

University of California
Santa Barbara

A search for R -parity violating supersymmetry at the 13 TeV LHC

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Rohan Bhandari

Committee in charge:

Professor David Stuart, Chair
Professor Harry Nelson
Professor Nathaniel Craig

September 2018



The Dissertation of Rohan Bhandari is approved.

Professor Harry Nelson

Professor Nathaniel Craig

Professor David Stuart, Committee Chair

May 2018

A search for R -parity violating supersymmetry at the 13 TeV LHC

Copyright © 2018

by

Rohan Bhandari

To my parents, Ramesh and Savita

Acknowledgements

Since my childhood, I have always been fascinated by big questions about the universe, and I feel blessed to have had the opportunity to immerse myself in them for the past five years. This dissertation encapsulates my career as a particle physicist and my contribution to answering those universal questions. I, however, could not have done this without many helping hands to whom I owe a great deal.

First, I must thank the RichStu group—my HEX family. In particular, I want to thank my advisor, David Stuart, for always being a fantastic mentor, and Jeff Richman, my de facto second advisor, for his extraordinary teaching. There is not enough space here to properly thank you for all your guidance and wisdom. I would not be the physicist I am today without it. To the postdocs—Ana Ovcharova, Chris West, Jae Hyeok Yoo, Manuel Franco Sevilla, and Matt Citron—and to my fellow graduate students in arms—Adam Dishaw, Alex Dorsett, Jack Bradmiller-Feld, and Ryan Heller—it was truly a pleasure to have gone through this Ph.D. with all of you. I could not have asked for a better group of colleagues and friends. You made grad school fun, and I’ll always have fond memories of the hour-long impromptu discussions we found ourselves in, whether we were debating physics, the news, or, far more often, color themes.

I also owe much gratitude to my parents, Ramesh and Savita, without whom none of this would be possible. It was only through their endless love and support and countless sacrifices that I was able to pursue my dreams. Included in this is my brother, Simit, who was always willing to make time for me and share his advice.

Finally, I would like to thank my partner, Mallorie Chase, for all her support over the last few years and for moving with me to Geneva for a year-long adventure. I am so thankful to have met you during my time here, and I am excited to tackle our next chapter in life together.

Curriculum Vitæ

Rohan Bhandari

Education

- 2018 Ph.D. in Physics, University of California, Santa Barbara.
2016 M.A. in Physics, University of California, Santa Barbara.
2013 B.A. in Physics, Columbia University.

Selected Publications

Search for R-parity violating supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV using b jets in a final state with a single lepton, many jets, and high sum of large-radius jet masses, CMS Collaboration, Physics Letters B (2018), arXiv:1712.08920

Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large-radius jets, CMS Collaboration, Physics Review Letters (2017), arXiv:1705.04673

Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large-radius jets, CMS Collaboration, Journal of High Energy Physics (2016), arXiv:1605.04608

Search for nonpointing photons in the diphoton and E_T^{miss} final state in $\sqrt{s} = 7$ TeV proton-proton collisions using the ATLAS detector, ATLAS Collaboration, Physical Review D (2013), arXiv:1304.6310

Abstract

A search for R -parity violating supersymmetry at the 13 TeV LHC

by

Rohan Bhandari

This dissertation describes a search for R -parity violating supersymmetry, motivated by the stringent limits set on R -parity conserving models from Run I and Run II of the LHC. These limits have excluded gluino masses up to approximately 2 TeV in mass, which is the rough scale expected for supersymmetry to “naturally” solve the Hierarchy Problem. These constraints, however, can be evaded by considering R -parity violating models, in which the lightest supersymmetric particle can decay to Standard Model particles and does not produce a large missing transverse momentum signature.

To avoid conflicts with experimental measurements, such as proton decay, the framework of Minimal Flavor Violation is applied, resulting in the largest R -parity violating coupling being between a top, bottom, and strange quark. Therefore, this search uses the pair production of gluinos that decay via $\tilde{g} \rightarrow t\bar{t} \rightarrow tbs$ as a benchmark model and generically looks for new physics with a signature of a single lepton, large jet and bottom quark jet multiplicities, and high sum of large-radius jet masses, without any requirement on the missing transverse momentum in an event.

The search is conducted with 35.9 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ proton-proton collisions collected by the CMS experiment in 2016. The background is estimated through a maximum-likelihood fit of the N_b distribution across bins of jet multiplicity and sum of large-radius jet masses. No evidence of new physics is observed, and limits on a simplified model, in which gluinos decay promptly via $\tilde{g} \rightarrow tbs$, are set, excluding gluino masses below 1610 GeV at the 95% confidence level.

Contents

Curriculum Vitae	vi
Abstract	vii
List of Figures	xi
List of Tables	xviii
0 Introduction for Non-physicists	1
1 Introduction	5
Part I Theoretical Context and Motivations	7
2 The Standard Model	8
2.1 Overview and Successes	8
2.2 The Standard Model as an Incomplete Theory	11
2.2.1 The Hierarchy Problem and Naturalness	12
3 Supersymmetry	15
3.1 Natural Supersymmetry	15
3.2 Phenomenological and Experimental Constraints	17
3.3 R -parity Violating Supersymmetry	22
3.3.1 Minimal Flavor Violating Supersymmetry	23
Part II Experimental Apparatus	26
4 Experimental Apparatus and Event Reconstruction	27
4.1 The Large Hadron Collider	27
4.2 Compact Muon Solenoid	30

4.2.1	Inner Tracking System	34
4.2.2	Electromagnetic Calorimeter	35
4.2.3	Hadronic Calorimeter	37
4.2.4	Muon System	39
4.2.5	Trigger System	40
5	Particle Reconstruction and Identification	43
5.1	Tracks	43
5.1.1	Vertices	44
5.2	Calorimeter Clusters	45
5.3	Particle Flow	46
5.3.1	Linking	46
5.3.2	PF Reconstruction and Identification	48
5.4	Leptons	49
5.4.1	Electrons	49
5.4.2	Muons	51
5.5	Jets	52
5.5.1	Clustering	53
5.5.2	Selection	54
5.5.3	b-tagging	55
5.6	Large-radius Jets	58
5.6.1	M_J — The Sum of Large-radius Jet Masses	59
Part III	The Search for SUSY	62
6	Data Samples and Simulation	63
6.1	Data	63
6.2	Monte Carlo Simulation	63
6.2.1	Background Samples	66
6.2.2	Signal Samples	66
7	Event Selection	70
7.1	Baseline Selection	70
7.2	Trigger Efficiency	73
7.3	Analysis Binning	75
8	Background Prediction	80
8.1	Overview	80
8.2	$t\bar{t}$ and QCD Normalizations	81
8.2.1	M_J Connection	83
8.3	W + jets Normalization	85
8.4	Other Normalization	85

9	Systematic Uncertainties	87
9.1	Gluon Splitting Rate	87
9.2	b-tagging Data-to-simulation Scale Factors	92
9.3	Lepton Fake Rate in QCD	96
9.4	Additional Systematic Uncertainties	97
9.4.1	N_b Distribution With NLO Precision	99
9.5	Signal Systematics	100
10	Fit Model and Validation	102
10.1	Description of the Fit Model	102
10.2	Validation of the Fit Model	104
10.2.1	Signal Injection Study	105
10.2.2	Control Region Fit	108
11	Results and Interpretation	112
11.1	Examination of Pre-fit Data	112
11.2	Results	114
11.2.1	Background-only Fit Results	114
11.2.2	Signal-plus-background Fit Results	120
11.3	Statistical Interpretation	123
11.4	Limits on the T1tbs Benchmark Model	125
12	Summary and Conclusions	126
Part IV	Appendix	128
A	DeepJEC: Deep Learning the Jet Response	129
A.1	Introduction	129
A.2	Jet Images	133
A.3	Network Architecture	136
A.4	Results	137
A.5	Conclusions	141
Bibliography		143

List of Figures

0.1	The discovery of the Higgs boson made the front page of many newspapers, including the New York Times [1].	3
2.1	The fundamental particles of the Standard Model and some of their properties.	9
2.2	The theoretical and experimental value for the production cross section of various processes [2].	10
2.3	Evidence for the existence of a 125 GeV Higgs boson in the di-photon channel (left) and in the four-lepton channel (right) [3, 4].	13
2.4	The fermionic one-loop Higgs mass correction.	14
3.1	The scalar one-loop Higgs mass correction.	17
3.2	An example spectrum for natural SUSY [5].	19
3.3	An overview of recent results from SUSY searches from the Compact Muon Solenoid experiment [6].	20
3.4	An overview of recent results from SUSY searches from the A Toroidal LHC Apparatus experiment [7].	21
3.5	Example diagram representing proton decay, where the particle X represents a down-type squark [8]. The left vertex is mediated by the baryon number violating coupling, λ'' , and the right vertex is mediated by the lepton number violating coupling, λ'	23
3.6	Example diagram of the pair production of gluinos that decay via $\tilde{g} \rightarrow t\bar{t} \rightarrow tbs$	25
4.1	A schematic of the CERN LHC accelerator complex [9].	28
4.2	A map that shows the physical extent of the Large Hadron Collider. . . .	29
4.3	Ratio of parton luminosities at $\sqrt{s} = 13$ and 8 TeV [10].	30
4.4	The CMS detector [11].	32
4.5	A diagram showing the various sub-detectors of the CMS detector [12]. .	33
4.6	A diagram of the cartesian and cylindrical coordinate systems used by CMS [13].	33

4.7	Layout of the CMS tracking system, showing both the pixel detector (blue) and the strips detector (red) [14].	35
4.8	A cross section of the ECAL, showing its geometry and layout [15]. . . .	36
4.9	The layout and geometry of a quarter of the HCAL detector [16].	37
4.10	The layout of the muon system within the CMS detector [17].	39
4.11	Flowchart depicting the generation of a L1 Accept [18].	42
5.1	The efficiency to select an analysis-level electron as a function of p_T and η . The low efficiency for $1.442 \leq \eta < 1.556$ corresponds to the ECAL “crack” region, the boundary between the EB and EE, in which electron reconstruction is particularly difficult.	51
5.2	The efficiency to select an analysis-level muon as a function of p_T and η	53
5.3	The distribution of the CSVv2 discriminator values for jets of different flavors. Jets are selected from $t\bar{t}$ events and required to have $p_T > 20$ GeV [19].	57
5.4	The efficiency of the CSVv2 algorithm as a function of jet p_T at the working point used in this analysis.	58
5.5	Distributions of M_J , normalized to the same area, for $t\bar{t}$ events (solid blue) and signal events with $m_{\tilde{g}} = 1000$ GeV (dashed red) and 1600 GeV (solid red) in a selection of $N_{\text{leps}} = 1$, $H_T > 1200$ GeV, $N_{\text{jets}} \geq 8$, $M_J > 500$ GeV, and $N_b \geq 1$	61
6.1	Delivered and recorded integrated luminosity by the LHC and CMS, respectively, over 2016 [20].	64
6.2	Example diagram of the T1tbs simplified model.	69
7.1	The N-1 plots for N_{leps} (top-left), H_T (top-right), M_J (middle-left), N_{jets} (middle-right), and N_b (bottom). The black dashed vertical line represents the value of the corresponding baseline requirement.	71
7.2	Trigger efficiency for HLT_PFHT900 as a function of H_T in Runs B-G (top-left), Run H (top-right), and full dataset (bottom). The efficiencies are measured using a data sample collected with the HLT_Ele27_WPTight trigger and an offline requirement of at least one electron and at least four jets.	74
7.3	Trigger efficiency as a function of H_T for HLT_PFJet450 in the full dataset (top-left) and for the combination of HLT_PFHT900 and HLT_PFJet450 in Run H (top-right) and the full dataset (bottom). The efficiencies are measured using a data sample collected with the HLT_Ele27_WPTight trigger and an offline requirement of at least one electron and at least four jets.	76
7.4	Illustration depicting the N_{jets} , M_J binning after the baseline selection, with control and signal region bins denoted by “CR” and “SR”, respectively.	77

7.5	The simulated N_b distribution for background and signal processes in the control region bins. The top-left plot corresponds to the $4 \leq N_{\text{jets}} \leq 5$, $500 < M_J \leq 800$ GeV bin, the top-right plot to the $5 \leq N_{\text{jets}} \leq 6$, $M_J > 800$ GeV bin, and the bottom plot to the $6 \leq N_{\text{jets}} \leq 7$, $500 < M_J \leq 800$ GeV bin.	78
7.6	The simulated N_b distribution for background and signal processes in the signal region bins. The top-left plot corresponds to the $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV bin, the top-right plot to the $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV bin, the middle-left plot to the $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV bin, the middle-right plot to the $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV, and the bottom plot to the $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV bin.	79
8.1	Diagram depicting the N_{jets} , M_J binning of the analysis including the $N_{\text{leps}} = 0$ QCD control region. The QCD control region bins are denoted by “QCD CR”, while the $N_{\text{leps}} = 1$ control and signal region bins are denoted by “CR” and “SR”, respectively. The pairing structure between lepton bins is indicated by arrows for the $500 < M_J \leq 800$ region and is the same for other M_J bins.	82
8.2	Distribution of post-fit yields of $t\bar{t}$ in the $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV bin for 1000 psuedodata experiments without (left) and with (right) constraints between adjacent M_J bins. The dotted black line indicates the pre-fit yield.	84
8.3	Data-to-simulation ratios as a function of M_J for different N_{jets} bins (data points) with a selection of $N_{\text{leps}} = 1$, $H_T > 1200$ GeV, and $N_b = 1$ applied. The shaded region corresponds to the size of the M_J connection in each M_J bin.	84
8.4	Jet multiplicity distribution for data and simulation in a Z + jets control sample selected by requiring $N_{\text{leps}} = 2$, $H_T > 1200$ GeV, $M_J > 500$ GeV, $N_b = 1$, and $80 < m_{\ell\ell} < 100$ GeV. The total yield from simulation is normalized to the number of events in data. The uncertainty in the ratio of data to simulation yields (lower panel) is statistical only.	86
9.1	The $\Delta R_{b\bar{b}}$ distribution shapes for the three gluon splitting categories: Events with a pair of b-tagged jets resulting from gluon splitting (green), events with a gluon splitting yielding fewer than 2 b-tagged jets (blue), and events without a gluon splitting to $b\bar{b}$. These events are selected by requiring $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 4$, and $N_b = 2$	88
9.2	The relative fraction of the possible final states that occur from gluon splitting to $b\bar{b}$ for events satisfying $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 4$, and $N_b = 2$	89
9.3	Post-fit $\Delta R_{b\bar{b}}$ distributions in a selection with $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 4$, and $N_b = 2$ with the post-fit uncertainty represented by a hatched band. The ratio of data to simulation yields is shown in the lower panel.	91

9.4	Effect of the ± 1 s.d. variations of the gluon splitting rate on the N_b distribution in $t\bar{t}$ events for the two most sensitive bins: $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right). Event yields are normalized to that expected in 35.9 fb^{-1} of data.	91
9.5	The data-to-simulation scale factors for the tagging efficiency of b-flavor jets (top-left), charm-flavor jets (top-right), and light-flavor or gluon jets (bottom) are shown as a function of jet p_T . The associated uncertainty with each scale factor is shown as a blue hashed band.	94
9.6	Effect of the ± 1 s.d. correlated variations of the b-flavor and c-flavor jet data-to-simulation scale factors on the N_b distribution in $t\bar{t}$ for the two most sensitive bins: $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right). Event yields are normalized to that expected in 35.9 fb^{-1} of data.	95
9.7	Effect of the ± 1 s.d. variations of the light-flavor jet data-to-simulation scale factors on the N_b distribution in $t\bar{t}$ for the two most sensitive bins: $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right). Event yields are normalized to that expected in 35.9 fb^{-1} of data.	95
9.8	The relative isolation distribution for electrons (left) and muons (right) in the analysis control region. The binning of the histograms is chosen such that the first bin corresponds to the relative isolation requirement for signal leptons (0.1 for electrons and 0.2 for muons). The normalizations of the QCD, $t\bar{t}$, and $W + \text{jets}$ processes are scaled to match the results of a control region fit described in Section 10.2.2.	97
9.9	Background systematic uncertainties (in percent) for the $N_{\text{jets}} \geq 8$, $500 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right) bins. The bottom row shows the total uncertainty for a given N_b bin by summing in quadrature all uncertainties. These values are similar for other bins. . . .	98
9.10	Comparison of the N_b distribution for a sample generated at LO precision with MADGRAPH5_aMC@NLO 2.2.2 (blue histogram) with that of one generated at NLO precision with POWHEG 2.0 (data points). The comparison is done after the baseline selection and in bins of $4 \leq N_{\text{jets}} \leq 5$ (left), $6 \leq N_{\text{jets}} \leq 7$ (middle), and $N_{\text{jets}} \geq 8$ (right). In order to evaluate only shape differences, the MADGRAPH5_aMC@NLO sample is normalized to match the normalization of the POWHEG sample.	99
9.11	Systematic uncertainties (in percent) for a $m_{\tilde{g}} = 1600$ GeV signal and for the $N_{\text{jets}} \geq 8$, $500 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right) bins. The bottom row shows the total uncertainty for a given N_b bin by summing in quadrature all uncertainties. These values are similar for other bins.	101

10.1	Median extracted signal strength of 1000 psuedodata experiments as a function of gluino mass. The uncertainties drawn are the median upper and lower errors of the fitted signal strengths per mass point.	106
10.2	Distribution of the fraction of signal extracted from 1000 psuedodata experiments for a 2000 GeV gluino at 1x (top-left), 3x (top-right), 5x (bottom-left), and 10x (bottom-right) the nominal cross section. The last bin includes the overflow contents, and the black dashed line represents extracting as much signal as was injected. The median extracted signal is 78%, 92%, 95%, and 98% the injected signal, respectively.	107
10.3	Post-fit N_b distributions of the control region fit with only statistical uncertainties shown.	109
10.4	Post-fit pulls of the background-only control region fit. The post-fit values of the nuisance parameters are indicated by data points, while the post-fit uncertainty is shown as a black line and is normalized by the pre-fit uncertainty depicted as the blue band.	110
10.5	Pre-fit N_b distributions of the control region bins with the background simulation normalized to the observed data yields with a single scaling factor. The ratio of data-to-simulation is shown in the lower panel. The pre-fit uncertainty is represented by the hatched band.	111
11.1	Pre-fit N_b distributions of the signal region bins with the background simulation normalized to the observed data yields with a single scaling factor: $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV (upper-left), $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV (upper-right), $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV (middle-left), $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (middle-right), and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (bottom-middle). The ratio of data-to-simulation is shown in the lower panel. The pre-fit uncertainty is represented by the hatched band.	113
11.2	Data and the background-only post-fit N_b distribution for the control region bins: $4 \leq N_{\text{jets}} \leq 5$, $500 < M_J \leq 800$ GeV (left), $4 \leq N_{\text{jets}} \leq 5$, $M_J > 800$ GeV (middle), and $6 \leq N_{\text{jets}} \leq 7$, $500 < M_J \leq 800$ GeV (right). The expected signal distribution is also shown for a gluino mass of 1600 GeV (red line). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band. . . .	116
11.3	Data and the background-only post-fit N_b distribution for the signal region bins: $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV (upper-left), $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV (upper-middle), $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (upper-right), $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV (bottom-left), and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (bottom-right). The expected signal distribution is also shown for a gluino mass of 1600 GeV (red line). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.	117

11.4	Post-fit pulls of the background-only fit. The post-fit values of the nuisance parameters are indicated by data points, while the post-fit uncertainty is shown as a black line and is normalized by the pre-fit uncertainty depicted as the blue band.	118
11.5	Data and the $m_{\tilde{g}} = 1600$ GeV signal-plus-background post-fit N_b distribution for the control region bins: $4 \leq N_{\text{jets}} \leq 5$, $500 < M_J \leq 800$ GeV (left), $4 \leq N_{\text{jets}} \leq 5$, $M_J > 800$ GeV (right), and $6 \leq N_{\text{jets}} \leq 7$, $500 < M_J \leq 800$ GeV (middle). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.	120
11.6	Data and the $m_{\tilde{g}} = 1600$ GeV signal-plus-background post-fit N_b distribution for the signal region bins: $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV (upper-left), $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV (upper-middle), $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (upper-right), $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV (bottom-left), and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (bottom-right). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.	121
11.7	Cross section upper limits at 95% CL for a model of gluino pair production with $\tilde{g} \rightarrow t\bar{b}s$ compared to the gluino pair production cross section. The theoretical uncertainties in the cross section are shown as a band around the red line [21]. The expected limits (dashed line) and their ± 1 s.d. and ± 2 s.d. variations are shown as green and yellow bands, respectively. The observed limit is shown by the solid line with dots.	125
A.1	The (planned) integrated luminosity delivered and operating center-of-mass energy for the LHC from 2010-2023.	130
A.2	The mean jet response as a function of jet p_T (left) and η (right).	132
A.3	The true jet response distribution (blue) and a model of the response as a function of jet p_T and η (green).	132
A.4	An example jet image.	134
A.5	The example jet image from Figure A.4 split by particle type: charged hadrons (left), neutral hadrons (right), and photons + electrons + muons (right).	135
A.6	Average unnormalized jet image for charged hadrons (left), neutral hadrons (right), and photons + electron + muons (right).	136
A.7	Schematic of the convolutional neural network architecture.	138
A.8	The true jet response distribution (blue), a model of the response trained only on jet p_T and η (green), and a model of the response additionally trained with jet images (orange).	139

A.9	The mean jet response as a function of the fractional jet energy from charged hadrons (left), neutral hadrons (middle), and photons + electrons + muons (right). The truth-level distributions are shown in blue, while the predictions from the p_T, η model and p_T, η + image model are shown in green and orange, respectively.	140
A.10	The true jet response distribution (blue), model of the response trained on jet p_T , η , and jet images (orange), and a model trained on jet p_T , η , and jet fragmentation information (purple).	141

List of Tables

2.1	The theoretical [22] and experimental [23, 24] values of the anomalous electron magnetic moment, a_e . The uncertainty in the last digits is shown in parantheses.	10
3.1	The additional SUSY particles in the Minimal Supersymmetric Standard Model.	16
3.2	Rough estimates for the sizes of the λ''_{ijk} MFV RPV couplings [25].	24
5.1	Identification criteria that a PF electron must pass in order to be considered an analysis-level electron.	50
5.2	Identification criteria that a PF muon must pass in order to be considered an analysis-level muon.	52
5.3	Identification criteria that a jet candidate must pass in order to be considered an analysis-level jet.	55
6.1	Data samples analyzed for this analysis. The corresponding integrated luminosity is 35.9 fb^{-1}	65
6.2	Simulated background samples used in this analysis with their corresponding sample size and equivalent luminosity.	67
6.3	Simulated signal samples used in this analysis with their corresponding sample size and equivalent luminosity.	69
7.1	Expected yields in 35 fb^{-1} from simulations of SM and signal processes. Rows above the horizontal line correspond to requirements in the baseline selection, while those below correspond to additional kinematic cuts. . . .	72
9.1	Gluon splitting weights derived in the nominal fit, a variation with a requirement of $M_J > 800 \text{ GeV}$, and 4 variations in bins of N_{jets} (with the nominal $M_J > 500 \text{ GeV}$ requirement.)	90
9.2	Comparison of the relative isolation distributions, as described in the caption of Figure 9.8, for electrons and muons between QCD and data with contributions from “all other” ($t\bar{t}$, $W + \text{jets}$, and Other) subtracted. . . .	97

9.3	The signal efficiency of the most sensitive bin $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV for a 1600 GeV gluino in various bins of the number of truth-level primary vertices.	100
10.1	Actual coverage probability of the 95% confidence interval of the fit for the mass points with a biased signal extraction.	108
10.2	Table comparing the post-fit normalizations of the control region fit to the pre-fit yields for the various background processes.	109
10.3	Table of post-fit pulls of the background-only and signal-plus-background control region fit. The last column, $\rho(\theta_m, \mu)$, lists the correlation between the corresponding nuisance parameter, θ_m , and the nuisance parameter controlling the signal strength, μ	111
11.1	Pre-fit data and simulation yields in the control region bins in 35.9 fb^{-1} of data.	114
11.2	Pre-fit data and simulation yields in the signal region bins in 35.9 fb^{-1} of data.	115
11.3	Post-fit yields of the background-only fit, observed data, and expected yields for $m_{\tilde{g}} = 1600$ GeV.	119
11.4	Table of post-fit pulls of the background-only and signal-plus-background fit. The last column, $\rho(\theta_m, \mu)$, lists the correlation between the corresponding nuisance parameter, θ_m , and the nuisance parameter controlling the signal strength, μ	122

Chapter 0

Introduction for Non-physicists

One of the primary tenets of science is to communicate results to as wide an audience as possible. But in this regard, dissertations can be funny things, as they tend to have the smallest audiences. For physicists, the results that make up this dissertation either A) have been already communicated to the physics community through publications in journals and/or B) will be outdated within a year or two (such is progress). For everyone else, I am sure it's just incomprehensible.

This introduction is my attempt to include the latter group and share with others, particularly family and friends, what I have been doing for the last five years. So bear with me, as I try to condense the next ~ 100 pages into this short introduction. Here goes...

The job of physicists is to model how the universe works. For particle physicists, in particular, we do this by understanding how fundamental particles interact with each other. This provides really profound insights into the universe because these particles and their interactions are the building blocks for the universe. A “fundamental particle” is any particle that is not made up of smaller particles. For example, neutrons and protons are not fundamental particles, as they are each composed of three “quarks”, but an electron, on the other hand, is fundamental, as it is not made up of any smaller particles.

There are two groups of fundamental particles: fermions and bosons. Fermions are the particles that make up matter and combine to form protons and neutrons (as mentioned above), which lead to atoms, molecules, etc. Bosons are the “force carrier” particles. They allow the particles to interact with each other and are responsible for transferring forces between them. For example, the photon mediates the electromagnetic force, i.e. electricity and magnetism. The gluon is responsible for the “strong force”, which is what holds nuclei together, and the W and Z bosons carry the “weak force”, which is what causes atoms to decay. In total there are 12 fermions and 5 bosons, shown in Figure 2.1, that have been discovered so far.

Additionally, there are rules for which particles can interact with which particles and through which forces and with what strength, etc. These rules are what particle physicists work to study and attempt to model. Our best understanding of these rules has been combined into the Standard Model of Particle Physics. The Standard Model is one of the greatest achievements of science. It has been able to predict and explain a wide variety of phenomena and has withstood a number of stress tests. To this day, it is the most precisely tested theory in physics, if not science as a whole.

It wasn’t until 2012, however, after almost 50 years of searching, that the Standard Model was completed with the discovery of the Higgs boson. The Higgs boson was initially theorized in the 1960’s and is the particle responsible for explaining why things have mass—very succinctly, the amount a particle interacts with the Higgs boson is proportional to the amount of mass it has. This was a profound discovery, and it showcased the incredible predictive power of the Standard Model, resulting in a Nobel prize and even making it to the front page of the New York Times, as shown in Figure 0.1.

There was, however, a downside to this discovery – the Standard Model was completed. Despite all the successes of the Standard Model, there are certain things we simply cannot explain with it. For example, you may have noticed earlier that when I



Figure 0.1: The discovery of the Higgs boson made the front page of many newspapers, including the New York Times [1].

described the fundamental forces, I did not mention gravity. This is because it turns out it is really, really difficult to combine the Standard Model and Einstein's theory of general relativity, which is our best description of gravity. Furthermore, from astronomical measurements, we know there is something called Dark Matter and that it comprises 4x more of the matter in the universe than do the Standard Model particles, but we can't really say much more about it. What this means is that the Standard Model cannot tell the full story of how the universe behaves and that there must be more undiscovered particles to fill in the gaps.

One way to get hints for what this new physics may look like is by examining where the Standard Model breaks down, as this would presumably be where new particles are needed to correct the theory. When you do the mathematical calculation of the Higgs boson mass, you get something close to infinity, while the measured mass of the Higgs boson is $125 \text{ GeV}/c^2$ (the units are unimportant here). Clearly, there is a discrepancy. One way to reconcile this difference with the Standard Model is if two independent parameters were perfectly tuned by nature to cancel at over 30 decimal places. This

corresponds to two random numbers agreeing at the trillionth trillionth millionth decimal place. Is it possible? Sure, but it's definitely not likely.

An alternate solution to this discrepancy comes from the theory of Supersymmetry, which extends the Standard Model, by positing that for every fermion there is a “superpartner” particle that is a boson and vice versa. This symmetry between fermions and bosons results in the Higgs mass calculation naturally being around $\sim 100 \text{ GeV}/c^2$. Additionally, in certain types of Supersymmetry, the additional particles may actually be what make up Dark Matter and there may even be a coherent framework for unifying gravity with the three other fundamental forces. Thus, Supersymmetry is a really appealing idea that seems like it could solve some of the biggest open questions in particle physics.

Nature, however, owes us nothing, and a beautiful idea is just that until proven otherwise. So what my dissertation boils down to is a detailed description of my search for Supersymmetry. Unfortunately, as of the writing of this, no evidence for Supersymmetry, by my search or any other, has been observed. But, this is an exploration, and as we all know, it's always the last place you look!

Chapter 1

Introduction

The discovery of the Higgs boson at the end of Run I of the Large Hadron Collider marked a transition in high energy physics from asking questions like “At what mass will we find the Higgs boson?” to “Something needs to be out there, but what?”. The different scopes of these questions reflect the different strategies of their respective search programs. For the Higgs bosons, its couplings and branching fractions as a function of mass were already well described by the Standard Model, and from these expectations, one could formulate a targeted, multi-channel approach for discovery. On the other hand, search strategies are now driven by motivated but still theoretical models and the goal is to cast as wide a net as possible in the hopes of finding hints of something.

To this end, an incredible amount of work has been done over the last few years to cover great amounts of model space. These initial searches rightfully worked under the principle of Occam’s razor and searched for the simplest models that could provide solutions to the most problems. Unfortunately, no evidence of new physics has been observed, and it appears that nature is not that kind and likely won’t take the simple forms hoped for. However, if one is willing to make sacrifices, in the form of increasing model complexity or giving up potential solutions, there is still much more intriguing phase space that needs to be covered.

This dissertation describes one such search using a dataset corresponding to 35.9 fb^{-1}

of $\sqrt{s} = 13$ TeV proton-proton collisions collected in 2016. A brief description of this search and its results has been published and is given in Reference [26] with additional information presented in Reference [27]. Preliminary results using 2.3 fb^{-1} are given in Reference [28], while results using similar methods at 8 TeV were published in Reference [29].

Chapter 2 of this dissertation discusses the current state of the Standard Model and its deficiencies, introducing the Hierarchy Problem. Chapter 3 then motivates minimal flavor violating supersymmetry as a class of models worth experimental attention in which the constraints on “natural” solutions to the Hierarchy Problem are reduced in exchange for disfavoring explanations for dark matter. Chapters 4 and 5 describe the Large Hadron Collider and Compact Muon Solenoid detector used for producing and collecting the relevant data samples, as well as how particles are identified for reconstructing collision events. Chapter 6 presents in more detail the data sample along with the generation of the simulated samples used.

At this point discussion of the search strategy begins with Chapter 7, which presents the event selection that defines the analysis region. Chapter 8 gives an overview of the maximum-likelihood fit used for predicting the background, while the systematic uncertainties that are assessed for this procedure are provided in Chapter 9. The technical aspects of the likelihood model along with extensive validation tests are detailed in Chapter 10. Finally, the results from this search are presented in Chapter 11, with a summary and conclusions given in Chapter 12.

Part I

Theoretical Context and Motivations

Chapter 2

The Standard Model

2.1 Overview and Successes

The Standard Model (SM) of particle physics is a wildly successful theory and one of the greatest accomplishments in science. Formulated in the second half of the 20th century, it describes 17 experimentally-observed, fundamental particles and their interactions through the electromagnetic, weak, and strong fundamental forces. These particles, shown in Figure 2.1, consist of six quarks and six leptons, known as fermions, which comprise the matter particles, along with four gauge bosons and one scalar boson, which mediate particle interactions.

Formally, the Standard Model is a quantum field theory with symmetries described by the group $SU_c(3) \times SU_L(2) \times U_Y(1)$ and with a Lagrangian of

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}_i \not{D}\psi_i \\ & + y_{ij}(\psi_i\psi_j + \bar{\psi}_i\bar{\psi}_j)\phi + |D_\mu\phi|^2 - V(\phi), \end{aligned} \tag{2.1}$$

where $F_{\mu\nu}$ is the field strength tensor, D_μ is the gauge covariant derivative, $\not{D} = \gamma^\mu D_\mu$, ψ_i are the fermion fields, ϕ is the Higgs field, and y_{ij} are the yukawa couplings. In the top line, the first term of the Lagrangian describes the interactions of gauge bosons, while

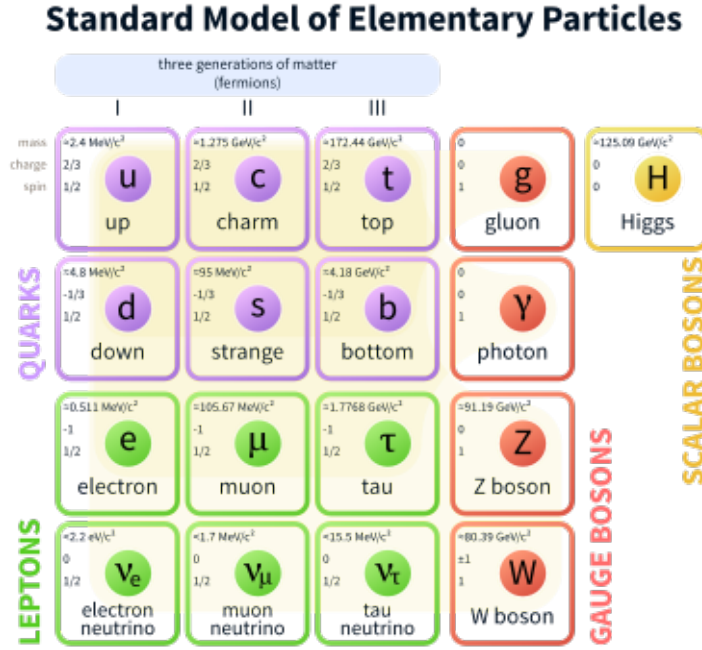


Figure 2.1: The fundamental particles of the Standard Model and some of their properties.

the second encodes the interactions between gauge bosons and fermions. The bottom line describes Higgs physics with the first term describing the Higgs-fermion interactions, the second term encoding the Higgs-gauge boson interactions, and finally, the third term representing the Higgs potential.

As the Standard Model is able to provide predictions across an enormous scope of physics, it has been rigorously tested throughout its history. For example, Figure 2.2, shows the agreement between the SM predictions and experimental results for the production cross section of a variety of processes. Amazingly, all of the measurements, spanning nine orders of magnitude, agree with SM predictions. At the same time, the Standard Model is the most precisely tested theory in physics with its prediction [22] of the anomalous electron magnetic moment incredibly agreeing with experimental measurements [23, 24] at up to twelve decimal places, as shown in Table 2.1.

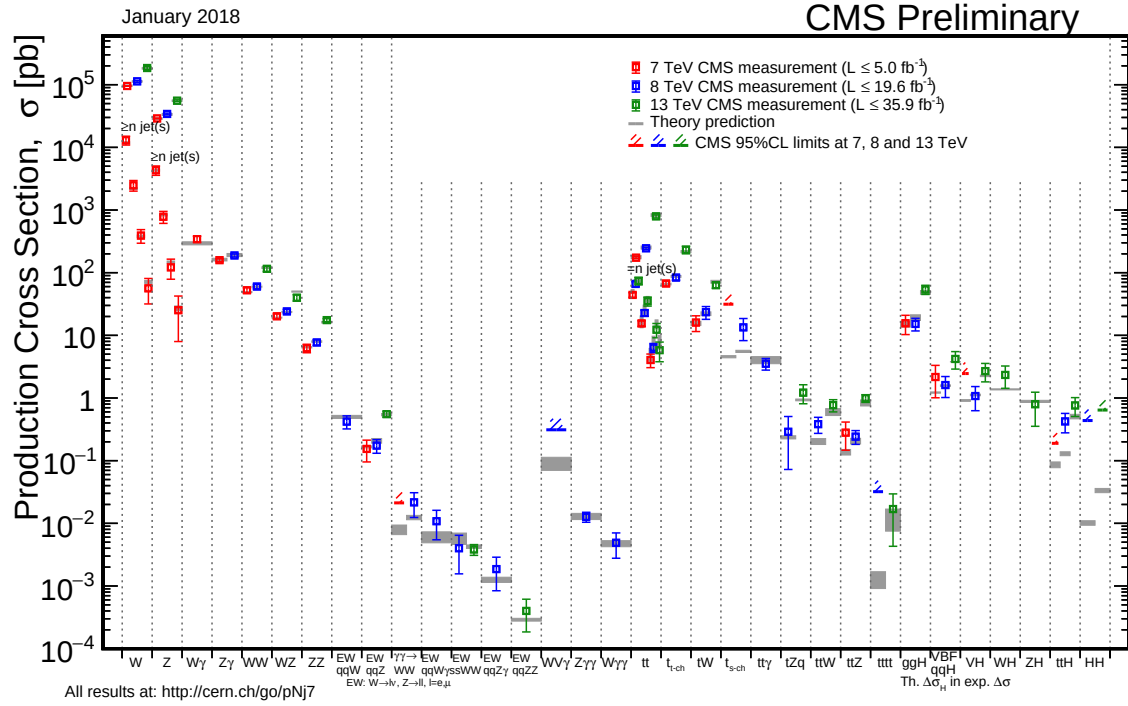


Figure 2.2: The theoretical and experimental value for the production cross section of various processes [2].

$a_e(\text{Theory})$	0.001 159 652 181 78(77)
$a_e(\text{Experiment})$	0.001 159 652 180 73(28)

Table 2.1: The theoretical [22] and experimental [23, 24] values of the anomalous electron magnetic moment, a_e . The uncertainty in the last digits is shown in parantheses.

2.2 The Standard Model as an Incomplete Theory

Despite the impressive successes of the Standard Model, it is incomplete as a fundamental description of the universe and many tensions exist between it and both experimental and theoretical concerns. For example, the Standard Model glaringly leaves out the fundamental force of gravity and attempts to construct a theory of quantum gravity have been fraught with difficulties. Furthermore, the Standard Model is unable to explain the substantial astrophysical evidence for dark matter [30, 31], which comprises $\sim 80\%$ of the matter in the universe, as it has no suitable candidate that can account for the observed mass density of dark matter.

Additionally, the Standard Model provides no mechanisms for:

- the origin of neutrino masses
- the matter-antimatter symmetry
- the presence of dark energy
- a grand unified theory of the strong and electroweak forces

Because of these issues, the Standard Model is believed to be a low-energy effective field theory with new physics entering at higher energies. It is, however, not easy to know what form this new physics may take, and thus solutions to the above problems have been used to drive much of the theoretical framework for extending the Standard Model. In particular, the Hierarchy Problem and the idea of “naturalness”, described in the section below, are perhaps the most significant “lampposts” in the search for new physics.

2.2.1 The Hierarchy Problem and Naturalness

The long-awaited discovery of the Higgs boson [32, 33, 34, 35, 36, 37], shown in Figure 2.3, confirmed the existence of a scalar boson with an observed mass of approximately 125 GeV. This relatively low mass of the Higgs boson, however, creates two related theoretical concerns. First, the Higgs mass is tied to the electroweak scale by the relation

$$v = \frac{m_H}{\sqrt{2}\lambda_H} \approx 246 \text{ GeV}, \quad (2.2)$$

where v is the higgs vacuum expectation value (vev), m_H is the physical Higgs mass, and λ_H is the Higgs Yukawa coupling. The vev directly dictates the electroweak scale and is only on the order of 100 GeV. The next largest known energy scale is that of quantum gravity and is typically defined as the Planck mass, which is on the order of 10^{19} GeV. The question as to why these two scales are so discrepant is known as the Hierarchy Problem [38].

The second question stems from trying to understand the observed mass of the Higgs boson in the presence of quantum corrections. The Higgs mass can be broken down into two components, its bare mass, $m_{H,0}$, and contributions from radiative corrections, δm_H^2 , as shown in the equation below

$$m_H^2 = m_{H,0}^2 + \delta m_H^2. \quad (2.3)$$

As a scalar particle, the Higgs boson mass receives contributions from all massive particles. In particular, for a fermion, f , the contribution to the Higgs mass at the one-loop

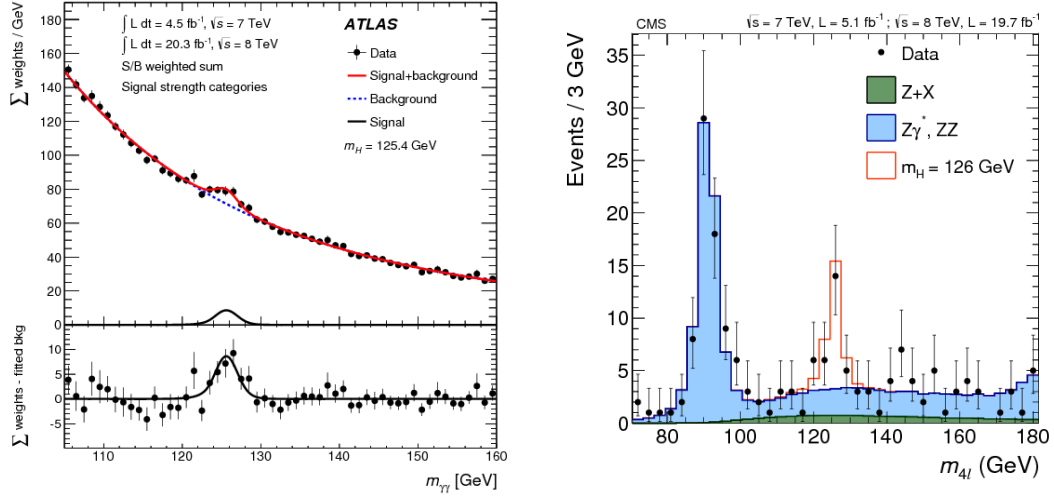


Figure 2.3: Evidence for the existence of a 125 GeV Higgs boson in the di-photon channel (left) and in the four-lepton channel (right) [3, 4].

level, corresponding to the the diagram shown in Figure 2.4, is of the form

$$\begin{aligned}
 \delta m_H^2|_f &= -\frac{\lambda_f^2}{8\pi^2} N_f \int^{\Lambda_{UV}} \frac{d^4 p}{p^2} \\
 &= -N_f \frac{\lambda_f^2}{8\pi^2} \left[\Lambda_{UV}^2 - 6m_f^2 \ln \left(\frac{\Lambda_{UV}}{m_f} \right) + 2m_f^2 \right] + \mathcal{O} \left(\frac{1}{\Lambda^2} \right), \quad (2.4)
 \end{aligned}$$

with λ_f the Yukawa coupling and N_f the number of fermionic degrees of freedom. The quantity Λ_{UV} is the cutoff of the momentum integral, which represents the approximate scale at which the Standard Model is no longer valid. In the case that there is no new physics beyond the Standard Model, this would correspond to $\Lambda_{UV} \sim$ Planck scale $\sim 10^{19}$ GeV, where effects due to quantum gravity are introduced.

In this scenario, the leading term of the Higgs mass corrections is $\mathcal{O}(10^{38} \text{ GeV}^2)$, while, as measured, the squared Higgs mass is only $\mathcal{O}(10^4 \text{ GeV}^2)$. For these two values to be consistent with each other, the Higgs bare mass parameter must exactly cancel the correction term at over 30 decimal places of precision. This high-level of fine-tuning is

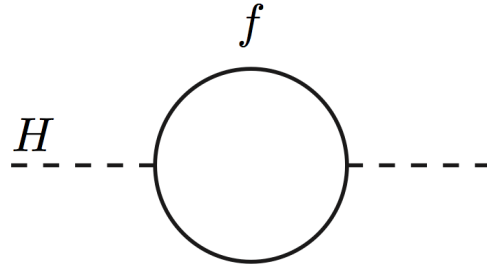


Figure 2.4: The fermionic one-loop Higgs mass correction.

considered to be “unnatural” and has been deemed the Naturalness Problem [39, 40, 5, 41]. While it is entirely plausible that such a fine-tuned cancellation of parameters occurs – there is no inherent theoretical reason against this, this problem motivates the need for new physics below the Planck scale, particularly physics that naturally incorporates a mechanism for cancelling the quadratic divergence of the Λ_{UV}^2 term.

Chapter 3

Supersymmetry

3.1 Natural Supersymmetry

Supersymmetry (SUSY) is an extension of the Standard Model that introduces a new symmetry that relates fermionic and bosonic degrees of freedoms [42, 43, 44, 45, 46, 47, 48, 49, 50, 51]. This symmetry imposes that for every fermionic degree of freedom in the Standard Model, there exists a “superpartner” bosonic degree of freedom, and vice versa. Furthermore, this symmetry dictates that both sets of particles couple to physics at the Λ_{UV} scale identitically.

In the Minimal Supersymmetric Standard Model (MSSM) [52], the Standard Model is extended to include a Higgs sector comprised of two scalar doublets and only the SM superpartners. The superpartners of fermions are scalars and labeled by prefixing an “s” to the beginning of the corresponding SM particle name, while the superpartners of bosons are spin-1/2, labeled by suffixing an “ino” to the end of the corresponding SM particle name, and collectively known as gauginos. Thus, the SUSY particles consist of 4 higgsinos, 12 squarks, 9 sleptons, and 7 gauginos, where there are 2 scalar particles for each SM fermion in order to preserve the number of degrees of freedom. The superpartners of SM particles are not necessarily mass eigenstates and can mix. The gauge and mass eigenstates of the superpartners, along with their properties, are shown in Table 3.1.

Name	Spin	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	$H_u^0 \ H_d^0 \ H_d^+ \ H_d^-$	$h^0 \ H^0 \ A^0 \ H^\pm$
squarks	0	$\tilde{u}_L \ \tilde{u}_R \ \tilde{d}_L \ \tilde{d}_R$	same
		$\tilde{c}_L \ \tilde{c}_R \ \tilde{s}_L \ \tilde{s}_R$	same
		$\tilde{t}_L \ \tilde{t}_R \ \tilde{b}_L \ \tilde{b}_R$	$\tilde{t}_1 \ \tilde{t}_2 \ \tilde{b}_1 \ \tilde{b}_2$
sleptons	0	$\tilde{e}_L \ \tilde{e}_R \ \tilde{\nu}_e$	same
		$\tilde{\mu}_L \ \tilde{\mu}_R \ \tilde{\nu}_\mu$	same
		$\tilde{\tau}_L \ \tilde{\tau}_R \ \tilde{\nu}_\tau$	$\tilde{\tau}_1 \ \tilde{\tau}_2 \ \tilde{\nu}_\tau$
neutralinos	1/2	$\tilde{B}^0 \ \tilde{W}^0 \ \tilde{H}_u^0 \ \tilde{H}_d^0$	$\tilde{N}_1 \ \tilde{N}_2 \ \tilde{N}_3 \ \tilde{N}_4$
charginos	1/2	$\tilde{W}^\pm \ \tilde{H}_u^\pm \ \tilde{H}_d^\pm$	$\tilde{C}_1^\pm \ \tilde{C}_2^\pm$
gluino	1/2	$\tilde{g}$	same

Table 3.1: The additional SUSY particles in the Minimal Supersymmetric Standard Model.

With the addition of these superpartners, a mechanism for naturally cancelling the radiative corrections to the Higgs mass is apparent, as the Higgs will couple to these new massive particles as well and provide additional radiative corrections that cancel the SM contributions. For example, for a fermion f , the radiative corrections from the corresponding sfermion $\tilde{f}$, shown in Figure 3.1, take the form

$$\begin{aligned}
\delta m_H^2|_{\tilde{f}} = & + N_{\tilde{f}} \frac{\lambda_{\tilde{f}}}{8\pi^2} \left[-\Lambda_{UV}^2 + 2m_{\tilde{f}}^2 \ln \left(\frac{\Lambda_{UV}}{m_{\tilde{f}}} \right) \right] \\
& - N_{\tilde{f}} \frac{\lambda_{\tilde{f}}}{8\pi^2} \left[-2m_f^2 + 4m_f^2 \ln \left(\frac{\Lambda_{UV}}{m_{\tilde{f}}} \right) \right] + \mathcal{O}\left(\frac{1}{\Lambda_{UV}^2}\right)
\end{aligned} \tag{3.1}$$

using the relation $\sqrt{2}m_f = \lambda_f \nu$. Supersymmetry imposes that $N_{\tilde{f}} = N_f$ and $\lambda_{\tilde{f}} = -\lambda_f^2$, and thus, in the case where $m_f = m_{\tilde{f}}$, this perfectly cancels not only the quadratic divergence in Equation 2.4 but also the higher-order corrections, solving the Hierarchy and Naturalness Problems.

Of course, no superparticles have been observed at the SM scale. This, however, does not imply that supersymmetry is incorrect but only that it may be a broken symmetry.

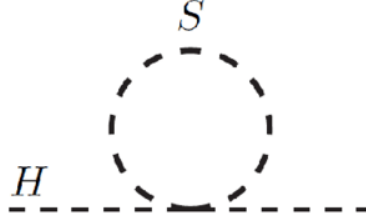


Figure 3.1: The scalar one-loop Higgs mass correction.

In this scenario where $m_f \neq m_{\tilde{f}}$, the quadratically divergent terms still cancel and what is left is only a logarithmic divergence, mediated by the squared mass difference of the partner particles:

$$\delta m_H^2 = N_f \frac{\lambda_f^2}{8\pi^2} (m_f^2 - m_{\tilde{f}}^2) \ln \left(\frac{\Lambda_{UV}}{m_{\tilde{f}}} \right) + \dots \quad (3.2)$$

This squared mass difference defines the degree to which the Higgs mass is fine-tuned.

3.2 Phenomenological and Experimental Constraints

While there is no theoretical prediction for the mass of the supersymmetric particles, there are certain guidelines for what the scale of these masses should be for a given level of fine-tuning that is deemed acceptable. Firstly, the Higgsino masses are directly controlled by the value of μ , which is related to the electroweak scale by

$$-\frac{m_Z}{2} = m_H^2 + |\mu|^2, \quad (3.3)$$

indicating that the Higgsino masses must be near the electroweak scale, ~ 100 GeV, in order to avoid a large fine-tuning of parameters.

The other superparticle masses are constrained by the size of their contributions to the Higgs mass, described in Equation 3.2. Not all superparticles, however, contribute equally to the Higgs mass, and thus the phenomenological constraints for particles are in proportion to the size of their correction to the Higgs mass. The largest constraint is on the stop squark masses due to the large Yukawa coupling of the top quark, which implies that the stop masses must be relatively light in order to keep the squared mass difference small and correspondingly the overall contribution small. This also constrains the mass of the left-handed sbottom squark as it is in a doublet with $\tilde{t}_L$ and thus must not be much heavier. Finally, the gluino couples to the squarks at the one-loop level, which means it still couples to the Higgs boson at the two-loop order, despite the Yukawa coupling for a gluon being zero. This constraint on the gluino is looser than for other particles described, but since it is strongly interacting, it has a high cross section at the Large Hadron Collider (LHC) and is thus an important experimentally accessible particle. Not considering model-dependent concerns, these are generically the only SUSY particles that are required to be light, and the rest of the superparticles may be decoupled with very high masses. A qualitative example of a natural SUSY spectrum is shown in Figure 3.2.

In order to gain a rough quantitative sense of what the mass constraints for these particles are, a measure of the Higgs mass fine-tuning can be constructed as

$$\mathcal{N} \equiv \frac{\delta m_H^2}{m_H^2}. \quad (3.4)$$

Thus for $\mathcal{N} = 10$, a fine tuning of 1 part in 10, the bounds are $m_{\tilde{t}} \lesssim 1$ TeV and correspondingly $m_{\tilde{b}} \lesssim 1$ TeV and $m_{\tilde{g}} \lesssim 2$ TeV.

Recent results from the LHC, however, are already starting to threaten these bounds, as shown in Figure 3.3 and Figure 3.4, with gluino and stop exclusion limits already surpassing these bounds for certain models. These constraints, however, are largely

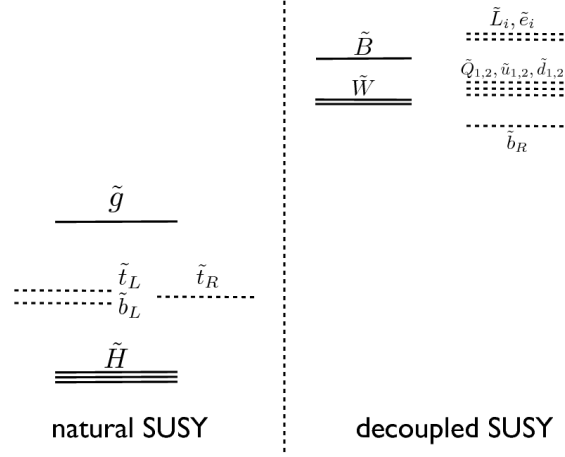


Figure 3.2: An example spectrum for natural SUSY [5].

in the context of R -parity conserving models, where a new quantum number, P_R , is conserved. R -parity is defined per particle as

$$P_R \equiv (-1)^{3(B-L)+2s}, \quad (3.5)$$

where B , L , and s is the baryon number, lepton number, and spin of the particle. This results in $P_R = +1$ for SM particles and $P_R = -1$ for SUSY particles. The motivation for requiring that R -parity be conserved is that the lightest supersymmetric particle (LSP), in this case, cannot decay to SM particles and is therefore stable. For models where the LSP is a neutralino, the LSP is then a dark matter candidate.

The phenomenology, however, for R -parity conserving (RPC) and R -parity violating (RPV) SUSY models is very different. For RPC scenarios, the stable LSP does not interact with the detector and escapes without depositing any energy. The presence of the LSPs, however, can be inferred by examining the missing transverse momentum (E_T^{miss}) in an event, which due to the negligible transverse momentum of the initial

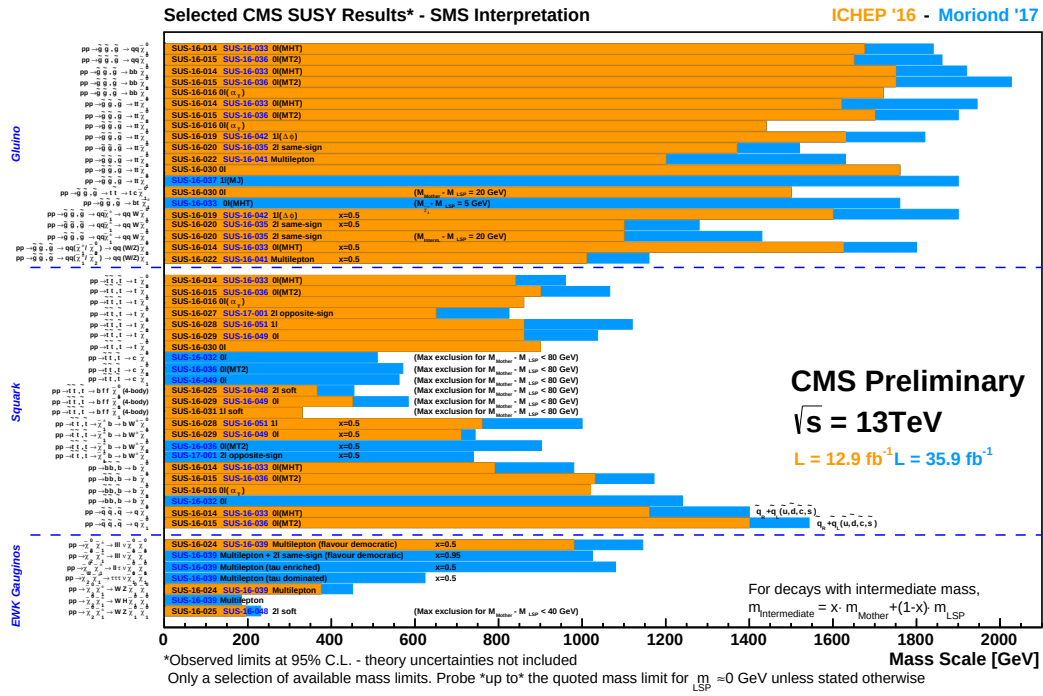


Figure 3.3: An overview of recent results from SUSY searches from the Compact Muon Solenoid experiment [6].

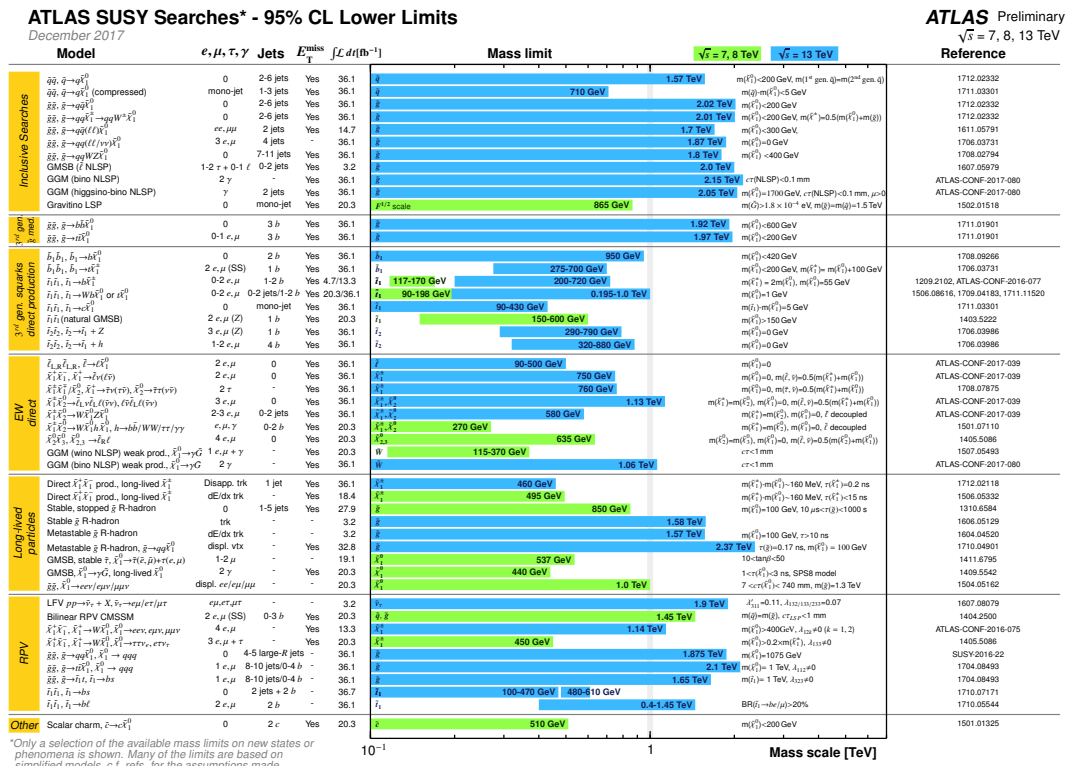


Figure 3.4: An overview of recent results from SUSY searches from the AToroidal LHC Apparatus experiment [7].

colliding particles, should be 0 in events without LSPs or neutrinos. Thus, the E_T^{miss} in an event provides a powerful handle for discriminating signal and background and most analyses search for signatures that include significant amounts of E_T^{miss} . In RPV models, however, the LSP is not stable and decays to SM particles, which does not produce a large E_T^{miss} signature. Although this disfavor the LSP as a dark matter candidate, it allows RPV models to evade the constraints from RPC searches. Subsequently, RPV SUSY yields an important class of models that can ease the tension between natural solutions to the hierarchy problem and current experimental limits.

3.3 R -parity Violating Supersymmetry

In the MSSM, the additional R -parity violating terms are

$$W_{RPV} = \frac{1}{2}\lambda^{ijk}L_iL_j\bar{e}_k + \lambda'^{ijk}L_iQ_j\bar{d}_k + \frac{1}{2}\lambda''^{ijk}\bar{u}_i\bar{d}_j\bar{d}_k + \mu'^iL_iH_u, \quad (3.6)$$

where the color indices have been suppressed and the letters i, j, k denote generation. The fields L_i , Q_j , and H_u are SU(2) doublets corresponding to leptons, quarks, and the Higgs boson, respectively, while $\bar{e}_k$, u_i , and d_j are the charged lepton, up-type quark, and down-type quark SU(2) singlets. The μ and various λ factors are coupling strengths for their corresponding interactions, where λ and λ'' must be antisymmetric in their first and last two indices, respectively, due to color conservation. A full description of RPV SUSY can be found in Reference [53].

While there is no fundamental theoretical reason forbidding R -parity violation, there are significant constraints on these interactions, primarily due to the lepton number violating (LNV) couplings, λ and λ' , and the baryon number violating (BNV) coupling, λ'' [54]. The most stringent of these constraints is from proton decay on which current

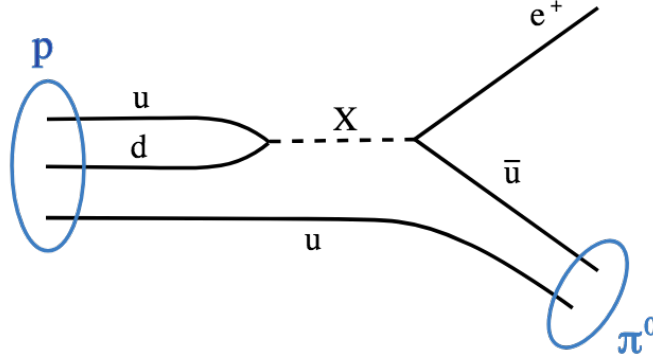


Figure 3.5: Example diagram representing proton decay, where the particle X represents a down-type squark [8]. The left vertex is mediated by the baryon number violating coupling, λ'' , and the right vertex is mediated by the lepton number violating coupling, λ' .

experimental results place a lower bound on the proton half-life of $\mathcal{O}(10^{34})$ years [55, 56]. Proton decay, however, requires both a lepton number and baryon number violating coupling, as shown in Figure 3.5. This constraint can be avoided if a mechanism exists to make one of these couplings zero or negligibly small. Additionally, though, there are strong limits on the individual LNV and BNV couplings, for example from neutron oscillation and muon-to-electron decay measurements, which are most stringent for the light generations. Thus, for any mechanism to evade these constraints, it must also motivate smaller couplings for the lighter generations.

3.3.1 Minimal Flavor Violating Supersymmetry

One way to avoid the constraints placed on the RPV couplings is to construct a model by following the structure of minimal flavor violation. In these Minimal Flavor Violating (MFV) SUSY models [57, 58, 25], the RPV couplings are related to the SM Yukawa couplings, making the third generation RPV couplings large and those of the first two

	ds	db	bs
u	3×10^{-12}	6×10^{-9}	5×10^{-7}
c	1×10^{-8}	1×10^{-5}	4×10^{-5}
t	4×10^{-5}	6×10^{-5}	2×10^{-4}

Table 3.2: Rough estimates for the sizes of the λ''_{ijk} MFV RPV couplings [25].

small. For example, the λ'' coupling can be written as

$$\lambda''_{ijk} = w'' y_i^{(u)} y_j^{(d)} y_k^{(d)} \epsilon_{jkl} V_{il}^* \quad (3.7)$$

with w'' an $\mathcal{O}(1)$ parameter, $y^{(u)}$ and $y^{(d)}$ the up- and down-type Yukawa couplings, and V the CKM matrix. From this, the sizes of the λ''_{ijk} couplings can be roughly estimated (using $w'' = 1$) and are shown in Table 3.2, which depicts the dependence of the coupling strength on generation. Additionally, in MFV scenarios, the LNV couplings are severely suppressed by neutrino masses, and in the limit of massless neutrinos, are exactly zero. Thus, the only relevant RPV coupling in MFV SUSY models is the BNV $\bar{u}\bar{d}\bar{d}$ coupling, which is small for the first two generations – meeting the exact criteria necessary to evade experimental constraints on RPV couplings. Furthermore, in the case where the LSP is a squark, it will decay promptly and not produce E_T^{miss} , allowing these models to even evade the constraints from RPC SUSY searches.

Because of these considerations, MFV SUSY is an intriguing class of models to investigate experimentally. In particular, due to the high $\tilde{g}\tilde{g}$ cross-section and large value of λ''_{tbs} , a search for the pair-production of gluinos that decay via $\tilde{g} \rightarrow t\bar{t} \rightarrow tbs$, as shown in Figure 3.6, is well-motivated.

The remainder of this dissertation is dedicated to describing such a search conducted with the Compact Muon Solenoid detector using $\sqrt{s} = 13$ TeV proton-proton collisions.

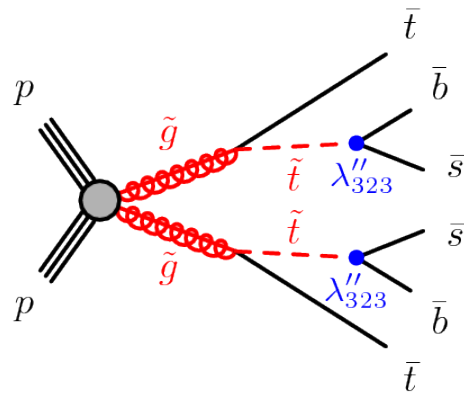


Figure 3.6: Example diagram of the pair production of gluinos that decay via $\tilde{g} \rightarrow t\bar{t} \rightarrow tbs$.

Part II

Experimental Apparatus

Chapter 4

Experimental Apparatus and Event Reconstruction

When construction of the Large Hadron Collider was approved over 20 years ago, much of its design was influenced by the need to search for physics beyond the Standard Model (BSM). The two most important properties of a collider are its center-of-mass energy and its (instantaneous) luminosity, and the LHC was designed to surpass all previous colliders in both aspects. The designed center-of-mass energy, $\sqrt{s}$, of 14 TeV allows for the production of particles heavier than ever before, while the designed luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ allows BSM searches to probe very rare processes.

In the same way, the design of the Compact Muon Solenoid (CMS) detector reflects the needs of BSM searches. In particular, an all-purpose, hermetic detector that can precisely measure a variety of particles is essential to fully search the parameter space of the many (un)theorized new physics models.

4.1 The Large Hadron Collider

The LHC since 2015 has been colliding protons together at $\sqrt{s} = 13 \text{ TeV}$, slightly below the designed specifications but still at an unsurpassed energy. In order to reach this center-of-mass energy, the LHC uses a large accelerator complex consisting of many

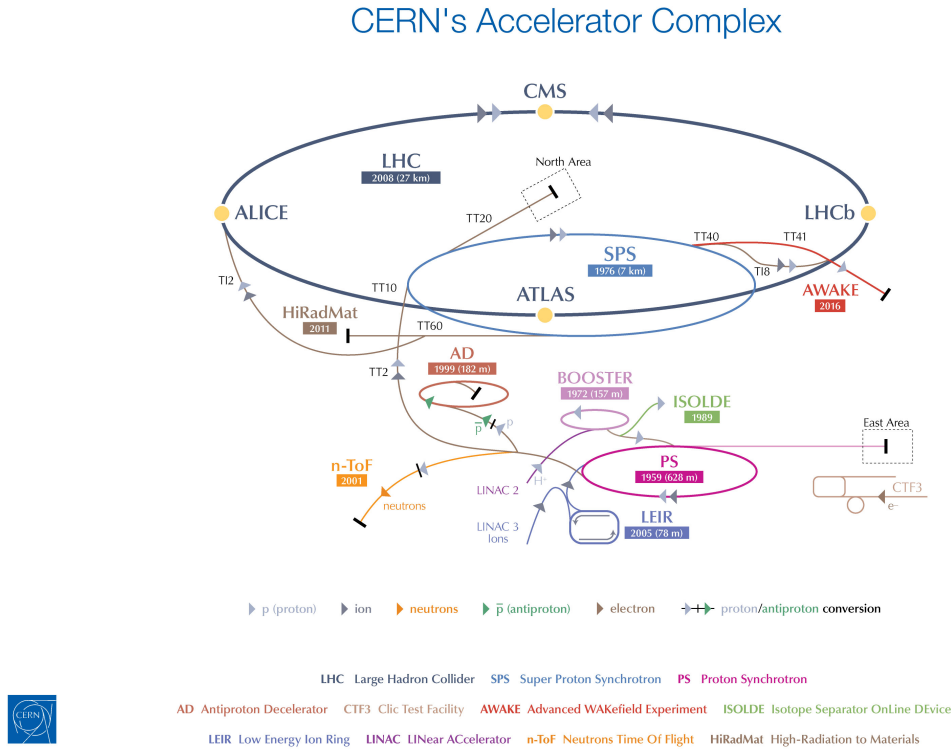


Figure 4.1: A schematic of the CERN LHC accelerator complex [9].

smaller particle accelerators, which is necessary to produce protons and bring them up to a speed such that they can be injected into the LHC ring. A diagram of the CERN accelerator complex is shown in Figure 4.1, while Figure 4.2 shows the physical size of the LHC.

This process first begins with a simple bottle of hydrogen gas, from which the hydrogen atoms are ionized by an electric field to produce the needed protons. These resulting protons are then fed into LINAC 2, the first accelerator in the chain, which accelerates them up to 50 MeV, creating a beam of protons. The proton beam is then passed successively to the Proton Synchrotron Booster, Proton Synchrotron, and Super Proton Synchrotron, where the beam reaches energies of 1.4 GeV, 25 GeV, and 450 GeV, respectively. At the

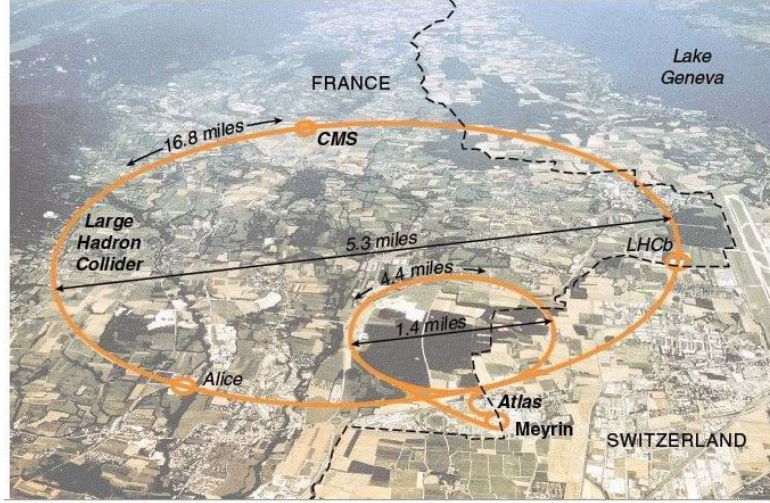


Figure 4.2: A map that shows the physical extent of the Large Hadron Collider.

Proton Synchrotron, the beams are additionally split into “bunches”, each consisting of $O(10^{11})$ protons and separated in time by 25 ns. Finally, the protons can be injected into the two beam pipes of the LHC, each circulating in opposite directions. These beams continue to be accelerated until they reach their final energy of 6.5 TeV, allowing for collisions at $\sqrt{s} = 13$ TeV. At this point, the proton beams are focused and fine-tuned at several stages in order to increase the luminosity. In 2016, the LHC was able to collide protons with an instantaneous luminosity of $1.4 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, which exceeded its designed specification and delivered a record-high integrated luminosity of 40.8 fb^{-1} .

These improvements from previous generations of colliders greatly increase the reach of searches for new, heavy particles. Since the center-of-mass energy of the actual colliding partons, $\sqrt{\hat{s}}$, is typically much less than the overall center-of-mass energy, raising the collider’s energy can greatly increase the production cross-section of heavy particles, especially of those around the TeV scale. For example, Figure 4.3, which depicts the ratio of parton luminosities at $\sqrt{s} = 13$ TeV and 8 TeV as a function of the characteristic mass scale of the event, shows that a 2000 GeV gluino will be produced through gg

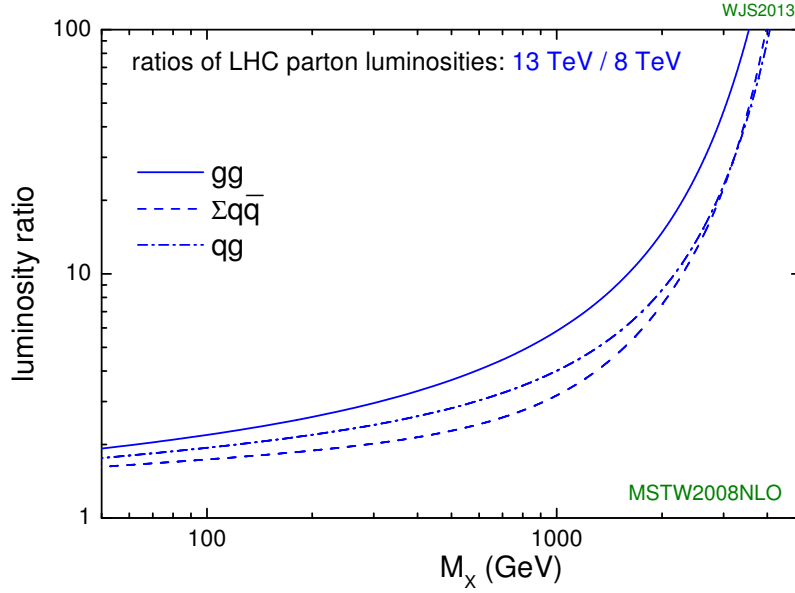


Figure 4.3: Ratio of parton luminosities at $\sqrt{s} = 13$ and 8 TeV [10].

scattering processes $\sim 15\times$ more often with less than a doubling of the collider energy. Increasing a collider's energy, however, is not always a practical option, involving new technologies, expensive upgrades, or even a new collider. When this is the case, the best alternative to continue to probe rare processes is to simply take more data, more quickly, which a high luminosity collider like the LHC allows for.

A more complete description of the LHC can be found in Reference [59].

4.2 Compact Muon Solenoid

Along the tunnels of the LHC, below Cessy, France, sits the CMS detector, shown in Figure 4.4, where proton-proton collisions are recorded. The overall shape of the detector is cylindrical with a length of 21.6 m and radius of 7.3 m, while weighing roughly 14,000 metric tons. The CMS detector is sometimes called a cylindrical onion,

as its shape is comprised of layers of specialized detectors, each designed to provide precise measurements for a particular particle type. Peeling back the layers from the outside-in, the first sub-detector is the muon system, which is represented by the “M” in CMS. Next is a superconducting solenoid of 6 m internal diameter that produces a magnetic field of 3.8 T, and perhaps most importantly provides the “S” in CMS. Placed within the solenoidal magnet, is the rest of the CMS detector, namely the Hadron Calorimeter (HCAL), Electromagnetic Calorimeter (ECAL), and a silicon tracker. The design challenge of compactly fitting most of the detector components within the solenoid is responsible for the “C” in CMS. A diagram of the layout of the CMS detector can be seen in Figure 4.5.

At the center of the silicon tracker is Interaction Point 5, the beam crossing which provides the proton-proton collisions to the CMS detector, and is the nominal origin of CMS’s coordinate system. The x -axis is defined to point towards the center of the LHC ring and the y -axis is defined to point up towards the surface, both of which are transverse to the proton beam. The z -axis points along the beamline with the positive direction given by the right-hand rule relative to the x - and y -axes. Due to the CMS detector shape, it is often useful to convert the cartesian coordinates to a cylindrical coordinate system. In this system, the azimuthal angle, ϕ , is measured from the x -axis in the xy -plane, and the polar angle, θ , is measured from the z -axis. The polar angle, however, is often replaced by pseudorapidity, defined as $\eta = -\ln(\theta/2)$. Thus, any point in the CMS coordinate system can be represented by (r, η, ϕ) . A diagram showing both the cartesian and cylindrical coordinate systems can be seen in Figure 4.6.

The remainder of this section briefly describes the main features of the various CMS sub-detectors. A detailed discussion can be found in Reference [18].

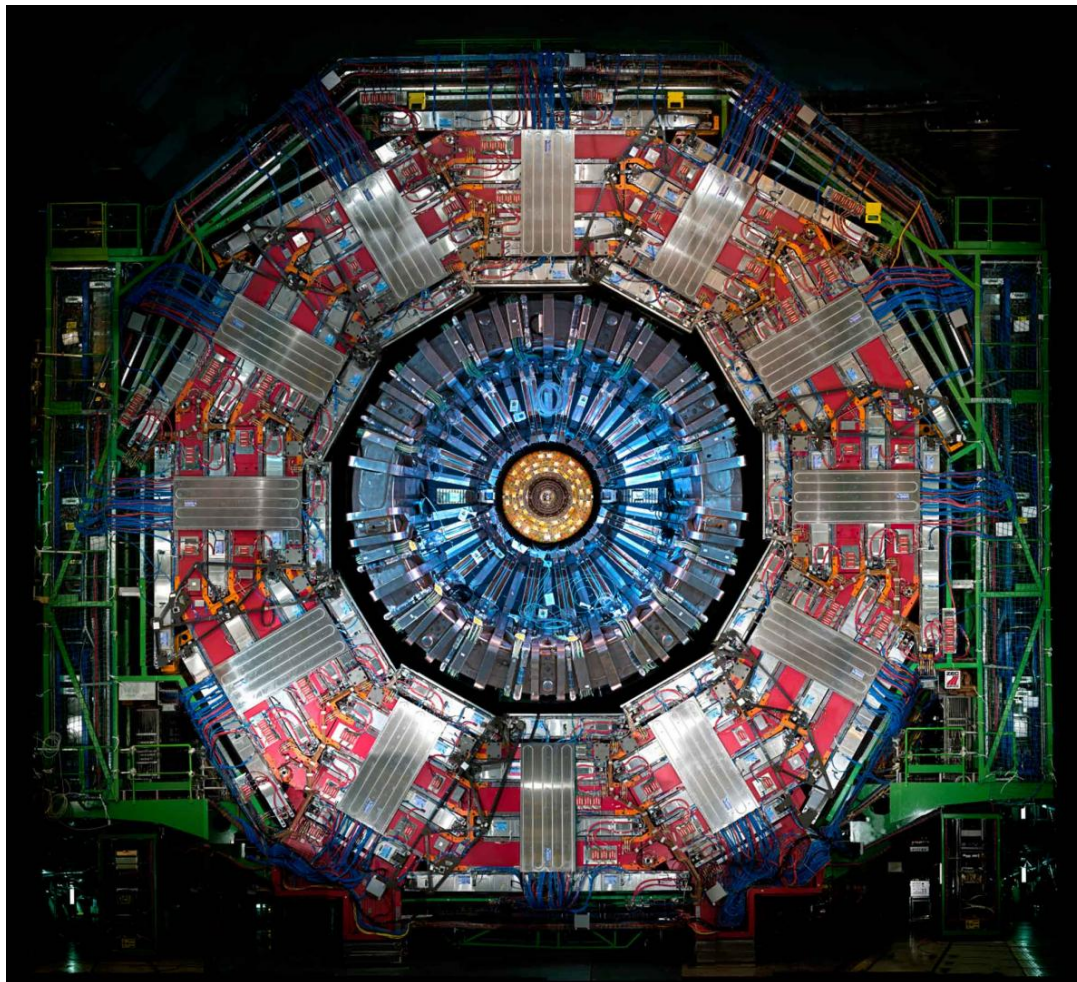


Figure 4.4: The CMS detector [11].

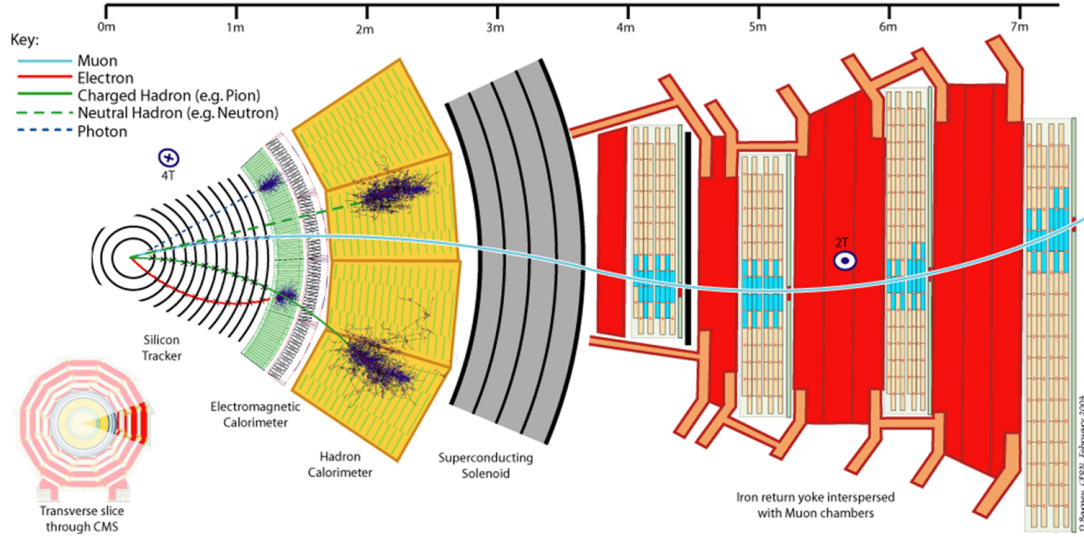


Figure 4.5: A diagram showing the various sub-detectors of the CMS detector [12].

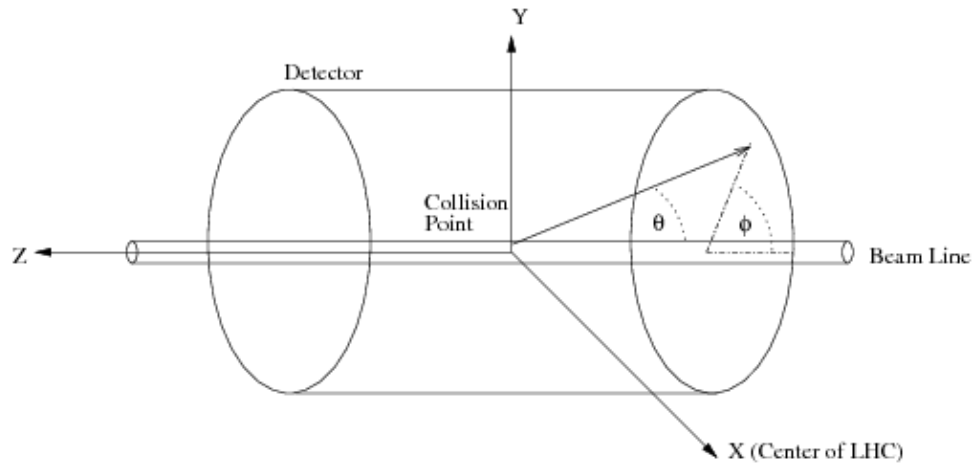


Figure 4.6: A diagram of the cartesian and cylindrical coordinate systems used by CMS [13].

4.2.1 Inner Tracking System

The tracking system is used for precise measurements of the trajectories of charged particles, as well as reconstruction of secondary vertices. As the tracking system is the closest subdetector to the interaction point, it faces a very large particle flux rate and so must be able to provide both high granularity and fast response, as well as be able to survive operating in those conditions with an expected lifetime of 10 years. At the same time, these features must be balanced with minimizing the amount of material in order to reduce unwanted interactions with the detector, such as multiple scattering, photon conversion, and nuclear interactions. These requirements lead to a tracking system composed entirely of silicon technology.

The CMS tracking system is actually composed of two parts. The first is the pixel detector, which surrounds the interaction point, and is composed of 3 barrel layers at radii between 4.4 and 10.2 cm and 2 endcap layers that extend the acceptance up to $|\eta| < 2.5$. In total, the pixel detector covers an area of roughly 1 m^2 with 66 million pixels and achieves a resolution of roughly 10 and 20 microns in the directions transverse and longitudinal to the beam line, respectively.

The second part of the tracking system is the strips detector which sits just outside the pixel detector. The strips detector is composed of 4 parts: the tracker inner barrel (TIB), tracker inner disks (TID), tracker outer barrel (TOB), and the tracker endcaps (TEC). The TIB and TID extend up to 55 cm in radius and are composed of 4 and 3 layers, respectively, while the TOB, which surrounds the TIB and TID, extends out to 116 cm and is composed of 6 layers. Lastly, the TEC which sits next to the other strip detector components, covers a radius of 22.5 to 113.5 cm and is composed of 9 disks. In total, the strips detector covers an area of 198 m^2 with 9.3 million strips. A layout of the tracking system including the pixel and strips detector is shown in Figure 4.7.

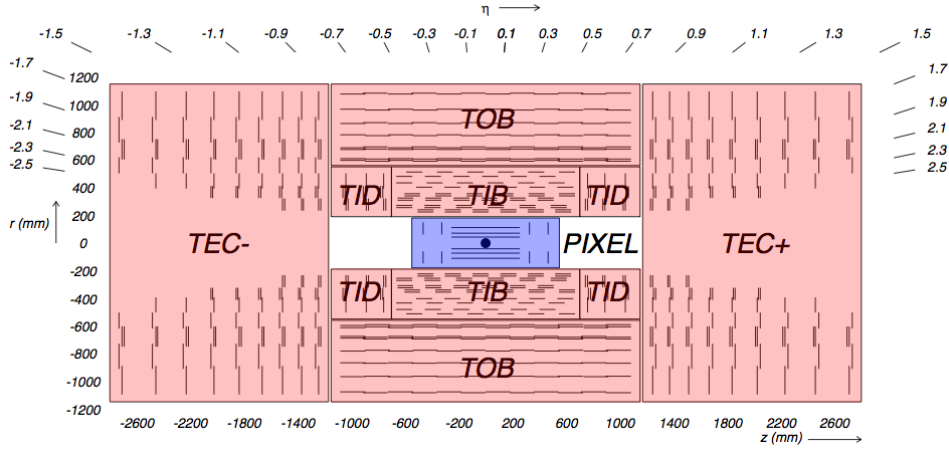


Figure 4.7: Layout of the CMS tracking system, showing both the pixel detector (blue) and the strips detector (red) [14].

4.2.2 Electromagnetic Calorimeter

The primary purpose of the electromagnetic calorimeter (ECAL) is to measure the energy of electrons and photons. The ECAL is a hermetic, homogenous detector made up of a barrel part, covering the $|\eta| < 1.479$ region, and two endcap parts that covers $1.566 < |\eta| < 3.0$. Both the barrel and endcap sections are comprised of lead tungstate (PbWO_4) crystals with 61,200 in the barrel and 7,324 in each of the endcaps. The use of the PbWO_4 crystals was motivated by their high density, short radiation length, small Molière radius, and radiation hardness, all of which allow for a fine granularity, radiation resistant, compact calorimeter.

The lead tungstate crystals act as scintillators, which produce an amount of light that is proportional to the energy of an incident particle. This light is then converted to an electrical signal by silicon photodetectors (avalanche photodetectors in the barrel and vacuum phototriodes in the endcaps), which is used for the final energy measurement.

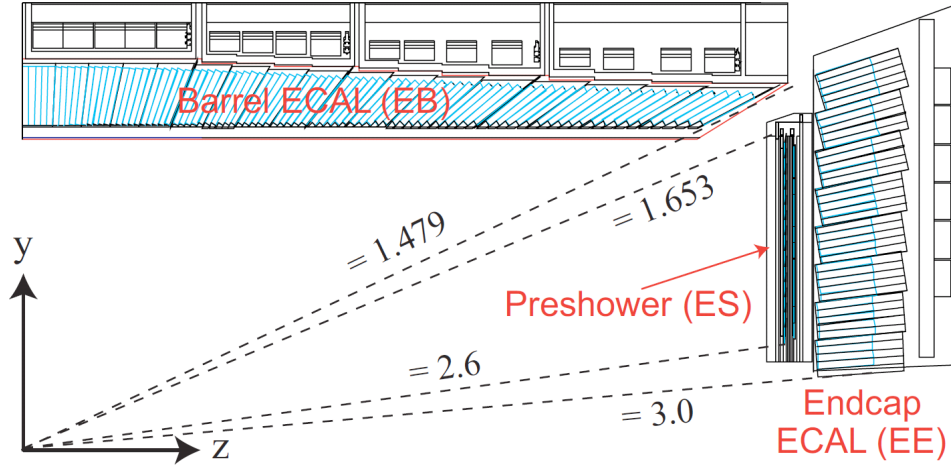


Figure 4.8: A cross section of the ECAL, showing its geometry and layout [15].

The resulting resolution on the energy measurements is given by

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \quad (4.1)$$

where S is the stochastic term, N the noise term, C the constant term, and E is in units of GeV. Typical values for S , N , and C , as measured in electron beam tests, are 2.8%, 12%, and 0.30%, respectively.

In addition to the ECAL barrel and endcaps is a preshower detector, which sits in front of the endcaps, covering $1.653 < |\eta| < 2.6$. The main purpose of the preshower detector is to identify neutral pions by improving the granularity, so as to be able resolve photon pairs from the decay of high energy pions that otherwise would be mismeasured as single photons. The preshower detector also provides improved position resolution for electrons and photons and helps identify electrons from minimum ionizing particles. The full layout of the ECAL is shown in Figure 4.8.

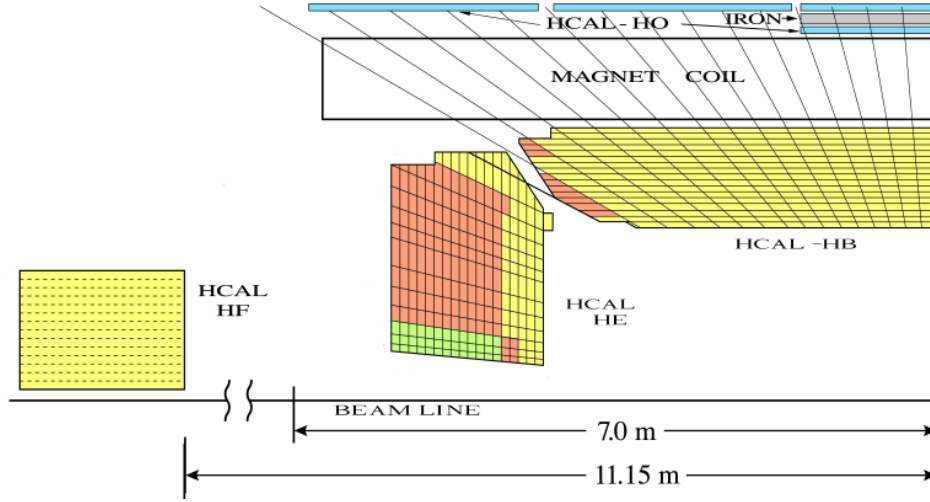


Figure 4.9: The layout and geometry of a quarter of the HCAL detector [16].

4.2.3 Hadronic Calorimeter

The primary purpose of the hadronic calorimeter (HCAL) is to measure the energy of hadrons, which can pass through the ECAL as they primarily interact through the strong force. The HCAL is a sampling calorimeter made up of either brass, iron, or steel absorbers and uses plastic scintillator tiles as the sampling material, which measures the energy of hadrons through scintillation, similarly to the ECAL. As a hadron travels through the HCAL, it interacts with one of the absorber layers, which results into a “shower” of particles that produces light in the scintillator tiles as the resulting particles pass through the sampling layers. These light pulses are carried by wavelength-shifting fibers to a hybrid photodiode, which converts the pulses to analog electrical signals that when summed have an amplitude proportional to the hadron’s energy.

The HCAL is separated into 4 components: the hadron barrel (HB), hadron endcap (HE), hadron forward (HF), and hadron outer (HO), the layouts of which are shown in Figure 4.9. The HB and HE completely surround the ECAL and were designed to

minimize any cracks between the two subdetectors with the HB covering $|\eta| < 1.3$ and the HE covering the rest up to $|\eta| = 3$. Both components function as sampling calorimeters with alternating absorber and sampling layers. In the HB, the first and last layers are made of steel while the 14 other absorber layers are made of brass, while the HE is made up of 18 brass absorber layers. For both components, there are sampling layers made of plastic scintillator tiles interspersed between each of the absorber layers.

The HF is used to measure the energy of the forward most hadrons in the pseudo-rapidity range of $3.0 < |\eta| < 5.0$. At this forward position, the HF faces extraordinary levels of particle flux and had to be designed to handle this radiation. Due to this constraint, the HF uses quartz fibers instead of plastic scintillator tiles as its active medium, as the quartz fiber are more radiation hard. The HF uses both long fibers, which run the full depth (165 cm) of the detector, and short fibers, which begin 22 cm from the front end of the detector. This geometry allows the HF to provide depth information of the energy deposits, which helps to identify electrons and photons from hadrons, as the former tends to deposit most of its energy in the first depth, while the latter deposits its energy more equally between the two depths. These fibers are embedded into the steel structure of the HF, which also acts as the absorber.

Lastly is the HO, whose main purpose is to act as a “tail catcher”. Due to the geometrical constraint that the HCAL fit within the CMS solenoid, the HB does not have enough material in the central η region to adequately contain hadron showers. So to provide extra sampling layers, the HO sits just beyond the solenoid and has 1 to 2 scintillator layers and uses the magnet as an extra absorber layer. At this position, the HO is able to identify late starting showers and measure the amount of energy that is deposited past the HB.

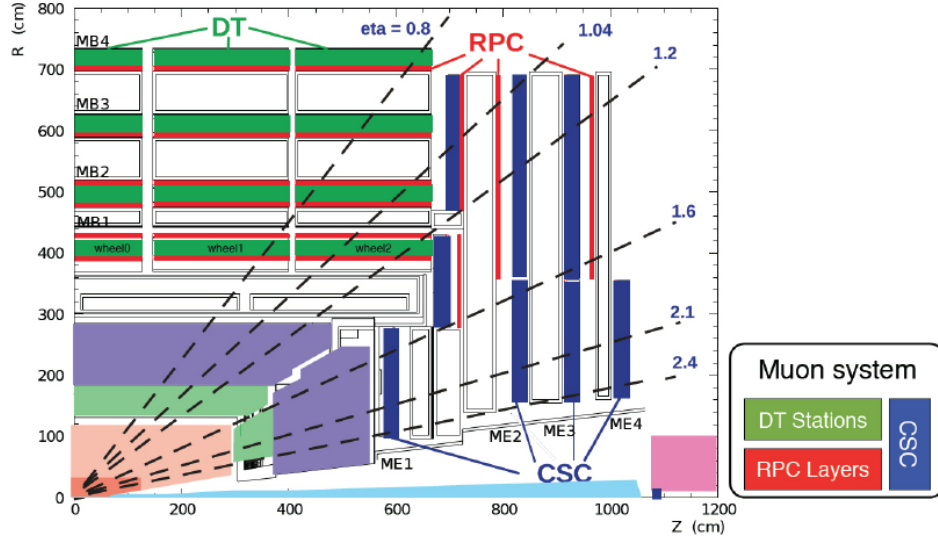


Figure 4.10: The layout of the muon system within the CMS detector [17].

4.2.4 Muon System

Muons, as implied by the name of CMS, are a central focus of the CMS detector, and the responsibility of identifying muons with high purity and providing precise momenta measurements falls to the muon system. To do this, the muon system is composed of three types of gaseous detectors, motivated by the need to cover a large area and varying radiation environments. Figure 4.10 shows the layout of the muon system within the CMS detector.

In the barrel region, $|\eta| < 1.2$, drift tube chambers (DTs) are used as the neutron-induced background is small, the muon rate is small, and the magnetic field is uniform. The DTs are organized into 4 stations, with three of the stations containing 8 chambers that measure position in the r - ϕ plane and 4 that measure position in the z -direction. The last station only contains the eight r - ϕ measuring chambers. In the endcap regions of CMS, $0.9 < |\eta| < 2.4$, the expected muon and background rates are higher and the magnetic field is large and non-uniform, both of which preclude the use of DTs. Instead,

the muon system endcaps are instrumented with cathode strip chambers (CSCs) that have a high response time, fine segmentation, and higher radiation resistance. The CSCs have 4 stations in each endcap with chambers that are aligned perpendicular to the beam line and are able to provide measurements in the r - ϕ plane and z -direction, along with the beam crossing time of a muon.

Both the DTs and CSC are capable of providing high efficiency and pure muon p_T triggers, independent of the rest of the detector. But in order to further improve this, particularly at the full LHC luminosity, another complementary trigger system consisting of Resistive Plate Chambers (RPCs) was added to both the barrel and endcap regions ($|\eta| < 1.6$). The RPCs are double-gap chambers that operate with a fast response and good time resolution. The spatial resolution, however, is coarser than the DTs or CSCs, though the extra hits in the RPC still help resolve ambiguities when making tracks. There are a total of 6 RPC layers in the barrel muon system, which help improve triggers for low p_T muons, and 3 layers in each of the endcaps that help reduce background and improve the time and p_T resolution of muons.

4.2.5 Trigger System

The high instantaneous luminosity of the LHC provides many technical challenges for the data acquisition system (DAQ), with proton-proton collisions occurring every 25 ns, corresponding to a frequency of 40 MHz. At this collision rate, it is unfeasible to process and store the data for each event. In order to reduce the rate, a two-stage trigger system is used to select only the most “interesting” events for processing.

The first stage is the Level-1 (L1) trigger system, which has approximately only 4 μ s to decide whether or not an event should be further processed. In order to operate at this timescale, the L1 trigger uses only coarse-grained information from the CMS calorimeters

and the muon system.

For the calorimeter set of data, the L1 first generates trigger primitives by looking for large energy deposits in the calorimeter. These trigger primitives are then passed to the Regional Calorimeter Trigger (RCT), which uses this information to determine electron/photon candidates and transverse energy sums per calorimeter region. In addition, the RCT also calculates information relevant for detecting minimally ionizing particles, vetoing tau leptons, and muon isolation. Lastly, the Global Calorimeter Trigger (GCT) uses the information from the RCT to construct jets and calculate the event-level transverse energy and missing transverse energy, along with the final isolated and non-isolated electron/photon candidates.

For the muon portion of the L1 trigger system, the DTs and CSCs both compute local trigger information which consists of two- and three-dimensional track segments, respectively. This information is then passed to a joint DT-CSC track finder, which connects these segments into a full candidate track. At the same time, the RPC constructs a separate, independent set of track candidates. Both sets of candidate tracks are sent to the Global Muon trigger (GMT), which also takes in the relevant information from the RCT, to construct muon candidates.

Lastly, the candidate particles and event-level information from the GCT and GMT are sent to the Global Trigger, which takes this information and checks to see if certain criteria are met. If so, a L1 Accept (L1A) is generated, which signals for the event to be fully read out. This process, shown in Figure 4.11, reduces the full readout rate to at most 100 kHz.

On the generation of an L1A, all the CMS subsystems read out the buffered data corresponding to the L1A event and pass it to an event builder. The data the event builder receives is both more complete and at a finer resolution, allowing it to construct more complex quantities before sending it to the High Level Trigger (HLT), the second

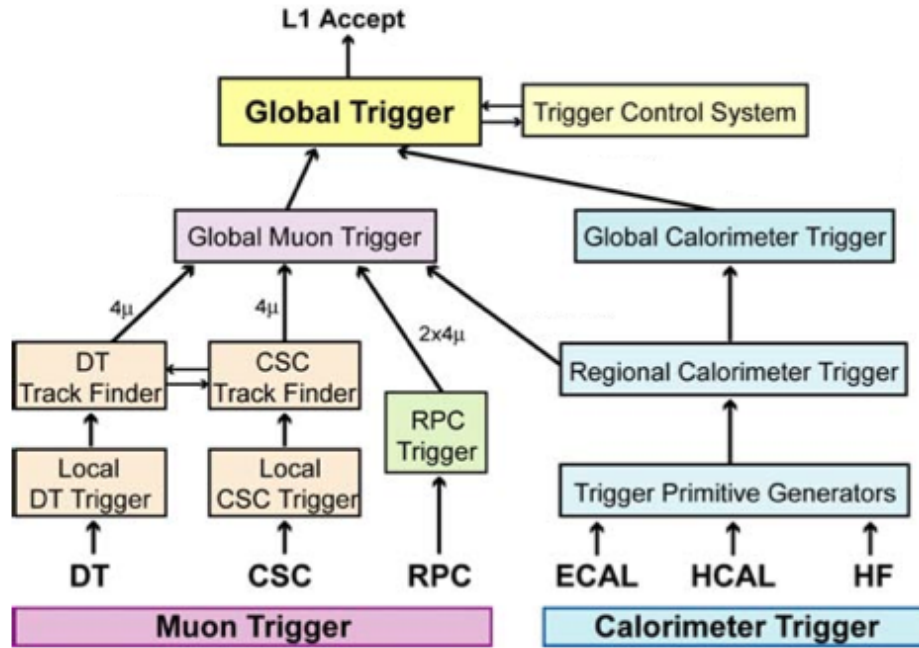


Figure 4.11: Flowchart depicting the generation of a L1 Accept [18].

stage of the trigger system.

The HLT software is run on a processor farm that reconstructs events in greater detail to decide whether or not they should be kept. This framework is flexible allowing both the HLT software and the processor farm to be updated in order to meet changing experimental needs. As such the exact criteria used by the HLT in its decision varies with time, but generally involves thresholding the p_T and/or multiplicity of particles along with event-level quantities. At the end of this process, the trigger rate is approximately 100 Hz.

Chapter 5

Particle Reconstruction and Identification

5.1 Tracks

Track reconstruction with the CMS detector faces many challenges, as at each bunch crossing $\mathcal{O}(10^3)$ particles are expected to pass through the CMS tracking system, all of which must be reconstructed in time to be inputted to the HLT. This constraint makes it immensely challenging to attain high track-finding efficiency, while minimizing the number of fake tracks.

The first step of track reconstruction is to reconstruct “hits” in a process called “local reconstruction”. In this step, signals in both the pixel and strip channels that are above some zero-suppression threshold are clustered together into hits, where the cluster positions and uncertainties are then estimated.

Next, tracks are reconstructed from these hits in order to provide estimates for the momentum and position of charged particles associated with the track. This is done using specialized software based on Kalman filters known as Combinatorial Track Finder (CTF). In order to reduce the combinatorial complexity of the problem, the CTF track reconstruction is performed six times. Each iteration attempts to reconstruct the most

easily-identifiable tracks, e.g. high- p_T tracks, and then removes the hits associated with those tracks. This helps simplify the track reconstruction in the following iterations.

Each iteration consists of four steps:

1. A seed is generated from a few (generally 2 or 3) hits. This seed provides an initial estimate of the trajectory and uncertainties associated with that track.
2. A Kalman-based track finder is used to extrapolate seed trajectories along their expected paths. Additional hits that are compatible with a path are assigned to that track candidate.
3. A track-fitting module uses a Kalman filter and smoother to provide estimates of the trajectory parameters for each track.
4. A track selection sets quality thresholds and discards tracks that fail the specified criteria.

A detailed description of track reconstruction can be found in Reference [60].

5.1.1 Vertices

An essential part of event reconstruction is identifying which particles were produced at parton-parton interaction vertices (primary vertices) and which were produced at a decay vertex of produced particles (secondary vertices). The process for selecting primary vertices consists of three steps: track selection, track clustering, and track fitting.

The track selection criteria is chosen in order to select tracks that are consistent with being produced promptly in the primary interaction region. Tracks are required to have a small transverse impact parameter relative to the beam spot, a certain number of hits in the pixel and strip detectors, and good fit quality when fitting to the trajectory.

No requirement is placed on the p_T of tracks, in order to ensure a high reconstruction efficiency, even for minimum-bias events.

After the track selection, a deterministic annealing clustering algorithm [61] is used in order to group together tracks that appear to originate from the same vertex. Here the selected tracks are originally all assigned to the same vertex and are then slowly divided into multiple vertices. This process continues until reaching a cutoff defined by balancing the risk of incorrectly splitting vertices and the resolving power of the algorithm.

Once the track clustering is completed, an “adaptive vertex fitter” is used to determine the 3D-position of vertices with at least two tracks [62], in which tracks corresponding to a vertex are each assigned a probability of correctly belonging to the vertex. The weighted sum of these probabilities is then used in the fitting algorithm to determine primary vertices.

In this process, many more than one primary vertex are reconstructed due to multiple parton-parton interactions in an event. There is, however, usually only one primary vertex of interest in the event, corresponding to the primary vertex with the highest sum of track p_T^2 . This primary vertex is commonly referred to as *the* primary vertex.

5.2 Calorimeter Clusters

Energy deposits in the various CMS calorimeters are clustered together to form “calorimeter clusters”. The purpose of the calorimeter clusters are to aid in, detecting and measuring the energy and direction of stable neutral particles, separating these neutral particles from the energy deposits of charged hadrons, reconstructing and identifying electrons, along with their corresponding bremsstrahlung photons, and measuring the energy of charged hadrons with low-quality track parameters.

The calorimeter clustering is done in three steps. First, cluster seeds are identified

as cells with an energy both larger than some threshold and larger than the energy of neighboring cells. Next, these cluster seeds are formed into “topological clusters” by iteratively merging together neighboring cells that have significant energy deposits. In this process, topological clusters can merge such that a cluster contains multiple cluster seeds. Lastly, the energy in a topological cluster is distributed among its seeds through a Gaussian-mixture model that results in the final calorimeter clusters.

This clustering is performed separately in each subdetector calorimeter, including the ECAL preshower detector, whose two layers are treated independently. There is no clustering performed in the HF, as each cell’s short- and long-fibers measure the electromagnetic- and hadronic-energy components, as described in Subsection 4.2.3. These components directly give rise to “HF EM” and “HF HAD” clusters.

5.3 Particle Flow

The particle-flow (PF) algorithm is a holistic approach towards event reconstruction. It combines the basic information of subdetectors, i.e. the tracks and clusters defined above, to identify each final-state particle and reconstruct their corresponding properties. By correlating measurements from the tracker and calorimeter, the PF algorithm is able to provide improved energy and momentum measurements. A complete, detailed presentation of the particle-flow algorithm is given in References [63, 64, 65].

5.3.1 Linking

As a particle traverses through the CMS detector, it is expected to generate multiple input elements to the PF algorithm. Thus, the first step of reconstructing particles is to “link” the various PF elements stemming from different subdetectors together. This is done by defining a “distance” between two linked elements, where the closer the distance

the more probable it is the two elements correspond to the same particle. The linking algorithm then creates “PF blocks” by associating directly or indirectly linked elements together. The exact criteria used to link elements together and define their distance depends on the type of elements being considered and are listed below.

A link between a track and calorimeter cluster is established by extrapolating the track trajectory through the ECAL and HCAL, up to a depth where energy deposits are most expected. If the extrapolated track falls within the area of the calorimeter cluster, the two elements are linked and the link distance is defined as the separation between their positions in the (η, ϕ) plane. In the case multiple clusters are linked to the same track only the link with the smallest distance is kept.

To link tracks with clusters from potential bremsstrahlung photons, tangents to the track are extrapolated to the ECAL. If a tangent line falls within the cluster, the track and cluster are linked with the η - ϕ separation used as the link distance.

Calorimeter clusters in the HCAL, ECAL, and preshower detector are linked together when the position of a cluster in a more granular calorimeter (preshower or ECAL) is within the boundaries of a cluster with less granularity (ECAL or HCAL). The distance between these two clusters is defined as either the η - ϕ or x - y separation for HCAL-ECAL and ECAL-preshower links, respectively. In the case where multiple HCAL(ECAL) clusters are linked to the same ECAL(preshower) cluster, only the link with the smallest distance is used.

Links between a track and a standalone-muon track, defined as a track segment constructed from hits in the muon system, are established when a global fit to the two tracks has an acceptable fit quality. The link distance in this case is defined as the χ^2 of the fit, and only the link with the smallest χ^2 is retained when there are multiple links to the same standalone-muon track. The resulting links are called “global muons”.

5.3.2 PF Reconstruction and Identification

Once the PF blocks have been constructed, the PF algorithm is applied to reconstruct and identify a set of particles from each PF block. This algorithm proceeds sequentially to hierarchically reconstruct particles as described below.

First, PF muons are formed from global muons whose momentum is compatible with that determined by only using the tracker. The corresponding tracks are then removed from the PF block.

Next, PF electrons are identified by using information from the inner tracker and calorimeters. Electron candidates in a PF block are seeded by tracks with links to ECAL clusters. These tracks are then re-extrapolated to the ECAL, and if a track is found to be compatible with ECAL energy deposits and consistent with an electron, the track and clusters are labelled a PF electron and are removed from the PF block.

The remaining elements in the PF block are used to form charged hadrons, photons, neutral hadrons, and, in rare cases, additional muons. PF elements are identified as one of these particle-types by comparing the track momentum to the linked cluster energies. The following scenarios define the identification process:

- If the total cluster energy is much smaller than the track momentum, the excess track momentum is labeled as a muon or fake track. This occurs for less than 0.03% of the tracks used in the algorithm.
- If the total cluster energy agrees within the uncertainty of the track momentum, a PF charged hadron is formed. The PF charged hadron is then assigned a mass equal to that of a charged pion and a momentum based on a fit of the tracker and calorimeter measurements.
- If the total cluster energy is significantly larger than the track momentum and the

excess is greater than the total ECAL energy, then the track is considered a PF charged hadron, as described above. The excess ECAL energy is labeled a PF photon, and the remaining energy is assigned to a PF neutral hadron. The excess ECAL energy is preferentially given to photons over neutral hadrons, as typically photons account for 25% of the energy of a jet, while neutral hadrons only account for 3%.

- If the total cluster energy is significantly larger than the track momentum and the excess is less than the total ECAL energy, the track is considered a PF charged hadron and the excess calorimeter energy is assigned as a PF photon.
- If there are ECAL or HCAL clusters without any linked tracks, the deposits are respectively treated as PF photons and PF neutral hadrons.

5.4 Leptons

As the identification criteria for selecting PF electrons and PF muons are loose, these objects serve as candidate particles. To increase the purity of true electrons/muon, more stringent criteria must be passed for a PF electron/muon to be considered an analysis-level electron/muon. The number of analysis-level leptons, where a lepton is defined as either an electron or muon, is denoted as N_{leps} .

5.4.1 Electrons

The electron candidates are required to have $p_T > 20$ GeV, $|\eta| < 2.5$, and to satisfy identification criteria [66] designed to remove hadrons misidentified as electrons, photon conversions, and electrons from heavy-flavor hadron decays. This criteria is shown in Table 5.1, where $\sigma_{i\eta i\eta}$ is a variable based on the width of the electron shower shape, and

Criteria	Barrel requirement	Endcap requirement
$\sigma_{in\eta}$	< 0.0101	< 0.0283
$\Delta\eta(\text{cluster, track})$	< 0.0103	< 0.07333
$\Delta\phi(\text{cluster, track})$	< 0.0336	< 0.114
$E_{\text{hadronic}}/E_{\text{electromagnetic}}$	< 0.876	< 0.0678
$\frac{1}{E} - \frac{1}{p} \text{ [GeV}^{-1}\text{]}$	< 0.0174	< 0.0898
$ d_0 \text{ [mm]}$	< 0.0118	< 0.0739
$ d_z \text{ [mm]}$	< 0.373	< 0.602
Missing hits	≤ 2	≤ 1
Pass photon conversion	True	True

Table 5.1: Identification criteria that a PF electron must pass in order to be considered an analysis-level electron.

d_0 and d_z are the transverse and longitudinal impact parameters of the associated electron track, respectively.

Additionally, to preferentially select electrons that originate from the decay of W and Z bosons, electrons are required to be isolated from other PF candidates. The relative isolation of a particle I^{rel} is quantified using an optimized version of the mini-isolation variable I^{mini} . Mini-isolation is computed as the scalar sum of the p_{T} of charged hadrons from the PV, neutral hadrons, and photons that are within a cone of radius $R^{\text{mini-iso}}$ surrounding the electron momentum vector $\vec{p}$ in η - ϕ space [67]. The cone radius $R^{\text{mini-iso}}$ varies with $1/p_{\text{T}}$ according to

$$R^{\text{mini-iso}} = \begin{cases} 0.2, & p_{\text{T}} \leq 50 \text{ GeV} \\ 10 \text{ GeV}/p_{\text{T}}, & 50 < p_{\text{T}} \leq 200 \text{ GeV} \\ 0.05, & p_{\text{T}} > 200 \text{ GeV}. \end{cases} \quad (5.1)$$

The p_{T} -dependent cone size reduces the rate of accidental overlaps between the electron and jets in high-multiplicity events or highly Lorentz-boosted decays and in particular overlaps between bottom quark jets and leptons originating from a boosted top quark.

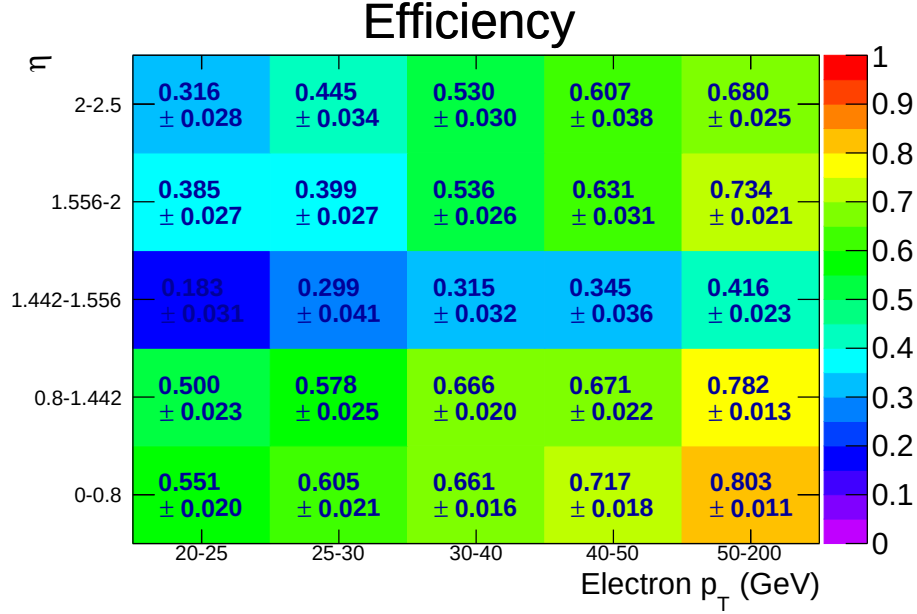


Figure 5.1: The efficiency to select an analysis-level electron as a function of p_T and η . The low efficiency for $1.442 \leq |\eta| < 1.556$ corresponds to the ECAL “crack” region, the boundary between the EB and EE, in which electron reconstruction is particularly difficult.

Relative isolation is computed as $I^{\text{rel}} = I^{\text{mini}}/p_T$ after subtraction of the average contribution from additional proton-proton collisions in the same bunch-crossing (pileup). To be considered isolated, electrons must satisfy $I^{\text{rel}} < 0.1$.

The combined efficiency for the electron reconstruction, identification, and isolation requirements, shown in Figure 5.1, is about 50% at p_T of 20 GeV, increasing to 65% at 50 GeV, and reaching a plateau of 80% above 200 GeV [68].

5.4.2 Muons

The muon candidates are required to have $p_T > 20$ GeV, $|\eta| < 2.4$, and to satisfy identification criteria [69] in order to select a high purity muon sample. This criteria is shown in Table 5.2, where d_0 and d_z are the transverse and longitudinal impact parameters of the associated muon track, respectively. Analogously to electron candidates, muon

Criteria	Requirement
Is PF muon	True
Fraction of valid tracker hits	> 0.8
AND	
$ d_0 $ [mm]	< 2
$ d_z $ [mm]	< 5
Is global muon	True
Normalized global-track χ^2	< 3
χ^2 of tracker-standalone position match	< 12
Track-kink χ^2	< 20
Segment compatibility	> 0.303
OR	
Segment compatibility	> 0.451

Table 5.2: Identification criteria that a PF muon must pass in order to be considered an analysis-level muon.

candidates are required to satisfy $I^{\text{rel}} < 0.2$, where the looser threshold is to account for purity differences between electrons and muons.

The efficiency for reconstructing muons, shown in Figure 5.2 is about 70% at p_T of 20 GeV, increasing to 80% at 50 GeV, and reaching a plateau of 95% for $p_T > 200$ GeV [70].

5.5 Jets

When a quark or gluon is produced at the LHC, they quickly hadronize due to color confinement and produce a collimated “spray” of particles, called a jet, which is the direct detector observable. The parameter of interest, however, is the momentum of the initial parton before hadronization. Thus, clustering the constituent jet particles in a way that accurately reconstructs the initial parton momentum is essential. Events, however, often contain multiple jets with each jet typically composed of some ~ 10 -100 particles that are incident on many detector channels across a large area. This makes the problem of jet clustering non-trivial and an important aspect of object reconstruction.

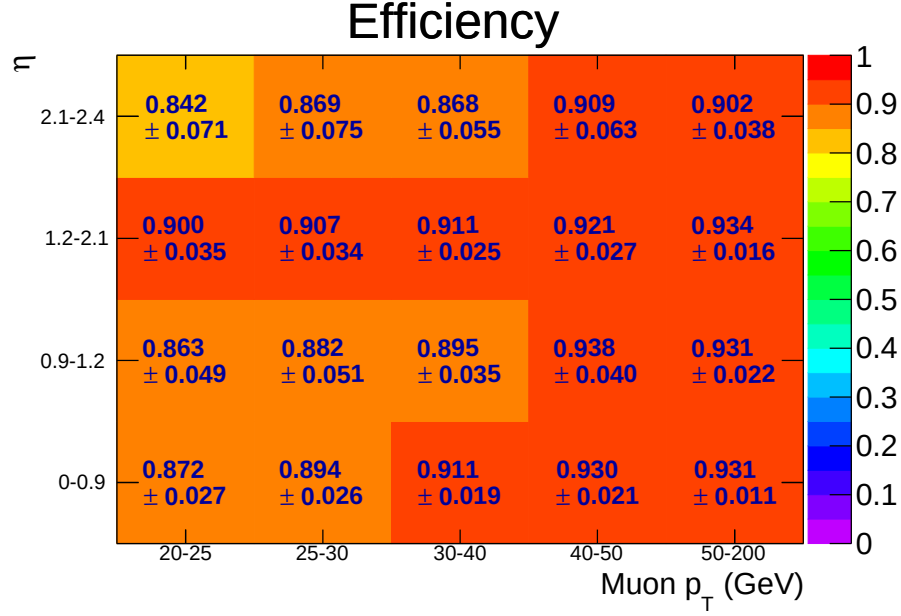


Figure 5.2: The efficiency to select an analysis-level muon as a function of p_T and η .

5.5.1 Clustering

Jets are not physically-defined objects and are instead defined by the clustering rules that form the jets. The primary class of clustering algorithms used are sequential recombination algorithms, which work by defining a distance measure between pairs of particles, typically based on energy and spatial-location, and then combining the closest pair of particles. This process proceeds sequentially and ends when some threshold condition is met.

Commonly, the distance measures, d_{ij} , which is the distance between two entities (particles or psuedojets), and d_{iB} , which is the distance between an entity and the beam, are defined to be

$$d_{ij} = \min(p_{Ti}^{2n}, p_{Tj}^{2n}) \frac{\Delta R_{ij}^2}{R^2} \quad (5.2)$$

and

$$d_{iB} = p_{Ti}^{2n}, \quad (5.3)$$

where $\Delta R_{ij}^2 = \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$, R is the jet radius parameter that sets the scale of the jet's size, and n is a parameter of the algorithm. The clustering procedure then proceeds by finding the smallest of the distances, and if it is a d_{ij} , the two entities i and j are merged, while if it is a d_{iB} , entity i is called a jet and is removed from the clustering list.

The parameter n determines the clustering order of the algorithm and thus different choices for the value of n lead to different clustering algorithms. For $n = 1$, the clustering follows the k_T algorithm [71], which prioritizes clustering softer particles first. Setting $n = 0$, reproduces the Cambridge-Aachen algorithm [72], which is an energy-independent approach, relying only on the spatial distances between particles. Lastly, choosing $n = -1$, gives the anti- k_T algorithm [73], which favors using the harder particles as the jet seeds and then clustering around them. A more detailed discussion of jets, including a comparison of these and other clustering algorithms can be found in Reference [74].

5.5.2 Selection

The jets used in this dissertation are constructed by clustering PF candidates with the anti- k_T algorithm and $R = 0.4$, using the FASTJET package [75]. To reduce the effect of pileup on the jet clustering and energy measurements, a process called ‘‘Charged Hadron Subtraction’’ is applied, where PF charged hadrons that do not originate from the PV are not included in the jet clustering. In addition, to remove the neutral energy component of pileup, the contribution from PF neutral hadrons produced by pileup is estimated based on the area of a jet and the energy density of the event and is subtracted from the jet [76]. Additionally, to prevent double-counting, any jets which contain a PF candidate

Criteria	Requirement
Number of constituents	> 1
Charged multiplicity	> 0
Neutral electromagnetic fraction	< 0.99
Neutral hadron fraction	< 0.99
Charged electromagnetic fraction	< 0.99
Chaged hadron fraction	> 0

Table 5.3: Identification criteria that a jet candidate must pass in order to be considered an analysis-level jet.

identified as an analysis-level electron or muon are removed from the jet collection.

Finally, to be considered an analysis-level jet, the jets must have $p_T > 30$ GeV, $|\eta| < 2.4$, and must pass a set of loose identification requirements [77, 78] to surpress, for example, calorimeter noise. These requirements are shown in Table 5.3. The resulting jets are considered to be “small- R ” jets and the variable N_{jets} represents the number of these jets in an event. Additionally, a proxy for the hadronic energy scale of an event, H_T , is defined as

$$H_T = \sum_{i=1}^{N_{\text{jets}}} p_{T,i}^{\text{jet}} \quad (5.4)$$

5.5.3 b-tagging

Jets that are formed through b-quark hadronization have several unique properties that allow them to be differentiated from jets formed from other quarks or gluons. This ability to “b-tag” jets is a very useful tool for determining what physics processes occurred in an event, as b quarks are associated with specific physics process, such as top quark decays. Similarly, many SUSY scenarios, where naturalness considerations motivate light third-generation squarks, result in either the direct or indirect production of b quarks through the decay of SUSY particles. Thus, b-tagging jets is not only often a crucial

component to reducing backgrounds from processes where no b quarks are expected, but also a powerful selector for potential signal events.

This analysis uses the Combined Secondary Vertex v2 (CSVv2) algorithm [79, 19], which utilizes the long lifetimes, large masses, high-momentum daughter particles, and frequent semi-leptonic decays typical of b hadrons to b-tag jets. As the b quark can only decay to an up or charm quark through highly Cabibbo suppressed weak interactions, b hadrons tend to have long lifetimes, typically on the order of 1.5 ps. Because of this, b mesons can travel a few mm to a cm from the PV before decaying and producing displaced tracks from which a secondary vertex can be reconstructed. In addition, due to the relatively large b-quark mass, b mesons tend to be heavy, which leads to both large secondary vertex masses and daughter particles with a hard momentum spectrum. Lastly, the weak decay of the b quark results in an associated electron or muon in about 20% of decays. The presence of these soft, nonisolated leptons provides an additional marker for the presence of a b-quark.

The CSVv2 algorithm exploits variables based on this information about secondary vertices, their associated displaced tracks, and the presence of soft leptons to accurately tag b-quark jets. A selection of the variables used that have high discrimination are listed below:

- The significance of the flight distance in the transverse plane
- The number of SV
- The SV mass
- The number of tracks associated with the SV
- The ratio of the transverse momentum of the SV tracks and the transverse momentum of the jet

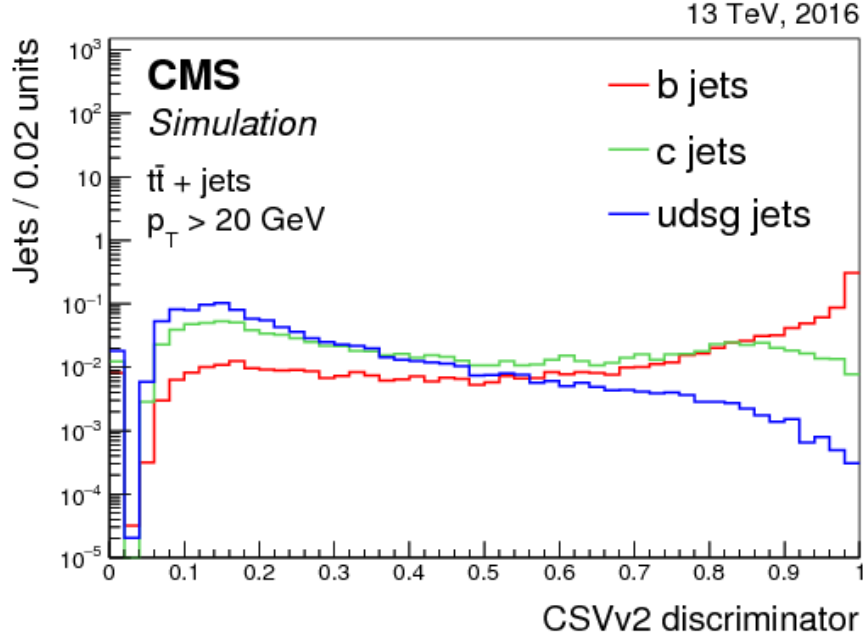


Figure 5.3: The distribution of the CSVv2 discriminator values for jets of different flavors. Jets are selected from $t\bar{t}$ events and required to have $p_T > 20$ GeV [19].

- The 3D impact parameter of soft leptons associated with the jet

These variables are then fed into a multilayer perceptron with one hidden layer that outputs a score between 0 and 1, indicating the likelihood the jet is a b-quark jet. The distribution of CSVv2 discriminator values for different flavor jets is shown in Figure 5.3.

A threshold score of 0.8484 is used to b-tag a jet and is chosen such that the mistag rate for light-flavor jets is 1%. This corresponds to a mistag rate for c-quark jets of 13–15% (11–13%) in the barrel (endcap) and a tagging efficiency for b jets of 60–67% (51–57%) in the barrel (endcap) for jets with p_T between 30–50 GeV. The tagging efficiency increases with p_T before decreasing to $\approx 50\%$ for jets above 150 GeV. The b-tagging efficiency as a function of jet p_T is shown in Figure 5.4.

The number of b-tagged jets in an event is denoted as N_b .

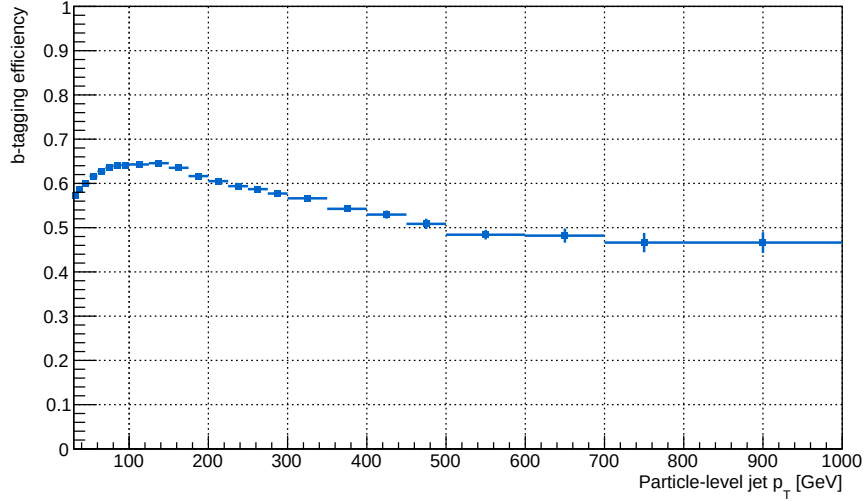


Figure 5.4: The efficiency of the CSVv2 algorithm as a function of jet p_T at the working point used in this analysis.

5.6 Large-radius Jets

While the distance parameter of small- R jets is optimized for clustering the hadronization products of a single parton, it is often useful to exploit information of physical processes on a scale larger than a single jet, such as top quark or W boson decays. One way to capture this information is to cluster jets with a large distance parameter, which can encode the momentum, angular, and multiplicity information of the partons contained in this larger jet.

This analysis constructs these “large- R ” jets by clustering small- R jets with a distance parameter of $R = 1.2$. Due to the relatively small number of small- R jets in an event ($\lesssim 10$), the construction of these large- R jets is insensitive to the clustering algorithm, and the anti- k_T algorithm is chosen. While these large- R jets could be constructed by performing the clustering at the PF candidate level, no significant improvement in performance was noted. Thus, clustering small- R jets were chosen for the following practical considerations. First, the FASTJET implementation of the anti- k_T algorithm

has complexity $\mathcal{O}(n \log n)$ [80], which results in a speed-up of on the order of 100x when clustering small- R jets ($\lesssim 10$ objects) compared to PF candidates (~ 1000 objects). Secondly, large- R jets clustered from PF candidates would require the computation of new energy-measurement and pileup-removal calibrations for jets of this specific radius. Small- R jets, however, already incorporate standardized calibrations and by clustering these calibrated small- R jets, the large- R jets correspondingly incorporate corrections without the need of any additional development.

In addition to small- R jets, the jets associated with selected leptons are included in the formation of large- R jets in order to capture as much of an event's kinematic information as possible. For example, this helps reduce the difference between large- R jets formed by clustering hadronic and leptonic top quark decays, as by including the lepton, the only information difference between the two scenarios is due to the undetected neutrino.

This technique of clustering small- R jets into large- R jets has been used previously by both the ATLAS and CMS collaborations, e.g. References [81, 82].

5.6.1 M_J — The Sum of Large-radius Jet Masses

A measure of the mass-scale of an event, M_J , can be constructed by summing the masses of large-radius jets, defined as

$$M_J = \sum_{J_i \in \text{large-}R \text{ jets}} m(J_i) \quad (5.5)$$

where $m(J)$ is the mass of a single large- R jet.

The quantity M_J has significant discriminating power between SM background processes and signal processes, as SM events tend to have significantly lower mass-scales than signal events that involve high mass particles. For example, M_J in events with only

a produced $t\bar{t}$ pair is limited to be $\lesssim 2m_{top} \approx 350$ GeV. This is because the top quarks decay back-to-back and if each top quark's decay products are captured in a single large- R jet, two large- R jets are clustered each with $m(J) = m_{top}$, resulting in $M_J = 2m_{top}$. In the case where a top quark is not fully contained in a large- R jet, the mass of that large- R jet, and correspondingly M_J , is even smaller. For signal events, however, the mass-scale is roughly set by the gluino mass, which is on the order of 1 TeV.

The M_J distribution in $t\bar{t}$ and signal events with $m_{\tilde{g}} = 1000$ GeV and 1600 GeV, selected with a $N_{\text{jets}} \geq 8$ requirement to ensure both processes have similar N_{jets} distributions, is shown in Figure 5.5. From the figure, it can be seen that the M_J distribution gets increasingly harder with higher gluino mass. Also of note is that the M_J distribution in $t\bar{t}$ events extends past $2m_{top}$. This is because, in the presence of significant initial state radiation (ISR), the ISR jets can either overlap with the $t\bar{t}$ daughter jets or boost the $t\bar{t}$ system such that the system is collimated, both which result in high-mass large- R jets and, correspondingly, high- M_J . Processes of this nature are responsible for generating high M_J background events.

The quantity M_J was proposed in phenomenological studies [83, 84, 85] and was first used for RPC SUSY searches by the ATLAS Collaboration in all-hadronic final states [86, 87] and by the CMS Collaboration in single-lepton events [82, 88]. Additionally, the basic properties and performance of the M_J variable were commissioned using early $\sqrt{s} = 13$ TeV data [89].

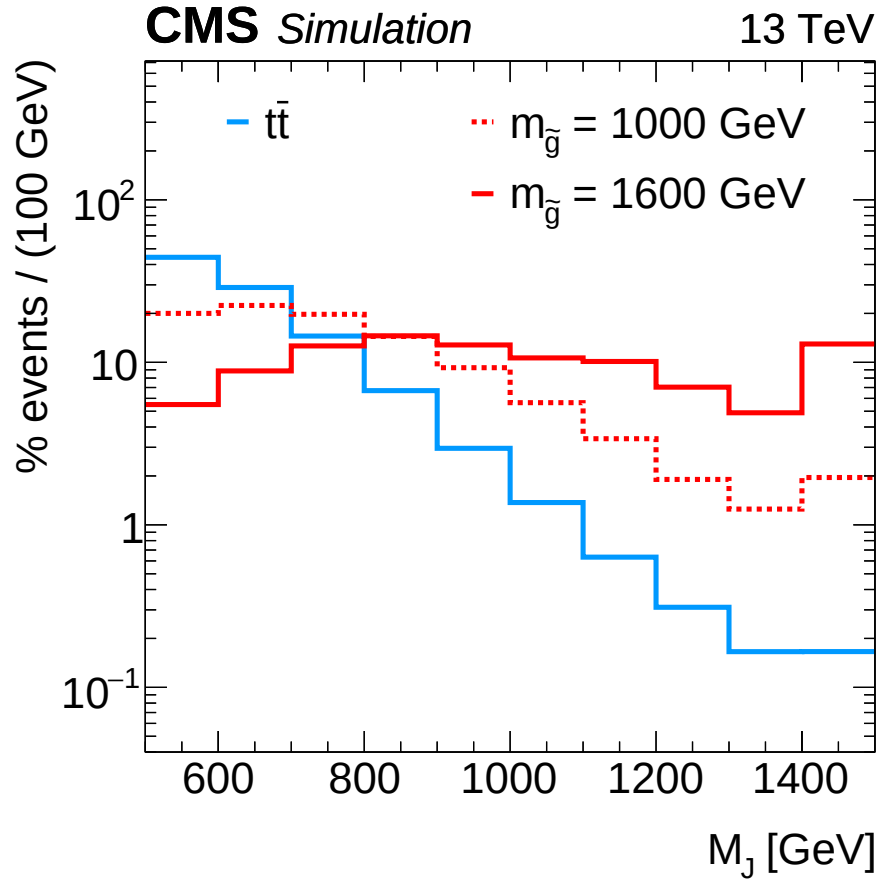


Figure 5.5: Distributions of M_J , normalized to the same area, for $t\bar{t}$ events (solid blue) and signal events with $m_{\tilde{g}} = 1000$ GeV (dashed red) and 1600 GeV (solid red) in a selection of $N_{\text{leps}} = 1$, $H_T > 1200$ GeV, $N_{\text{jets}} \geq 8$, $M_J > 500$ GeV, and $N_b \geq 1$.

Part III

The Search for SUSY

Chapter 6

Data Samples and Simulation

6.1 Data

The dataset used in this search corresponds to 35.9 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ collected by the CMS detector over the year 2016. This is a subset of the 40.8 fb^{-1} delivered by the LHC and selected to correspond to when all sub-detectors were fully-operational. A plot of the cumulative delivered and recorded integrated luminosity by the LHC and CMS, respectively, is shown in 6.1.

The data samples used in this analysis are shown in Table 6.1. The **JetHT** datasets are primarily used to populate the analysis region with events, while the **Single Electron** and **Single Muon** datasets are primarily used for trigger efficiency studies.

6.2 Monte Carlo Simulation

Monte Carlo simulations are used to model both SM and BSM physics processes and are extremely useful in the design and optimization of the analysis strategy of new-physics searches. These simulated samples allow for studying processes that in data would have control regions with impure and/or small sample sizes or, in the case of signal processes, may not even exist. This allows for both the optimization and validation of the (signal-

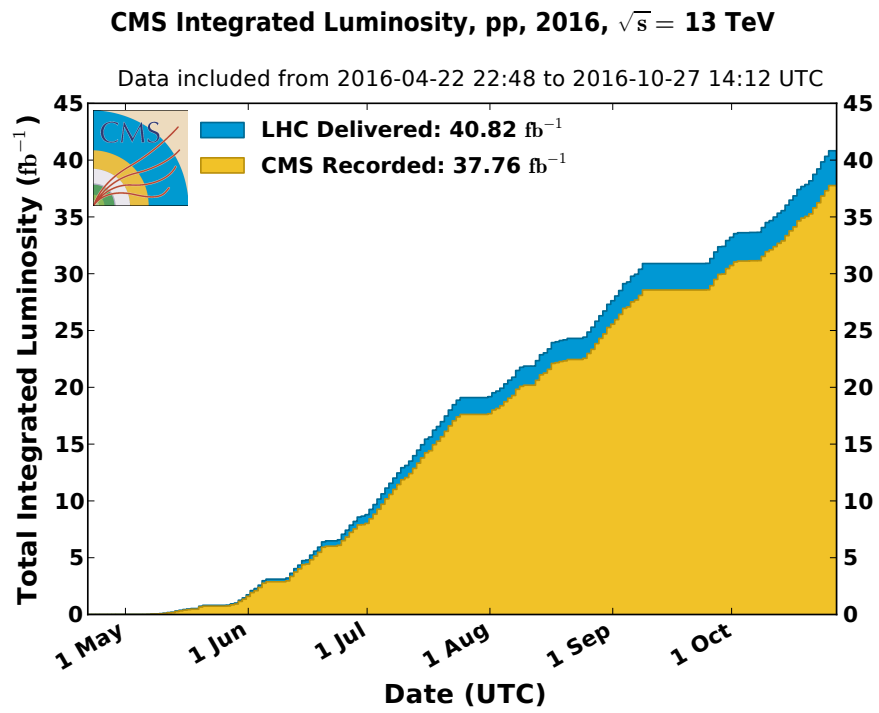


Figure 6.1: Delivered and recorded integrated luminosity by the LHC and CMS, respectively, over 2016 [20].

Dataset name
/JetHT/Run2016B-03Feb2017_ver2-v2/MINIAOD
/JetHT/Run2016C-03Feb2017-v1/MINIAOD
/JetHT/Run2016D-03Feb2017-v1/MINIAOD
/JetHT/Run2016E-03Feb2017-v1/MINIAOD
/JetHT/Run2016F-03Feb2017-v1/MINIAOD
/JetHT/Run2016G-03Feb2017-v1/MINIAOD
/JetHT/Run2016H-03Feb2017_ver2-v1/MINIAOD
/JetHT/Run2016H-03Feb2017_ver3-v1/MINIAOD
/SingleElectron/Run2016B-03Feb2017_ver2-v2/MINIAOD
/SingleElectron/Run2016C-03Feb2017-v1/MINIAOD
/SingleElectron/Run2016D-03Feb2017-v1/MINIAOD
/SingleElectron/Run2016E-03Feb2017-v1/MINIAOD
/SingleElectron/Run2016F-03Feb2017-v1/MINIAOD
/SingleElectron/Run2016G-03Feb2017-v1/MINIAOD
/SingleElectron/Run2016H-03Feb2017_ver2-v1/MINIAOD
/SingleElectron/Run2016H-03Feb2017_ver3-v1/MINIAOD
/SingleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD
/SingleMuon/Run2016C-03Feb2017-v1/MINIAOD
/SingleMuon/Run2016D-03Feb2017-v1/MINIAOD
/SingleMuon/Run2016E-03Feb2017-v1/MINIAOD
/SingleMuon/Run2016F-03Feb2017-v1/MINIAOD
/SingleMuon/Run2016G-03Feb2017-v1/MINIAOD
/SingleMuon/Run2016H-03Feb2017_ver2-v1/MINIAOD
/SingleMuon/Run2016H-03Feb2017_ver3-v1/MINIAOD

Table 6.1: Data samples analyzed for this analysis. The corresponding integrated luminosity is 35.9 fb^{-1} .

plus-)background prediction methods, as the sensitivity of the analysis to particular signal models can be estimated and pathologies in the prediction methodology can be unearthed and investigated. Additionally, the simulated samples can be used to help commission and understand the collected data by comparing expectations from simulation to what is actually observed, particularly in the initial data taking periods.

6.2.1 Background Samples

The MADGRAPH5_aMC@NLO 2.2.2 event generator is used in leading-order (LO) mode [90, 91] to generate the $t\bar{t}$, quantum chromodynamics multijet (QCD), W + jets and Drell–Yan background processes with extra partons, while the $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}t\bar{t}$, and t -channel single top quark production backgrounds are generated with MADGRAPH5_aMC@NLO 2.2.2 in next-to-leading order (NLO) mode [92]. The POWHEG 2.0 event generator [93, 94, 95] is used to generate the tW , $\bar{t}W$, and s -channel single top quark processes at NLO precision.

The $t\bar{t}$, W + jets, and QCD samples are generated with up to 2, 4, and 2 extra partons, respectively, and all samples are generated using a top quark mass of 172.5 GeV and with the NNPDF3.0 set of parton distribution functions (PDF) [96]. For the fragmentation and showering of partons, the generated samples are interfaced with PYTHIA 8.205 [97] and use the CUETP8M1 tune to describe the underlying event [98]. The detector response is simulated with GEANT4 [99], and the simulated samples are processed through the same reconstruction algorithms as the data, as discussed in Chapter 5. The background samples are normalized to the highest precision cross sections available [100, 101, 102, 103, 104, 105, 106].

The background samples used, along with their sample size and equivalent luminosity, are shown in Table 6.2.

6.2.2 Signal Samples

For ease of generation and interpretation, signal samples are produced according to the Simplified Model Spectra (SMS) paradigm [107, 108], where all but one or two mass parameters in a particular decay chain are fixed. Due to their “simplicity”, SMS models can be interpreted generally and have reduced sensitivity to model specifics, allowing for

Simulated sample name	Events	L [fb^{-1}]
TTJets_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	10,259,790	12.57
TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	53,056,561	296.90
TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	60,282,318	337.34
TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	30,681,952	358.18
TTJets_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	13,838,472	5,291.21
TTJets_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	10,506,985	9,753.77
TTJets_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2,913,606	14,943.68
TTJets_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	523,826	225,454.60
QCD_HT100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	82,072,813	0.00
QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	57,336,294	0.03
QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	54,706,023	0.15
QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	63,336,989	2.16
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	45,232,792	6.93
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	15,314,987	14.39
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	11,647,431	95.86
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6,004,039	236.19
WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	28,210,241	0.46
WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	27,546,847	16.90
WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	19,851,490	45.57
WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	7,432,643	125.41
WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18,132,628	1,243.62
WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1,540,354	231.42
WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	7,012,526	4,360.78
WJetsToLNu_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2,505,140	64,376.98
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	49,190,673	8.17
DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	7,558,769	44.08
DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	8,683,638	165.14
DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	396,532	58.65
DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	8,231,815	4,910.15
DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2,650,562	3,188.24
DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	616,458	4,320.56
DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	375,865	117,894.02
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	999,995	116.20
ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	1,682,394	64.14
ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	3,279,179	74.41
ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	5,388,666	343.66
ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	5,405,674	344.74
ttHJetTobb_M125_13TeV_amcatnloFXFX-madspin-pythia8	9,823,967	2,957.09
TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	4,664,534	132.40
TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8	752,497	14,412.60
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	252,664	328.75
TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	833,278	547.03
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	398,596	340.36
TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	749,386	310.64

Table 6.2: Simulated background samples used in this analysis with their corresponding sample size and equivalent luminosity.

results to be applicable to a wide variety of new-physics models. Additionally, the use of SMS models results in significantly reduced computing time for model generation, as models need to be produced by scanning across at most only two parameters.

The assumptions used in the simplified model of the $\tilde{g} \rightarrow t\bar{t} \rightarrow tbs$ process, denoted as T1tbs, are given below:

- squarks other than the top squark are much heavier than the gluino, so they do not affect the gluino decay
- the branching ratio of $\tilde{g} \rightarrow t\bar{t} \rightarrow tbs$ is 100%
- the top squark is virtual in its decay. This results in a three-body decay, so searches for dijet resonances, i.e., $\tilde{t} \rightarrow bs$, are not applicable in this scenario.
- the gluinos decay promptly

These assumptions correspond to a model of direct gluino pair production with each gluino decaying to a top, bottom, and strange quark. An example diagram of the T1tbs model is shown in Figure 6.2. The only free parameter in this model is the mass of the gluino.

These signal models are generated between $m_{\tilde{g}} = 1000$ and 2000 GeV in steps of 100 GeV with up to two extra partons in leading-order mode and dynamic factorization and renormalization scales by MADGRAPH5_aMC@NLO 2.2.2. The rest of the signal sample production follows the same procedure used for the background samples, which includes the parton fragmentation and showering, simulation, and event reconstruction. Lastly, the samples are normalized to the NLO + next-to-leading logarithmic cross sections for gluino pair production [21].

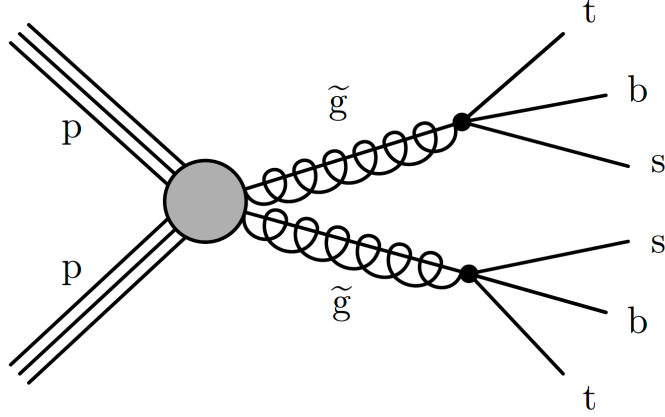


Figure 6.2: Example diagram of the T1tbs simplified model.

Dataset name	Events	L [fb^{-1}]
SMS-T1tbs_mGluino-1000_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	130,281	400.39
SMS-T1tbs_mGluino-1100_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	71,488	437.26
SMS-T1tbs_mGluino-1200_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	50,739	592.46
SMS-T1tbs_mGluino-1300_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	31,815	690.84
SMS-T1tbs_mGluino-1400_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	21,392	845.61
SMS-T1tbs_mGluino-1500_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	16,702	1,177.00
SMS-T1tbs_mGluino-1600_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	10,101	1,246.92
SMS-T1tbs_mGluino-1700_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	10,380	2,206.99
SMS-T1tbs_mGluino-1800_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	9,865	3,572.55
SMS-T1tbs_mGluino-1900_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	10,071	6,157.86
SMS-T1tbs_mGluino-2000_mLSP-0_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	10,556	10,759.60

Table 6.3: Simulated signal samples used in this analysis with their corresponding sample size and equivalent luminosity.

Chapter 7

Event Selection

7.1 Baseline Selection

One of the main challenges for a SUSY search is that the ratio of SM-to-SUSY events is $\mathcal{O}(10^{12} - 10^{14})$. To surmount this problem, it is paramount to develop highly efficient signal-to-background discriminators. Fortunately, SUSY signatures typically have characteristics unlike most SM processes. For the T1tb process, events are expected to have a large number of jets, many of which are b-quark jets, resulting in a large amount of H_T and many b-tagged jets. Additionally, the mass scale of the event is expected to be larger than most SM events due to the high masses of the gluinos ($\mathcal{O}(1 \text{ TeV})$), as discussed in Section 5.6.1.

These features are used to construct the “baseline selection”, defined as a basic set of requirements that events must pass, which reduces the background-to-signal ratio to a more manageable value of $\mathcal{O}(10^2)$. In this analysis, the baseline selection is defined as $N_{\text{leps}} = 1$, $H_T > 1200 \text{ GeV}$, $M_J > 500 \text{ GeV}$, $N_{\text{jets}} \geq 4$, and $N_b \geq 1$. To see the separation of the signal and background distributions in these variables, the corresponding “N-1” distributions, i.e. the distribution of a variable with the baseline selection applied except for the requirement corresponding to the binned variable, are depicted in Figure 7.1 with the black dashed vertical line representing the value of the baseline cut. Additionally, a

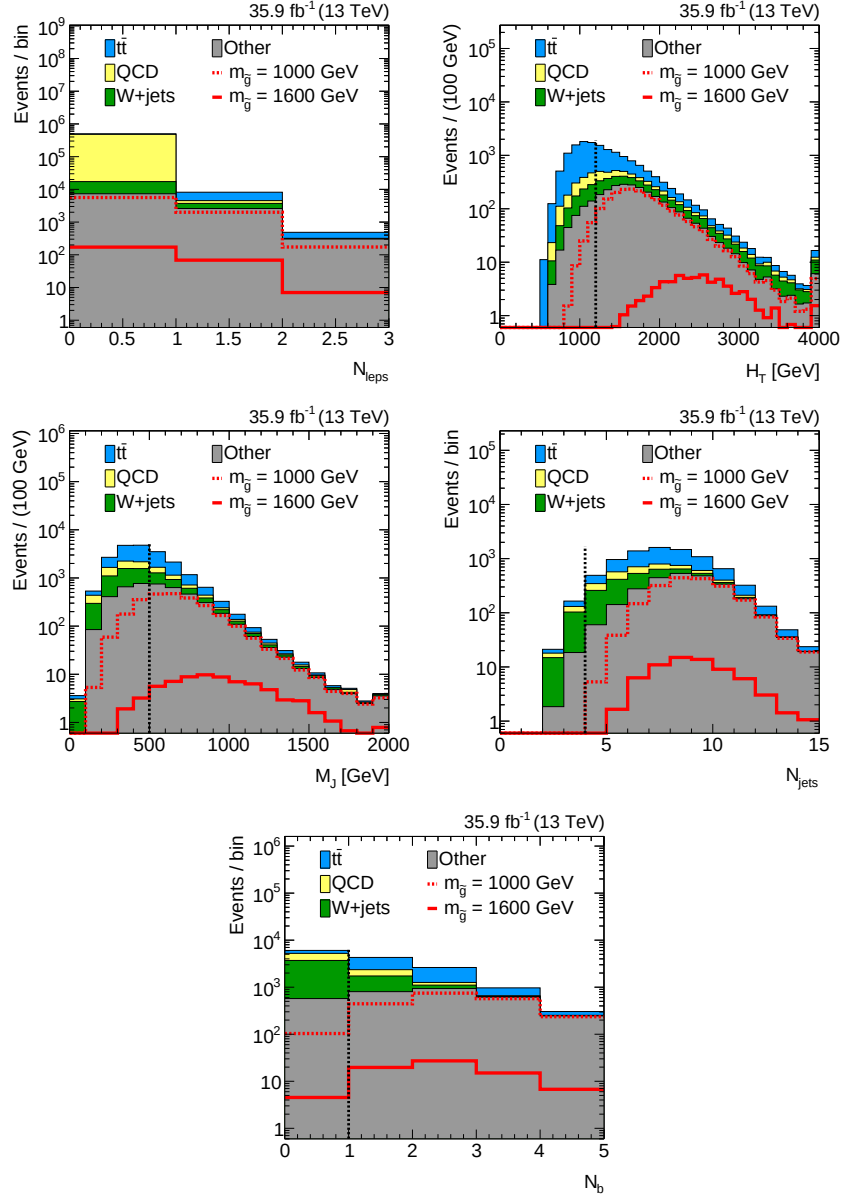


Figure 7.1: The N-1 plots for N_{leps} (top-left), H_T (top-right), M_J (middle-left), N_{jets} (middle-right), and N_b (bottom). The black dashed vertical line represents the value of the corresponding baseline requirement.

“cutflow table” that depicts the expected yields for background and signal processes as each baseline selection requirement is individually applied is shown in Table 7.1. Rows above the horizontal line correspond to requirements in the baseline selection, while those below correspond to additional kinematic cuts.

$\mathcal{L} = 35.9 \text{ fb}^{-1}$	QCD	$t\bar{t}$	W+jets	Other	All bkg.	$m_{\tilde{g}} = 1600 \text{ GeV}$
$H_T > 1200 \text{ GeV}$	1.615×10^7	9.76×10^4	2.718×10^5	2.965×10^4	$(1.655 \pm 0.007) \times 10^7$	2.8×10^2
$N_{\text{leps}} = 1$	1.11×10^4	1.292×10^4	2.502×10^4	4.48×10^3	$(5.35 \pm 0.01) \times 10^4$	7.9×10^1
$N_b \geq 1$	3240	10340	4990	2150	20725 ± 80	74
$N_{\text{jets}} \geq 4$	2770	9920	3740	1870	18304 ± 70	74
$M_J > 500 \text{ GeV}$	810	3658	1120	574	6162 ± 40	67
$M_J > 800 \text{ GeV}$	99	360	150	75	685 ± 9	47
$N_{\text{jets}} \geq 8$	38	200	42	29	309 ± 7	36
$M_J > 1000 \text{ GeV}$	11	43.0	11.3	7.9	73 ± 2	22.6
$N_b \geq 3$	0.7	6.2	0.5	1.1	8.5 ± 0.6	8.9
	QCD	$t\bar{t}$	W+jets	Other	All bkg.	$m_{\tilde{g}} = 1600 \text{ GeV}$

Table 7.1: Expected yields in 35 fb^{-1} from simulations of SM and signal processes. Rows above the horizontal line correspond to requirements in the baseline selection, while those below correspond to additional kinematic cuts.

The $H_T > 1200 \text{ GeV}$ cut sets the hadronic energy scale of the selected events and reduces contributions from processes other than QCD, $t\bar{t}$, and W + jets. At this point, QCD is the dominate background and is reduced significantly by selecting events with $N_{\text{leps}} = 1$, resulting in the QCD, $t\bar{t}$, and W + jets contributions being similar in scale. In order to create a $t\bar{t}$ dominant background, $N_b \geq 1$ is required as $t\bar{t}$ events have at least two b-quark jets, while QCD and W + jets contributions largely pass the selection by the mistagging of at least one jet. Finally, the $N_{\text{jets}} \geq 4$ and $M_J > 500 \text{ GeV}$ cuts are applied in order to further increase the $t\bar{t}$ purity of the analysis region.

A benefit of these selections is that the background is dominated by a single process, $t\bar{t}$, which reduces the complexity of the background prediction. This is especially necessary as, after these cuts, the analysis region is on the kinematic tails of SM distributions where the physics is less well-modelled, typically requiring data-driven background predictions.

Lastly, the baseline selection requires that events pass a series of filters designed to remove poorly reconstructed events. These filters remove events with noise in the HCAL or ECAL, beam halo effects, jets that fail to pass quality criteria, and events with zero good PVs.

A final note of interest is that there is no requirement on the E_T^{miss} , making this

analysis sensitive to BSM models other than MFV SUSY that produce either little or no E_T^{miss} in an event. While the T1tb model is used as a benchmark for interpreting results, this search is structured to take advantage of this feature and be generically sensitive to high-mass signatures with large jet and bottom quark jet multiplicities, which are potential features of other BSM models.

7.2 Trigger Efficiency

The data sample used in this analysis is obtained by selecting events that pass a loose HLT selection. In order to avoid biasing the selected sample, the HLT requirements must be loose enough that the selection efficiency is as high as possible and independent of any kinematic properties. In particular, events must pass an OR of the HLT_PFHT900 trigger, which requires an online H_T of at least 900 GeV and the HLT_PFJet450 trigger, which at least one jet with online p_T above 450 GeV.

Figure 7.2 shows the performance of the HLT_PFHT900 trigger as a function of H_T during the first 27.3 fb⁻¹ (Runs B-G, top-left), last 8.7 fb⁻¹ (Run H, top-right), and full dataset (Runs B-H, bottom). The trigger performances are measured in a data sample collected using the HLT_Ele27_WPTight trigger and offline requirements of at least one electron and at least 4 jets. While the trigger efficiency for Runs B-G is 100% after the trigger plateau of roughly $H_T = 1000$ GeV, the trigger efficiency only performs with 80% efficiency in Run H. This inefficiency was caused by an issue with an updated trigger implementation that erroneously excluded high p_T jets from the online H_T calculation. This effect corresponds to an overall trigger efficiency of 95%, corresponding to a loss of about 2 fb⁻¹ of data.

In order to recover this inefficiency, events passing the HLT_PFJet450 trigger are included in the collected data sample. To pass this trigger, events must have at least

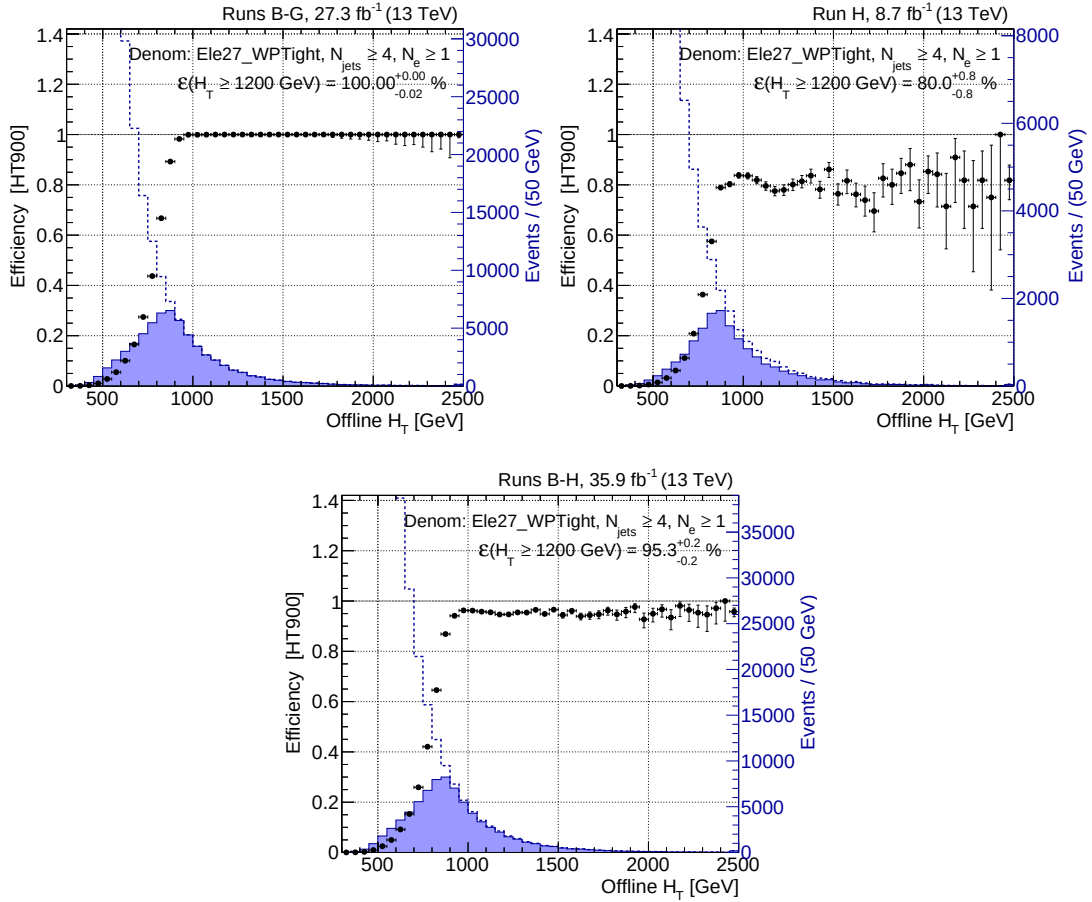


Figure 7.2: Trigger efficiency for HLT_PFHT900 as a function of H_T in Runs B-G (top-left), Run H (top-right), and full dataset (bottom). The efficiencies are measured using a data sample collected with the HLT_Ele27_WPTight trigger and an offline requirement of at least one electron and at least four jets.

one very high p_T jet, i.e. $\gtrsim 450$ GeV, which provides complementary efficiency where the HLT_PFHT900 trigger is inefficient. Figure 7.3 shows the performance as a function of H_T of the HLT_PFJet450 trigger in Runs B-H (top-left) and the combination of the HLT_PFHT900 and HLT_PFJet450 in Run H (middle) and in the full dataset (bottom), measured with a dataset collected with HLT_Ele27_WPTight and offline requirements of at least one electron and at least 4 jets. The inclusion of the HLT_PFJet450 trigger restores the overall trigger efficiency to essentially 100% in both Run H and the entire dataset.

This trigger efficiency, however, does not necessarily correspond to the efficiency for signal events. A lower bound on the signal efficiency can be estimated by considering that the HLT_PFJet450 trigger is fully efficient for jets with $p_T > 500$ GeV and 80%-95% of simulated signal events (depending on the mass of the gluino) have a jet with $p_T > 500$ GeV. Thus, in the worst case scenario, the addition of the HLT_PFJet450 trigger is still expected to recover at least 85% of the lost signal efficiency in Run H. This results in an efficiency of at least 97% in Run H and over 99% for the full dataset for signal events.

Lastly, to ensure that there is no kinematic bias in the trigger efficiency either inherently or from residual effects of the online H_T calculation issue, the trigger efficiency is measured as a function of M_J , N_{jets} , and N_b . The measurements are done in a data sample collected with the HLT_Ele27_WPTight trigger and an offline requirement of at least one electron, at least four jets, and $H_T > 1200$ GeV. No kinematic bias is observed.

7.3 Analysis Binning

After the baseline selection, the background is dominated by $t\bar{t}$ events with small contributions from $W + \text{jets}$ and QCD production. There are additional rare background

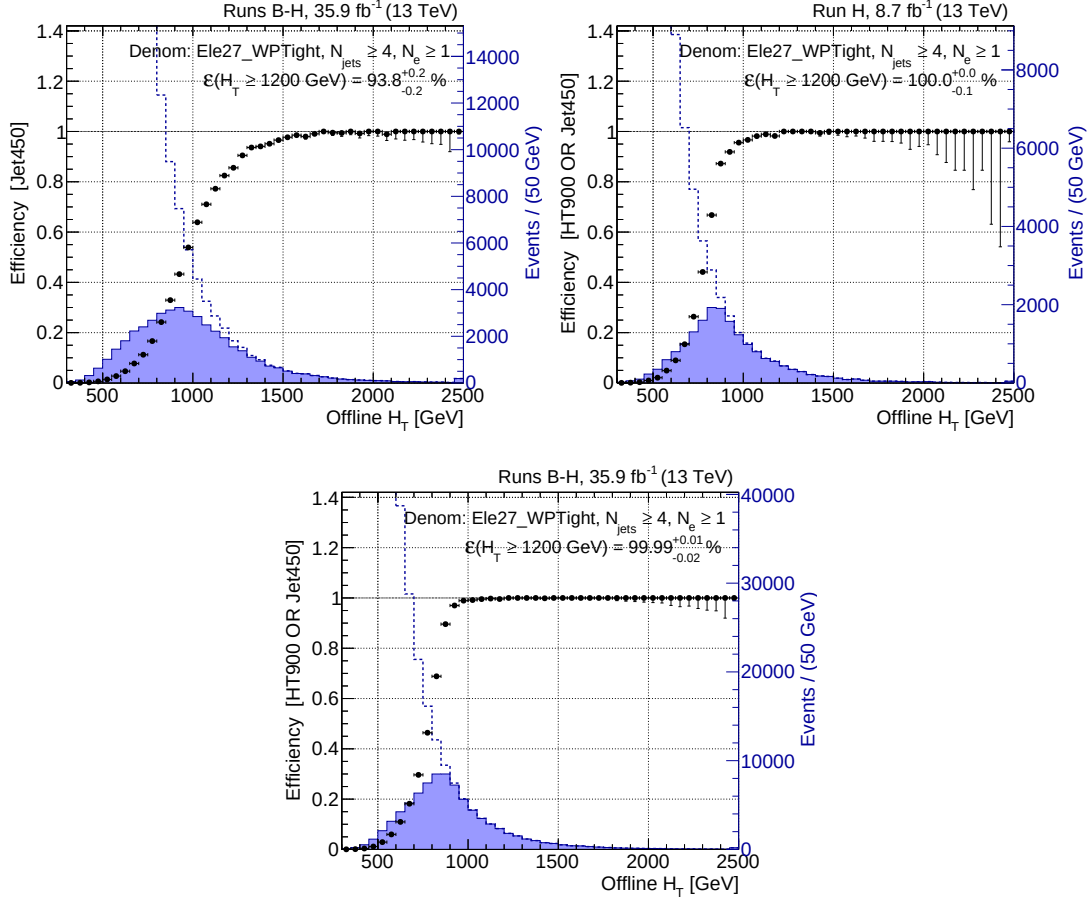


Figure 7.3: Trigger efficiency as a function of H_T for HLT_PFJet450 in the full dataset (top-left) and for the combination of HLT_PFHT900 and HLT_PFJet450 in Run H (top-right) and the full dataset (bottom). The efficiencies are measured using a data sample collected with the HLT_Ele27_WPTight trigger and an offline requirement of at least one electron and at least four jets.

	M_J [GeV]	N_{jets}		
		4–5	6–7	≥ 8
$N_{\text{leps}} = 1$	500 – 800	CR	CR	SR
	800 – 1000	CR	SR	SR
	> 1000		SR	SR

Figure 7.4: Illustration depicting the N_{jets} , M_J binning after the baseline selection, with control and signal region bins denoted by “CR” and “SR”, respectively.

processes, jointly noted as “Other”, with tiny, but non-zero contributions that arise from single top quark, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H$, $t\bar{t}t\bar{t}$, and Drell-Yan production.

In order to further increase the signal-to-background ratio, as well as create background-dominated control regions, the analysis region is binned with respect to N_{jets} and M_J . The N_{jets} bins are defined as $4 \leq N_{\text{jets}} \leq 5$, $6 \leq N_{\text{jets}} \leq 7$, and $N_{\text{jets}} \geq 8$. Each N_{jets} bin is further split into bins of $500 < M_J \leq 800$ GeV, $800 < M_J \leq 1000$ GeV, and $M_J > 1000$ GeV, with the exception of the $4 \leq N_{\text{jets}} \leq 5$ bin for which the two highest M_J bins are combined due to the limited data sample size in the $M_J > 1000$ GeV bin. A diagram representing this binning is shown in Figure 7.4. The low- N_{jets} , low- M_J bins are expected to be background-dominated and are used as control regions for constraining systematics and for validating the prediction methodology, while the high- N_{jets} , high- M_J bins are used as signal regions.

Within each N_{jets} and M_J bin, the N_b distribution is examined for evidence of new physics and is separated into bins of $N_b = 1, 2, 3$, and ≥ 4 . The two lowest N_b bins are used to provide constraints on the background normalizations and systematic uncertainties, while the higher N_b bins are the most sensitive to potential signals due to its larger signal-to-background ratios.

In total, this analysis has 8 kinematic regions—3 control and 5 signal regions with four N_b bins per kinematic region. The simulated N_b distribution of the SM background

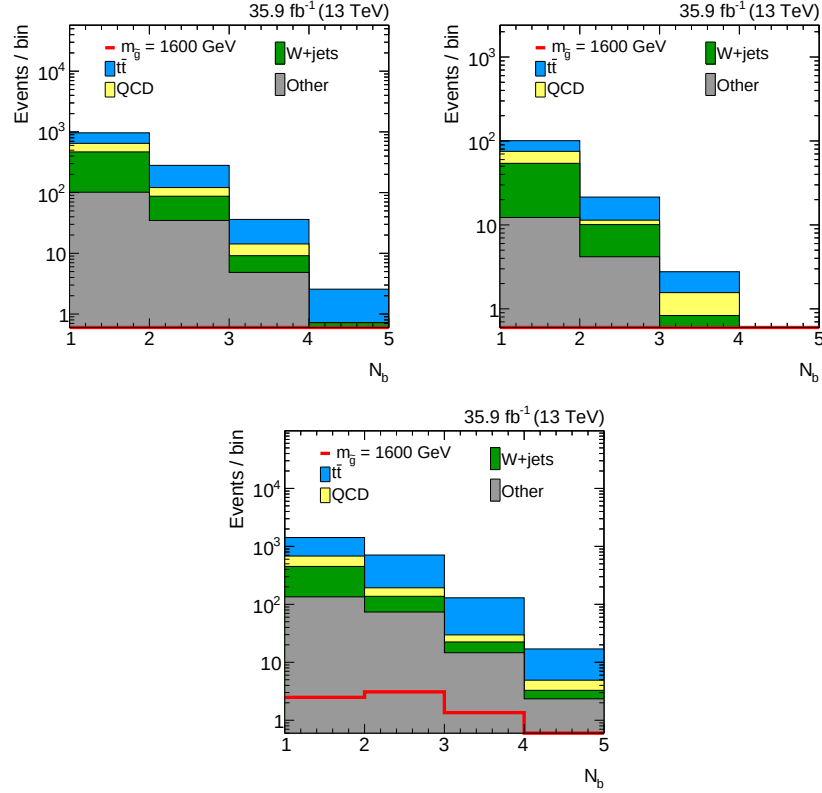


Figure 7.5: The simulated N_b distribution for background and signal processes in the control region bins. The top-left plot corresponds to the $4 \leq N_{\text{jets}} \leq 5$, $500 < M_J \leq 800$ GeV bin, the top-right plot to the $5 \leq N_{\text{jets}} \leq 6$, $M_J > 800$ GeV bin, and the bottom plot to the $6 \leq N_{\text{jets}} \leq 7$, $500 < M_J \leq 800$ GeV bin.

processes and a signal model with $m_{\tilde{g}} = 1600$ GeV for the control and signal regions are shown in Figure 7.5 and Figure 7.6, respectively.

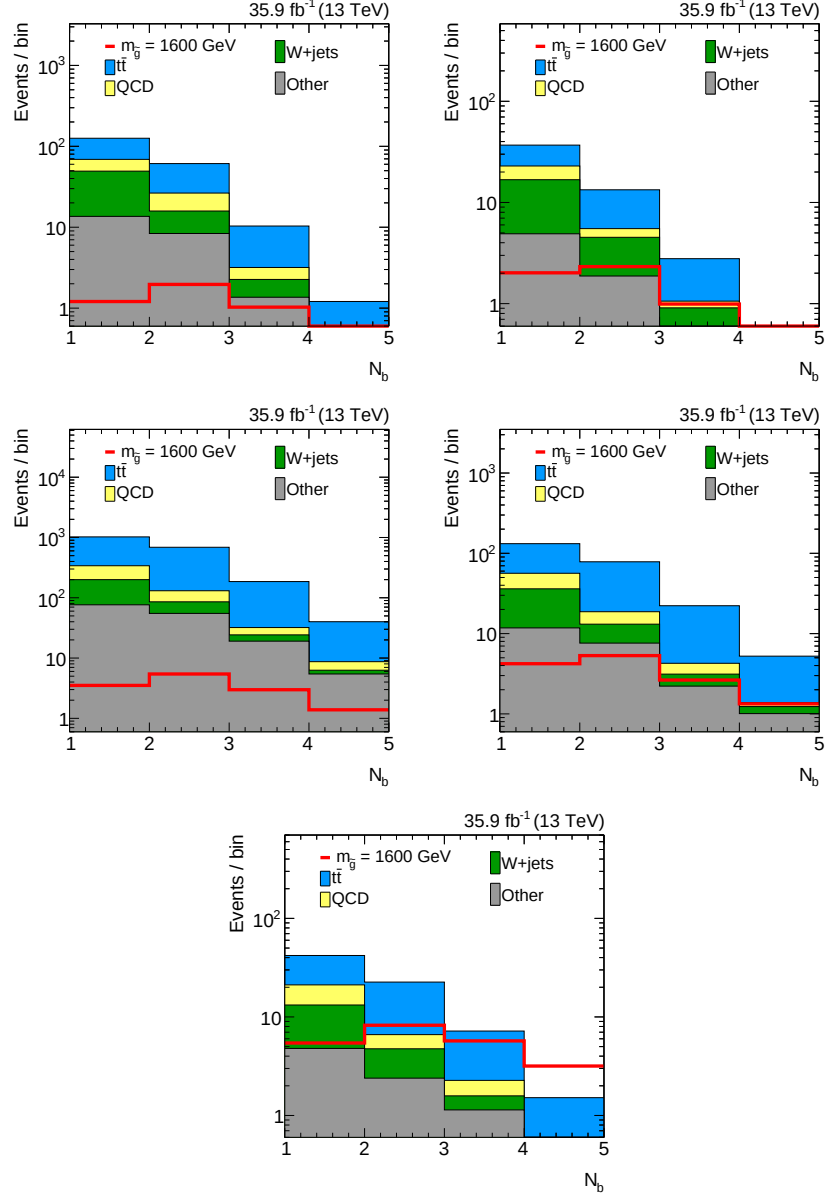


Figure 7.6: The simulated N_b distribution for background and signal processes in the signal region bins. The top-left plot corresponds to the $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV bin, the top-right plot to the $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV bin, the middle-left plot to the $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV bin, the middle-right plot to the $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV, and the bottom plot to the $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV bin.

Chapter 8

Background Prediction

8.1 Overview

This analysis seeks to find evidence of new physics by searching for deviations from the SM in the N_b distribution. In order to do this, it is essential to be able to robustly and accurately predict both the normalization and shape of the N_b distribution. To obtain these predictions, a global maximum-likelihood fit is performed. This fit is carried out both for a background-only hypothesis and for signal-plus-background hypotheses, in which a signal contribution is extracted in addition to the contributions of SM background processes. The model is constructed using the poisson probabilities of the bin contents of the N_b distribution for all N_{jets} , M_J regions, while systematic uncertainties are applied as nuisance parameters.

As the kinematic tails of the N_{jets} and M_J variables are difficult to model reliably, the $t\bar{t}$ and QCD normalizations are individually allowed to (almost) freely vary in each N_{jets} , M_J bin. The $t\bar{t}$ normalizations are constrained in each bin by the low- N_b bins, while the QCD normalizations are constrained by control regions with no identified leptons ($N_{\text{leps}} = 0$). The overall $W + \text{jets}$ normalization is determined from data and is allowed to vary across N_{jets} bins by amounts measured using a kinematically similar $Z + \text{jets}$ sample, while the normalization of Other is largely taken from simulation, as its contribution is

small in the regions considered. Further details on the measurement of the normalizations are given in the following sections.

Once the SM background processes are normalized accordingly, further corrections to the N_b shape are relatively small. The nominal N_b shape prediction for each process is taken from simulation with data-to-simulation correction factors (SFs) applied for the b-tagging efficiency of heavy- and light-flavor jets [79, 19]. This shape is allowed to vary in order to assess the impact of mismodeling of relevant parameters, such as the rate of gluon splitting to $b\bar{b}$ and the b-tagging SFs. The appropriate ranges for these parameters are determined based on measurements in dedicated control samples and then constrained by a simultaneous fit across all bins of N_{jets} and M_J in a correlated manner. A detailed discussion of these variations and their measurements is given in Chapter 9.

8.2 $t\bar{t}$ and QCD Normalizations

The $t\bar{t}$ and QCD normalizations are allowed to float in each N_{jets} , M_J bin but with a loose constraint across M_J bins discussed in the following subsection. The largest constraint on the $t\bar{t}$ normalization in each bin is the background-dominated $N_b \leq 2$ bins, while the QCD normalization in each bin is mostly constrained by corresponding bins in a similar 0-lepton kinematic region selected by requiring $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 6$, and $N_b \geq 1$. The higher H_T requirement compared to the analysis's baseline selection is imposed in order to account for the extra energy in an event carried by the lepton in the $N_{\text{leps}} = 1$ selection, while the higher N_{jets} selection is imposed in order to account for differences in the N_{jets} distribution between the $N_{\text{leps}} = 1$ and $N_{\text{leps}} = 0$ samples. This control sample follows the same kinematic binning as the $N_{\text{leps}} = 1$ regions, except that the N_b distribution in each bin is integrated in N_b for $N_b \geq 1$ and each bin's N_{jets} requirement is increased by two. A diagram representing

	M_J [GeV]	N_{jets}			
		4–5	6–7	8–9	≥ 10
$N_{\text{leps}} = 0$	500 – 800		QCD CR	QCD CR	QCD CR
	800 – 1000			QCD CR	QCD CR
	> 1000		QCD CR	QCD CR	QCD CR
$N_{\text{leps}} = 1$	500 – 800	CR	CR	SR	
	800 – 1000	CR	SR	SR	
	> 1000		SR	SR	

Figure 8.1: Diagram depicting the N_{jets} , M_J binning of the analysis including the $N_{\text{leps}} = 0$ QCD control region. The QCD control region bins are denoted by “QCD CR”, while the $N_{\text{leps}} = 1$ control and signal region bins are denoted by “CR” and “SR”, respectively. The pairing structure between lepton bins is indicated by arrows for the $500 < M_J \leq 800$ region and is the same for other M_J bins.

the binning of the $N_{\text{leps}} = 0$ control sample is shown in Figure 8.1, where the black arrows indicate the pairing structure between lepton bins and is representative for all M_J bins. The QCD contribution in a particular $N_{\text{leps}} = 1$ bin is then constrained by the corresponding $N_{\text{leps}} = 0$ bin. To avoid biasing the normalization measurement, the small contribution of $t\bar{t}$ background to the $N_{\text{leps}} = 0$ control regions is simultaneously included using the normalization from the corresponding $N_{\text{leps}} = 1$ bins, while contributions from other processes are taken from simulation.

8.2.1 M_J Connection

Due to the large freedom of unconstrained normalization parameters, the fit can be sensitive to rare statistical fluctuations and return unphysical normalization values particularly in bins dominated by $t\bar{t}$ events. For example, in psuedodata experiments, where psuedo-datasets are randomly generated according to the statistical and systematic uncertainties of the pre-fit values, the fit reduced the $t\bar{t}$ contribution in the $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV bin (where statistical uncertainties are largest) to ~ 0 in about $\sim 1\%$ of the experiments. This can be seen in Figure 8.2 (left) which shows a low tail in the distribution of post-fit $t\bar{t}$ yields in the $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV bin for 1000 psuedodata experiments. This effect occurs when yields in a bin have a large fluctuation downwards, as the fit must lower the normalization of a process to compensate. The QCD, W + jets, and Other contributions, however, are largely constrained by other data control samples or taken from simulation, and so the fit uses the freedom to adjust the $t\bar{t}$ normalization in order to model the fluctuation, leading to the unphysically small values.

In order to avoid this instability, the normalizations of $t\bar{t}$ and QCD are (independently) connected by log-normal constraints between adjacent M_J bins. By correlating the normalizations across M_J bins, the fit's sensitivity to large fluctuations in a single bin is greatly reduced. The size of these connections is motivated by measurements of the data-to-simulation ratio with $N_b = 1$ events (in order to avoid potential signal contamination and unblinding the data) and is particularly chosen to be significantly larger than the uncertainty on the data-to-simulation ratios in order to avoid over-constraining the normalization parameters, while still providing some constraint against unphysical fits. Based on these measurements, shown in Figure 8.3, and criteria, a connection size between adjacent bins of [50%-200%] is chosen.

Figure 8.2 (right) shows the results of the same 1000 psuedodata experiments but

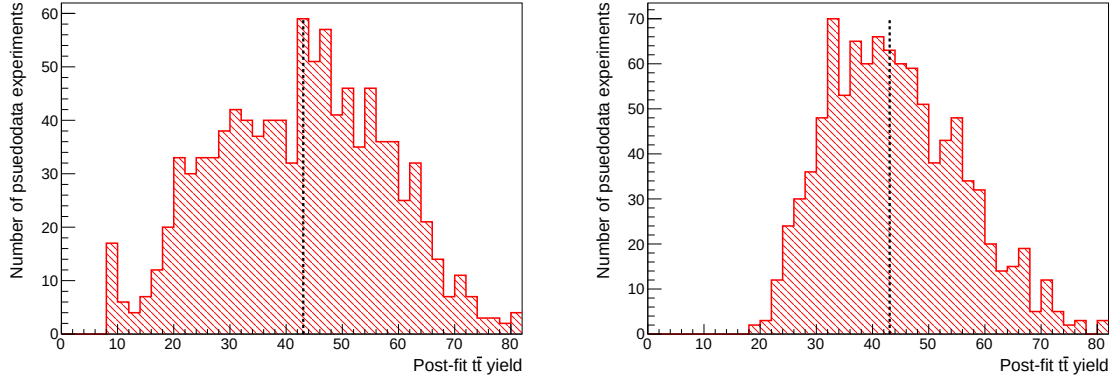


Figure 8.2: Distribution of post-fit yields of $t\bar{t}$ in the $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV bin for 1000 pseudodata experiments without (left) and with (right) constraints between adjacent M_J bins. The dotted black line indicates the pre-fit yield.

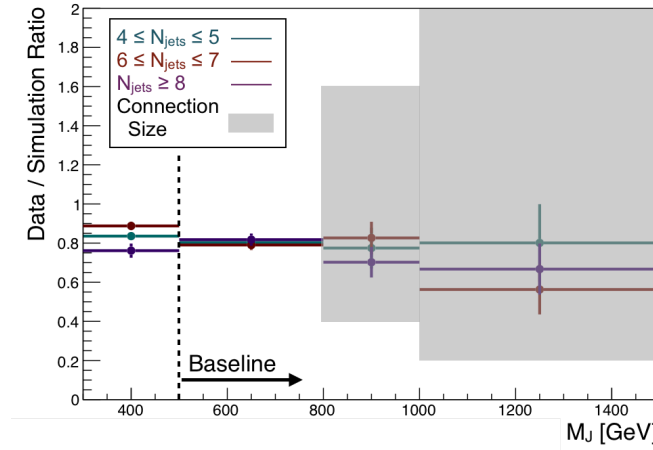


Figure 8.3: Data-to-simulation ratios as a function of M_J for different N_{jets} bins (data points) with a selection of $N_{\text{leps}} = 1$, $H_T > 1200$ GeV, and $N_b = 1$ applied. The shaded region corresponds to the size of the M_J connection in each M_J bin.

now with this constraint across M_J bins applied. The resulting distribution of post-fit $t\bar{t}$ yields now shows no evidence of unphysical normalizations and appears better behaved.

8.3 W + jets Normalization

The W + jets background is determined in the fit with one global normalization parameter and two parameters to adjust the bin-to-bin normalization of adjacent N_{jets} bins, since the N_{jets} shape may not be well-modelled by simulation. The amount the N_{jets} shape may vary is based on the data-to-simulation agreement in a kinematically similar Z + jets sample, which is selected by requiring $N_{\text{leps}} = 2$ (ee or $\mu\mu$), $H_{\text{T}} > 1200$ GeV, $M_{\text{J}} > 500$ GeV, $N_{\text{b}} = 1$, and $80 < m_{\ell\ell} < 100$ GeV, where $m_{\ell\ell}$ is the invariant mass of the two leptons. This sample is used due to the similar electroweak processes involved in W + jets and Z + jets production, which allows for the ISR component of the Z + jets sample to be used as a probe of that of W + jets production. The N_{jets} distribution and data/simulation yields ratio for this sample are shown in Figure 8.4. The resulting uncertainties are 17% between $4 \leq N_{\text{jets}} \leq 5$ and $6 \leq N_{\text{jets}} \leq 7$ and 62% between $6 \leq N_{\text{jets}} \leq 7$ and $N_{\text{jets}} \geq 8$. After correcting the N_{jets} spectrum, the residual M_{J} mismodeling is expected to be small, so no further correction is applied.

8.4 Other Normalization

The nominal normalization for Other is largely taken from simulation, as its contribution is less than 20% in every bin with typical values $\lesssim 5\%$. It is, however, allowed to vary according to its statistical and systematic uncertainties.

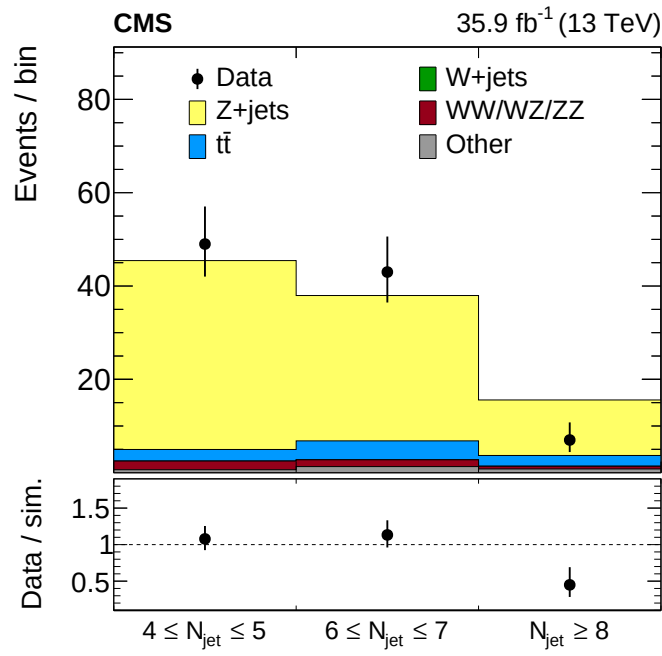


Figure 8.4: Jet multiplicity distribution for data and simulation in a Z + jets control sample selected by requiring $N_{\text{leps}} = 2$, $H_T > 1200$ GeV, $M_J > 500$ GeV, $N_b = 1$, and $80 < m_{\ell\ell} < 100$ GeV. The total yield from simulation is normalized to the number of events in data. The uncertainty in the ratio of data to simulation yields (lower panel) is statistical only.

Chapter 9

Systematic Uncertainties

The nominal simulated shape of the N_b distribution is allowed to vary by the inclusion of systematic uncertainties. Each uncertainty is incorporated in the fit with template N_b histograms to account for the effects of the systematic variation and a nuisance parameter θ to control the variation amplitude. The nuisance parameters are subject to Gaussian constraints, normalized so that $\theta = 0$ corresponds to the nominal N_b shape and $\theta = \pm 1$ corresponds to ± 1 standard deviation (s.d.) variation of the systematic uncertainty. These uncertainties affect only the N_b shape for $t\bar{t}$, QCD, and W+jets backgrounds, because their normalizations are determined from data, while for the other (subleading) backgrounds the uncertainties affect both the N_b shape and normalization.

9.1 Gluon Splitting Rate

The primary source of systematic uncertainty is on the modelling of the rate of gluon splitting (GS), as events with a gluon splitting to $b\bar{b}$ provide an additional source of b quarks in events. As this process may not be properly simulated, constraining the splitting rate in data is crucial for establishing a robust prediction of the N_b distribution. The dominant contribution of this effect is due to gluons that split specifically to b quark pairs, so the phrase “gluon splitting” will hereafter refer exclusively to gluon splitting

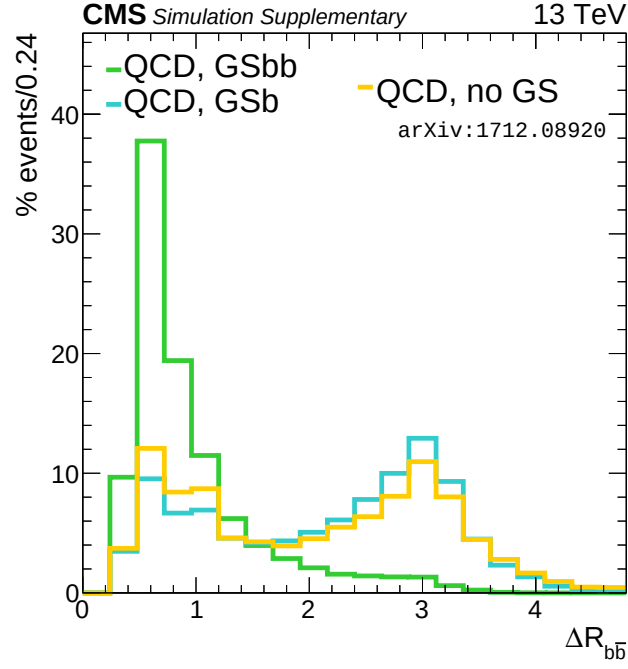


Figure 9.1: The $\Delta R_{b\bar{b}}$ distribution shapes for the three gluon splitting categories: Events with a pair of b-tagged jets resulting from gluon splitting (green), events with a gluon splitting yielding fewer than 2 b-tagged jets (blue), and events without a gluon splitting to $b\bar{b}$. These events are selected by requiring $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 4$, and $N_b = 2$.

to $b\bar{b}$. One way to select a data sample enriched in gluon splitting events is to use the $\Delta R_{b\bar{b}}$ distribution, where $\Delta R_{b\bar{b}}$ is defined as the ΔR between two b-tagged jets, as pairs of b-tagged jets resulting from the same gluon splitting tend to have smaller values of $\Delta R_{b\bar{b}}$ than pairs resulting from hard scatter b-quarks or mis-tagged jets. This can be seen in Figure 9.1, which shows the $\Delta R_{b\bar{b}}$ distribution in simulated QCD events with $N_b = 2$ for three important categories. Events that have a correlated pair of b-tagged jets originating from a gluon splitting (green, denoted GSbb) populate the low- $\Delta R_{b\bar{b}}$ region, while events without gluon splitting (yellow, denoted noGS) or where the splitting yields one or fewer b-tagged jets (blue, denoted GSb) populate the low-and high- $\Delta R_{b\bar{b}}$ regions roughly equally.

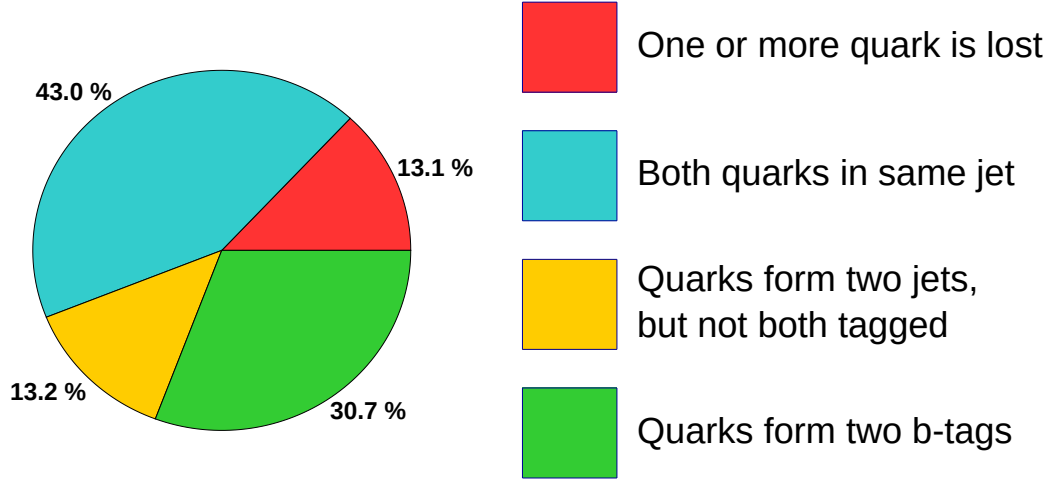


Figure 9.2: The relative fraction of the possible final states that occur from gluon splitting to $b\bar{b}$ for events satisfying $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 4$, and $N_b = 2$.

Gluon splittings can contribute less than 2 b-tagged jets either because the quarks are collimated into a single jet, one of the b-tagged jets is not tagged, or because one of the jets fails to pass the jet selection criteria, typically because it is too soft. The relative fractions of these contributions is shown in Figure 9.2.

The gluon splitting rate is then constrained by fitting the $\Delta R_{b\bar{b}}$ distributions to data by using the difference in shapes of the GSbb, GSb, and noGS categories. This fit varies the normalization of the GSbb and GSb contributions (varied together) and the noGS contributions in order to extract the relative contributions of events with and without a gluon splitting. It is performed in four equal bins in the range of $0 \leq \Delta R_{b\bar{b}} < 4.8$ with events selected by requiring $N_{\text{leps}} = 0$, $H_T > 1500$ GeV, $N_b = 2$, $N_{\text{jets}} \geq 4$, and $M_J > 500$ GeV, as the gluon splitting signal in a $N_{\text{leps}} = 1$ control sample is contaminated by b quarks from the decay of top quarks. Additionally, the $N_{\text{leps}} = 0$ control sample is formed from a subset of the data that is selected to be most stable in the b tagging algorithm performance, since the precision of the $\Delta R_{b\bar{b}}$ fit is not

	Nominal	$M_J > 800 \text{ GeV}$	$4 \leq N_{\text{jets}} \leq 5$	$6 \leq N_{\text{jets}} \leq 7$	$8 \leq N_{\text{jets}} \leq 9$	$N_{\text{jets}} \geq 10$
GS	0.77 ± 0.09	0.70 ± 0.38	0.80 ± 0.32	0.76 ± 0.14	0.75 ± 0.16	0.95 ± 0.36
No GS	1.21 ± 0.08	1.28 ± 0.35	1.15 ± 0.26	1.22 ± 0.13	1.24 ± 0.15	1.05 ± 0.36

Table 9.1: Gluon splitting weights derived in the nominal fit, a variation with a requirement of $M_J > 800 \text{ GeV}$, and 4 variations in bins of N_{jets} (with the nominal $M_J > 500 \text{ GeV}$ requirement.)

limited by the data sample size. This choice isolates the physical effects of gluon splitting from the potential time dependence of the b tagging performance due to variations in experimental conditions, which are separately incorporated by the uncertainties on the b-tagging data-to-simulation scale factors, as described in Section 9.2.

The $\Delta R_{b\bar{b}}$ fit extracts a weight of 0.77 ± 0.09 for gluon splitting events and a weight of 1.21 ± 0.08 for non-gluon splitting events. The post-fit distributions are shown in Figure 9.3. The GSbb and GSb categories are plotted separately to demonstrate the difference in shapes. The discrepancy in the last bin does not significantly impact the fit because the higher yield bins at lower values of $\Delta R_{b\bar{b}}$ constrain the fit. The deviations of these weights from unity, summed in quadrature with their post-fit uncertainty, are used to form the ± 1 s.d. variations of the gluon splitting rate nuisance parameter by applying weights of 1 ± 0.25 to gluon splitting events and 1 ∓ 0.22 to non-gluon splitting events in an anti-correlated manner. The fit results are used as a measure of the uncertainty on modelling of the GS rate as opposed to a correction to the central value, since the $\Delta R_{b\bar{b}}$ variable may not be a perfect proxy for the GS rate. Figure 9.4 shows the effect of the ± 1 s.d. variations on the N_b distribution of $t\bar{t}$ for the two most sensitive bins.

In order to test the stability of the fit results and the dependence of the gluon splitting weights across kinematic regions, the $\Delta R_{b\bar{b}}$ fit is repeated both with a higher M_J threshold and with different N_{jets} bins. The resulting weights are shown in Table 9.1 and are all consistent with those of the nominal fit.

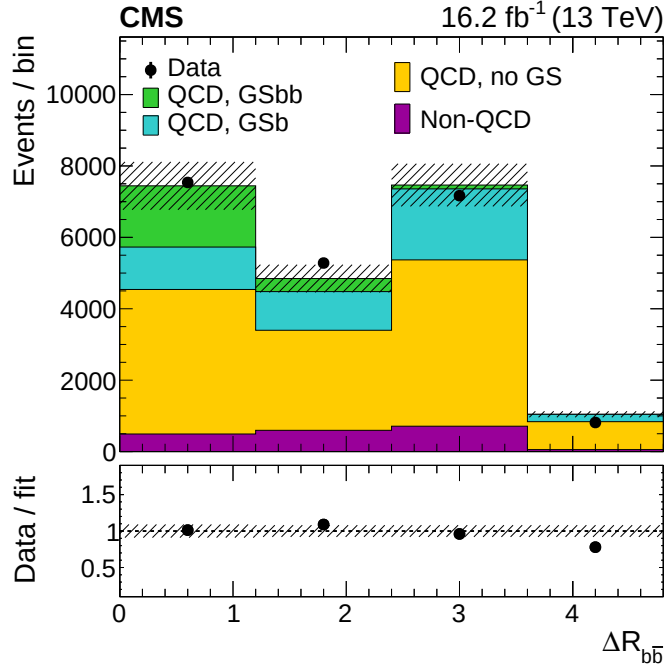


Figure 9.3: Post-fit $\Delta R_{b\bar{b}}$ distributions in a selection with $N_{\text{lept}} = 0$, $H_T > 1500$ GeV, $M_J > 500$ GeV, $N_{\text{jets}} \geq 4$, and $N_b = 2$ with the post-fit uncertainty represented by a hatched band. The ratio of data to simulation yields is shown in the lower panel.

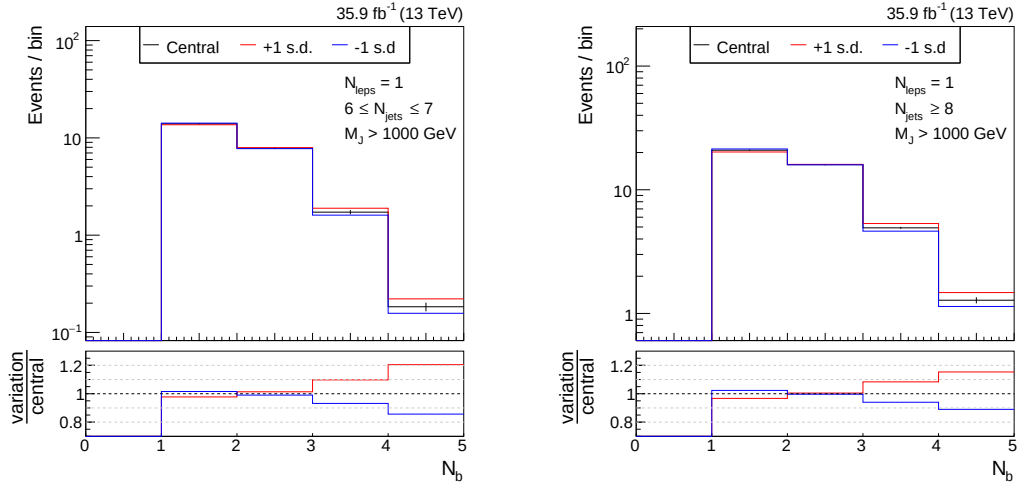


Figure 9.4: Effect of the ± 1 s.d. variations of the gluon splitting rate on the N_b distribution in $t\bar{t}$ events for the two most sensitive bins: $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right). Event yields are normalized to that expected in 35.9 fb^{-1} of data.

9.2 b-tagging Data-to-simulation Scale Factors

Another significant systematic uncertainty is the uncertainty in the data-to-simulation scale factors (SF) for b tagging efficiency and mistag rates. Simulating the b-tagging algorithm relies on understanding the detailed behavior of the detector and also accurate modelling of the parton shower and hadronization, both of which are non-trivial. Therefore, it is important to measure the b tagging efficiencies and mistag rates in data and correct the simulation to match.

The difference between data and simulation is corrected for by using a per jet data-to-simulation SF

$$SF_f = \varepsilon_f^{data}(p_T) / \varepsilon_f^{sim}(p_T). \quad (9.1)$$

where $\varepsilon_f^{data}(p_T)$ and $\varepsilon_f^{sim}(p_T)$ are the tagging efficiencies for a jet with flavor f as a function of p_T in data and simulation, respectively. No dependence on η is derived due to limited data sample sizes. In simulation, the efficiency is determined by matching jets to their generated hadron to determine their flavor and then measuring how many of those jets are correctly tagged. In data, this is done by using control regions determined by specific selection requirements that produce pure samples of a certain flavor of jets while not biasing the jets with respect to variables used in the b tagging algorithm.

The probability to tag a light-flavor or gluon jet (light jet) is measured in an inclusive QCD sample. This sample is selected through a series of triggers that require at least one jet over a certain p_T threshold, the lowest being 40 GeV.

The probability to tag a charm-flavor jet (c jet) is determined by measurements in two charm-enriched control regions. The first control region is formed by selecting events in which a charm quark is produced in association with a W boson. The main contributions

to this process is from $s+g \rightarrow W^-+c$ and $\bar{s}+g \rightarrow W^++\bar{c}$, where a key property is that the W boson and quark have oppositely signed electrical charges. The dominant background is $W + q\bar{q}$ events, which produce an equal amount of events with same- and opposite-signed W boson and quark pairs. Thus, this background is removed by measuring its contribution in the same-sign channel and then subtracting it from the opposite-sign channel, resulting in a pure $W + c$ channel. The second charm-enriched control region is created by selecting single-lepton $t\bar{t}$ events. As hadronically-decaying W bosons decay to a charm quark about 50% of the time, about one of two single-lepton $t\bar{t}$ events will contain a charm quark. Finally, measurements in these two regions are combined using the best linear unbiased estimator (BLUE) method described in Reference [109].

The probability to tag a b-flavor jet (b jet) is computed using QCD and $t\bar{t}$ control regions. The QCD control regions are enriched in b quarks by requiring that at least one jet contains a muon with $p_T > 5$ GeV, which takes advantage of the high branching fraction to leptons of b hadrons. In the $t\bar{t}$ -dominated regions, there are two b quarks per event, due to the decay of the two top quarks, and the b quark purity is further enhanced by limiting the number of non-b jets in the event through the requirement that either one or both of the W bosons decays leptonically. This creates independent single-lepton and di-lepton control regions where multiple measurements are made and then combined through the BLUE method.

The resulting SFs, including their uncertainties, are shown in Figure 9.5. A complete discussion of how the SF measurements are made can be found in Reference [19].

The systematic uncertainties on the N_b shape are assessed by the ± 1 s.d. N_b templates resulting from varying the SFs according to their uncertainties. Because the b and c jet SFs have correlated uncertainties, they are conservatively varied together and form one set of templates. The light-flavor SFs are uncorrelated with the b and c jet SFs and are varied independently. The effects of these variations on the N_b distribution in $t\bar{t}$ events

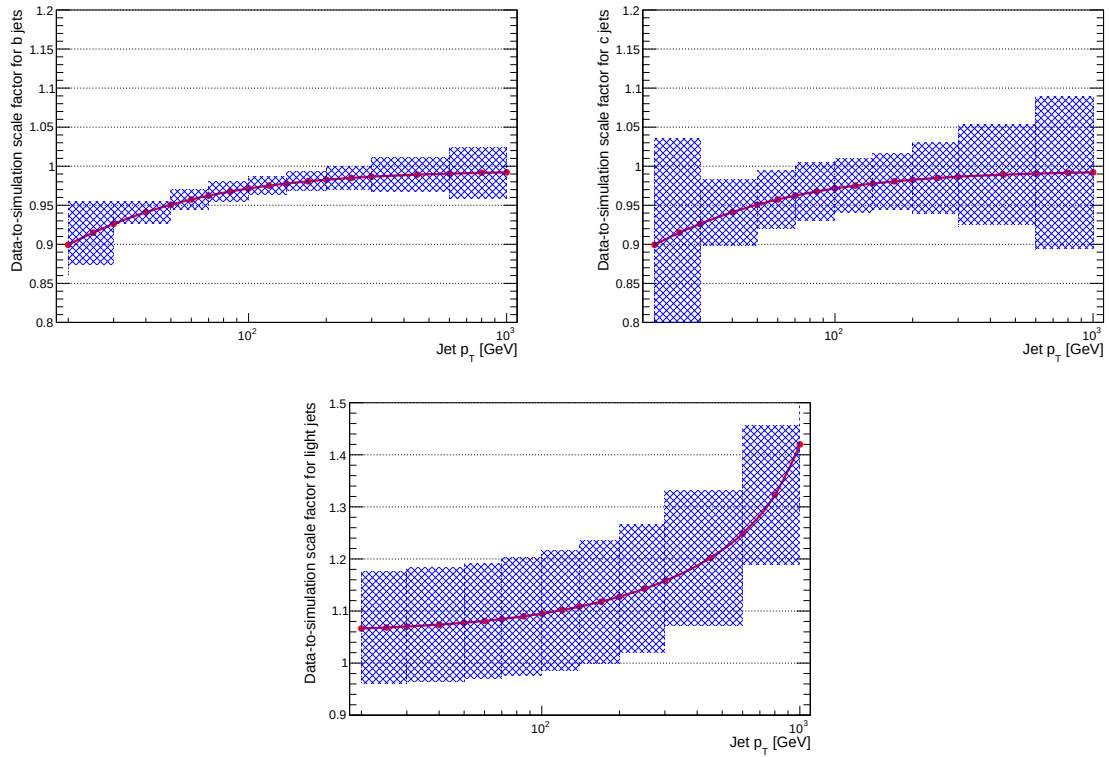


Figure 9.5: The data-to-simulation scale factors for the tagging efficiency of b-flavor jets (top-left), charm-flavor jets (top-right), and light-flavor or gluon jets (bottom) are shown as a function of jet p_T . The associated uncertainty with each scale factor is shown as a blue hashed band.

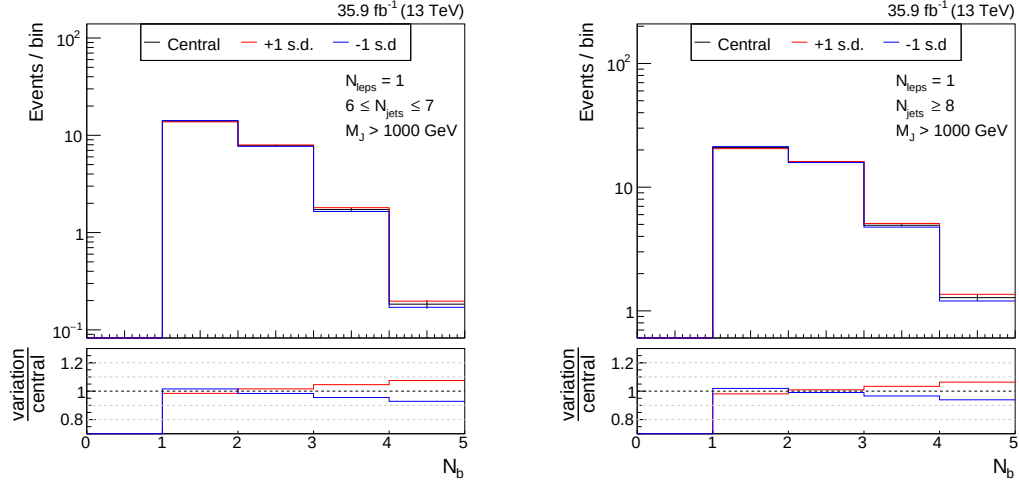


Figure 9.6: Effect of the ± 1 s.d. correlated variations of the b-flavor and c-flavor jet data-to-simulation scale factors on the N_b distribution in $t\bar{t}$ for the two most sensitive bins: $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right). Event yields are normalized to that expected in 35.9 fb $^{-1}$ of data.

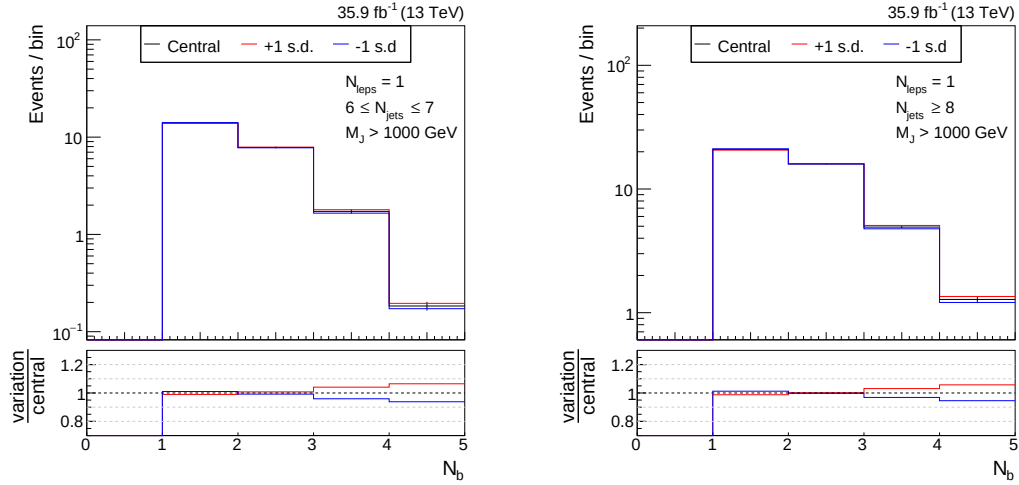


Figure 9.7: Effect of the ± 1 s.d. variations of the light-flavor jet data-to-simulation scale factors on the N_b distribution in $t\bar{t}$ for the two most sensitive bins: $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right). Event yields are normalized to that expected in 35.9 fb $^{-1}$ of data.

are shown in Figure 9.6 and Figure 9.7.

9.3 Lepton Fake Rate in QCD

While the QCD normalization is measured from data, it is mostly constrained by the $N_{\text{leps}} = 0$ selection and applied in a $N_{\text{leps}} = 1$ region. If the simulated N_{leps} distribution is not modelled perfectly, there may be residual differences between the normalizations of these two regions. For processes that have true prompt leptons, such as $t\bar{t}$ and $W + \text{jets}$, the N_{leps} distribution is well modelled, because the dominant effects are the W branching fractions and the acceptance (including selection efficiency), both of which are well understood. For QCD, however, this is less well modelled as the simulation of the tail of the jet fragmentation function, as well as detector effects that can produce fake leptons, are not as well understood.

To assign a systematic uncertainty on the modelling of the ratio of 0-lepton to 1-lepton events in QCD, the lepton isolation distributions are studied. Figure 9.8 shows the relative isolation distributions for electrons (left) and muons (right) in a data sample corresponding to the analysis control regions. The binning of the histograms are chosen such that the first bin corresponds to the relative isolation requirement for signal leptons (0.1 for electrons and 0.2 for muons). The normalizations of the QCD, $t\bar{t}$, and $W + \text{jets}$ processes are scaled to match the results of a control region fit described in Section 10.2.2.

Table 9.2 shows the ratio of $I^{\text{rel}} < 0.1(0.2)$ to $I^{\text{rel}} \geq 0.1(0.2)$ for electrons(muons) in QCD and data with contributions for all other processes ($t\bar{t}$, $W + \text{jets}$, Other) subtracted. As the ratio in data agrees to that in QCD simulation within 20%, an additional 20% log-normal uncertainty is assigned to the QCD normalization, which is fully correlated across the 1-lepton bins.

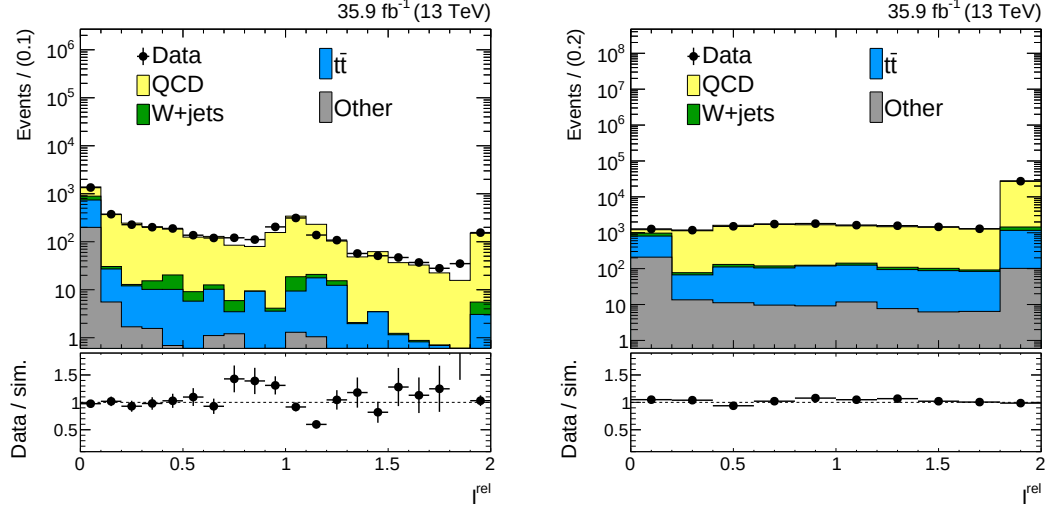


Figure 9.8: The relative isolation distribution for electrons (left) and muons (right) in the analysis control region. The binning of the histograms is chosen such that the first bin corresponds to the relative isolation requirement for signal leptons (0.1 for electrons and 0.2 for muons). The normalizations of the QCD, $t\bar{t}$, and $W + \text{jets}$ processes are scaled to match the results of a control region fit described in Section 10.2.2.

Process	Electrons			Muons		
	$I^{\text{rel}} < 0.1$	$I^{\text{rel}} \geq 0.1$	ratio	$I^{\text{rel}} < 0.2$	$I^{\text{rel}} \geq 0.2$	ratio
QCD	496.8	2455.5	0.20	219.8	36553.3	0.0060
Data - all other	452.8	2500.5	0.18	275.4	37497.7	0.0073

Table 9.2: Comparison of the relative isolation distributions, as described in the caption of Figure 9.8, for electrons and muons between QCD and data with contributions from “all other” ($t\bar{t}$, $W + \text{jets}$, and Other) subtracted.

9.4 Additional Systematic Uncertainties

Other experimental uncertainties are small and include lepton selection efficiency, jet energy scale, jet energy resolution, and integrated luminosity. The uncertainty associated with lepton selection efficiency is determined by varying the efficiency to select a lepton within its uncertainty determined from data. Jet energy scale uncertainties [110, 111] are assessed by varying the p_T of small- R jets as a function of p_T and η . The uncertainty arising from jet energy resolution [110, 111] is determined by applying an $|\eta|$ -dependent

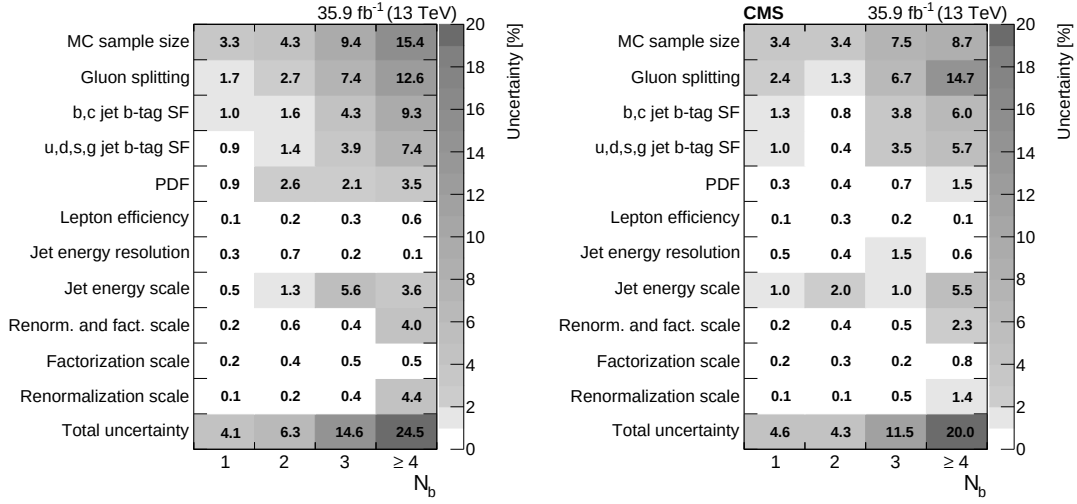


Figure 9.9: Background systematic uncertainties (in percent) for the $N_{\text{jets}} \geq 8$, $500 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right) bins. The bottom row shows the total uncertainty for a given N_b bin by summing in quadrature all uncertainties. These values are similar for other bins.

factor to the jet p_T to match the jet energy resolution observed in data. The integrated luminosity is varied according to its uncertainty of 2.5% [112], affecting only the backgrounds estimated from simulation. No uncertainty is applied for the amount of pileup as studies have shown its effect to be negligible in this high- H_T selection. The uncertainties due to the limited size of simulation samples are incorporated as uncorrelated nuisance parameters in the fit.

Theoretical systematic uncertainties are applied and include independent and correlated variations of the renormalization and factorization scales. Additionally, uncertainties on the PDF are incorporated by considering variations in the NNPDF 3.0 scheme [96]. The size of these uncertainties is typically small as the effect of these variations is largely to modify the cross section of processes, which for the main backgrounds are constrained by data.

The background systematic uncertainties that affect the N_b shape are shown in Figure 9.9 for the two most sensitive search bins.

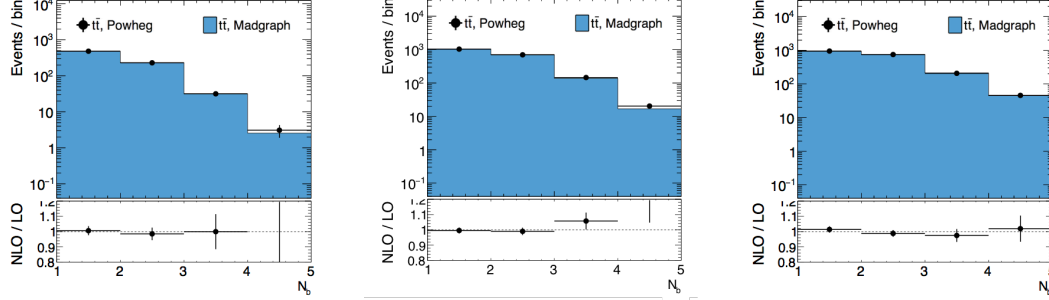


Figure 9.10: Comparison of the N_b distribution for a sample generated at LO precision with MADGRAPH5_aMC@NLO 2.2.2 (blue histogram) with that of one generated at NLO precision with POWHEG 2.0 (data points). The comparison is done after the baseline selection and in bins of $4 \leq N_{\text{jets}} \leq 5$ (left), $6 \leq N_{\text{jets}} \leq 7$ (middle), and $N_{\text{jets}} \geq 8$ (right). In order to evaluate only shape differences, the MADGRAPH5_aMC@NLO sample is normalized to match the normalization of the POWHEG sample.

9.4.1 N_b Distribution With NLO Precision

As the nominal N_b distribution shape is generated at LO precision, it is important to verify that the size of NLO effects is small and covered by the experimental uncertainties. To do this, the N_b distributions of $t\bar{t}$ samples generated at LO precision with the default MADGRAPH5_aMC@NLO 2.2.2 generator and samples generated at NLO precision with POWHEG 2.0 are compared after the baseline selection. In order to evaluate only the shape differences between the two samples, the normalization of the MADGRAPH5_aMC@NLO sample is scaled to match that of the POWHEG sample. The comparison of the two distributions, shown in Figure 9.10, is conducted in bins of N_{jets} and indicates a disagreement of only about 5%. This disagreement is small and subdominant to the experimental uncertainties previously discussed and no additional uncertainty is applied.

Given the level of agreement between the two samples, the MADGRAPH5_aMC@NLO

$N_{PV}^{true} \leq 20$	$20 < N_{PV}^{true} \leq 40$	$N_{PV}^{true} > 40$
$8.0 \pm 0.5\%$	$8.1 \pm 0.4\%$	$7.5 \pm 1.5\%$

Table 9.3: The signal efficiency of the most sensitive bin $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV for a 1600 GeV gluino in various bins of the number of truth-level primary vertices.

sample is used in this analysis as the sample size is significantly larger than that for the POWHEG sample.

9.5 Signal Systematics

Several of the systematic uncertainties affecting the signal yield are evaluated in the same way as the background yield. These are the uncertainties due to gluon splitting, lepton selection efficiency, jet energy scale, jet energy resolution, b tagging scale factors, simulation sample size, integrated luminosity, and theoretical uncertainties. All systematic variations affect both the N_b shape and normalization, except for the gluon splitting uncertainty, which is taken to affect only the N_b shape.

The number of jets from ISR produced in the signal simulation is reweighted based on comparisons between data and simulated $t\bar{t}$ samples. The reweighting factors vary between 0.92 and 0.51 for the number of ISR jets between 1 and ≥ 6 . One half of the deviation from unity is taken as the systematic uncertainty in these reweighting factors.

The systematic uncertainties affecting the signal N_b shape are shown in Figure 9.11 for the most sensitive bins in a model with $m_{\tilde{g}} = 1600$ GeV. The dominant signal systematic uncertainties arise from the limited simulation sample size, the b tagging efficiency scale factors, and the ISR modeling. There is no systematic uncertainty taken for pileup reweighting, as the signal efficiency is found to be insensitive to the number of pileup interactions, which is shown in Table 9.3.

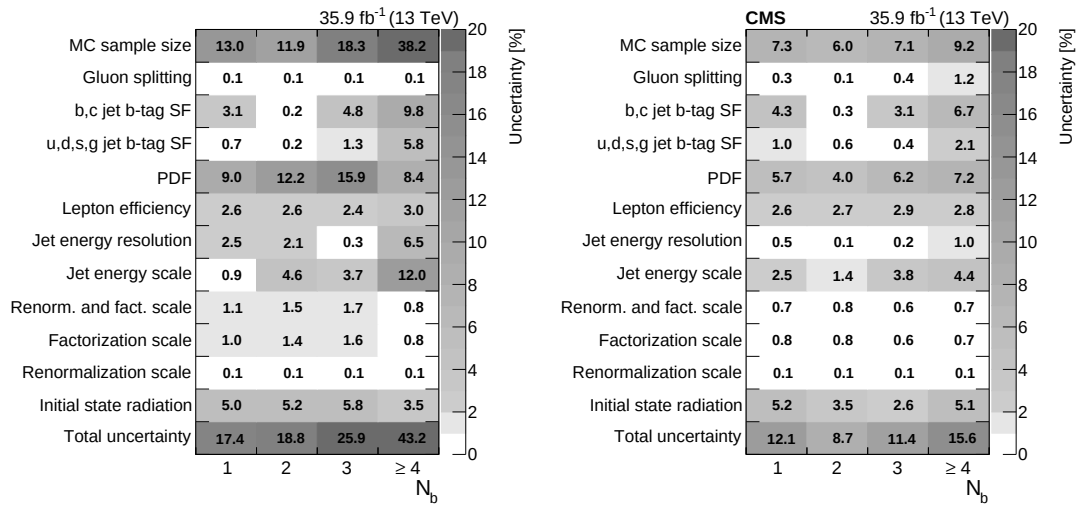


Figure 9.11: Systematic uncertainties (in percent) for a $m_{\tilde{g}} = 1600$ GeV signal and for the $N_{\text{jets}} \geq 8$, $500 < M_J \leq 1000$ GeV (left) and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (right) bins. The bottom row shows the total uncertainty for a given N_b bin by summing in quadrature all uncertainties. These values are similar for other bins.

Chapter 10

Fit Model and Validation

10.1 Description of the Fit Model

In order to determine the values of the background normalizations and nuisance parameters, as well as extract any signal that may be present, a maximum-likelihood fit to the observed data is performed. The full likelihood model can be written as the multiplication of a likelihood constructed from only the $N_{\text{leps}} = 1$ bins and a likelihood constructed from only the $N_{\text{leps}} = 0$ bins:

$$\mathcal{L} = \mathcal{L}_{N_{\text{leps}}=1} \mathcal{L}_{N_{\text{leps}}=0} \quad (10.1)$$

where

$$\begin{aligned} \mathcal{L}_{N_{\text{leps}}=1} = & \prod_{\substack{i \in N_{\text{jets}} \\ j \in M_J \\ k \in N_b}} P(N_{ijk} | \mu^{\text{signal}} \nu_{ijk}^{\text{signal}} + \mu_i^{\text{t}\bar{\text{t}}} \nu_{ijk}^{\text{t}\bar{\text{t}}} + \mu_i^{\text{QCD}} \nu_{ijk}^{\text{QCD}} + \mu^{\text{W+jets}} \nu_{ijk}^{\text{W+jets}} + \nu_{ijk}^{\text{Other}}) \\ & \prod_{l \in \text{t}\bar{\text{t}} \text{ norm.}} P(N_{ijk} | \theta_l) \prod_{l' \in \text{QCD norm.}} P(N_{ijk} | \theta_{l'}) \prod_{l'' \in \text{W+jets norm.}} P(N_{ijk} | \theta_{l''}) \prod_{m \in \text{shape}} P(N_{ijk} | \theta_m), \end{aligned} \quad (10.2)$$

and

$$\begin{aligned} \mathcal{L}_{N_{\text{leps}}=0} = & \prod_{\substack{i \in N_{\text{jets}} \\ j \in M_J \\ i' = i+1}} P(N_{ij} | \mu^{\text{signal}} \nu_{ij}^{\text{signal}} + \mu_i^{\text{t}\bar{\text{t}}} \nu_{ij}^{\text{t}\bar{\text{t}}} + \mu_{i'}^{\text{QCD}} \nu_{ij}^{\text{QCD}} + \nu_{ij}^{\text{W+jets}} + \nu_{ij}^{\text{Other}}) \\ & \prod_{l \in \text{t}\bar{\text{t}} \text{ norm.}} P(N_{ij} | \theta_l) \prod_{l' \in \text{QCD norm.}} P(N_{i'j} | \theta_{l'}) \prod_{m \in \text{shape}} P(N_{ij} | \theta_m). \end{aligned} \quad (10.3)$$

Here, the indices i , j , and k respectively run over the appropriate N_{jets} , M_J , and N_b bins of the analysis. The value of N_{ijk} is the data yield in the corresponding bin, the μ factors correspond to unconstrained normalization parameters, the ν factors are the simulated yields, and the θ factors represent nuisance parameters. The sets of nuisance parameters “ $\text{t}\bar{\text{t}}$ norm.” and “QCD norm.” correspond to the M_J -connection normalization parameters discussed in Subsection 8.2.1, while the nuisance parameter set “W + jets norm.” corresponds to the normalization parameters described in Section 8.3. The “shape” nuisance parameter set is comprised of those described in Chapter 9. For conciseness, the dependencies of the ν terms and elements of the θ terms have been dropped but are kept for the discussion below.

In more detail, for the $N_{\text{leps}} = 1$ likelihood, the $\mu^{\text{signal}} \nu_{ijk}^{\text{signal}}(\theta_m)$ term represents the signal yield in the corresponding bin, as a function of θ_m , multiplied by a global signal strength parameter. The $\mu_i^{\text{t}\bar{\text{t}}} \nu_{ijk}^{\text{t}\bar{\text{t}}}(\theta_m, \theta_{l,ij})$ term encodes the floating normalizations in each N_{jets} bin through $\mu_i^{\text{t}\bar{\text{t}}}$, while the effects of shape systematics and the M_J -connection normalizations are represented by the dependence of $\nu_{ijk}^{\text{t}\bar{\text{t}}}$ on θ_m and $\theta_{l,ij}$, respectively. The QCD term behaves analogously to the $\text{t}\bar{\text{t}}$ term, except that it has separate M_J -connection normalizations and, thus, is dependent on $\theta_{l',ij}$ instead of $\theta_{l,ij}$. The $\mu^{\text{W+jets}} \nu_{ijk}^{\text{W+jets}}(\theta_m, \theta_{l'',i})$ term controls the W + jets yield in the corresponding bin, which is dependent on an overall unconstrained normalization parameter, $\mu^{\text{W+jets}}$, on shape systematics, θ_m , and on N_{jets} -dependent normalizations, $\theta_{l'',i}$. Finally, the term $\nu_{ijk}^{\text{Other}}(\theta_m)$ does not have a

corresponding μ factor as the normalization of Other is taken from simulation. The dependence on θ_m , however, does allow the normalization to vary according to systematic uncertainties.

The $N_{\text{leps}} = 0$ likelihood largely follows the same structure as the $N_{\text{leps}} = 1$ likelihood, though with some important differences and features. First, the data yields, N_{ij} , and simulation yields, ν_{ij} , do not depend on the index k , reflecting the fact that the $N_{\text{leps}} = 0$ regions are integrated in N_b for $N_b \geq 1$. Second, the likelihood depends not on index i but on i' , where $i' = i + 1$. This represents the correspondence of N_{jets} bins between the 0-lepton and 1-lepton regions, where a higher N_{jets} bin in the $N_{\text{leps}} = 0$ sample is used to constrain an N_{jets} bin in the $N_{\text{leps}} = 1$ sample. Third, the $\mu^{\text{t}\bar{\text{t}}}$ and μ^{QCD} terms appear in this likelihood as well, indicating the simultaneous fit of their contributions across the two N_{leps} regions, while the $\mu^{\text{W+jets}}$ term is not present, as the W + jets normalization is only determined in the $N_{\text{leps}} = 1$ region. The W + jets normalization is not simultaneously fit in both N_{leps} regions, because it is negligibly small in the 0-lepton region. This fact also explains why a product analogous to $\prod_{l'' \in \text{W+jets norm.}} P(N_{ijk}|\theta_{l''})$ is not present. Lastly, the product $\prod_{m \in \text{shape}} P(N_{ij}|\theta_m)$ only affects the term ν_{ij}^{Other} , as for all other processes the shape variations have no effect due to the integration in N_b .

10.2 Validation of the Fit Model

Due to the fit's complexity and many adjustable parameters, it is important to verify the fit model and that it behaves as intended. Signal injection tests are simulation-based studies that allow for the assumptions of the model to be tested and, importantly, for the behavior of the fit in the presence of signal to be evaluated. A control region fit is a data-driven validation test, which is needed to ensure that the method is not overly-reliant on

simulation and is able to model the actual experimental conditions. Both of these tests were performed to validate the fit model and are described in fuller detail below.

10.2.1 Signal Injection Study

Signal injection studies are a useful way to quantify how well the maximum-likelihood fit can extract a potential signal if it is indeed present. These studies rely on the use of psuedodata experiments. A single experiment consists of generating psuedodata by fluctuating bin yields around their pre-fit values according to their statistical and systematic uncertainties. This pseudodata can then be treated as observations and can be fit with the results examined. As many psuedodata experiments are generated, the collection of observations approximates the distribution of possible observations as defined by the fit model and as a result the distribution of post-fit results approximates the distribution of possible post-fit results.

For the signal injection study, 1000 experiments are generated by fluctuating both the expected background and signal yields (with signal strength = 1) for each gluino mass point. Figure 10.1 shows the median fitted signal strength of the 1000 experiments for each gluino mass point. For gluino masses between 1000 – 1700 GeV the fit shows no evidence of a bias and has a median extracted signal strength of ~ 1 , while for higher gluino masses, the fit tends to under-extract the signal contribution (up to $\sim 25\%$ for $m_{\tilde{g}} = 2000$ GeV). These biased mass points correspond to models where the number of signal events is very low. For example, there are only 8.6 events expected for the $m_{\tilde{g}} = 2000$ GeV model, summing over all analysis bins. This low yield means that gaussian-approximations of the poisson-distributed bin yields used in the fit model are no longer appropriate, leading to the bias in the fit.

In order to test this hypothesis, additional signal injection studies, each consisting

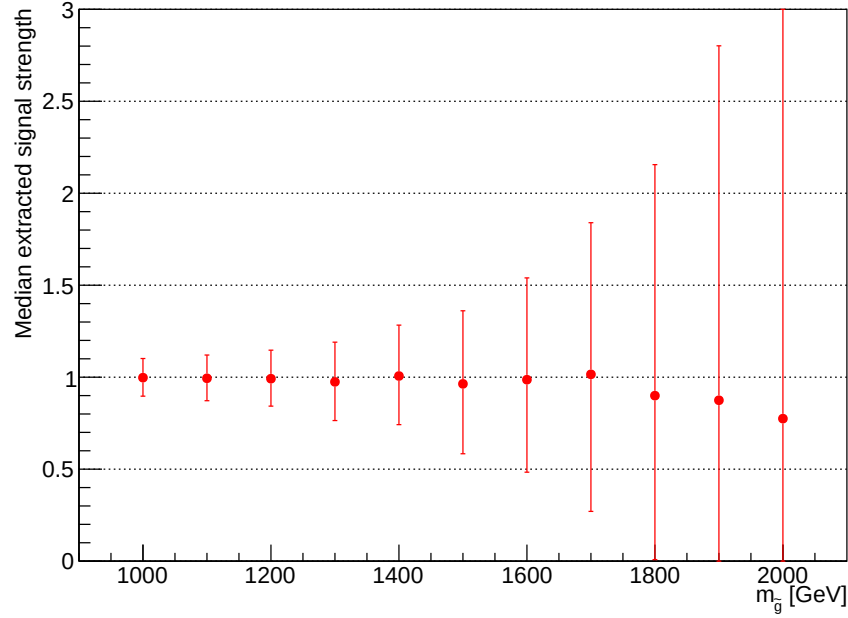


Figure 10.1: Median extracted signal strength of 1000 psuedodata experiments as a function of gluino mass. The uncertainties drawn are the median upper and lower errors of the fitted signal strengths per mass point.

of 1000 experiments, are performed for the $m_{\tilde{g}} = 2000$ GeV mass point, where the injected signal strengths are 1x, 3x, 5x, 10x the nominal cross section. The resulting median extracted signal is 78%, 92%, 95%, and 98% of the injected signal, respectively. These results support the hypothesis, as, with increasing signal strength, the gaussian-approximations become increasingly accurate, allowing for the fit to properly extract the signal contributions. The distributions of the fractional fitted signal strength for these tests are shown in Figure 10.2.

No modifications to the fit model are made to correct for this issue. This is because the fit bias only affects mass points that are far above the highest mass (1650 GeV) expected to be excluded by this analysis, and Figure 10.1 shows that the fit bias is much smaller than the precision of the fit for those mass points. Additionally, the coverage of the 95% confidence intervals of the fit is tested using the signal injection experiments

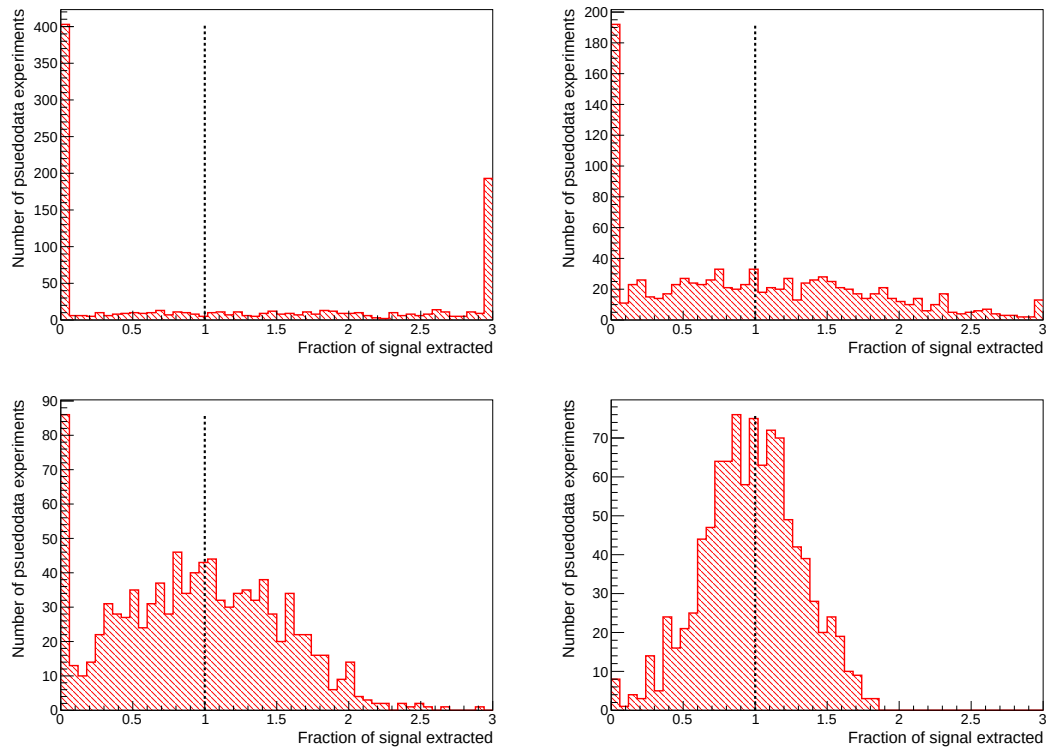


Figure 10.2: Distribution of the fraction of signal extracted from 1000 pseudodata experiments for a 2000 GeV gluino at 1x (top-left), 3x (top-right), 5x (bottom-left), and 10x (bottom-right) the nominal cross section. The last bin includes the overflow contents, and the black dashed line represents extracting as much signal as was injected. The median extracted signal is 78%, 92%, 95%, and 98% the injected signal, respectively.

$m_{\tilde{g}} = 1800 \text{ GeV}$	$m_{\tilde{g}} = 1900 \text{ GeV}$	$m_{\tilde{g}} = 2000 \text{ GeV}$
96%	95%	96%

Table 10.1: Actual coverage probability of the 95% confidence interval of the fit for the mass points with a biased signal extraction.

and found to be either correct or slightly conservative, as shown in Table 10.1.

10.2.2 Control Region Fit

While the signal injection studies are a useful validation of the fit model, it is important to validate the model using data in order to test for unmodeled effects. This is done by performing the maximum-likelihood fit with only the low- N_{jets} , low- M_J control regions, as defined in Table 7.4. These bins are chosen due to their low-expected signal yields, which avoids signal contamination effects and unblinding the high-expected signal regions in the case further modifications of the fit are required.

The control region fit, under the background-only hypothesis, is able to model the observed data well without needing large adjustments to the nuisance parameters, as seen in the post-fit N_b distributions shown in Figure 10.3. The change between the pre- and post-fit normalizations of the background processes is shown in Table 10.2, while the pulls of the nuisance parameters corresponding to the systematic uncertainties (largely controlling the shape of the N_b distribution) are shown in Figure 10.4. Both sets of values are well-behaved, as the largest change in normalization is less than 50% with typical values around 10-15%, while the nuisance parameters are all consistent with their pre-fit uncertainties, with most shifted less than 0.05 s.d. The largest pulls correspond to nuisance parameters controlling the gluon splitting rate (gs, +0.42 s.d.), the light-flavor b-tag SFs (btag_udsg, +0.37 s.d.), and the heavy-flavor b-tag SFs (btag_bc, +0.13 s.d.). These nuisances are expected to be shifted up as the observed data is higher than simulation in the tail of the pre-fit N_b distributions, as seen in Figure 10.5, where

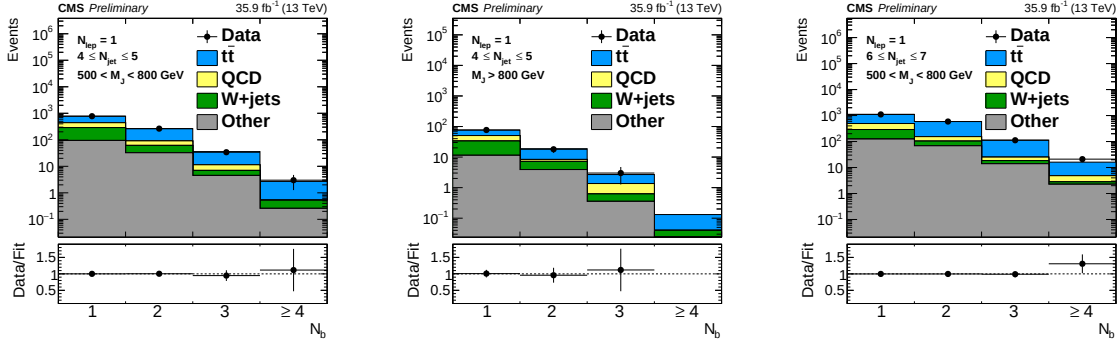


Figure 10.3: Post-fit N_b distributions of the control region fit with only statistical uncertainties shown.

Process	Pre-fit Yield	Post-fit Yield (b-only)	% change
$4 \leq N_{\text{jets}} \leq 5, 500 \leq M_J \leq 800$			
$t\bar{t}$	501.4	533.3 ± 80.7	+6.3
QCD	218.8	186.7 ± 36.8	-14.7
W + jets	400.4	225.5 ± 100.0	-43.7
Other	141.4	131.8 ± 34.5	-6.8
$4 \leq N_{\text{jets}} \leq 5, M_J \geq 800$			
$t\bar{t}$	36.9	37.5 ± 13.3	+1.6
QCD	23.1	18.9 ± 4.5	-18.2
W + jets	45.7	25.7 ± 11.4	-43.8
Other	16.8	15.7 ± 3.8	-6.5
$6 \leq N_{\text{jets}} \leq 7, 500 \leq M_J \leq 800$			
$t\bar{t}$	1370.4	1148.3 ± 78.0	-16.2
QCD	293.9	262.6 ± 52.1	-10.6
W + jets	367.7	205.6 ± 92.2	-44.1
Other	225.2	209.7 ± 58.5	-6.9

Table 10.2: Table comparing the post-fit normalizations of the control region fit to the pre-fit yields for the various background processes.

the background simulation is normalized to match the observed data yields with a single scaling factor and with the pre-fit uncertainty represented by a hatched band.

Lastly, Table 10.3 compares the post-fit pulls of the background-only and signal-plus-background control region fit. The post-fit pulls between the two fits are fully consistent with each other, as is expected for these signal-poor regions.

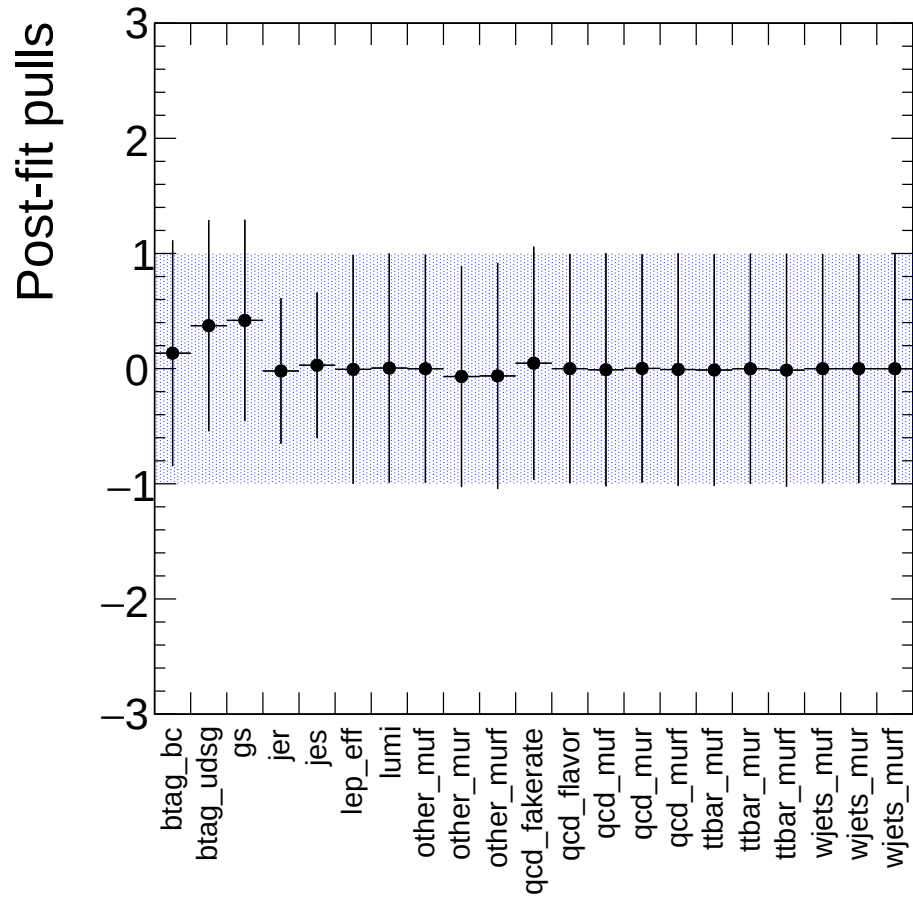


Figure 10.4: Post-fit pulls of the background-only control region fit. The post-fit values of the nuisance parameters are indicated by data points, while the post-fit uncertainty is shown as a black line and is normalized by the pre-fit uncertainty depicted as the blue band.

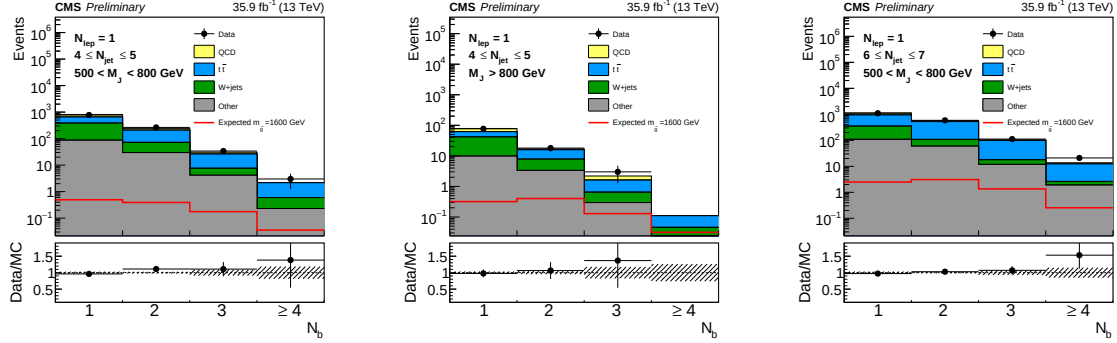


Figure 10.5: Pre-fit N_b distributions of the control region bins with the background simulation normalized to the observed data yields with a single scaling factor. The ratio of data-to-simulation is shown in the lower panel. The pre-fit uncertainty is represented by the hatched band.

Nuisance parameter	Post-fit pull (b -only fit)	Post-fit pull ($s + b$ fit)	$\rho(\theta_m, \mu)$
b,c jet b-tag SF (btag_bc)	$+0.13 \pm 0.98$	$+0.07 \pm 1.05$	-0.18
u,d,s,g jet b-tag SF (btag_udsg)	$+0.37 \pm 0.92$	$+0.28 \pm 0.95$	-0.26
Gluon splitting (gs)	$+0.42 \pm 0.87$	$+0.22 \pm 1.12$	-0.43
Jet energy resolution (jer)	-0.02 ± 0.63	-0.02 ± 0.60	-0.01
Jet energy scale (jes)	$+0.03 \pm 0.63$	$+0.03 \pm 0.61$	-0.03
Lepton efficiency (lep_eff)	-0.01 ± 0.99	-0.01 ± 0.99	+0.01
Luminosity (lumi)	$+0.00 \pm 0.99$	$+0.00 \pm 0.99$	-0.01
Fact. scale for Other (other_muf)	-0.00 ± 0.99	-0.00 ± 0.99	+0.00
Renorm. scale for Other (other_mur)	-0.07 ± 0.96	-0.06 ± 1.02	+0.02
Renorm. and Fact. scale for Other (other_murf)	-0.06 ± 0.98	-0.08 ± 0.96	+0.01
QCD fake rate (qcd_fakrate)	$+0.05 \pm 1.01$	$+0.09 \pm 1.14$	+0.09
Fact. scale for QCD (qcd_muf)	-0.01 ± 1.01	-0.01 ± 1.01	-0.00
Renorm. scale for QCD (qcd_mur)	$+0.00 \pm 0.99$	$+0.00 \pm 0.99$	-0.00
Renorm. and Fact. scale for QCD (qcd_murf)	-0.01 ± 1.01	-0.01 ± 1.01	-0.00
Fact. scale for $t\bar{t}$ (ttbar_muf)	-0.01 ± 1.01	-0.01 ± 1.00	+0.00
Renorm. scale for $t\bar{t}$ (ttbar_mur)	-0.00 ± 1.00	$+0.00 \pm 0.99$	+0.01
Renorm. and Fact. scale for $t\bar{t}$ (ttbar_murf)	-0.01 ± 1.01	-0.01 ± 1.00	+0.01
Fact. scale for W + jets (wjets_muf)	-0.00 ± 0.99	$+0.00 \pm 0.99$	+0.00
Renorm. scale for W + jets (wjets_mur)	-0.00 ± 0.99	-0.00 ± 0.99	-0.00
Renorm. and Fact. scale for W + jets (wjets_murf)	-0.00 ± 1.00	-0.00 ± 1.00	+0.00

Table 10.3: Table of post-fit pulls of the background-only and signal-plus-background control region fit. The last column, $\rho(\theta_m, \mu)$, lists the correlation between the corresponding nuisance parameter, θ_m , and the nuisance parameter controlling the signal strength, μ .

Chapter 11

Results and Interpretation

11.1 Examination of Pre-fit Data

Before proceeding with the full maximum likelihood fit using all the analysis bins, it is helpful to evaluate the observed pre-fit data distributions. The pre-fit N_b distributions for the control region bins are those shown in Figure 10.5 and have corresponding yields given in Table 11.1, while those for the signal region bins are shown in Figure 11.1 and Table 11.2. In Figure 10.5 and Figure 11.1, the background normalization is scaled to match the observed data yields in each bin with a scaling factor that controls all processes equally. This approximates the effect of the various normalization parameters in the fit and allows for a rough comparison of the background and observed N_b distribution shapes. By examining this comparison, a qualitative understanding of the results and expected fit behavior can be gained.

In particular, a generally good level of agreement is seen between the background processes and observed data, suggesting that the background-only fit should be able to describe the data well. There are, however, some small trends that suggest the observed N_b distribution may be slightly harder than that of the simulated background processes. Thus, the background-only fit is expected to increase the nuisance parameters controlling the GS rate, heavy-flavor, and light-flavor b-tag SFs in order to correct the background

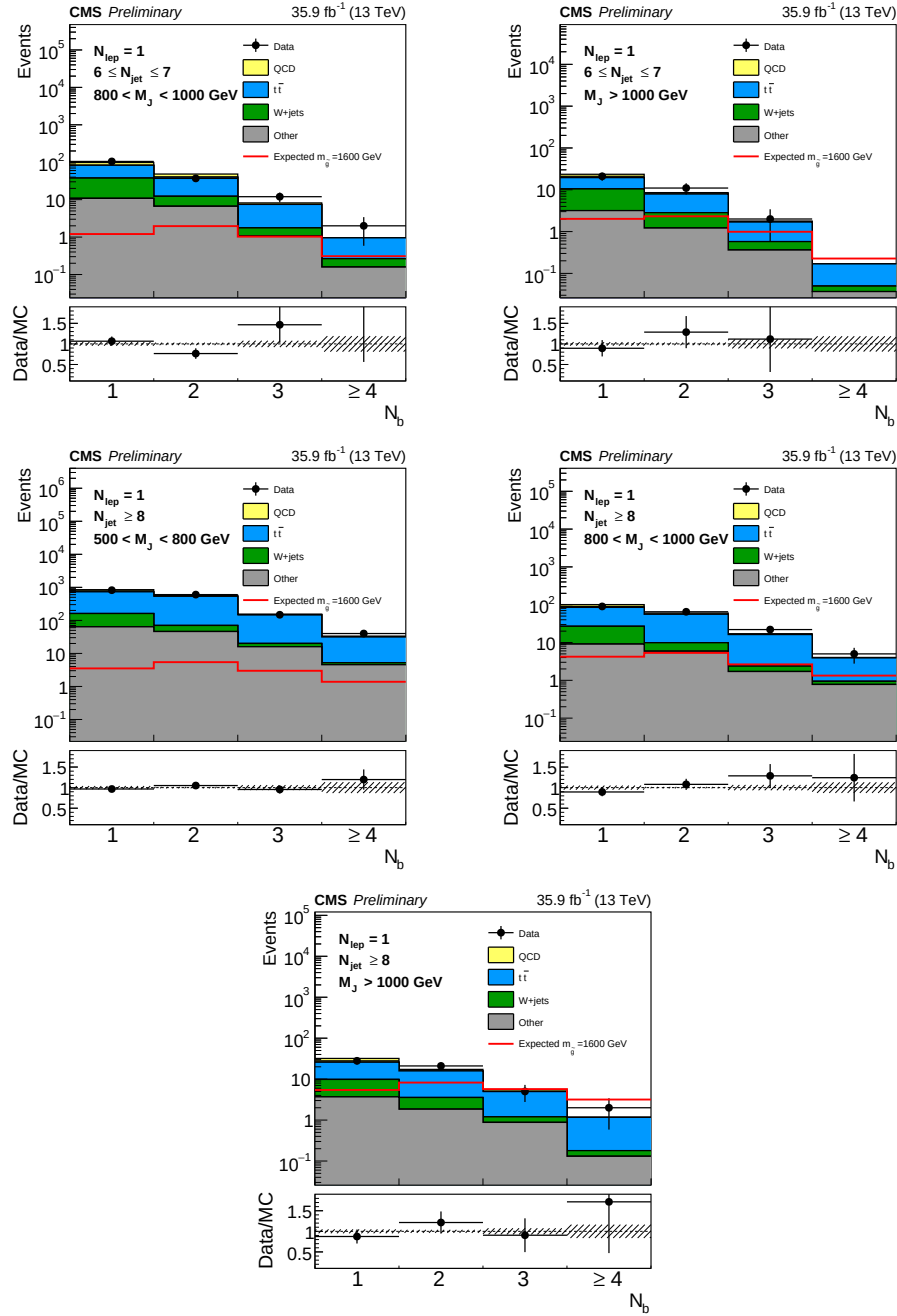


Figure 11.1: Pre-fit N_b distributions of the signal region bins with the background simulation normalized to the observed data yields with a single scaling factor: $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV (upper-left), $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV (upper-right), $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV (middle-left), $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (middle-right), and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (bottom-middle). The ratio of data-to-simulation is shown in the lower panel. The pre-fit uncertainty is represented by the hatched band.

N_b	QCD	$t\bar{t}$	W + jets	Other	All bkg.	Data	Expected $m_{\tilde{g}} = 1600$ GeV
$4 \leq N_{\text{jets}} \leq 5, 500 < M_J \leq 800$ GeV							
1	179.8	317.7	346.5	101.5	945.4	777	0.5
2	33.9	160.0	49.5	34.8	278.1	264	0.4
3	5.1	21.9	4.1	4.8	35.9	34	0.2
≥ 4	0.0	1.8	0.4	0.3	2.5	3	0.0
$4 \leq N_{\text{jets}} \leq 5, M_J > 800$ GeV							
1	21.1	25.6	39.6	12.3	98.5	77	0.3
2	1.3	10.1	5.6	4.2	21.1	18	0.4
3	0.7	1.2	0.4	0.4	2.7	3	0.1
≥ 4	0.0	0.1	0.0	0.0	0.1	0	0.0
$6 \leq N_{\text{jets}} \leq 7, 500 < M_J \leq 800$ GeV							
1	229.8	741.5	299.0	134.6	1404.9	1105	2.5
2	55.3	516.8	60.4	73.7	706.1	588	3.1
3	7.3	100.0	7.4	14.6	129.3	112	1.4
≥ 4	1.6	12.1	0.9	2.3	17.0	21	0.3

Table 11.1: Pre-fit data and simulation yields in the control region bins in 35.9 fb^{-1} of data.

shape accordingly. In the case of the signal-plus-background fit, the fit may also shift these nuisance parameters, or it may extract a signal contribution, depending on whether the observed trends are consistent with variations in systematic uncertainties or with a signal presence.

The quantitative results of the background-only and signal-plus-background fit are described in the section below.

11.2 Results

11.2.1 Background-only Fit Results

The results of a background-only fit of the observed N_b distributions are shown in Figures 11.2 and 11.3. These figures show the $N_{\text{leps}} = 1$ control and signal regions separately although the fit includes all bins simultaneously. The observed N_b distributions are well described by the fit, and an examination of the nuisance parameters, displayed

N_b	QCD	$t\bar{t}$	W + jets	Other	All bkg.	Data	Expected $m_{\tilde{g}} = 1600$ GeV
$6 \leq N_{\text{jets}} \leq 7, 800 < M_J \leq 1000$ GeV							
1	19.6	56.9	33.8	13.6	123.9	105	1.2
2	10.6	34.7	7.1	8.4	60.8	37	2.0
3	0.9	7.2	0.9	1.4	10.3	12	1.0
≥ 4	0.0	0.9	0.1	0.2	1.2	2	0.3
$6 \leq N_{\text{jets}} \leq 7, M_J > 1000$ GeV							
1	6.2	14.0	11.2	4.9	36.2	21	2.0
2	1.0	7.8	2.5	1.9	13.2	11	2.3
3	0.1	1.7	0.3	0.6	2.8	2	1.0
≥ 4	0.0	0.2	0.0	0.1	0.3	0	0.2
$N_{\text{jets}} \geq 8, 500 < M_J \leq 800$ GeV							
1	140.1	683.3	116.6	76.4	1016.4	821	3.5
2	45.1	557.7	28.9	55.0	686.8	603	5.4
3	7.8	153.7	4.9	19.1	185.5	148	3.0
≥ 4	2.4	31.4	0.8	5.4	40.0	40	1.4
$N_{\text{jets}} \geq 8, 800 < M_J \leq 1000$ GeV							
1	20.3	75.5	23.0	11.8	130.5	90	4.2
2	5.6	59.7	5.2	7.6	78.1	65	5.3
3	1.1	18.0	0.9	2.2	22.2	22	2.6
≥ 4	0.1	3.9	0.2	1.0	5.2	5	1.3
$N_{\text{jets}} \geq 8, M_J > 1000$ GeV							
1	7.9	20.9	8.0	4.8	41.6	28	5.4
2	1.9	16.0	2.2	2.4	22.4	21	8.2
3	0.7	4.9	0.4	1.1	7.2	5	5.7
≥ 4	0.0	1.3	0.1	0.2	1.5	2	3.2

Table 11.2: Pre-fit data and simulation yields in the signal region bins in 35.9 fb^{-1} of data.

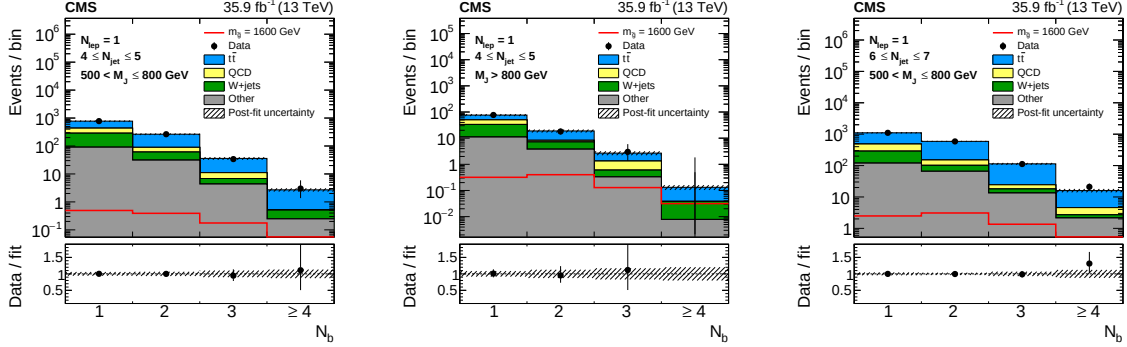


Figure 11.2: Data and the background-only post-fit N_b distribution for the control region bins: $4 \leq N_{\text{jets}} \leq 5$, $500 < M_J \leq 800$ GeV (left), $4 \leq N_{\text{jets}} \leq 5$, $M_J > 800$ GeV (middle), and $6 \leq N_{\text{jets}} \leq 7$, $500 < M_J \leq 800$ GeV (right). The expected signal distribution is also shown for a gluino mass of 1600 GeV (red line). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.

in Figure 11.4, shows that none of them are significantly changed by the fit, with typical deviations less than 0.05 s.d. The largest shifts in the nuisance parameters correspond to those controlling the gluon splitting rate (+0.50 s.d.), light-flavor b-tag SFs (+0.25 s.d.), and heavy-flavor b-tag SFs (+0.14 s.d.) and are well within their pre-fit uncertainties. The post-fit yields are presented in Table 11.3.

The good description of the observed data as well as the behavior of the nuisance parameters match the pre-fit expectations outlined in Section 11.1. An additional indication that the background-only fit is well-behaved is that the post-fit nuisance parameter values are in good agreement with those of the control-region fit. This suggests that the measurements of the nuisance parameters in the background-dominated control regions are able to largely describe the difference in N_b shape between simulation and data in the signal regions without the need of a signal contribution.

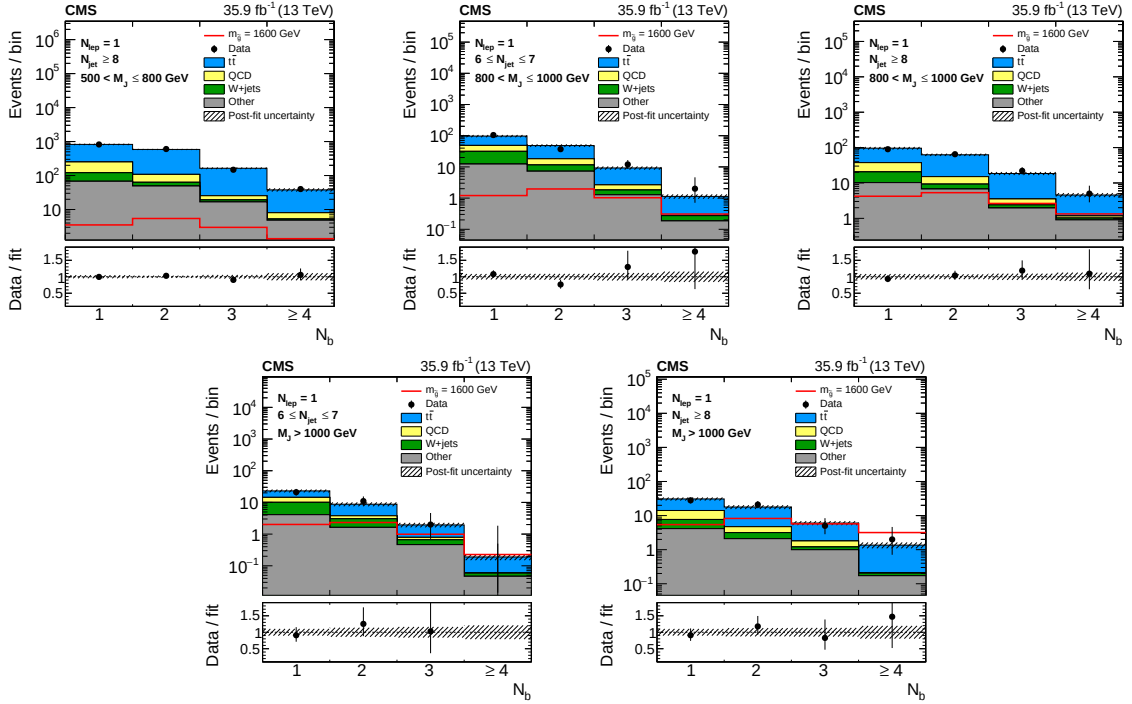


Figure 11.3: Data and the background-only post-fit N_b distribution for the signal region bins: $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV (upper-left), $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV (upper-middle), $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (upper-right), $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV (bottom-left), and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (bottom-right). The expected signal distribution is also shown for a gluino mass of 1600 GeV (red line). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.

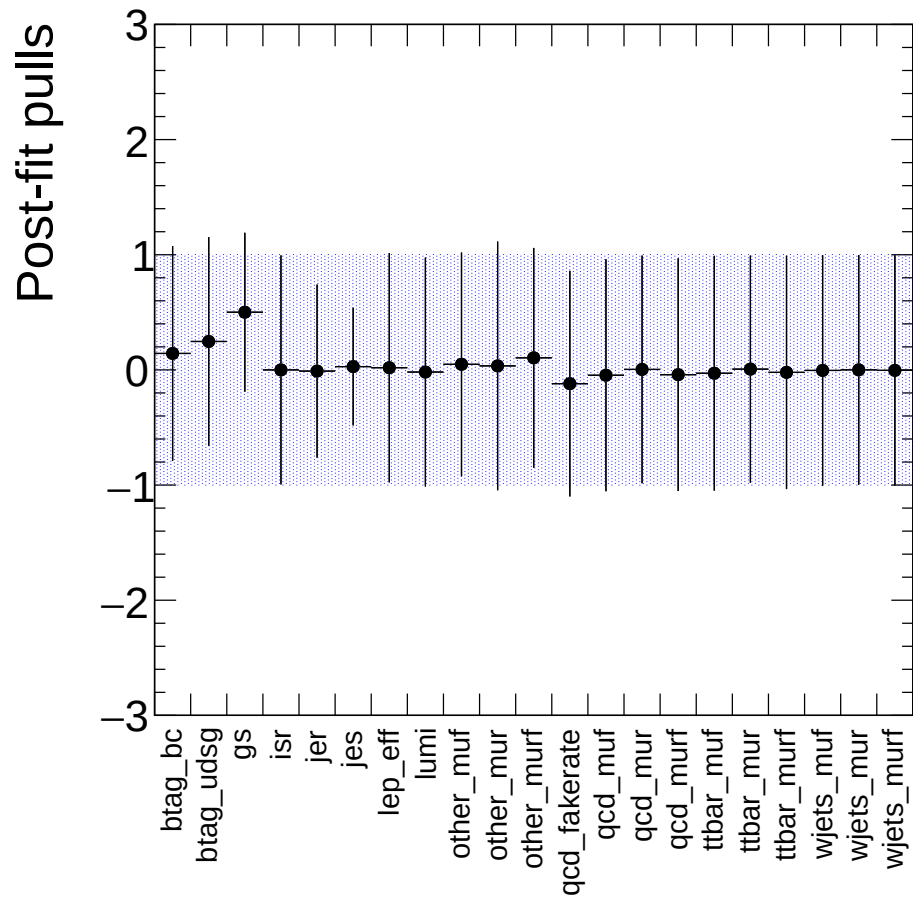


Figure 11.4: Post-fit pulls of the background-only fit. The post-fit values of the nuisance parameters are indicated by data points, while the post-fit uncertainty is shown as a black line and is normalized by the pre-fit uncertainty depicted as the blue band.

N_b	QCD	$t\bar{t}$	W + jets	Other	All bkg.	Data	Expected $m_{\tilde{g}} = 1600$ GeV
$4 \leq N_{\text{jets}} \leq 5, 500 < M_J \leq 800$ GeV							
1	148	340	196	91	775 ± 43	777	0.50 ± 0.13
2	29	175	30	31	264 ± 17	264	0.39 ± 0.11
3	4.3	24.8	2.5	4.4	36 ± 4	34	0.18 ± 0.08
≥ 4	0.0	2.2	0.3	0.2	2.7 ± 0.4	3	0.04 ± 0.04
$4 \leq N_{\text{jets}} \leq 5, M_J > 800$ GeV							
1	16.5	26.3	22.5	11.0	76 ± 6	77	0.32 ± 0.11
2	1.1	10.6	3.4	3.8	19 ± 2	18	0.40 ± 0.12
3	0.7	1.3	0.3	0.3	2.7 ± 0.5	3	0.13 ± 0.06
≥ 4	0.00	0.09	0.03	0.01	0.13 ± 0.03	0	0.03 ± 0.03
$6 \leq N_{\text{jets}} \leq 7, 500 < M_J \leq 800$ GeV							
1	197	620	169	120	1106 ± 48	1105	2.5 ± 0.3
2	49	440	36	66	591 ± 21	588	3.1 ± 0.3
3	6.4	89.2	4.6	13.4	114 ± 8	112	1.4 ± 0.2
≥ 4	1.9	11.4	0.6	2.1	16 ± 2	21	0.25 ± 0.09
$N_{\text{jets}} \geq 8, 500 < M_J \leq 800$ GeV							
1	130	574	53	68	825 ± 38	821	3.5 ± 0.3
2	45	478	14	49	586 ± 20	603	5.4 ± 0.4
3	6.3	138.1	2.5	16.7	164 ± 9	148	3.0 ± 0.3
≥ 4	2.8	29.8	0.4	4.8	38 ± 4	40	1.4 ± 0.2
$6 \leq N_{\text{jets}} \leq 7, 800 < M_J \leq 1000$ GeV							
1	17.3	48.4	19.2	12.3	97 ± 8	105	1.2 ± 0.2
2	6.6	30.1	4.3	7.3	48 ± 4	37	2.0 ± 0.3
3	0.8	6.6	0.5	1.3	9.3 ± 1.0	12	1.0 ± 0.2
≥ 4	0.0	0.9	0.1	0.2	1.1 ± 0.2	2	0.31 ± 0.09
$N_{\text{jets}} \geq 8, 800 < M_J \leq 1000$ GeV							
1	17.0	58.7	10.3	10.2	96 ± 8	90	4.2 ± 0.4
2	5.8	47.5	2.5	6.8	63 ± 5	65	5.3 ± 0.4
3	1.1	15.0	0.4	2.0	19 ± 2	22	2.6 ± 0.3
≥ 4	0.2	3.4	0.1	0.9	4.6 ± 0.6	5	1.3 ± 0.2
$6 \leq N_{\text{jets}} \leq 7, M_J > 1000$ GeV							
1	4.4	8.7	6.0	4.1	23 ± 2	21	2.0 ± 0.3
2	0.7	5.0	1.4	1.6	8.8 ± 1.2	11	2.3 ± 0.3
3	0.1	1.2	0.2	0.5	1.9 ± 0.3	2	1.0 ± 0.2
≥ 4	0.00	0.13	0.01	0.05	0.19 ± 0.04	0	0.23 ± 0.08
$N_{\text{jets}} \geq 8, M_J > 1000$ GeV							
1	6.4	16.7	3.5	4.1	31 ± 3	28	5.4 ± 0.4
2	1.6	13.1	1.1	2.1	18 ± 2	21	8.2 ± 0.5
3	0.6	4.2	0.2	1.0	6.0 ± 0.8	5	5.7 ± 0.4
≥ 4	0.0	1.2	0.0	0.2	1.4 ± 0.3	2	3.2 ± 0.3

Table 11.3: Post-fit yields of the background-only fit, observed data, and expected yields for $m_{\tilde{g}} = 1600$ GeV.

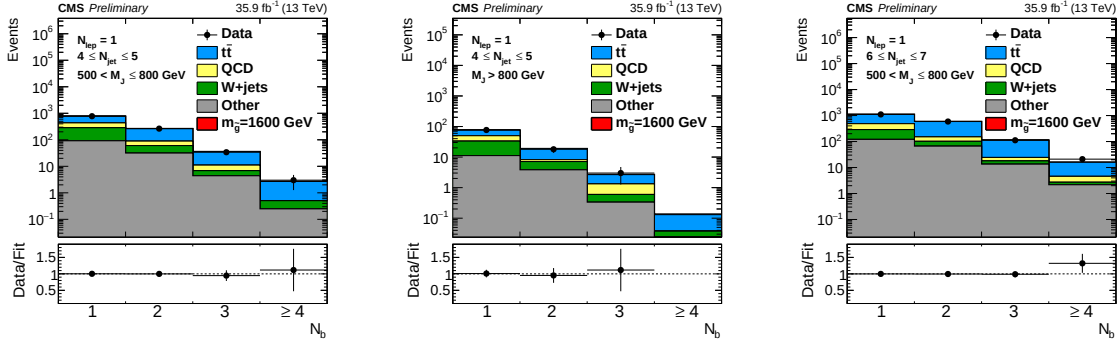


Figure 11.5: Data and the $m_{\tilde{g}} = 1600$ GeV signal-plus-background post-fit N_b distribution for the control region bins: $4 \leq N_{\text{jets}} \leq 5$, $500 < M_J \leq 800$ GeV (left), $4 \leq N_{\text{jets}} \leq 5$, $M_J > 800$ GeV (right), and $6 \leq N_{\text{jets}} \leq 7$, $500 < M_J \leq 800$ GeV (middle). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.

11.2.2 Signal-plus-background Fit Results

Despite the good description of the observed data by the background-only fit, performing a signal-plus-background fit is necessary to see if the observed data is also consistent with the presence of signal. Thus, a signal-plus-background fit is performed for gluino masses ranging from 1000 to 2000 GeV, and, for all masses, the post-fit N_b distribution describes the data well with the fits extracting at most a small and insignificant signal contribution.

For example, the fitted signal strength for a model corresponding to a 1600 GeV gluino is $\mu = 0.18^{+0.41}_{-0.18}$ with the post-fit N_b distributions shown in Figure 11.5 and Figure 11.6 for the control and signal region bins, respectively. The post-fit values of the nuisance parameters are small and consistent with those of the background-only fit, as shown in Table 11.4.

Given the post-fit agreement of the N_b distributions, well-behaved nuisance parameters, and insignificant fitted signal strength, no evidence for a signal corresponding to the T1tbts model is observed.

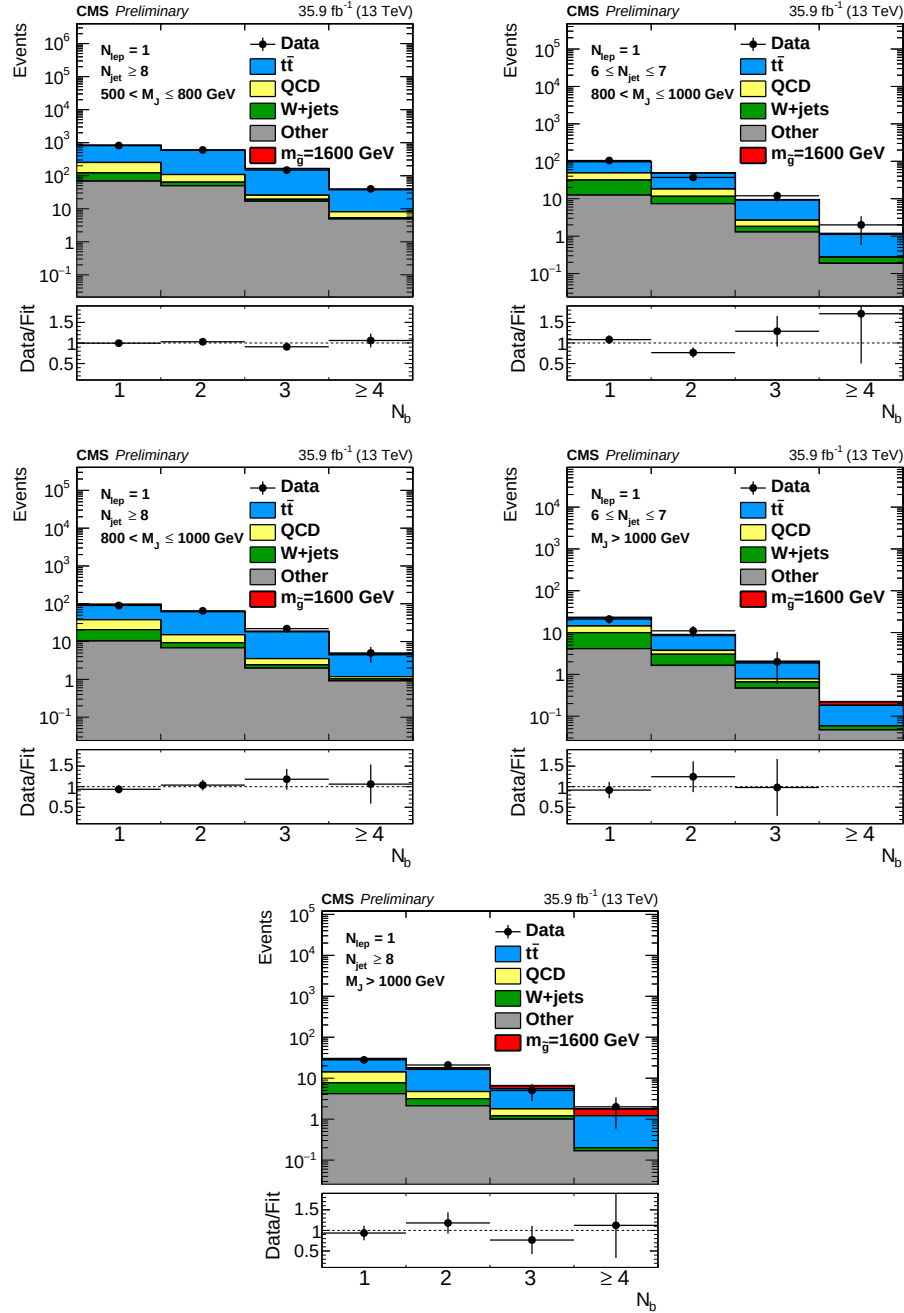


Figure 11.6: Data and the $m_{\tilde{g}} = 1600$ GeV signal-plus-background post-fit N_b distribution for the signal region bins: $N_{\text{jets}} \geq 8$, $500 < M_J \leq 800$ GeV (upper-left), $6 \leq N_{\text{jets}} \leq 7$, $800 < M_J \leq 1000$ GeV (upper-middle), $N_{\text{jets}} \geq 8$, $800 < M_J \leq 1000$ GeV (upper-right), $6 \leq N_{\text{jets}} \leq 7$, $M_J > 1000$ GeV (bottom-left), and $N_{\text{jets}} \geq 8$, $M_J > 1000$ GeV (bottom-right). The ratio of data to post-fit yields is shown in the lower panel. The post-fit uncertainty is depicted as a hatched band.

Nuisance parameter	Post-fit pull (b -only fit)	Post-fit pull ($s + b$ fit)	$\rho(\theta_m, \mu)$
b,c jet b-tag SF (btag_bc)	+0.14, 0.93	+0.12, 0.94	-0.09
u,d,s,g jet b-tag SF (btag_udsg)	+0.25, 0.91	+0.22, 0.91	-0.05
Gluon splitting (gs)	+0.50, 0.69	+0.45, 0.70	-0.14
Jet energy resolution (jer)	-0.01, 0.76	-0.02, 0.73	-0.01
Jet energy scale (jes)	+0.03, 0.53	+0.01, 0.41	-0.06
Lepton efficiency (lep_eff)	+0.02, 1.00	+0.02, 1.00	+0.02
Luminosity (lumi)	-0.02, 0.99	-0.02, 0.99	-0.02
Fact. scale for Other (other_muf)	+0.05, 0.97	+0.04, 0.97	-0.00
Renorm. scale for Other (other_mur)	+0.03, 1.09	+0.04, 1.09	+0.00
Renorm. and Fact. scale for Other (other_murf)	+0.11, 0.96	+0.09, 0.94	-0.00
QCD fake rate (qcd_fakerate)	-0.12, 0.98	-0.10, 0.99	+0.03
Fact. scale for QCD (qcd_muf)	-0.05, 1.01	-0.05, 1.01	+0.00
Renorm. scale for QCD (qcd_mur)	+0.00, 0.99	+0.00, 0.99	+0.00
Renorm. and Fact. scale for QCD (qcd_murf)	-0.04, 1.01	-0.04, 1.01	+0.00
Fact. scale for $t\bar{t}$ (ttbar_muf)	-0.03, 1.02	-0.03, 1.02	+0.00
Renorm. scale for $t\bar{t}$ (ttbar_mur)	+0.01, 0.99	+0.01, 0.98	+0.00
Renorm. and Fact. scale for $t\bar{t}$ (ttbar_murf)	-0.02, 1.01	-0.02, 1.01	+0.00
Fact. scale for $W + \text{jets}$ (wjets_muf)	-0.00, 1.00	-0.00, 1.00	+0.00
Renorm. scale for $W + \text{jets}$ (wjets_mur)	+0.00, 1.00	+0.00, 1.00	-0.00
Renorm. and Fact. scale for $W + \text{jets}$ (wjets_murf)	-0.00, 1.00	-0.00, 1.00	+0.00

Table 11.4: Table of post-fit pulls of the background-only and signal-plus-background fit. The last column, $\rho(\theta_m, \mu)$, lists the correlation between the corresponding nuisance parameter, θ_m , and the nuisance parameter controlling the signal strength, μ .

11.3 Statistical Interpretation

In the absence of significant evidence of signal, limits on the production cross section of a model can be set through the CL_s procedure [113, 114, 115]. The procedure is given by the following steps:

First, construct likelihood functions for the background-only and signal-plus-background, which for this search corresponds to the likelihood presented in Section 10.1.

Next, in order to evaluate the compatibility of the observed data with the background-only and signal-plus-background hypotheses, a test statistic, q_μ , is defined as,

$$q_\mu = -2 \ln \frac{\mathcal{L}(N_{obs}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(N_{obs}|\hat{\mu}, \hat{\theta})}, \quad 0 \leq \hat{\mu} \leq \mu \quad (11.1)$$

Here, N_{obs} is the observed data, μ is the signal strength, θ represents the full set of nuisance parameters. Additionally, $\hat{\theta}_\mu$ is the value of θ that maximizes the likelihood conditioned on a specified value of μ , i.e. for a fixed signal strength, while $\hat{\mu}$ and $\hat{\theta}$ respectively correspond to the values of μ and θ that globally maximize the likelihood. The structure of q_μ as a likelihood ratio is motivated by the Neyman-Pearson lemma [116], which states that the ratio of likelihoods is the most powerful discriminator. The lower constraint of $0 \leq \hat{\mu}$ is imposed on the physical consideration that the signal cannot contribute a negative event yield, while the upper constraint of $\hat{\mu} \leq \mu$ is imposed so that upward fluctuations that result in $\hat{\mu} > \mu$ cannot be interpreted as evidence against the signal-plus-background hypothesis.

For some choice of signal strength, μ' , and the observed data, calculate the observed value of $q_{\mu'}^{obs}$.

Then compute the values θ_0^{obs} and $\theta_{\mu'}^{obs}$, which are the values of nuisance parameters that maximize the background-only ($\mu = 0$) and signal-plus-background ($\mu = \mu'$)

likelihoods under the observed data, respectively.

From the values of θ_0^{obs} and $\theta_{\mu'}^{obs}$, construct the probability density functions (pdfs) for $q_{\mu'}$ under the background-only hypothesis, $f(q_{\mu}|0, \theta_0^{obs})$, and under the signal-plus-background with μ' , $f(q_{\mu}|\mu', \theta_{\mu'}^{obs})$. These pdfs can either be determined by generating and fitting psuedoexperiments or by using asymptotic formulae for approximating the distribution of q_{μ} [117]. The latter option is chosen, as with the former method, the computational time and power needed for generating pseudoexperiemnts for each signal mass point can be prohibitive.

Using these distributions, one can calculate the probability of observing data at least as extreme as $q_{\mu'}^{obs}$ under the background hypothesis, CL_b , and under the signal-plus-background hypothesis, CL_{s+b} . Formally,

$$CL_b = P(q_{\mu} \geq q_{\mu'}^{obs} | b) = \int_{q_0^{obs}}^{\infty} f(q_{\mu}|0, \theta_0^{obs}) dq_{\mu}, \quad (11.2)$$

and

$$CL_{s+b} = P(q_{\mu} \geq q_{\mu'}^{obs} | \mu' s + b) = \int_{q_{\mu'}^{obs}}^{\infty} f(q_{\mu}|\mu', \theta_{\mu'}^{obs}) dq_{\mu}. \quad (11.3)$$

Finally, the CL_s quantity can be calculated as the ratio of these two probabilities, with

$$CL_s(\mu) = \frac{CL_{s+b}}{CL_b}. \quad (11.4)$$

Using the CL_s quantity, a model can be said to be excluded at the $(1 - \alpha)$ confidence level (CL), if for $\mu = 1$, $CL_s \leq \alpha$. Convention is for the 95% CL upper limit on μ to be quoted, corresponding to $\alpha = 0.05$, in which case the value of μ can be scanned until reaching $CL_s = 0.05$. The resulting value of μ then represents the 95% CL upper limit on the signal strength of the model and thus on its production cross section.

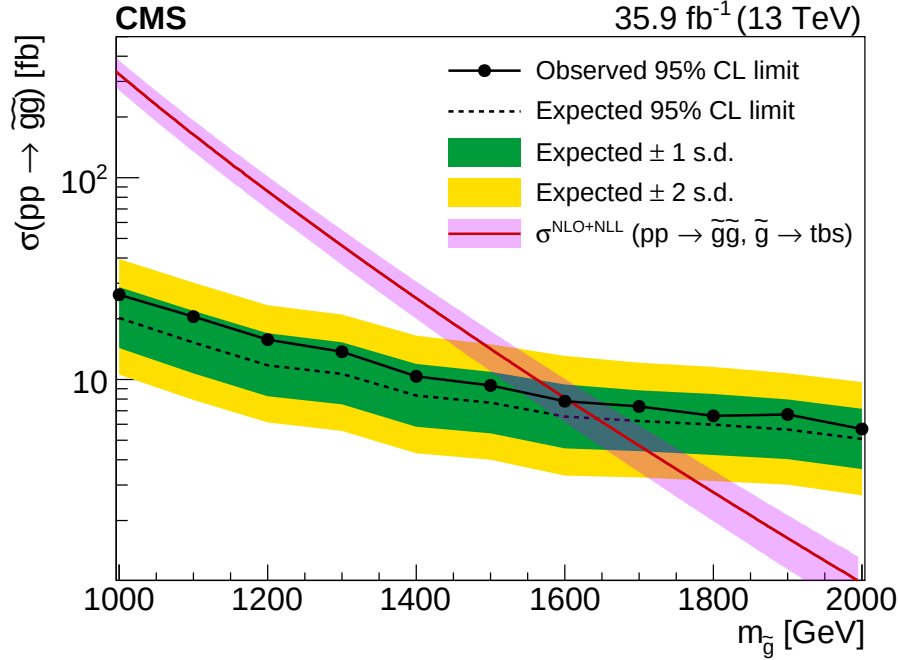


Figure 11.7: Cross section upper limits at 95% CL for a model of gluino pair production with $\tilde{g} \rightarrow tbs$ compared to the gluino pair production cross section. The theoretical uncertainties in the cross section are shown as a band around the red line [21]. The expected limits (dashed line) and their ± 1 s.d. and ± 2 s.d. variations are shown as green and yellow bands, respectively. The observed limit is shown by the solid line with dots.

11.4 Limits on the T1tbs Benchmark Model

The resulting 95% upper limits on the signal production cross section from the CL_s procedure are shown in Figure 11.7, which includes the expected and observed limits along with the gluino pair production cross section. Comparing the observed limits to the gluino pair production cross section [21] indicates that gluino masses below 1610 GeV are excluded in the benchmark $\tilde{g} \rightarrow tbs$ model. The observed limits are slightly less stringent than the expected limits, which is largely due to the insignificant excess in the $M_J > 1000$ GeV, $N_{\text{jets}} > 8$ and $N_b \geq 4$ bin.

Chapter 12

Summary and Conclusions

This dissertation presented a search for new physics in 35.9 fb^{-1} of data produced by $\sqrt{s} = 13 \text{ TeV}$ proton-proton collisions from the LHC and collected by the CMS detector in 2016. The search investigated events with a final state of a single lepton, large jet and b-tagged jet multiplicities, and high sum of large-radius jet masses. This final state is motivated by a R -parity and minimal-flavor violating supersymmetric model, in which gluinos are pair produced and decay via $\tilde{g} \rightarrow t\bar{t} \rightarrow tb$ s. This search, however, is structured to be generically sensitive to models with high-mass signatures and many b-tagged jets, while the lack of an explicit $E_{\text{T}}^{\text{miss}}$ requirement increases the search coverage to even R -parity conserving models in which there is little $E_{\text{T}}^{\text{miss}}$ produced.

The background is predicted through a global maximum-likelihood fit of the distribution of number of b-tagged jets across bins of jet multiplicity and sum of masses of large-radius jets. The normalizations of the dominant backgrounds are measured in data, while their shapes are taken from simulation with corrections measured in data control samples and are allowed to vary in the fit according to their uncertainties.

The main uncertainty in the background prediction method is the statistical uncertainty due to the data sample size, while the largest systematic uncertainties arise from the modeling of the gluon splitting rate and the b quark tagging efficiency and mistag rate.

Results from the background-only fit found the observed data to be well modelled and consistent with the background-only hypothesis. Accordingly, limits are set on a benchmark simplified model where pair produced gluinos each decay via $\tilde{g} \rightarrow t\bar{b}$. An upper limit of approximately 10 fb is measured for the pair production of gluinos in this scenario, which corresponds to excluding gluino masses below 1610 GeV at a 95% confidence level.

These limits represent a significant improvement on the coverage of R-parity violating supersymmetric models, improving on results obtained at $\sqrt{s} = 8$ TeV [29, 86] by approximately 600 GeV and are among the most stringent limits set by both the CMS and ATLAS Collaborations [118, 119].

The original hopes for Run II of the LHC were for a quick discovery of new physics, after which these new particles could be studied in detail and solutions to universal questions obtained. Of course, this has not been the case, and instead many limits have been set. The idea of naturalness, however, remains a highly-compelling reason for new physics to be just around the corner, and despite the lack of evidence, there is still significant phase space left in which (*R*-parity violating) supersymmetric models may be hiding, while still meeting the naturalness guidelines outlined in Section 3.2. Thus, supersymmetric models are an important class of signatures for which searches need to be continued, and, if one truly believes in naturalness, comfort can be taken in that every null result is one step closer to a discovery.

I am proud to know that the contents of this dissertation represent my contribution of one such step to the corpus of particle physics. I sincerely hope that the next generation of SUSY-searchers find themselves describing their results not in terms of limits but of significances. Good luck and happy hunting!

Part IV

Appendix

Appendix A

DeepJEC: Deep Learning the Jet Response

A.1 Introduction

The beginning of the LHC program has been marked by either large increases in center-of-mass energy or integrated luminosity, as shown in Figure A.1, with, for example, the collision energy almost doubling from 7 to 13 TeV and the integrated luminosity collected increasing 10-fold from 2015 to 2016. With each increase, large amounts of new phase space were opened and inclusive searches attempted to cover this space by broadly looking for signs of new physics. However, with the LHC's energy capabilities nearly saturated and the increase in data collection rate slowing down, more targeted searches that try to extract hints of new physics from precision measurements are becoming more well-motivated. These searches will require a more precise understanding of our detectors.

One physics object that is particularly difficult to measure is jets, as each jet is composed of roughly ~ 10 -100 correlated particles that are of various particle-types (charged hadrons, neutral hadrons, photons, etc), each with their own energy, and incident on many different detector channels. All of these factors affect how well a detector measures a jet and thus contribute to a wide detector response for jets, where response is defined

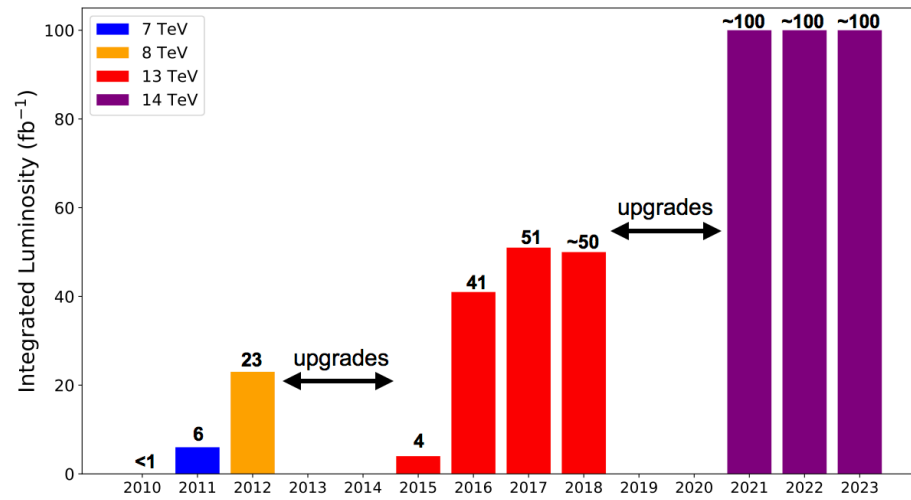


Figure A.1: The (planned) integrated luminosity delivered and operating center-of-mass energy for the LHC from 2010-2023.

as

$$\mathcal{R} = \frac{p_{\text{T}}^{\text{reconstructed}}}{p_{\text{T}}^{\text{true}}}. \quad (\text{A.1})$$

By understanding how the jet response depends on these factors, jet measurements can be better corrected. This can provide both better resolution for measurements, like those that reconstruct the Higgs mass from $b\bar{b}$ pairs, as well as better background-rejection for searches where, for example, events with mis-measured jets comprise the largest background.

Current methods [120, 121] only measure the jet response as a function of jet p_{T} and η , which are important first-order effects to consider. Figure A.2 shows that the mean response does vary with both variables, where the disjoint η -dependence is due to different detector technologies used at different values of η . The main goal of measuring the jet response is to determine the mean of the response distribution correctly, as after correcting for this bias, jets will be measured correctly on average. Figure A.3 shows the jet response distribution (blue) and a model¹ of the distribution as a function of only jet p_{T} and η (green, denoted $p_{\text{T},\eta}$) overlaid and depicts the limitations of using only these variables: While the green distribution does capture the mean of the blue distribution, it does not model much of the distribution’s width, which results in a wide jet resolution for the detector. Although much of this width is due to inherently stochastic processes (e.g. photon production in scintillators), a question worth investigating is how much of the jet response can be modeled by including the particle-level information of jets.

¹This model was generated by training a simple 3-hidden layer neural network with 32 hidden units each on only the jet p_{T} and η with the jet response as the target. The training and test set details are the same as those described in Section A.3. The model complexity here far exceeds that which is necessary for a two-variable regression but is used for equal comparison to later results.

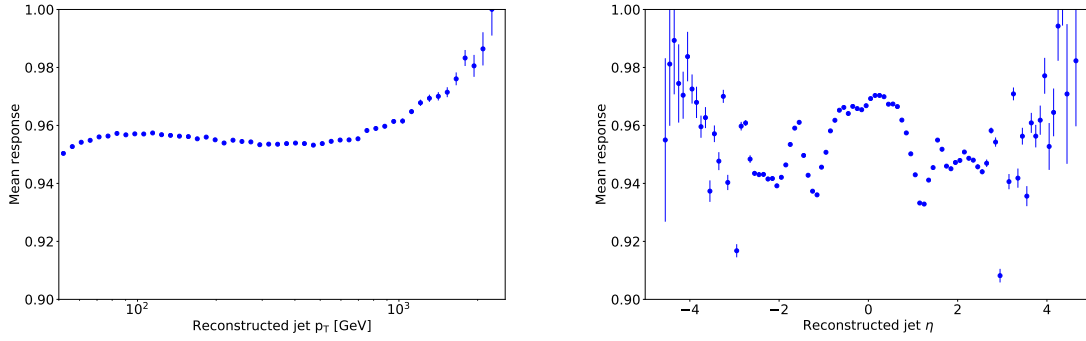


Figure A.2: The mean jet response as a function of jet p_T (left) and η (right).

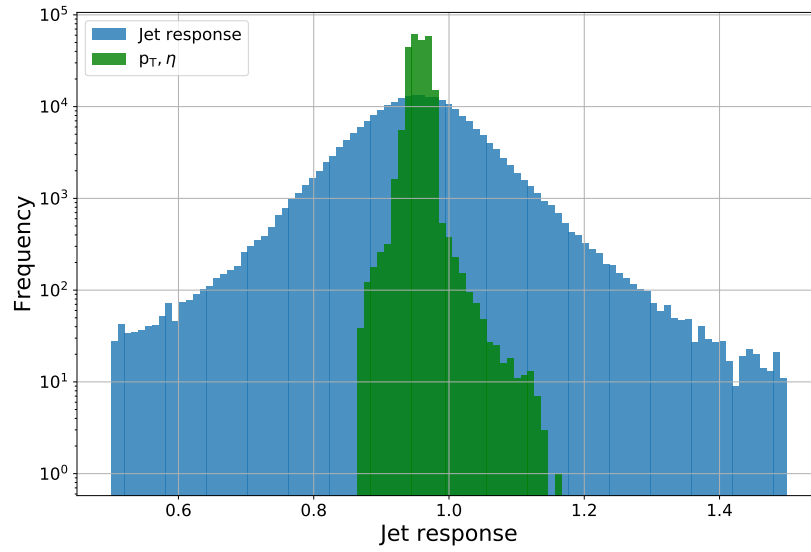


Figure A.3: The true jet response distribution (blue) and a model of the response as a function of jet p_T and η (green).

A.2 Jet Images

One way to capture this particle-level information is by using the “jet images” technique [122, 123], which is a way of converting the detector signals of a jet into a 2D image that encodes this information. Sophisticated image processing techniques can then be used on these images in order to extract the dependence of the detector response on the spatial and energy correlations of the jet fragmentation. The use of jet images has been demonstrated to be effective for a variety of tasks, such as quark-gluon discrimination, W boson and top tagging, and in jet quenching studies with heavy-ion collisions [122, 123, 124, 125, 126, 127].

The process used in this study for generating is carried out for each jet in an event and is as follows:

- First, define a 2D histogram in η - ϕ space centered around the jet’s η and ϕ . The exact binning choice is arbitrary, but it is helpful that the segmentation of the histogram approximates the segmentation of the detector and that the range of the axes depend on the radius of the jet clustering algorithm. Here, for $R = 0.5$ jets, a range of two with 25 bins is used for both the η and ϕ dimensions.
- Next, fill the histogram based on the η and ϕ of the constituent jet particles with the weight of each entry equal to the particle’s p_T .
- Finally, center the image such that the origin ($\eta = 0$, $\phi = 0$) corresponds to the highest p_T bin and normalize the image such that the value of the highest p_T bin equals one.

Once completed, each bin of the histogram can be thought of as a “pixel” of an image with an intensity equal to the weight of that bin. These images encode the particle-level spatial and energy information of a jet, and an example jet image is shown in Figure A.4.

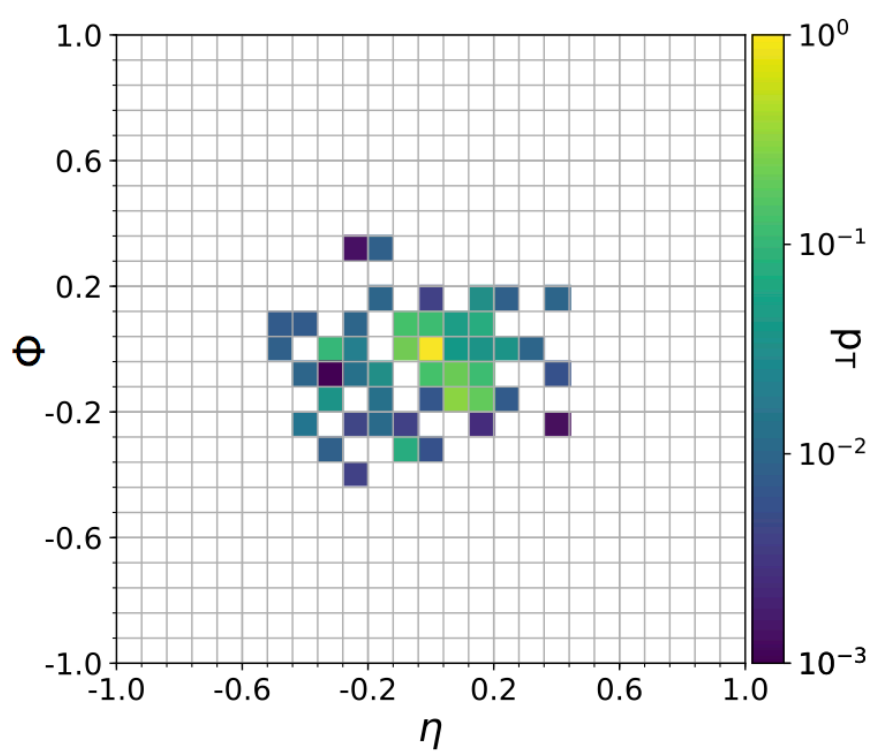


Figure A.4: An example jet image.

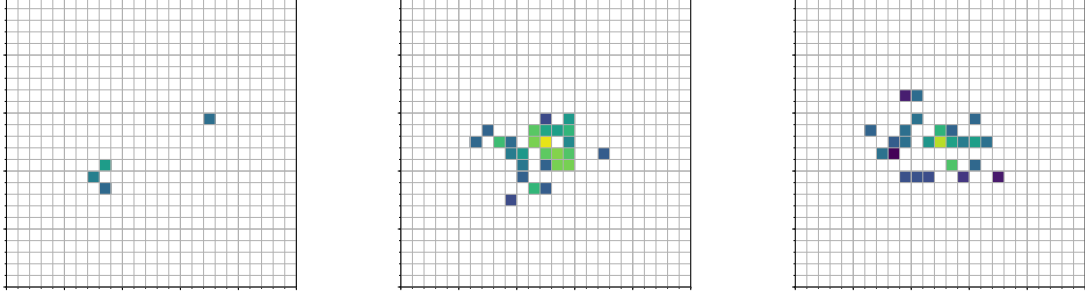


Figure A.5: The example jet image from Figure A.4 split by particle type: charged hadrons (left), neutral hadrons (right), and photons + electrons + muons (right).

Detectors, however, have different responses to different particles, and this information can be additionally encoded by splitting the jet images by particle type, where a separate jet image is constructed for each particle type. Figure A.5 shows the example jet image now split into three categories: charged hadrons, neutral hadrons, and photons + electrons + muons (which is dominated by, and hereafter referred to as, photons). The resulting three images can then be combined into a single 3-channel jet image, completely analogous to how a pixel encodes an R, G, and, B value in traditional colored images. One can see that there is an information gain by comparing the average unnormalized jet images for each category, as shown in Figure A.6. For example, while subtle, it can be gleamed just by eye that the center of the average neutral hadron jet image is “brighter” than that of the average charged hadron image, which indicates that on average neutral hadrons are higher in p_T than are charged hadrons. This information, along with other much lower-level information, is now made accessible. The idea of “colorizing” jet images was first presented in Reference [127].

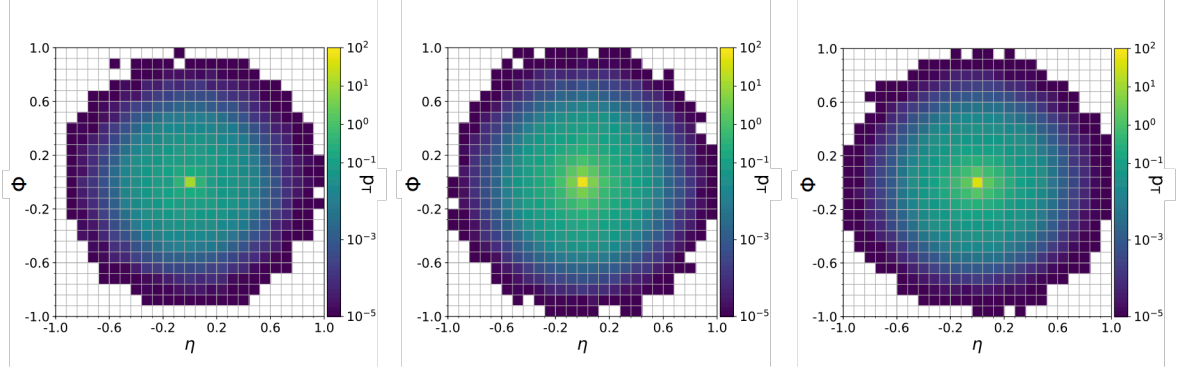


Figure A.6: Average unnormalized jet image for charged hadrons (left), neutral hadrons (middle), and photons + electron + muons (right).

A.3 Network Architecture

Now that jet images have been constructed that encode the particle-level jet information, the correlations between this information and the detector response needs to be extracted. This can be done by training a deep convolutional neural network (CNN) and having it additionally “learn” the jet response as a function of jet fragmentation. The network architecture is composed of two main parts: a set of convolutional layers that are used as feature extractors and a set of fully-connected layers that build low-level features from those extracted by the convolutional layer.

Because these jet images are sparse, i.e. only 5-10% of pixels are non-zero, the CNN architecture deviates from typical image processing heuristics. For example, instead of standard 3x3 filters, the network uses relatively large filters in early layers in order to encapsulate more of the image and improve the network training. Then, as the network reduces the image size through max-pooling layers, the filter sizes are correspondingly reduced, while increasing the number of filters in each layer to allow the network to learn many complex low-level features.

The exact network architecture is shown in Figure A.7 and is detailed below:

- Input layer that accepts 25x25x3 jet images
- Four convolutional layers with tanh activations followed by 2x2 max-pooling layers
 - 4 - 9x9 filters
 - 8 - 6x6 filters
 - 16 - 3x3 filters
 - 32 - 1x1 filters
- Flatten and merge jet p_T and η features
- Three fully connected layers, each with 32 hidden-units, ReLu activations, and 10% dropout
- Output layer with linear activation that outputs the predicted jet response

The network is trained using ~ 2 million jets and tested on $\sim 200,000$ jets. These jets are taken from a 2012, $\sqrt{s} = 8$ TeV simulation of QCD events obtained from CMS Open Data [128] that does not include pileup interactions. Jets in this sample are formed using the anti- k_T clustering algorithm with $R = 0.5$ and are selected by requiring that they have $p_T > 50$ GeV. The mean squared error (MSE) loss function is optimized with the ADAM optimizer (learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$) using stochastic gradient descent with a batch size of 256.

A.4 Results

The results from the training of this network are presented in Figure A.8 which shows the true jet response distribution of the test set (blue) and the corresponding predictions of the $p_{T,\eta}$ model (green) and the CNN model (orange, denoted $p_{T,\eta} +$

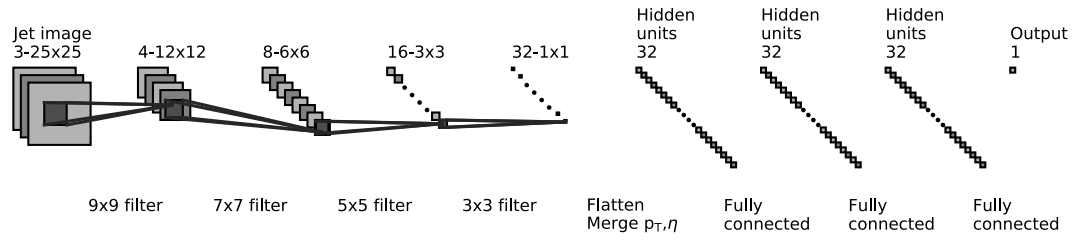


Figure A.7: Schematic of the convolutional neural network architecture.

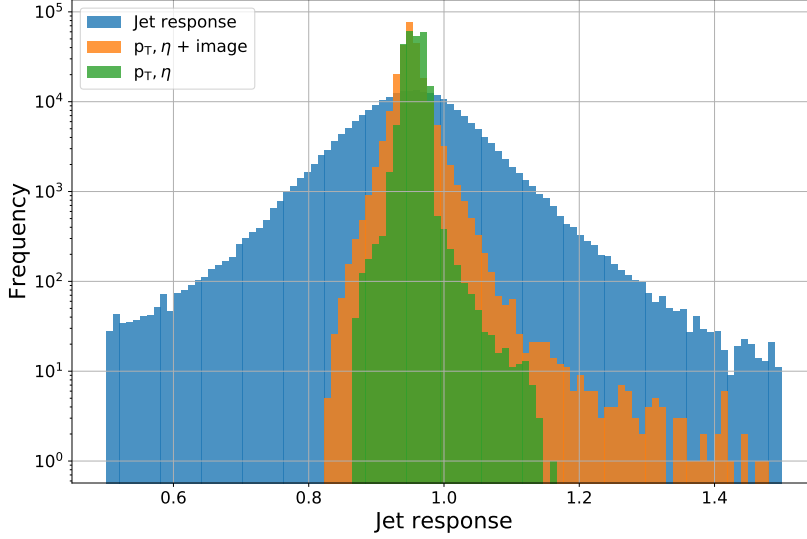


Figure A.8: The true jet response distribution (blue), a model of the response trained only on jet p_T and η (green), and a model of the response additionally trained with jet images (orange).

image). This figure shows that the CNN is able to capture a wider range of the jet response, corresponding to a $\sim 10\%$ improvement with respect to the MSE. Thus, one can see that the jet images not only contain extra relevant information but also that the CNN is able to extract it. One thing of note is that the CNN seems to model the high-response tail better than the low-response tail. This behavior is not yet fully understood but believed to be due to biases in the jet selection of the training set.

One way to see what the CNN is learning is by examining how the response is modeled as a function of particle-type. Figure A.9 shows the mean response in bins of the fractional jet energy due to charged hadrons (left), neutral hadrons (middle), and photons (right). The mean response of the model using only p_T and η , depicted in green data points, is flat as a function of jet energy in each of the plots, which is expected because the network was not given any information about particle type. The mean response of the

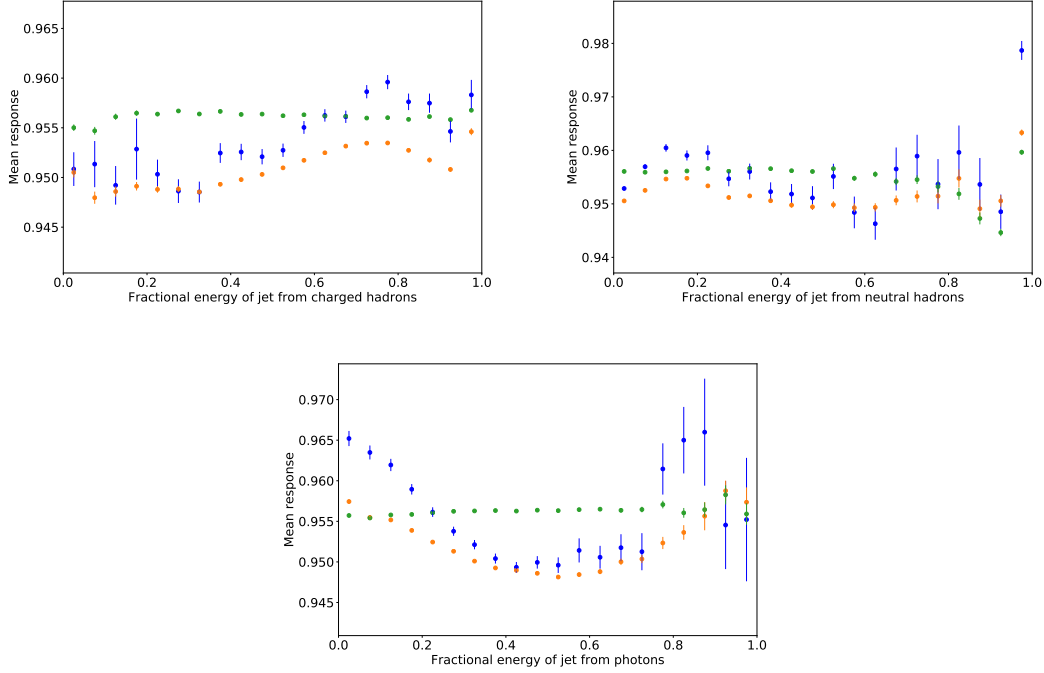


Figure A.9: The mean jet response as a function of the fractional jet energy from charged hadrons (left), neutral hadrons (middle), and photons + electrons + muons (right). The truth-level distributions are shown in blue, while the predictions from the $p_{T,\eta}$ model and $p_{T,\eta}$ + image model are shown in green and orange, respectively.

CNN model (orange), however, is starting to show the same dependencies on particle type as the true jet response (blue). It is believed that with a larger training set this modelling will improve further.

At this point, one may wonder how much of this improvement is due to the jet images versus the network simply having more information. To separate the two effects, a dense neural network (corresponding to just the fully connected portion of the CNN network) is trained in which information on the jet fragmentation is included, i.e. the multiplicity and fractional jet energy of the particle types. This model, shown as the purple distribution in Figure A.10, is able to explain only roughly half the improvement of the model trained on jet images, as it improves over the $p_{T,\eta}$ model by about 5% with

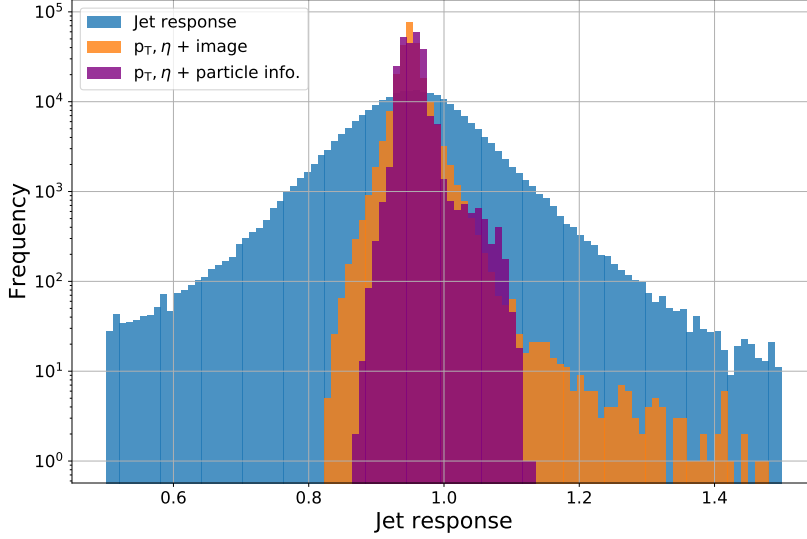


Figure A.10: The true jet response distribution (blue), model of the response trained on jet p_T , η , and jet images (orange), and a model trained on jet p_T , η , and jet fragmentation information (purple).

respect to the MSE. Thus, it appears that the spatial and energy correlations encoded in jet images carry additional information important for predicting the jet response.

A.5 Conclusions

While these results are certainly preliminary and much work is left to do, there is early indication that by encoding the particle-level information of jets through jet images, one can significantly improve the jet response measurement. This is because these jet images not only include the individual particle information but also the energy and spatial correlations of the jet fragmentation. This has the potential to help improve both the core and the tail of the jet resolution.

There are, however, many steps necessary before this potential can be realized. This

includes increasing the training set size to ~ 10 million jets, as well as including pileup effects. Additionally, a more detailed investigation is needed to understand what the network is learning. Lastly, this method will have to be validated in data as the details of the jet fragmentation may not be well simulated, and in this case, features developed by training on simulated jets may not correspond to useful features for jets in data. This can be done through tag-and-probe methods using, for example, photon + jets or QCD di-jet events.

Bibliography

- [1] D. Overbye, *Physicists find elusive particle seen as key to universe*, *The New York Times* (Jul, 2012).
- [2] CMS Collaboration, “Summaries of CMS cross section measurements.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>, Jan, 2018. Accessed: May 08, 2018.
- [3] ATLAS Collaboration, *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, *Phys. Rev.* **D90** (2014), no. 11 112015, [arXiv:1408.7084].
- [4] CMS Collaboration, *Measurement of the properties of a Higgs boson in the four-lepton final state*, *Phys. Rev.* **D89** (2014), no. 9 092007, [arXiv:1312.5353].
- [5] M. Papucci, J. T. Ruderman, and A. Weiler, *Natural SUSY endures*, *JHEP* **09** (2012) 035, [arXiv:1110.6926].
- [6] CMS Collaboration, “CMS supersymmetry physics results.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS>, April, 2018. Accessed: May 09, 2018.
- [7] ATLAS Collaboration, “Summary plots from the ATLAS Supersymmetry physics group.” <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY/>, Dec, 2017. Accessed: May 09, 2018.
- [8] B. C. Allanach, *Beyond the standard model lectures for the 2016 European School of High-Energy Physics*, in *Proceedings, 2016 European School of High-Energy Physics (ESHEP2016): Skeikampen, Norway, June 15-28 2016*, (2017), pp. 123–152, 2017. arXiv:1609.0201.
- [9] F. Marcastel, *CERN’s accelerator complex. La chaîne des accélérateurs du CERN, General Photo* (Oct, 2013).
- [10] W. J. Stirling, *13/8 TeV LHC luminosity ratios*, *Private Communication* (2013).

- [11] M. Hoch and M. Brice. https://home.cern/sites/home.web.cern.ch/files/image/update-for_the_public/2013/12/cms-detector.jpg, May, 2017. Accessed: May 20, 2018.
- [12] L. R. F. Castillo, *The ATLAS and CMS detectors*, in *The search and discovery of the Higgs boson*, 2053-2571, pp. 4–1 to 4–8. Morgan & Claypool Publishers, 2015.
- [13] M. Schott and M. Dunford, *Review of single vector boson production in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **74** (May, 2014) 60 p.
- [14] T. Lenzi, *Development and study of different muon track reconstruction algorithms for the level-1 trigger for the CMS muon upgrade with GEM detectors*, Master’s thesis, U. Brussels, 2013.
- [15] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV*. PhD thesis, Bogazici U., 2011. arXiv:1308.6064.
- [16] CMS Collaboration, *Identification and filtering of uncharacteristic noise in the cms hadron calorimeter*, *Journal of Instrumentation* **5** (2010), no. 03 T03014.
- [17] P. Paolucci *et. al.*, *CMS resistive plate chamber overview, from the present system to the upgrade phase I*, *Journal of Instrumentation* **8** (2013), no. 04 P04005.
- [18] CMS Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [19] CMS Collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, *Submitted to: JINST* (2017) [arXiv:1712.0715].
- [20] CMS Collaboration, “Public CMS Luminosity Information.” https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016_proton_proton_13_TeV_collis, April, 2018. Accessed: April 23, 2018.
- [21] C. Borschensky, M. Krämer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, *Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$, and 100 TeV*, *Eur. Phys. J. C* **74** (2014) 3174, [arXiv:1407.5066].
- [22] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, *Tenth-order QED contribution to the electron $g-2$ and an improved value of the fine structure constant*, *Phys. Rev. Lett.* **109** (Sep, 2012) 111807.
- [23] D. Hanneke, S. Fogwell, and G. Gabrielse, *New measurement of the electron magnetic moment and the fine structure constant*, *Phys. Rev. Lett.* **100** (Mar, 2008) 120801.

- [24] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse, *Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment*, *Phys. Rev. A* **83** (May, 2011) 052122.
- [25] C. Csaki, Y. Grossman, and B. Heidenreich, *MFV SUSY: A natural theory for R-parity violation*, *Phys. Rev.* **D85** (2012) 095009, [arXiv:1111.1239].
- [26] CMS Collaboration, *Search for R-parity violating supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV using b jets in a final state with a single lepton, many jets, and high sum of large-radius jet masses*, *Physics Letters B* **783** (2018) 114 – 139, [arXiv:1712.0892].
- [27] CMS Collaboration, *Search for R-parity violating supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV using b jets in a final state with a single lepton, many jets, and high sum of large-radius jet masses*, . Accessed: May 17, 2018.
- [28] CMS Collaboration, *Search for R-parity-violating SUSY in final states with zero or one lepton and large multiplicity of jets and b-tagged jets*, Tech. Rep. CMS-PAS-SUS-16-013, CERN, Geneva, 2016.
- [29] CMS Collaboration, *Searches for R-parity-violating supersymmetry in pp collisions at $\sqrt{s} = 8$ TeV in final states with 0-4 leptons*, *Phys. Rev.* **D94** (2016), no. 11 112009, [arXiv:1606.0807].
- [30] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, *Phys. Rept.* **405** (2005) 279–390, [hep-ph/0404175].
- [31] V. C. Rubin and W. K. Ford, Jr., *Rotation of the Andromeda nebula from a spectroscopic survey of emission regions*, *Astrophys. J.* **159** (1970) 379–403.
- [32] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [arXiv:1207.7214].
- [33] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [arXiv:1207.7235].
- [34] CMS Collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **06** (2013) 081, [arXiv:1303.4571].
- [35] CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the Standard Model predictions using proton collisions at 7 and 8 TeV*, *Eur. Phys. J.* **C75** (2015), no. 5 212, [arXiv:1412.8662].

- [36] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb⁻¹ of pp collision data*, *Phys. Rev.* **D90** (2014), no. 5 052004, [arXiv:1406.3827].
- [37] ATLAS and CMS Collaborations, *Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803, [arXiv:1503.0758].
- [38] R. Barbieri and G. F. Giudice, *Upper bounds on supersymmetric particle masses*, *Nucl. Phys.* **B306** (1988) 63–76.
- [39] J. L. Feng, *Naturalness and the status of supersymmetry*, *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 351–382, [arXiv:1302.6587].
- [40] N. Craig, *The state of supersymmetry after Run I of the LHC*, in *Beyond the Standard Model after the first run of the LHC Arcetri, Florence, Italy, May 20-July 12, 2013*, 2013. arXiv:1309.0528.
- [41] J. A. Casas, J. Moreno, S. Robles, K. Rolbiecki, and B. Zaldivar, *What is a natural SUSY scenario?*, *JHEP* **06** (2015) 070, [arXiv:1407.6966].
- [42] E. Witten, *Dynamical breaking of supersymmetry*, *Nucl. Phys. B* **188** (1981) 513.
- [43] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and SU(5)*, *Nucl. Phys. B* **193** (1981) 150.
- [44] P. Ramond, *Dual theory for free fermions*, *Phys. Rev. D* **3** (1971) 2415.
- [45] Y. A. Gol’fand and E. P. Likhtman, *Extension of the algebra of Poincaré group generators and violation of P invariance*, *JETP Lett.* **13** (1971) 323.
- [46] A. Neveu and J. H. Schwarz, *Factorizable dual model of pions*, *Nucl. Phys. B* **31** (1971) 86.
- [47] D. V. Volkov and V. P. Akulov, *Possible universal neutrino interaction*, *JETP Lett.* **16** (1972) 438.
- [48] J. Wess and B. Zumino, *A lagrangian model invariant under supergauge transformations*, *Phys. Lett. B* **49** (1974) 52.
- [49] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*, *Nucl. Phys. B* **70** (1974) 39.
- [50] P. Fayet, *Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino*, *Nucl. Phys. B* **90** (1975) 104.
- [51] H. P. Nilles, *Supersymmetry, supergravity and particle physics*, *Phys. Rep.* **110** (1984) 1.

- [52] C. Csaki, *The minimal supersymmetric standard model (MSSM)*, *Mod. Phys. Lett.* **A11** (1996) 599, [hep-ph/9606414].
- [53] R. Barbier *et. al.*, *R-parity violating supersymmetry*, *Phys. Rept.* **420** (2005) 1–202, [hep-ph/0406039].
- [54] B. C. Allanach, A. Dedes, and H. K. Dreiner, *Bounds on R-parity violating couplings at the weak scale and at the GUT scale*, *Phys. Rev.* **D60** (1999) 075014, [hep-ph/9906209].
- [55] B. Bajc, J. Hisano, T. Kuwahara, and Y. Omura, *Threshold corrections to dimension-six proton decay operators in non-minimal SUSY SU(5) GUTs*, *Nucl. Phys.* **B910** (2016) 1–22, [arXiv:1603.0356].
- [56] Super-Kamiokande Collaboration, *Search for proton decay via $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ in a large water Cherenkov detector*, *Phys. Rev. Lett.* **102** (2009) 141801, [arXiv:0903.0676].
- [57] C. Csaki and B. Heidenreich, *A complete model for R-parity violation*, *Phys. Rev.* **D88** (2013) 055023, [arXiv:1302.0004].
- [58] G. Krnjaic and D. Stolarski, *Gauging the way to MFV*, *JHEP* **04** (2013) 064, [arXiv:1212.4860].
- [59] L. Evans and P. Bryant, *LHC machine*, *JINST* **3** (2008) S08001.
- [60] CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *JINST* **9** (2014), no. 10 P10009, [arXiv:1405.6569].
- [61] K. Rose, *Deterministic annealing for clustering, compression, classification, regression, and related optimization problems*, *Proceedings of the IEEE* **86** (Nov, 1998) 2210–2239.
- [62] Früwirth, R and Waltenberger, Wolfgang and Vanlaer, Pascal, *Adaptive vertex fitting*, Tech. Rep. CMS-NOTE-2007-008, CERN, Geneva, Mar, 2007.
- [63] CMS Collaboration, *Particle-flow reconstruction and global event description with the cms detector*, *Journal of Instrumentation* **12** (2017), no. 10 P10003.
- [64] CMS Collaboration, *Particle-flow event reconstruction in CMS and performance for jets, taus, and MET*, Tech. Rep. CMS-PAS-PFT-09-001, CERN, Geneva, Apr, 2009.
- [65] CMS Collaboration, *Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector*, Tech. Rep. CMS-PAS-PFT-10-001, 2010.

- [66] CMS Collaboration, “Cut-based electron identification in Run 2.” https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2Archive#Spring15_selection_25ns, May, 2017. Accessed: May 01, 2018.
- [67] K. Rehermann and B. Tweedie, *Efficient identification of boosted semileptonic top quarks at the LHC*, *JHEP* **03** (2011) 059, [arXiv:1007.2221].
- [68] CMS Collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *JINST* **10** (2015) P06005, [arXiv:1502.0270].
- [69] CMS Collaboration, “Baseline muon selections for Run-II.” https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2#Medium_Muon, April, 2018. Accessed: May 01, 2018.
- [70] CMS Collaboration, *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV*, *JINST* **7** (2012) P10002, [arXiv:1206.4071].
- [71] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D48** (1993) 3160–3166, [hep-ph/9305266].
- [72] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [hep-ph/9707323].
- [73] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_T jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189].
- [74] G. P. Salam, *Towards Jetography*, *Eur. Phys. J.* **C67** (2010) 637–686, [arXiv:0906.1833].
- [75] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, *Eur. Phys. J.* **C72** (2012) 1896, [arXiv:1111.6097].
- [76] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett.* **B659** (2008) 119–126, [arXiv:0707.1378].
- [77] CMS Collaboration, “Jet identification for the 13 TeV data Run2016.” <https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2016>, February, 2018. Accessed: May 01, 2018.
- [78] The CMS Collaboration, *Determination of jet energy calibration and transverse momentum resolution in CMS*, *Journal of Instrumentation* **6** (2011), no. 11 P11002.
- [79] CMS Collaboration, *Identification of b-quark jets with the CMS experiment*, *JINST* **8** (2013) P04013, [arXiv:1211.4462].

- [80] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B641** (2006) 57–61, [hep-ph/0512210].
- [81] ATLAS Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb⁻¹ of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, arXiv:1711.1152.
- [82] CMS Collaboration, *Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large-radius jets*, *JHEP* **08** (2016) 122, [arXiv:1605.0460].
- [83] A. Hook, E. Izaguirre, M. Lisanti, and J. G. Wacker, *High multiplicity searches at the LHC using jet masses*, *Phys. Rev.* **D85** (2012) 055029, [arXiv:1202.0558].
- [84] T. Cohen, E. Izaguirre, M. Lisanti, and H. K. Lou, *Jet substructure by accident*, *JHEP* **03** (2013) 161, [arXiv:1212.1456].
- [85] S. El Hedri, A. Hook, M. Jankowiak, and J. G. Wacker, *Learning how to count: a high multiplicity search for the LHC*, *JHEP* **08** (2013) 136, [arXiv:1302.1870].
- [86] ATLAS Collaboration, *Search for massive supersymmetric particles decaying to many jets using the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Rev.* **D91** (2015), no. 11 112016, [arXiv:1502.0568]. [Erratum: *Phys. Rev.* **D93**, no. 3, 039901 (2016)].
- [87] ATLAS Collaboration, *Search for new phenomena in final states with large jet multiplicities and missing transverse momentum at $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS experiment*, *JHEP* **10** (2013) 130, [arXiv:1308.1841]. [Erratum: *JHEP* **01**, 109 (2014)].
- [88] CMS Collaboration, *Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large-radius jets*, *Phys. Rev. Lett.* **119** (2017), no. 15 151802, [arXiv:1705.0467].
- [89] CMS Collaboration, *Commissioning the performance of key observables used in SUSY searches with the first 13 TeV data*, .
- [90] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301].
- [91] J. Alwall et. al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur. Phys. J. C* **53** (2008) 473, [arXiv:0706.2569].

- [92] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061, [arXiv:1209.6215].
- [93] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].
- [94] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [arXiv:0709.2092].
- [95] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581].
- [96] NNPDF Collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849].
- [97] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159.
- [98] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 tune*, *Eur. Phys. J. C* **74** (2014) 3024.
- [99] GEANT4 Collaboration, *GEANT4—a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [100] M. Czakon, P. Fiedler, and A. Mitov, *Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_s^4)$* , *Phys. Rev. Lett.* **110** (Jun, 2013) 252004.
- [101] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *W physics at the LHC with FEWZ 2.1*, *Comput. Phys. Commun.* **184** (2013) 208, [arXiv:1201.5896].
- [102] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, *JHEP* **09** (2009) 111, [arXiv:0907.4076]. [Erratum: DOI10.1007/JHEP02(2010)011].
- [103] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547, [arXiv:1009.2450].
- [104] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, *JHEP* **06** (2015) 184, [arXiv:1504.0344].
- [105] G. Bevilacqua and M. Worek, *Constraining BSM physics at the LHC: Four top final states with NLO accuracy in perturbative QCD*, *JHEP* **07** (2012) 111, [arXiv:1206.3064].

- [106] Z. Nagy, *Three jet cross-sections in hadron hadron collisions at next-to-leading order*, *Phys. Rev. Lett.* **88** (2002) 122003, [hep-ph/0110315].
- [107] J. Alwall, P. Schuster, and N. Toro, *Simplified models for a first characterization of new physics at the LHC*, *Phys. Rev.* **D79** (2009) 075020, [arXiv:0810.3921].
- [108] LHC New Physics Working Group, *Simplified models for LHC new physics searches*, *J. Phys.* **G39** (2012) 105005, [arXiv:1105.2838].
- [109] L. Lyons, D. Gibaut, and P. Clifford, *How to combine correlated estimates of a single physical quantity*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **270** (1988), no. 1 110 – 117.
- [110] CMS Collaboration, *Determination of jet energy calibration and transverse momentum resolution in CMS*, *JINST* **6** (2011) P11002, [arXiv:1107.4277].
- [111] CMS Collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** (2017), no. 02 P02014, [arXiv:1607.0366].
- [112] CMS Collaboration, *CMS luminosity measurements for the 2016 data taking period*, CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [113] A. L. Read, *Presentation of search results: the CL_s technique*, *J. Phys. G* **28** (2002) 2693.
- [114] ATLAS and CMS Collaborations, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.
- [115] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl. Instrum. Meth. A* **434** (1999) 435, [hep-ex/9902006].
- [116] J. Neyman and E. S. Pearson, *On the problem of the most efficient tests of statistical hypotheses*, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **231** (1933), no. 694-706 289–337, [<http://rsta.royalsocietypublishing.org/content/231/694-706/289.full.pdf>].
- [117] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554. [Erratum: DOI10.1140/epjc/s10052-013-2501-z].
- [118] ATLAS Collaboration, *Search for new phenomena in a lepton plus high jet multiplicity final state with the ATLAS experiment using $\sqrt{s} = 13$ TeV proton-proton collision data*, *JHEP* **09** (2017) 088, [arXiv:1704.0849].

- [119] ATLAS Collaboration, *Search for R-parity-violating supersymmetric particles in multi-jet final states produced in p-p collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Submitted to: *Phys. Lett. B* (2018) [arXiv:1804.0356].
- [120] CMS Collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** (2017), no. 02 P02014, [arXiv:1607.0366].
- [121] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev.* **D96** (2017), no. 7 072002, [arXiv:1703.0966].
- [122] J. Cogan, M. Kagan, E. Strauss, and A. Schwartzman, *Jet-images: Computer vision inspired techniques for jet tagging*, *JHEP* **02** (2015) 118, [arXiv:1407.5675].
- [123] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, *Jet-images – deep learning edition*, *JHEP* **07** (2016) 069, [arXiv:1511.0519].
- [124] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, *Parton shower uncertainties in jet substructure analyses with deep neural networks*, *Phys. Rev.* **D95** (2017), no. 1 014018, [arXiv:1609.0060].
- [125] G. Kasieczka, T. Plehn, M. Russell, and T. Schell, *Deep-learning top taggers or the end of QCD?*, *JHEP* **05** (2017) 006, [arXiv:1701.0878].
- [126] Y.-T. Chien and R. Kunnawalkam Elayavalli, *Probing heavy ion collisions using quark and gluon jet substructure*, arXiv:1803.0358.
- [127] P. T. Komiske, E. M. Metodiev, and M. D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination*, *JHEP* **01** (2017) 110, [arXiv:1612.0155].
- [128] “CMS Open Data.” <http://opendata.cern.ch/>.