons are a combination of tracks reconstructed by the Muon Spectrometer (MS) with tracks fitted by the inner detectors after taking the magnetic field profile and

ionization in the Calorimeter into account. Consistency between the momenta measured by the two systems is used to suppress muons from light hadron. For muons originating from decays of heavy hadron (with a bottom quark or a charm quark) that occur much closer to the beamline, the closed approach of the muon flight path to the primary vertex is distinctly larger than a hard-scatter muon on average (the same idea as identifying a small-radius jet with a B -hadron decay, as discussed in Chapter 6).

This analysis only considers muons with $|\eta| < 2.5$ with the inner detector coverage. The rate of muon reconstruction and identification is uniformly 98% across a wide p_T range up to 50 GeV and almost the entire η plane [139]. A drop to around 60% of muon acceptance occurs at $|\eta| < 0.1$ due to a gap between sections of MS. Muons from light hadron decays have merely less than 1% of accepting probabilities which decrease further with p_T .

Appendix B

Insignificant Background Processes

This appendix concerns the Standard Model backgrounds much less prominent than the QCD multi-jet and the $t\bar{t}$ production. It starts with the single-top-quark production and the production of a vector boson in association with jets. The predicted yields for these backgrounds will be shown to be similar to the systematic uncertainties of the data-driven and $t\bar{t}$ backgrounds, and thus requiring no independent estimation. The last paragraph will suggest that other background processes like the di-boson productions are even more negligible.

Single-top

The first type of subdominant SM background is the single-top, meaning only a top quark is produced. It could be classified into three kinds – the s -channel, the t -channel, and the tW -associated production (Figure B.1). The s -channel is the SM counterpart of the left-handed W' signal. Therefore, it should be preferred by the event selections. However, the channel's production cross-section is dwarfed by other SM processes, canceling out the advantage in selection probabilities. The cross-section, computed by HATHOR [140, 141] at the next-to-leading order in QCD, is 10.32 pb. In comparison, the cross-section of the $t\bar{t}$ is 831.76 pb, as calculated by the TOP++ program (citations in Chapter 8).

Furthermore, the MADGRAPH+PYTHIA (the same event generators used for simulating

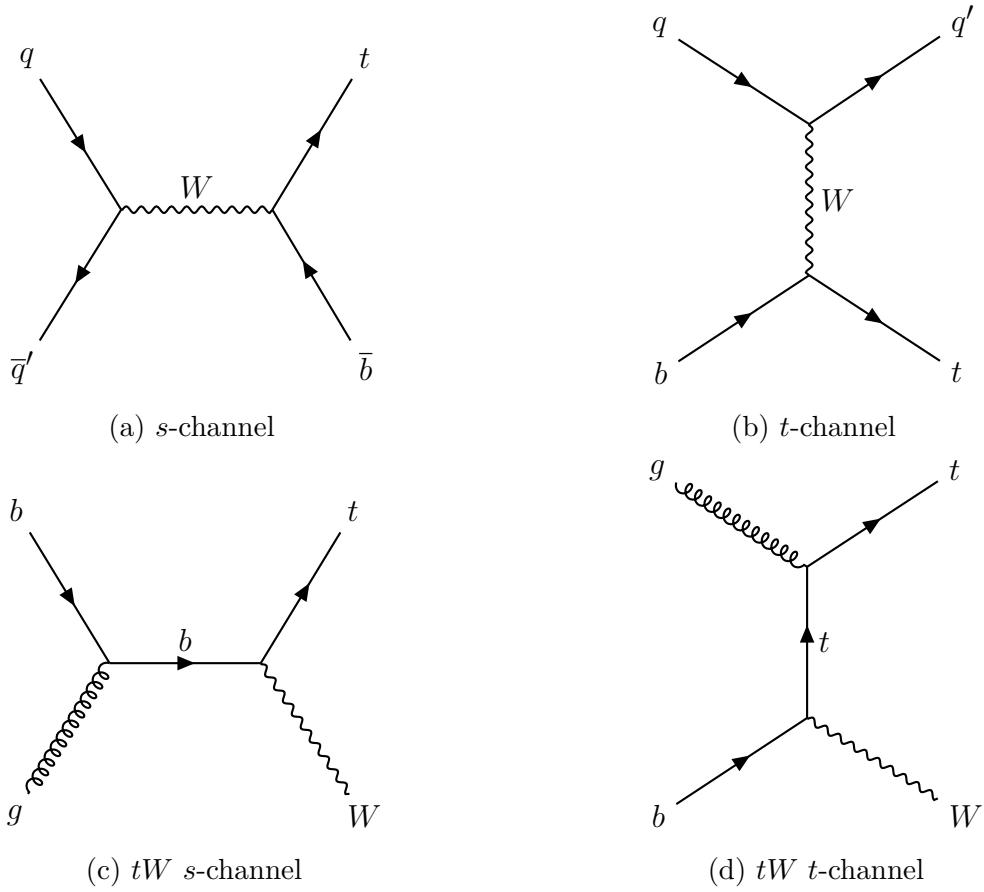


Figure B.1: The Feynman diagrams for hard-scatter matrix elements of the single-top processes at the LHC. The vertical axis represents space and the horizontal axis represents time. The diagrams above implicitly include three colors of quarks, eight types of gluons, and the combination rules between colors and gluon types. The q and q' is a duo of up-type and down-type quarks in the first two generation. The four diagrams are labelled by the virtual particle's Mandelstam variables and the final state particles. Feynman diagrams are made with TikZ-Feynman [64].

the W' events) reveals that less than 0.1% of the particle level jets from the s -channel could pass the $p_T > 500$ GeV and $|\eta|$ requirements of the event selection (Chapter 7). Considering the optimal top-tagging and b -tagging probabilities – 50% and 85% – only about 500 events remain in the 50% working point signal region in the “1 b -tag in” category. By contrast, $t\bar{t}$ has more than 7000 events in that category and is about one-third of the data-driven estimate (Figure 8.2). A surge in higher invariant mass is also unlikely as one of the s -channel initial parton is an antiquark, which is unlikely to be energetic.

The second way to produce a single top quark is through the t -channel, where a first- or second-generation parton exchanges energy with a bottom quark via a W boson. This mode has the highest production cross-section, 216.99 pb, according to HATHOR. Still, this value is less than the $t\bar{t}$ production cross-section. Moreover, this final state contains only one bottom particle, which is fewer than $t\bar{t}$ and the s -channel single-top. Therefore, the b -candidate jet from a t -channel single-top event is likely to have a similar b -tagging ratio as the QCD multi-jet (with some enhancement by the charm-quark final state). Resultingly, the data-driven background estimation does not underestimate its contribution in the signal region.

Lastly, the top quark could be accompanied by a W -boson when a bottom quark is scattered by a gluon. The initial gluon has a much lower probability than an up or down quark to possess high momentum fraction [20]. The suppression from initial states explains why the cross-section is merely 71.7 pb [142, 143]. The partner W -boson’s hadronic decay products seldomly contain a bottom quark. Thus, similar to the t -channel, the Wt events in the signal region are well-accounted for in extrapolating backgrounds from data in the template region. In the end, the single-top background does not need to be subtracted in the template region for the data-driven background, nor is their MC estimate in the signal

region necessary.

$W/Z/\gamma* + \text{jets}$

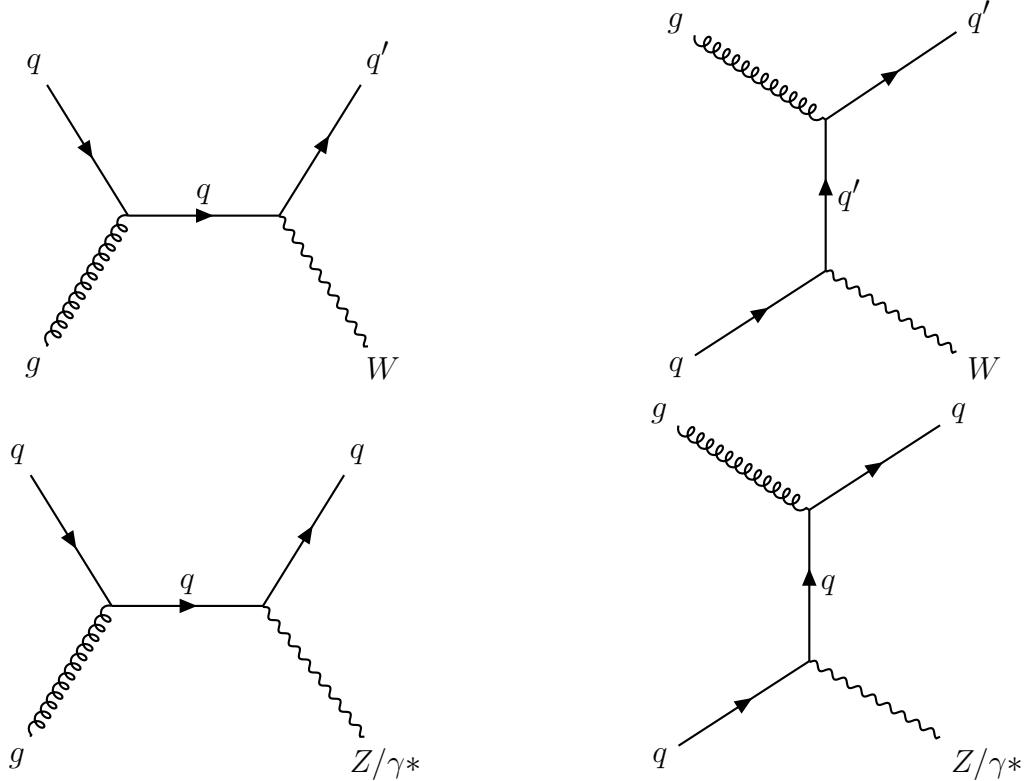


Figure B.2: The Feynman diagrams for hard-scatter matrix elements of the $V + \text{jet}$ processes at the LHC. The vertical axis represents space and the horizontal axis represents time. The diagrams above implicitly include three colors of quarks, eight types of gluons, and the combination rules between colors and gluon types. For the W -boson diagrams, q and q' is a duo of one of the two lighter up-type quark and one down-type quark. For the neutral boson diagrams, q is any quark but the top quark. γ^* means a photon carrying nonzero mass, i.e., being off-shell. Even though the initial state's flavor and color degrees of freedom are reminiscent of the QCD multi-jet events, the small electroweak coupling suppresses the vector boson production relative to the former background. Feynman diagrams are made with TikZ-Feynman [64].

While less frequently produced than quarks and gluons, vector bosons are ubiquitous in proton collisions at 13 TeV. Since the W and Z bosons carry mass ~ 100 GeV, not far from the top quark mass, vector bosons decaying into quarks have better chances than mere

partons to be top-tagged. If there is a recoiling parton possessing considerable transverse momentum, the b -tagging algorithm always has a finite chance to identify it as a b-candidate jet. The lowest order hard-scattering matrix elements for such events are drawn in Figure B.2. Furthermore, if a parton radiated by the strong force is close to the vector boson, which decays hadronically, the three quarks resemble a hadronic top-quark jet.

The number of the vector boson plus jets ($V+\text{jets}$) events in the signal and template regions is estimated by MC simulation. SHERPA v2.2.5 [144] generates matrix elements of a vector boson with none to four additional partons at the leading order of the strong interaction. It also dresses the partons in the matrix element with showering partons having lower transverse momenta. The momenta of the initial state partons are drawn from Parton Distribution Function NNPDF3.0, the same as the PDF sets used for the $t\bar{t}$ simulation.

The $V+\text{jets}$ invariant distributions are compared with data in the “0 b-tag in”, 50%-working-point template region in Figure B.3(a). This category is selected rather than the “1 b-tag in”, as the latter has too few events to have numerous bins with small statistical errors. The lower panel shows that the fraction of $V+\text{jets}$ events in data stays below 3% for $M_{tb} < 1.8$ TeV and then jumps to around 5% for $M_{tb} > 2.4$ TeV. Figure B.3(b) shows the b -tagging ratio that carries the template region counts to the corresponding signal region. The ratio of $V+\text{jets}$ is about 40% higher than the data-driven estimation obtained from the control regions. Therefore, if the $V+\text{jets}$ event number in the signal region is part of the data-driven estimation, the data-driven background events will be underestimated by roughly 2% in the high mass regime. Since the systematic uncertainties of the data-driven estimation and the propagated $t\bar{t}$ modeling uncertainties are above 2% in high M_{tb} , the $V+\text{jets}$ contribution to the data-driven background estimate has been incorporated by the statistical analysis.

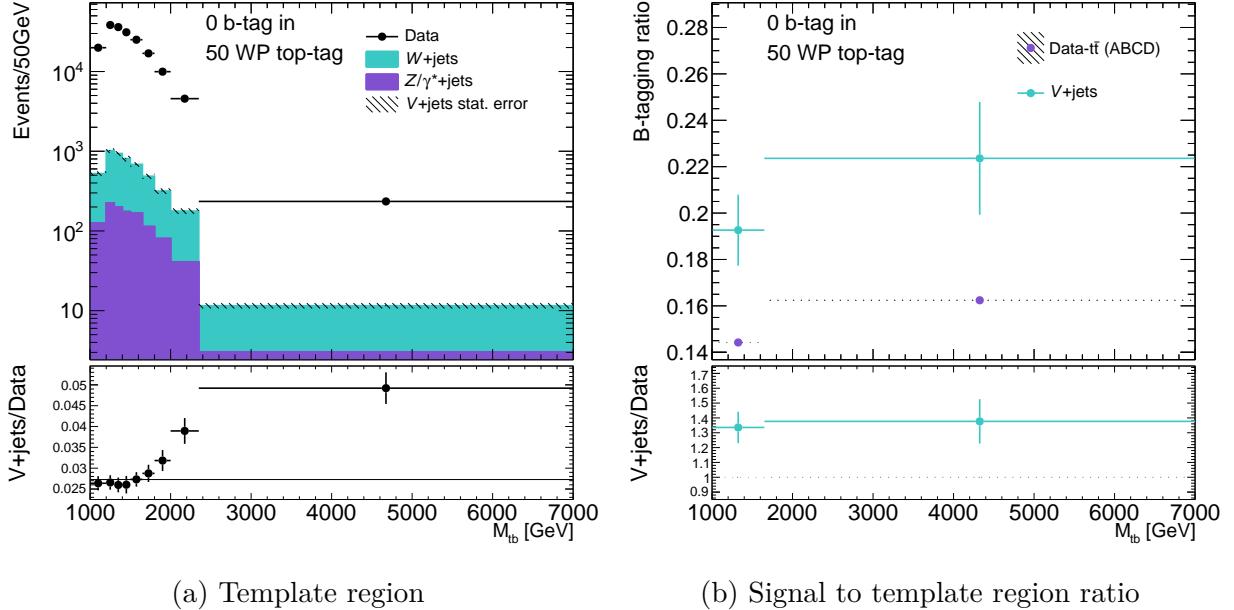


Figure B.3: (a) The invariant mass distribution of a large-radius top-candidate jet and a small-R b-candidate jet in one of the template regions. The template region has the data plotted with two types of V +jets distributions. The W +jets channel, which has a W boson and some high p_T partons in the final state, is stacked on top of the Z/γ^*+ jets channel, where the W boson is replaced with a neutral vector boson. The statistical uncertainties of the two channels are summed in quadrature into a single V +jets statistical error. The ratio of the V +jets to data is shown in the lower panel with the statistical errors of V +jets and data. There is a horizontal line given by the ratio between total area. (b) The b -tagging ratio – the ratio of the signal region to the template regions shown in (a) – is shown as a function of M_{tb} . Since simulated events of V +jets are scarce, the two channels are added together, and the number of bins is reduced. The b -tagging ratio obtained in the upper control region (top-tagging DNN $> e^{-4}$ while below the 80% working point) with $t\bar{t}$ -subtracted data is compared. The hatched area represent the latter’s statistical error, which is minuscule compared with the V +jets. The bottom panels shows the division of the two ratios with the statistical uncertainties propagated.

Diboson and Others

Instead of one vector boson, two vector bosons with quark decays – the diboson process – could end up in the signal regions, too. But these events have smaller production cross-sections than W -boson + jets or even the t -channel single-top background. These cross-sections have been measured by the ATLAS collaboration using data at 8 TeV of \sqrt{s} [145]. Following the same logic regarding the Wt associated production, the di-boson background comprises a tiny portion of the data-driven background estimate, which could be easily covered by systematic uncertainties. In contrast to the QCD multi-jet, $t\bar{t}$, V +jets, single-top, and diboson background, the ATLAS measurements show that the remaining SM background has even fewer events. In conclusion, the SM background events are well estimated by the data-driven method and the MC simulation of the $t\bar{t}$ events.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] S. Tomonaga. *Quantum mechanics. Translated from the Japanese by Koshiba*. Amsterdam: North-Holland Pub. Co., 1962 (cit. on p. 1).
- [2] H. Newman and T. Ypsilantis, eds. *History of Original Ideas and Basic Discoveries in Particle Physics*. New York: NATO ASI Series. Plenum Press, 1996 (cit. on p. 1).
- [3] CDF Collaboration. “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”. In: *Phys. Rev. Lett.* 74.14 (1995), p. 2626 (cit. on p. 2).
- [4] DØ Collaboration. “Observation of the Top Quark”. In: *Phys. Rev. Lett.* 74.14 (1005), p. 2632 (cit. on p. 2).
- [5] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39.5 (1977), p. 252 (cit. on p. 2).
- [6] UA1 Collaboration. “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV”. In: *Phys. Lett. B* 122 (1983), p. 103 (cit. on p. 2).
- [7] UA1 Collaboration. “Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider”. In: *Phys. Lett. B* 126 (1983), p. 398 (cit. on p. 2).
- [8] UA2 Collaboration. “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider”. In: *Phys. Lett. B* 122 (1983), p. 476 (cit. on p. 3).
- [9] UA2 Collaboration. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider”. In: *Phys. Lett. B* 129 (1983), p. 130 (cit. on p. 3).
- [10] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716.1 (2012), p. 1 (cit. on p. 3).
- [11] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett. B* 716.1 (2012), p. 30 (cit. on p. 3).
- [12] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics*. Vol. 1. CRC Press, 2012 (cit. on p. 3).

- [13] D. E. Morrissey et al. “Physics searches at the LHC”. In: *Physics Reports* 515 (2012), p. 1 (cit. on pp. 3, 4, 6, 12, 13).
- [14] G. Panico and A. Wulzer. *The Composite Nambu-Goldstone Higgs*. Springer International Publishing, 2016 (cit. on pp. 3, 4).
- [15] R. Barbieri. “Electroweak theory after the first Large Hadron Collider phase”. In: *Physica Scripta* T158 (2013), p. 014006 (cit. on p. 4).
- [16] J. H. Christenson et al. “Evidence for the 2π Decay of the K_2^0 Meson”. In: *Phys. Rev. Lett.* 13.4 (1964), p. 138 (cit. on p. 4).
- [17] M. Kobayashi and T. Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (1973), p. 652 (cit. on p. 4).
- [18] J. –E. Augustin et al. “Discovery of a Narrow Resonance in e^+e^- Annihilation”. In: *Phys. Rev. Lett.* 33.23 (1974), p. 1406 (cit. on p. 5).
- [19] J. J. Aubert et al. “Experimental Observation of a Heavy Particle J ”. In: *Phys. Rev. Lett.* 33.23 (1974), p. 1404 (cit. on p. 5).
- [20] P.A. Zyla et al. [Particle Data Group]. “Review of Particle Physics”. In: *Prog. Theor. Exp. Phys.* 083C01 (2020) (cit. on pp. 5, 6, 13, 20, 28, 30, 40, 43–45, 73, 76, 101, 143).
- [21] T. Ohl. “Drawing Feynman diagrams with *FX340-1* and *METAFONT*”. In: *Comput. Phys. Commun.* 90.2-3 (1995), p. 340 (cit. on pp. 5, 40).
- [22] Universität Zürich. *How to make Feynman diagrams in LaTeX*. URL: <https://wiki.physik.uzh.ch/cms/latex:feynman> (visited on 11/05/2020) (cit. on p. 5).
- [23] H. Georgi et al. “The un-unified standard model”. In: *Nucl. Phys. B* 331.3 (1990), p. 541 (cit. on p. 6).
- [24] D. J. Muller and S. Nandi. “Topflavor: a separate $SU(2)$ for the third family”. In: *Phys. Lett. B* 383.3 (1996), p. 345 (cit. on pp. 6, 14).
- [25] G. P. Salam. “Towards jetography”. In: *Eur. Phys. J. C* 67.3-4 (2010), p. 637 (cit. on pp. 6, 54, 82).
- [26] CDF Collaboration. “Search for the Production of Narrow $t\bar{b}$ Resonances in 1.96 fb^{-1} of $p\bar{p}$ Collisions at $\text{sqrt}s = 1.96\text{ TeV}$ ”. In: *Phys. Rev. Lett.* 103.4 (2009), p. 041801 (cit. on p. 6).
- [27] DØ Collaboration. “Search for $W' \rightarrow tb$ resonances with left- and right-handed couplings to fermions”. In: *Phys. Lett. B* 699.3 (2011), p. 145 (cit. on p. 6).

[28] ATLAS Collaboration. “Search for vector-boson resonances decaying to a top quark and bottom quark in the lepton plus jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 788 (2019), p. 347 (cit. on p. 6).

[29] CMS Collaboration. “Search for heavy resonances decaying to a top quark and a bottom quark in the lepton+jets final state in proton-proton collisions at 13 TeV”. In: *Phys. Lett. B* 777 (2018), p. 39 (cit. on p. 6).

[30] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge University Press, 2003 (cit. on pp. 10, 44).

[31] M. Perelstein. “Little Higgs models and their phenomenology”. In: *Progress in Particle and Nuclear Physics* 58.1 (2007), p. 247 (cit. on p. 13).

[32] L. Randall and R. Sundrum. “Large Mass Hierarchy from a Small Extra Dimension”. In: *Phys. Rev. Lett.* 83.17 (1999), p. 3370 (cit. on p. 13).

[33] K. Agashe et al. “LHC signals for warped electroweak charged gauge bosons”. In: *Phys. Rev. D* 80.7 (2009), p. 075007 (cit. on p. 13).

[34] R. N. Mohapatra and J. C. Pati. “Left-right gauge symmetry and an “isoconjugate” model of CP violation”. In: *Phys. Rev. D* 11.3 (1975), p. 566 (cit. on p. 14).

[35] Y. Nagashima. *Beyond the Standard Model of Elementary Particle Physics*. Wiley-VCH, 2012 (cit. on p. 14).

[36] E. Malkawi, T. Tait, and C.-P. Yuan. “A model of strong flavor dynamics for the top quark”. In: *Phys. Lett. B* 385.1-4 (1996), p. 304 (cit. on p. 14).

[37] L. Evans and P. Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001 (cit. on pp. 19, 20).

[38] *A Large Ion Collider Experiment*. URL: <http://aliceinfo.cern.ch/Public/Welcome.html> (visited on 04/28/2020) (cit. on p. 19).

[39] *ATLAS*. URL: <https://atlas.cern/> (visited on 04/28/2020) (cit. on p. 19).

[40] *LHCb - Large Hadron Collider beauty experiment*. URL: <https://lhcb-public.web.cern.ch/> (visited on 04/28/2020) (cit. on p. 19).

[41] *CMS*. URL: <https://cms.cern/> (visited on 04/28/2020) (cit. on p. 19).

[42] CERN. *The CERN accelerator complex*. 2008. URL: <https://cds.cern.ch/record/1260465> (visited on 04/28/2020) (cit. on p. 21).

[43] CERN. *CERN's accelerator complex*. URL: <https://home.cern/science/accelerators/accelerator-complex> (visited on 01/14/2021) (cit. on p. 20).

[44] *Luminosity Public Results Run 2*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (visited on 03/25/2020) (cit. on pp. 22, 23).

[45] ATLAS Collaboration. “Luminosity Determination in pp Collisions at $\sqrt{s} = 13$ TeV using the ATLAS Detector at the LHC”. In: *ATLAS-CONF-2019-021* (2019) (cit. on pp. 23, 94).

[46] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08003 (cit. on pp. 24, 29, 34).

[47] I. C. Brock and T. Schorner-Sadenius, eds. *Physics at the Terascale*. Wiley-VCH Verlag GmbH, 2011 (cit. on p. 25).

[48] A. Lechner. “Particle interactions with matter”. In: *CERN Yellow Rep. School Proc.* 5 (2018), p. 47 (cit. on p. 25).

[49] L. B. Navarro. “Alignment of the ATLAS Inner Detector in the LHC Run II”. In: *Proceedings of the 3rd Annual Large Hadron Collider Physics Conference ATL-PHYS-PROC-2015-190* (2015) (cit. on p. 27).

[50] ATLAS Collaboration. *A new subdetector for ATLAS*. URL: <https://home.cern/news/news/experiments/new-subdetector-atlas> (visited on 01/20/2021) (cit. on p. 28).

[51] ATLAS LAr project leader. *ATLAS Liquid Argon Calorimeter 2m prototype*. URL: <https://cds.cern.ch/record/2254084?ln=en> (visited on 05/01/2020) (cit. on p. 29).

[52] F. Ariztizabal et al. [RD-34]. “Construction and performance of an iron-scintillator hadron calorimeter with longitudinal tile configuration”. In: *Nucl. Instrum. Methods A* 349.2-3 (1994), p. 384 (cit. on p. 30).

[53] ATLAS Collaboration. “ATLAS detector and physics performance : Technical Design Report. Volume 1”. In: *CERN-LHCC-1999-014, ATLAS-TDR-014* (1999) (cit. on p. 31).

[54] ATLAS Collaboration. “In situ calibration of large- R jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector”. In: *Eur. Phys. J. C* 79.2 (2019), p. 135 (cit. on pp. 32, 57, 93).

[55] ATLAS Collaboration. *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. 2020. arXiv: 2007.02645 [hep-ex] (cit. on pp. 33, 64, 93).

[56] ATLAS Collaboration. *Performance of the ATLAS muon triggers in Run 2*. 2020. arXiv: 2004.13447 [hep-ex] (cit. on p. 33).

[57] ATLAS Collaboration. *ATLAS : a 25-year insider story of the LHC experiment*. World Scientific Publishing Co. Pte. Ltd., 2019 (cit. on pp. 35, 36).

[58] E. F. Eisenhandler. “ATLAS Level-1 Calorimeter Trigger Algorithms”. In: *ATL-DAQ-2004-011* (2004) (cit. on p. 36).

[59] ATLAS Collaboration. “Performance of the ATLAS trigger system in 2015”. In: *Eur. Phys. J. C* 77.5 (2017), p. 317 (cit. on p. 36).

[60] P. Sebastien. “The updated ATLAS Jet Trigger for the LHC Run II”. In: *Meeting of the Division of Particles and Fields*. American Physical Society. Ann Arbor, Michigan, Aug. 2015, p. 360 (cit. on p. 37).

[61] ATLAS Collaboration. “Trigger Menu in 2017”. In: *ATL-DAQ-PUB-2018-002* (2018) (cit. on pp. 37, 57).

[62] ATLAS Collaboration. “Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System”. In: *CERN-LHCC-2013-018, ATLAS-TDR-023* (2013) (cit. on p. 37).

[63] R. Achenbach. “The ATLAS Level-1 Calorimeter Trigger”. In: *Journal of Instrumentation* 3.03 (2008), P03001 (cit. on p. 37).

[64] J. P. Ellis. “TikZ-Feynman: Feynman diagrams with TikZ”. In: *Comput. Phys. Commun.* 210 (2017), p. 103 (cit. on pp. 41, 77, 142, 144).

[65] Z. Sullivan. *ZTOP Fully differential NLO s-channel and t-channel single-top-quark production*. URL: <http://www.hep.anl.gov/zack/ZTOP/ZTOP.html> (visited on 05/08/2020) (cit. on p. 42).

[66] Z. Sullivan. “Fully differential W' production and decay at next-to-leading order in QCD”. In: *Phys. Rev. D* 66.7 (2002), p. 075011 (cit. on p. 42).

[67] D. Duffty and Z. Sullivan. “Model independent reach for W' bosons at the LHC”. In: *Phys. Rev. D* 86.7 (2012), p. 075018 (cit. on p. 42).

[68] A. Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: *Eur. Phys. J. C* 75.3 (2015), p. 132 (cit. on p. 42).

[69] J. Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *J. Phys. G* 43.2 (2016), p. 023001 (cit. on pp. 42, 44).

[70] S. Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Phys. Rev. D* 93.3 (2016), p. 033006 (cit. on p. 42).

[71] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C* 75.5 (2015), p. 204 (cit. on p. 42).

[72] R. D. Ball et al. “Parton Distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040 (cit. on pp. 42, 78).

[73] G. Watt and R. S. Thorne. “Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs”. In: *JHEP* 08 (2012), p. 052 (cit. on p. 42).

[74] S. Carrazza et al. “A compression algorithm for the combination of PDF sets”. In: *Eur. Phys. J. C* 75.10 (2015), p. 474 (cit. on p. 42).

[75] S. Heim et al. “Next-to-leading order QCD corrections to s -channel single top quark production and decay at the LHC”. In: *Phys. Rev. D* 81.3 (2010), p. 034005 (cit. on pp. 43, 51).

[76] ATLAS Collaboration. “Measurement of the top quark mass in the $t\bar{t} \rightarrow \text{lepton+jets}$ channel from $\sqrt{s} = 8$ TeV ATLAS and combination with previous results”. In: *Eur. Phys. J. C* 79.3 (2019), p. 290 (cit. on p. 45).

[77] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079 (cit. on p. 48).

[78] G. Mahlon and S. Parke. “Angular correlations in top quark pair production and decay at hadron colliders”. In: *Phys. Rev. D* 53.9 (1996), p. 4886 (cit. on p. 51).

[79] T. Sjöstrand, S. Mrenna, and P. Skands. “A brief introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178.11 (2008), p. 852 (cit. on p. 51).

[80] ATLAS Collaboration. “Performance of top-quark and W -boson tagging with ATLAS in Run 2 of the LHC”. In: *Eur. Phys. J. C* 79.5 (2019), p. 375 (cit. on pp. 52, 55, 62, 63, 73).

[81] S. Agostinelli et al. “GEANT4 – a simulation toolkit”. In: *Nucl. Instrum. Methods A* 506.3 (2003), p. 250 (cit. on p. 52).

[82] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70.3 (2010), p. 823 (cit. on p. 52).

[83] ATLAS Collaboration. *athena - The ATLAS Experiment’s main offline software repository*. URL: <https://gitlab.cern.ch/atlas/athena> (visited on 05/19/2020) (cit. on p. 52).

[84] J. Campbell, J. Huston, and F. Krauss. *The black book of quantum chromodynamics : a primer for the LHC era*. Oxford University Press, 2018 (cit. on pp. 53, 76).

[85] M. Cacciari, G. Salam, and G. Soyez. “The anti- k_t jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063 (cit. on p. 54).

[86] S. Catani et al. “New clustering algorithm for multijet cross sections in e^+, e^- annihilation”. In: *Phys. Lett. B* 269.3-4 (1991), p. 432 (cit. on p. 54).

[87] S. D. Ellis and D. E. Soper. “Successive combination jet algorithm for hadron collisions”. In: *Phys. Rev. D* 48.7 (1993), p. 3160 (cit. on p. 54).

[88] S. Catani et al. “Longitudinally-invariant k_\perp -clustering algorithms for hadron-hadron collisions”. In: *Nucl. Phys. B* 406.1-2 (1993), p. 187 (cit. on p. 54).

[89] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *Eur. Phys. J. C* 77.7 (2017), p. 490 (cit. on p. 56).

[90] Z. Marshall. “Simulation of Pile-up in the ATLAS Experiment”. In: *Proceedings of CHEP2013* 513.2 (2014), p. 022024 (cit. on p. 56).

[91] D. Krohn, J. Thaler, and L.-T. Wang. “Jet trimming”. In: *JHEP* 02 (2010), p. 084 (cit. on p. 56).

[92] M. Cacciari, G. P. Salam, and G. Soyez. “The catchment area of jets”. In: *JHEP* 04 (2008), p. 005 (cit. on pp. 60, 64).

[93] ATLAS Collaboration. “Measurement of k_T splitting scales in $W \rightarrow \ell\nu$ events at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 73.5 (2013), p. 2432 (cit. on p. 59).

[94] J. Thaler and K. V. Tilburg. “Identifying boosted objects with N-subjettiness”. In: *JHEP* 03 (2011), p. 015 (cit. on p. 61).

[95] J. Thaler and K. V. Tilburg. “Maximizing boosted top identification by minimizing N-subjettiness”. In: *JHEP* 02 (2012), p. 093 (cit. on p. 61).

[96] A. J. Larkoski, D. Neill, and J. Thaler. “Jet shapes with the broadening axis”. In: *JHEP* 04 (2014), p. 017 (cit. on p. 62).

[97] A. J. Larkoski, I. Moult, and B. Nachman. “Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning”. In: *Physics Reports* 841 (2020), p. 1 (cit. on p. 63).

[98] P. Mehta et al. “A high-bias, low-variance introduction to Machine Learning for physicists”. In: *Physics Reports* 810 (2019), p. 1 (cit. on p. 63).

[99] D. E. Soper and M. Spannowsky. “Finding physics signals with shower deconstruction”. In: *Phys. Rev. D* 84.7 (2011), p. 074002 (cit. on p. 63).

[100] D. E. Soper and M. Spannowsky. “Finding top quarks with shower deconstruction”. In: *Phys. Rev. D* 87.5 (2013), p. 054012 (cit. on p. 63).

[101] ATLAS Collaboration. “Search for $W' \rightarrow tb$ decays in the hadronic final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 781 (2018), p. 327 (cit. on pp. 63, 67).

[102] ATLAS Collaboration. “Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data”. In: *ATL-PHYS-PUB-2020-017* (2020) (cit. on p. 63).

[103] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77.7 (2017), p. 466 (cit. on pp. 63, 64).

[104] ATLAS Collaboration. “Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”. In: *Eur. Phys. J. C* 76.11 (2016), p. 581 (cit. on p. 64).

[105] ATLAS Collaboration. “Jet global sequential corrections with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV”. In: *ATLAS-CONF-2015-002* (2015) (cit. on p. 65).

[106] ATLAS Collaboration. *Expected performance of the 2019 ATLAS b-taggers*. URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-005/> (visited on 06/24/2020) (cit. on p. 65).

[107] ATLAS Collaboration. “ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79.11 (2019), p. 970 (cit. on pp. 66, 67).

[108] ATLAS Collaboration. “Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC”. In: *Eur. Phys. J. C* 77.5 (2017), p. 332 (cit. on p. 66).

[109] ATLAS Collaboration. “Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV”. In: *ATL-PHYS-PUB-2015-026* (2015) (cit. on p. 66).

[110] ATLAS Collaboration. “Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run”. In: *ATL-PHYS-PUB-2017-013* (2017) (cit. on p. 67).

[111] ATLAS Collaboration. “Measurement of b -tagging Efficiency of c -jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector”. In: *ATLAS-CONF-2018-001* (2018) (cit. on p. 67).

[112] ATLAS Collaboration. “Calibration of light-flavour b -jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV”. In: *ATLAS-CONF-2018-006* (2018) (cit. on p. 67).

[113] M. L. Mangano and T. J. Stelzer. “TOOLS FOR THE SIMULATION OF HARD HADRONIC COLLISIONS”. In: *Annual Review of Nuclear and Particle Science* 55.1 (2005), p. 555 (cit. on p. 77).

[114] M. Cacciari et al. “Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation”. In: *Phys. Lett. B* 710.4-5 (2012), p. 612 (cit. on p. 78).

[115] M. Beneke et al. “Hadronic top-quark pair production with NNLL threshold resummation”. In: *Nucl. Phys. B* 855.3 (2012), p. 695 (cit. on p. 78).

[116] P. Baernreuther, M. Czakon, and A. Mitov. “Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$ ”. In: *Phys. Rev. Lett.* 109.13 (2012), p. 132001 (cit. on p. 78).

[117] M. Czakon and A. Mitov. “NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels”. In: *JHEP* 12 (2012), p. 054 (cit. on p. 78).

[118] M. Czakon and A. Mitov. “NNLO corrections to top-pair production at hadron colliders: the quark-gluon reaction”. In: *JHEP* 01 (2013), p. 080 (cit. on p. 78).

[119] M. Czakon, P. Fiedler, and A. Mitov. “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$ ”. In: *Phys. Rev. Lett.* 110.25 (2013), p. 252004 (cit. on p. 78).

[120] M. Czakon and A. Mitov. “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”. In: *Comput. Phys. Commun.* 185.11 (2014), p. 2930 (cit. on p. 78).

[121] P. Baernreuther. “Top Quark Pair Production at the LHC”. Diplom–Physiker [Thesis]. RWTH Aachen University, 2012 (cit. on p. 78).

[122] P. Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 11 (2004), p. 040 (cit. on p. 78).

[123] S. Frixione, P. Nason, and C. Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070 (cit. on p. 78).

[124] S. Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043 (cit. on pp. 78, 95).

[125] ATLAS Collaboration. “Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s}=13$ TeV”. In: *ATL-PHYS-PUB-2020-023* (2020) (cit. on pp. 79, 95, 111).

[126] ATLAS Collaboration. “Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s} = 13$ TeV pp data”. In: *ATLAS-CONF-2020-022* (2020) (cit. on p. 93).

[127] ATLAS Collaboration. “Simulation-based extrapolation of b -tagging calibrations towards high transverse momenta in the ATLAS experiment”. In: *ATL-PHYS-PUB-2021-003* (2021) (cit. on pp. 94, 112).

[128] G. Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *Journal of Instrumentation* 13.7 (2018), P07017 (cit. on p. 94).

[129] ATLAS Collaboration. “Determination of the parton distribution functions of the proton from ATLAS measurements of differential W^\pm and Z boson production in association with jets”. In: *ATLAS-CONF-2020-057* (2020) (cit. on p. 95).

[130] S. Frixione and B. R. Webber. “Matching NLO QCD computations and parton shower simulations”. In: *JHEP* 06 (2002), p. 029 (cit. on p. 95).

[131] M. Bahr et al. “Herwig++ physics and manual”. In: *Eur. Phys. J. C* 58.4 (2008), p. 639 (cit. on p. 96).

[132] J. Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196 (cit. on p. 96).

[133] G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71.2 (2011), p. 1554 (cit. on p. 99).

[134] A. L. Read. “Presentation of search results: the CL_s technique”. In: *J. Phys. G* 28.10 (2002), p. 2693 (cit. on p. 99).

[135] E. Gross. “Practical Statistics for High Energy Physics”. In: *Proceedings of the 2015 European School of High-Energy Physics* 4 (2017), p. 165 (cit. on p. 99).

[136] R. D. Cousins, J. T. Linnemann, and J. Tucker. “Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process”. In: *Nucl. Instrum. Methods A* 595.2 (2008), p. 480 (cit. on p. 120).

[137] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79.8 (2019), p. 639 (cit. on p. 139).

[138] ATLAS Collaboration. “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data”. In: *ATLAS-CONF-2016-024* (2016) (cit. on p. 139).

[139] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 76.5 (2016), p. 292 (cit. on p. 140).

[140] M. Aliev et al. “HATHOR - HAdronic Top and Heavy quarks crOss section calculaToR”. In: *Comput. Phys. Commun.* 182.4 (2011), p. 1034 (cit. on p. 141).

[141] P. Kant et al. “HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions”. In: *Comput. Phys. Commun.* 191 (2015), p. 74 (cit. on p. 141).

[142] N. Kidonakis. “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^- ”. In: *Phys. Rev. D* 82.5 (2010), p. 054018 (cit. on p. 143).

[143] N. Kidonakis. *Top Quark Production*. 2013. arXiv: 1311 . 0283 [hep-ph] (cit. on p. 143).

[144] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007 (cit. on p. 145).

[145] ATLAS Collaboration. “Standard Model Summary Plots Spring 2020”. In: *ATL-PHYS-PUB-2020-010* (2020) (cit. on p. 147).