

**PAIRS OF BOOSTED HIGGS
AND WHERE TO FIND THEM**

by
D. Brehm

A dissertation submitted to The Johns Hopkins University in conformity
with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
August, 2021

© 2021 D. Brehm
All rights reserved

Abstract

A search is presented for pair production of the standard model Higgs boson using data from proton-proton collisions at a centre-of-mass energy of 13 TeV, collected by the CMS experiment at the CERN LHC in 2016–2018, and corresponding to an integrated luminosity of 137 fb^{-1} . The final state consists of two b quark-antiquark pairs. The search is conducted in the region of phase space where at least one pair is highly Lorentz-boosted and is reconstructed as a single large-area jet. Depending on the resonance mass, the other Higgs boson may be sufficiently boosted to be also reconstructed as another large-area jet or using two b -tagged jets. The results are obtained by statistically combining the analysis channels. Limits on the product of the cross sections and branching fractions are set for narrow and wide bulk gravitons and radions in warped extra-dimensional models having a mass in the range 1000–3000 GeV . Extensions of this search to include Vector Boson Fusion and substituting a Higgs boson for a light scalar Y in the final state of two b quark-antiquark pairs are also discussed. New limits are set on the quartic HHVV coupling constant. Expected limits on the mass of the light scalar and heavy resonance interpreted in the context of scalar resonances are presented. Upper limits are placed on the production cross section of the process as a function of the masses of X and Y .

Thesis Readers

Dr. Petar Maksimovic (Primary Advisor)
Professor
Department of Physics and Astronomy
Johns Hopkins University

Dr. Morris Swartz
Professor
Department of Physics and Astronomy
Johns Hopkins University

For Karen.

Wit beyond measure is man's greatest treasure.

Acknowledgements

I first need to thank family for putting up with my academic saga. To my Mom and Dad, thanks so much for your unwavering support in helping me finish. I am sure I would have quit long ago without your help. Your individual examples inspired much of who I am today. To Kristi, thanks for being willing to step in to my life, even when it was hard, and guide me with your wisdom. To my siblings, Michael, Zach, Ashley, and Caroline, I appreciate all of the teasing, without which I would not have developed the skin needed to survive this journey. I also thank you for your friendship. I appreciate everything about being in your lives and look forward to our adventures together in the coming years. To my grandparents, who gave advice and love and provided environments of peace for me to rest, thank you.

I also had the support of many JHU friends Alice Cocoros, Patrick Bryesse, Rachael Alexandroff, Kevin Nash, Dave Fehling, Keith Redwine, Erini Lambrides, Kirsten Hall, Dan Pfeffer, Tanvi Karwal (pew pew), Kim Berghaus, Charles Hussong, Nikhil Anand, Julian Munoz, Sean Cantrell, and many others who helped me along my way. I would also like to acknowledge JoVone Pender, Keno Spruill, Matt Collinson, Charles Smith, Kim B. Miller, Debbie Camper and everyone at The Well who helped me when finishing seemed so impossibly hard. I appreciate you all for your love and support.

I would like to specially recognize a few people. Raimon Jackson, thank you for choosing me as a brother and supporting me in my crazy ideas. Thank you for seeing my potential when I could not and fostering it. Thank you for supporting my family when it seemed like it would all fall apart. Matt Morris, thank you for your friendship

and for all of the joys it has brought. I appreciate all of the time we have spent together and the times you have made my life possible. I hope to be able to return the favor one day. Marc Osherson, thank you for all that you have given me. From day one, when I was touring JHU with you and the group, I knew we would be friends for life. From working out, discussing builds, drinking beer, and trying to solve the worlds problems, I have been and will remain eternally grateful for your friendship. So sayeth the wise Alaundo!

To my children, Helen and Alice, I hope you look back on my time in grad school with fond memories. I also hope you can forgive any bad temper I may or may not have developed while trying to finish this degree. I love you both very much and look forward to getting to know you as our family adventure continues.

To Karen, my partner, my love, my light, the smartest person I know, the best mother I could ask for, thank you! Thank you for putting up with late nights, with me hunkering down and forgetting to look up, with commuting an hour and a half one way for 7 years, and for thinking I knew more than you about academia. I am beyond grateful for the love you have showered me with and am excited to see where the rest of our journey together takes us. I am blessed to wake up with you everyday. Thank you!

Contents

Abstract	ii
Dedication	iv
Acknowledgements	v
Contents	vii
List of Figures	xii
Chapter 1 The Standard Model of Particle Physics	1
1.1 Standard Model Particles	2
1.1.1 Leptons	2
1.1.2 Hadrons	5
1.1.3 Bosons	6
1.2 Standard Model Forces	7
1.2.1 Electromagnetic Force	9
1.2.2 Strong Force	10
1.2.3 Weak Force	13
1.2.4 The Higgs Boson and Mass	15
1.3 Gauge Field	17
1.4 What's Left?	17
Chapter 2 Beyond the Standard Model	19

2.1	The Hierarchy Problem	19
2.2	Possible Solutions to the Hierarchy Problem	21
2.2.1	Warped Extra Dimensions	21
2.2.2	Minimal Supersymmetry	23
2.3	Predictions of WED and MSSM	24
Chapter 3	The CMS Detector at the LHC, CERN	25
3.1	The Large Hadron Collider	27
3.1.1	Proton-Proton Collisions at the LHC	29
3.2	The Compact Muon Solenoid Detector	30
3.2.1	Coordinate System	31
3.2.2	The Tracker	31
3.2.3	The Calorimeters	32
3.2.4	The Solenoid	33
3.2.5	The Muon Chambers	33
3.3	Jets	34
3.3.1	The Particle Flow Algorithm	34
3.3.2	Anti-KT Algorithm	35
3.4	The Trigger	36
3.5	Pileup	40
3.6	Samples	40
3.6.1	Data samples	40
3.6.2	MC simulation	41
3.6.3	Theory predictions	41
Chapter 4	Algorithms that Detect Interesting Higgs Boson Candidates	43
4.1	Soft-Drop Mass Algorithm	43

4.2	Deep AK8 Mass Decorrelated Tagger	44
4.2.1	Custom Neural Network	46
4.2.2	Custom Neural Network	46
4.2.3	Advantages	47
4.3	Deep Jet Tagger	47
Chapter 5	HH4b: Event Selection	49
5.1	Event Selection	51
5.1.1	Fully Merged Topology	51
5.1.2	Jet kinematics selection	51
5.1.3	H mass selection	52
5.1.4	Deep AK8 Mass De-correlated H Tagger	53
5.1.5	Boosted Event Selection	57
5.1.6	Identification of Higgs jets in boosted analysis	57
5.1.7	MC distributions before applying Selections	59
5.2	Semi-Resolved Topology	59
5.2.1	Event Selection	61
5.2.2	Identification of Higgs jets in semi-resolved analysis	62
Chapter 6	HH4b: Background Estimate and Systematic Uncertainties	65
6.1	Background Estimation	65
6.1.1	Alphabet Background Estimate	65
6.1.1.1	Fitting procedure	69
6.1.2	Closure Test in Data	70
6.1.3	Top Control Region	74
6.1.4	Signal Regions Prediction	79
6.2	Systematic Uncertainties	88
6.2.1	$t\bar{t}$ -Tagging Scale Factor Derivation	92

Chapter 7 HH4b: Results	95
7.1 Results	95
Chapter 8 Extensions of the HH4b Search	101
8.1 VBF Production of HH in the 4b Final State	101
8.2 A Massive Resonance decaying to a light scalar and a Higgs Boson in the 4b Final State	103
Conclusions	108
References	109
Chapter A	112
A.1 Statistical Tests	112
A.1.1 F-Test	118
A.2 2D Alphabet Validation Study	120
A.3 Example Combine Card	120
A.4 τ_{21} Study	121
A.5 By Year Fits	122
A.5.1 Signal Regions Prediction By Year	122
A.5.2 MC distributions before applying Selections	136
A.6 Limits by Year	143
A.7 Deep AK8 MD Hbb Tagger Working Point Study	145
A.8 Trigger Scale Factor	148
A.9 Jet Energy Resolution Template Examples	150
A.10 M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut studies	158
A.11 Wide Resonances	161

List of Figures

Figure 1-1	The lagrangian form of the the Standard Model	3
Figure 1-2	A graphical depiction of the Standard Model	4
Figure 1-3	An example Feynman Diagram	8
Figure 1-4	Electron positron annihilation.	10
Figure 1-5	Electron positron annihilation into a muon and anti-muon pair.	10
Figure 1-6	Strong force Vertices.	11
Figure 1-7	Weak force Vertices.	14
Figure 1-8	Higgs Potential where the y-axis has units of energy and the x-axis is the scalar field.	16
Figure 1-9	Higgs Boson Vertex.	16
Figure 2-1	A tree level Higgs diagram	19
Figure 2-2	A loop level Higgs diagram	20
Figure 3-1	Left: A diagram of the LHC and its rings. Right: An overhead picture of the LHC where you can see Lake Geneva on top of it.	27
Figure 3-2	The quadrupole magnetic field that allows the proton beams to be focused in both horizontal and vertical directions.	28
Figure 3-3	A detailed cross section of the CMS detector.	30
Figure 3-4	2016 2-Dimensional Trigger Efficiency Scale Factor.	38
Figure 3-5	2017 2-Dimensional Trigger Efficiency Scale Factor.	38
Figure 3-6	2018 2-Dimensional Trigger Efficiency Scale Factor.	39

Figure 4-1	A diagram of how quarks behave in the detector.	45
Figure 4-2	A diagram of the network architecture of the DeepAK8 Algorithm.	46
Figure 4-3	A diagram of the network architecture for the Mass-Decorrelated version of the DeepAK8 Algorithm.	47
Figure 4-4	A diagram of the network architecture for the Deep Jet Algorithm.	48
Figure 5-1	Cutflow Diagram for Tight Tight QCD Selection	54
Figure 5-2	Cutflow Diagram for Loose Loose QCD Selection	54
Figure 5-3	Cutflow Diagram for Tight Tight $t\bar{t}$ Selection	55
Figure 5-4	Cutflow Diagram for Loose Loose $t\bar{t}$ Selection	55
Figure 5-5	Cutflow Diagram for Tight Tight Radion 1500 GeV Selection	56
Figure 5-6	Cutflow Diagram for Loose Loose Radion 1500 GeV Selection	56
Figure 5-7	Preliminary limits used to optimize the soft drop mass window.	57
Figure 5-8	A pictorial representation of the semaphoring of Loose Loose and Tight Tight events.	58
Figure 5-9	Pre deepTagMD_HbbvsQCD selection p_T distribution of leading fat jet	59
Figure 5-10	Pre deepTagMD_HbbvsQCD selection p_T distribution of sub-leading fat jet	59
Figure 5-11	Pre deepTagMD_HbbvsQCD selection η distribution of leading fat jet	60
Figure 5-12	Pre deepTagMD_HbbvsQCD selection η distribution of sub-leading fat jet	60
Figure 5-13	Pre deepTagMD_HbbvsQCD selection $\Delta\eta$ distribution	60
Figure 5-14	Pre deepTagMD_HbbvsQCD selection m_j distribution	61
Figure 5-15	Pre deepTagMD_HbbvsQCD selection m_{jj} distribution	61

Figure 5-16	Cutflow Diagram for Semi-resolved for QCD Selection	62
Figure 5-17	Cutflow Diagram for Semi-resolved for $t\bar{t}$ Selection	63
Figure 5-18	Cutflow Diagram for Semi-resolved for Radion 1500 GeV Selection	64
Figure 6-1	Un-smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Loose Loose.	66
Figure 6-2	Smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Loose Loose.	66
Figure 6-3	Un-smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Tight Tight.	67
Figure 6-4	Smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Tight Tight.	67
Figure 6-5	Un-smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Semi-resolved.	67
Figure 6-6	Smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Semi-resolved.	68
Figure 6-7	Full Run 2 CR background fits for M_j axis.	72
Figure 6-8	Full Run 2 CR background fits for M_{jj} axis.	73
Figure 6-9	Full Run 2 Loose Loose $t\bar{t}$ Control Region fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	75
Figure 6-10	Full Run 2 Loose Loose $t\bar{t}$ Control Region fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	76
Figure 6-11	Full Run 2 Tight Tight $t\bar{t}$ Control Region fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	77
Figure 6-12	Full Run 2 Tight Tight $t\bar{t}$ Control Region fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	78
Figure 6-13	Full Run 2 Loose Loose fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	80

Figure 6-14 Full Run 2 Loose Loose fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	81
Figure 6-15 Full Run 2 Tight Tight fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	82
Figure 6-16 Full Run 2 Tight Tight fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	83
Figure 6-17 Full Run 2 2+1 fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	84
Figure 6-18 Full Run 2 2+1 fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	85
Figure 6-19 Full Run 2 Nuisance Pulls Plot.	86
Figure 6-20 Full Run 2 Impacts plot. The impact of a nuisance parameter moving in one direction or another is in some sense a convention. A nuisance parameter that pulls “negative” can always be redefined such that the pull will be positive. What one should do is to check to make sure all the nuisance parameters “pull” in the correct direction relative to their definition.	87
Figure 6-21 KDE bandwidth nominal, up, and down distributions for Tight Tight.	92
Figure 7-1 Full Run 2 Radion Limit for 1+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.	96

Figure 7-2 Full Run 2 Graviton Limit for 1+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.	96
Figure 7-3 Full Run 2 Radion Limit for 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.	97
Figure 7-4 Full Run 2 Graviton Limit for 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.	97
Figure 7-5 Upper limits at 95% confidence level on $\sigma(pp \rightarrow X)\mathcal{B}(X \rightarrow b\bar{b}b\bar{b})$ for the narrow spin-0 radion (left) and the spin-2 bulk graviton (right) models. The predicted theoretical cross sections for the narrow radion and bulk graviton are also shown.	98
Figure 7-6 Upper limits at 95% confidence level on $\sigma(pp \rightarrow X)\mathcal{B}(X \rightarrow b\bar{b}b\bar{b})$ for the 10% width spin-0 radion (left) and the spin-2 bulk graviton (right) models.	99
Figure 8-1 Feynman diagrams for Vector Boson Fusion production of HH in the 4b Final State.	101
Figure 8-2 Expected and observed 95% CL exclusion limit on the product of the HH production cross section and the branching fraction into the 4b final state, as a function of the 2V coupling, obtained using the Asimov dataset. The crossings of expected median (dashed black line) and the theoretical cross section (red line) indicate the expected range of the κ_{2V} values to be excluded at 95% CL	102
Figure 8-3 Two-dimensional likelihood scan where both κ_{2V} and κ_V couplings are varied simultaneously, for $\kappa_V < 0$ (left) and $\kappa_V > 0$ (right)	103
Figure 8-4 Feynman diagram for gluon gluon fusion production of HY in the 4b Final State.	103

Figure 8-5 The expected upper limit of $pp \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at 95% confidence level for a fixed mass of Y, $M_Y = 125GeV$. Calculated using toy data for full RunII.	105
Figure 8-6 The expected upper limit of $pp \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at 95% confidence level for seven different M_Y . Limits are calculated using toy data for full RunII	106
Figure 8-7 The expected upper limit of $pp \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at 95% confidence level in the M_X, M_Y plane. Limits are calculated using toy data for full RunII.	107
Figure A-1 Full Run 2 Goodness of Fit test plot with 500 toys.	112
Figure A-2 Full Run 2 Correlation Matrix.	113
Figure A-3 Signal Injection (Left Column is Pull and Right is Strength) for Full Run 2 at 1 TeV for r = 0, 1, 3 with 500 toys.	114
Figure A-4 Signal Injection (Left Column is Pull and Right is Strength) for Full Run 2 at 2 TeV for r = 0, 1, 3 with 500 toys.	115
Figure A-5 Signal Injection (Left Column is Pull and Right is Strength) for Full Run 2 at 3 TeV for r = 0, 1, 3 with 500 toys.	116
Figure A-6 Signal Injection for Full Run 2 at 2 TeV for r = 0 with 500 toys and a simulated higher integrated luminosity.	117
Figure A-7 F-test plots showing that the 2x2 fit function is the most optimal. When reading these plot, p1 is the base fit function and p2 is the function that is being compared as an improvement. Whether for p1 or p2, the designation pol gives the order of the polynomial fit function in M_j , M_{jj}^{red} in order. Pol11 means 1st order polynomial in M_j and 1st order polynomial in M_{jj}^{red}	119
Figure A-8 Comparison plot of various limits.	120

Figure A-9 Comparison of limits with and without τ_{21} cut. The plot on the right is with the cut and the plot on the left is without the cut.	121
Figure A-10 2016 Loose Loose background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	123
Figure A-11 2016 Tight Tight background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	124
Figure A-12 2016 2+1 background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	125
Figure A-13 2017 Loose Loose background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	126
Figure A-14 2017 Tight Tight background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	127
Figure A-15 2017 2+1 background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	128
Figure A-16 2018 Loose Loose background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	129
Figure A-17 2018 Tight Tight background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	130

Figure A-18 2018 2+1 background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	131
Figure A-19 Full Run 2 Nuissance Pulls Plot.	132
Figure A-20 Full Run 2 Impacts plot.	132
Figure A-21 2016 Loose Loose background fit for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	133
Figure A-22 2016 Tight Tight background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	133
Figure A-23 2017 Loose Loose background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	134
Figure A-24 2017 Tight Tight background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	134
Figure A-25 2018 Loose Loose background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	135
Figure A-26 2018 Tight Tight background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.	135
Figure A-27 Pre deepTagMD_HbbvsQCD selection p_T distribution of leading jet	136
Figure A-28 Pre deepTagMD_HbbvsQCD selection p_T distribution of sub-leading jet	137

Figure A-29 Pre deepTagMD_HbbvsQCD selection η distribution of leading jet	138
Figure A-30 Pre deepTagMD_HbbvsQCD selection η distribution of sub-leading jet	139
Figure A-31 Pre deepTagMD_HbbvsQCD selection $\Delta\eta$ distribution	140
Figure A-32 Pre deepTagMD_HbbvsQCD selection m_j distribution	141
Figure A-33 Pre deepTagMD_HbbvsQCD selection m_{jj} distribution	142
Figure A-34 2016,2017,2018 Radion Limits for 1+1 and 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.	143
Figure A-35 2016,2017,2018 Graviton Limits for 1+1 and 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.	144
Figure A-36 Optimization of dak8MDHbb tagger using S/\sqrt{B} for Narrow Radion and Bulk Graviton 1500 Gev.	146
Figure A-37 Optimization of dak8MDHbb tagger using S/\sqrt{B} for Narrow Radion and Bulk Graviton 1500 Gev with just $t\bar{t}$ as a background.	147
Figure A-38 Comparison of trigger scale factors for 2016.	148
Figure A-39 Comparison of trigger scale factors for 2017.	148
Figure A-40 Comparison of trigger scale factors for 2018.	149
Figure A-41 2016 Narrow Radion 1500 GeV JER pass and fail templates for M_j axis and M_{jj}^{red} axis.	150
Figure A-42 2016 $t\bar{t}$ JER pass and fail templates for M_j axis and M_{jj}^{red} axis.	151
Figure A-43 2016 Narrow Radion 1500 GeV JES pass and fail templates for M_j axis and M_{jj}^{red} axis.	152
Figure A-44 2016 $t\bar{t}$ JES pass and fail templates for M_j axis and M_{jj}^{red} axis.	153
Figure A-45 2016 Narrow Radion 1500 GeV JMR pass and fail templates for M_j axis and M_{jj}^{red} axis.	154
Figure A-46 2016 $t\bar{t}$ JMR pass and fail templates for M_j axis and M_{jj}^{red} axis.	155

Figure A-47 2016 Narrow Radion 1500 GeV JMS pass and fail templates for M_j axis and M_{jj}^{red} axis.	156
Figure A-48 2016 $t\bar{t}$ JMS pass and fail templates for M_j axis and M_{jj}^{red} axis.	157
Figure A-49 Radion Signal Efficiency for full Run 2.	158
Figure A-50 Radion Signal Efficiency for full Run 2.	158
Figure A-51 Radion Signal Efficiency for full Run 2 with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.	159
Figure A-52 Radion Signal Efficiency for full Run 2 with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.	159
Figure A-53 Full Run 2 Radion Limit comparison with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.	159
Figure A-54 Full Run 2 Graviton Limit comparison with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.	160
Figure A-55 Full Run 2 Loose Loose fits for M_j axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.	162
Figure A-56 Full Run 2 Loose Loose fits for M_{jj}^{red} axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.	163
Figure A-57 Full Run 2 Tight Tight fits for M_j axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.	164
Figure A-58 Full Run 2 Tight Tight fits for M_{jj}^{red} axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.	165

Figure A-59 Full Run 2 2+1 fits for M_j axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.	166
Figure A-60 Full Run 2 2+1 fits for M_{jj}^{red} axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.	167
Figure A-61 Full Run 2 Nuissance Pulls Plot for 2 TeV wide Radion signal.	168
Figure A-62 Full Run 2 Goodness of Fit test plot with 500 toys for 2 TeV wide Radion signal.	168
Figure A-63 Full Run 2 Correlation Matrix for 2 TeV wide Radion signal.	169
Figure A-64 Wide Signal (Left Column is Pull and Right is Strength) Injection for Full Run 2 at 1 TeV for $r = 0, 1, 3$ with 500 toys.	170
Figure A-65 Wide Signal (Left Column is Pull and Right is Strength) Injection for Full Run 2 at 2 TeV for $r = 0, 1, 3$ with 500 toys.	171
Figure A-66 Wide Signal (Left Column is Pull and Right is Strength) Injection for Full Run 2 at 3 TeV for $r = 0, 1, 3$ with 500 toys.	172

Chapter 1

The Standard Model of Particle Physics

The first recorded use of the word “elements” is by Empedocles in the 5th Century B.C. He proposed that 4 substances; earth air, fire, water, make up all of nature and can explain all the complexity of matter. This constitutes one of the first attempts at reductionist thinking in order to explain the world. Reductionist thinking has been very influential in modern physics, especially in the area of particle physics. The ability to break down what we previously thought were “elementary” particles into the constituent parts has led to many great discoveries. The electron, the first truly elementary particle to be discovered, was first hinted at in the 1869 discovery of cathode rays by Johann Wilhelm Hittorf. In 1897 , JJ. Thompson showed that cathode rays are streams of a previously unidentified negatively charged particle, which would later be named the electron. In 1911, Charles Wilson developed a cloud chamber which allowed for the first photographic evidence of electrons. Protons would be discovered in 1917 and then the neutron would be discovered in 1935. This launched a century of discovery that would culminate in 2012 with the discovery of the Higgs Boson, which was one of the last pieces of the puzzle in order to explain 3 of the 4 fundamental forces in nature. We call it the Standard Model of Particle Physics.

The Standard Model(SM) of Particle Physics is the theory that explains all the

known particles and the forces that govern their behavior, with the VERY notable exception of gravity. This is quite an astonishing statement because there are hundreds of known particles and to be able to explain them in a relatively succinct manner is quite the accomplishment of modern science. One way to represent this is a useful diagram that has become popular in explaining the SM, see Figure 1-2.

1.1 Standard Model Particles

1.1.1 Leptons

Starting with the electron, we can begin to fill out the SM with two other particles that are in some sense just heavier version, the muon (μ) and the tau (τ). They all have charge of -1^1 . One can see that the muon and the tau have masses roughly 200 and 4000 times that of the electron. There also exists a pair neutrino, denoted as ν , for each of these leptons which will complete our lepton table in the SM. The electron, muon, and tau neutrinos all have 0 charge and are supposed to be massless in the SM. You will see that on the table in Figure 1-2 each of the flavors of neutrino have a mass bound of less than some value. This is because various experiments have shown that the neutrinos have some mass. Since the mass of the neutrino is not predicted by the SM, we expect that some beyond Standard Model (BSM) explanation exists. Each of these particles has an anti-particle which is notated with a bar over the symbol. These are the \bar{e} , $\bar{\mu}$, $\bar{\tau}$, $\bar{\nu}_e$, $\bar{\nu}_\mu$, $\bar{\nu}_\tau$.

All the leptons also have other properties which are important to mention. The first is something called “spin”. This represents the intrinsic angular momentum of the particle. It is a purely quantum property so one should not think of it as a measurement of how fast the particle rotates. Rather it is one component of the total angular momentum of the particle. However, the classical analogy is useful. Just as something can spin in multiple directions classically, this quantum notion of spin has

¹The unit of 1 here represents $-1.602 \times 10^{-19} C$ oulombs

$$\begin{aligned}
& -\frac{1}{2}\partial_\nu g_\mu^a \partial_\nu g_\mu^a - g_s f^{abc} \partial_\mu g_\nu^a g_\mu^b g_\nu^c - \frac{1}{4}g_s^2 f^{abc} f^{ade} g_\mu^b g_\nu^c g_\mu^d g_\nu^e + \\
& \frac{1}{2}ig_s^2 (\bar{q}_i^\sigma \gamma^\mu q_j^\sigma) g_\mu^a + \bar{G}^a \partial^2 G^a + g_s f^{abc} \partial_\mu \bar{G}^a G^b g_\mu^c - \partial_\nu W_\mu^+ \partial_\nu W_\mu^- - \\
& M^2 W_\mu^+ W_\mu^- - \frac{1}{2} \partial_\nu Z_\mu^0 \partial_\nu Z_\mu^0 - \frac{1}{2c_w^2} M^2 Z_\mu^0 Z_\mu^0 - \frac{1}{2} \partial_\mu A_\nu \partial_\mu A_\nu - \frac{1}{2} \partial_\mu H \partial_\mu H - \\
& \frac{1}{2} m_h^2 H^2 - \partial_\mu \phi^+ \partial_\mu \phi^- - M^2 \phi^+ \phi^- - \frac{1}{2} \partial_\mu \phi^0 \partial_\mu \phi^0 - \frac{1}{2c_w^2} M \phi^0 \phi^0 - \beta_h [\frac{2M^2}{g^2} + \\
& \frac{2M}{g} H + \frac{1}{2}(H^2 + \phi^0 \phi^0 + 2\phi^+ \phi^-)] + \frac{2M^4}{g^2} \alpha_h - ig c_w [\partial_\nu Z_\mu^0 (W_\mu^+ W_\nu^- - \\
& W_\nu^+ W_\mu^-) - Z_\nu^0 (W_\mu^+ \partial_\nu W_\mu^- - W_\mu^- \partial_\nu W_\mu^+) + Z_\mu^0 (W_\nu^+ \partial_\nu W_\mu^- - \\
& W_\nu^- \partial_\nu W_\mu^+)] - igs_w [\partial_\nu A_\mu (W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) - A_\nu (W_\mu^+ \partial_\nu W_\mu^- - \\
& W_\mu^- \partial_\nu W_\mu^+) + A_\mu (W_\nu^+ \partial_\nu W_\mu^- - W_\nu^- \partial_\nu W_\mu^+)] - \frac{1}{2} g^2 W_\mu^+ W_\mu^- W_\nu^+ W_\nu^- + \\
& \frac{1}{2} g^2 W_\mu^+ W_\nu^- W_\mu^+ W_\nu^- + g^2 c_w^2 (Z_\mu^0 W_\mu^+ Z_\nu^0 W_\nu^- - Z_\mu^0 Z_\mu^0 W_\nu^+ W_\nu^-) + \\
& g^2 s_w^2 (A_\mu W_\mu^+ A_\nu W_\nu^- - A_\mu A_\nu W_\mu^+ W_\nu^-) + g^2 s_w c_w [A_\mu Z_\nu^0 (W_\mu^+ W_\nu^- - \\
& W_\nu^+ W_\mu^-) - 2A_\mu Z_\mu^0 W_\nu^+ W_\nu^-] - g\alpha [H^3 + H\phi^0 \phi^0 + 2H\phi^+ \phi^-] - \\
& \frac{1}{8} g^2 \alpha_h [H^4 + (\phi^0)^4 + 4(\phi^+ \phi^-)^2 + 4(\phi^0)^2 \phi^+ \phi^- + 4H^2 \phi^+ \phi^- + 2(\phi^0)^2 H^2] - \\
& g M W_\mu^+ W_\mu^- H - \frac{1}{2} g \frac{M}{c_w^2} Z_\mu^0 Z_\mu^0 H - \frac{1}{2} ig [W_\mu^+ (\phi^0 \partial_\mu \phi^- - \phi^- \partial_\mu \phi^0) - \\
& W_\mu^- (\phi^0 \partial_\mu \phi^+ - \phi^+ \partial_\mu \phi^0)] + \frac{1}{2} g [W_\mu^+ (H \partial_\mu \phi^- - \phi^- \partial_\mu H) - W_\mu^- (H \partial_\mu \phi^+ - \\
& \phi^+ \partial_\mu H)] + \frac{1}{2} g \frac{1}{c_w} (Z_\mu^0 (H \partial_\mu \phi^0 - \phi^0 \partial_\mu H) - ig \frac{s_w^2}{c_w} M Z_\mu^0 (W_\mu^+ \phi^- - W_\mu^- \phi^+) + \\
& igs_w M A_\mu (W_\mu^+ \phi^- - W_\mu^- \phi^+) - ig \frac{1-2c_w^2}{2c_w} Z_\mu^0 (\phi^+ \partial_\mu \phi^- - \phi^- \partial_\mu \phi^+) + \\
& igs_w A_\mu (\phi^+ \partial_\mu \phi^- - \phi^- \partial_\mu \phi^+) - \frac{1}{4} g^2 W_\mu^+ W_\mu^- [H^2 + (\phi^0)^2 + 2\phi^+ \phi^-] - \\
& \frac{1}{4} g^2 \frac{1}{c_w^2} Z_\mu^0 Z_\mu^0 [H^2 + (\phi^0)^2 + 2(2s_w^2 - 1)^2 \phi^+ \phi^-] - \frac{1}{2} g^2 \frac{s_w^2}{c_w} Z_\mu^0 \phi^0 (W_\mu^+ \phi^- + \\
& W_\mu^- \phi^+) - \frac{1}{2} ig^2 \frac{s_w^2}{c_w} Z_\mu^0 H (W_\mu^+ \phi^- - W_\mu^- \phi^+) + \frac{1}{2} g^2 s_w A_\mu \phi^0 (W_\mu^+ \phi^- + \\
& W_\mu^- \phi^+) + \frac{1}{2} ig^2 s_w A_\mu H (W_\mu^+ \phi^- - W_\mu^- \phi^+) - g^2 \frac{s_w}{c_w} (2c_w^2 - 1) Z_\mu^0 A_\mu \phi^+ \phi^- - \\
& g^1 s_w^2 A_\mu A_\mu \phi^+ \phi^- - \bar{e}^\lambda (\gamma \partial + m_e^\lambda) e^\lambda - \bar{\nu}^\lambda \gamma \partial \nu^\lambda - \bar{u}_j^\lambda (\gamma \partial + m_u^\lambda) u_j^\lambda - \\
& \bar{d}_j^\lambda (\gamma \partial + m_d^\lambda) d_j^\lambda + igs_w A_\mu [-(\bar{e}^\lambda \gamma^\mu e^\lambda) + \frac{2}{3} (\bar{u}_j^\lambda \gamma^\mu u_j^\lambda) - \frac{1}{3} (\bar{d}_j^\lambda \gamma^\mu d_j^\lambda)] + \\
& \frac{ig}{4c_w} Z_\mu^0 [(\bar{\nu}^\lambda \gamma^\mu (1 + \gamma^5) \nu^\lambda) + (\bar{e}^\lambda \gamma^\mu (4s_w^2 - 1 - \gamma^5) e^\lambda) + (\bar{u}_j^\lambda \gamma^\mu (\frac{4}{3}s_w^2 - \\
& 1 - \gamma^5) u_j^\lambda) + (\bar{d}_j^\lambda \gamma^\mu (1 - \frac{8}{3}s_w^2 - \gamma^5) d_j^\lambda)] + \frac{ig}{2\sqrt{2}} W_\mu^+ [(\bar{\nu}^\lambda \gamma^\mu (1 + \gamma^5) e^\lambda) + \\
& (\bar{u}_j^\lambda \gamma^\mu (1 + \gamma^5) C_{\lambda\kappa} d_j^\kappa)] + \frac{ig}{2\sqrt{2}} W_\mu^- [(\bar{e}^\lambda \gamma^\mu (1 + \gamma^5) \nu^\lambda) + (\bar{d}_j^\kappa C_{\lambda\kappa}^\dagger \gamma^\mu (1 + \\
& \gamma^5) u_j^\lambda)] + \frac{ig}{2\sqrt{2}} \frac{m_e^\lambda}{M} [-\phi^+ (\bar{\nu}^\lambda (1 - \gamma^5) e^\lambda) + \phi^- (\bar{e}^\lambda (1 + \gamma^5) \nu^\lambda)] - \\
& \frac{g}{2} \frac{m_e^\lambda}{M} [H (\bar{e}^\lambda e^\lambda) + i\phi^0 (\bar{e}^\lambda \gamma^5 e^\lambda)] + \frac{ig}{2M\sqrt{2}} \phi^+ [-m_d^\kappa (\bar{u}_j^\lambda C_{\lambda\kappa} (1 - \gamma^5) d_j^\kappa)] + \\
& m_u^\lambda (\bar{u}_j^\lambda C_{\lambda\kappa} (1 + \gamma^5) d_j^\kappa) + \frac{ig}{2M\sqrt{2}} \phi^- [m_d^\lambda (\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger (1 + \gamma^5) u_j^\kappa) - m_u^\kappa (\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger (1 - \\
& \gamma^5) u_j^\kappa)] - \frac{g}{2} \frac{m_e^\lambda}{M} H (\bar{u}_j^\lambda u_j^\lambda) - \frac{g}{2} \frac{m_d^\lambda}{M} H (\bar{d}_j^\lambda d_j^\lambda) + \frac{ig}{2} \frac{m_u^\lambda}{M} \phi^0 (\bar{u}_j^\lambda \gamma^5 u_j^\lambda) - \\
& \frac{ig}{2} \frac{m_d^\lambda}{M} \phi^0 (\bar{d}_j^\lambda \gamma^5 d_j^\lambda) + \bar{X}^+ (\partial^2 - M^2) X^+ + \bar{X}^- (\partial^2 - M^2) X^- + \bar{X}^0 (\partial^2 - \\
& \frac{M^2}{c_w^2}) X^0 + \bar{Y} \partial^2 Y + ig c_w W_\mu^+ (\partial_\mu \bar{X}^0 X^- - \partial_\mu \bar{X}^+ X^0) + igs_w W_\mu^+ (\partial_\mu \bar{Y} X^- - \\
& \partial_\mu \bar{X}^+ Y) + ig c_w W_\mu^- (\partial_\mu \bar{X}^- X^0 - \partial_\mu \bar{X}^0 X^+) + igs_w W_\mu^- (\partial_\mu \bar{X}^- Y - \\
& \partial_\mu \bar{Y} X^+) + ig c_w Z_\mu^0 (\partial_\mu X^+ X^- - \partial_\mu \bar{X}^- X^+) + igs_w A_\mu (\partial_\mu \bar{X}^+ X^+ - \\
& \partial_\mu \bar{X}^- X^-) - \frac{1}{2} g M [\bar{X}^+ X^+ H + \bar{X}^- X^- H + \frac{1}{c_w^2} \bar{X}^0 X^0 H] + \\
& \frac{1-2c_w^2}{2c_w} ig M [\bar{X}^+ X^0 \phi^+ - \bar{X}^- X^0 \phi^-] + \frac{1}{2c_w} ig M [\bar{X}^0 X^- \phi^+ - \bar{X}^0 X^+ \phi^-] + \\
& ig M s_w [\bar{X}^0 X^- \phi^+ - \bar{X}^0 X^+ \phi^-] + \frac{1}{2} ig M [\bar{X}^+ X^+ \phi^0 - \bar{X}^- X^- \phi^0]
\end{aligned}$$

Figure 1-1. The lagrangian form of the the Standard Model

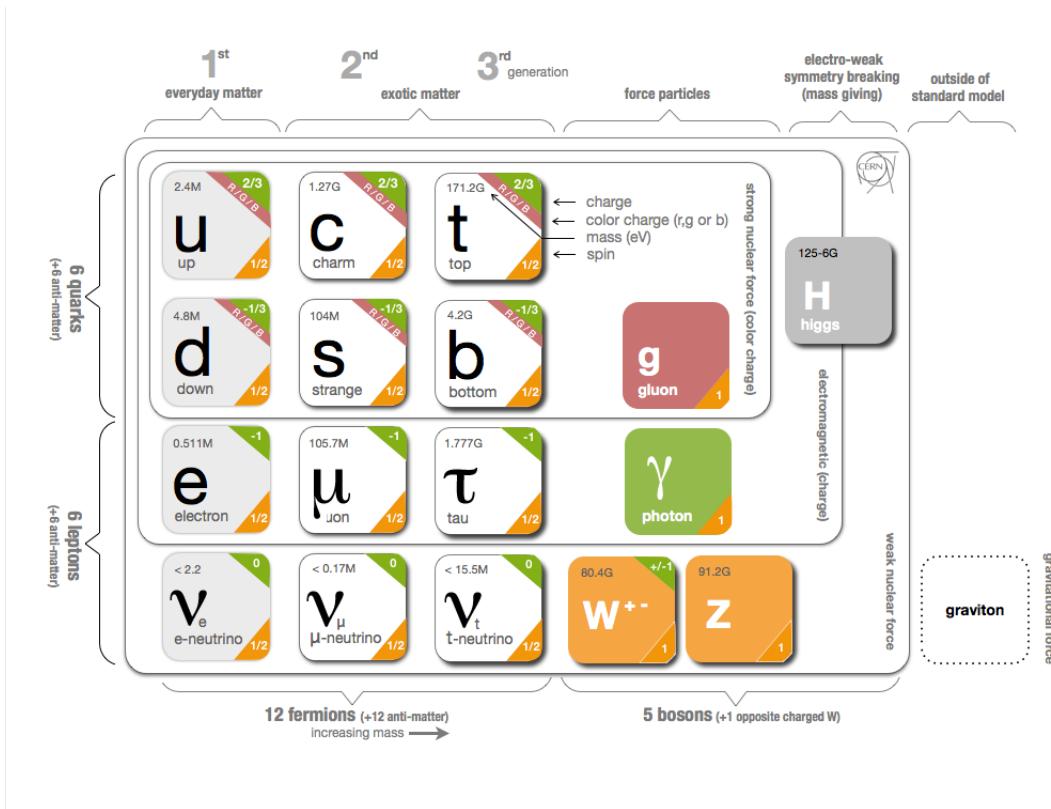


Figure 1-2. A graphical depiction of the Standard Model

two directions we denote as positive and negative. This spin can come in quantities which have whole units, 0,1,2, and in half units of, $\frac{1}{2}$. The leptons mentioned above all have spin of positive or negative $\frac{1}{2}$. Particles with integer half spin, like leptons, are called fermions.

There is another important property which we call helicity (also called handedness), defined as the sign of the particles spin vector projected onto its momentum vector. Usually, negative helicity is referred to as left-handed and positive helicity is referred to as right-handed. This property plays an important part in the calculations of the SM and therefore is not a trivial quantity. Particles with different handedness can behave differently. For example, while the massive leptons can all be either left or right handed, the neutrinos are all left handed and the anti-neutrinos are all right handed. This phenomenon currently has no explanation in the SM.

1.1.2 Hadrons

There are many more particles that have been discovered in the past 100 years. These have been meticulously detailed in a reference made by the Particle Data Group (PDG) and can be bought in a large textbook form or a small quick reference. However, the majority of these particles are composite particles, called Hadrons, that are made up of one of the six quarks seen in Figure 1-2. These quarks also come in three groupings or “generations”.² The quarks are the up (u) and down (d), strange (s) and charm (c), and bottom (b) and top (t). The u, c, and t quarks all have charge of $+\frac{2}{3}$ and the d, s, and b quarks all have charge of $-\frac{1}{3}$. They have their own masses, spins, and anti-particles (also denoted with a bar over the usual symbol), like the leptons do. They all have spin $\pm\frac{1}{2}$ so they are also fermions.

The quarks have an extra property that is unique, called color. This is another

²Why only 3? The SM theory requires 3 generations to preserve CP violation. However, this is a minimum and there is nothing preventing a 4th or 5th generation except we have not found them.

quantum property but unlike quantum spin, there is not a very good classical analogy. It can be thought of as a kind of “charge”, but it has three different types, red, green, and blue. There also exists anti-red, anti-green, and anti-blue for the anti-quarks. The color of a quark is not directly detectable and must be determined through the quarks interactions. It is also not possible for quarks to exist in anything but color neutral combinations. This will be further explained in the section on forces. These color neutral combinations make up a lot of the particles in nature. For example, the proton is made of two u and one d quark. Therefore, it has a charge of $\frac{2}{3} + \frac{2}{3} - \frac{1}{3} = +1$ because of this fact. The neutron on the other hand, is made of two d and one u quark. This combination creates a neutral electric charge. Since the quarks always come in color neutral combinations, the fractional charges of the quarks also cannot be directly detected because all of the color neutral combinations also have whole numbers of electric charge.

1.1.3 Bosons

There are 5 bosons listed in figure 1-2. They are the photon, the gluon, the W^\pm , the Z and the Higgs bosons. Unlike the fermions, which all have spin $\pm\frac{1}{2}$, the bosons all have a unit spin of 1^3 . The photon and the gluon are massless, while the W, Z, and Higgs bosons have mass. This is not an accident and will be further explained in the following section on forces. So why do we distinguish fermions from bosons? The reason is that they fundamentally behave differently. The fermions are what everyday matter is made of but the bosons are what mediate the interactions between fermions. To put it simply, the bosons act on the fermions in what we call the forces of the SM.

³There exists the theoretical possibility of spin 0 or 2 bosons but none has been detected so far

1.2 Standard Model Forces

There are three fundamental forces that the SM explains, electromagnetism, the weak force, and the strong force, with the obvious absence of gravity. While gravity was one of the earliest forces to be quantified, it has resisted our attempts at understanding it at a quantum level. To date, our inability to reconcile the current theory of gravity with the quantum nature of the SM is one of the most vexing problems facing physics.

The forces of the SM interacting with particles are mediated by “force carrying” particles. We can consider a very basic situation where two particles are launched at each other. Let us consider like charges so that the two particles will repel each other. Then we can construct an infinite sum of all the possible interactions between the two particles that start in an initial state, i , and end in some final state f as:

$$S_{i \rightarrow f} = \sum_{n=0}^{\infty} I^n \quad (1.1)$$

where each term I^n is a different possible interaction between the two particles. The first term is not interesting because it is the term that denotes the two particles not interacting. So the first interacting term is then:

$$I^2 = (\pm)^n \bar{\psi}(x) V \psi(x) \bar{\psi}(x') V \psi(x') \int \left(\frac{dk}{2\pi} \right)^4 \frac{-g_{\mu\nu}}{k^2 - i0} e^{-ik(x-x')} \quad (1.2)$$

where the terms $\psi, \bar{\psi}$ refer to the two particles initial and final states and the V term is the interaction vertex. The $\frac{g_{\mu\nu}}{(k^2 - i0)}$ term in the integral is a propagator which denotes the force interaction between the two particles (in this case the electromagnetic force is mediated with a photon). The integral is computed over all incoming and outgoing momenta, k .

This equation, while precise and useful for direct computation, is not very instructive and can be put in a pictorial representation that is much easier to read. These representations are called Feynman Diagrams. Each term in the infinite sum above

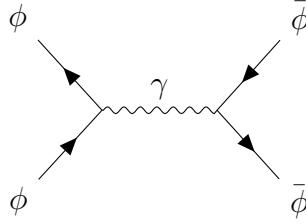


Figure 1-3. An example Feynman Diagram

can be drawn with a Feynman diagram allowing us to see all the necessary components for the interaction. More importantly, this representation does not lose any of the precision of the equations, they are essentially one and the same.

The diagram shows a virtual photon γ (wavy line) interacting with a field V (solid line). The interaction is represented by a vertex where the γ line meets the V line.

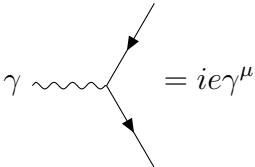
$$\gamma \text{ } \text{ } \text{ } \text{ } = V \quad (1.3)$$

Equation 1.3 is the mathematical expression for a vertex. A vertex is essentially the interaction point between two particles. One can build the diagram in Figure 1-3 by combining two vertices and correctly writing the corresponding propagator.

The simplest Feynman diagram (which is the first term in the sum) allows us to build an intuition of what is dominantly happening, which is that two particles are coming together, combining into an intermediate particles, and then that intermediate particle is decaying into two particles. This general description is actually very powerful. One can construct any number of particle combinations in such a way as they annihilate (or merge) into an intermediary particle and then decay into two (maybe even more) new particles. This ability to write down particle interactions based on a set of rules, which we will describe in following sections, is the foundation of the SM. We can classify three of the four fundamental forces of nature by constructing rules around the combination of vertex diagrams with particular intermediate particles. These intermediate particles are usually the bosons that govern a force that the particles are interacting with each other through.

1.2.1 Electromagnetic Force

Classically, Maxwell's equations do a very good job describing classical electromagnetic experiments. However, to describe phenomena involving elementary particles that are moving with relativistic energies, we need a quantum theory. The quantum theory of electromagnetism is called *quantum electrodynamics*. This theory governs all of the possible interactions between the photon, the propagating particle of the electromagnetic force, and any particle that interacts with the photon. Recall that the interaction vertex is the base of the rules for constructing Feynman diagrams and the interaction vertex for the photon is given in equation 1.4 as:



$$\gamma \sim \text{wavy line} = ie\gamma^\mu \quad (1.4)$$

The equation for the propagating photon is:

$$\gamma = \frac{g_{\mu\nu}}{(k^2 - i0)} \quad (1.5)$$

which when combined together and integrated over all momenta, can be made into an equation which is similar in form to 1.2.

A quick example of how powerful these tools can be is electron-positron annihilation. If a positron and an electron interact through a photon we will get the diagram shown in Figure 1-4. We can also use some rules from classical electromagnetism, i.e charge is conserved, to tell us if this process is possible in nature. In this example, the total charge of the incoming particles is $1 + (-1) = 0$. The photon has charge = 0 so then this will conserve charge and is possible in nature. Using charge conservation as another rule we can construct other diagrams by just switching out the electron-positron pair with another lepton, anti-lepton pair. One classic example is to use the muon and anti-muon which yields the diagram seen in Figure 1-5. QED is very well tested in modern day experiments and there are no known expected differences between it and the SM.

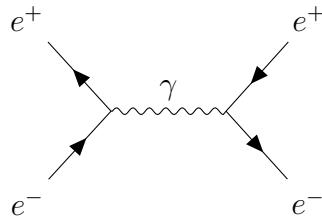


Figure 1-4. Electron positron annihilation.

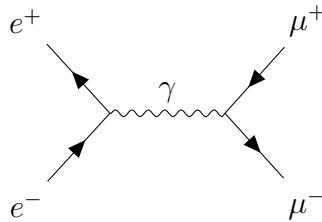


Figure 1-5. Electron positron annihilation into a muon and anti-muon pair.

1.2.2 Strong Force

The Strong Nuclear Force, or strong force for short, governs the interactions inside of hadrons like the proton and neutron. The mediator, or propagator, for the strong nuclear force is the gluon. The theory governing this force is called *Quantum Chromodynamics*. It can be said that QED is the theory of electrically charged interaction and the analogy for QCD is that it is the theory for colored interactions⁴. Unlike the photon, which since it has no charge it cannot couple to itself, the gluon has a color charge and so the interaction vertices possible to make Feynman diagrams are slightly more complicated and are shown in Figure 1-6. Just like in QED, if we apply a conservation law we can constrain what diagrams we can draw. However, unlike with QED, the color charge is not as straightforward. For the diagrams in Figure 1-6 to work, we need each gluon to carry one color and one anti color since each quark carries only a color or anti color. Naively, one might think to just combine a color-anti color pair like red anti-red, but gluons cannot carry both the color and corresponding anti color. This is due to the nature of the behavior of the strong force.

⁴Recall that I described quarks as also having a color charge.

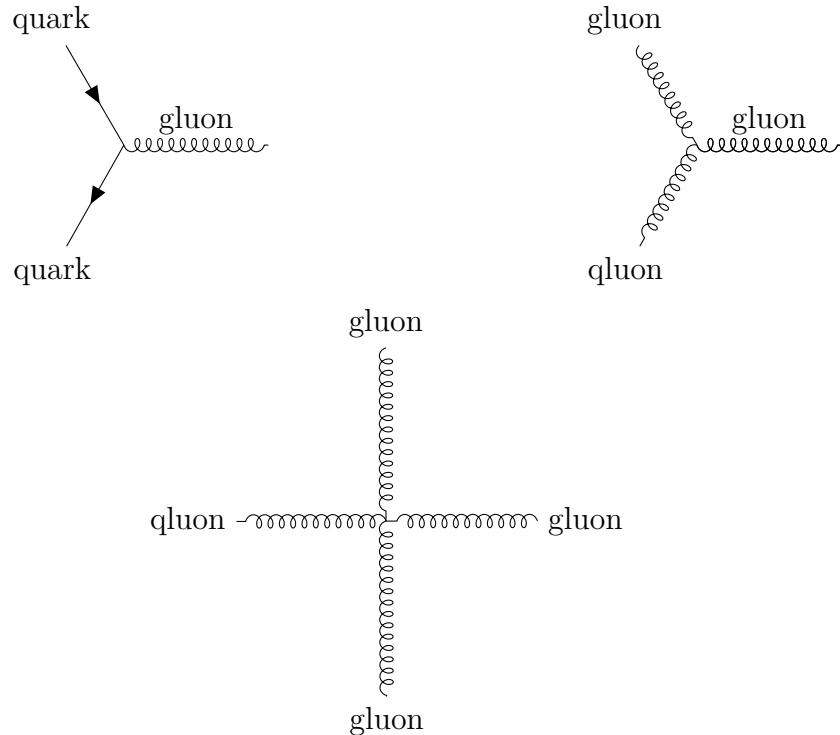


Figure 1-6. Strong force Vertices.

To give a little more context, protons and neutrons (sometimes called nucleons to refer to both of them) are confined to the nucleus by the strong force. However, they do not carry color charge in the same way that quarks do. The distinction here is that while the gluon is force carrier particle binding the quarks together inside the nucleons, it does not do that between nucleons. Two quark combinations called mesons are the mediators between nucleons. This difference is due to the emergent properties of complex systems of colored particles and is sometimes called the *residual strong force*. Since the mesons are not massless, the range for which the strong force can be felt is much smaller than say electromagnetism. In the case of these nucleons, they form what is called a color singlet. The color singlet is the superposition of quantum states as:

$$\frac{r\bar{r} + b\bar{b} + g\bar{g}}{\sqrt{3}} \quad (1.6)$$

Singlet states interact with each other. This allows us to make an empirically driven

statement that since the gluon is massless, and force mediated by the singlet state has a range limit, no gluons will carry this color singlet state. This then allows us to construct the remaining color combinations for the gluons. There are eight of them and so they are called the color *octet*. They are as follows:

$$\begin{array}{cccc} \frac{r\bar{b} + b\bar{r}}{\sqrt{2}} & \frac{r\bar{g} + g\bar{r}}{\sqrt{2}} & \frac{b\bar{g} + g\bar{b}}{\sqrt{2}} & \frac{r\bar{r} + b\bar{b}}{\sqrt{2}} \\ \frac{-i(r\bar{b} + b\bar{r})}{\sqrt{2}} & \frac{-i(r\bar{g} + g\bar{r})}{\sqrt{2}} & \frac{-i(b\bar{g} + g\bar{b})}{\sqrt{2}} & \frac{r\bar{r} + b\bar{b} - 2g\bar{g}}{\sqrt{6}} \end{array} \quad (1.7)$$

Interestingly, this is not just the only set of possible combinations, but it is also the least complex. One might ask what happens if there is sufficient energy to force something into a color singlet state? This becomes an interesting question because physics not only prevents this from happening but does so in a way that allows us to further explore the SM.

For the sake of an argument, consider an electron colliding with a proton with sufficient energy that one of the quarks in the proton was ejected. Any other parts of the proton that are ejected as well will interact through the strong force also. As the constituent quarks drift, the gluons that had bound them inside the proton form a web of self interacting connection called the color tube. These tubes exert constant force when stretched and increase in energy until they reach a characteristic size where it becomes more energetically favorable to create two new quarks out of the vacuum. This means that the lone quark that we ejected out of the proton will at the distance of about 10^{-15} meters, roughly the radius of atomic nuclei, become a new bound state meson because the lone quark cannot be in a color singlet state. If the quark still has too much energy to be bound in the meson, the process continues creating new quark anti quark pairs out of the vacuum until all free quarks are in bound states. This property is called *confinement* and the method by which the new quarks are produced in the vacuum is called *hadronization*. Hadronization in

particular is important because it creates objects called “jets” which will be crucial to our experiment.

1.2.3 Weak Force

When we say the very early universe was very hot, what we mean is that the ambient temperature was extremely hot because of how energetic the particles were during that time. During this time, the electromagnetic force was unified with what we know as the weak nuclear force into what is called the Electroweak interaction. This force was mediated by 4 massless bosons named W_1 , W_2 , W_3 , and B . Another field that matters in this story is the Higgs field which will be more completely described in the following section. At these high energies, the 4 original bosons did not interact with the Higgs field. However, as the universe cooled, this picture changed. As these original 4 bosons started interacting with the Higgs field, their interactions with the fermions changed. The fermions would no longer interact with the individual bosons but with superpositions of them. These superpositions give rise to 4 new bosons. The Z and γ bosons replaced the W_3 and B bosons. This is written as:

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & \sin(\theta_W) \\ -\sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad (1.8)$$

Here θ_W is called the Weinberg Angle or Mixing Angle and is a measurable parameter of the SM. The two remaining bosons combine to make two new bosons as:

$$W^\pm = \frac{W_1 \mp iW_2}{\sqrt{2}} \quad (1.9)$$

Now that we know where the three bosons of the weak force come from, their interaction vertices will be in Figure 1-7. These two bosons, the W^\pm and Z are not massless like the other bosons. Their masses are $80.4 \frac{\text{GeV}}{c^2}$ and $90.4 \frac{\text{GeV}}{c^2}$ ⁵ respectively. Some of the interactions include a particle interacting with its anti-particle through a “virtual”

⁵GeV stands for 10^9 eV which stands for electron volt. The unit eV/c^2 is equivalent to 1.78×10^{-36} kg. An electron has a mass of about $0.5 \text{ MeV}/c^2$ and a proton has a mass of about $1 \text{ MeV}/c^2$



Figure 1-7. Weak force Vertices.

boson. So $Z \rightarrow \mu^- \mu^+$ is a valid interaction, if the Z boson is virtual. W bosons can couple with pairs of fermions but exchanges them in terms of their flavor. For example, $W \rightarrow d\bar{u}$ is allowed. You can also have a W decay into an electron and its neutrino as $W^- \rightarrow e^- \nu_\mu$ as long as the charge has the same sign.

If you recall that the proton and neutron are just combinations of up and down quarks, you might wonder why we don't see other combinations of quarks in nature. The reason is that the W boson can interact with quarks of different generations⁶. Through the W boson, all other quarks end up decaying to the lighter first generation of quarks. This process can be expressed mathematically as the CKM Matrix. The CKM Matrix encodes the probability of a W decaying into a pair of quarks.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix} \quad (1.10)$$

where $|V_{ij}|^2$ is the probability of the i th quark decaying into the j th quark through emitting a W boson. It is through this mechanism that all the various other particles that have been detected in the last century will decay into lighter ones. For example, if there weren't for the CKM matrix, the b quark would have been stable, and we would be living in a very very different universe! However, because of the CKM matrix, b quarks can decay to c quarks, then c quarks can decay to s quarks, then to light quarks that eventually form mesons and baryons

⁶Recall Figure 1-2 and that there are 3 generations of particles.

1.2.4 The Higgs Boson and Mass

The Higgs Boson was discovered at the Large Hadron Collider at CERN in 2012. Its discovery rounded out the SM because, as we mentioned previously, the Higgs Boson mediates the Higgs field and it is a particle's interaction with the Higgs Field that gives it mass. We said in the previous section that the Higgs Field interactions did not matter at high energies and now we will detail why. If you consider the form of the Higgs potential, $\approx (\phi^2 - \eta^2)^2$, where ϕ represents the Higgs field, you can see in Figure 1-8 that at high energies there is no interaction. As cooling happens, particles now may interact with the “bump” seen in the Higgs potential. When this happens, a choice needs to be made about which valley things must settle into. This is a phenomenon called *spontaneous symmetry breaking*.

Originally, there is a symmetry of the Higgs potential about the y-axis but the choice of valley or minima “breaks” that symmetry. This does not mean that the SM is asymmetric. It just means that there are energy ranges in the SM that give rise to asymmetric like behavior. Now that a choice has been made for a minima, this symmetry breaking causes the Higgs field to take a vacuum expectation value or VEV. This can be thought of as the average value of the field in empty space. When we say it “takes” a value for the VEV, what we mean is that it deviates from 0. In the case of the Higgs field, its VEV is 246 GeV. The VEV is what is coupling to the electroweak interactions creating the photon and weak force bosons we see today. Similarly, though definitely not identically, the Higgs field couples to weak bosons, quarks, electrons, muons, and tau particles and gives them mass. We cannot say that it couples to all fermions because the neutrino ends up being the odd particle out here and does not couple to the Higgs field. In order to finish off our rules for creating Feynman diagrams we can draw the interaction vertex for the Higgs Boson in Figure 1-9. It can interact with any massive particle, including itself. The Higgs Boson has no charge, no spin and a mass of $126 \frac{\text{GeV}}{c^2}$.

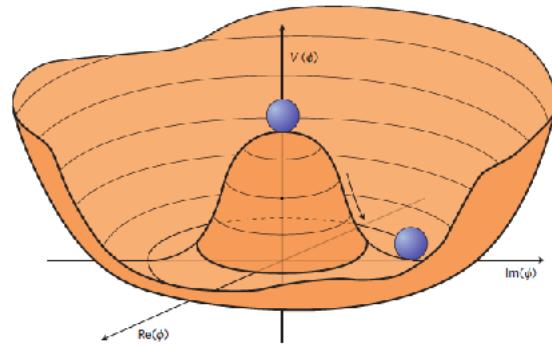


Figure 1-8. Higgs Potential where the y-axis has units of energy and the x-axis is the scalar field.

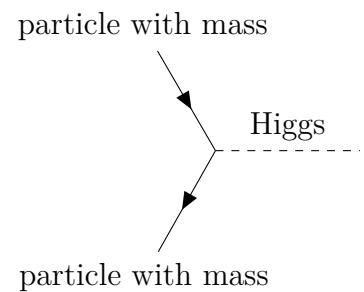


Figure 1-9. Higgs Boson Vertex.

1.3 Gauge Field

A gauge field is the mathematical mechanism behind the SM forces. A gauge theory is a type of field theory for which the Lagrangian⁷ does not change for local transformations. QCD, and QED are gauge field theories. The local symmetry that the SM is invariant under is $U(1) \times SU(2) \times SU(3)$ where U stands for unitary and S stands for special. Although the mathematics of this local symmetry are beyond the scope of this thesis, the consequences are important to note.

For example, the mathematical reason that there are eight gluons in QCD is because $SU(3)$, the symmetry of QCD, has eight generators. QED is the $U(1)$ symmetry and the weak force is under the $SU(2)$ symmetry. Therefore, the $U(1) \times SU(2) \times SU(3)$ symmetry encompasses all three fundamental forces of the SM. The combined symmetry generates 12 gauge bosons, the bosons which corresponds to the gauge field, which we now know as the photon, the 3 bosons of the weak force, and the 8 gluons.

1.4 What's Left?

Now that we have a good understanding of the SM, we can make a vast number of theoretical predictions. All such predictions have, so far, been confirmed by experiments. Unfortunately, it isn't complete. Gravity is noticeably absent. Our current understanding of gravity, through the theory of General Relativity, does not mesh nicely with the equations of the SM. Large and complicated theoretical efforts have been working very hard to try to reconcile this difference but it still remains today.

Beyond that, the visible matter in our universe only makes up about 4% of the stuff in it. Physicists have good evidence that the universe is roughly 25% matter so

⁷The lagrangian is a mathematical formalism for writing out all the interactions in a field theory through the action (the spatial integration of the lagrangian density) of the fields. The SM lagrangian is shown in Figure 1-1.

what is the other 21%? This extra stuff is theorized to be “dark matter”, i.e. not interacting in a way we can directly detect, and is not currently included in the SM. There are many theories and experiments working on this question. After that, we still don’t understand the other approximately 75% of the rest of the universe. Is it the so called “dark energy”? We know even less about this than we do about dark matter.

Another problem is that we don’t even fully understand the roughly 4% of baryonic matter we do know about. The majority of this matter is made of particles but why? Why are there not more anti particles? It is not clear why there would be an asymmetry to the amount of matter vs anti matter. Some progress has been made recently on this question but the answer is still incomplete. It was previously stated that we don’t have a good measurable of the neutrino masses. This is also a open question because it is not clear whether or not they are massless.

There are also theoretical inconsistencies in the SM. One example is that the quantum corrections to the Higgs mass are 18 order of magnitude larger than the measured value of the Higgs mass. This is called the Hierarchy problem and will be detailed more in the next chapter. There is also an unmeasured Higgs quantity, the shape of the curvature of the Higgs potential. This is the last unmeasured parameter in the SM.

So given that we seem to be missing many answers to questions about known measurable physical phenomenon, how do we explain them? Theorists spend their days creating answers for these questions in a manner that would be consistent with the current SM. These theories are known as “Beyond Standard Model” (BSM) and can address one or all of the known issues with the SM. Now we shall talk about a few BSM theories that will be relevant to this experiment.

Chapter 2

Beyond the Standard Model

In the previous chapter we discussed the major parts of the SM, detailed some of the rules for working with it, and then talked briefly about what has been left out. The SM itself actually imposes many constraints, so any BSM theory must conform to those constraints. It should also be pointed out that there have been many searches, but very little evidence, of BSM physics. This does not exclude BSM physics, it just creates another set of constraints that must be satisfied along with the constraints already set by the SM. One of the most pressing questions about the SM, and one of the most relevant for the following experiment, is called the hierarchy problem.

2.1 The Hierarchy Problem

To understand the hierarchy problem, we need to revisit the Higgs boson feynman diagrams. If you recall, we can build up diagrams by adding vertices together. This then is a valid diagram between 4 particles that is mediated by a Higgs boson:

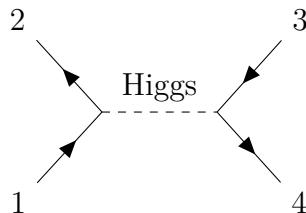


Figure 2-1. A tree level Higgs diagram

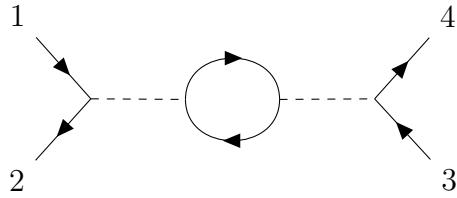


Figure 2-2. A loop level Higgs diagram

Figure 2-1 is what we call a “tree level” diagram. We can add possible interactions in the middle of the tree diagram creating what are called “loops”, like in figure 2-2.

Since all we do is measure the incoming and outgoing particles, we do not know which of the diagrams is physically happening. The SM starts with what we call a “bare” mass term, in this case m_H^{bare} , and then when adding loops¹, it makes quantum corrections to the bare mass. The Higgs boson is a scalar, which, for our purposes, means that when adding up the corrective terms, they do not cancel each other out in a well-controlled manner. Since we do measure a consistent Higgs mass, theoretically it is not satisfactory that we cannot predict this mass in the SM due to the lack of control of the quantum correction terms. The reason that this is a problem is that in the theoretical calculations of these quantum corrections, the Higgs mass is quadratically sensitive to any scale that you introduce to the theory. Energy scales are introduced to quantum corrections in order to carry out a process called renormalization. Simply stated, renormalization is the process by which we can remove any non-physical values, i.e. infinities, that show up in our calculations. This happens frequently when calculating quantum corrections so if when we try to renormalize our quantum corrections but the terms only grow in relative size, there will be a problem when we try to take a large sum of ALL possible corrections to the Higgs bare mass. In this case, the infinities that we tried to regulate will show up all over again. Again, since we can measure the Higgs mass, and it is definitely NOT infinity, we need to find a way to control these quantum corrections, which is where BSM physics comes

¹This process of adding loops can create some wacky looking diagrams. However, even wacky diagrams can be experimentally relevant.

in. The mass we measure is:

$$m_H = m_H^{bare} \times (1 + \sum_{\text{all loops}} \text{quantum loop corrections}) \quad (2.1)$$

2.2 Possible Solutions to the Hierarchy Problem

There are several theoretical models that propose solutions that solve the hierarchy problem. We will not detail all of these proposals, as that would take a thesis of its own. Here we will give an overview of the two solutions that are most relevant to our experiment. They are Warped Extra Dimensions (WED) and Minimal Supersymmetric extensions to the SM (MSSM).

2.2.1 Warped Extra Dimensions

In classical physics, the notions of space and time are fixed to 4 dimensions, 3 space and 1 time. In quantum physics, there is no a priori reason we cannot consider additions to the amount of spacetime dimensions. After Einstein's newly introduced theory of 4-D spacetime, the General Theory of Relativity, researchers started asking if it was possible that what we see as 4-D spacetime is really 5-D, where the 5th dimension is compactified or warped in some way as to avoid easy detection. The first of these theories is called Kaluza-Klein Theory and was originally published by Kaluza in 1921. This theory was originally unused as it did not offer many testable predictions. However, there have been many improvements in recent decades.

One of the recent improvements, indeed the one that is most experimentally relevant to us, is WED. The WED models have an extra spatial dimension compactified between two branes, with the region between (called the bulk) warped via an exponential metric κl , κ being the warp factor and l the coordinate of the extra spatial dimension [1]. For our purposes, we can think of branes as the “boundaries” to the spacetime in which we live. They are lower dimensional spaces that are dynamic and can effect the

fields that propagate the forces we see in nature. In these WED models, we can think of the “low energy” versions of our theories, i.e. the SM, as “living” on one brane. Here living on a brane means that the fields in the field theory do not propagate into all of the available spatial dimensions, in this case the bulk. Then the “high energy” versions will live on the other and the bulk will mediate between them. The original version of these theories had only gravity propagating in the bulk, but that it is not the only thing that may propagate in the bulk of a WED theory ². In the literature, the “low energy” brane is called the infrared (IR) brane and the “high energy” brane is called the ultraviolet (UV) brane. These two scales, UV and IR, are very important for theoretical discussions of the behavior of SM theories.

In WED models, there can exist excitations, which is the mathematical way we think of particles in their respective fields, that are of the type described in Kaluza-Klein theory, so called KK excitations, and which propagate in the bulk. The prediction from WED is that these excitations will be spin-2 bosons called KK gravitons. These would mediate gravity in the bulk and be a way to incorporate gravity into the SM. They also predict spin-0 particles, called radions, which are scalar versions of gravitons. These KK gravitons and radions are predicted to interact with the weak force which in turn allows them to interact with the SM Higgs boson and give them mass.

The radion serves another purpose. One question you might have thought to ask is, what is the size of the extra dimension that you keep talking about? If it is so small that we do not detect it, then how small is small enough to be undetectable? This is not currently known, however, the radion is produced from spontaneous symmetry breaking and therefore, takes a VEV. This radion VEV then sets the scale of the size of the extra dimension or bulk that is between the UV and IR branes.

²Other particles, namely SM fermions and bosons can propagate in the bulk.

2.2.2 Minimal Supersymmetry

Another way we can attempt to solve the hierarchy problem is to introduce a new class of theories called Supersymmetry. In supersymmetric theories, each particle has a superpartner particle that is different from the anti-particle. These extra particles would add cancelling terms to the equation in 2.1 which would allow the quantum correction terms to be controlled. There are many supersymmetric models, each with small differences to account for the issues seen in the SM. We will not be able to cover all of the known supersymmetry models and will focus on the MSSM models.

MSSM models make what is called the “minimal” extension to the standard model. These minimal extensions take each SM fermion and add a superpartner, known as a sfermion (squarks or sleptons). They also take each gauge field and add a gaugino, which is a propagating fermion for the field. While this sounds overly simplistic, it adds its own version of complications after solving the hierarchy problem. For example, if these so called superpartners are just different version of their respective particles, and with the same mass, then we should have discovered them long ago. Since we haven’t, we must assume that there is a mass scale above which all of these superpartners live.

Again, there are several ways to accomplish this addition of mass. However, they all involve a form of spontaneous symmetry breaking of the supersymmetry that generates all of these new particles. This symmetry breaking is the underlying cause of the masses that these superpartners have. This is also very similar to the case of the electroweak interaction. In both of these cases, in the UV limit of these theories, the symmetry is preserved. As soon as we move from the UV to the IR, the symmetry is broken.

2.3 Predictions of WED and MSSM

In the standard model (SM), the pair production of Higgs bosons (H) [2–4] in proton-proton (pp) collisions at $\sqrt{s} = 13\text{ TeV}$ is a rare process [5]. However, the existence of massive resonances decaying to Higgs boson pairs (HH) in many new physics models may enhance this rate to observable levels, even with current experimental data. For instance, WED models [6] contain new particles such as the spin-0 radion [7–9] and the spin-2 first KK excitation of the graviton [10–12], which have sizable branching fractions to HH .

In WED models, the reduced Planck scale ($\overline{M_{\text{Pl}}} \equiv M_{\text{Pl}}/8\pi$, M_{Pl} being the Planck scale) is considered a fundamental scale. The free parameters of the model are $\kappa/\overline{M_{\text{Pl}}}$ and the UV cutoff of the theory $\Lambda_R \equiv \sqrt{6}e^{-\kappa l}\overline{M_{\text{Pl}}}$ [7]. In pp collisions, the graviton and the radion are produced primarily through gluon-gluon fusion and are predicted to decay to HH [13].

Other scenarios, such as the two-Higgs doublet models [14] (in particular, the minimal supersymmetric model [15]) and the Georgi-Machacek model [16] predict spin-0 resonances that are produced primarily through gluon-gluon fusion, and decay to an HH pair. These particles have the same Lorentz structure and effective couplings to the gluons and, for narrow widths, result in the same kinematic distributions as those for the bulk radion. We will focus on the decays of $H \rightarrow bb$, although other decays are also possible, but this one has the largest BR (58%) We need to create BSM particle X that decays to HH pair, so we need a high energy collider (the higher the energy, the better). We als need to reconstruct each H->bb, so we need a detector. In the next chapter we will discuss the collider and detector that we use.

Chapter 3

The CMS Detector at the LHC, CERN

The way to probe the mysteries of particle physics has, in principle, not changed much since it started. The basic premise is what has been called the “reductionist” approach. This approach tries to reduce everything to its most fundamental parts. It has been very successful in the field of particle physics. Protons were once thought of as fundamental particles. However, scientists were able to discover they are composite by continually colliding them together at higher energies.

The act of smashing particles together is the mechanism of particle physics. One needs to decide at what energies they will be smashed together, how the particles will be controlled, whether or not one is fixed in place initially, and a whole host of other factors. These will effect what particles you can smash and what new particles you might expect to see from this collision.

Nature itself actually provides a useful laboratory to start this type of research program. One of the early particles that started helping physicists start filling out the SM was the muon (μ). Muons are produced in nature when a positron is accelerated by the sun’s atmosphere and launched at earth. The positron then collides with one of the electrons in an atom. These two would then annihilate and become a very energetic photon. This photon will have so much energy that it is very likely to decay

into something heavier than the electron/positron pair, a pair muons for example. Through this mechanism, muons actually are constantly being made in the upper atmosphere and shooting down to earth. We are able to detect them because they travel close enough to the speed of light that time dilation causes them to last long enough before they decay. In theory, other heavy particles could be produced this way but they do not last long enough in order for them to reach the surface ¹.

Since we cannot rely on nature to make experimental labs for us, we need to make them ourselves. The first thing to consider is how to construct the collider. In linear colliders, one side fires a beam of particles at a beam of particles fired from the other side. The downside here is that the majority of the particles will pass by each other and not collide. The next thing we can do to alleviate this is to turn to circular colliders. In circular colliders, if some particles are not used in the initial collision, they can be recycled until they are used. This efficiency comes at a cost because particles radiate when being turned by an electromagnetic field. The equation for this is

$$P = \frac{e^4}{6\pi m^4 c^5} E^2 B^2 \quad (3.1)$$

where E and B are the strengths of the respective electric and magnetic fields used to turn and focus the beams of particles. Since the power here is inversely proportional to mass, it is actually better to collide heavier particles rather than lighter ones ². This is the idea behind the Large Hadron Collider. By balancing the energy needed to collide a heavier particle with the energy saved by colliding a heavier particle, it was determined that colliding protons together would allow physicists to reach energies needed to probe the scale of physics that CERN was hoping to achieve.

¹One can even turn a modern cell phone into a muon detector. There are apps that can register the passing of a muon through the CMOS sensor in your phone's camera.

²For example, if we collide protons instead of electrons, the power produced by radiation is reduced by 10^{13}

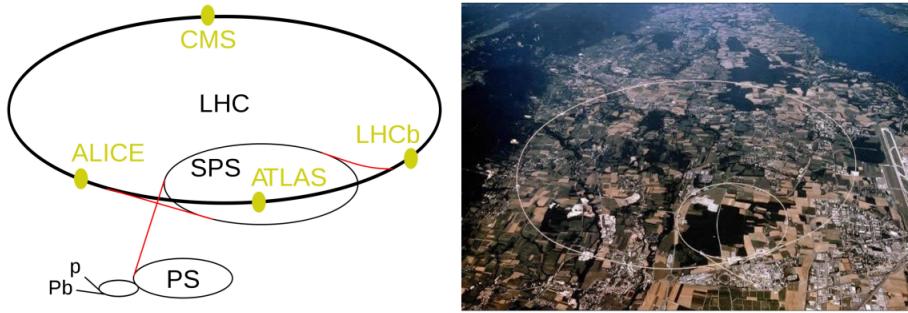


Figure 3-1. Left: A diagram of the LHC and its rings.
Right: An overhead picture of the LHC where you can see Lake Geneva on top of it.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the biggest particle collider, indeed the biggest machine, ever built. It has produced more data than all other particle experiments combined ³. It is a 27 km ring with several smaller rings which feed the proton beams into the LHC in stages. The main ring has 1,232 dipole magnets that keep the protons in the ring and accelerate them. It also has 392 quadrupole magnets that focus the beams of protons. The magnetic fields needed to achieve this are some of the strongest ever made coming in at 7 Tesla. These can only be made by superconducting magnets that are cooled to below 2 Kelvin. This amount of cooling requires around 100 tonnes of superfluid liquid helium to achieve. The LHC currently collides protons at a center of mass energy of 13 TeV but will be upgraded to achieve higher energies in the future. It should be noted that the beams of protons are not continuous. They come in “bunches” which consist of about 115 billion protons in each bunch. The bunches collide at 4 different points, corresponding to the 4 experiments that the LHC powers, and collide every 25 nanoseconds. This means there are roughly 40 Million collisions every second.

³The LHC has produced over 130 PB of data so far.

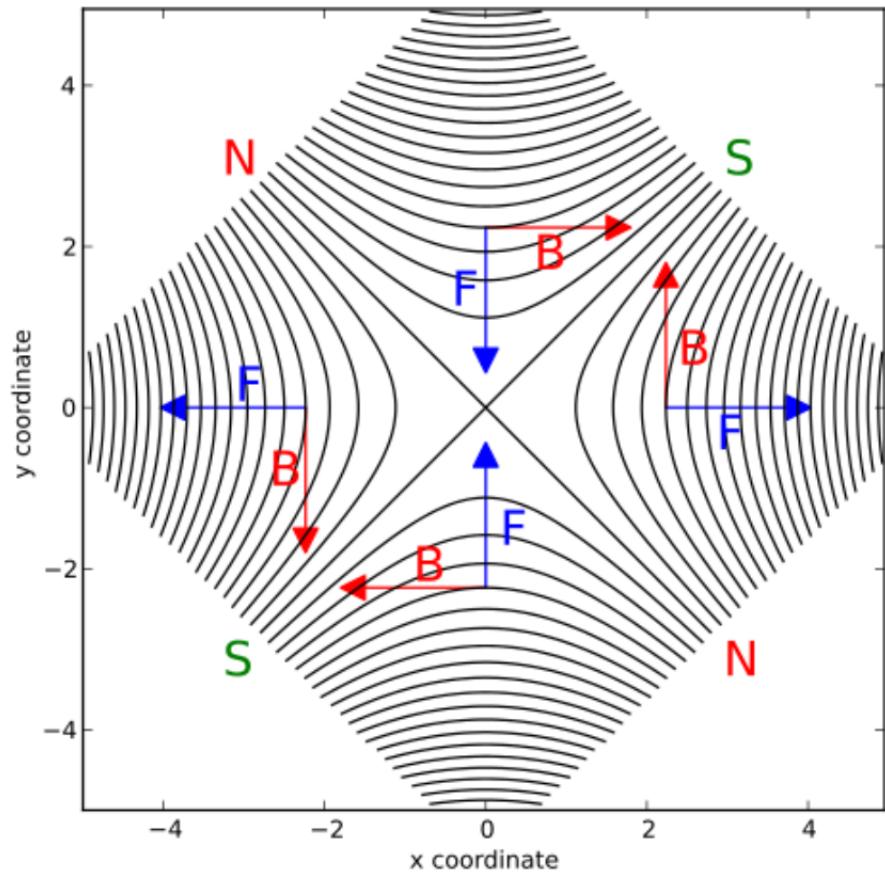


Figure 3-2. The quadrupole magnetic field that allows the proton beams to be focused in both horizontal and vertical directions.

3.1.1 Proton-Proton Collisions at the LHC

I previously had said that the proton was made up of a combination of up and down quarks. This is actually an oversimplification of the inside of a proton. In order for us to understand why one would want to smash protons together we need to get into some of the details about the proton. While it is true that the proton has two up and one down quark, the interactions between those quarks are dynamic and make up the rest of the picture of the inside of the proton. There is not actually one gluon per pair of quarks, as one might have thought, but a whole web of gluons that interact with each other and the up and down quarks. Gluons interacting with each other creates pairs of quarks of all types that then decay back into gluons. This is a very important distinction because it powers the modern program of particle physics. When two protons collide, at the energies of the LHC, we are actually smashing together all of the quarks and gluons currently present in those protons. In this “soup” of interactions, we collectively call the gluons and quarks “partons”.

The next important question is to ask what was actually collided in a given proton-proton collision. As previously stated, it is a complicated interaction of quarks and gluons that is not predetermined. In order to probe this complicated set of interactions, the Parton Distribution Function⁴ must be computed. Instead of trying to find out collision by collision what has happened, the PDF is computed to create a statistical model that can provide an expectation. The PDF models are very important for discovering new physics because many predicted new particles happen under only specific interactions. Therefore, if without knowing what to expect in a given proton-proton collision at a given energy, then one cannot know where to look for any undiscovered particle. Unfortunately, PDFs are very hard to compute. They are approximated by combining partial combinations with high precision measurements from fixed target colliders in order to approach the true PDF. Since it is only an

⁴This is usually abbreviated as PDF.

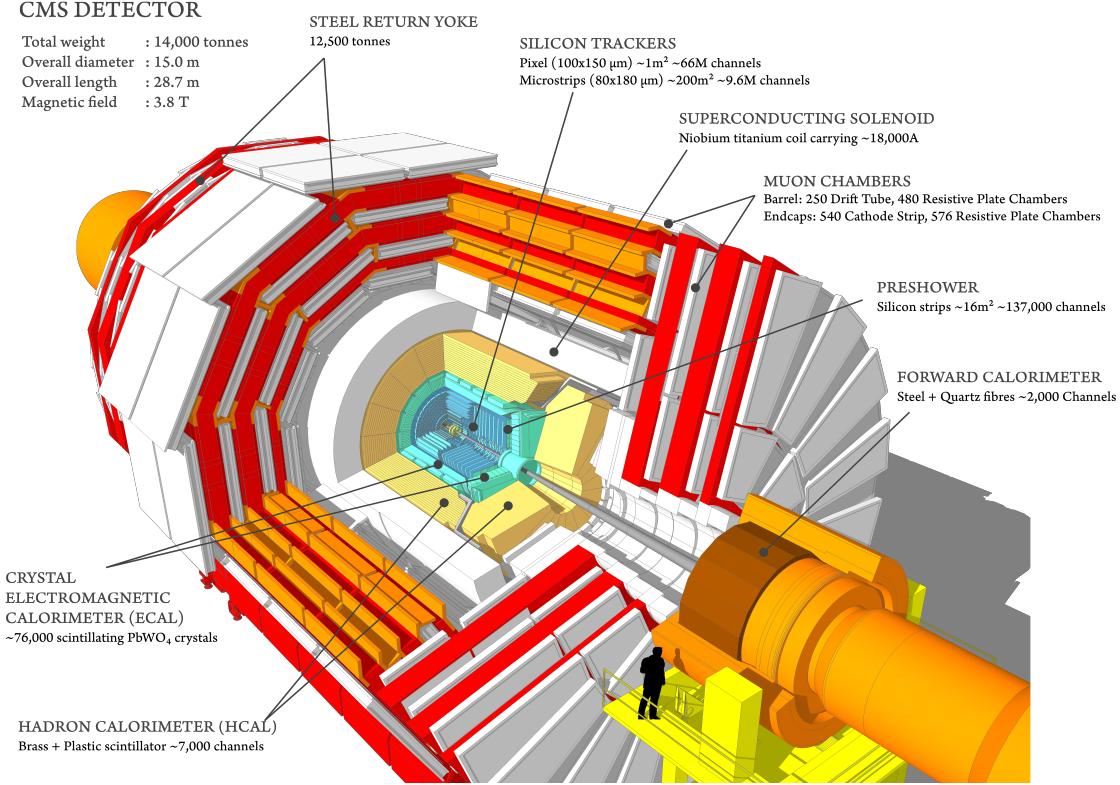


Figure 3-3. A detailed cross section of the CMS detector.

approximation, one needs to assign a systematic error to the PDF based on the variance of likely PDFs. This will be a theme revisited later in the thesis. In order to make a measurement, it will need to be approximated, and therefore it will be necessary to assign systematic uncertainty.

3.2 The Compact Muon Solenoid Detector

There are 4 detector experiments at the LHC. The Compact Muon Solenoid (CMS), shown in Figure 3-3, is the detector that took the data we are using to perform this analysis. While it is quite massive compared to a human, weighing in at 14,000 tons, it is actually smaller than the other detectors. Once again, a full description of CMS and the work done by over 5000 scientists would take many more pages to write so we will confine our discussion to a small overview just to make sure the main parts of the detector are understood.

3.2.1 Coordinate System

CMS is essentially a large tube where both ends are covered. In order to describe this geometry, we use a modified spherical coordinate system. The angles used are η and ϕ where ϕ creates the circular component of the cylinder, and will always be perpendicular to the beam. η is defined as:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \quad (3.2)$$

and is also known as the *pseudorapidity*. Here θ is the usual angle from spherical coordinates. In these coordinates, $\eta = 0$ points straight out of the detector and $\eta = \infty$ points directly along the beampipe. One reason to use these coordinates is that η is approximately Lorentz invariant⁵. This is necessary because while the protons themselves will have the same energy and likely be symmetric, there is no such guarantee with the partons. When we are talking about particles in the detector, we usually talk about them in terms of three values: radial angle ϕ , pseudorapidity η , and transverse momentum p_T . If one is interested in returning to Cartesian coordinates, which we are almost never going to do, the following equations will accomplish this task:

$$p_x = p_T \cos(\theta) \quad p_y = p_T \sin(\theta) \quad p_z = p_T \cosh(\eta) \quad (3.3)$$

3.2.2 The Tracker

The first layer that a particle shot out of a proton-proton collision would encounter is called the Silicon Tracker. This layer is directly on the beampipe and consists of 13 individual layers surrounding the beampipe and 14 layers in the endcaps. These individual layers are bunched together and the first 4 are made up of 66 million silicon pixels that are $100 \times 150 \mu m$ in area. The rest of the layers are made of strip pixels

⁵This means that it does not change depending on the reference frame in which we are looking at the interaction. It is always good to work with Lorentz invariant quantities in physics.

which are longer than the first set. All of this makes for 200 square meters of silicon that will measure the momentum of any charged particle moving through it.

A magnetic field is applied throughout the silicon tracker which causes any charged particle to curve. We can then measure the momentum because it will deposit charge on the way through each layer of the silicon tracker. With these “hits”, we can reconstruct the path and therefore the curvature which will allow us to calculate its charge and energy. The reconstruction of the hits in the silicon tracker of individual particles are called tracks and are determined to an accuracy of about $10 \mu m$. The silicon tracker, unlike the rest of the detector, does not try to stop any particle, just measure it as it is shooting towards the calorimeters

3.2.3 The Calorimeters

Recall that different types of particles will interact in different ways, most notably, some particles are much heavier than others. The first layer outside of the tracker is called the Electromagnetic Calorimeter (ECAL) and is designed to allow photons and electrons to deposit their energy in this layer. Heavier particles, namely the hadrons, will move through this part of the detector and not deposit any energy. The Hadron Calorimeter (HCAL) is the next layer and is designed to stop the heavier particles.

The ECAL is composed of around 80,000 lead-tungstate ($PbWO_4$) crystals that are a type of scintillator. Scintillators emit light when a particle deposits energy into it. So the ECAL will scintillate when light charged particles impact with the crystals. Since photons have no mass, you might wonder how they interact. The photon can either produce a pair of electrons that will interact with the scintillator or interact with an electron in the crystal itself. Since these photons and electrons are typically very energetic, they can interact multiple times and create cascades of light or “showers” as byproducts of the initial interaction. The light that is produced is directly proportional to the energy of the initial particle and therefore gives us the

missing piece of the puzzle to identify photons and electrons in the detector.

The HCAL is setup a little differently. Instead of crystal scintillators, it has plastic ones that are layered in between brass plates. The heavy particle will be stopped by the brass plate. This will cause a shower of secondary particles that are measured by the scintillators. Knowing what kinds of particles can decay to what final states, through the careful study of the rules of the SM, it is then possible to reconstruct what particle hit the HCAL. It should be noted that the HCAL is the only part of the detector that can stop neutrally charged particles, like neutrons and some mesons, so it is very important for the CMS research program.

3.2.4 The Solenoid

All of the previous layers are contained in a large cylindrical electromagnet called the solenoid. This electromagnet provides a very large magnetic field (3.8 Tesla ⁶) which causes electrically charged particles to bend due to the lorentz force. This effect is dependant on the energy of the particle so the information can be combined with the silicon tracker to measure the momentum of the particle. This magnetic field is strong enough to shift the alignment of the whole detector and this effect has to be accounted for. The last layer before the solenoid, the HCAL, uses brass specifically because it is non magnetic.

3.2.5 The Muon Chambers

After the solenoid there is one final set of detectors. The issue these solve is that muons⁷ are too heavy to be stopped by the ECAL and not heavy enough to be stopped by the HCAL. Muons will go through both and then be stopped by the aptly named Muon Chambers. There are three kinds of muon chambers; drift tubes, cathode strip chambers, and resistive plate chambers, all of which work under the principle that

⁶A normal bar magnet is measured in millitesla!

⁷Muons are always problematic!

as a muons traverses them, an electron is knocked off of gas atoms. The amount of electrons that a muon displaces is proportional to its energy and so we can measure the energy of muons with the muons chambers. Another important thing to note is that, due to the many layers of this detector and the specialities of each type, the muons chamber system is very good at reducing and filtering background noise.

3.3 Jets

In describing the strong force, I mentioned that colored particles cannot exist outside of a color singlet state, meaning they must be in pairs. In proton-proton collisions, bare quarks (quarks NOT in a color singlet state) can be produced. As soon as they are produced they begin the process of hadronization: creating new particles out of the vacuum until no bare quarks remain. Gluons may also be created in proton-proton collisions but they will decay to quarks which must then undergo hadronization. Due to hadronization, the LHC is not able to see individual quarks and gluons. All that can be seen is the shower of hadronized quarks in the direction that the quark or gluon was moving. The showers are called “Jets”. The jets are composed of constituent particles, which may in turn decay and leave traces in parts of the detector. These secondary particles are also considered constituents of the jet. The only particles that are not able to be counted as constituents are neutrinos as their energy is lost in the detector.

3.3.1 The Particle Flow Algorithm

All of the information from the subdetectors is analyzed by the *Particle Flow Algorithm*. This algorithm allows us to reconstruct jets with a high degree of precision. It starts in the silicon tracker. This subdetector is the crucial part of the algorithm because it makes the initial measurement. If it misses a charged particle, it will bias any reconstruction of that particle. Accordingly, great care has been taken to

achieve high efficiencies. An iterative tracking strategy allows for extremely high efficiencies in the first pass with softer acceptance for follow up iterations. Next, a clustering algorithm is used in the calorimeters to detect and measure the energy of stable neutral particles, separate neutral from charged particles, reconstruct charged electromagnetically charged particles, and aid in the energy measurements for low-quality tracks. Finally, a link algorithm links the track detected in the silicon tracker with the appropriate clustering in the calorimeter if the extrapolation of the track fits within a given cluster’s boundaries. This is a quick overview of the algorithm because, as is usual in most high energy physics topics, a full description would be beyond the scope of this thesis.

3.3.2 Anti-KT Algorithm

Since most of the events at CMS will generate more than one quarks or gluon, a number of algorithms exists to correctly “cluster” the constituents into the right jet. The one we use and will describe is called the *Anti- K_t* Algorithm. The algorithm is an iterative algorithm and runs in the following way. First, every PF (Particle Flow) candidate is compared against every other candidate with a distance like parameter, d_{ij} , given as:

$$d_{ij} = \min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right) \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2} \quad (3.4)$$

where R is a predetermined distance parameter that sets the size of the jet. The two closest constituents are paired and become a new constituent. This process continues, generating new “pseudo-jets”, until the distance of the pseudo-jet from the beam is $1/p_T^2$ of the pseudo-jet. It is then removed from consideration and the process starts again with the remaining constituents until none remain.

This algorithm creates what are called “conical” jets, with smooth, rounded edges⁸. We call these jets AKR jets, where we use the actual value of R in the naming.

⁸This does not always happen, especially with jets that are very close to each other.

Accordingly, an AK8 jet is a jet constructed with the R value of 0.8. We use AK8 jets and AK4 jets in this analysis. For smaller particles, jets are useful for reconstructing kinematic properties. For the larger particles, heavy bosons and quarks, AKR jets can reconstruct the particle with a good degree of accuracy. A simplistic rule with this algorithm is the R value the larger the particle that can be reconstructed.

3.4 The Trigger

Each crossing of proton-proton bunches generates about 1 Megabyte of data. At the rate of 1 crossing every 25 nanoseconds, or 40 Megahertz, no modern computing system can actually keep up with this rate of data generation. This means the majority of the data is thrown out. The system responsible for filtering through all of the data is called the “trigger”. The trigger operates in two stages, the Level 1 trigger and the High Level trigger.

The level 1 trigger is a hardware trigger. Output from the detector is stored in a buffer and then it is analyzed by custom circuits. These circuits look for “interesting” physics, such as especially large deposits in the calorimeters. This stage is very useful because it allows the rejection of all but about 0.1 % of all events. The rate of release to the next stage is around 50 kilohertz.

The High Level trigger takes the output from level 1 and analyzes the data further in order to find interesting events. At this level, there are many available triggers depending on what kind of event you are looking for. The selections are mostly kinematic differences. This is one of the first steps in conducting an analysis. One studies the available triggers in order to understand how your kinematic selections will effect the trigger efficiency.

For our analysis, the trigger algorithm used places requirements on the scalar sum of the jet transverse energy, H_t , jet p_T , and the jet groomed mass. We compensate

for the difference in the trigger response between data and simulation by applying a trigger scale factor, defined as the ratio of trigger efficiency in data divided by the trigger efficiency in qcd, to simulated events. The trigger efficiency is defined as the ratio of the number of events passing the combined triggers and a pre-trigger to the number that pass the pre-trigger and is parameterized as a function of our measurement variable. The trigger efficiency in simulation is modeled by weighting simulated events by this data-derived trigger efficiency. We select events from the 2016 dataset that pass the `HLT-HT650` trigger, the `HLT-PFHT800` trigger or the `HLT-PFHT900` trigger, and the `HLT-AK8PFJet360_TrimMass30` trigger. For 2017 and 2018, we select events that pass the `HLT-PFHT1050` and the `HLT-AK8PFJet400_TrimMass30` trigger are used to select events for the trigger efficiency measurement. The pre-trigger for 2016, 2017, and 2018 is the `HLT-Jet260` trigger. Trigger scale factors are measured as a function of the reduced mass, which will be further discussed in Chapter 5. After passing the trigger, the events are required to have at least one reconstructed pp collision vertex satisfying the following criteria:

- Vertex number of degrees of freedom > 4 ;
- Absolute displacement from the beamspot position along the z direction < 4 cm;
- Absolute displacement from the beamspot position along the transverse direction < 2 cm.

Trigger scale factors are shown here in Figures 3-4, 3-5, 3-6.

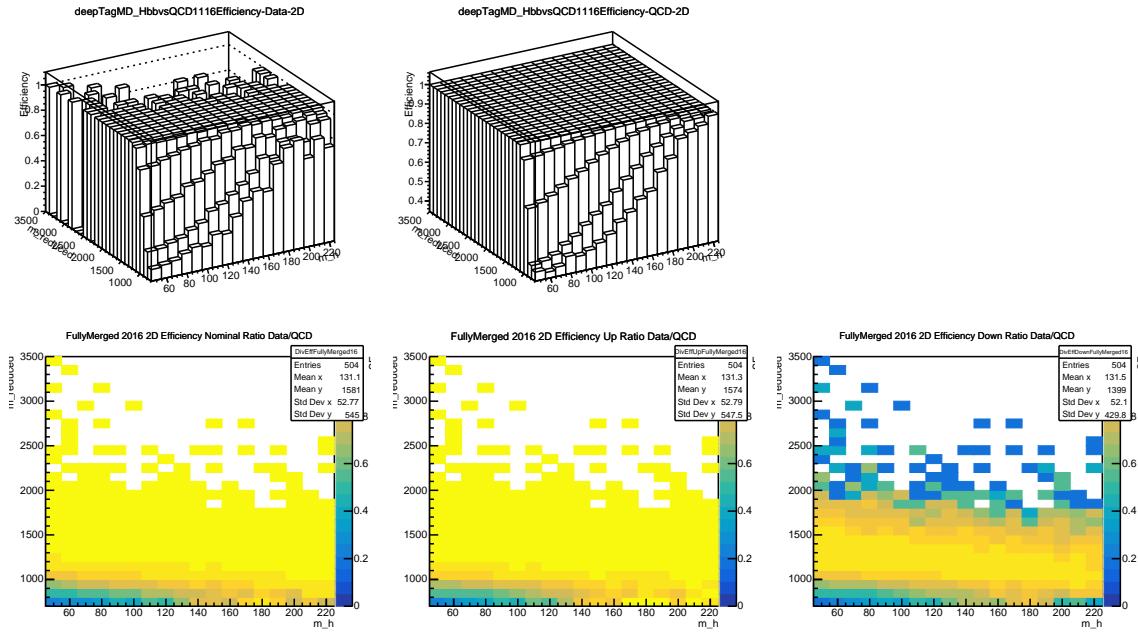


Figure 3-4. 2016 2-Dimensional Trigger Efficiency Scale Factor.

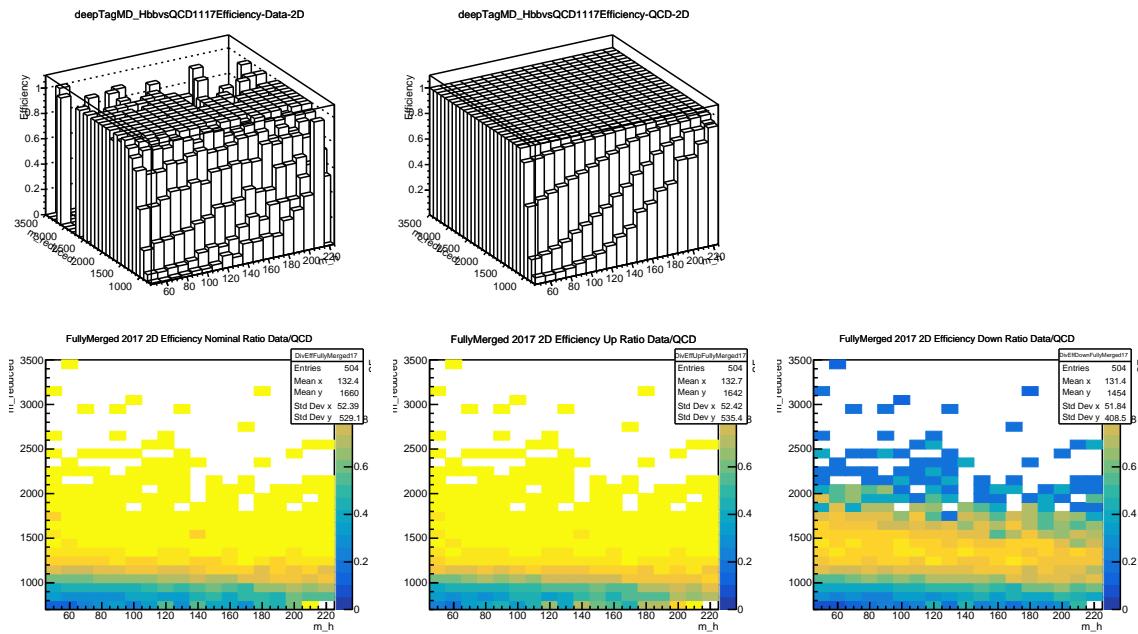


Figure 3-5. 2017 2-Dimensional Trigger Efficiency Scale Factor.

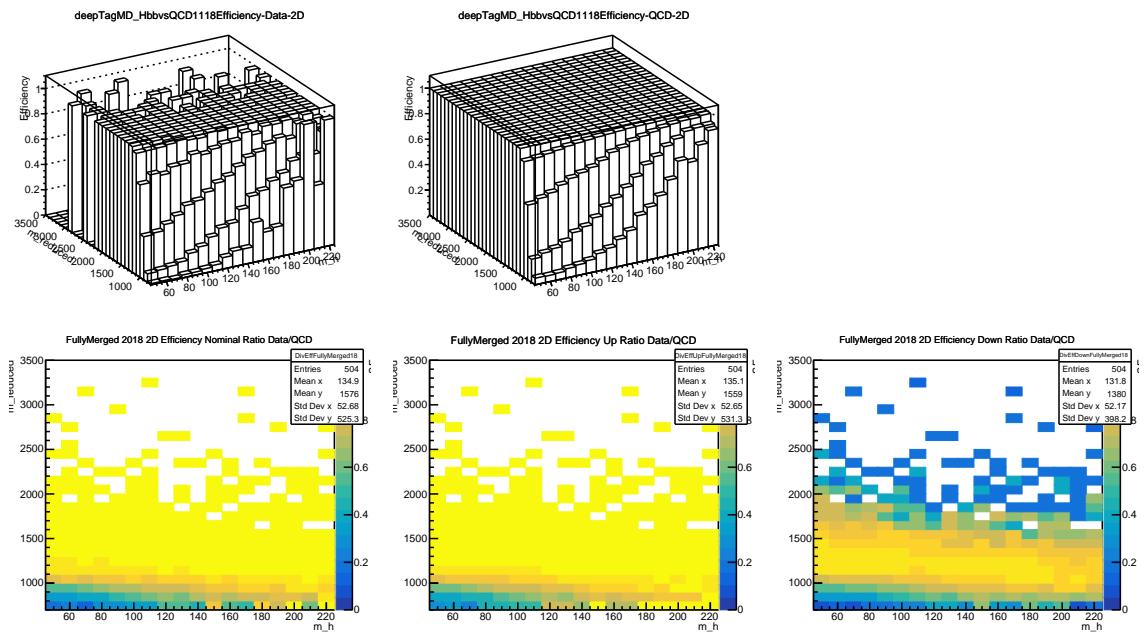


Figure 3-6. 2018 2-Dimensional Trigger Efficiency Scale Factor.

3.5 Pileup

Proton-proton collision events are ideally one proton colliding into one proton. However, since they are collided in bunches of protons, this is not the reality of what happens. Events may consist of up to 40 such proton-proton collisions. Therefore, it is necessary to define the primary vertex. A vertex is a point along the beam from which some number of particle flow candidates originate from. The primary vertex is selected as the vertex with the highest value for the sum of the square of the transverse momenta of tracks and candidates associated.

Pileup is when the jets coming from the primary vertex contain particle flow candidates that actually originate from a different vertex. This effect ends up smearing the actual measurement of jet mass and momentum and so we must take this effect into account. The choice of the AK8 and AK4 clustering algorithm is partly motivated by the fact that the Anti- k_T algorithm resists these effects better than other jet clustering algorithms.

3.6 Samples

3.6.1 Data samples

The analysis is performed using pp interactions collected with the CMS detector at $\sqrt{s} = 13$ TeV. The data correspond an integrated luminosity of 137 fb^{-1} , as measured using the golden JSON files as follows:

- 2016: [Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt](#)
- 2017: [Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON.txt](#)
- 2018: [Cert_314472-325175_13TeV_17SeptEarlyReReco2018ABC_PromptEraD_Collisions18_JSON.txt](#)

The data samples are summarized in Table 3-I. Unless otherwise noted, all data, QCD, $t\bar{t}$, and signal samples are processed from nanoAOD version 7.

Table 3-I. List of primary datasets for the $p p$ collisions at $\sqrt{s} = 13$ TeV and their corresponding integrated luminosities. The golden JSON files were used are given above.

Dataset	Processing	Int. lumi. (fb^{-1})
JetHT/Run2016B2	01June2019	5.746
JetHT/Run2016C	01June2019	2.573
JetHT/Run2016D	01June2019	4.242
JetHT/Run2016E	01June2019	4.025
JetHT/Run2016F	01June2019	3.105
JetHT/Run2016G	01June2019	7.576
JetHT/Run2016H	01June2019	8.651
JetHT/Run2017B	01June2019	4.794
JetHT/Run2017C	01June2019	9.631
JetHT/Run2017D	01June2019	4.248
JetHT/Run2017E	01June2019	9.314
JetHT/Run2017F	01June2019	13.53
JetHT/Run2018A	01June2019	14.03
JetHT/Run2018B	01June2019	7.061
JetHT/Run2018C1	01June2019	6.895
JetHT/Run2018D	01June2019	31.71
Total		137 fb^{-1}

3.6.2 MC simulation

The MC samples used for this analysis includes spin-2 bulk graviton and spin-0 radion resonances, as given in Table 3-III, as well as multijet and $t\bar{t}\text{MC}$, as given in Table 3-II. We use the puWeightProducer module in “auto” mode form the NanoAOD-tools to perform the calculation of the weight.

3.6.3 Theory predictions

The NLO bulk graviton and radion cross sections are provided in Table 3-III for the mass points used in the analysis. Lower mass points were considered but the analysis is not sensitive below 750 GeV, where the trigger efficiency is extremely poor. The bulk graviton production cross sections are evaluated for $k/\overline{M_{\text{Pl}}} = 0.1$ [18] then multiplied by 25 to get the values for $k/\overline{M_{\text{Pl}}} = 0.5$, and the radion cross sections for $\Lambda_R = 3 \text{ TeV}$ and $kl = 35$ [19]. The bulk graviton and radion to HH decay branching fractions are taken from Refs. [20] and [21], respectively.

Table 3-II. List of background Monte Carlo samples used. The two $t\bar{t}$ +jets samples correspond to two different productions with the same generator parameters, but the latter with a much higher statistics. The cross sections σ and number of events generated are also given. The cross sections for the QCD processes are at LO. The other SM background cross sections are taken from Ref. [17].

Background		
2016 Process	σ (pb)	size
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.831×10^3 (LO)	15,629,253
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.207×10^3 (LO)	48,50,746
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	119.9 (LO)	3,970,819
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	25.24 (LO)	1,991,645
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8_ext1-v1	6.831×10^3 (LO)	21,604,533
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8_ext1-v1	1.207×10^3 (LO)	10,360,193
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8_ext1-v1	119.9 (LO)	7,868,538
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8_ext1-v1	25.24 (LO)	3,812,534
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.76 (NNLO)	76,367,863
2017 Process	σ (pb)	size
QCD_HT700to1000_TuneCP5_13TeV-madgraphMLM-pythia8	6.831×10^3 (LO)	47,610,552
QCD_HT1000to1500_TuneCP5_13TeV-madgraphMLM-pythia8	1.207×10^3 (LO)	16,595,628
QCD_HT1500to2000_TuneCP5_13TeV-madgraphMLM-pythia8	119.9 (LO)	11,634,434
QCD_HT2000toInf_TuneCP5_13TeV-madgraphMLM-pythia8	25.24 (LO)	5,941,306
TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8	377.96 (NNLO)	130,091,218
TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8	365.34 (NNLO)	109,663,792
2018 Process	σ (pb)	size
QCD_HT700to1000_TuneCP5_13TeV-madgraphMLM-pythia8	6.831×10^3 (LO)	48,158,738
QCD_HT1000to1500_TuneCP5_13TeV-madgraphMLM-pythia8	1.207×10^3 (LO)	15,466,225
QCD_HT1500to2000_TuneCP5_13TeV-madgraphMLM-pythia8	119.9 (LO)	10,955,087
QCD_HT2000toInf_TuneCP5_13TeV-madgraphMLM-pythia8	25.24 (LO)	5,475,677
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	377.96 (NNLO)	128,640,000
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	365.34 (NNLO)	100,790,000

Table 3-III. NLO Bulk graviton and radion cross sections times the decay rate for the samples that are used to set a limit (lower mass samples were studied but the analysis is not sensitive below 750 GeV). The bulk graviton production cross sections are evaluated for $k/\overline{M_{Pl}} = 0.1$ [18] then multiplied by 25 to get the values for $k/\overline{M_{Pl}} = 0.5$, and the radion cross sections for $\Lambda_R = 3$ TeV and $kl = 35$ [19]. The bulk graviton and radion to HH decay branching fractions are taken from Refs. [20] and [21], respectively. All cross sections listed in Ref. [22].

M_X (GeV)	Bulk graviton (fb)		Radion (fb)	
	$\sigma(pp \rightarrow X)$	$\sigma(pp \rightarrow X \rightarrow HH)$	$\sigma(pp \rightarrow X)$	$\sigma(pp \rightarrow X \rightarrow HH)$
1000	141.65	13.97	261.89	62.057
1500	14.325	1.425	54.428	12.897
2000	2.265	0.22575	14.293	3.3868
2500	0.4675	0.04625	4.3017	1.0193
3000	0.11025	0.0110	1.3843	0.3280

Chapter 4

Algorithms that Detect Interesting Higgs Boson Candidates

The analysis we perform will be looking for 2 Higgs bosons which decay into 4 total b quarks. The 4b final state will decay into jets. Jets resulting from pairs of b quarks are identifiable because the parent Higgs jets are identifiable with a high degree of accuracy. The following sections will detail how this is accomplished.

4.1 Soft-Drop Mass Algorithm

As previously stated, the decay signature of the Higgs boson we are looking for is a pair of b quarks. When heavy particles, like the Higgs boson, decay to a pair of quarks, the angle between those two quarks is dependant only on the velocity of the parent heavy particle¹. Low energy Higgs will produce two distinct b quark jets where high energy Higgs will produce a collimated, also called “boosted”, single jet. Since we then are concerned with the energy of the parent particle, it is important for us to accurately determine if we are looking at the parent particle or not. If you recall, pileup causes jets from background processes to appear to have the same mass² as a jet from a heavier particle. Of the grooming algorithms that exist to differentiate

¹This is for the lab frame. In the rest frame of the parent particle the two quarks will decay back to back.

²We will sometimes substitute mass and energy freely since the lorentz invariant mass is also useful to measure as a proxy for energy.

these two cases, we will use the *Soft-Drop Mass Algorithm*.

The algorithm starts by unclustering the jet, recall we cluster jet with the Anti- k_T algorithm, and then categorizing the constituents as pseudo-jets. These pesudo-jets are then compared to each other using the formula:

$$\frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,3}} > z \quad (4.1)$$

where z determines the strength of the cut. Our analysis uses $z = 0.1$ **CHECK THIS** as the cut. If this condition is true, then the jet is kept. Otherwise, it is thrown out for consideration as a heavy particle jet. This helps mimic the conditions that a real heavy particle would cause rather than those that come from pileup. Since we use this as a discriminator for heavy particle jets and their decays, we will use it as a tagging variable in our analysis.

4.2 Deep AK8 Mass Decorrelated Tagger

Anotehr way to tag Higgs bosons is through exploiting the information that is gained when creating particle flow candidates and using a machine learning algorithm to identify hadronically decaying heavy particles, like the Higgs boson. The algorithm also further delineates the decay product into decay modes, i.e. a Higgs to two b quarks. This algorithm is called *Deep-AK8*. The algorithm begins by defining two lists of inputs. The first list is a list of 100³ jet constituent particles list in decreasing p_T . Measured properties of each particle, p_T , the energy deposit, the charge, the angular separation between the particle and the jet axis, etc., are used to help the algorithm extract features related to the substructure of the jet. Charged particle will also have information from tracking including track quality and displacement. These features are especially useful for identifying heavy flavour quarks, like the b quark. In total there are 42 pieces of information for each particle in the list.

³Typically, jets do not have more than 100 constituent particles so using this cap contributes to a negligible loss of information.

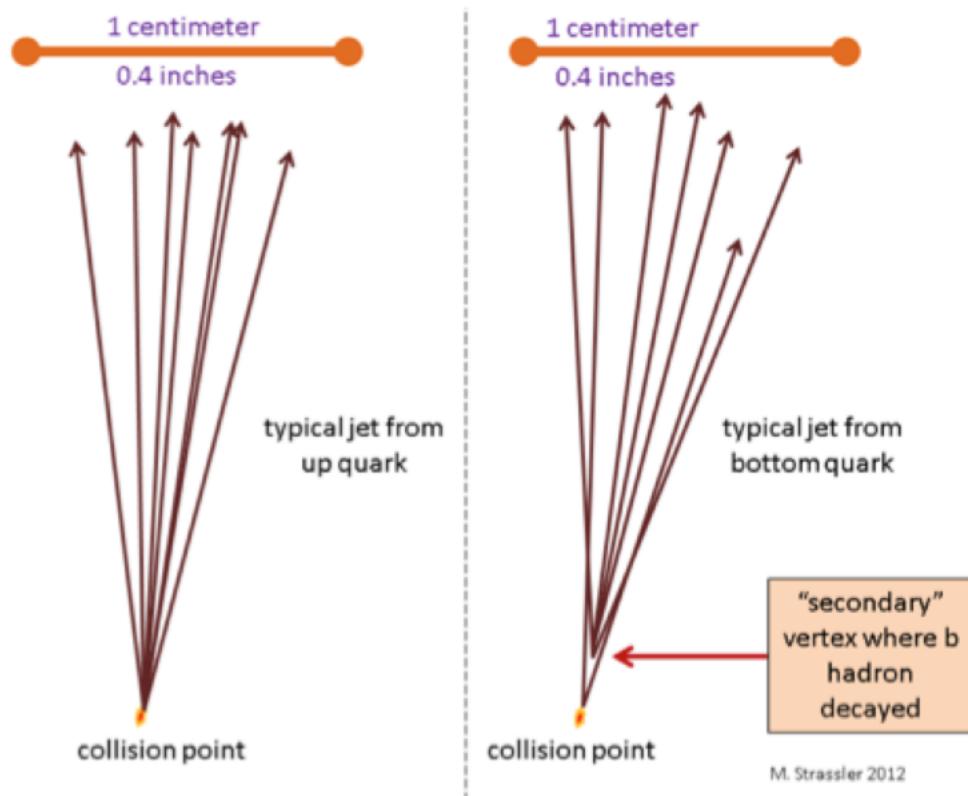


Figure 4-1. A diagram of how quarks behave in the detector.

The second list is comprised of secondary vertex information. Recall that the primary vertex relates to the location of the initial proton-proton collision. The secondary vertex relates to the location of the next decay. So then this can be useful for identifying decaying products. Notably, the b quark has a longer lifetime than other quarks, so its decay vertex, the secondary vertex, is very useful in its identification. This is shown pictorially in Figure 4-1. This list will contain up to 7 secondary vertices as well as kinematic information about the vertices, displacement, and quality of the vertices. Since this is a large amount of information, it poses a challenge to directly using it. The correlation between these inputs is very important for identifying particles so a custom neural network architecture is used.

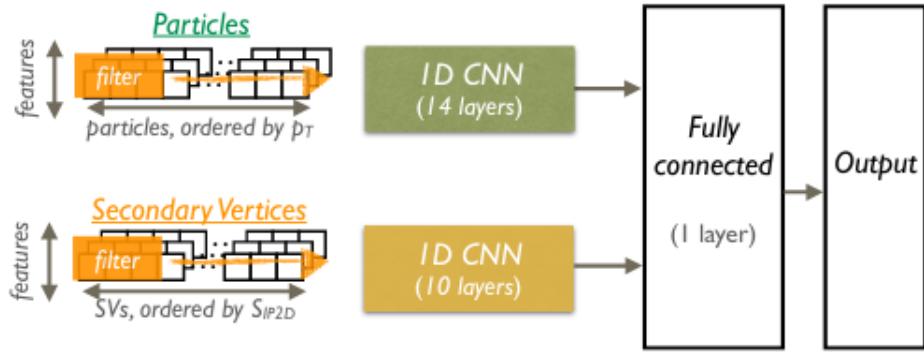


Figure 4-2. A diagram of the network architecture of the DeepAK8 Algorithm.

4.2.1 Custom Neural Network

A custom deep neural network (DNN) is constructed in order to handle the complicated correlations between inputs. This consists of two steps. The first step is to apply a convolutional neural network (CNN) is used to process each of the two lists in parallel. Then in the second step, the output of the CNN is combined by a simple, fully connected network to perform the classification of the jet. A re-weighting is used avoid any dependance on jet p_T that can occur when training the network with a mix of background and signal samples. The network architecture is shown in Figure 4-2.

4.2.2 Custom Neural Network

If one wants to use mass as a discriminating variable, then a mass-decorrelated version of the DeepAK8 algorithm is used. This will add a feature to the network architecture that acts as a mass prediction score. This score then acts as a penalty weight to prevent the network from extracting features that correlate with mass. This allows the algorithm to become largely mass independent. This will decrease the power of the algorithm. The network architecture is shown in Figure 4-3. The training of the DeepAK8-MD tagger was conducted on jets with a softdrop mass (m_{SD}) between 30 and 250 GeV so any jet outside of that range should not be used in conjunction with this algorithm.

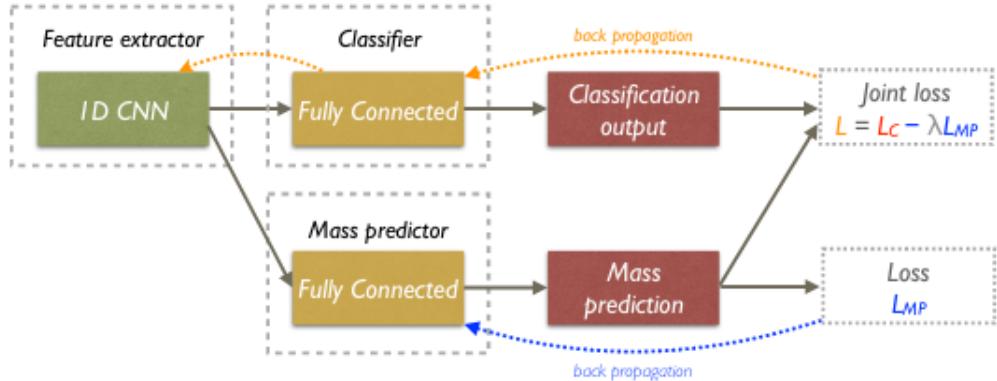


Figure 4-3. A diagram of the network architecture for the Mass-Decorrelated version of the DeepAK8 Algorithm.

4.2.3 Advantages

A similar analysis to the one performed in this analysis was completed with a different tagging algorithm and on only the 2016 data. This “Double B” tagger worked by using a boosted decision tree learning algorithm to assign a score to the jet measuring the likelihood that it contains two b quarks. It is also mass and p_T independent. At the time of that analysis, it was the best performing tagger available. However, it had drawbacks. The biggest drawback is that, while it attempted to use jet substructure information, it needed to be supplemented with directly measured substructure variables. The penalty is paid in the systematic uncertainty of those variables, which unfortunately is relatively high. The DeepAK8 tagger uses those substructure variables more efficiently so we can actually drop them as an extra discriminator and avoid paying the same penalty.

4.3 Deep Jet Tagger

The Deep Jet algorithm is used to find jets that are not pairs of b quarks but individual b quarks, i.e AK4 jets. It uses a similar two network structure like the DeepAK8 algorithm but substitutes a recurrent neural network (RNN) instead of using the simple fully connected network that DeepAK8 uses. It starts by training

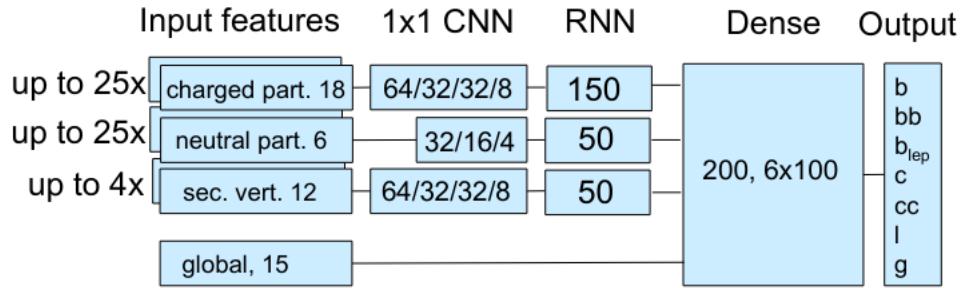


Figure 4-4. A diagram of the network architecture for the Deep Jet Algorithm.

the CNN on separate collections of charged and neutral jet particle flow candidates. The outputs are then fed into the RNN. After that training, the resultant output is combined with variables such as p_T and η of each jet and then processed by a dense layer with 7 hidden layers. The network architecture is shown in Figure 4-4. A score is then given based on the likelihood of the decay product being a heavy flavour quark. The improvement over the algorithm previously used to identify AK4 jets, called Deep CSV, is gained by using a larger set of inputs and a better neural network model.

Chapter 5

HH4b: Event Selection

In this search, the $X \rightarrow HH$ decay of a very heavy new resonance X would result in two highly Lorentz-boosted Higgs bosons. Due to its large Lorentz boost, at least one of the decay products of each $H \rightarrow b\bar{b}$ decay are collimated, and reconstructed within a single AK8 jet. These are reconstructed using jet substructure and jet flavour-tagging techniques [23–25] and, once they pass this selection, referred to as H jets.

The standard model background consists mostly of multijet events, and is estimated using several control regions defined in the phase space of the masses and flavour-tagging discriminators of the two H jets, and the HH dijet invariant mass, allowing the background to be predicted over the entire M_X range. The final event selection also contains a smaller amount of $t\bar{t} + \text{jets}$, which is modeled by 2D templates from the simulation; these templates are allowed to morph within uncertainties, and this morphing is governed by a number of nuisance parameters. The dijet M_{jj} mass distribution of the two leading Higgs-tagged jets corresponds to the invariant mass of the resonance searched for.

The effectiveness of constraining the mass of each Higgs candidate to a M_H window [26] has been studied extensively. This technique was validated in the resolved and boosted searches [26]. One needs a variable which ‘corrects’ the dijet invariant mass by the amount by which the individual H-jet masses are above or below the mass of

the Higgs boson. This approach is similar to a mass constraint in a χ^2 fit but it does not bias the H-jet mass distribution. Therefore, the variable:

$$M_{jj}^{red} = M_{jj} - (M_{jet}^1 - M_H) - (M_{jet}^2 - M_H) \quad (5.1)$$

is used to provide the best resolution of M_{jj} .

The signal would appear as a peak in the M_{jj} spectrum on top of a smooth background distribution. With this in mind, our analysis strategy is as follows:

- The signal region has two H tagged jets
- We will be using the reduced mass, defined in Equation 5.1:
- The main backgrounds are (i) QCD and (ii) $t\bar{t}$. Their ratio depends on the H-tagger used (double-b or Deep AK8).
- $t\bar{t}$ +jets is obtained from template-morphed MC shapes ($t\bar{t}$ the nuisances are constrained in the fit to data)
- QCD is obtained from H mass sidebands, and a ratio of pass and fail events, which is a smooth analytical function of m_H and $m_{reduced}$ ($R_{p/f}$) (w.r.t. H-tag). This is accomplished by assigning one AK8 jet to pass the H-tagging first, and be ‘preselection side’, and the other one is used in the 2DAlphabet background estimate and is called ‘Alphabet side’.
- Then an Alphabet style procedure is run in order to estimate the background and provide the $R_{p/f}$. The passing distribution in the signal region is modeled inside Combine by multiplying the failing distribution in data by the $R_{p/f}$.

5.1 Event Selection

5.1.1 Fully Merged Topology

Events passing the baseline triggers are further required to pass selection criteria close to the signal selection in the actual analysis:

- Leading two AK8 jets in the event with $p_T > 300\text{GeV}$ and $|\eta| < 2.4$;
- $\Delta\eta_{jj} < 1.3$ for the leading two AK8 jets, where $\Delta\eta_{jj} = |\eta_{1stFatJet} - \eta_{2ndFatJet}|$.

The details of these variables and selections are later described in Sections 5.1.2 and 5.1.3.

For the $\Delta\eta_{jj}$ cut, the rationale is to suppress QCD, since the production of a heavy $X \rightarrow HH$ will be mostly central (as the bulk of the energy of the incoming partons would be used to create X), and thus the X will usually have low boost along the z axis. In contrast, in QCD, the valence quarks that glance off of each other and each go at very high eta to produce very large dijet invariant mass events. So this cut is a natural way to suppress QCD for very high masses of X. Events are required to have at least two jets, of which the two leading jets each need to have $p_T > 300\text{GeV}$ and pseudorapidity $|\eta| < 2.4$. These two jets also need to be relatively close, $\Delta\eta_{jj} < 1.3$ in order to reduce any contribution from multijet events. A detailed study of this is given in Appendix B of [26].

5.1.2 Jet kinematics selection

Individual particles are reconstructed using a particle flow (PF) algorithm [27, 28], that combines the information from all the CMS detector components. Each such particle is referred to as a PF candidate. The five classes of PF candidates are muons, electrons, photons, and charged and neutral hadrons.

The anti- k_t algorithm [29], implemented in FASTJET [30], clusters PF candidates [27, 28] into jets using a distance parameter $R = 0.8$ (referred to as AK8 jets). In order to mitigate the effect of pileup on the different jet observables, we take advantage of the available pileup per particle identification (PUPPI) [31]. This method uses the local shape information, event pileup properties, and tracking information together in order to compute a weight describing how pileup-like a particle is.

The jet 4-momenta are corrected to account for the difference between expected and measured momentum at the particle level, using a standard CMS correction procedure described in [32]. We use the Jet Energy Correction (JEC) and Jet Energy Resolution (JER) corrections as implemented by the NanoAODtools JetMetUncertainties module¹. All jets are further required to pass tight jet identification requirements provided by JetMET POG².

Additionally, all events are required to pass all of the recommended³ filters which account for, among other things, missing energy (MET).

5.1.3 H mass selection

The masses of the two leading jets can be used to suppress the multijet and $t\bar{t}$ backgrounds. The jet-grooming algorithm called Soft Drop (cite) is used to remove the contributions from the underlying event (UE) activity and pileup, as well as remove soft and wide-angle radiation. The jet grooming leaves the hard prongs from a $H \rightarrow b\bar{b}$ unaffected, whereas it strips away most of the soft radiation from a QCD shower. As a result, the masses of QCD jets are pushed lower, whereas the jet masses from true Higgs jets are largely unchanged. The invariant mass, M_{jj} , is calculated.

¹See here <https://github.com/cms-nanoAOD/nanoAOD-tools/tree/master/python/postprocessing/modules/jme>

²<https://twiki.cern.ch/twiki/bin/view/CMS/JetMET>

³https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2#Analysis_Rec-ommendations_for_ana

5.1.4 Deep AK8 Mass De-correlated H Tagger

In order to identify the two jets most likely to contain two b quarks, we use the deep AK8 mass decorrelated Hbb tagger, labeled `deep_TagMDHbbvsQCD` [33] in nanoAOD version 5. This tagger uses customized machine learning methods and what is called the “DeepAK8” algorithm using particle level information and secondary vertex information. Each is split into a separate “set” and a 1D CNN is applied to each set. The output of that is fed into a fully connected network to perform the jet classification. A mass prediction network is added for the version we are using because it is not desirable to use the “DeepAK8” algorithm when the mass distribution is used to discriminate between signal and background, as is in this analysis. Further details can be found in [33].

We have made a S/\sqrt{B} study in order to optimize the working points for this tagger, where we measure the data-to-MC efficiency SF ourselves. It should be noted that, in the optimization, the change in the H-jet selection affects also the ‘preselection’ H-jet.” At each signal MC mass point, S/\sqrt{B} is calculated and the working points are chosen as the average over all the signal MC mass points.⁴ We are using 0.9 as the tight working point and 0.8 as the loose working point⁵. We also show cutflow diagrams to illustrate the working point efficiencies for these points chosen.

⁴In the previous analysis [26], τ_N was used to quantify the degree to which jet constituents could be arranged into N subjets. The ratio $\tau_{21} = \tau_2/\tau_1$ was calculated for both jets and contributed a systematic uncertainty of 5 – 15% [26]. This requirement has been dropped for this analysis. We are using more modern taggers, `deepAK8MD_HbbvsQCD` [33] for example, which have the substructure information already included. Since this selection criteria is now redundant, we drop it and gain by reducing our systematic uncertainty. This study is documented in Appendix C.

⁵See the Appendix for previous studies on the working point.

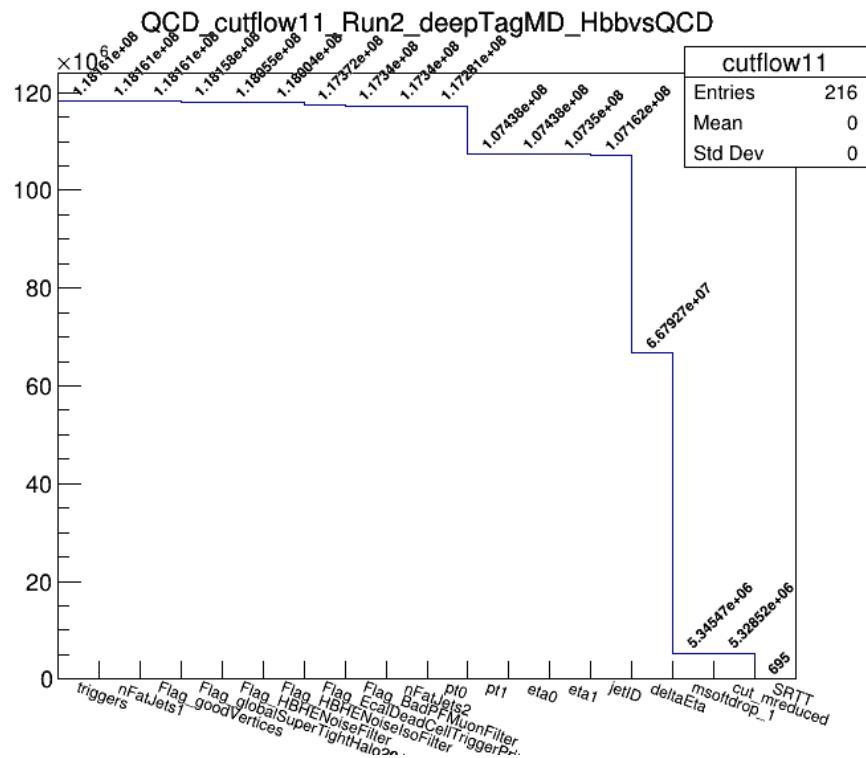


Figure 5-1. Cutflow Diagram for Tight Tight QCD Selection

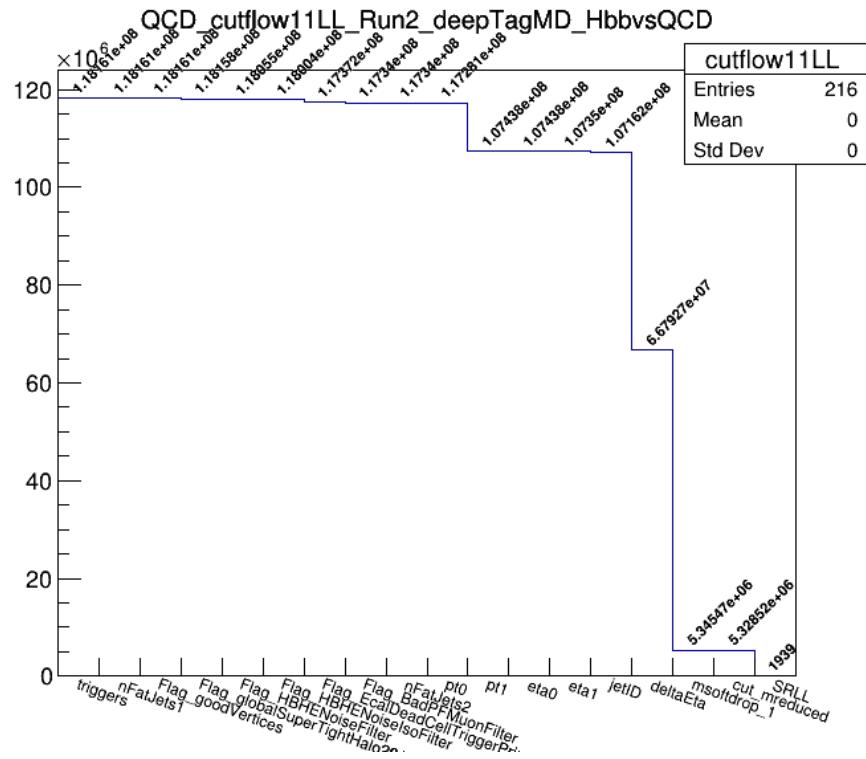


Figure 5-2. Cutflow Diagram for Loose Loose QCD Selection

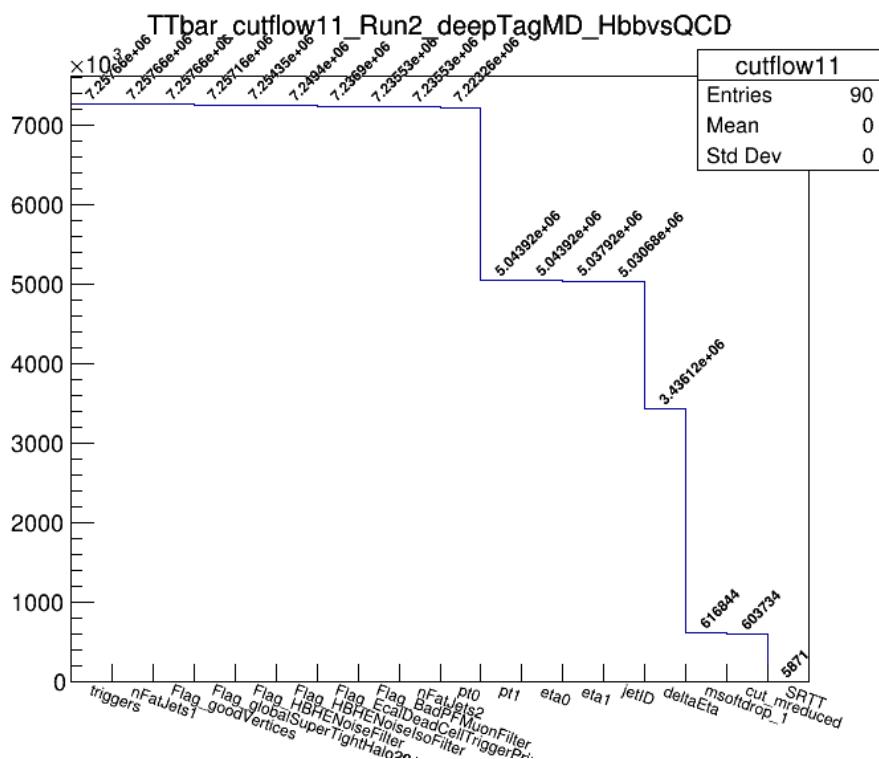


Figure 5-3. Cutflow Diagram for Tight Tight $t\bar{t}$ Selection

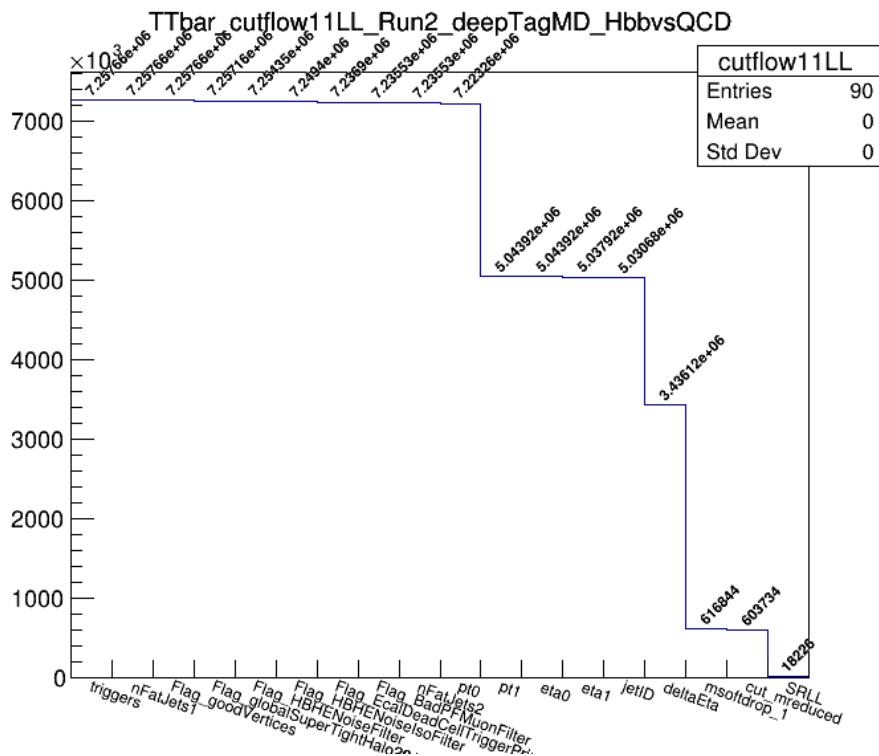


Figure 5-4. Cutflow Diagram for Loose Loose $t\bar{t}$ Selection

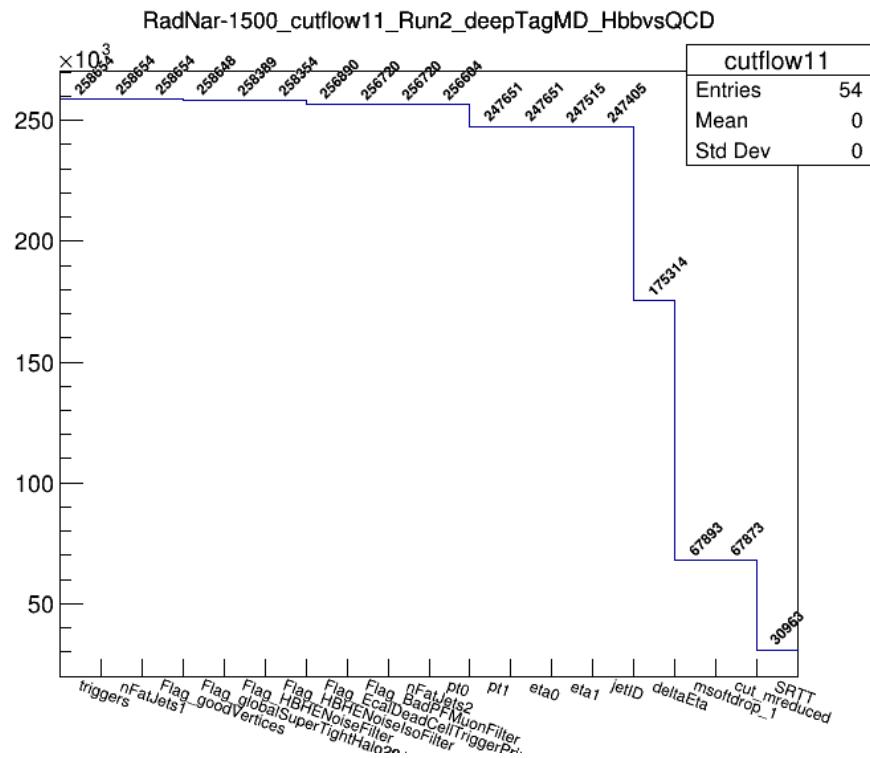


Figure 5-5. Cutflow Diagram for Tight Tight Radion 1500 GeV Selection

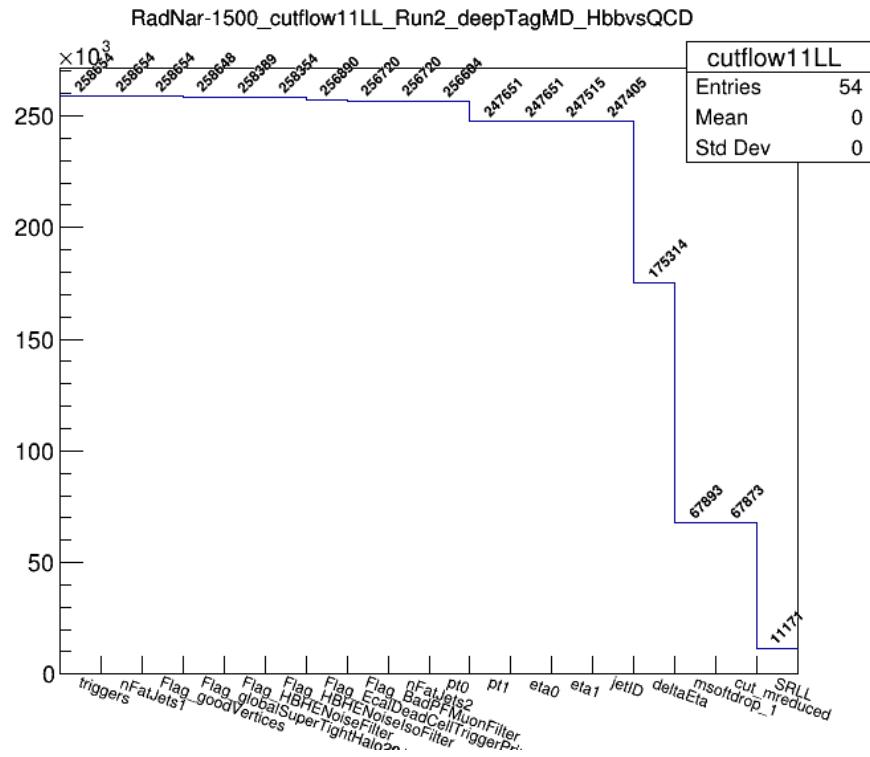


Figure 5-6. Cutflow Diagram for Loose Loose Radion 1500 GeV Selection

5.1.5 Boosted Event Selection

The event selection follows the same as the trigger above but adding the following criteria:

- The soft drop mass of the two jets are $110 < M_{softdrop} < 140 \text{ GeV}$, with all necessary jet mass corrections applied
- The dijet invariant mass, $M_{jj}^{red}, > 750$.

The soft drop mass window was optimized by making preliminary limits for three different windows. The results are shown in Figure 5-7

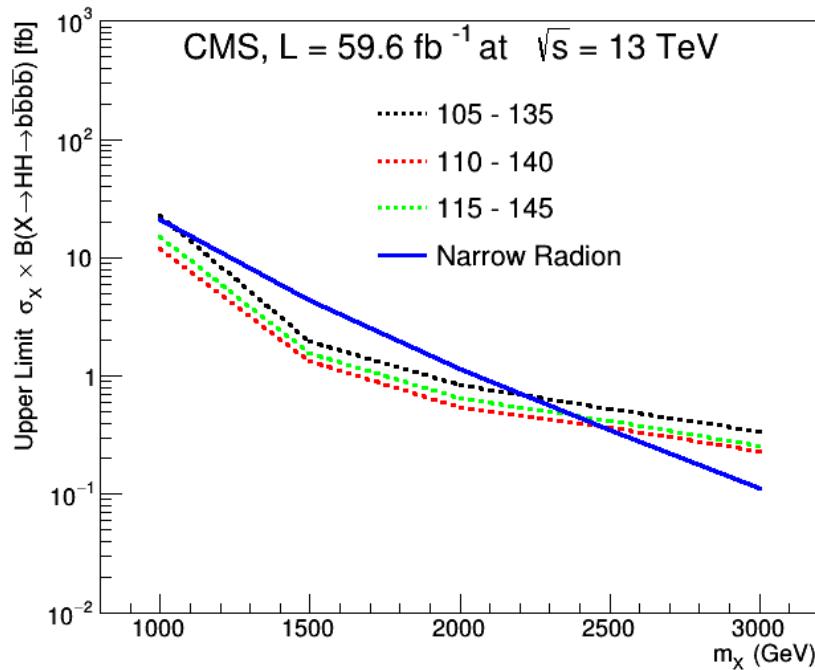


Figure 5-7. Preliminary limits used to optimize the soft drop mass window.

5.1.6 Identification of Higgs jets in boosted analysis

For the boosted analysis, events are required to pass the above selection criteria. The double b-tagger described in Section 5.1.4 is used to identify the boosted Higgs

jet in the boosted analysis. We then split the candidate events into tight-tight (TT) and loose-loose (LL) pass and fail regions. The algorithm to do this is as follows: The 2D Alphabet method, with the signal region blinded, is then used to obtain the background estimate in the signal region. The possible combinations taken into account are:

- TT Pass: Both H-jets pass the tight operating point;
- TT Fail: The leading H-jet fails the tight operating point and the subleading passes.
- LL Pass: Both H-jets pass the loose working point but both fail the tight working point. This also means that if one jet passes the loose working point and the other passes the tight working point, the event will be classified as LL.
- LL Fail: The leading H-jet fails the loose operating point and the subleading passes the loose operating point. The subleading H-jet must also fail the tight working point.

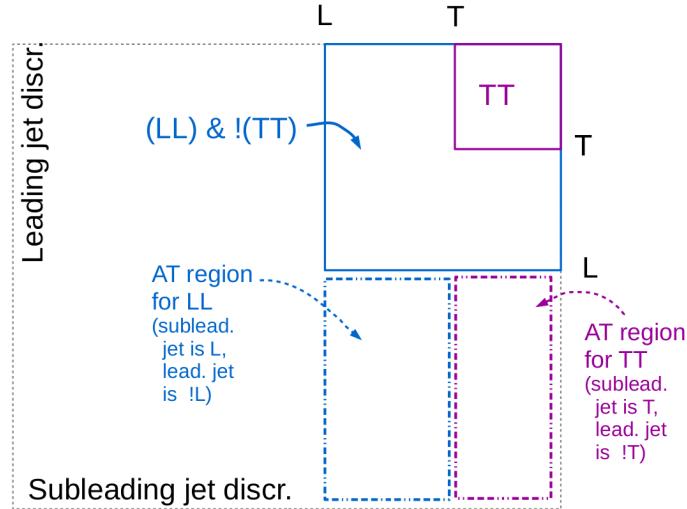


Figure 5-8. A pictorial representation of the semaphoring of Loose Loose and Tight Tight events.

5.1.7 MC distributions before applying Selections

The kinematic variables before the application of $H \rightarrow bb$ tagger(s), are shown in Figs.5-9 through 5-15. Here, the $t\bar{t}$ distributions are normalized to luminosity where signal and qcd is normalized to $t\bar{t}$ for ease of viewing. Also, the difference between TT and LL distributions is in the weight used for each. Since the deepTagMD_HbbvsQCD tagger has not been applied yet, we do not split the plots into Tight Tight and Loose

Loose.

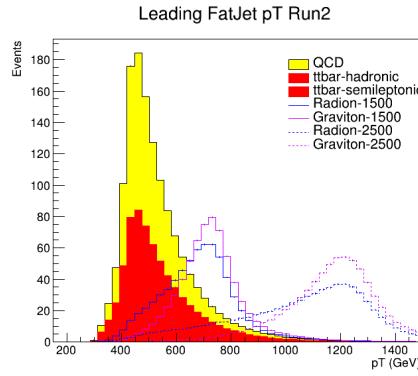


Figure 5-9. Pre deepTagMD_HbbvsQCD selection p_T distribution of leading fat jet

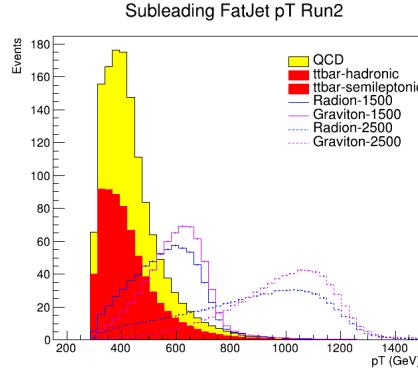


Figure 5-10. Pre deepTagMD_HbbvsQCD selection p_T distribution of subleading fat jet

5.2 Semi-Resolved Topology

The semi-resolved case bridges the gap between the fully resolved analysis for $HH \rightarrow 4b$ [34] and the boosted analysis, presented in Section 5.1. It assumes one H is

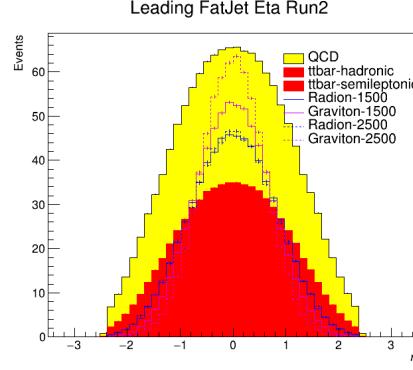


Figure 5-11. Pre deepTagMD_HbbvsQCD selection η distribution of leading fat jet

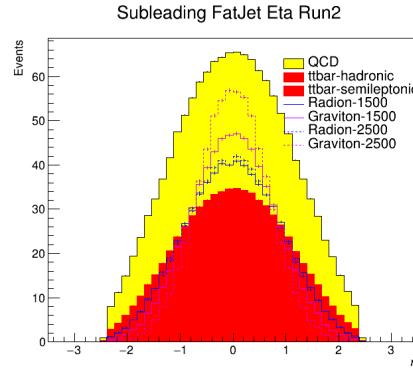


Figure 5-12. Pre deepTagMD_HbbvsQCD selection η distribution of subleading fat jet

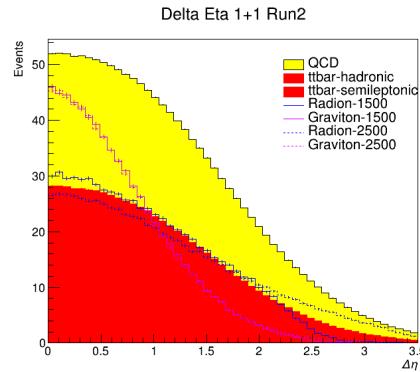


Figure 5-13. Pre deepTagMD_HbbvsQCD selection $\Delta\eta$ distribution

boosted enough to be contained within an AK8 jet with two subjets and the other H is not, resulting in two AK4 jets, one for each b quark. We use techniques similar to the boosted analysis to identify the boosted Higgs and use techniques similar to the resolved analysis to identify the two resolved b-jets. For this case, the mass range is

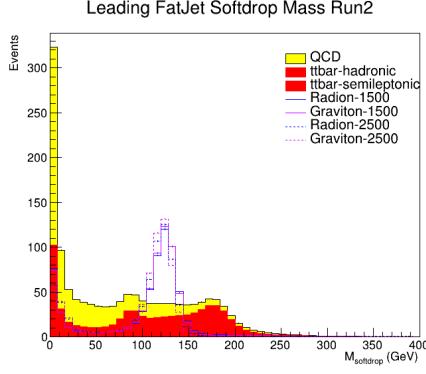


Figure 5-14. Pre deepTagMD_HbbvsQCD selection m_j distribution

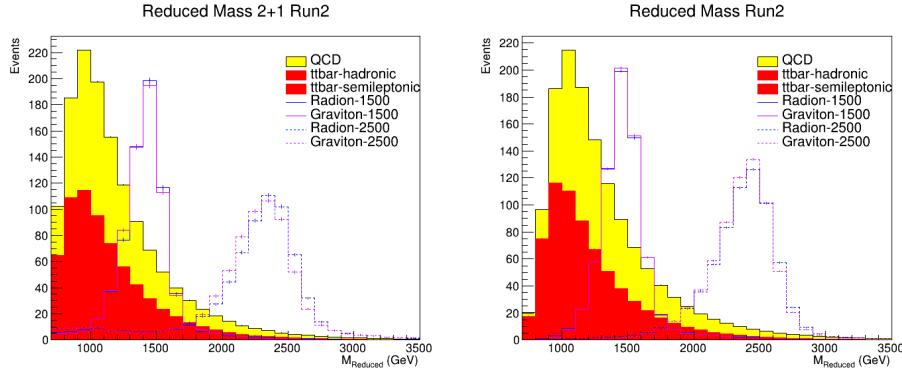


Figure 5-15. Pre deepTagMD_HbbvsQCD selection m_{jj} distribution

sensitive as well.

5.2.1 Event Selection

The jet kinematics selection is the same as that of the boosted analysis for the AK8 jet. The AK4 jets also have a kinematic selection. Events are required to have:

- 2 AK4 jets, $p_T > 30$ GeV, $|\eta| < 2.0$, deep Jet > medium WP by year [2016:0.3093 ,2017:0.3033 ,2018:0.2770]
- 1 AK8 jet, $p_T > 300$ GeV, $|\eta| < 2.4$
- Trijet mass, defined and studied in [26], > 200.0 GeV

We choose the appropriate AK4 jets by requiring that the AK8 jet is at least $\Delta R > 0.8$ away from the candidate AK4 jets. We also require that the AK4 jets are $\Delta\phi > \frac{\pi}{2}$ away

from the candidate AK8 jet in order to guarantee that the AK4 jets are in a different hemisphere than the AK8 jet. For each candidate AK4 jet pair, we require that the jets are within $\Delta R < 1.5$. After this, the dijet mass is required to be between 90 and 140 GeV, and the DeepAK8-discriminant is used to determine the control region (DeepAK8 discriminant < 0.9 , Tight WP) and signal region (DeepAK8 discriminant > 0.9).

5.2.2 Identification of Higgs jets in semi-resolved analysis

For the semi-resolved analysis, events are required to have at least one jet as described Section 5.1. The DeepAK8-tagger described in Section 5.1 is used to identify the boosted Higgs jet in the semi-resolved analysis, requiring the fatjet pass the working point of DeepAK8 discriminant > 0.9 . We show the cutflow diagram here.

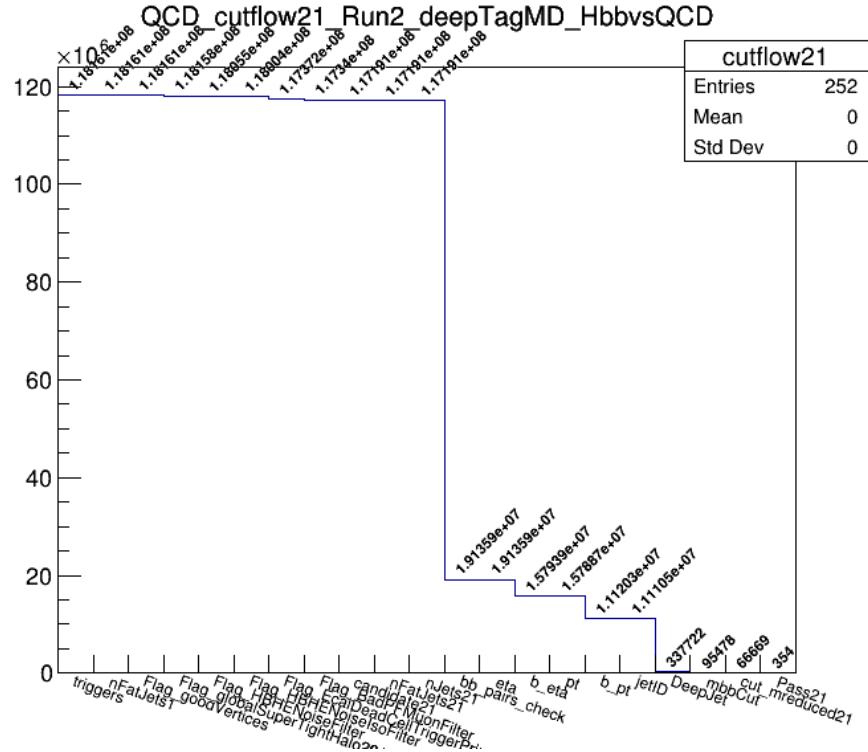


Figure 5-16. Cutflow Diagram for Semi-resolved for QCD Selection

The invariant mass of all three jets is used in the same way as the reduced dijet

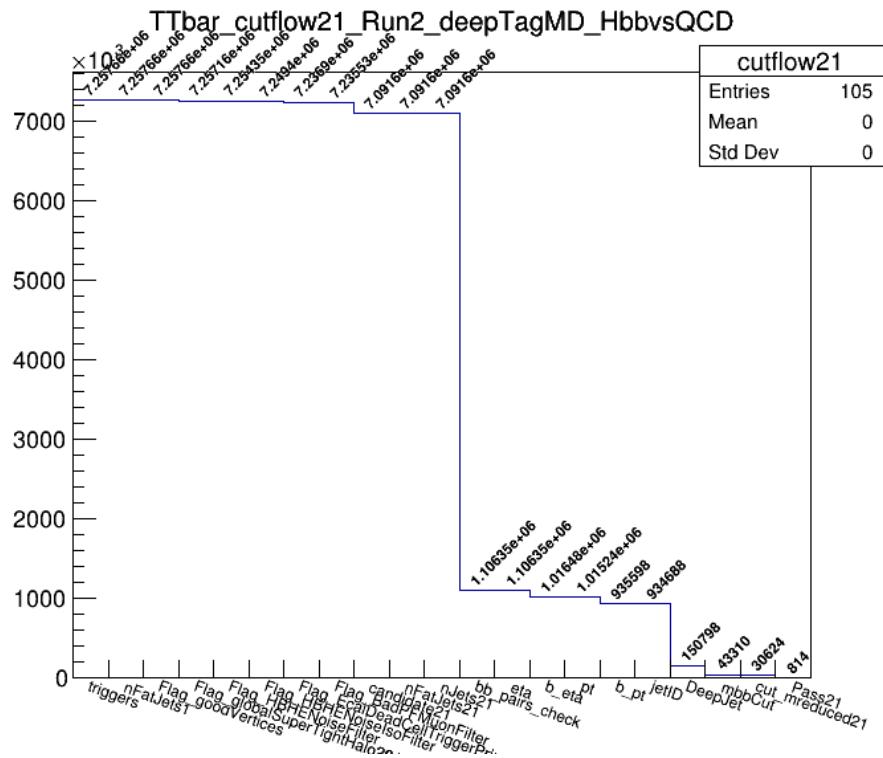


Figure 5-17. Cutflow Diagram for Semi-resolved for $t\bar{t}$ Selection

mass is used for the boosted analysis and is given by:

$$M_{jjj}^{red} = M_{jjj} - (M_{Fatjet} - M_H) - (M_{jj(Ak4jets)} - M_H) \quad (5.2)$$

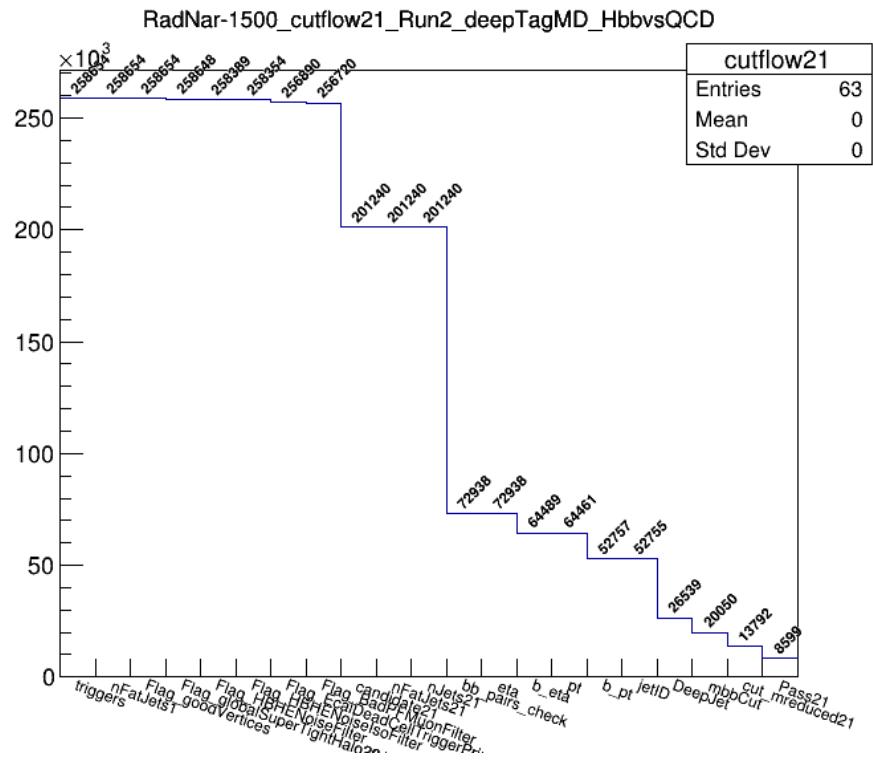


Figure 5-18. Cutflow Diagram for Semi-resolved for Radion 1500 GeV Selection

Chapter 6

HH4b: Background Estimate and Systematic Uncertainties

6.1 Background Estimation

6.1.1 Alphabet Background Estimate

After applying the kinematic selection, we define the ratio of multijet distributions that pass and fail the H tagging requirement in data and QCD multijet MC simulation as $R_{\text{P/F}}^{\text{data}}(M_j, M_{jj}^{red})$ and $R_{\text{P/F}}^{\text{MC}}(M_j, M_{jj}^{red})$, respectively. Because of the combinatorial nature of multijet processes, $R_{\text{P/F}}^{\text{data}}(M_j, M_{jj}^{red})$ and $R_{\text{P/F}}^{\text{MC}}(M_j, M_{jj}^{red})$ are both smooth in the (M_j, M_{jj}^{red}) plane. The ratio of these ratios ($R_{\text{ratio}}(M_j, M_{jj}^{red})$) is then also smooth and can be used to correct for differences in simulation and data by parameterizing it with an analytic function. While $R_{\text{P/F}}^{\text{data}}(M_j, M_{jj}^{red})$ could also be described by an analytic function, it requires more parameters and a more complex shape relative to $R_{\text{ratio}}(M_j, M_{jj}^{red})$. The number of events in a given bin in the passing category can then be estimated by:

$$n_{\text{P,QCD}}(i) = n_{\text{F,QCD}}(i) \cdot R_{\text{P/F}}^{\text{MC}}(M_j, M_{jj}^{red}) \cdot R_{\text{ratio}}(M_j, M_{jj}^{red}), \quad (6.1)$$

where $f(M_j, M_{jj}^{red})$ has been replaced by $R_{\text{P/F}}^{\text{MC}}(M_j, M_{jj}^{red}) \cdot R_{\text{ratio}}(M_j, M_{jj}^{red})$ and $R_{\text{ratio}}(M_j, M_{jj}^{red})$ is a surface parameterized by a two-dimensional polynomial in the (M_j, M_{jj}^{red}) plane with coefficients whose values are a priori unknown to the fit. A second-order polyno-

mial was chosen for the M_{jj}^{red} axis and the M_j axis. The parameters of the polynomials are uncorrelated between regions. To eliminate the effect of statistical fluctuations when calculating $R_{P/F}^{\text{MC}}(M_j, M_{jj}^{red})$ in the QCD multijet simulation, pass and fail distributions are smoothed by using an adaptive kernel estimate [35] (KDE). Figures 6-2,6-4,6-6 show the smoothed and un-smoothed distributions for $R_{P/F}^{\text{MC}}(M_j, M_{jj}^{red})$. The mostly uniform distributions that we get after smoothing are what we expect because we are trying to factor out the statistical fluctuations in QCD due to low statistics.

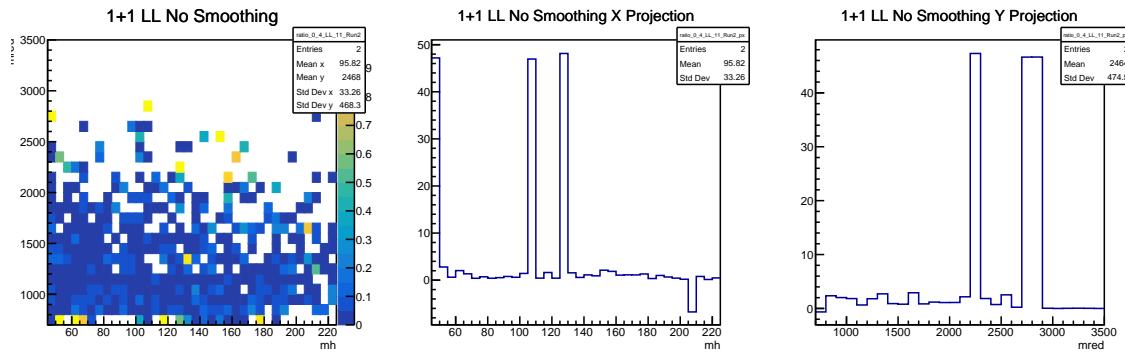


Figure 6-1. Un-smoothed $R_{P/F}^{\text{MC}}(M_j, M_{jj}^{red})$ distributions for Loose Loose.

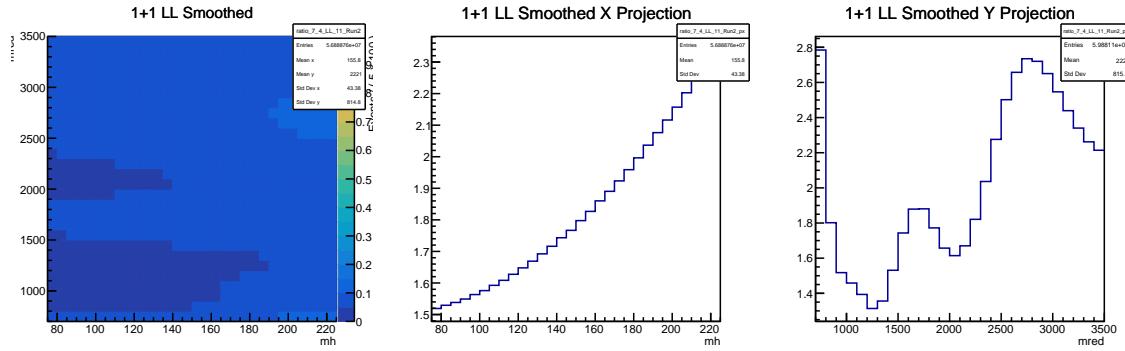


Figure 6-2. Smoothed $R_{P/F}^{\text{MC}}(M_j, M_{jj}^{red})$ distributions for Loose Loose.

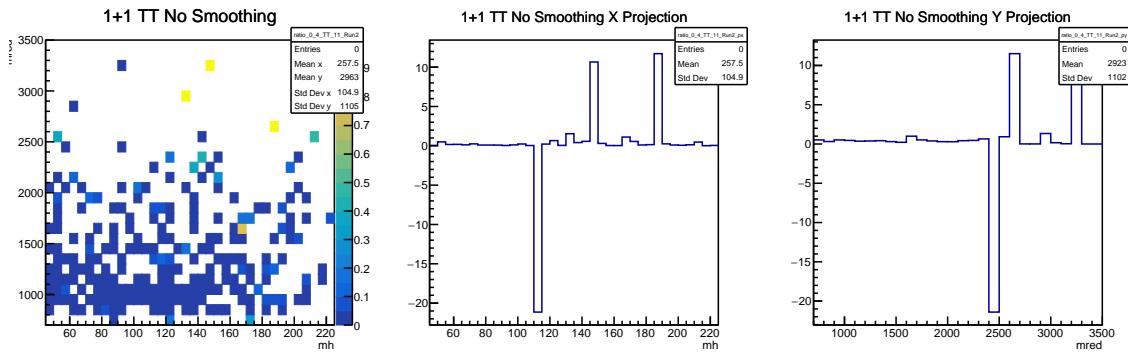


Figure 6-3. Un-smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Tight Tight.

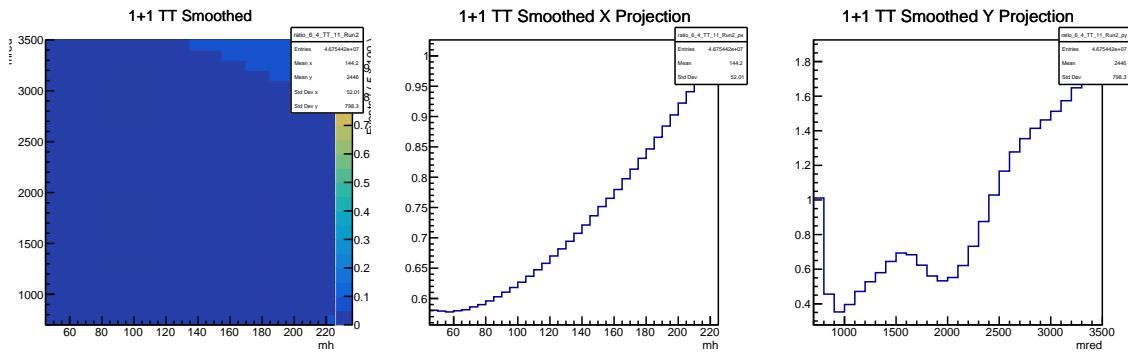


Figure 6-4. Smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Tight Tight.

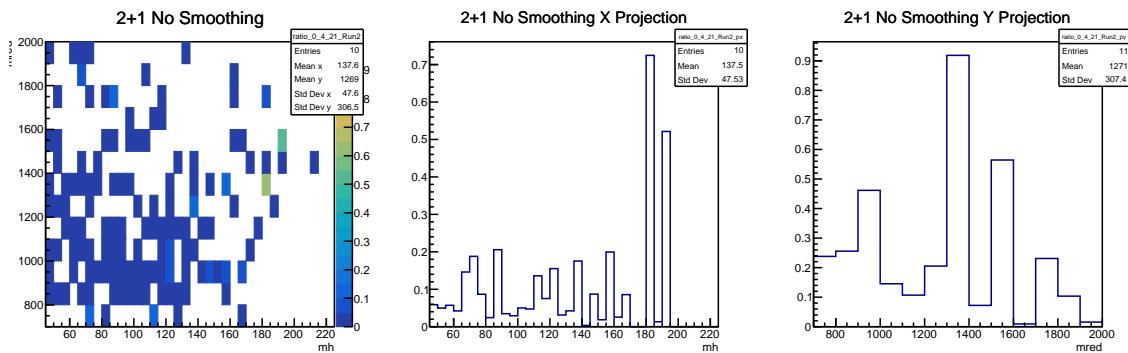


Figure 6-5. Un-smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Semi-resolved.

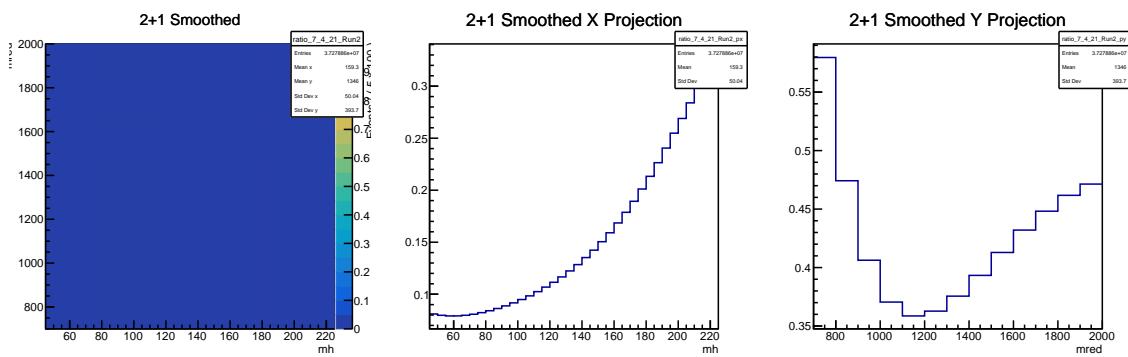


Figure 6-6. Smoothed $R_{P/F}^{MC}(M_j, M_{jj}^{red})$ distributions for Semi-resolved.

Due to the fact that the QCD MC samples are binned in HT, they need to be normalized to each other’s cross sections before they are combined. While also normalizing to data luminosity, samples could be scaled by factors between 0.15 and 8.43. These have significant impacts on the effective yield per-bin passed to the Kernel Density Estimate and can thus change the bandwidth of the kernels which adapts to the statistics per-bin. Since we want to smooth based on the generated statistics, a KDE PDF is built per HT-bin and samples *before* the scaling and stitching. One billion events are generated for the pass and fail distribution of each HT-binned sample and then normalized to the relative cross sections and data luminosity. The KDE generated distributions are then stitched to together and the ratio of pass and fail is taken to create the $R_{\text{P/F}}^{\text{MC}}(M_j, M_{jj}^{\text{red}})$. We also perform the same smoothing with the $t\bar{t}$ distributions in order to eliminate any statistical fluctuations that cause the fit to try to enhance an un-physical peak. Two parameters ρ and σ control the smoothing as follows. ρ is a scale factor that is applied to the bandwidth calculated for each kernel. σ determines the size of the box that is used to search for contributing kernels around a given point and is also used for the 1st non-adaptive pass for the calculation of adaptive keys pdfs. The various values for the KDE bandwidth, ρ and σ , are shown here:

Region	Features
TT	QCD $\rho = 6$ and $\sigma = 4$, $t\bar{t}$ $\rho = 1$ and $\sigma = 2$
LL	QCD $\rho = 7$ and $\sigma = 4$, $t\bar{t}$ $\rho = 1$ and $\sigma = 2$
2+1	QCD $\rho = 7$ and $\sigma = 4$, $t\bar{t}$ $\rho = 1$ and $\sigma = 2$

6.1.1.1 Fitting procedure

The production cross section of $\sigma(pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}b\bar{b})$ is evaluated by comparing for each bin in the two-dimensional $(M_j, M_{jj}^{\text{red}})$ distribution, the number of observed and expected events, given the expected background and the theoretical cross section. The expected number of events is calculated as $N_{\text{expected}} = \sigma_{HH \rightarrow b\bar{b}b\bar{b}} \times \varepsilon \times \mathcal{L}$, where $\sigma_{HH \rightarrow b\bar{b}b\bar{b}}$ is the $HH \rightarrow b\bar{b}b\bar{b}$ cross-section, ε is the acceptance times the efficiency,

and L is the integrated luminosity of our dataset. A likelihood fit to data is used to test the signal hypothesis where the total background model is constructed as a sum of the individual background contributions using a Poisson model for each bin of the (M_j, M_{jj}^{red}) distribution. The number of expected failing and passing events in a given bin i is given by

$$n_F(i, \vec{\theta}) = n_F^{\text{QCD}}(i) + n_F^{t\bar{t}}(i, \vec{p}) + n_F^{\text{signal}}(i, \vec{r}) \quad (6.2)$$

and

$$n_P(i, \vec{\theta}) = n_P^{\text{QCD}}(i) + n_P^{t\bar{t}}(i, \vec{p}) + n_P^{\text{signal}}(i, \vec{r}), \quad (6.3)$$

where i is a bin in the (M_{jet}, M_{jj}^{red}) plane, \vec{p} , \vec{q} , and \vec{r} are the nuisance parameters, and $\vec{\theta}$ is the union set of all nuisance parameters. The variable $n_{i,F,\text{QCD}}$ is an unconstrained positive real number. Finally, $n_{P,\text{QCD}}(i)$ is given by

$$n_{P,\text{QCD}}(i) = n_{F,\text{QCD}}(i) \cdot f(M_j, M_{jj}^{red}), \quad (6.4)$$

where $f(M_j, M_{jj}^{red})$ is a transfer function used in the data-driven multijet background estimate and described fully in Section 6.1.1. The negative log-likelihood is then

$$\begin{aligned} -\ln L(\vec{d}; \vec{\theta}) &= \sum_{i=1}^{N_{\text{bins},F}} \left[n_F(i, \vec{\theta}) - d_F(i) \ln n_F(i, \vec{\theta}) + \ln d_F(i)! \right] \\ &\quad + \sum_{i=1}^{N_{\text{bins},P}} \left[(n_P(i, \vec{\theta}) - d_P(i) \ln n_P(i, \vec{\theta}) + \ln d_P(i)!) \right], \end{aligned} \quad (6.5)$$

where $N_{\text{bins},F}$ and $N_{\text{bins},P}$ are the total number of bins in the fail and pass distributions, respectively, and $d_{i,F}$ and $d_{i,P}$ are the number of observed events in a given bin in the fail and pass distributions, respectively.

6.1.2 Closure Test in Data

The double-b tag inverted control sample has been selected by applying all the event selection criteria but the discriminator, deepAK8MDHbb, tag requirement. We check for a leading and subleading jet that “pass” if the leading H-jet passes the LL

working point and the subleading H-jet fails the LL working point. The “fail” region is then when both leading and subleading jets fail the LL working point. The results are shown in the following plots:

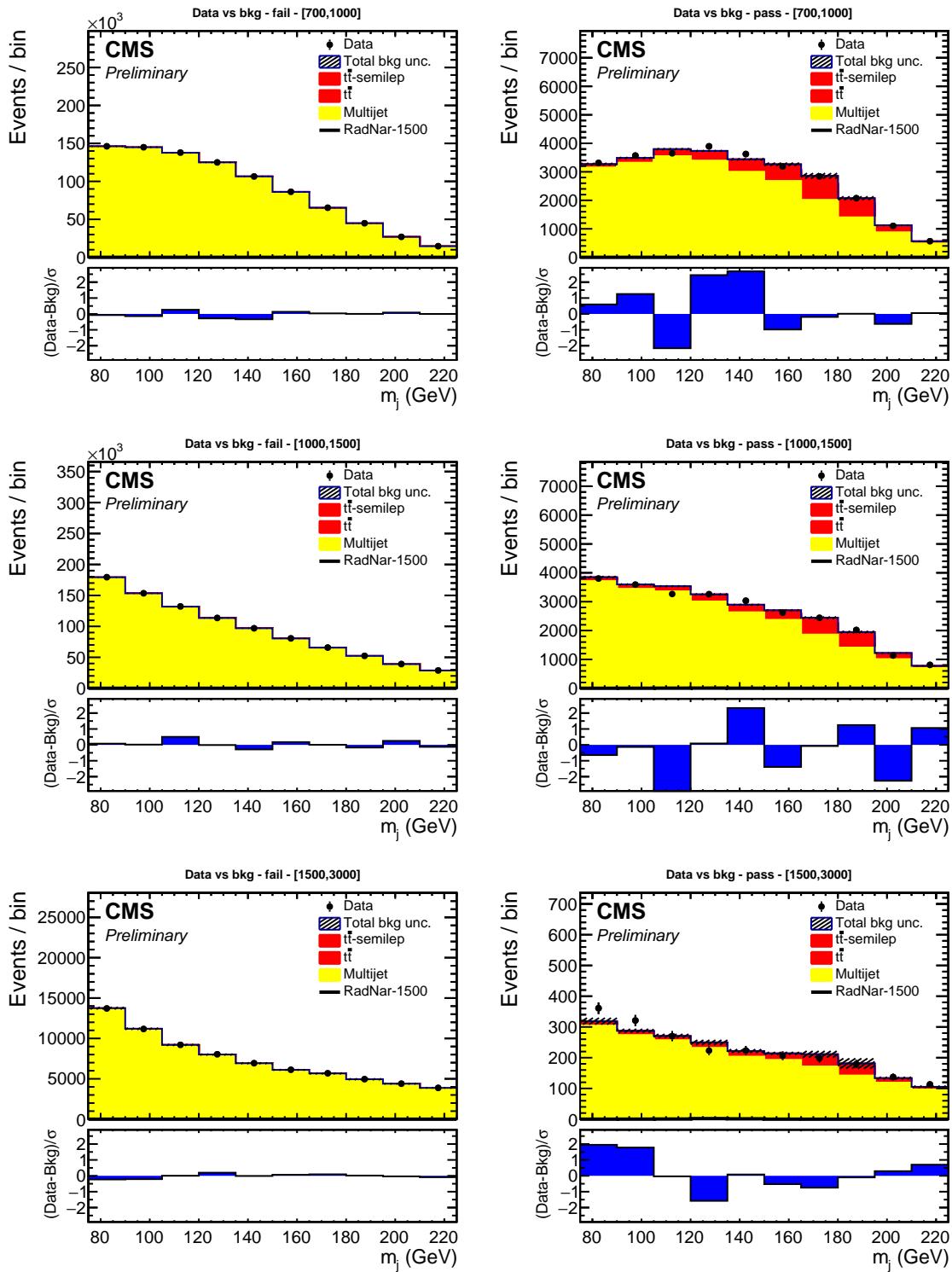


Figure 6-7. Full Run 2 CR background fits for M_j axis.

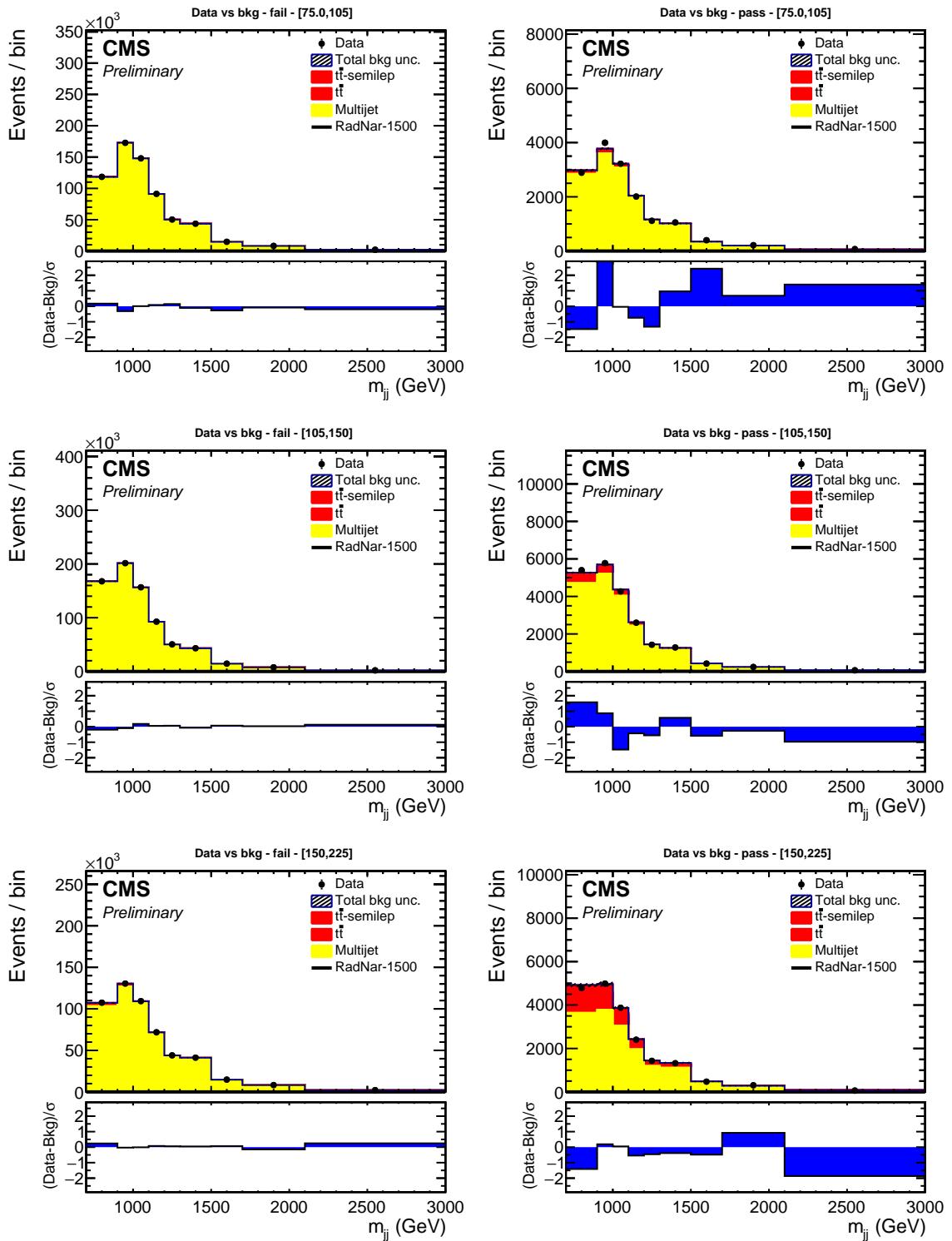


Figure 6-8. Full Run 2 CR background fits for M_{jj} axis.

6.1.3 Top Control Region

We define a $t\bar{t}$ control region in order to help constrain the fit to the $t\bar{t}$ MC background. The event selection is the same as the fully merged topology except we change the soft drop mass window to be $140 < M_{softdrop} < 210 GeV$ in order to select for the top mass in the subleading H-jet.

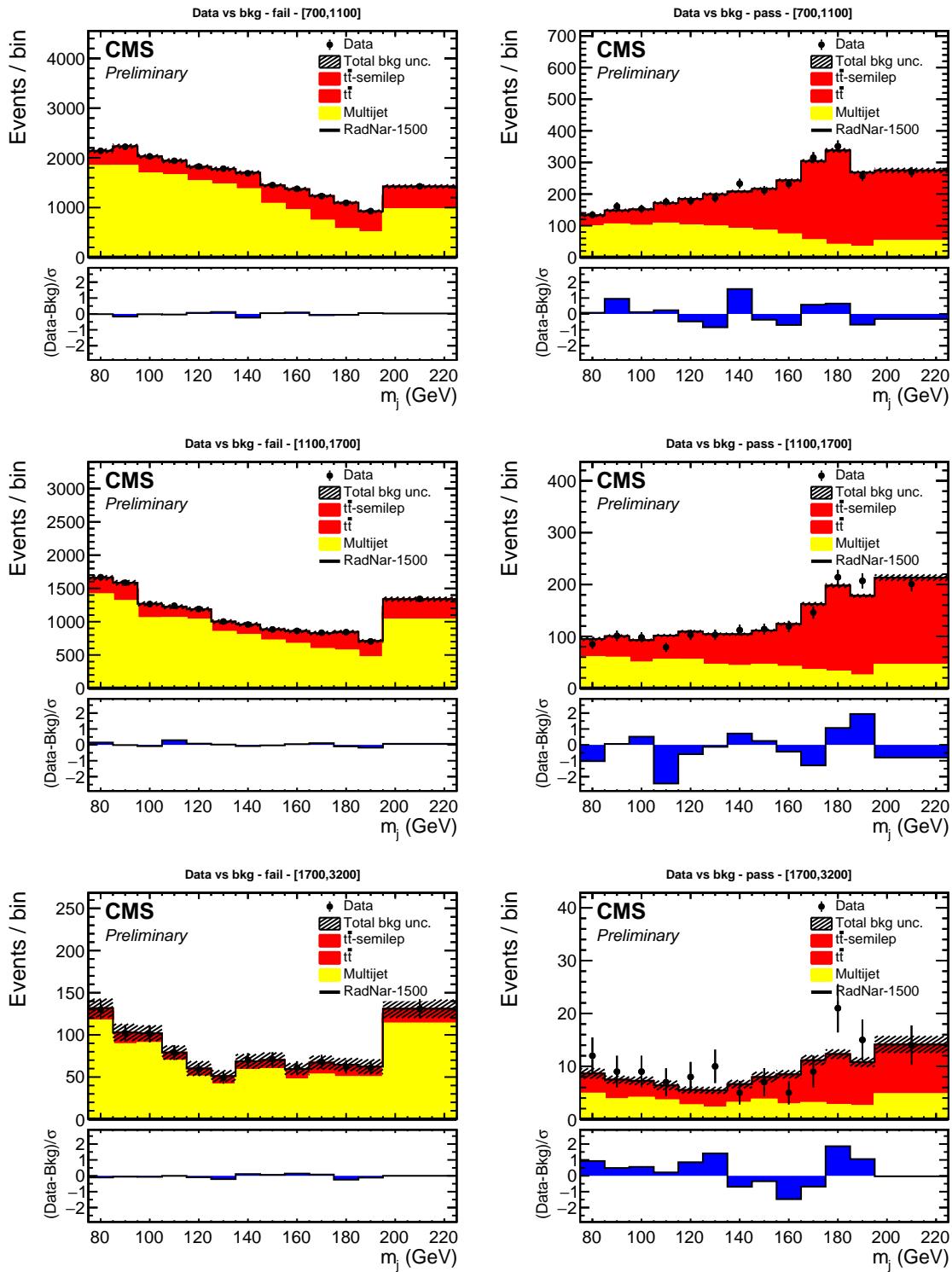


Figure 6-9. Full Run 2 Loose Loose $t\bar{t}$ Control Region fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

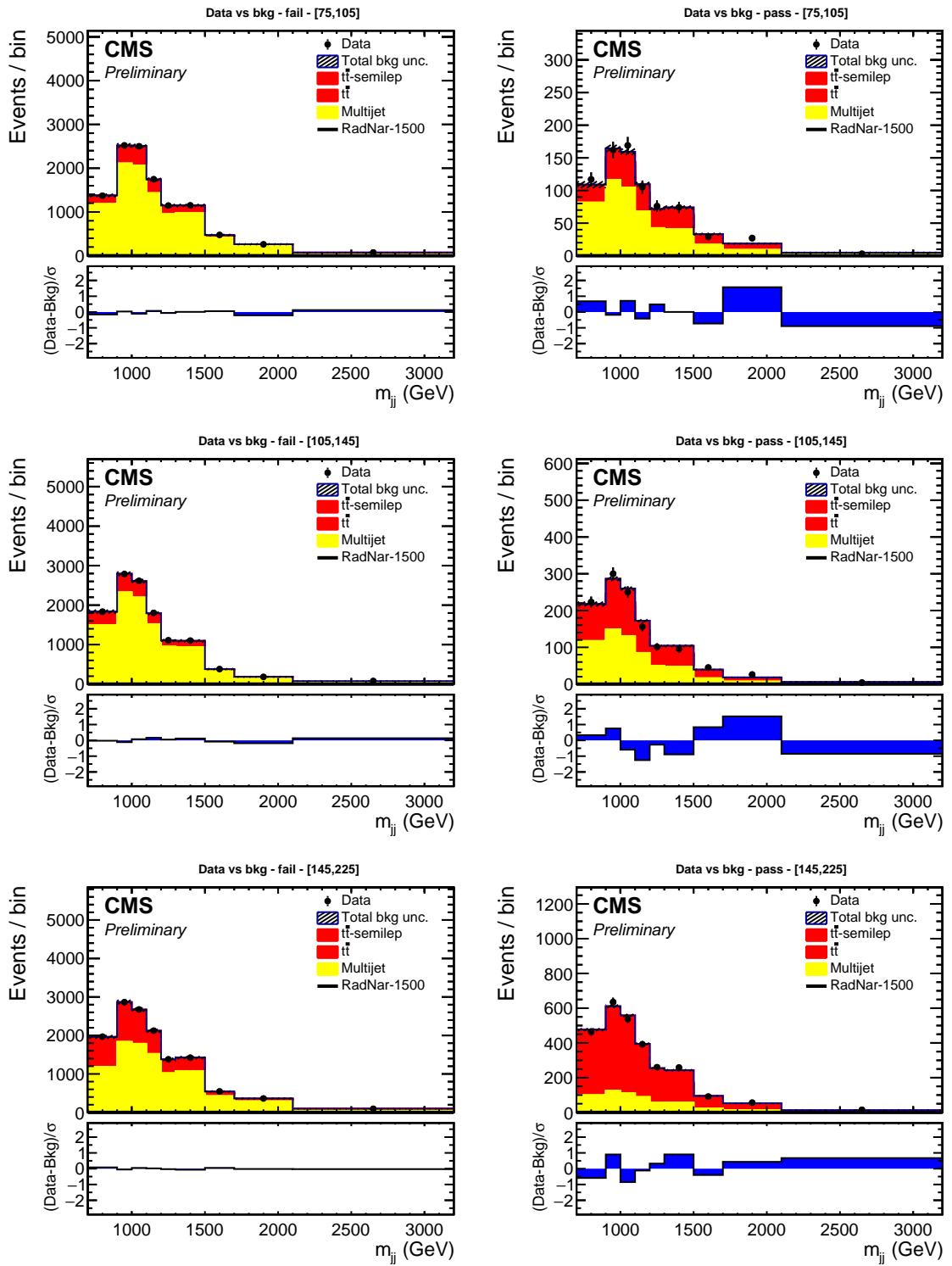


Figure 6-10. Full Run 2 Loose Loose $t\bar{t}$ Control Region fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

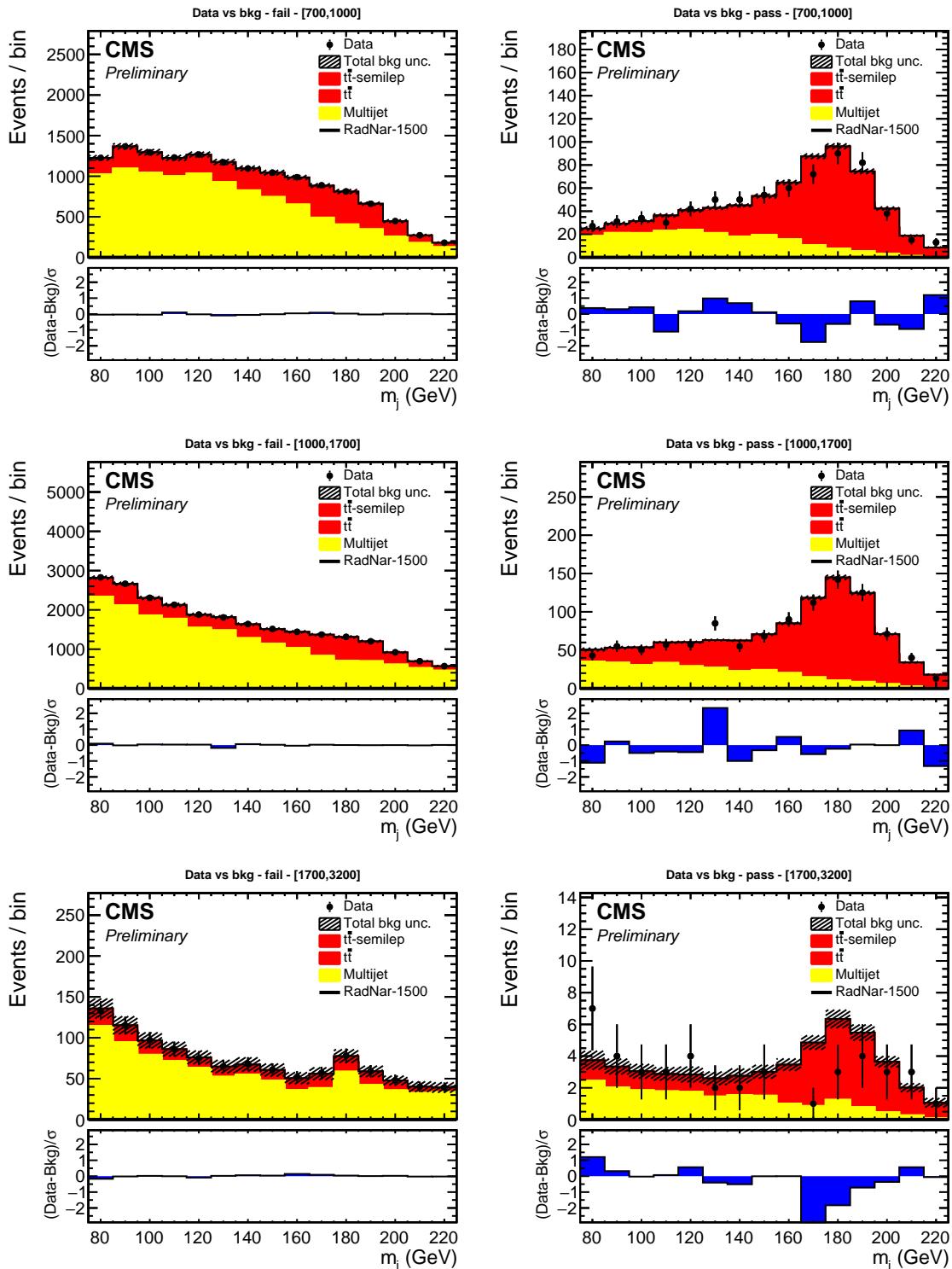


Figure 6-11. Full Run 2 Tight Tight $t\bar{t}$ Control Region fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

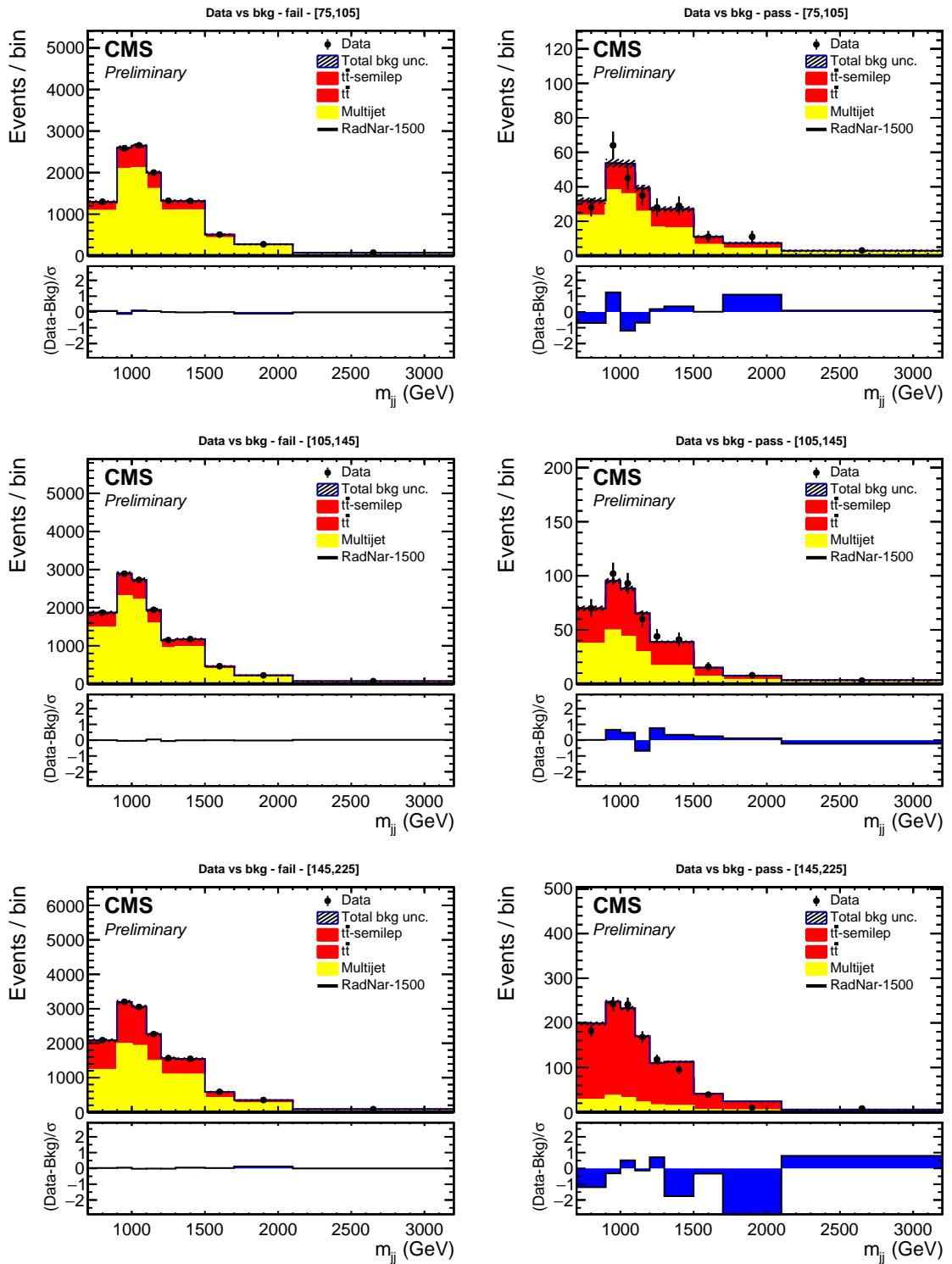


Figure 6-12. Full Run 2 Tight Tight $t\bar{t}$ Control Region fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

6.1.4 Signal Regions Prediction

The following section displays the post-fit distributions from the full simultaneous fits performed across the 2016, 2017, and 2018 distributions using the DeepAK8 Mass Decorrelated Hbb tagger.

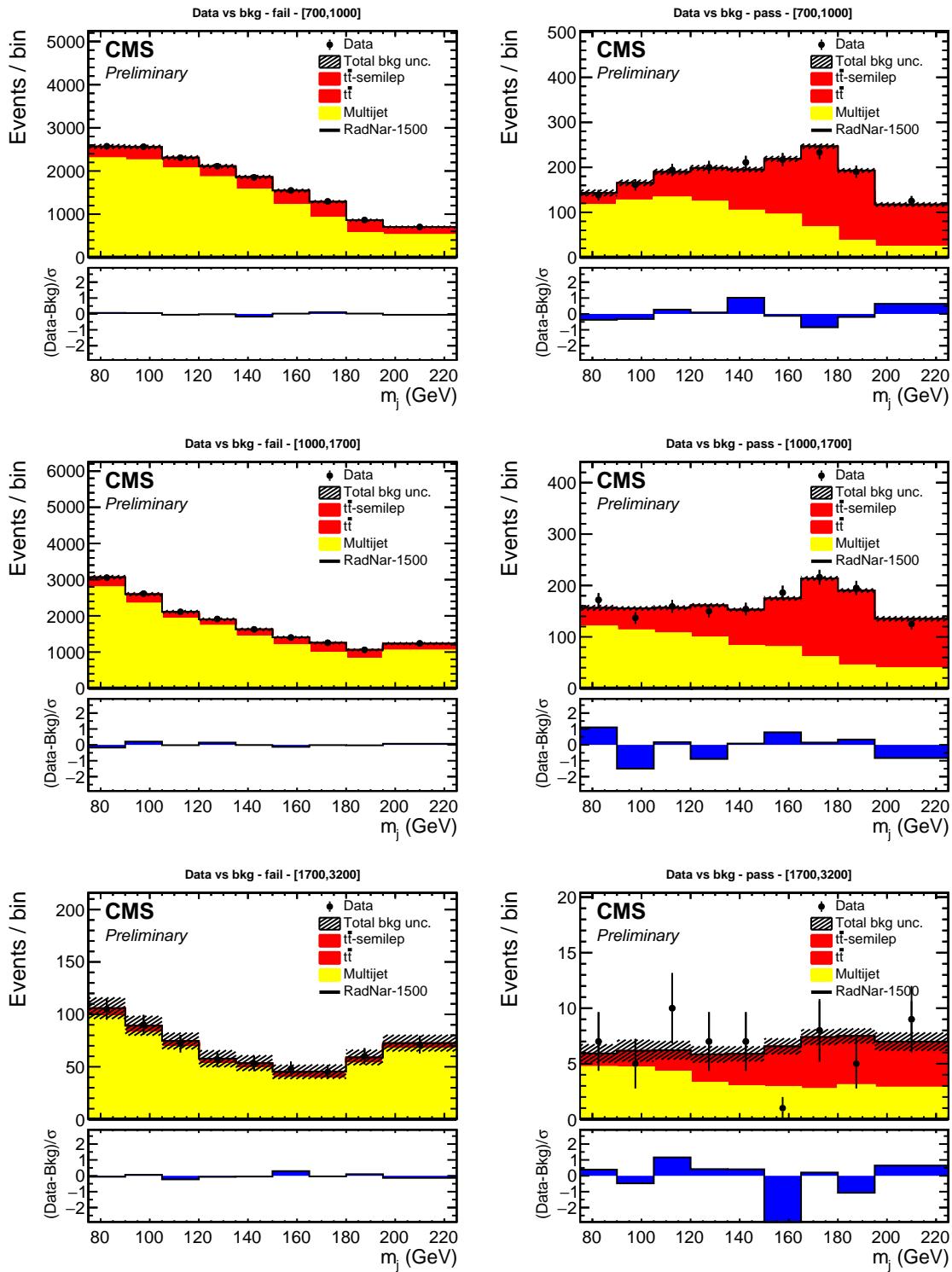


Figure 6-13. Full Run 2 Loose Loose fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

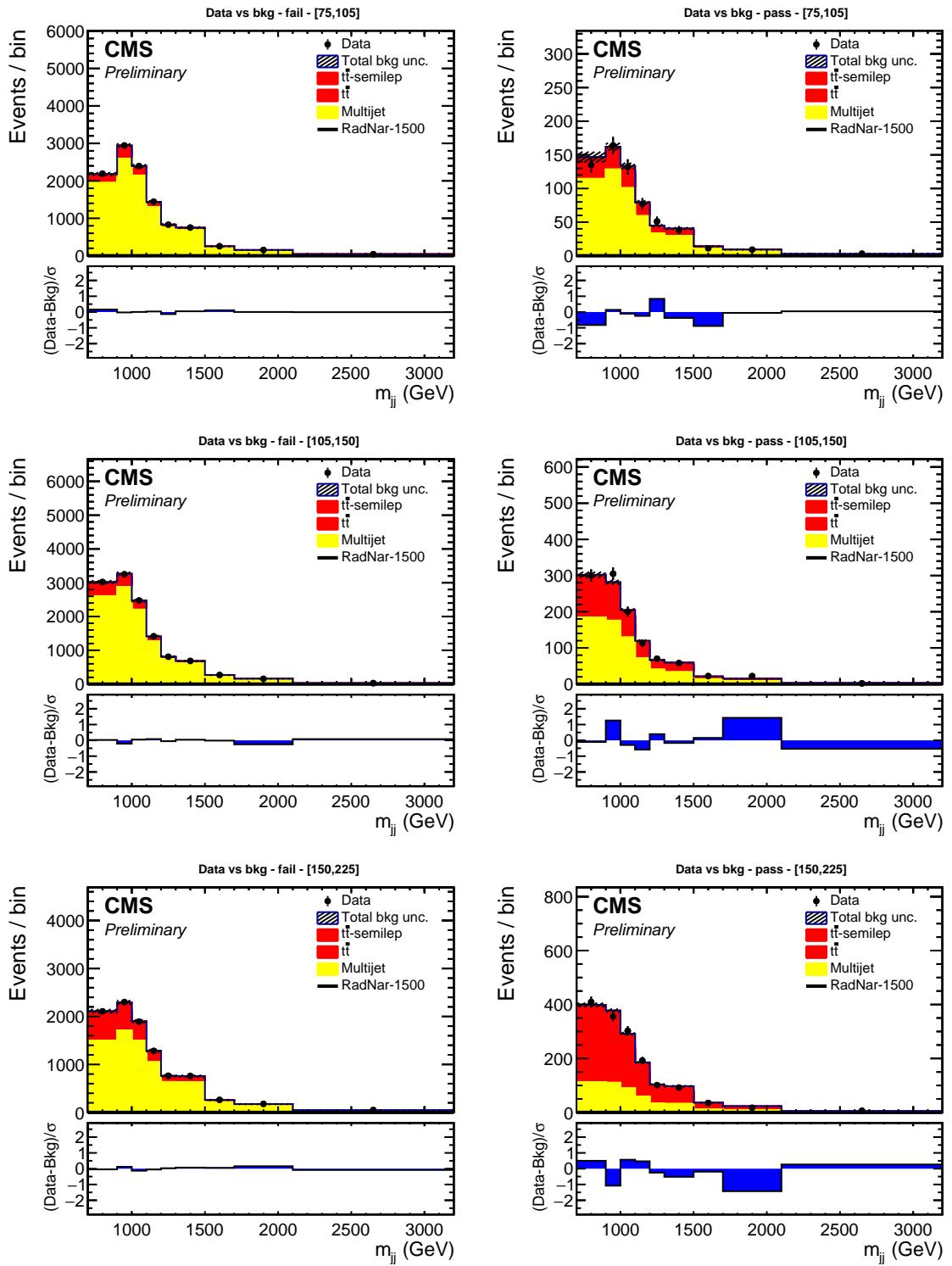


Figure 6-14. Full Run 2 Loose Loose fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

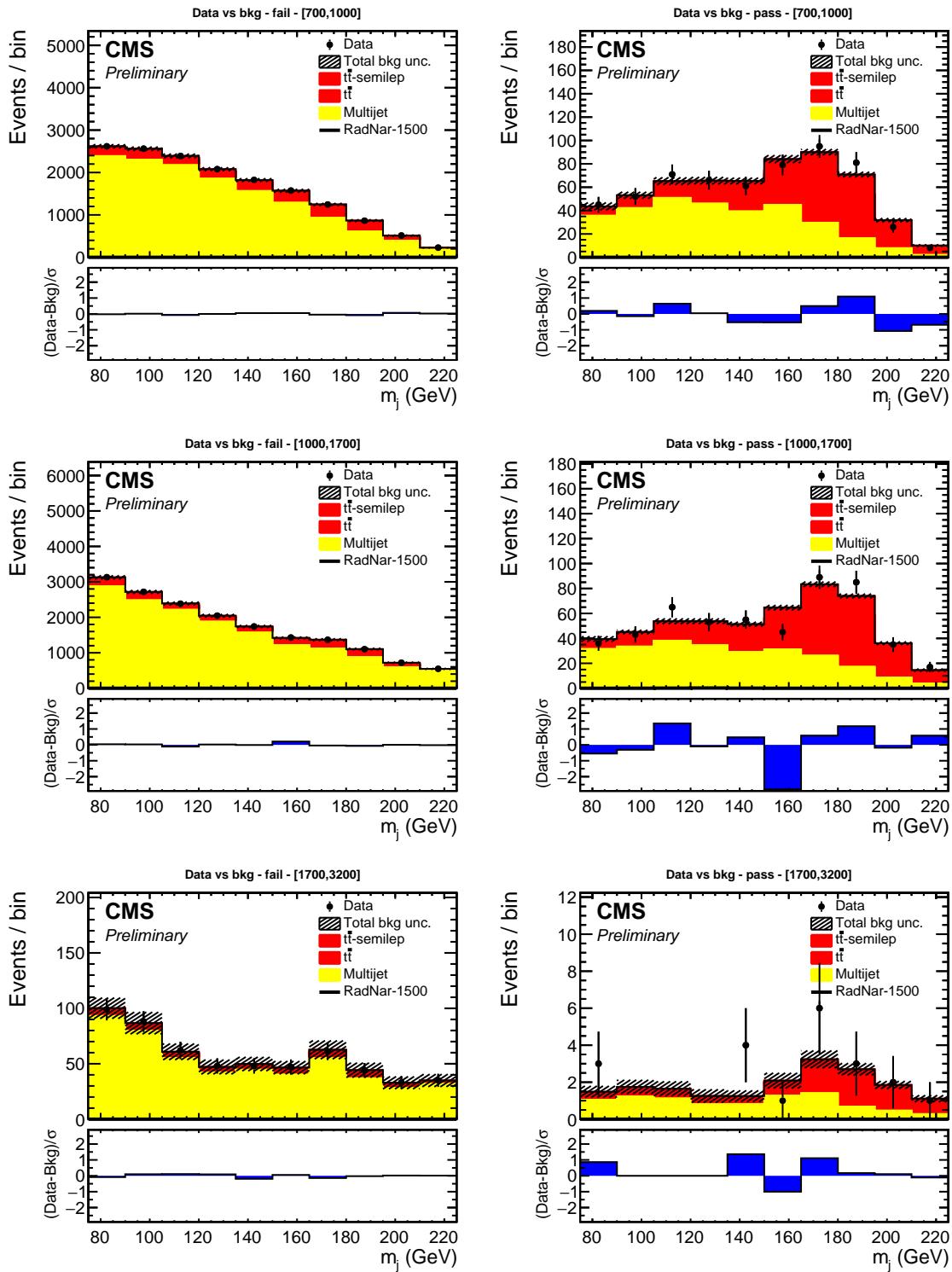


Figure 6-15. Full Run 2 Tight Tight fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

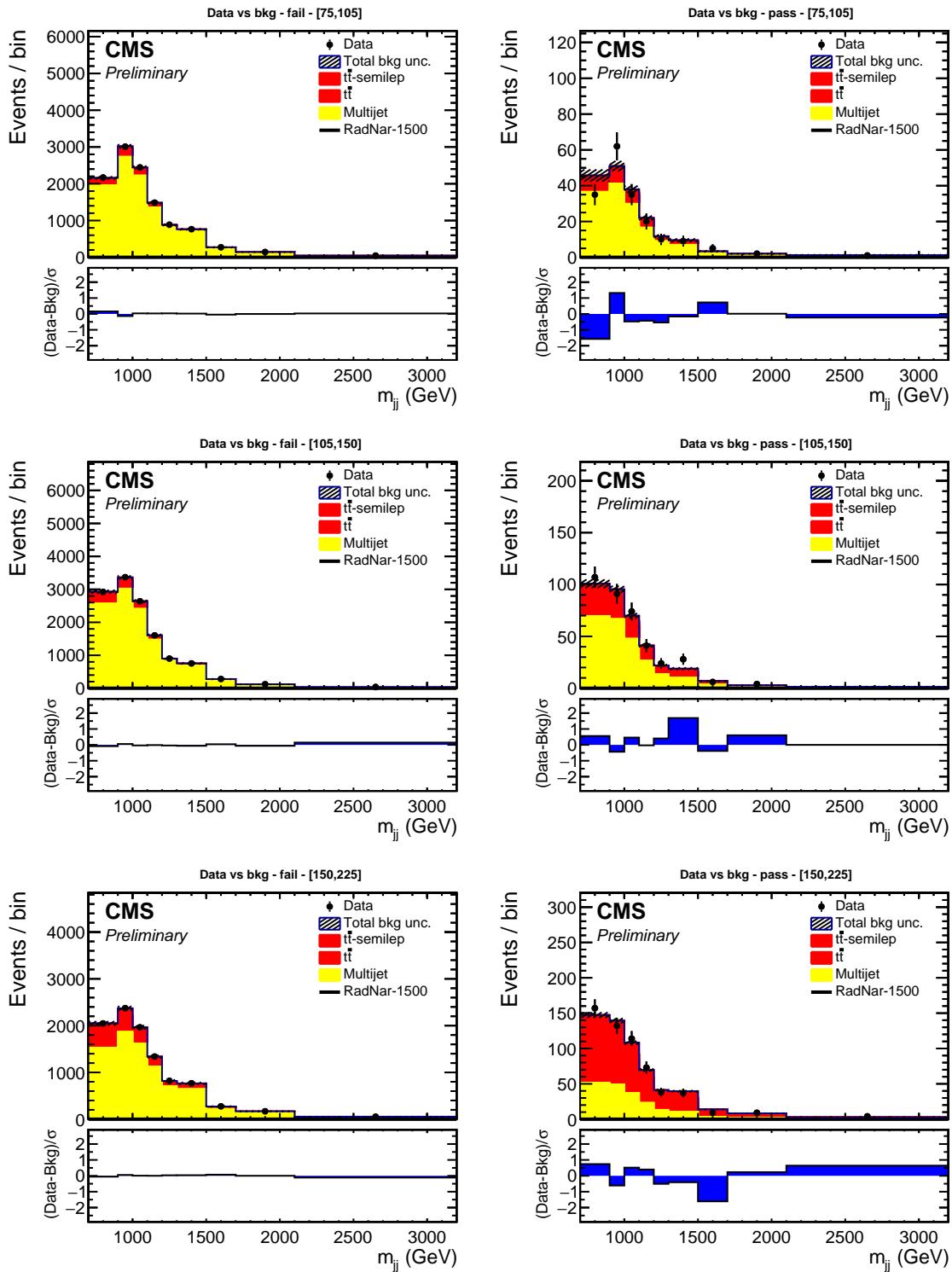


Figure 6-16. Full Run 2 Tight Tight fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

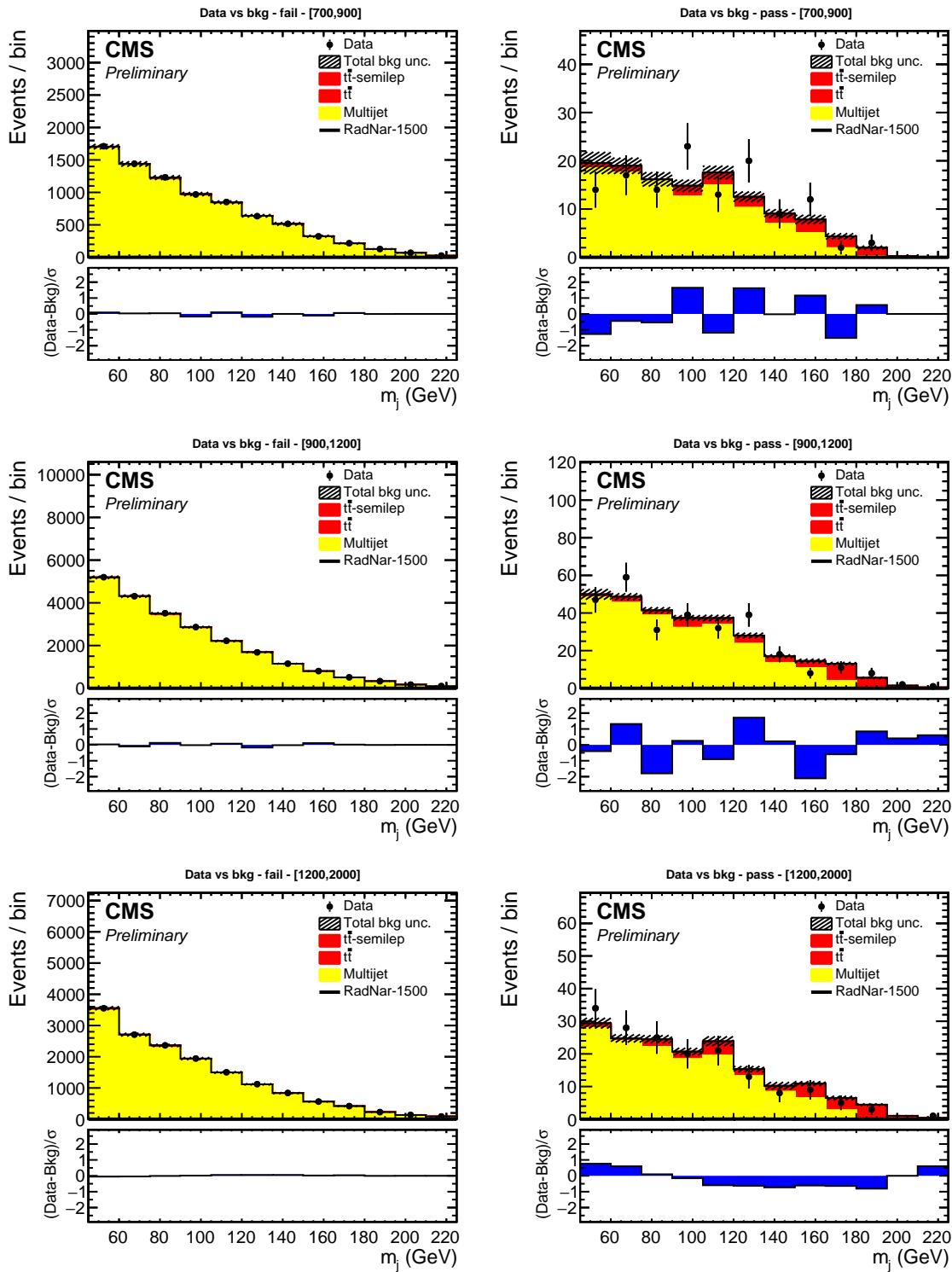


Figure 6-17. Full Run 2 2+1 fits for M_j axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

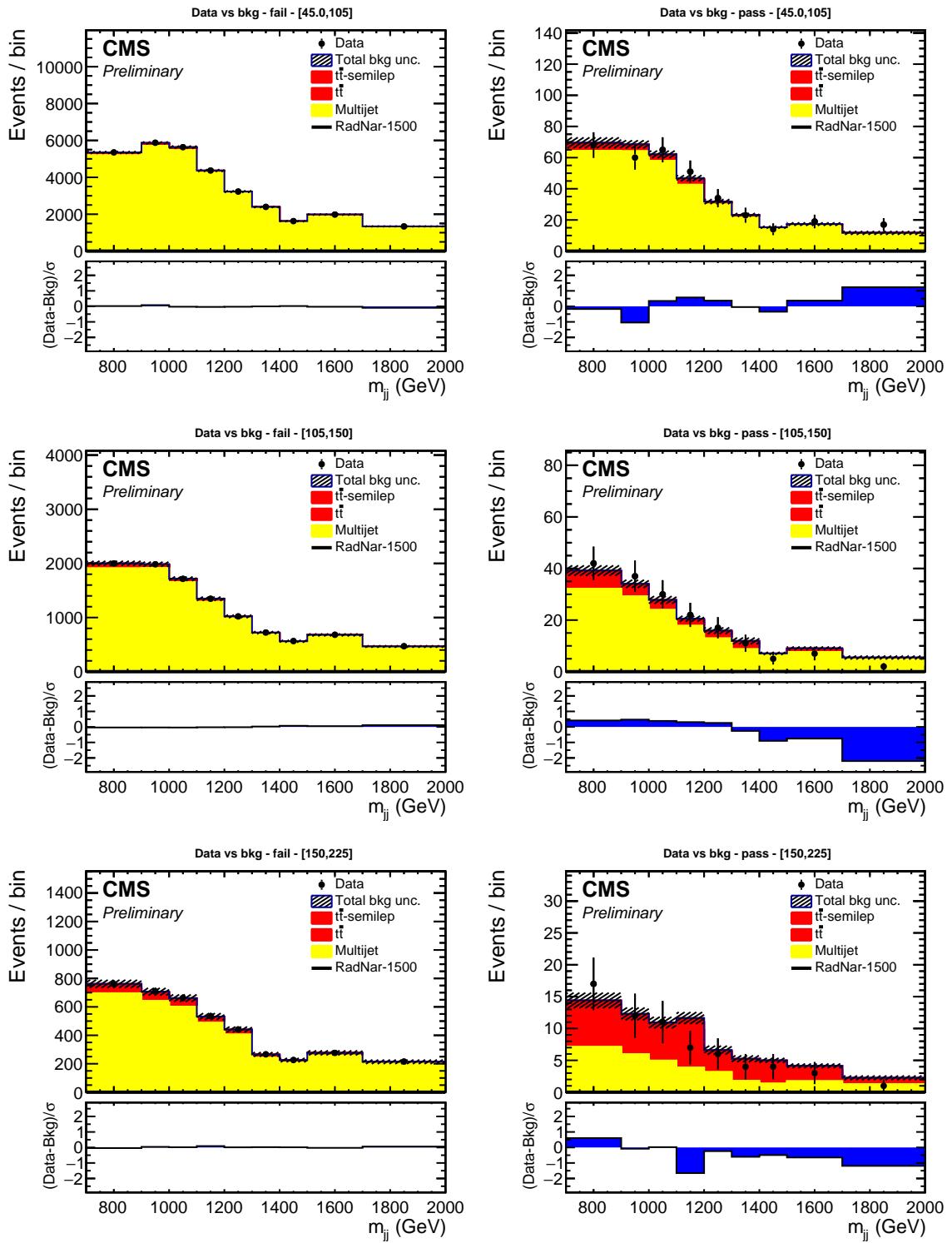


Figure 6-18. Full Run 2 2+1 fits for M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

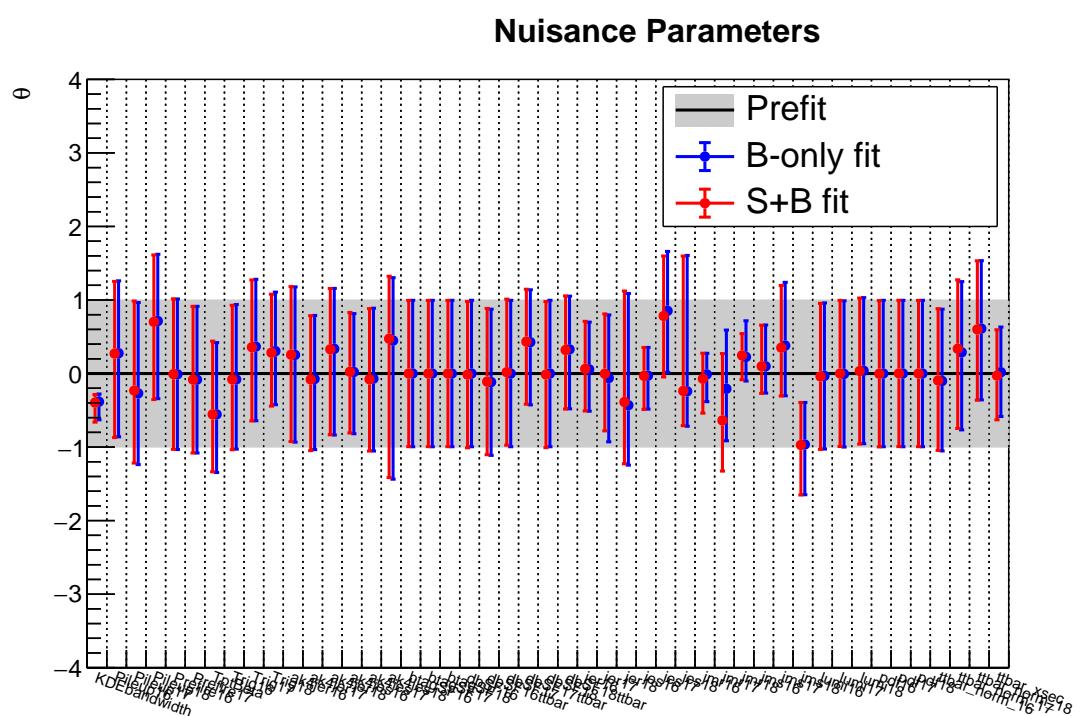


Figure 6-19. Full Run 2 Nuisance Pulls Plot.

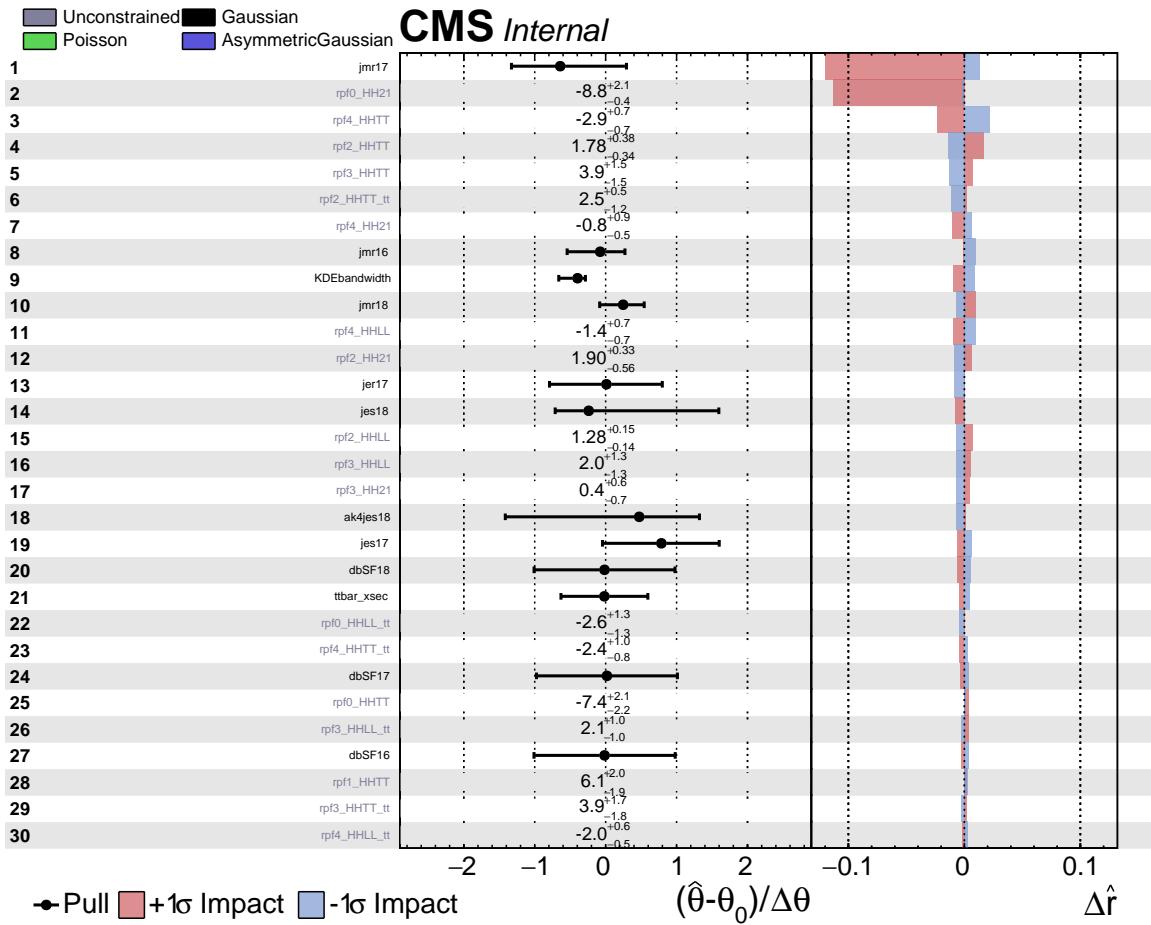


Figure 6-20. Full Run 2 Impacts plot. The impact of a nuisance parameter moving in one direction or another is in some sense a convention. A nuisance parameter that pulls “negative” can always be redefined such that the pull will be positive. What one should do is to check to make sure all the nuisance parameters “pull” in the correct direction relative to their definition.

6.2 Systematic Uncertainties

Each of the sources of the systematic uncertainty listed below is represented by at least one nuisance parameter in the fit, controlling either a normalization or a shape of the signal or the $t\bar{t}$ +jets background. The background estimation of QCD is unaffected by them because it is computed entirely from data, which brings with it different procedural sources of systematic uncertainty as described in the following text. We first present the uncertainties that impact the normalization of the signal and $t\bar{t}$, followed by those which impact the background estimate.

- **Luminosity:** An uncertainty of 2.5% [36] is applied to 2016 and 2018 and 2.3% is applied to 2017.
- **Pileup:** An uncertainty up to 5% for signal and 1% for $t\bar{t}$ is associated with pileup impact on M_{jj}^{red} by varying the estimated minimum bias cross section of pp collisions at 13 TeV ($= 69.2$ mb) by $\pm 4.6\%$ [37].
- **PDF and scale hypotheses impact:** To estimate the uncertainty from parton distribution functions, we evaluate the RMS of the distribution of the PDF MC replicas. For Hessian sets we compute the square root of the sum of differences squared. Each is used as the $\pm 1\sigma$ shape variation due to parton distribution functions for their respective sets. Typical uncertainties are $< 1\%$

The PDF set for each sample is as follows:

- 2016 All Signal Samples: MMHT2015qed_nlo_inelastic (MC-Set)
 - 2017 All Signal Samples: NNPDF31_nnlo_hessian_pdfs (Hessian)
 - 2018 All Signal Samples: NNPDF31_nnlo_hessian_pdfs (Hessian)
- **$t\bar{t}$ Cross Section Uncertainty:** The total uncertainty, calculated as the sum in quadrature of the scale uncertainty and the PDF+ α_S uncertainty on the cross

section of $t\bar{t}$ is applied, amounting to 6%, as prescribed by the TOP PAG group:

$$\sigma_{t\bar{t}} = 831.76 + 19.77 - 29.20 \text{ (scale)} + 35.06 - 35.06 \text{ (PDF+alpha s) pb.}$$

- **Trigger efficiency uncertainty:** The trigger strategy is described in Section 3.4. The uncertainty of the scale factor is treated as a shape based uncertainty and the template shapes are derived using the maximum of either $\Delta TriggerEff = 0.05 * (1.0 - jetTriggerWeight)$ or the Clopper-Pearson error derived from the trigger efficiency for each jet to calculate the $\pm 1\sigma$ distributions.
- **Top p_T re-weighting:** In order to account for differences between the shapes of the measured and simulated top p_T spectra, due to the absence of NNLO in the simulation, for $t\bar{t}$ samples [38], we nominally re-weight our $t\bar{t}$ MC samples using the TOP groups p_T -dependent scale factor. For a given event, we apply the nominally derived weight as:

$$SF(p_t) = e^{(\alpha - \beta p_T)} \quad (6.6)$$

and the nominal weight is given as: $W = \sqrt{SF(t)SF(\bar{t})}$ and the nominal parameters are $\alpha = 0.0615$ and $\beta = 0.0005$. Then to create the 1σ up and down templates we vary α and β up and down independently giving us 4 up and down weights. We use a 1σ upper uncertainty of 2.0 times the nominal weight and a 1σ lower uncertainty bound of 0.5 times the nominal weight when varying α and β .

- **dak8MDHbb-tagging:** Scale factors for the dak8MDHbb tagger are computed similarly to the double-b tagger scale factors¹. We create a template for each p_T bin range and the values are as follows:
- **$t\bar{t}$ -tagging** The $t\bar{t}$ deepAK8 tagger scale factor has been split into SF and normalization and both are applied to the $t\bar{t}$ distributions. The DeepAK8-MD

¹<https://indico.cern.ch/event/853828/contributions/3723593/attachments/1977626/3292045/lgbtv-deepak8v2-sf-20200127.pdf>

Table 6-I. List of Scale Factors for LP [0.80-0.90] values of the Deep AK8 MD Hbb Tagger

Year	200-500	500-600	600-Inf
2016	1.260 $-0.118/+0.103$	1.113 $-0.102/+0.100$	0.954 $-0.064/+0.063$
2017	1.364 $-0.194/+0.179$	1.114 $-0.108/+0.119$	1.114 $-0.055/+0.090$
2018	1.233 $-0.5543/+0.269$	0.975 $-0.115/+1.025$	0.979 $-0.058/+0.072$

Table 6-II. List of Scale Factors for HP [0.90-Inf] values of the Deep AK8 MD Hbb Tagger

Year	200-500	500-600	600-Inf
2016	0.951 $-0.045/+0.047$	0.974 $-0.044/+0.046$	1.049 $-0.051/+0.045$
2017	0.937 $-0.064/+0.114$	0.922 $-0.054/+0.060$	0.914 $-0.089/+0.091$
2018	0.966 $-0.091/+0.101$	0.958 $-0.065/+0.068$	0.920 $-0.080/+0.089$

(bb) mis-tagging scale factors (SF) for top events are derived in a single-muon sample dominated by $t\bar{t}$ events. See section 6.2.1 for details.

Table 6-III. List of $t\bar{t}$ Scale Factors for the Deep AK8 MD Hbb Tagger

Year	p_T Range [GeV]	Scale Factor	Normalization
2016	300 - 600	$1.039^{+0.061}_{-0.058}$	$0.72^{+0.05}_{-0.05}$
2016	600 - 800	$1.035^{+0.105}_{-0.098}$	$0.65^{+0.06}_{-0.06}$
2016	> 800	$1.301^{+0.325}_{-0.266}$	$0.52^{+0.07}_{-0.07}$
2017	300 - 600	$0.91^{+0.05}_{-0.05}$	$0.85^{+0.06}_{-0.06}$
2017	600 - 800	$0.93^{+0.11}_{-0.09}$	$0.87^{+0.08}_{-0.08}$
2017	> 800	$1.07^{+0.28}_{-0.25}$	$0.74^{+0.09}_{-0.09}$
2018	300 - 600	$0.89^{+0.04}_{-0.05}$	$0.83^{+0.06}_{-0.06}$
2018	600 - 800	$0.94^{+0.08}_{-0.08}$	$0.89^{+0.08}_{-0.08}$
2018	> 800	$1.05^{+0.21}_{-0.19}$	$0.86^{+0.09}_{-0.09}$

- **DeepJet b-tagging:** This is treated as a shape based uncertainty where the nominal and $\pm 1\sigma$ distributions are calculated with the recommended BTagCalibration reader class.²
- **Jet Energy Scale:** To estimate the uncertainty in the jet energy scale, we use the $\pm 1\sigma$ variations on the jet energy scale derived as a shape based uncertainty

²For 2016:<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation2016Legacy>
For 2017:<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation94X> For 2018:<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation102X>

from jet energy corrections for AK8/AK4 PFPUPPI jets. The variations are η and p_T dependent, and result is a scaling of the jet four momentum prior to the analysis selection.

- **Jet Energy Resolution:** We apply the approved jet energy resolution correction, which is η dependent and is used to correct for differences in p_T resolution between data and Monte Carlo. The recommended $\pm 1\sigma$ variations are also taken into account³ to make a shape based uncertainty.
- **Jet Mass Scale:** We use the $\pm 1\sigma$ variations on the jet mass scale derived as a shape based uncertainty and do not apply the nominal jet mass correction as it is not appropriate for this analysis.
- **Jet Mass Resolution:** We use $\pm 1\sigma$ variations of the jet mass resolution as a shape based uncertainty. We also do not apply the nominal jet mass resolution correction as it is not appropriate. Examples of all JECs can be seen in appendix **FIXME**

- **Multijets background from the data:** The main source of uncertainty for the non-top background is propagated directly from the error of the fit in mass-sidebands from which we have estimated the transfer factor. This uncertainty can be treated as a shape based uncertainty. We must further account for the statistical uncertainty in the anti-tag region which is propagated to the signal region when the estimate is made.
- **KDE Bandwidth:** The bandwidth used for to smooth qcd has an uncertainty that is treated as shape based uncertainty where the $\pm 1\sigma$ distributions are made from smoothed qcd distributions made with the value ± 1 used as the bandwidth. Figure 6-21 shows the up,down, and nominal distributions as an example.

³<https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution> and stored in a private github repo here: <https://github.com/lcorcodilos/nanoAOD-tools/tree/master> where we updated to the latest v7

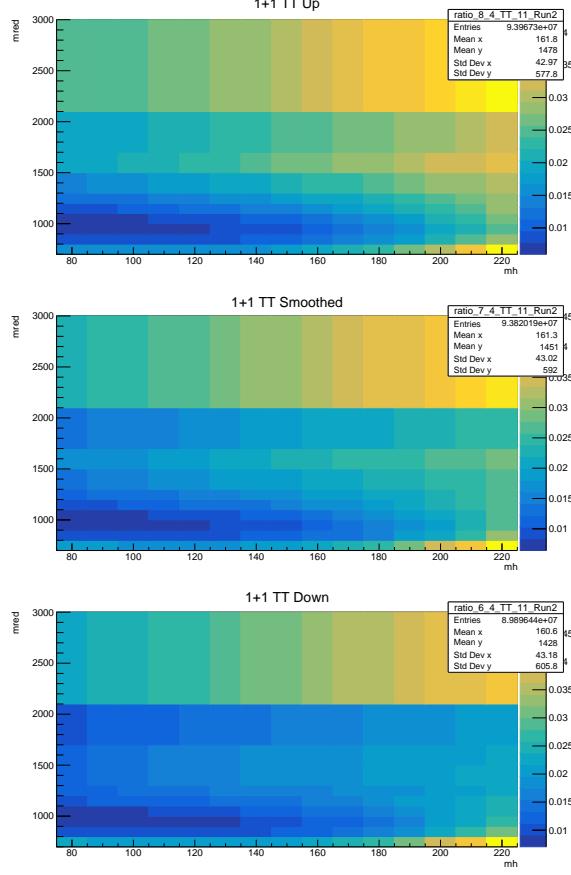


Figure 6-21. KDE bandwidth nominal, up, and down distributions for Tight Tight.

- **Prefire Corrections** We apply the recommended 2016/2017 prefire corrections using the nanoAOD tools module to create a nominal correction and $\pm 1\sigma$ distributions to be used as a shape based uncertainty.

6.2.1 $t\bar{t}$ -Tagging Scale Factor Derivation

The measurement of the t quark mis-tagging efficiency in data is performed using a “tag-and-probe” method [39]. The muon, in combination with the b-tagged jet, is used as the tag”. In the opposite hemisphere of the event, the jet is considered as the “probe jet”.

The total SM sample is decomposed into three categories based on the spatial nanoAOD tools.

separation of the partons from the t quark decay with respect to the AK8 jet. The merged “t quark” category includes cases where the three partons and the jet have $\Delta R < 0.6$. The merged “W boson” category includes cases where only the two partons from the W boson decay are within $\Delta R < 0.6$ of the jet and the b quark from the top quark decay is outside the jet cone. Any other topology falls in the “non-merged” category.

The jet mass distributions in simulation of each one of the three categories are used to derive templates to fit the jet mass distribution in data. For a given working point, the fit is done for all three categories simultaneously for both the “passing” and “failing” events. The fit is performed in the range from 50 to 250 GeV with a bin width of 10 GeV.

A number of sources of systematic effects can affect the modeling of the performance of the algorithms in data by the simulation. These include systematic uncertainties in the parton showering model, renormalization and factorization scales, PDFs, jet energy scale and resolution, p_T^{miss} un-clustered energy, trigger and lepton identification, pileup modeling, and integrated luminosity, as well as statistical uncertainties of simulated samples.

Parton shower uncertainties are evaluated using samples with the same event generators but a different choice for the modeling of the parton showering (i.e. pythia vs. herwig). Changes in renormalization (μ_R) and factorization (μ_F) scales are estimated by varying the scales separately by a factor of two up and down, relative to the choices of the scale values used in the sample generation. The uncertainty related to the choice of the PDFs is obtained from the standard deviation in 100 replicas of the NNPDF3.0 PDF set [40]. The jet energy scale and resolution are changed within their p_T - and η -dependent uncertainties, based on the studies presented in Ref. [41]. Their effects are also propagated to p_T^{miss} . The effect of the uncertainty in the measurement of the un-clustered energy (i.e., contribution of PF candidates not associated to

any of the physics objects) is evaluated based on the momentum resolution of each PF candidate, which depends on the type of the candidate [42]. Uncertainties in the measurement of the trigger efficiency and in the energy/momentum scale and resolution of the leptons are propagated in the SF extraction. The uncertainty in the pileup weighting procedure is determined by varying the minimum bias cross section used to produce the pileup profile by $\pm 5\%$ from the measured central value of 69.2 mb [43, 44]. The limited size of the simulated samples and the size of the data control samples are also considered.

A logNormal uncertainty of the 10% is applying to account for possible missing effects on the normalization of the three templates obtained from the $t\bar{t}$ sample. An uncertainty of 50% is applied to the other subdominant contributions.

All sources of systematic uncertainties are treated as nuisance parameters in the fit. After calculating the efficiencies in data (ϵ_{Data}) and simulation ($\epsilon_{\text{Simulation}}$), the SF is determined as the ratio of ϵ_{Data} over $\epsilon_{\text{Simulation}}$.

The $t\bar{t}$ normalization and the SFs are extracted deferentially in jet p_T for the t quark tagging working point: $\text{deepAK8-MD(bb)} > 0.90$. The following exclusive jet p_T regions are considered: 450–600, 600–800, and 800–1200 GeV.

Chapter 7

HH4b: Results

7.1 Results

Upper limits at 95% confidence level are set on the product of the production cross section and the branching fractions $\sigma(pp \rightarrow X)\mathcal{B}(X \rightarrow b\bar{b}b\bar{b})$. They are obtained using the profile likelihood as a test statistic [45]. The systematic uncertainties are treated as nuisance parameters and are profiled in the minimization of the negative of the logarithm of the profile likelihood ratio and the distributions of the likelihood ratio are calculated using the asymptotic approximation [46] of the procedure reported in Refs. [47, 48].

As shown in Fig. 7-5, left, a narrow radion with mass between 1000 and 2600 GeV is excluded at 95% confidence level for $\Lambda_R = 3$ TeV. Narrow bulk graviton for $k/\overline{M_{Pl}} = 0.5$ is excluded at 95% confidence level only for masses between 1000 and 1200 GeV, as shown in Fig. 7-5, right. The deviation in observed limit at 1300 and 1500 GeV resonance masses is driven by an upward fluctuation of data over the background prediction at $M_{jj}^{red} \sim 1400$ GeV, as can be seen in Fig. 6-16, middle row.

The expected limits are reported in Tables 7-I and 7-II for radions and bulk gravitons of different assumed masses, respectively. The upper limits range from 4.94 to 0.19 fb for the bulk graviton and from 9.74 to 0.29 fb for the radion for the mass range 1–3 TeV.

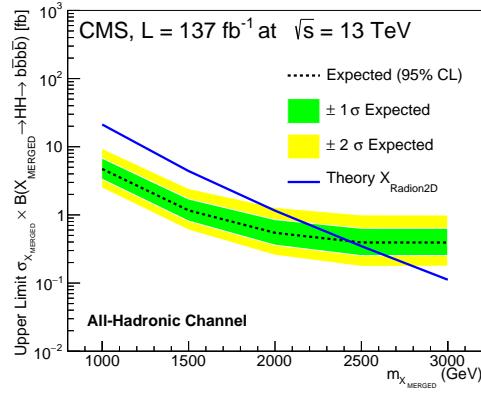


Figure 7-1. Full Run 2 Radion Limit for 1+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.

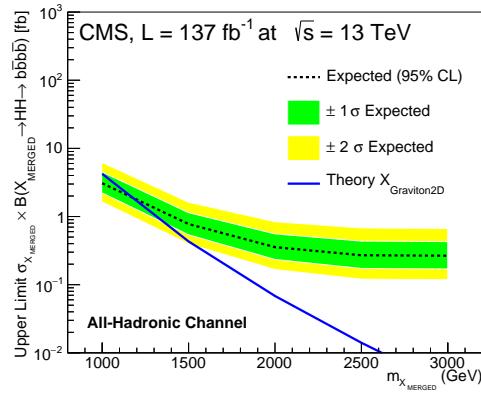


Figure 7-2. Full Run 2 Graviton Limit for 1+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.

Table 7-1. Radion observed 95% CL exclusion limits.

Mass (GeV)	Obs. (fb)	Exp. lim. (fb)	+Exp (68%) (fb)	-Exp (68%) (fb)	+Exp (95%) (fb)	-Exp (95%) (fb)
1000	9.74	5.12	7.41	3.68	10.26	2.74
1100	3.37	3.20	4.58	2.28	6.31	1.69
1200	2.80	2.20	3.18	1.55	4.40	1.15
1300	3.24	1.72	2.50	1.21	3.5	0.89
1500	1.46	1.10	1.75	0.83	2.48	0.61
2000	0.47	0.53	0.83	0.34	1.26	0.24
2500	0.26	0.42	0.66	0.26	1.02	0.18
3000	0.29	0.46	0.74	0.29	1.13	0.20

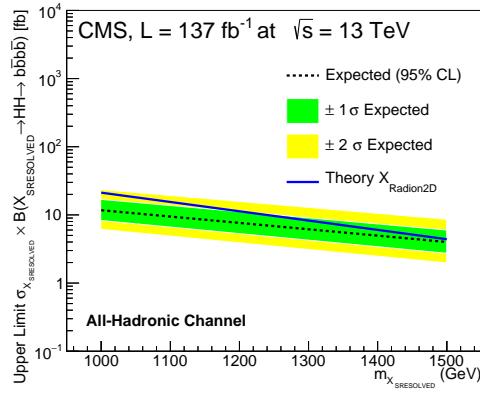


Figure 7-3. Full Run 2 Radion Limit for 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.

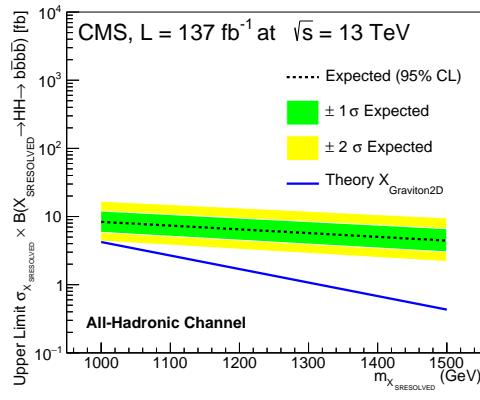


Figure 7-4. Full Run 2 Graviton Limit for 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.

Table 7-II. Graviton observed 95% CL exclusion limits.

Mass (GeV)	Obs. (fb)	Exp. lim. (fb)	+Exp (68%) (fb)	-Exp (68%) (fb)	+Exp (95%) (fb)	-Exp (95%) (fb)
1000	4.94	3.47	4.93	2.46	6.82	1.81
1100	2.13	2.12	3.10	1.52	4.26	1.14
1200	1.42	1.54	2.20	1.10	3.07	0.79
1300	1.76	1.20	1.71	0.82	2.42	0.62
1500	1.37	0.83	1.21	0.57	1.71	0.42
2000	0.31	0.35	0.55	0.23	0.83	0.16
2500	0.18	0.28	0.45	0.18	0.69	0.13
3000	0.19	0.31	0.49	0.20	0.77	0.14

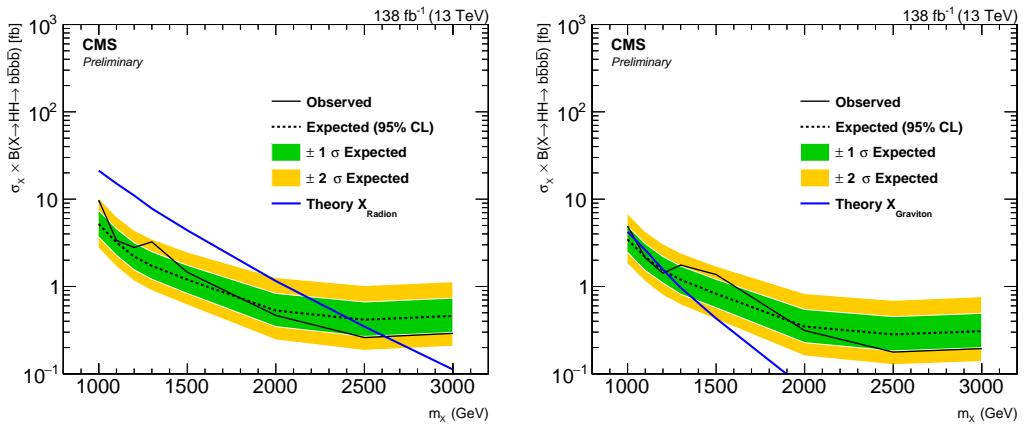


Figure 7-5. Upper limits at 95% confidence level on $\sigma(pp \rightarrow X)\mathcal{B}(X \rightarrow b\bar{b}b\bar{b})$ for the narrow spin-0 radion (left) and the spin-2 bulk graviton (right) models. The predicted theoretical cross sections for the narrow radion and bulk graviton are also shown.

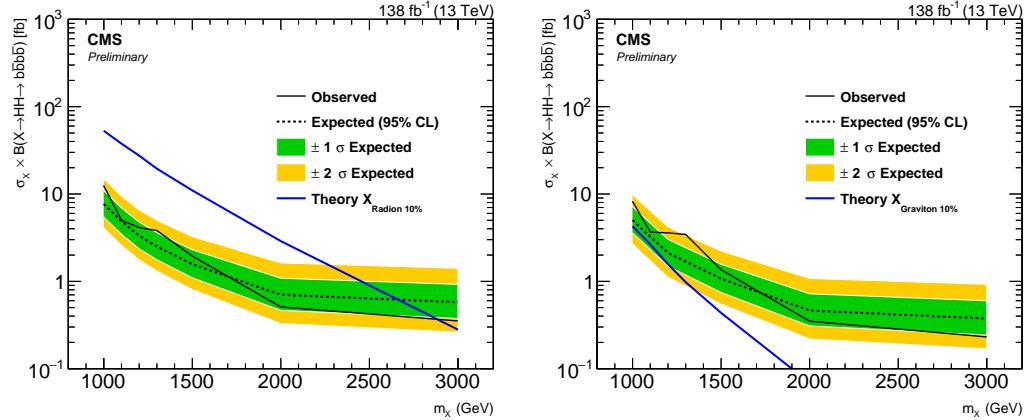


Figure 7-6. Upper limits at 95% confidence level on $\sigma(pp \rightarrow X)\mathcal{B}(X \rightarrow b\bar{b}b\bar{b})$ for the 10% width spin-0 radion (left) and the spin-2 bulk graviton (right) models.

Table 7-III. 10%-width Radion observed 95% CL exclusion limits.

Mass (GeV)	Obs. (fb)	Exp. lim. (fb)	+Exp (68%) (fb)	-Exp (68%) (fb)	+Exp (95%) (fb)	-Exp (95%) (fb)
1000	12.48	7.70	10.96	5.47	15.00	4.10
1100	5.01	4.88	6.94	3.47	9.50	2.59
1200	4.13	3.33	4.76	2.36	6.57	1.76
1300	3.82	2.51	3.60	1.77	5.00	1.31
1500	1.95	1.57	2.30	1.10	3.25	0.80
2000	0.51	0.70	1.08	0.47	1.63	0.33
3000	0.35	0.58	0.92	0.37	1.41	0.26

Table 7-IV. 10%-width Graviton observed 95% CL exclusion limits.

Mass (GeV)	Obs. (fb)	Exp. lim. (fb)	+Exp (68%) (fb)	-Exp (68%) (fb)	+Exp (95%) (fb)	-Exp (95%) (fb)
1000	8.23	5.03	7.15	3.59	9.86	2.67
1100	3.67	3.25	4.64	2.32	6.38	1.73
1200	3.59	2.12	3.02	1.49	4.22	1.11
1300	3.44	1.68	2.42	1.18	3.37	0.88
1500	1.36	1.07	1.57	0.75	2.22	0.55
2000	0.35	0.46	0.72	0.31	1.08	0.22
3000	0.23	0.36	0.60	0.24	0.92	0.17

Chapter 8

Extensions of the HH4b Search

8.1 VBF Production of HH in the 4b Final State

The previously described analysis of boosted HH in the 4b final state focuses on the dominant HH production mode, gluon-gluon fusion (ggF). This is not the only production mode that can be studied. The next most common production mode is through vector boson fusion (VBF). The importance of this production mode is that the leading order diagrams provide a way to measure the Higgs self-coupling (HHH), the vector boson coupling (VVH), and the quartic HHVV coupling. These are called, κ_λ , κ_V , and κ_{2V} respectively. While the other two couplings, κ_λ and κ_V , are able to be probed by other methods, the quartic coupling, κ_{2V} is uniquely probed by the VBF production mode. According to the SM, all of these coupling should all be equal to 1.

At energy scales available at the LHC, the leading contribution for longitudinal

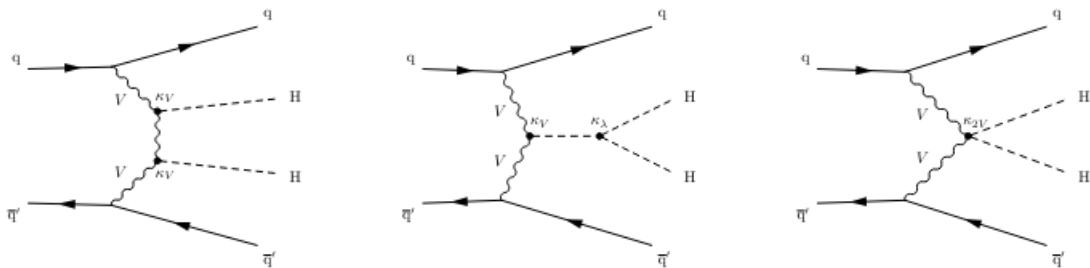


Figure 8-1. Feynman diagrams for Vector Boson Fusion production of HH in the 4b Final State.

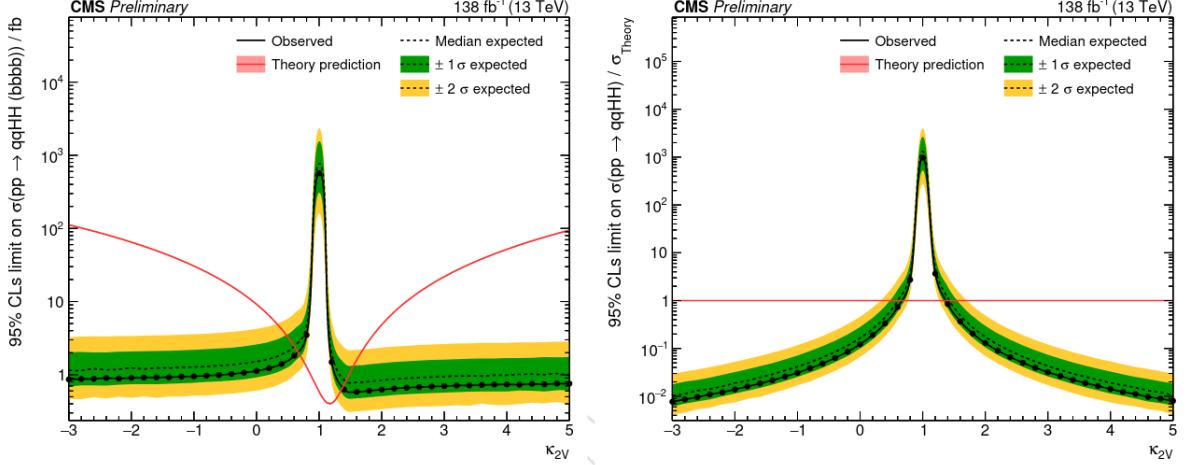


Figure 8-2. Expected and observed 95% CL exclusion limit on the product of the HH production cross section and the branching fraction into the 4b final state, as a function of the 2V coupling, obtained using the Asimov dataset. The crossings of expected median (dashed black line) and the theoretical cross section (red line) indicate the expected range of the κ_{2V} values to be excluded at 95% CL

scattering amplitude of the VBF production is found to be proportional to m_{HH} times $\kappa_{2V} - \kappa_V^2$, where m_{HH} is the invariant mass of the di-Higgs system. This means that in the SM, the two leftmost diagrams of Figure 8-1 cancel each other almost completely (since $\kappa_{2V} = \kappa_V = 1$), yielding a small production cross section. However, in more generic models where $\kappa_{2V} \neq 1$ and/or $\kappa_V \neq 1$, this is not the case and the VBF cross section is dramatically increased, especially in phase space region characterized by large di-Higgs invariant mass.

This analysis then attempts to set stricter limits on κ_{2V} by using a novel, powerful multivariate classifier based on graph neural networks, the *ParticleNet* jet tagger. In addition to flavor tagging, the ParticleNet DNN architecture is also applied to train a jet mass regression algorithm, allowing efficient background suppression based on the regressed mass of the $H \rightarrow bb$ candidate jets. The statistical tests and systematic uncertainties used in this analysis are adapted in part or whole from the HH to 4b search.

Here we show the limits placed on κ_{2V} .

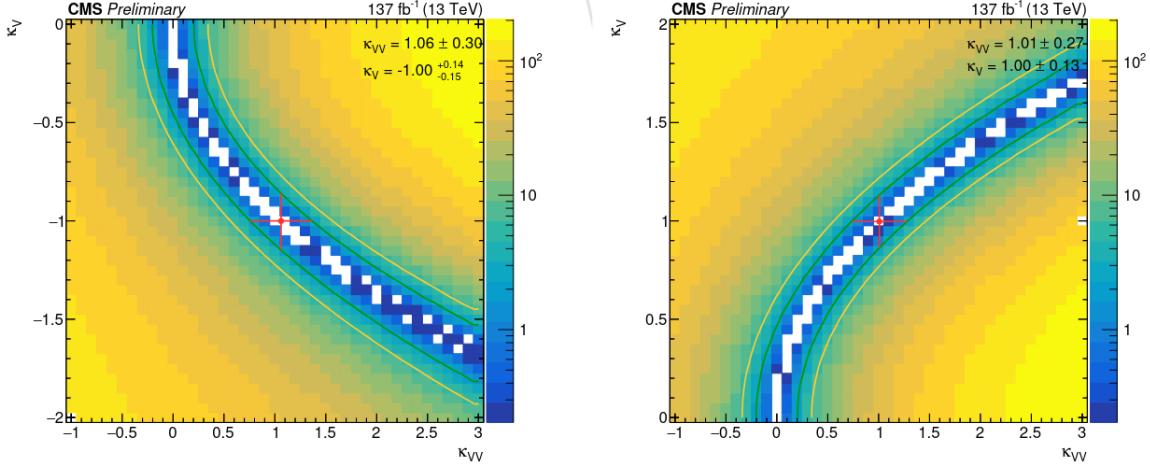


Figure 8-3. Two-dimensional likelihood scan where both κ_{2V} and κ_V couplings are varied simultaneously, for $\kappa_V < 0$ (left) and $\kappa_V > 0$ (right)

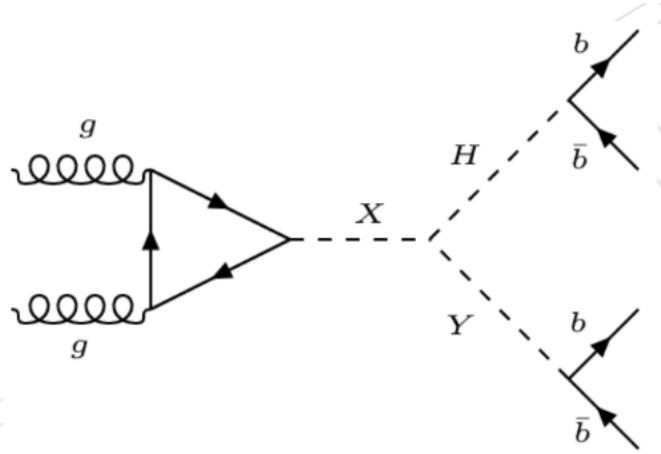


Figure 8-4. Feynman diagram for gluon gluon fusion production of HY in the 4b Final State.

8.2 A Massive Resonance decaying to a light scalar and a Higgs Boson in the 4b Final State

One class of searches is the search for additional spin-0 (scalar) particles. These are motivated in the two-Higgs doublet models which extend the SM scalar sector by one additional scalar doublet. In particular, the type II two-Higgs doublet model such as the Minimal Supersymmetric extension of the standard model (MSSM) postulates the existence of extra CP even scalars (H and h) and CP-odd pseudo-scalars (A)

with sizable production cross sections. One of these could be the discovered Higgs boson H_{125} . The next-to-Minimal Supersymmetric extension of the standard model (NMSSM) contains an additional scalar complex singlet which mixes with the existing scalar doublets to give rise to extra CP-even and CP-odd states with masses below the scale of the heavier scalars. Thus, in these models, cascade decays such as $H \rightarrow H_{125}H_{125}$ or $H \rightarrow hH_{125}$ are possible, if kinematically favorable.

As the $H \rightarrow H_{125}H_{125}$ is the main search of this thesis, one can then probe the $H \rightarrow hH_{125}$ decay. The cascade production and decay processes involving two new scalars of unequal masses $pp \rightarrow X \rightarrow HY$, X being the heavier and Y being the lighter scalars, are yet to be explored at the LHC, though several searches are ongoing or planned. One of the main difficulties of such a search, compared with final states having two particles of equal masses, is the reconstruction of the two different mass particles. Overall, the physics signature will have two unknown masses, that of the parent massive resonance X and that of the undiscovered scalar Y. The H and Y candidates are selected by employing jet tagging techniques, including the previously mentioned ParticleNet, to identify jets consistent with the decay of a massive resonance into a pair of b quarks. The background estimate method, statistical tests, and systematic uncertainties used in this analysis are adapted in part or whole from the HH to 4b search.

As of the writing of this thesis, the final results of this HY to 4b search have not been made public so here we show the expected limits on both M_X and M_Y .

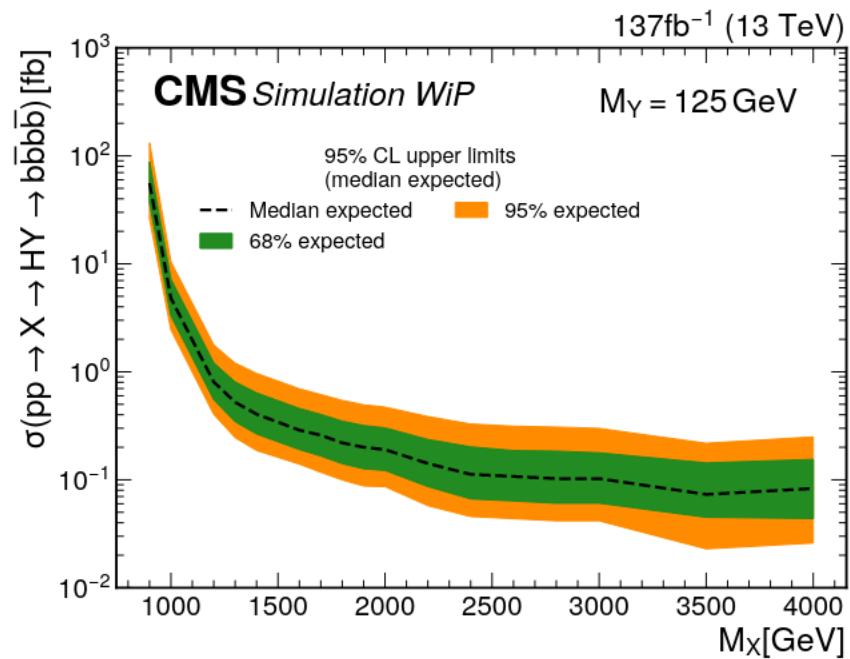


Figure 8-5. The expected upper limit of $pp \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at 95% confidence level for a fixed mass of Y , $M_Y = 125\text{GeV}$. Calculated using toy data for full RunII.

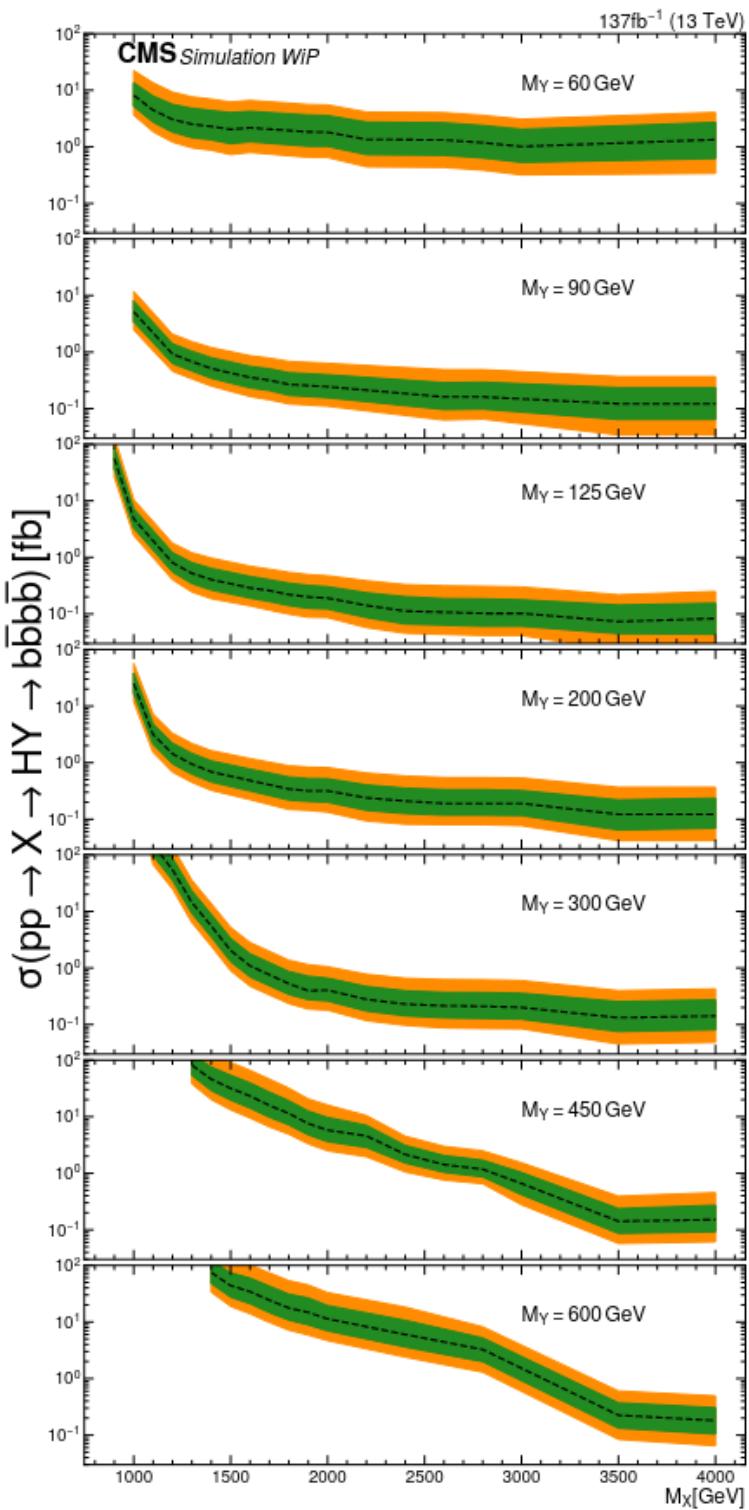


Figure 8-6. The expected upper limit of $pp \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at 95% confidence level for seven different M_Y . Limits are calculated using toy data for full RunII

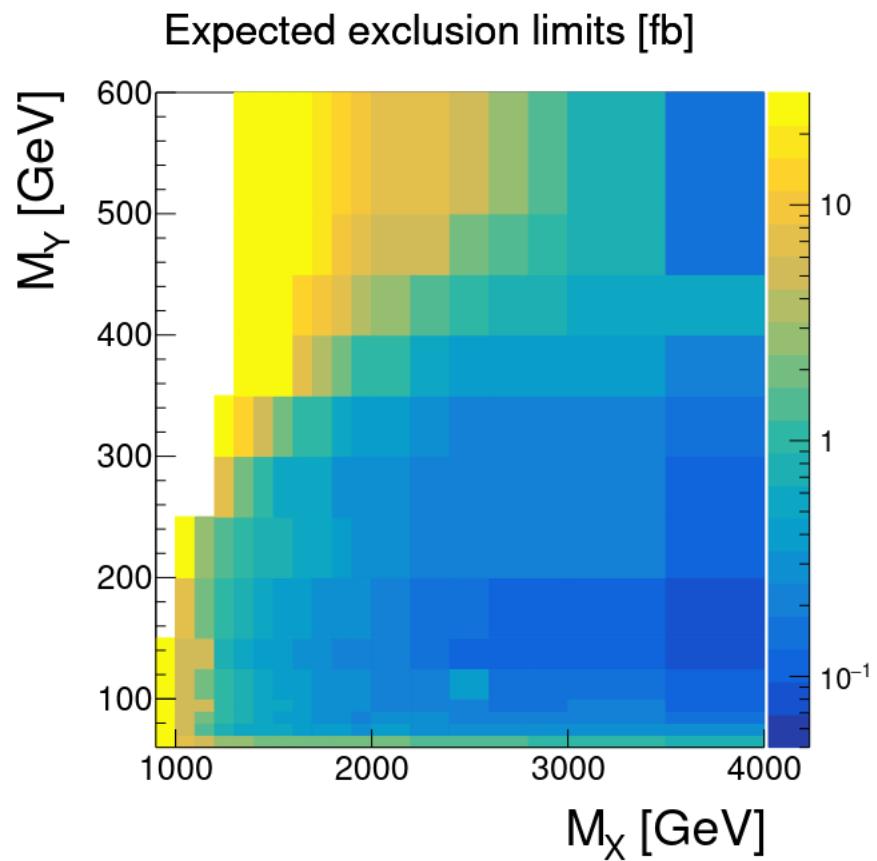


Figure 8-7. The expected upper limit of $pp \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at 95% confidence level in the M_X, M_Y plane. Limits are calculated using toy data for full RunII.

Conclusions

We have presented an analysis using the full Run 2 data probing several BSM models using the all hadronic channel of the $HH \rightarrow b\bar{b}b\bar{b}$ search. We have also presented two extensions of this search that compliment its predictive abilities. As we have previously stated, there are many BSM models and trying to check every possible prediction is quite a large task. The Higgs boson continues to be a unique lens into BSM models so there are many other extensions that have not been explored that should be in the future. For example, there are other decay channels for a Higgs boson. A search where one of the Higgs bosons first decays to two W bosons is possible. It is also possible to search for the CP-odd pseudo scalar, A , in a $Ha \rightarrow b\bar{b}b\bar{b}$ analysis. The search, $aa \rightarrow b\bar{b}b\bar{b}$, has recently been completed so $Ha \rightarrow b\bar{b}b\bar{b}$ would compliment the main analysis and the previously mentioned one nicely.

The main search of this thesis is not complete however. In Run 3 it is possible that enough data will be taken to directly measure the Higgs boson self coupling with the $HH \rightarrow b\bar{b}b\bar{b}$ search. This measurement is a good test of the SM in conjunction with the quartic coupling measurement that was presented in Section 8.1. One of the basic principles of science is to continuously verify that your theory is accurate. This accomplishes two main tasks. The first is it serves to push science to do an even more accurate job of describing our physical universe. The second is it provides the chief avenue for discovery as discrepancies between theory and experiment can sometimes widen into chasms of new understanding. In this approach, science never rests.

References

1. Giudice, G. F., Rattazzi, R. & Wells, J. D. Graviscalars from higher dimensional metrics and curvature Higgs mixing. *Nucl. Phys. B* **595**, 250. arXiv: [hep-ph/0002178 \[hep-ph\]](#) (2001).
2. Aad, G. *et al.* Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B* **716**, 01. arXiv: [1207.7214 \[hep-ex\]](#) (2012).
3. Chatrchyan, S. *et al.* Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B* **716**, 30. arXiv: [1207.7235 \[hep-ex\]](#) (2012).
4. Chatrchyan, S. *et al.* Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP* **06**, 081. arXiv: [1303.4571 \[hep-ex\]](#) (2013).
5. De Florian, D. & Mazzitelli, J. Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD. *Phys. Rev. Lett.* **111**, 201801. arXiv: [1309.6594 \[hep-ph\]](#) (2013).
6. Randall, L. & Sundrum, R. A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.* **83**, 3370. arXiv: [hep-ph/9905221 \[hep-ph\]](#) (1999).
7. Goldberger, W. D. & Wise, M. B. Modulus stabilization with bulk fields. *Phys. Rev. Lett.* **83**, 4922. arXiv: [hep-ph/9907447 \[hep-ph\]](#) (1999).
8. Csaki, C., Graesser, M., Randall, L. & Terning, J. Cosmology of brane models with radion stabilization. *Phys. Rev. D* **62**, 045015. arXiv: [hep-ph/9911406 \[hep-ph\]](#) (2000).
9. Csaki, C., Graesser, M. L. & Kribs, G. D. Radion dynamics and electroweak physics. *Phys. Rev. D* **63**, 065002. arXiv: [hep-th/0008151 \[hep-th\]](#) (2001).
10. Davoudiasl, H., Hewett, J. L. & Rizzo, T. G. Phenomenology of the Randall-Sundrum Gauge Hierarchy Model. *Phys. Rev. Lett.* **84**, 2080. arXiv: [hep-ph/9909255 \[hep-ph\]](#) (2000).
11. DeWolfe, O., Freedman, D. Z., Gubser, S. S. & Karch, A. Modeling the fifth dimension with scalars and gravity. *Phys. Rev. D* **62**, 046008. arXiv: [hep-th/9909134 \[hep-th\]](#) (2000).
12. Agashe, K., Davoudiasl, H., Perez, G. & Soni, A. Warped Gravitons at the LHC and Beyond. *Phys. Rev. D* **76**, 036006. arXiv: [hep-ph/0701186 \[hep-ph\]](#) (2007).
13. Oliveira, A. *Gravity particles from Warped Extra Dimensions, a review. Part I - KK Graviton* Submitted on arXiv. 2014. arXiv: [1404.0102 \[hep-ph\]](#).

14. Branco, G., Ferreira, P., Lavoura, L., Rebelo, M., Sher, M., *et al.* Theory and phenomenology of two-Higgs-doublet models. *Phys.Rept.* **516**, 1–102. arXiv: [1106.0034 \[hep-ph\]](#) (2012).
15. Djouadi, A. The anatomy of electroweak symmetry breaking. Tome II: the Higgs bosons in the minimal supersymmetric model. *Phys. Rept.* **459**, 1. arXiv: [hep-ph/0503173 \[hep-ph\]](#) (2008).
16. Georgi, H. & Machacek, M. Doubly charged Higgs bosons. *Nuclear Physics B* **262**, 463 (1985).
17. CMS Collaboration. *Standard Model Cross Sections for CMS at 13 TeV*. CMS TWiki r24 (CERN, 2016).
18. A. Oliveira. *Bulk graviton NLO cross sections in WED at 13 TeV* 2016.
19. A. Oliveira. *Bulk radion NLO cross sections in WED at 13 TeV* 2016.
20. A. Oliveira. *Bulk graviton to HH decay branching fractions in WED*
21. A. Oliveira. *Bulk radion to HH decay branching fractions in WED*
22. Carvalho, A. *CrossSectionsLHC/WED* https://github.com/CrossSectionsLHC/WED/blob/master/KKGraviton_Bulk/GF_NLO_13TeV_ktilda_0p1.txt.
23. Butterworth, J. M., Davison, A. R., Rubin, M. & Salam, G. P. Jet substructure as a new Higgs search channel at the LHC. *Phys. Rev. Lett.* **100**, 242001. arXiv: [0802.2470 \[hep-ph\]](#) (2008).
24. Cooper, B., Konstantinidis, N., Lambourne, L. & Wardrope, D. Boosted $hh \rightarrow b\bar{b}b\bar{b}$: a new topology in searches for TeV-scale resonances at the LHC. *Phys.Rev.* **D88**, 114005. arXiv: [1307.0407 \[hep-ex\]](#) (2013).
25. Gouzevitch, M. *et al.* Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production. *JHEP* **07**, 148. arXiv: [1303.6636 \[hep-ph\]](#) (2013).
26. *Search for heavy resonances decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV* tech. rep. CMS-PAS-B2G-16-026 (CERN, Geneva, 2017).
27. CMS Collaboration. *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET* CMS Physics Analysis Summary CMS-PAS-PFT-09-001 (CERN, 2009).
28. CMS Collaboration. *Commissioning of the Particle-Flow Event Reconstruction with the first LHC collisions recorded in the CMS detector* CMS Physics Analysis Summary CMS-PAS-PFT-10-001 (CERN, 2010).
29. Cacciari, M., Salam, G. P. & Soyez, G. The anti- k_t jet clustering algorithm. *JHEP* **04**, 063. arXiv: [0802.1189 \[hep-ph\]](#) (2008).
30. Cacciari, M., Salam, G. P. & Soyez, G. FastJet User Manual. *Eur. Phys. J. C* **72**, 1896. arXiv: [1111.6097 \[hep-ph\]](#) (2012).
31. Bertolini, D., Harris, P., Low, M. & Tran, N. Pileup Per Particle Identification. *JHEP* **10**, 059. arXiv: [1407.6013 \[hep-ph\]](#) (2014).
32. CMS Collaboration. *Jet Performance in pp Collisions at $\sqrt{s}=7$ TeV* CMS Physics Analysis Summary CMS-PAS-JME-10-003 (CERN, 2010).

33. Collaboration, C. Machine learning-based identification of highly Lorentz-boosted hadronically decaying particles at the CMS experiment (2019).
34. Das, S., Konigsberg, J., Donato, S., Rizzi, A. & Vernieri, C. *Search for di-Higgs resonances decaying to 4 b-jets in pp collisions at 13 TeV* CMS Note 2015/108 (CERN, 2015).
35. Cranmer, K. S. Kernel estimation in high-energy physics. *Comput. Phys. Commun.* **136**, 198. arXiv: [hep-ex/0011057](#) (2001).
36. CMS Collaboration. *CMS Luminosity Measurement for the 2016 Data Taking Period* CMS Physics Analysis Summary CMS-PAS-LUM-17-001 (CERN, 2017).
37. *Utilities for Accessing Pileup Information for Data: 2015 Pileup JSON Files* https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData#2015_Pileup_JSON_Files. 2015.
38. Khachatryan, V. e. a. Measurement of differential cross sections for top quark pair production using the lepton + jets final state in proton-proton collisions at 13 TeV. *Phys. Rev. D* **95**, 092001 (9 May 2017).
39. Khachatryan, V. et al. Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP* **01**, 080. arXiv: [1012.2466 \[hep-ex\]](#) (2011).
40. Ball, R. D. et al. Parton distributions with LHC data. *Nucl. Phys. B* **867**, 244. arXiv: [1207.1303 \[hep-ph\]](#) (2013).
41. Khachatryan, V. et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST* **12**, P02014. arXiv: [1607.03663 \[hep-ex\]](#) (2017).
42. Sirunyan, A. M. et al. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector. *JINST* **14**, P07004. arXiv: [1903.06078 \[hep-ex\]](#) (2019).
43. Sirunyan, A. M. et al. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV. *JHEP* **07**, 161. arXiv: [1802.02613 \[hep-ex\]](#) (2018).
44. Aaboud, M. et al. Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC. *Phys. Rev. Lett.* **117**, 182002. arXiv: [1606.02625 \[hep-ex\]](#) (2016).
45. ATLAS and CMS Collaborations, The LHC Higgs Combination Group. *Procedure for the LHC Higgs boson search combination in Summer 2011* tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11 (CERN, 2011).
46. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C* **71**. [Erratum: DOI10.1140/epjc/s10052-013-2501-z], 1554. arXiv: [1007.1727 \[physics.data-an\]](#) (2011).
47. Read, A. L. Presentation of search results: the CL_s technique. *J. Phys. G* **28**, 2693 (2002).
48. Junk, T. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A* **434**, 435. arXiv: [hep-ex/9902006 \[hep-ex\]](#) (1999).

Appendix A

A.1 Statistical Tests

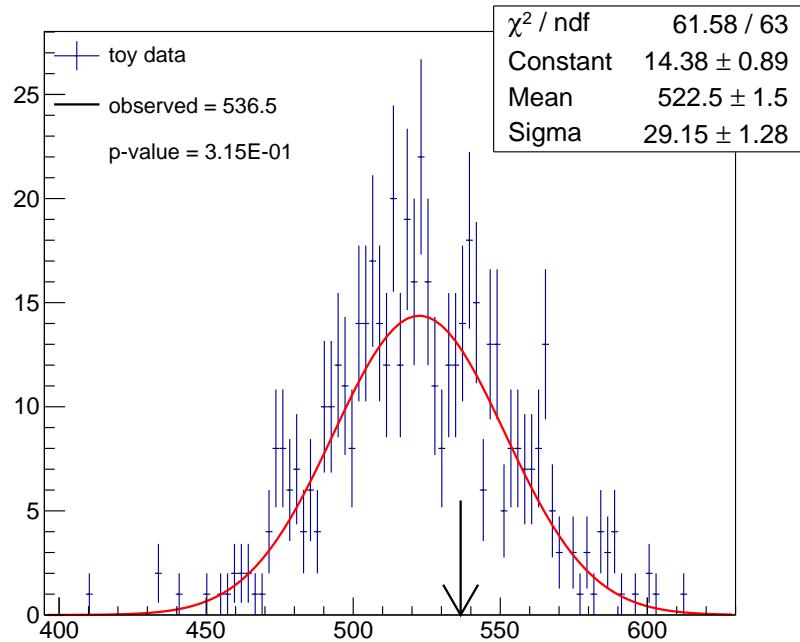


Figure A-1. Full Run 2 Goodness of Fit test plot with 500 toys.

It is obvious that there may be some bias at the 2 and 3 TeV points when we inject $r = 0$. However, this is actually due to there being poor statistics at those mass points. When we simulate a higher integrated luminosity we get a plot with a more reasonable shape.

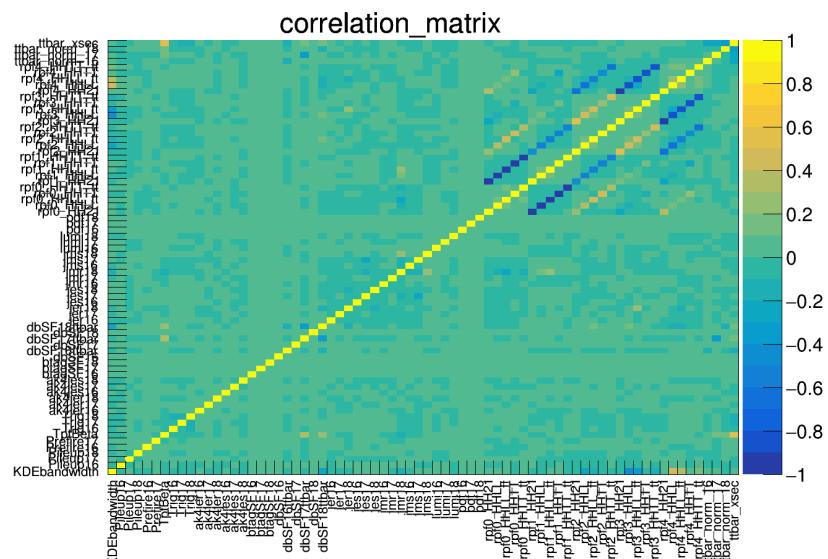


Figure A-2. Full Run 2 Correlation Matrix.

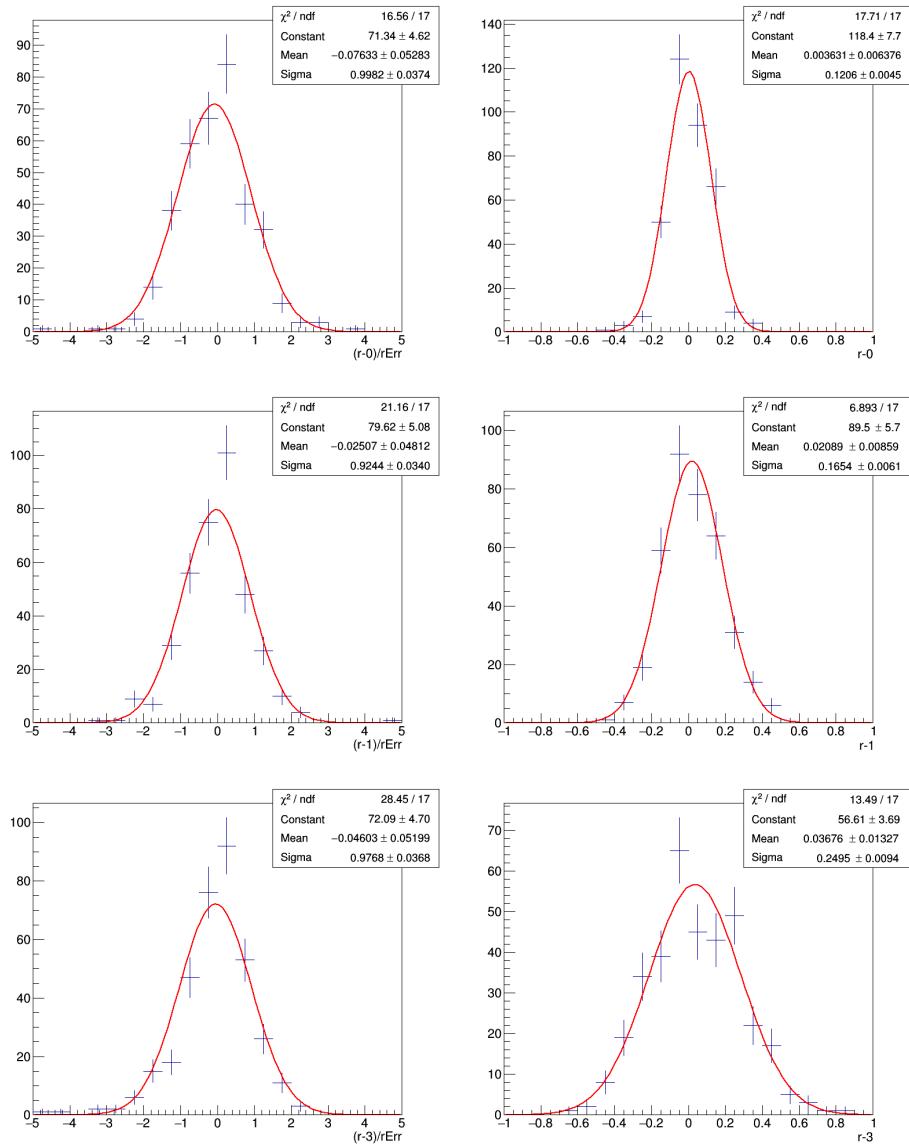


Figure A-3. Signal Injection (Left Column is Pull and Right is Strength) for Full Run 2 at 1 TeV for $r = 0, 1, 3$ with 500 toys.

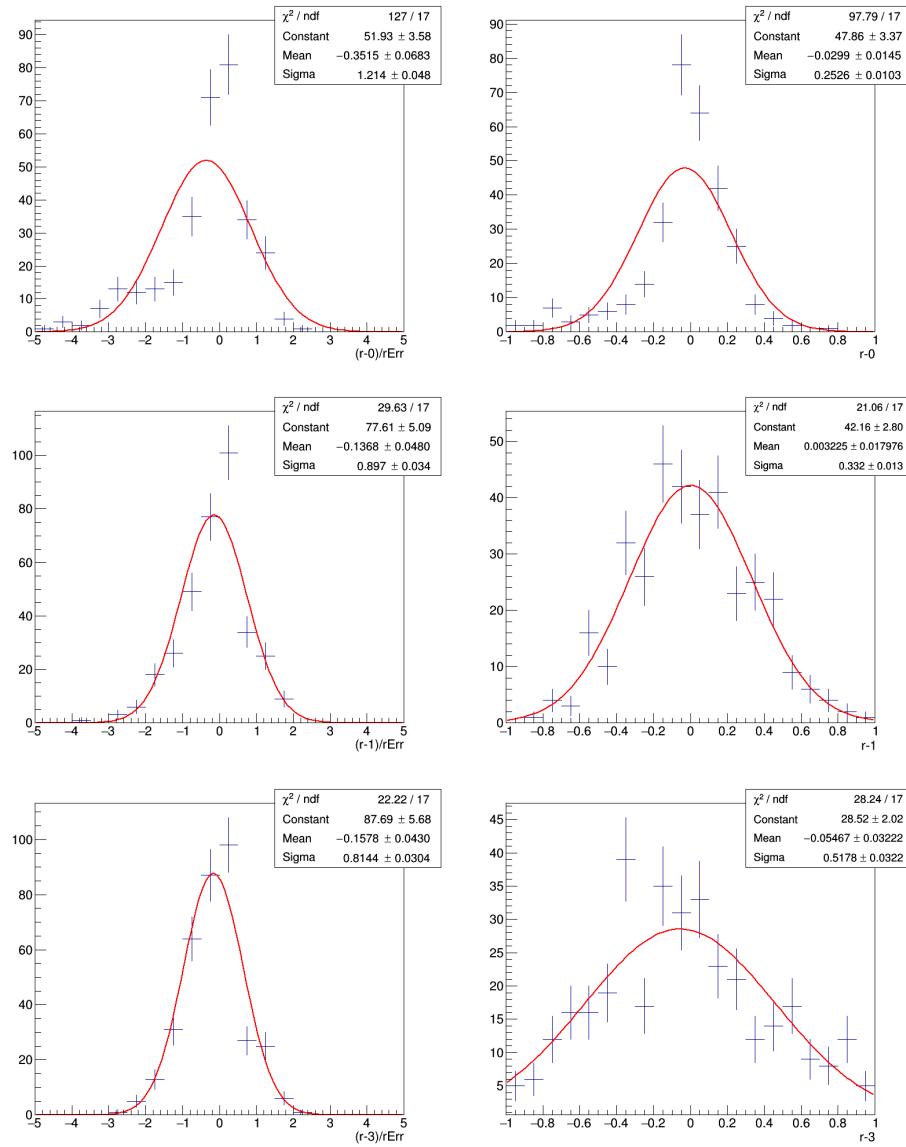


Figure A-4. Signal Injection (Left Column is Pull and Right is Strength) for Full Run 2 at 2 TeV for $r = 0, 1, 3$ with 500 toys.

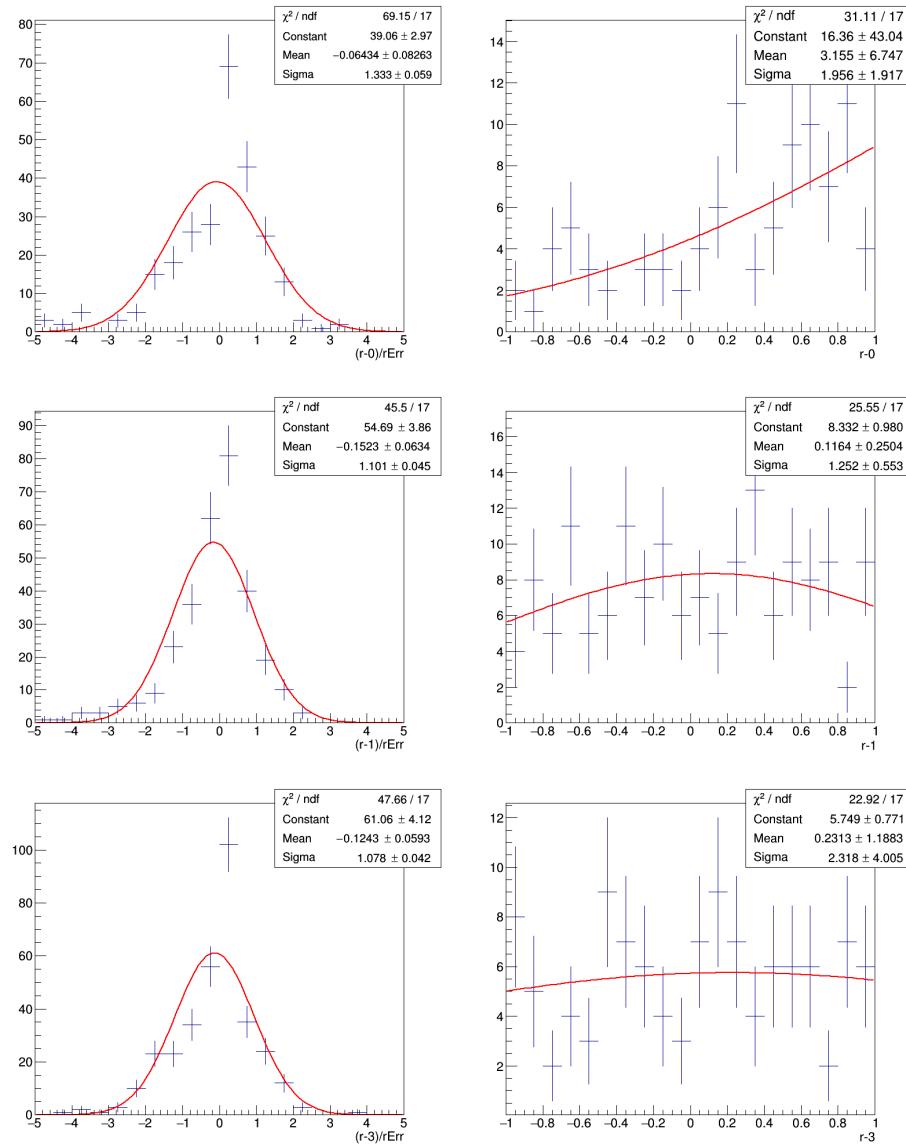


Figure A-5. Signal Injection (Left Column is Pull and Right is Strength) for Full Run 2 at 3 TeV for $r = 0, 1, 3$ with 500 toys.

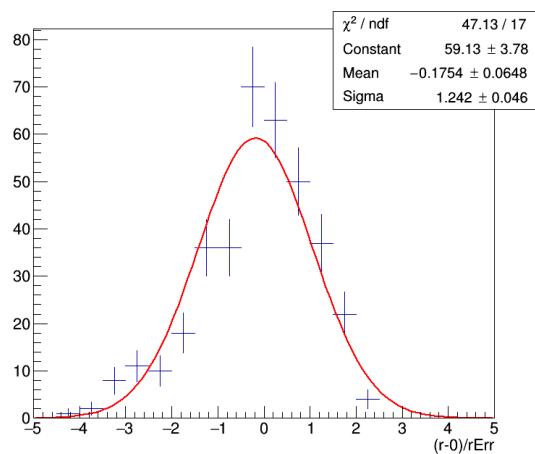


Figure A-6. Signal Injection for Full Run 2 at 2 TeV for $r = 0$ with 500 toys and a simulated higher integrated luminosity.

A.1.1 F-Test

We performed an F-test to optimize the function used to perform the fit. The results show that a 2nd order in M_j and 2nd order in M_{jj}^{red} fit function is most optimal.

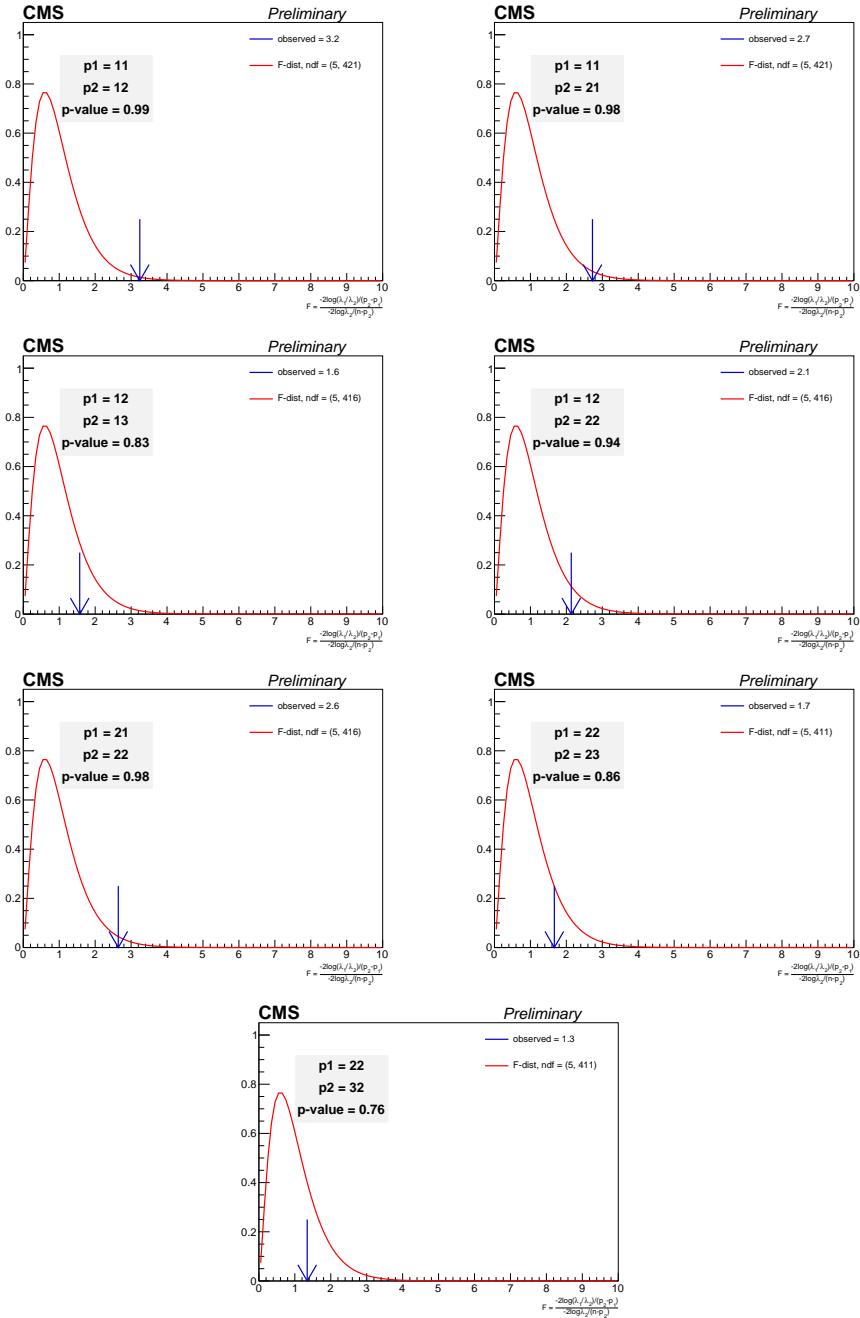


Figure A-7. F-test plots showing that the 2x2 fit function is the most optimal. When reading these plot, p1 is the base fit function and p2 is the function that is being compared as an improvement. Whether for p1 or p2, the designation pol gives the order of the polynomial fit function in M_j , M_{jj}^{red} in order. Pol11 means 1st order polynomial in M_j and 1st order polynomial in M_{jj}^{red} .

A.2 2D Alphabet Validation Study

We are interested in knowing how well the 2D Alphabet method compares to the previous method used in [26]. Here we present multiple limits:

- The 2016 Double B tagger limit from [26].
- The 2016 Double B tagger limit made by using 2DAlphabet in 1D and including the τ_{21} cut.
- The 2016 Double B Tagger limit made with 2D Alphabet.
- The full run 2 Double B Tagger limit made with 2D Alphabet
- The full run 2 dak8MDHbb limit made with 2D Alphabet

We present the comparison in the following plot:

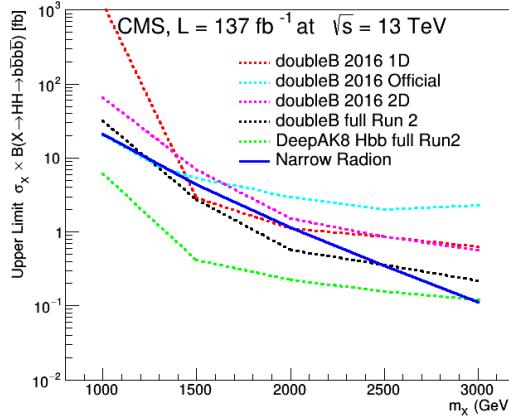


Figure A-8. Comparison plot of various limits.

A.3 Example Combine Card

The combine cards for this analysis are stored in git and may be viewed here:

<https://github.com/dbrehm/JHUAnalyzer/tree/rdfHH4b/Framework/AnalysisModules/B2G-20-004/DataCards>

A.4 τ_{21} Study

We studied the effect of the τ_{21} cut used in the original analysis. We plot the results of this study: It is easy to see here that there is very little difference in the above

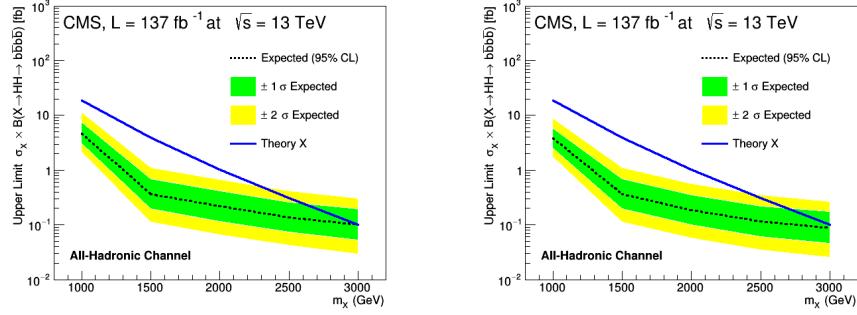


Figure A-9. Comparison of limits with and without τ_{21} cut. The plot on the right is with the cut and the plot on the left is without the cut.

plots. This is helpful because we then gain on statistics and systematic uncertainties as the τ_{21} scale factor uncertainty is not applied to the above limit plots.

A.5 By Year Fits

Previous version of this analysis fit each year by itself. The current version of the fitting procedure combined each year so that there are only the three regions to be fit. These are the plots generated when the fit was done by year.

These plots don't necessarily correspond to the latest fixes/updates/optimizations that were made to the full run 2 fit. Until we are able to confirm that the by year fit can happen with our latest methodology, the plots in this section should be considered to be in need of fixing.

A.5.1 Signal Regions Prediction By Year

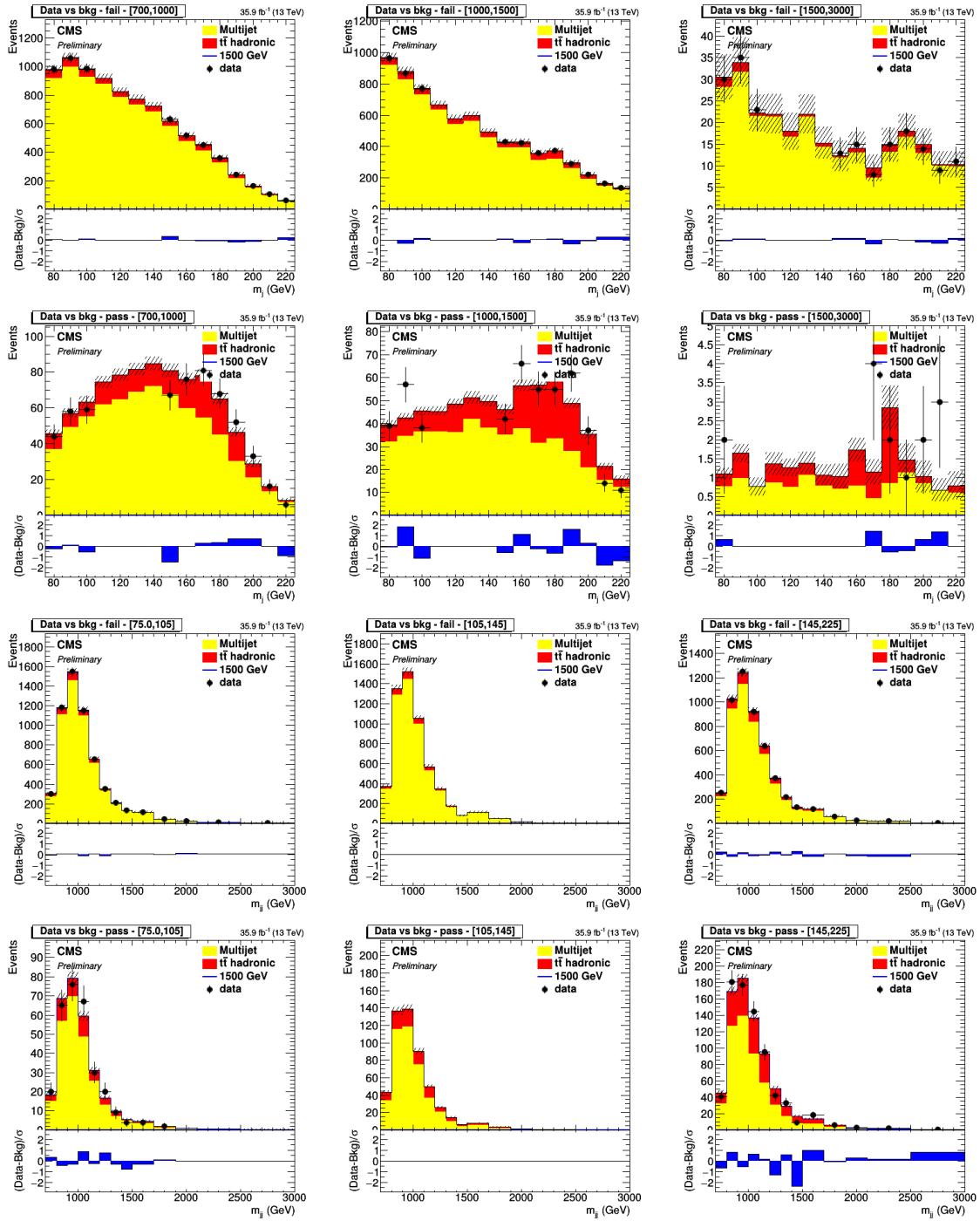


Figure A-10. 2016 Loose Loose background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

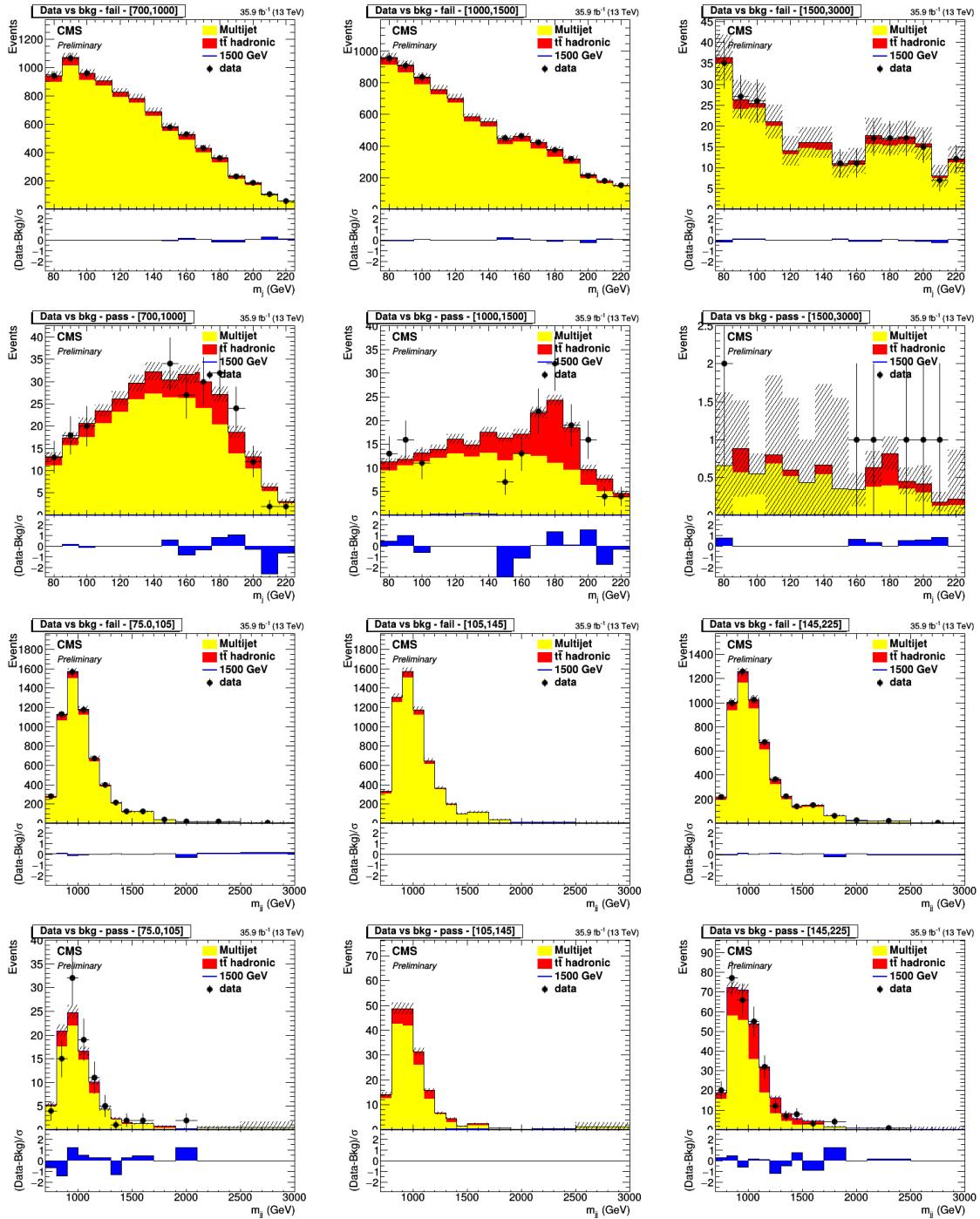


Figure A-11. 2016 Tight Tight background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

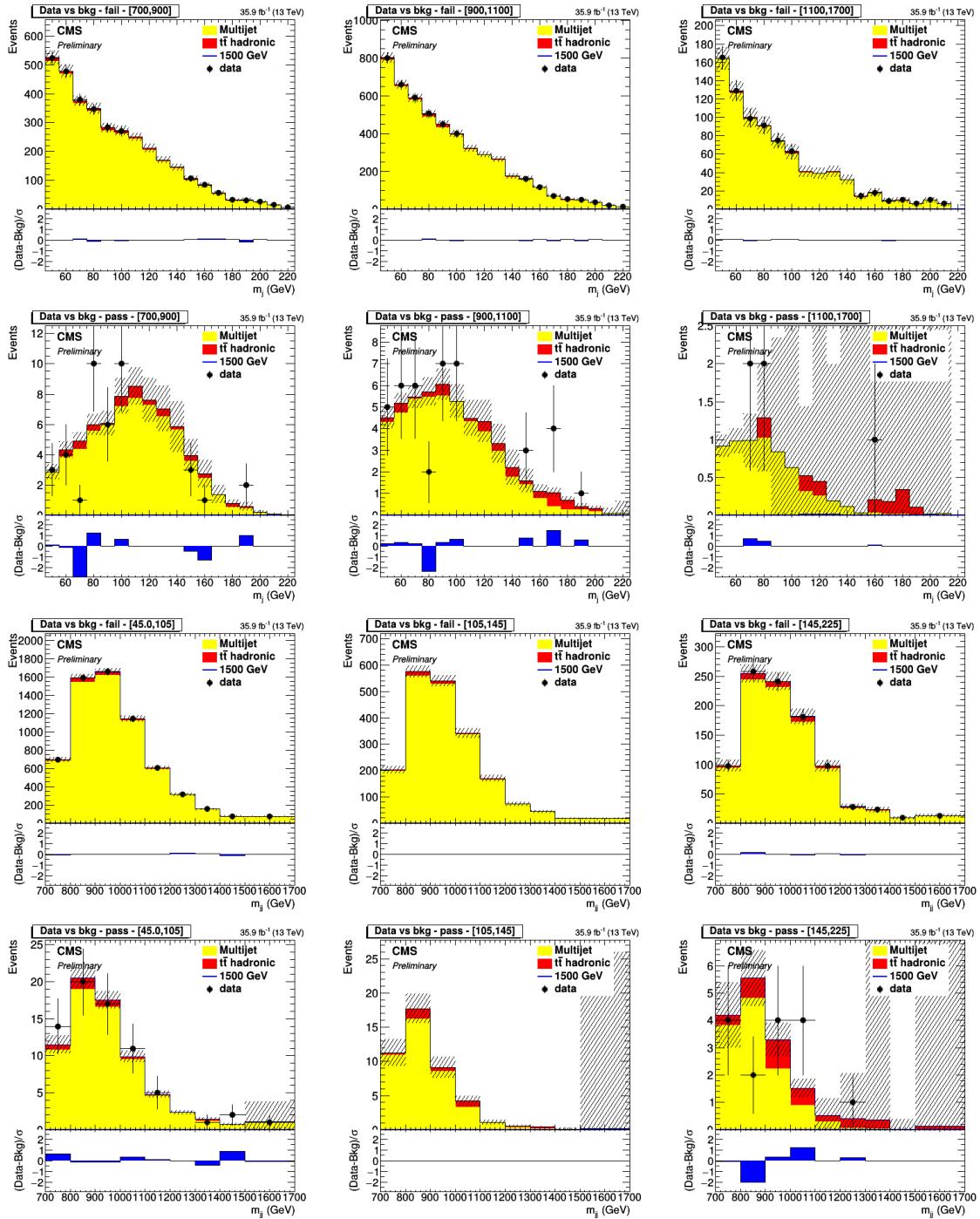


Figure A-12. 2016 2+1 background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

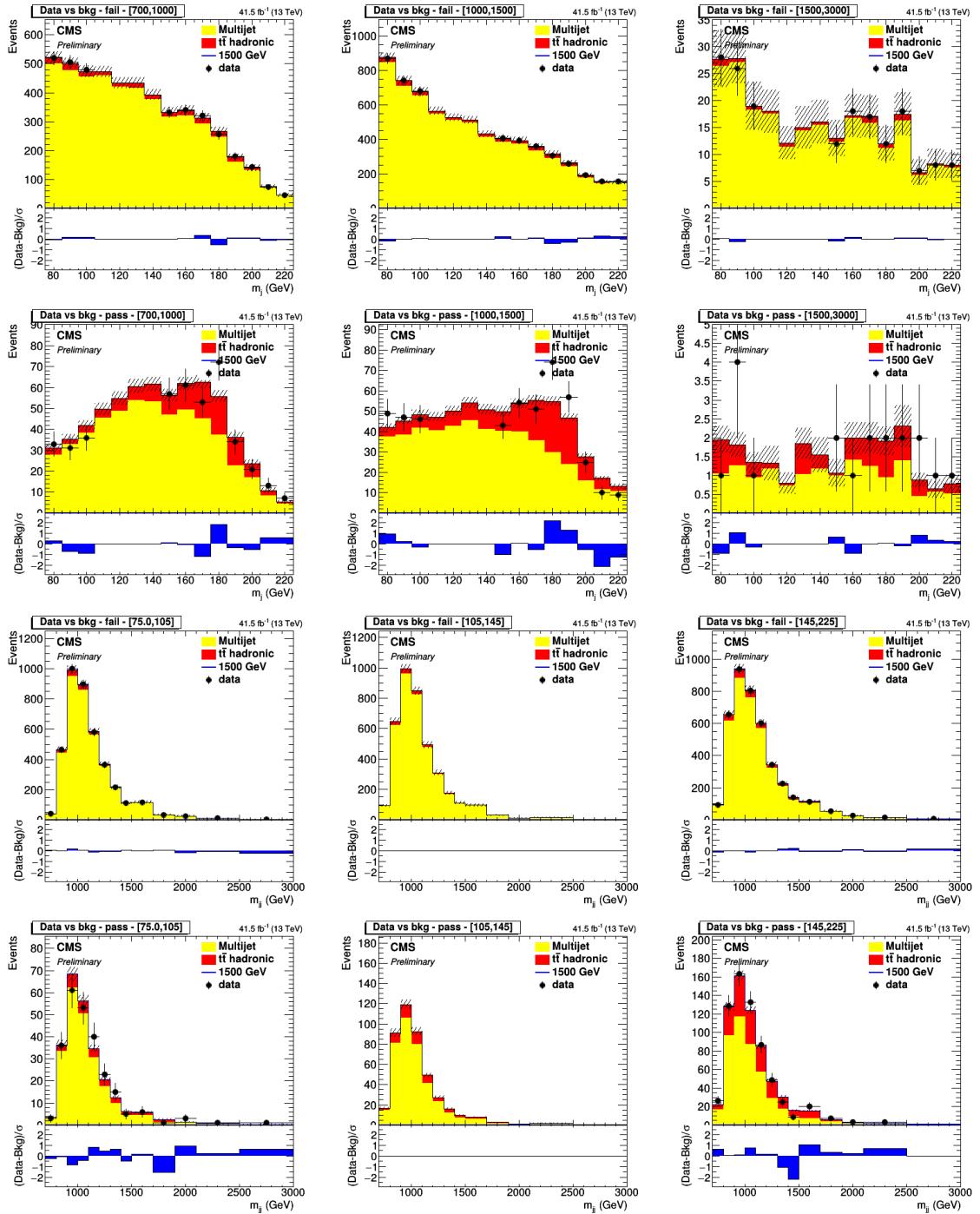


Figure A-13. 2017 Loose Loose background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

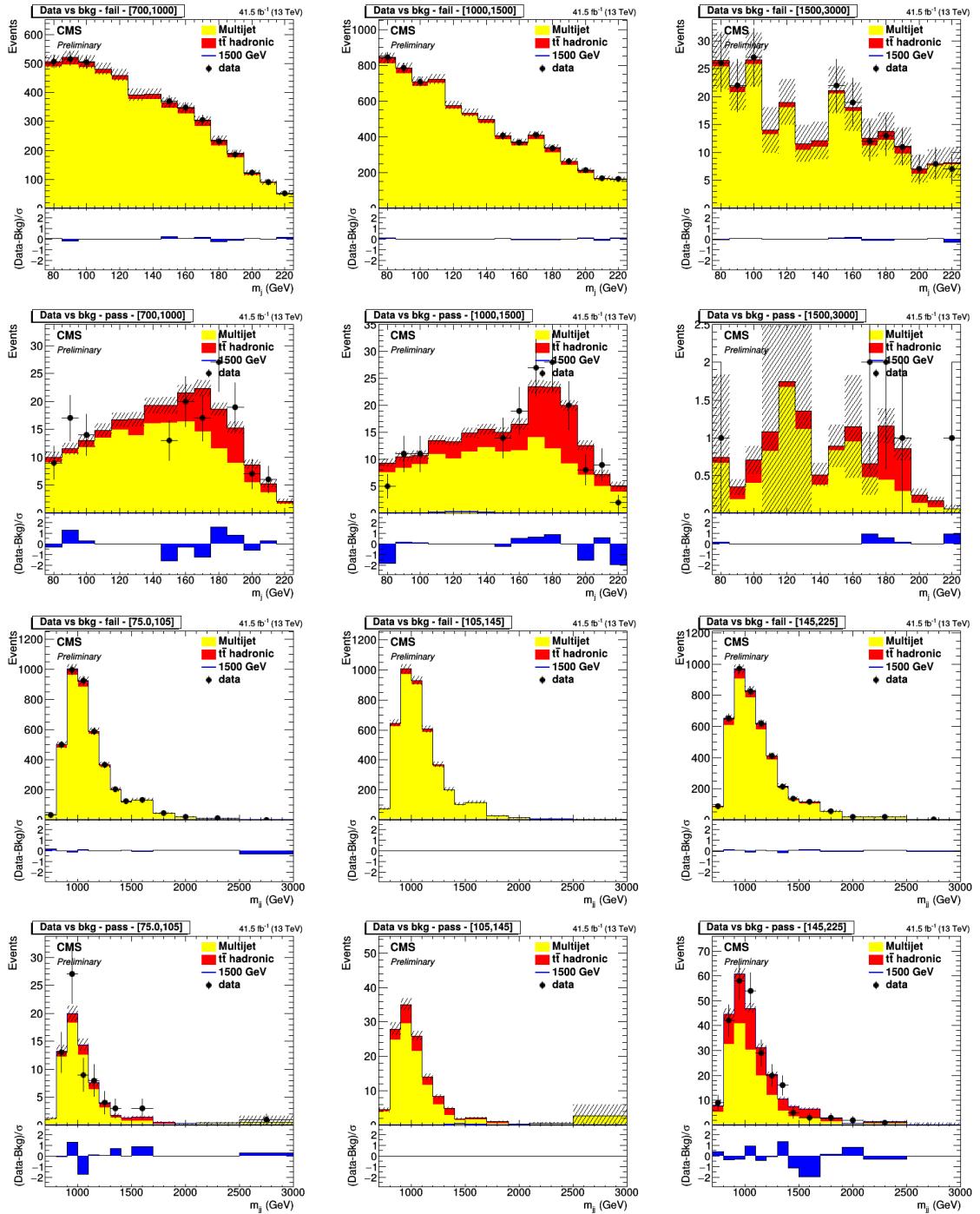


Figure A-14. 2017 Tight Tight background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

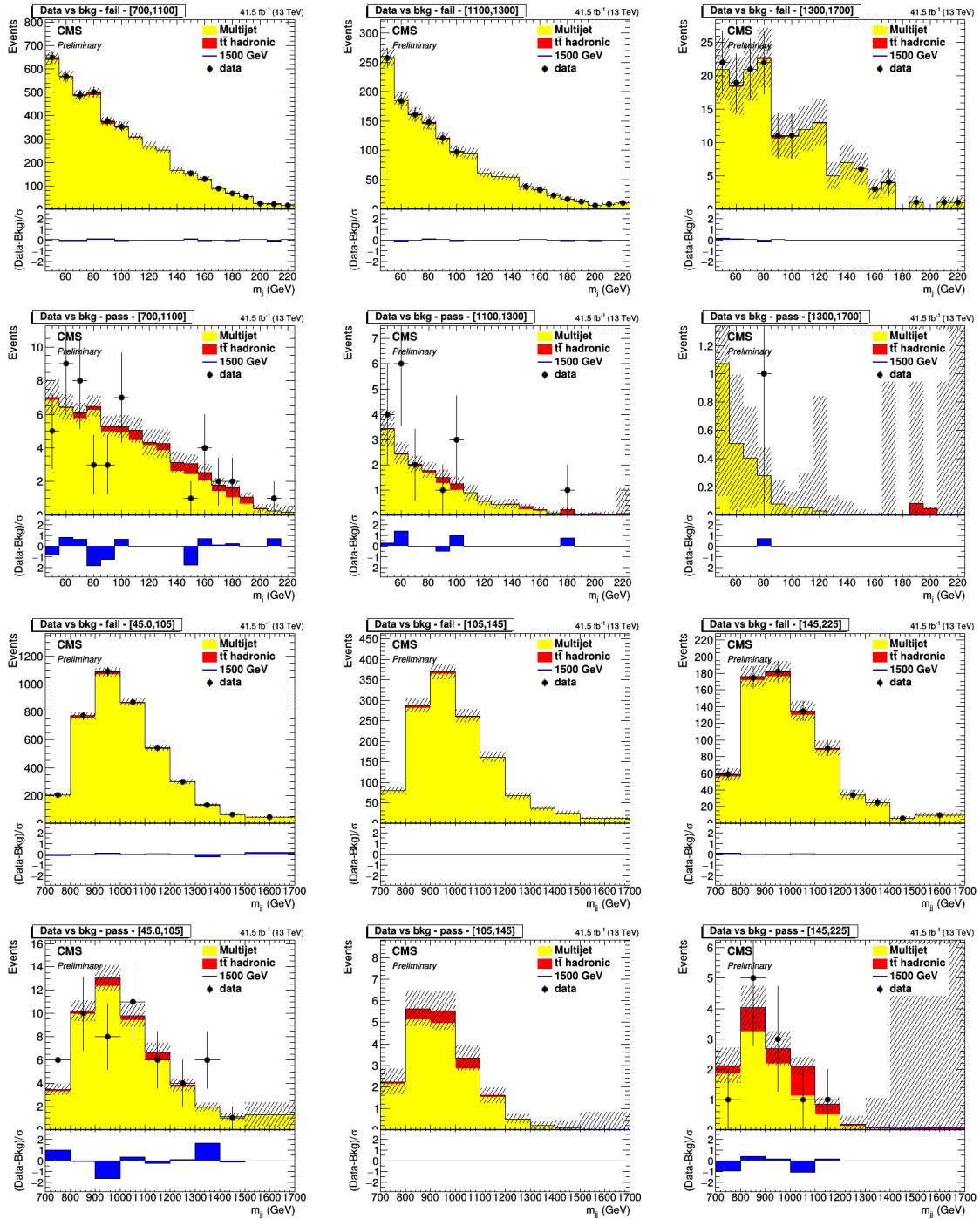


Figure A-15. 2017 2+1 background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

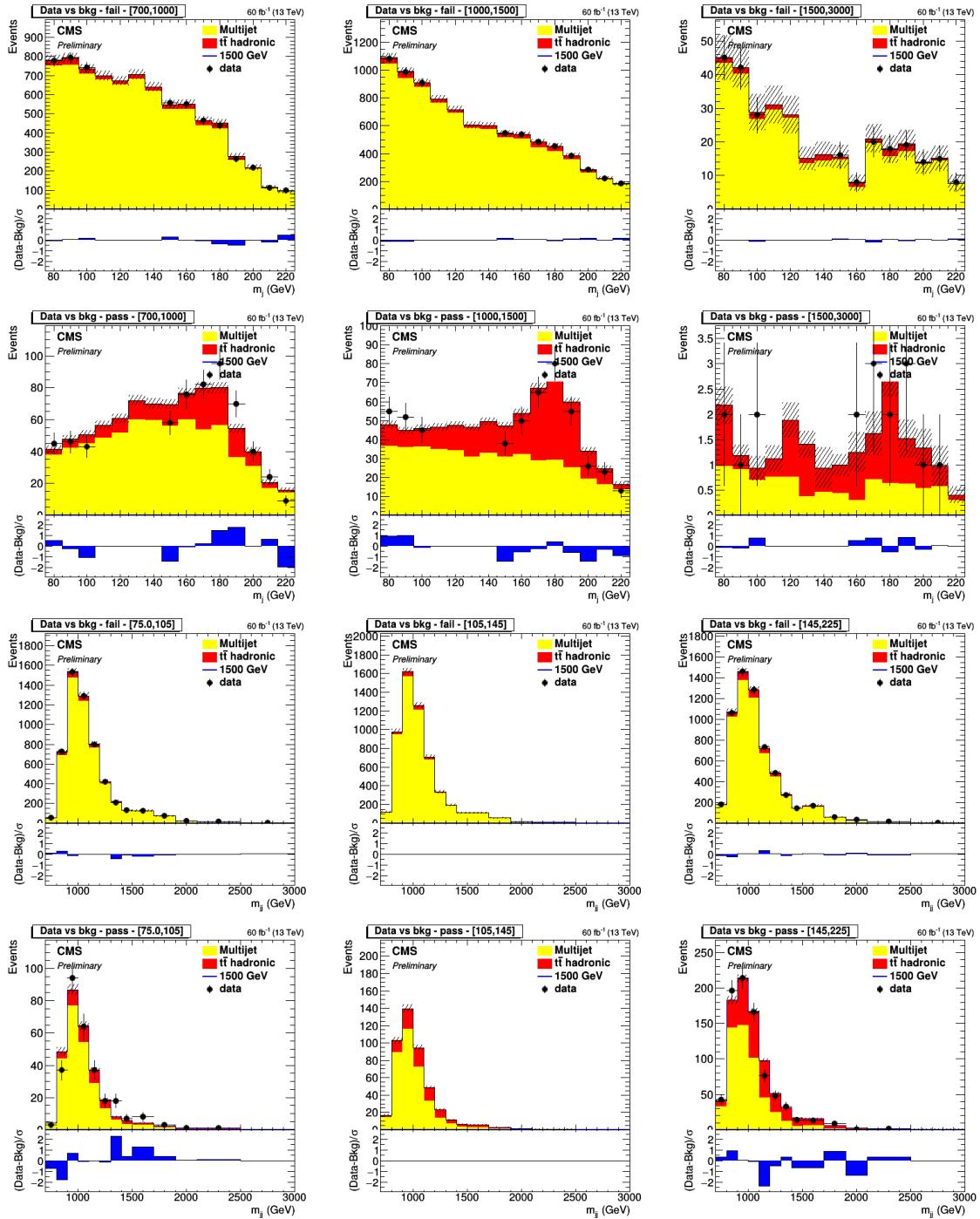


Figure A-16. 2018 Loose Loose background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

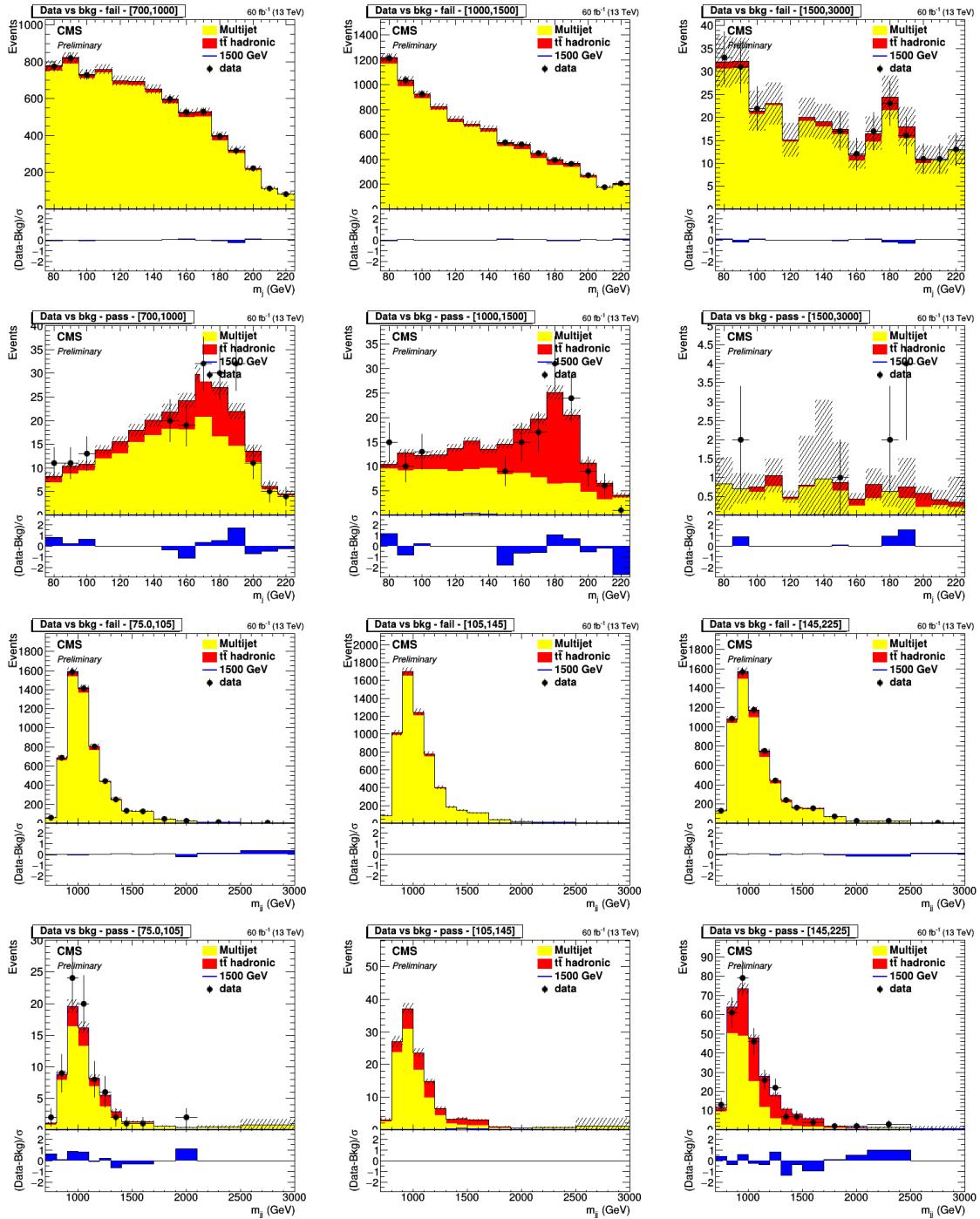


Figure A-17. 2018 Tight Tight background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

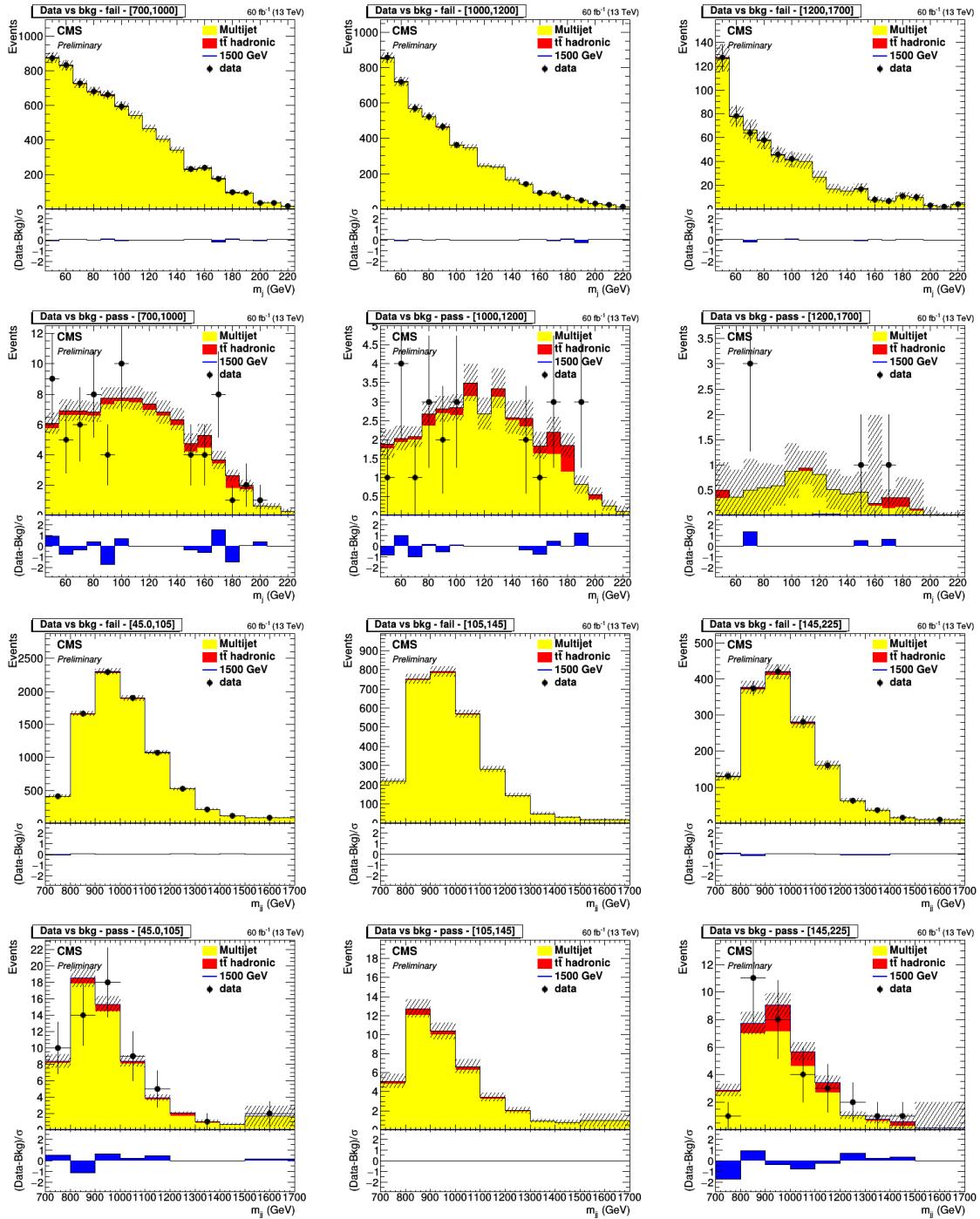


Figure A-18. 2018 2+1 background fits for M_j axis and M_{jj}^{red} axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

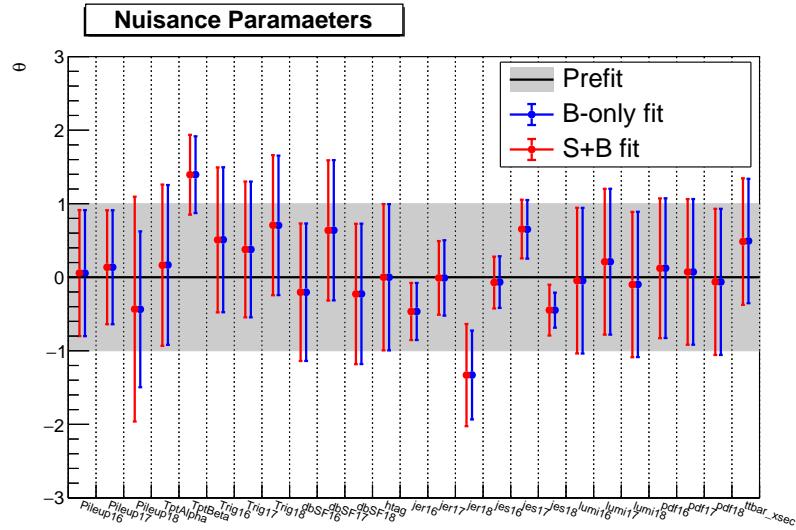


Figure A-19. Full Run 2 Nuissance Pulls Plot.

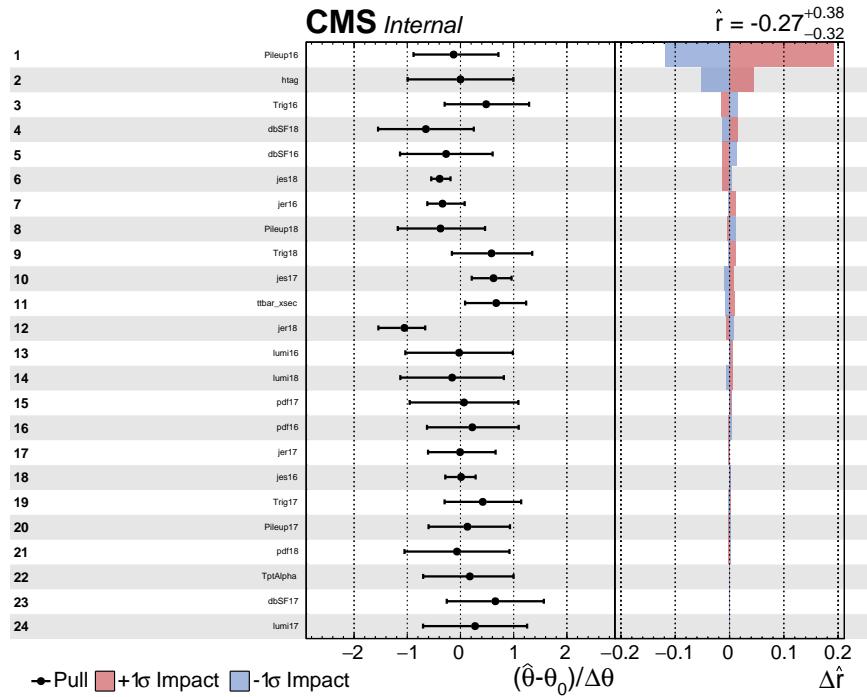


Figure A-20. Full Run 2 Impacts plot.

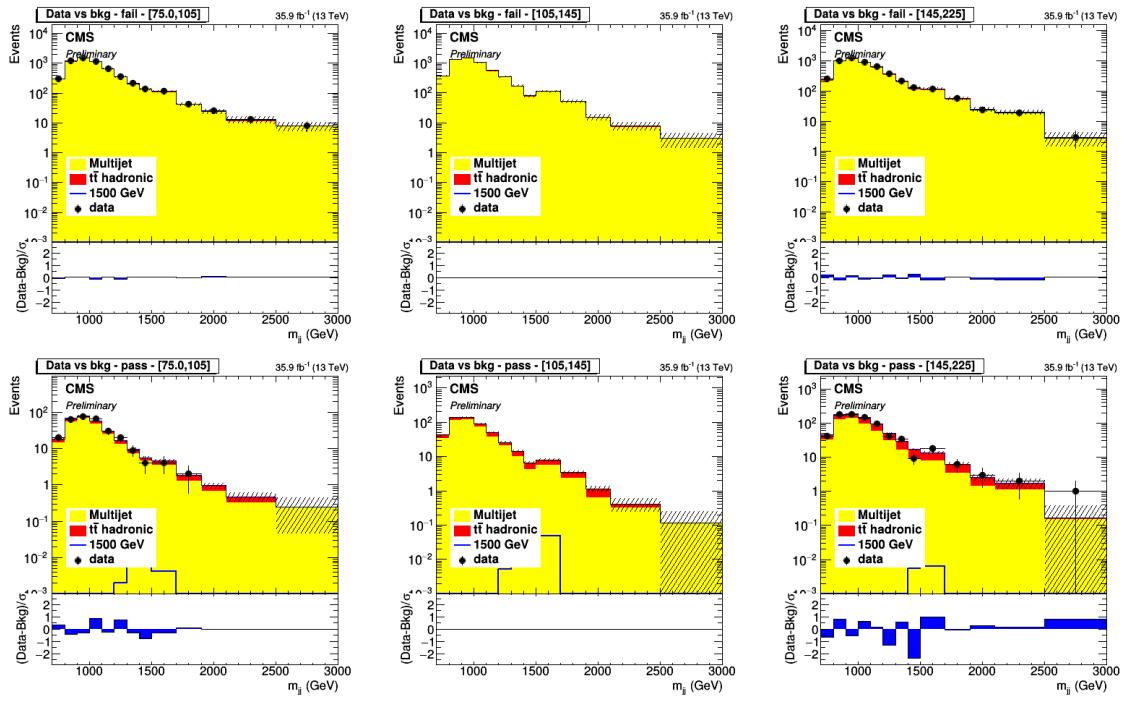


Figure A-21. 2016 Loose Loose background fit for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

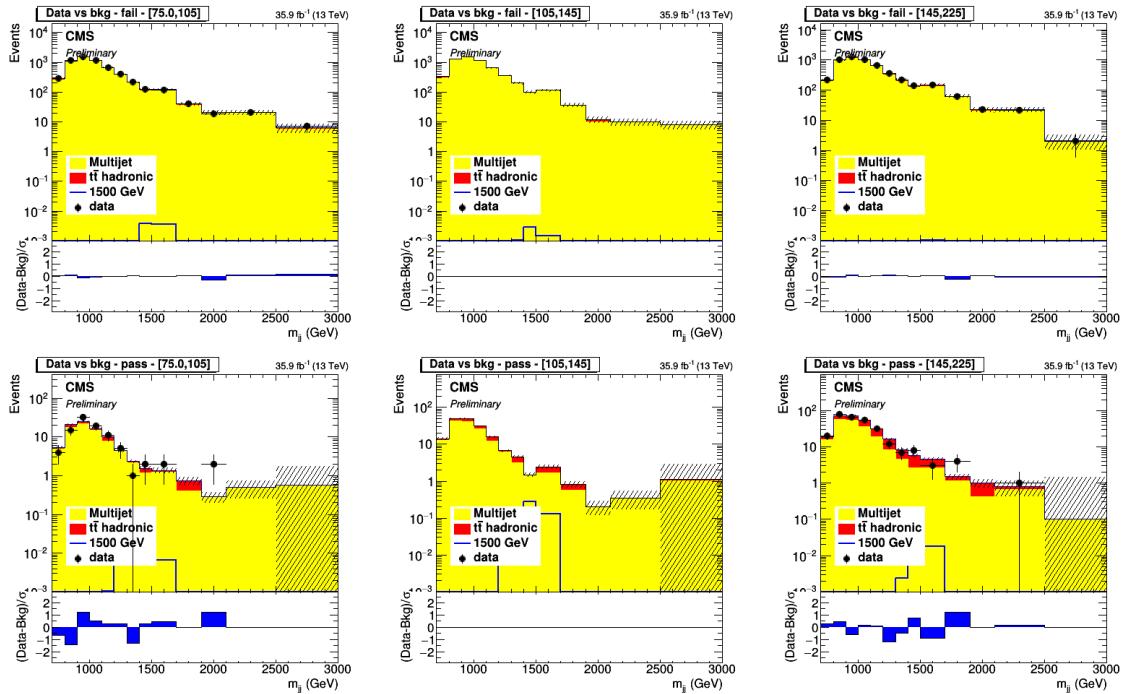


Figure A-22. 2016 Tight Tight background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

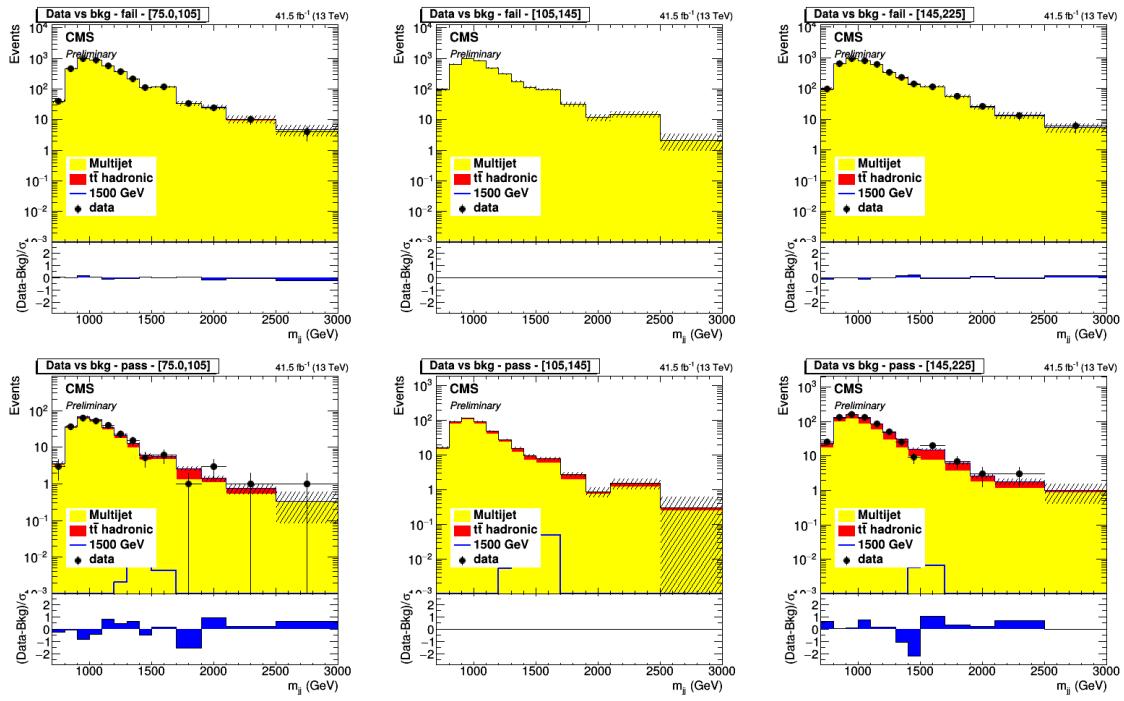


Figure A-23. 2017 Loose Loose background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

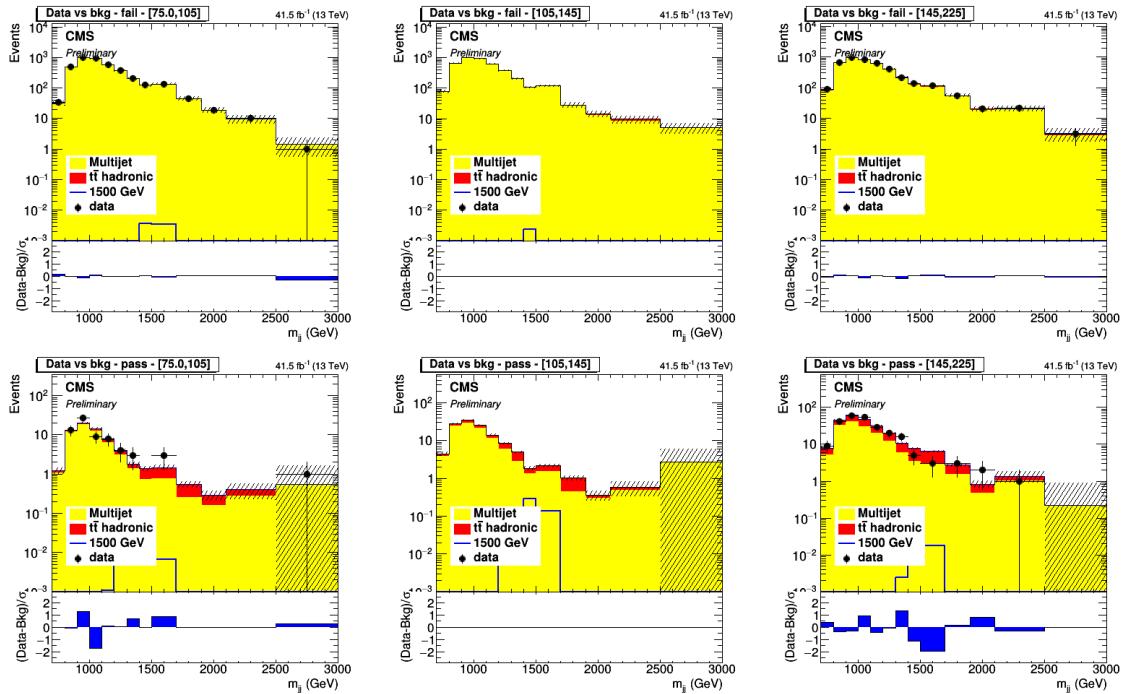


Figure A-24. 2017 Tight Tight background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

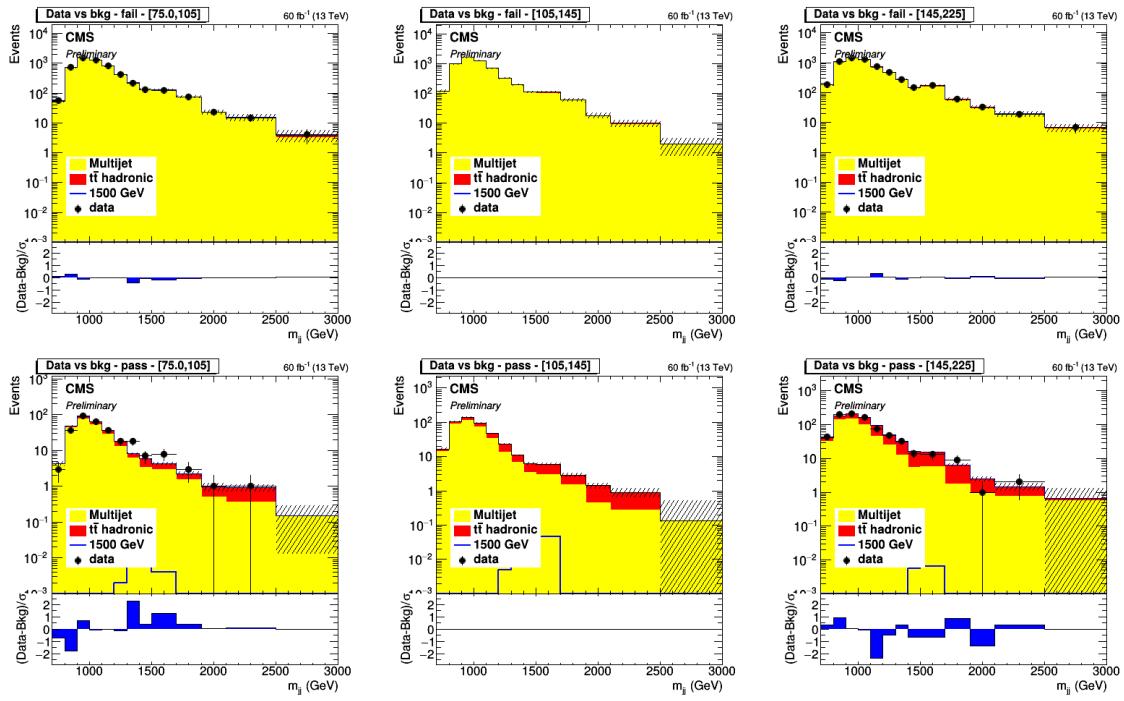


Figure A-25. 2018 Loose Loose background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

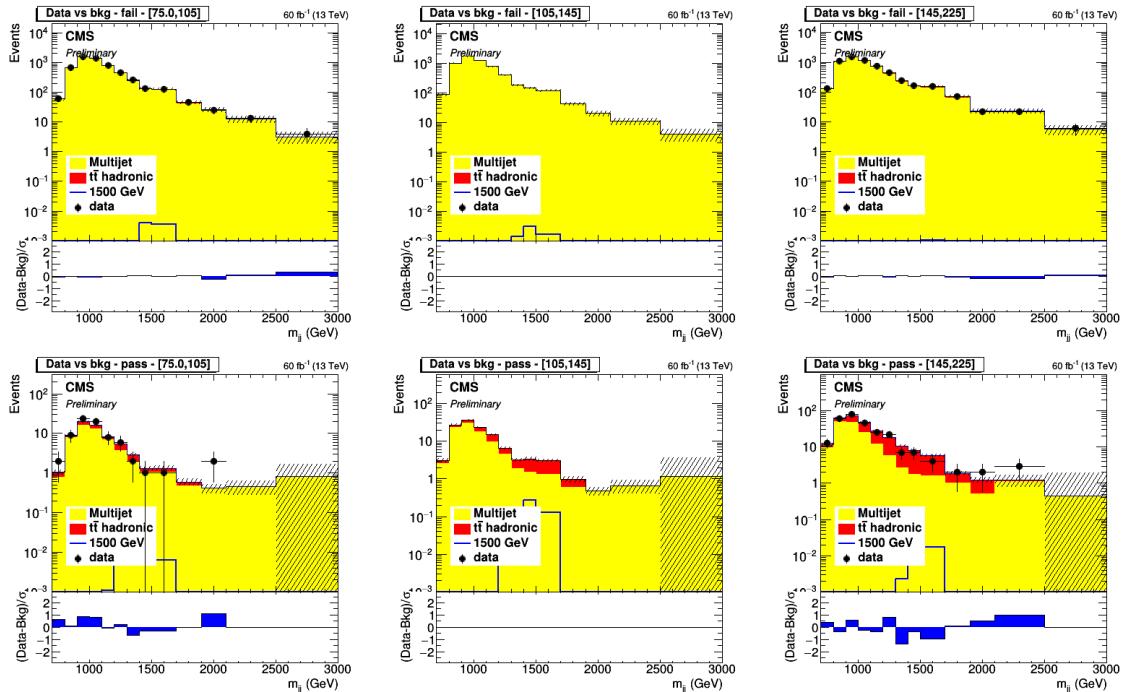


Figure A-26. 2018 Tight Tight background fits for M_{jj}^{red} log-scale axis including expected Radion 1500 GeV signal, normalized to the signal strength found by the fit.

A.5.2 MC distributions before applying Selections

The kinematic variables before the application of H->bb tagger(s), are shown in Figs.5-9 through 5-15. Here, the distributions are normalized to luminosity and the signal is normalized to 50%. Also, the difference between TT and LL distributions is in the weight used for each. Since the deepTagMD_HbbvsQCD tagger has not been applied yet, these differences may be small and not show up in these plots.

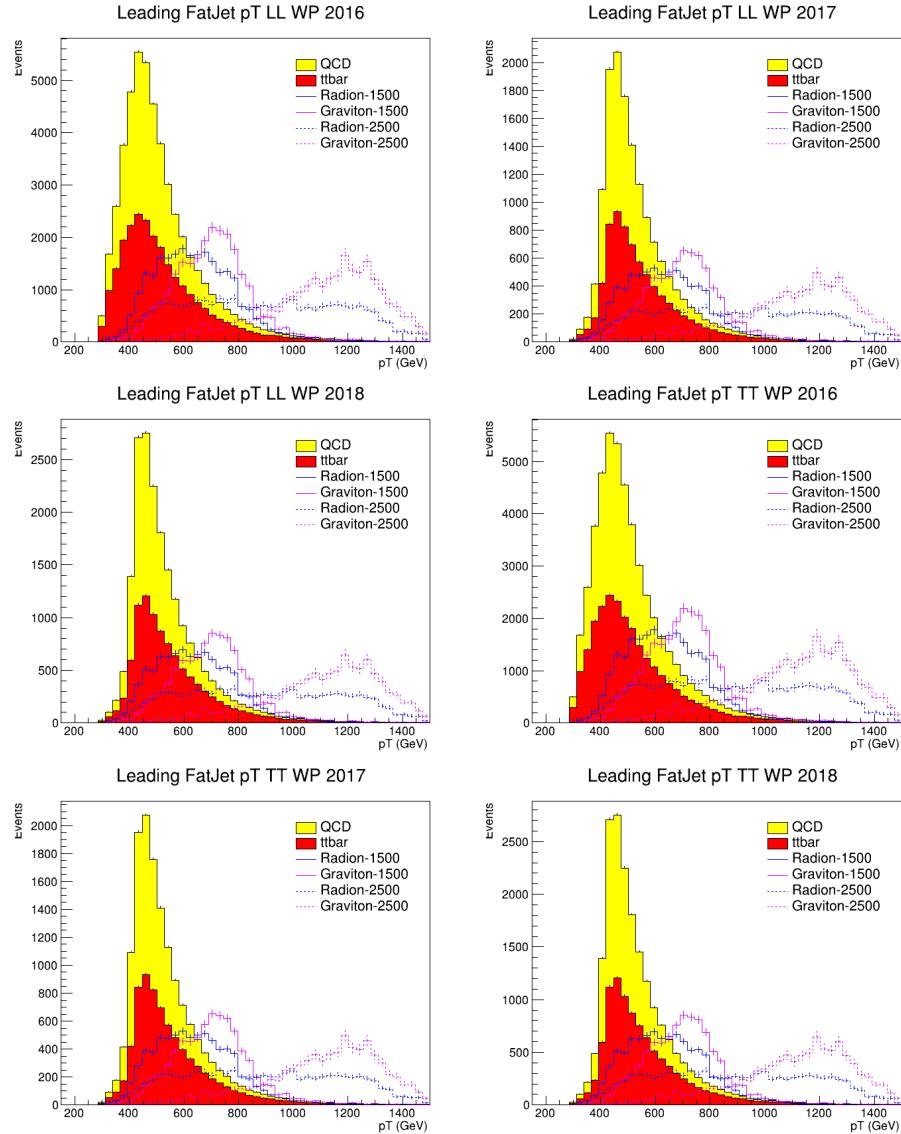


Figure A-27. Pre deepTagMD_HbbvsQCD selection p_T distribution of leading jet

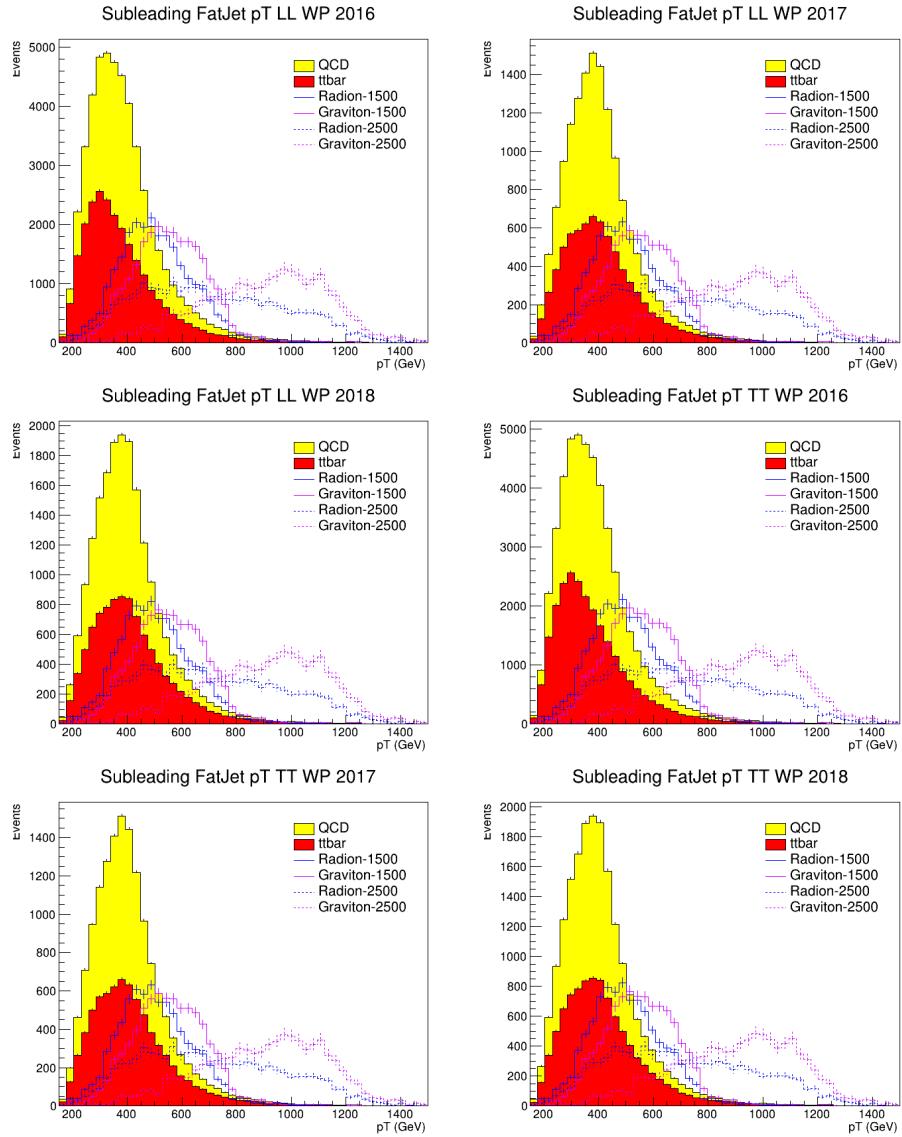


Figure A-28. Pre deepTagMD_HbbvsQCD selection p_T distribution of subleading jet

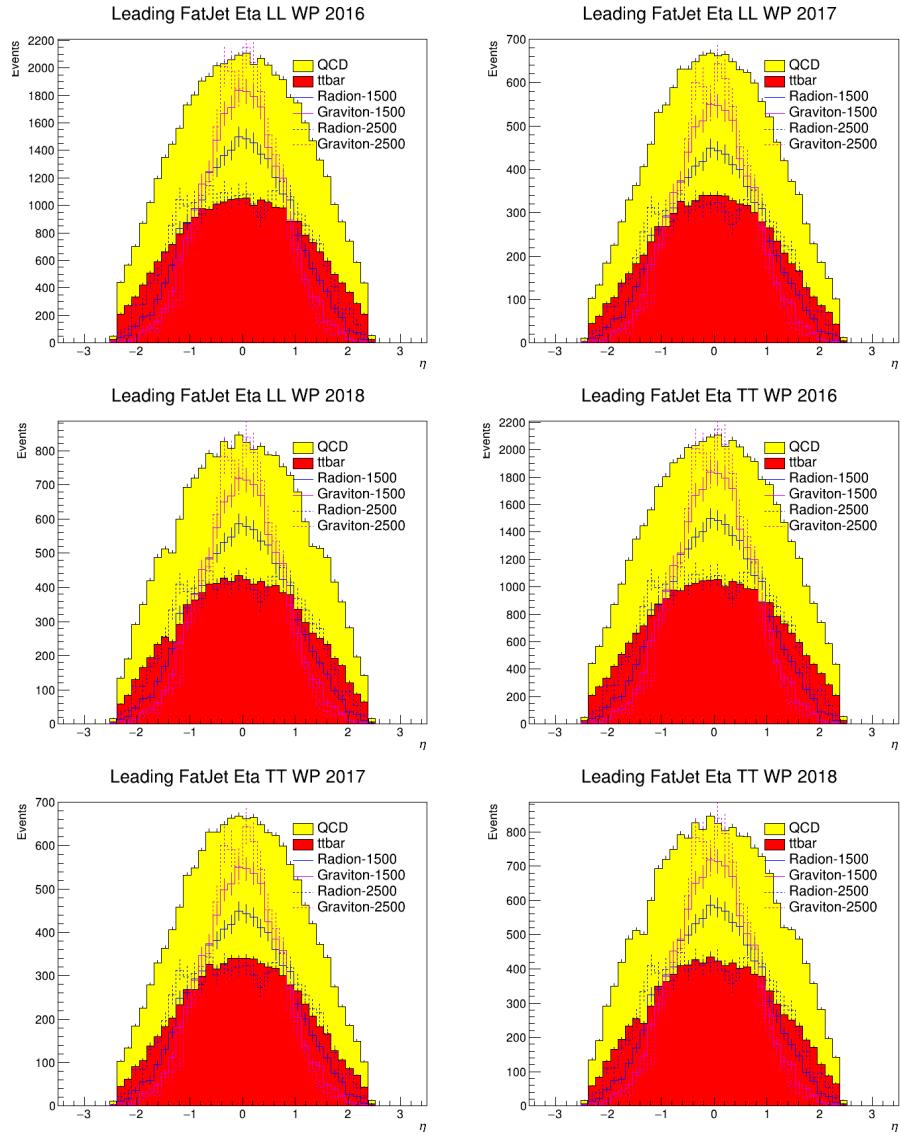


Figure A-29. Pre deepTagMD_HbbvsQCD selection η distribution of leading jet

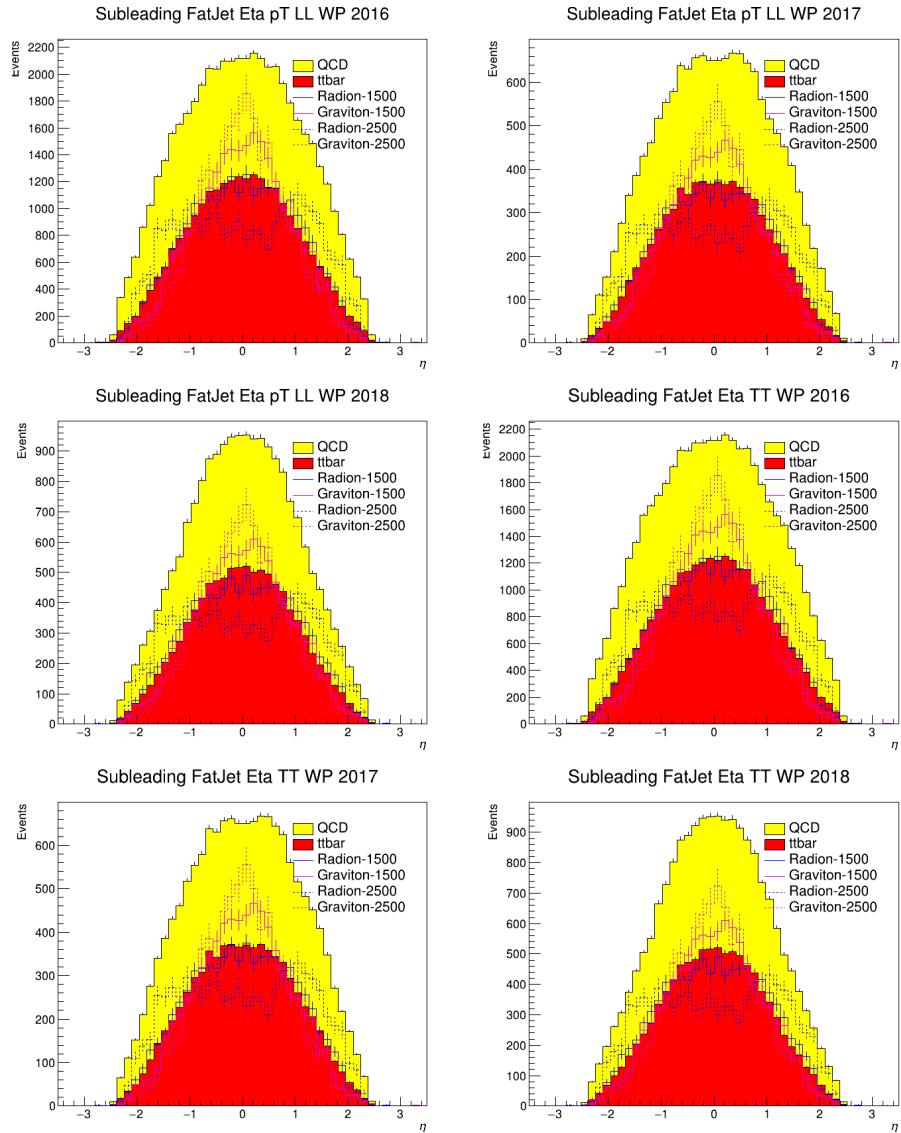


Figure A-30. Pre deepTagMD_HbbvsQCD selection η distribution of subleading jet

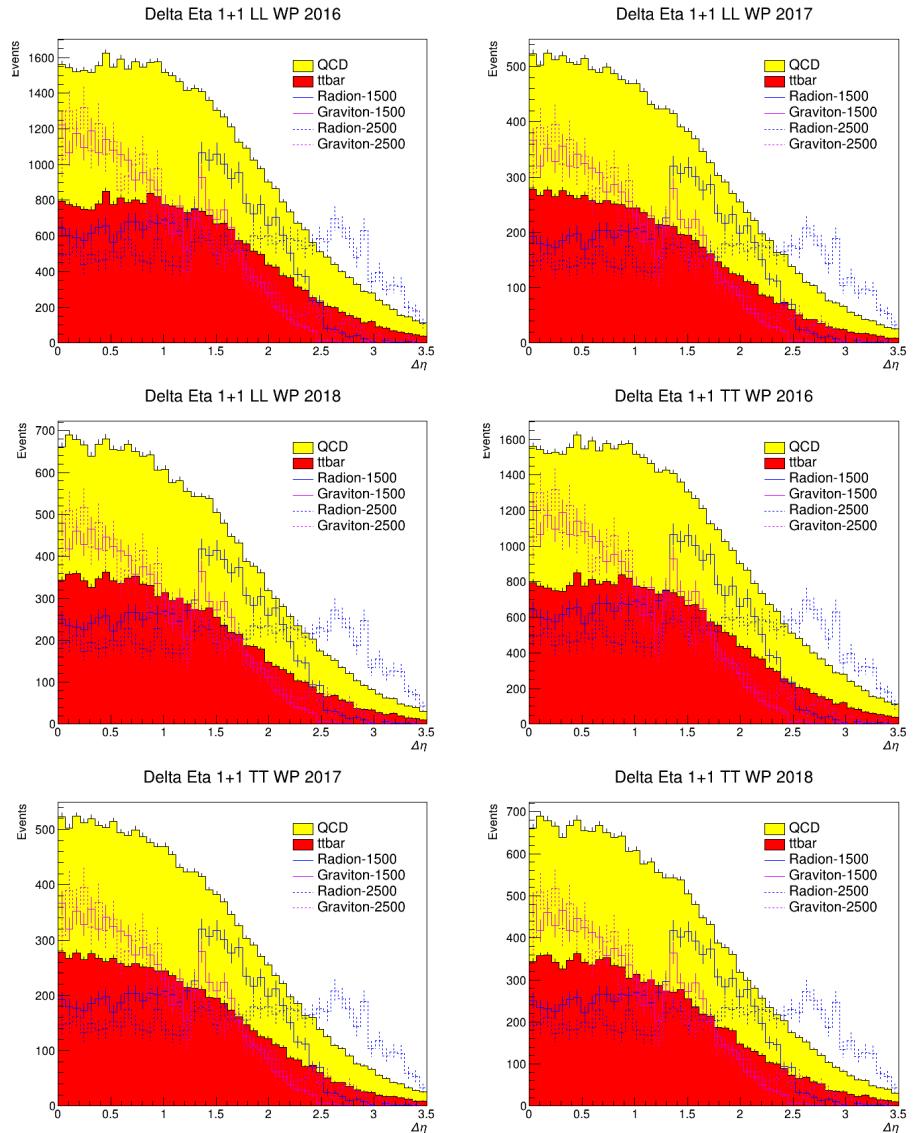


Figure A-31. Pre deepTagMD_HbbvsQCD selection $\Delta\eta$ distribution

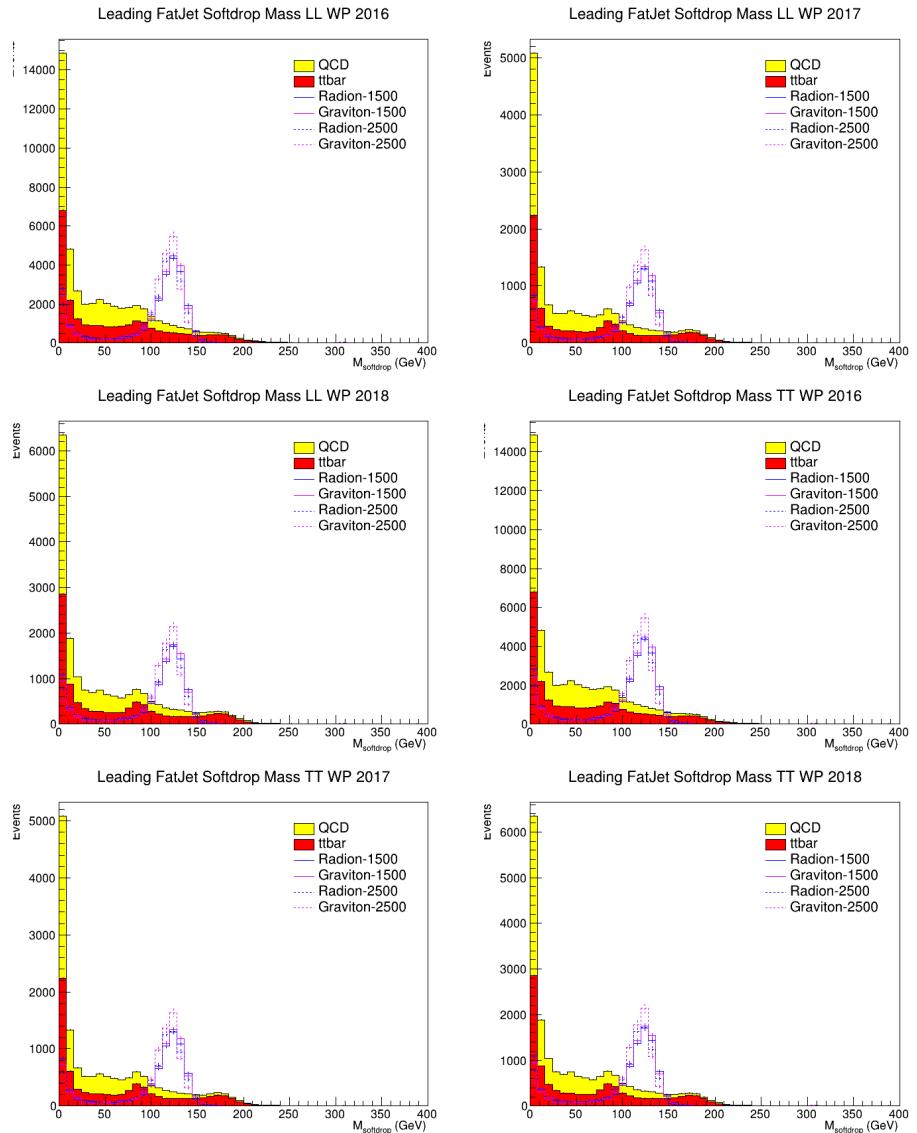


Figure A-32. Pre deepTagMD_HbbvsQCD selection m_j distribution

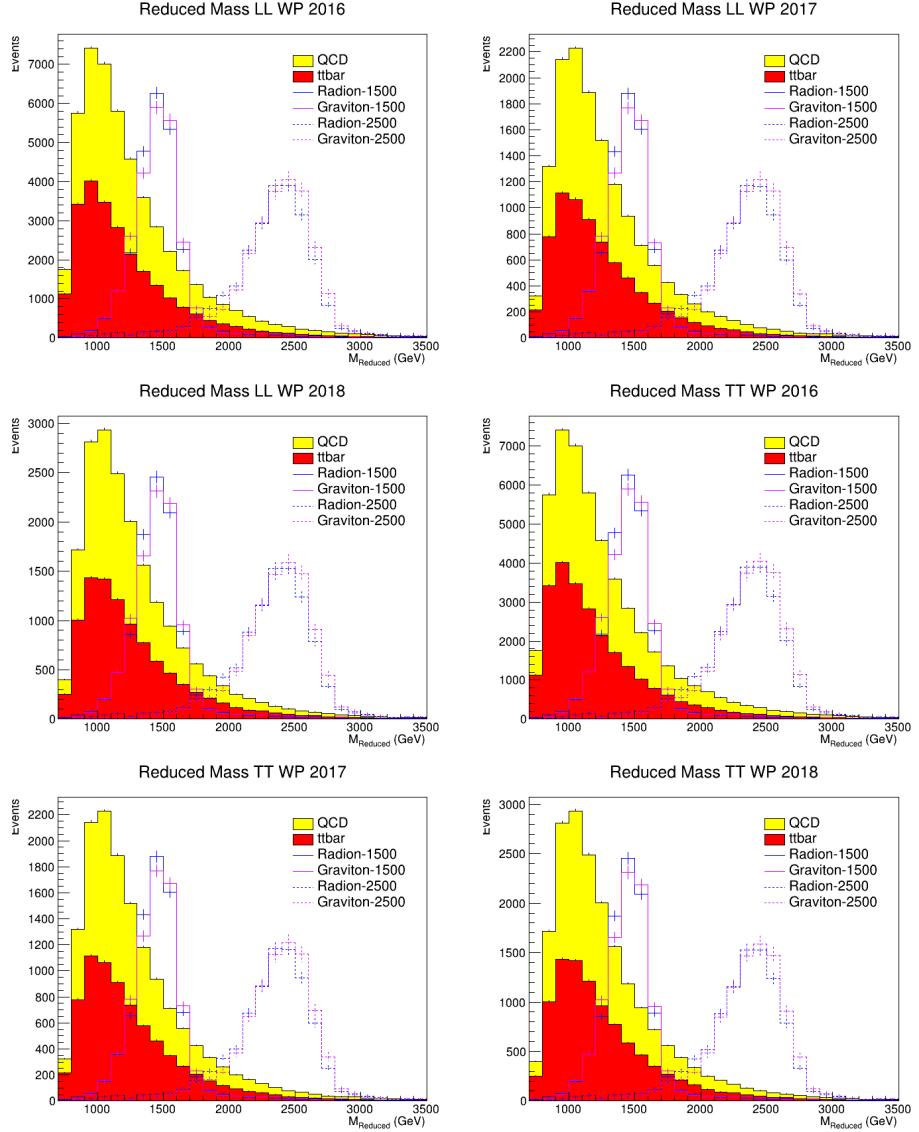


Figure A-33. Pre deepTagMD_HbbvsQCD selection m_{jj} distribution

A.6 Limits by Year

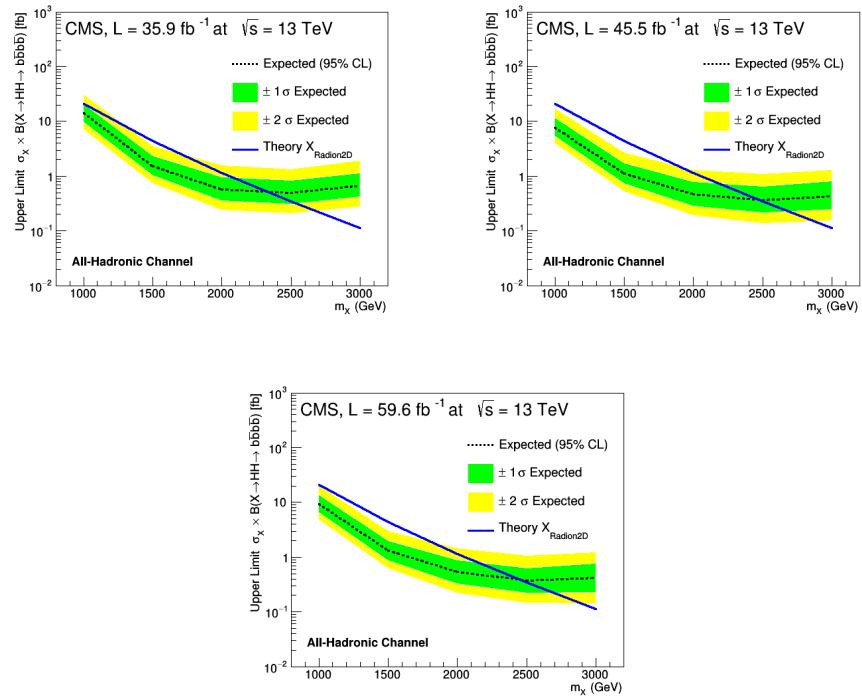


Figure A-34. 2016,2017,2018 Radion Limits for 1+1 and 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.

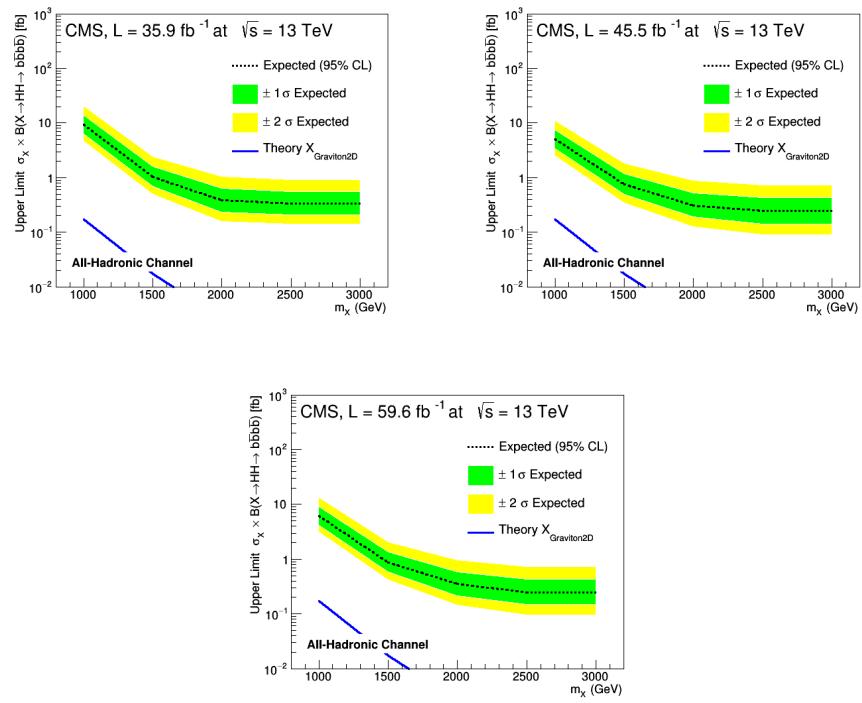


Figure A-35. 2016,2017,2018 Graviton Limits for 1+1 and 2+1 using the Deep AK8 Mass Decorrelated Hbb Tagger.

A.7 Deep AK8 MD Hbb Tagger Working Point Study

The following study was conducted to optimize the working points for the boosted and semi-resolved analysis since official working points are not available for this tagger. We scan the deepAK8 score and find that we should choose a working point as close to 1 as possible without totally eliminating the background in order to do a background estimate. We originally chose 0.95 for the TT working point but ended up relaxing it to 0.9 because of the extreme nature of the cut. Here we present just the 1500 GeV mass point for Narrow Radion and Bulk Graviton but the other mass points have the same conclusion. We also do the same study with just $t\bar{t}$ as a background to make sure the same conclusion is reached, which it is.

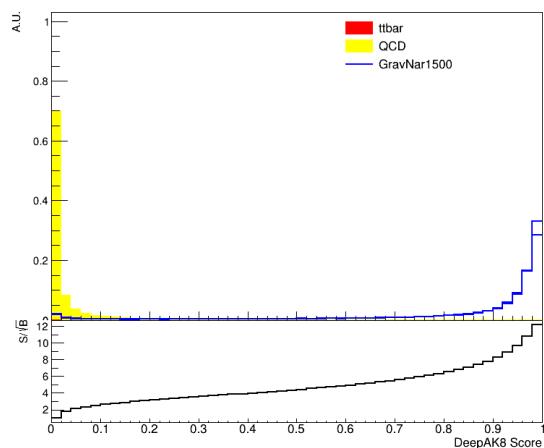
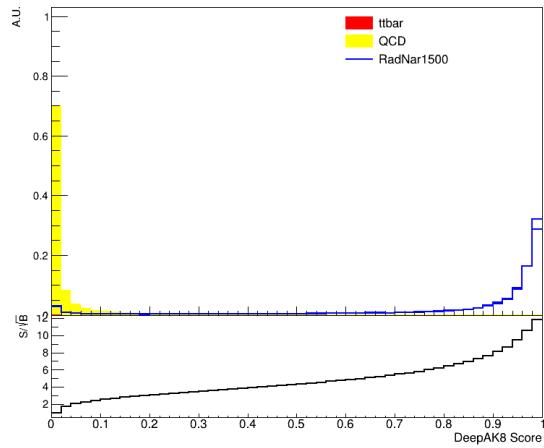


Figure A-36. Optimization of dak8MDHbb tagger using S/\sqrt{B} for Narrow Radion and Bulk Graviton 1500 Gev.

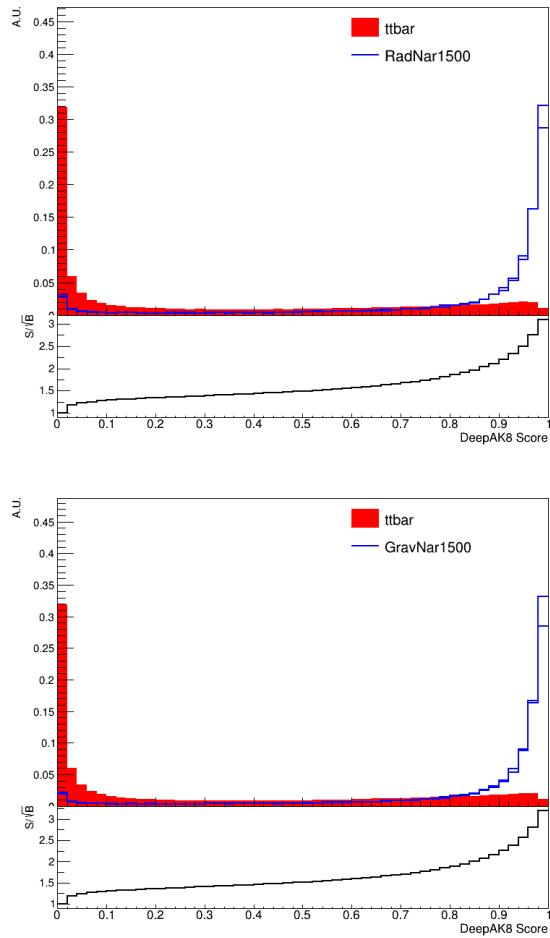


Figure A-37. Optimization of dak8MDHbb tagger using S/\sqrt{B} for Narrow Radion and Bulk Graviton 1500 GeV with just $t\bar{t}$ as a background.

A.8 Trigger Scale Factor

In previous version of this analysis, the trigger efficiency in data was used to correct the trigger response in MC. The latest method is documented in 3.4. Here we show the comparison of the two methods.

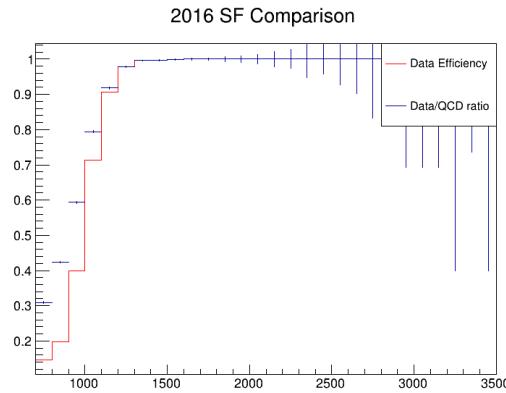


Figure A-38. Comparison of trigger scale factors for 2016.

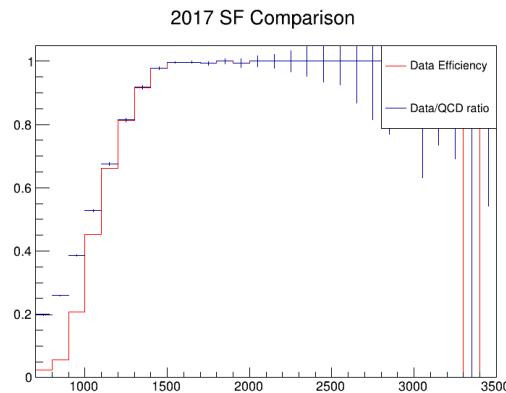


Figure A-39. Comparison of trigger scale factors for 2017.

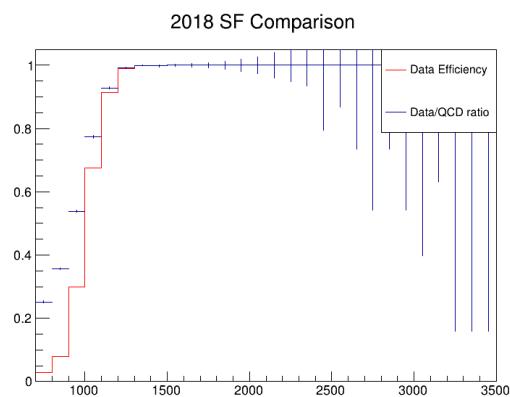


Figure A-40. Comparison of trigger scale factors for 2018.

A.9 Jet Energy Resolution Template Examples

Here we show example Jet Energy and Mass correction templates of both resolution and scale for Radion at 1500 GeV and $t\bar{t}$, for the year 2016 and the TT region as an example.

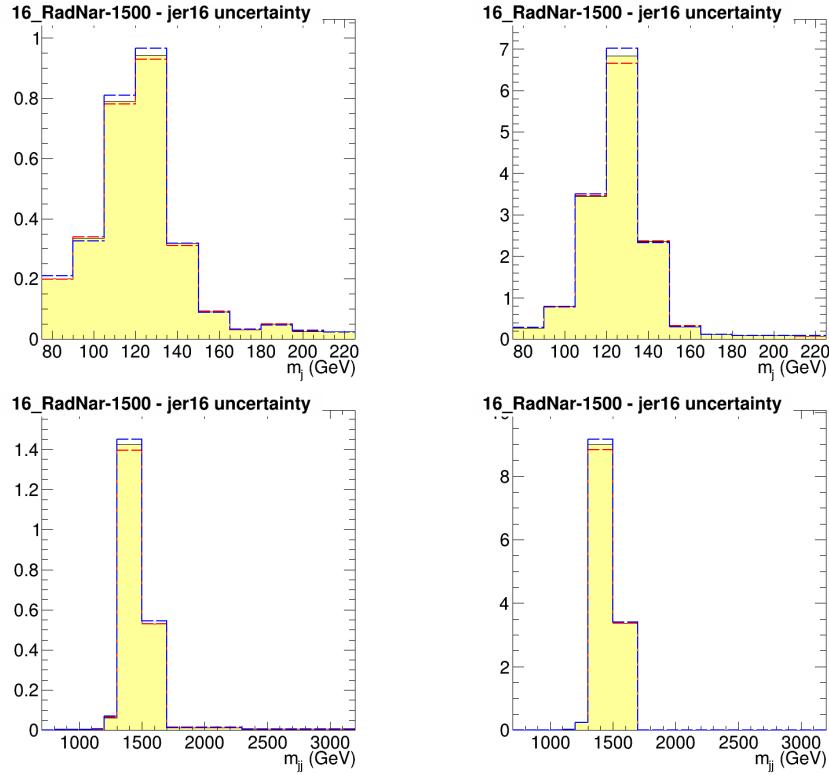


Figure A-41. 2016 Narrow Radion 1500 GeV JER pass and fail templates for M_j axis and M_{jj}^{red} axis.

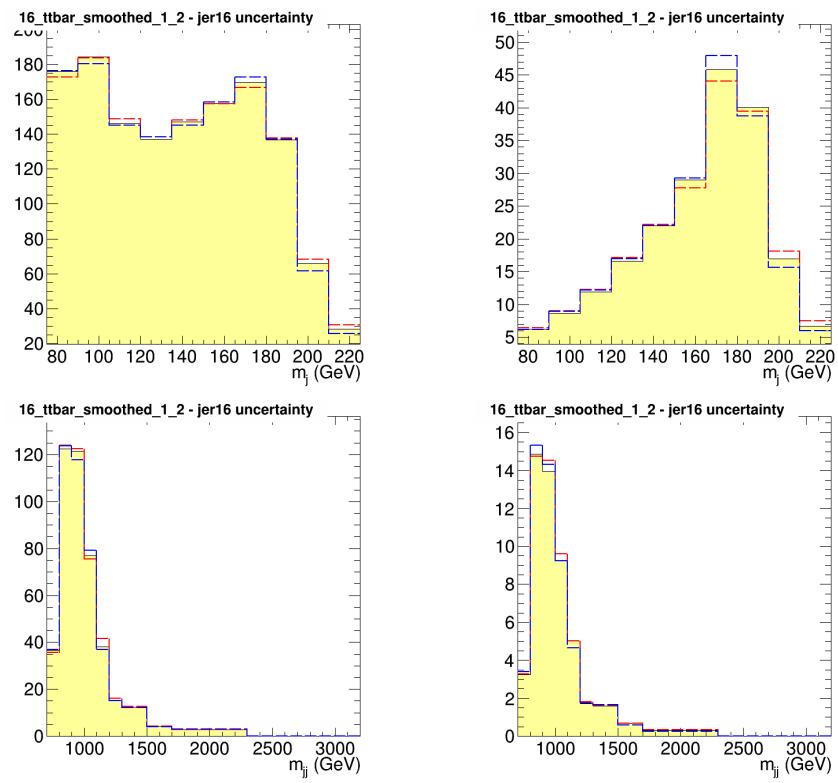


Figure A-42. 2016 $t\bar{t}$ JER pass and fail templates for M_j axis and M_{jj}^{red} axis.

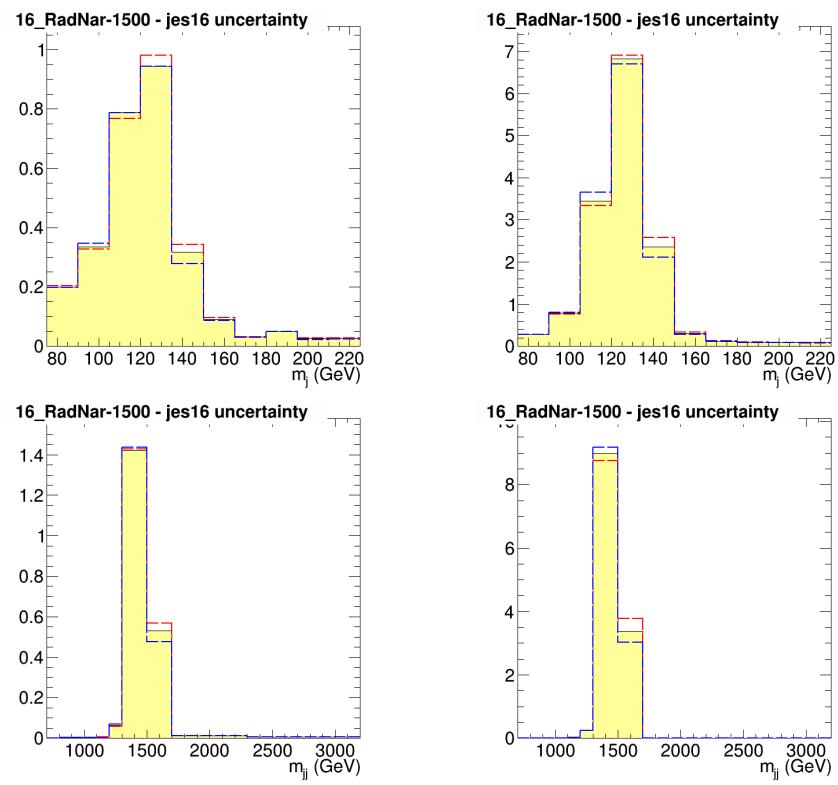


Figure A-43. 2016 Narrow Radion 1500 GeV JES pass and fail templates for M_j axis and M_{jj}^{red} axis.

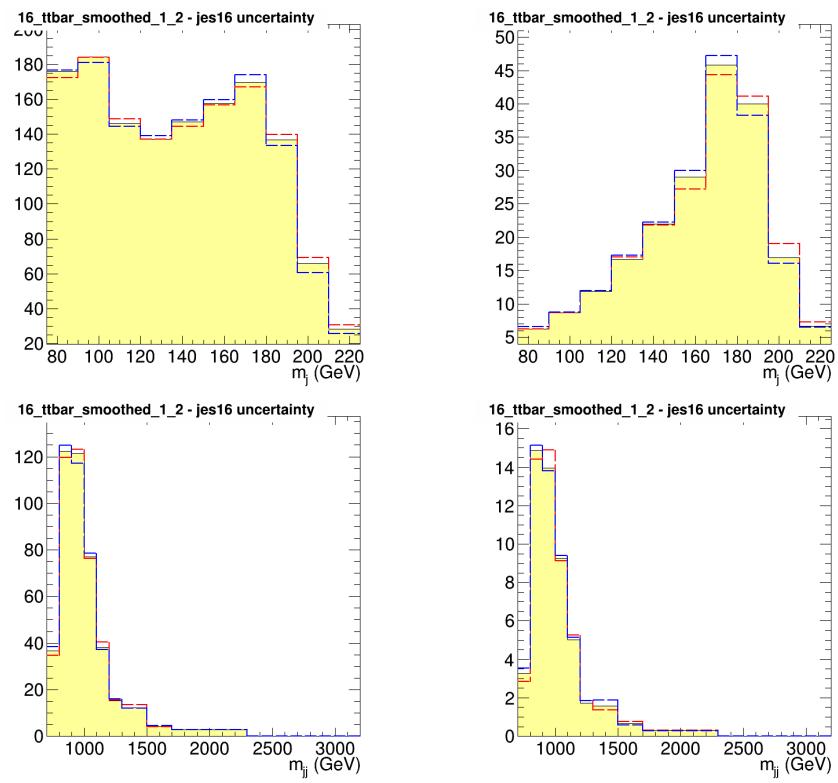


Figure A-44. 2016 $t\bar{t}$ JES pass and fail templates for M_j axis and M_{jj}^{red} axis.

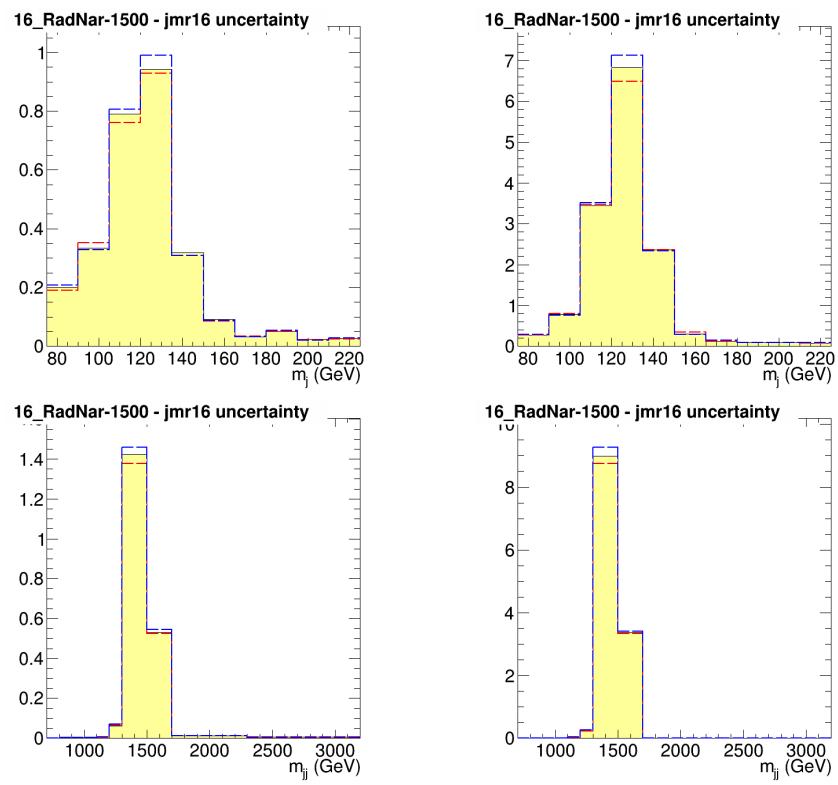


Figure A-45. 2016 Narrow Radion 1500 GeV JMR pass and fail templates for M_j axis and M_{jj}^{red} axis.

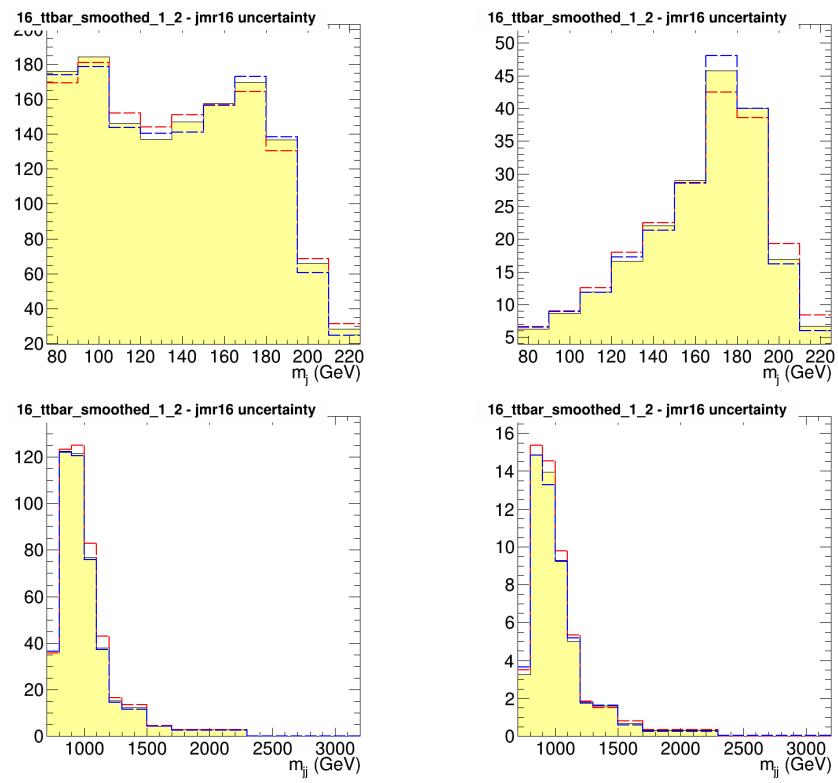


Figure A-46. 2016 $t\bar{t}$ JMR pass and fail templates for M_j axis and M_{jj}^{red} axis.

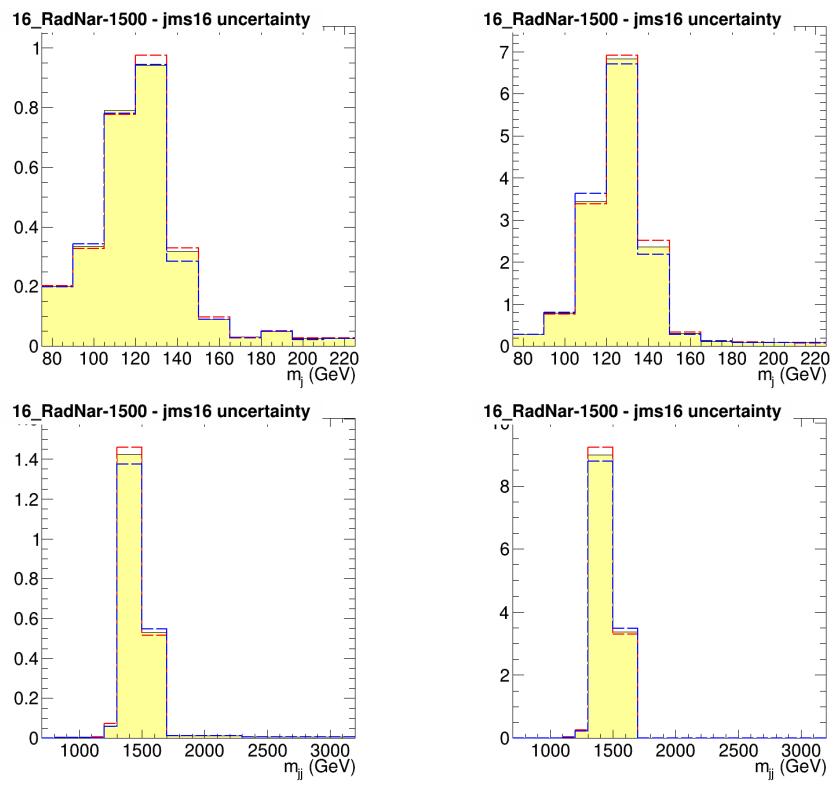


Figure A-47. 2016 Narrow Radion 1500 GeV JMS pass and fail templates for M_j axis and M_{jj}^{red} axis.

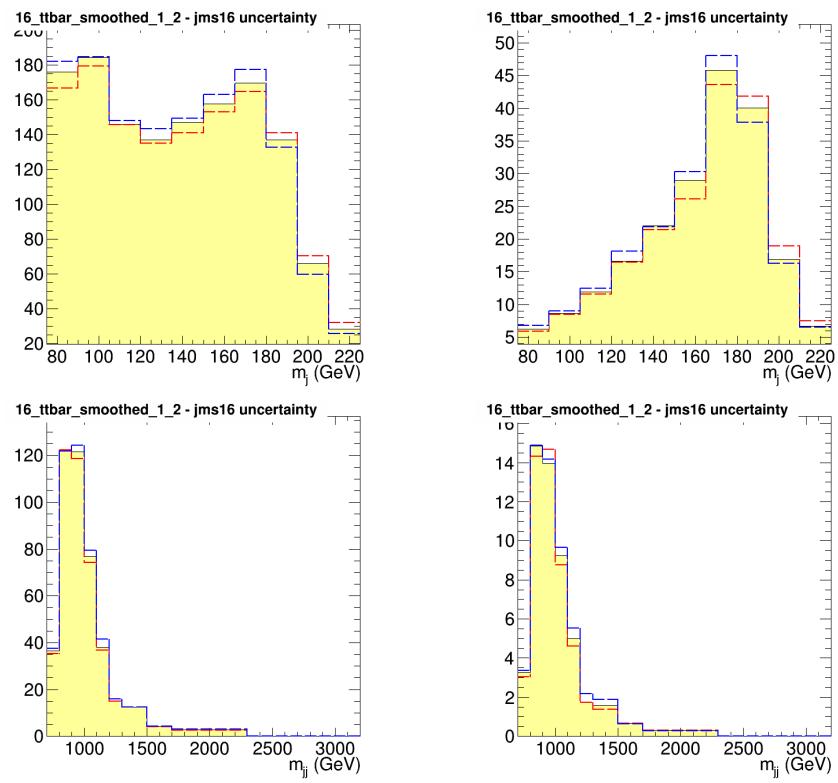


Figure A-48. 2016 $t\bar{t}$ JMS pass and fail templates for M_j axis and M_{jj}^{red} axis.

A.10 M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut studies

We studied the effect on both the signal efficiency and the limits of introducing a M_{jj}^{red} dependent $\Delta\eta_{jj}$ cut in order to try to capture any lost signal, particularly at higher resonance masses. We concluded that although the signal efficiency for each region and resonance mass point increases, it does not do so enough to make a difference in the expected limits and therefore we do not implement this change.

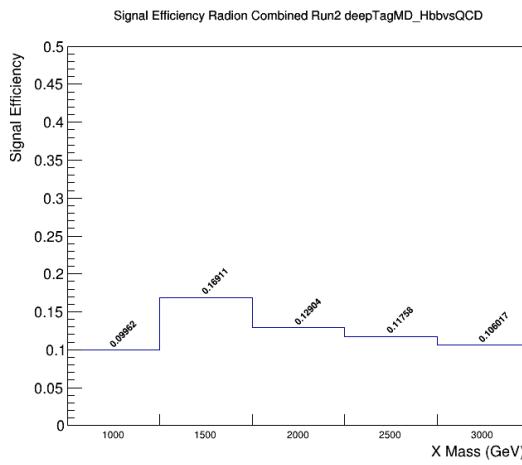


Figure A-49. Radion Signal Efficiency for full Run 2.

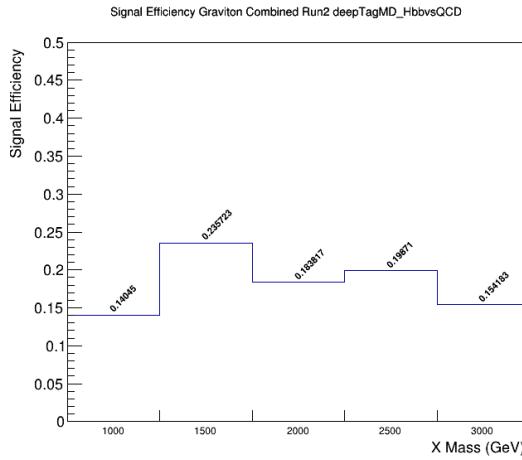


Figure A-50. Radion Signal Efficiency for full Run 2.

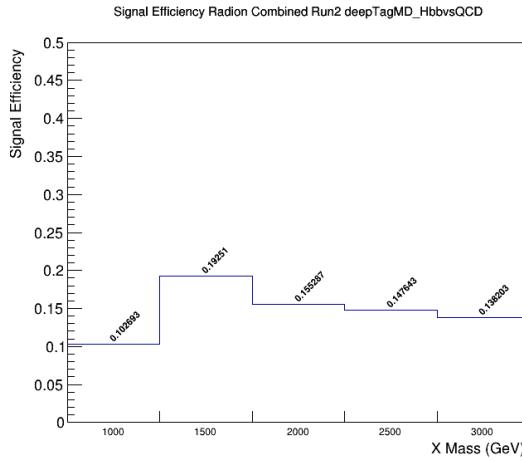


Figure A-51. Radion Signal Efficiency for full Run 2 with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.

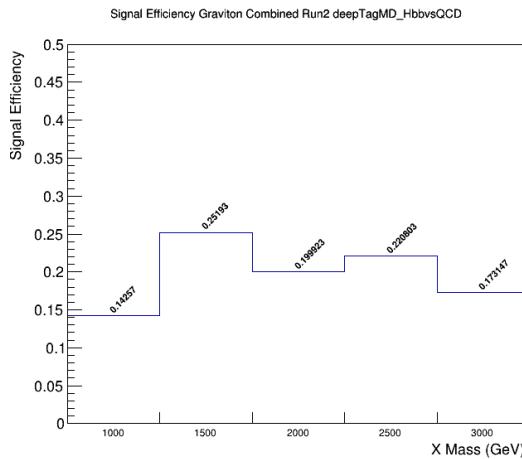


Figure A-52. Radion Signal Efficiency for full Run 2 with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.

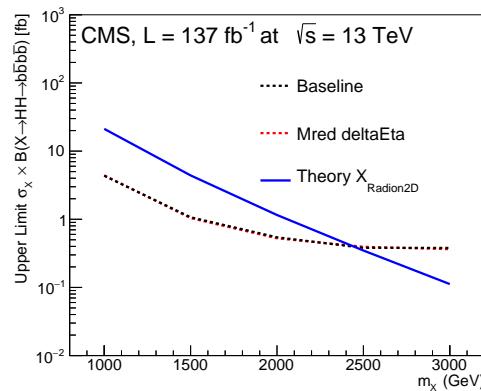


Figure A-53. Full Run 2 Radion Limit comparison with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.

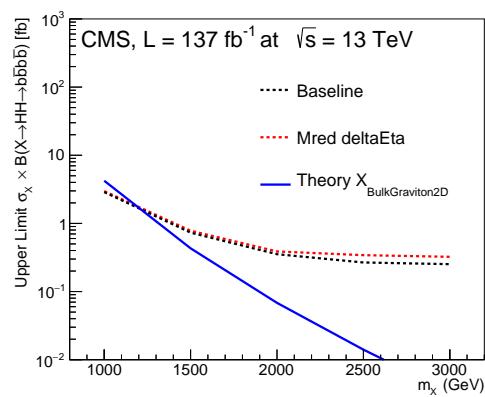


Figure A-54. Full Run 2 Graviton Limit comparison with M_{jj}^{red} Dependent $\Delta\eta_{jj}$ Cut.

A.11 Wide Resonances

We perform this analysis for resonances with a higher width as a check of the validity of the background estimate. Here we show both the unblinded results for the bulk graviton and radion signal samples with a 10% width.

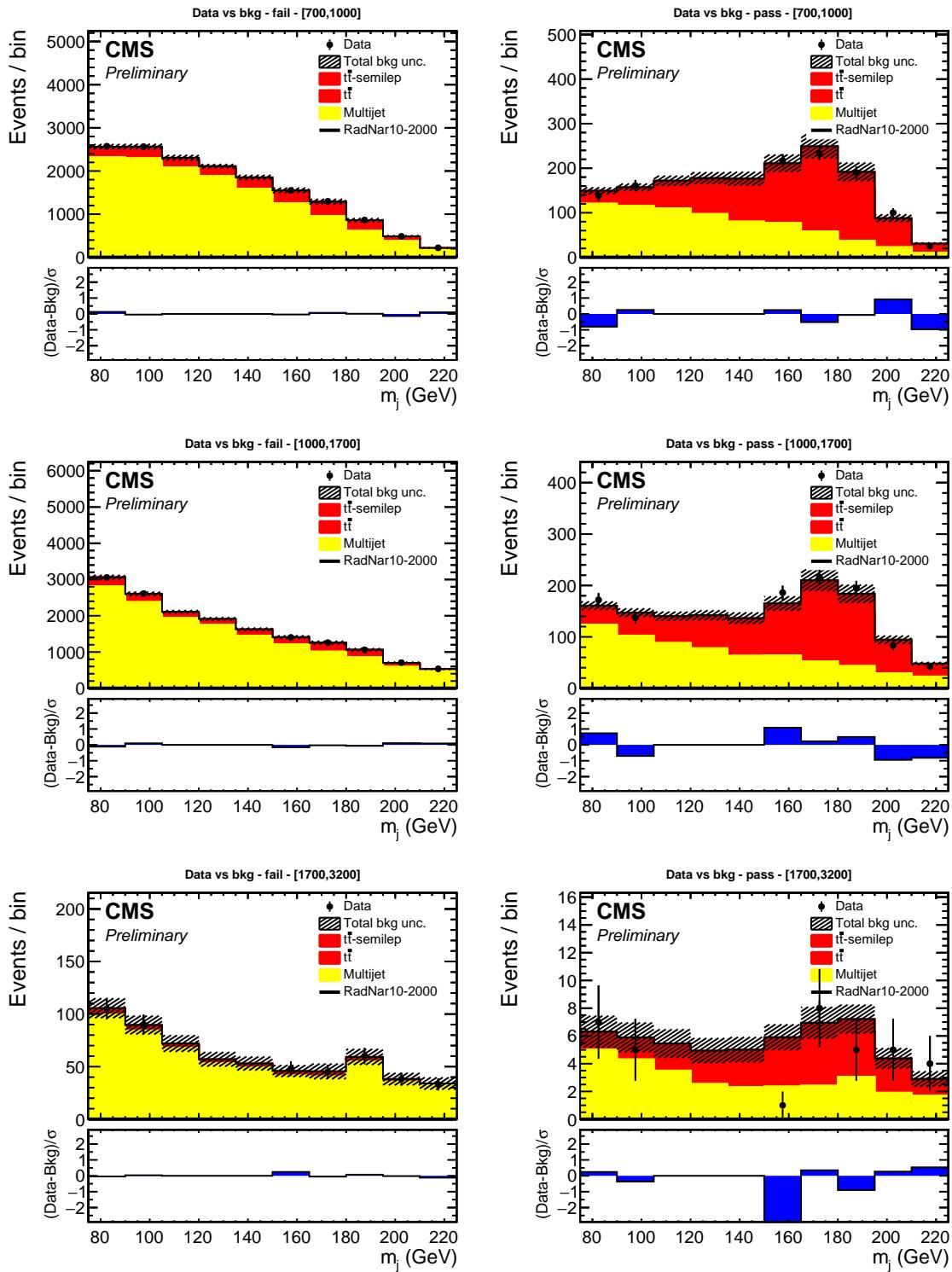


Figure A-55. Full Run 2 Loose Loose fits for M_j axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.

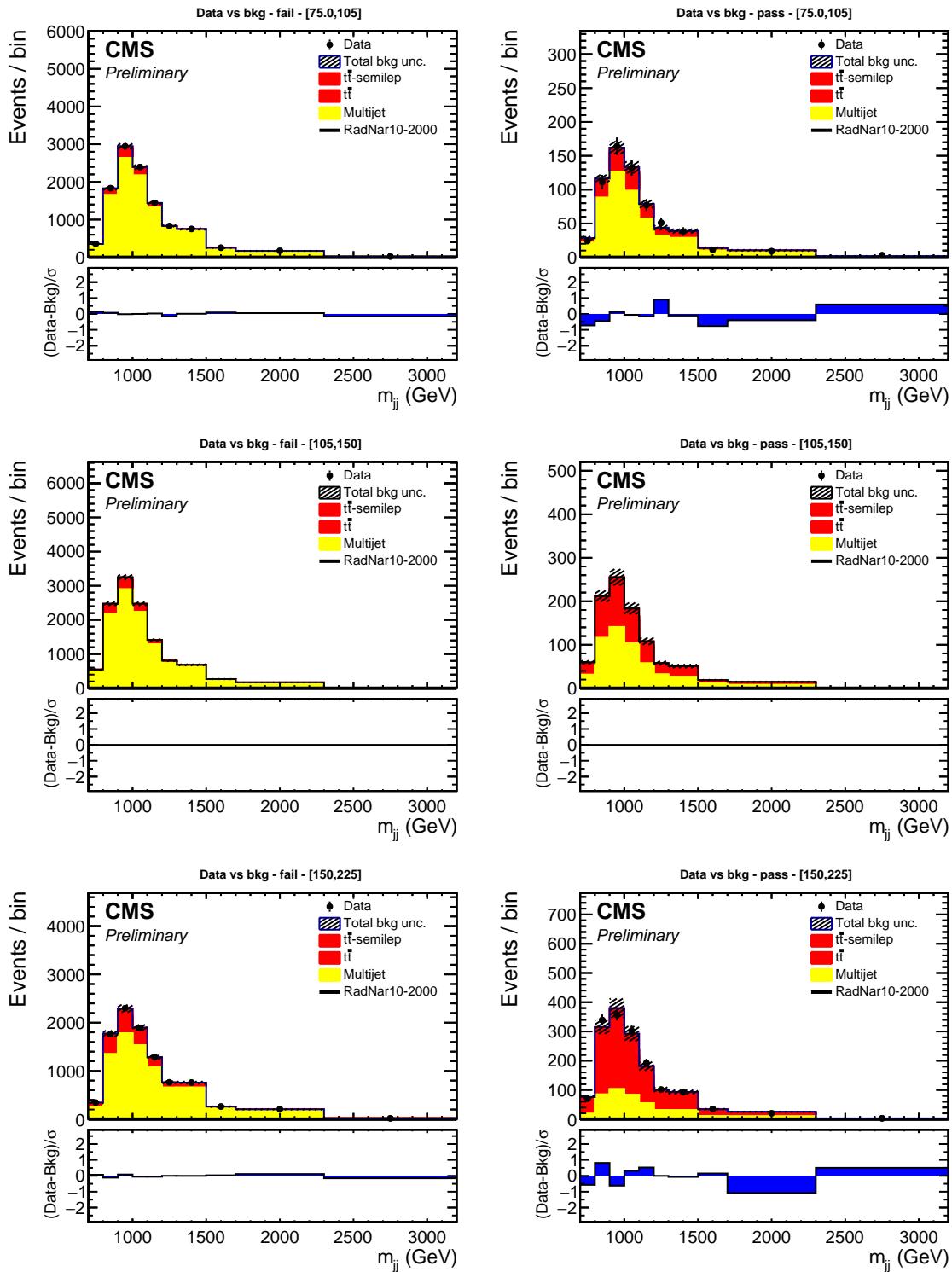


Figure A-56. Full Run 2 Loose Loose fits for M_{jj}^{red} axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.

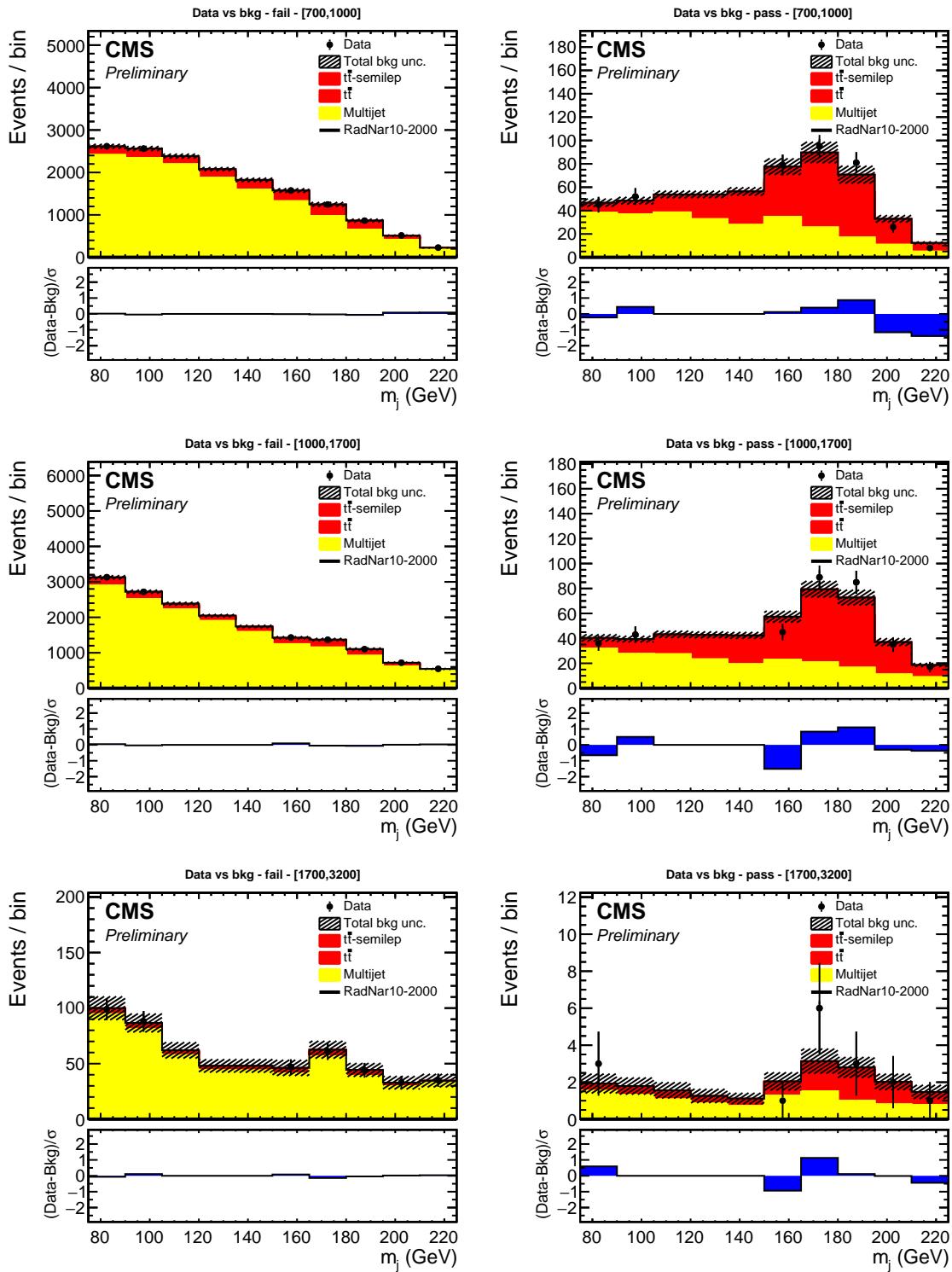


Figure A-57. Full Run 2 Tight Tight fits for M_j axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.

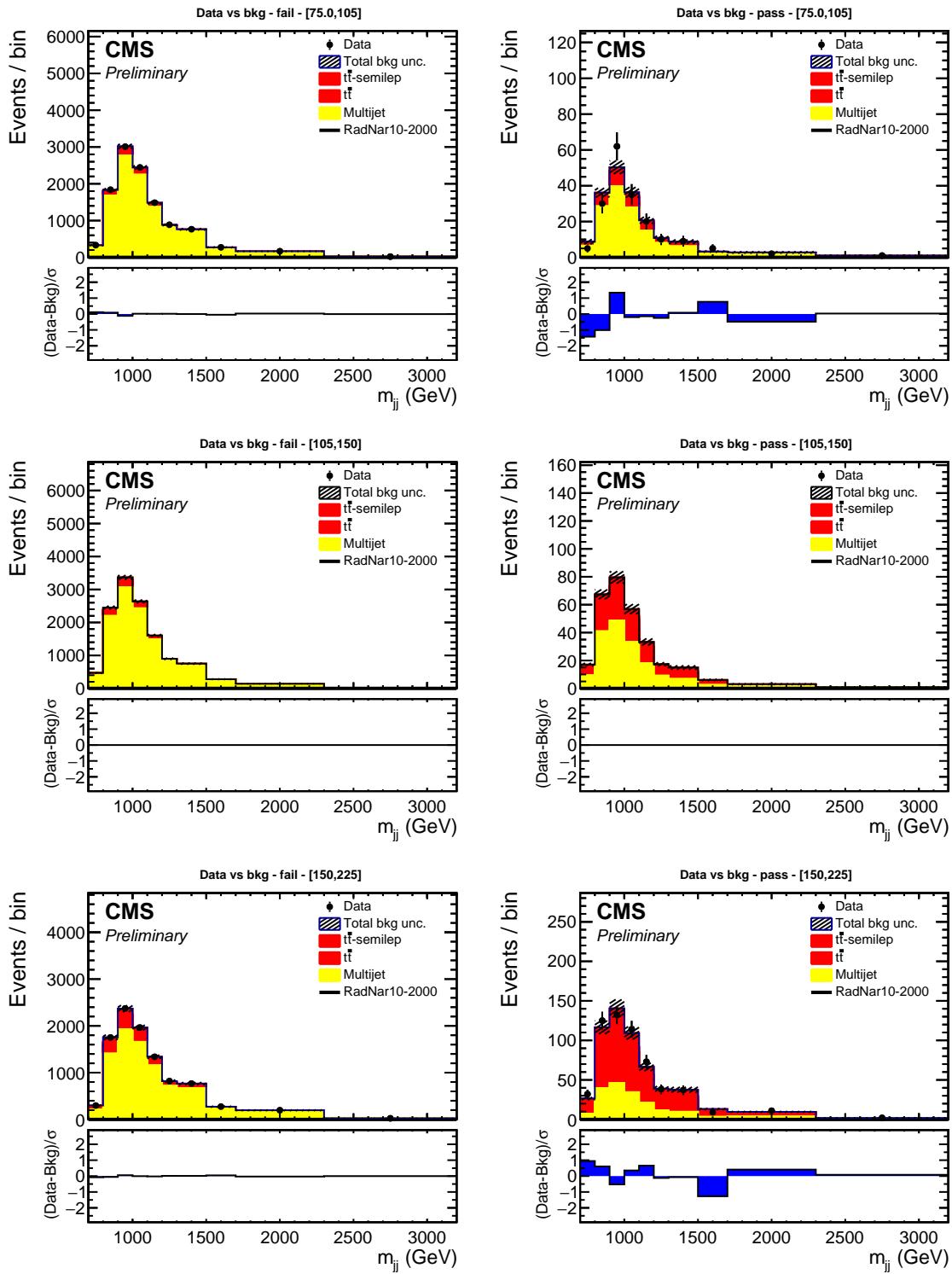


Figure A-58. Full Run 2 Tight Tight fits for M_{jj}^{red} axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.

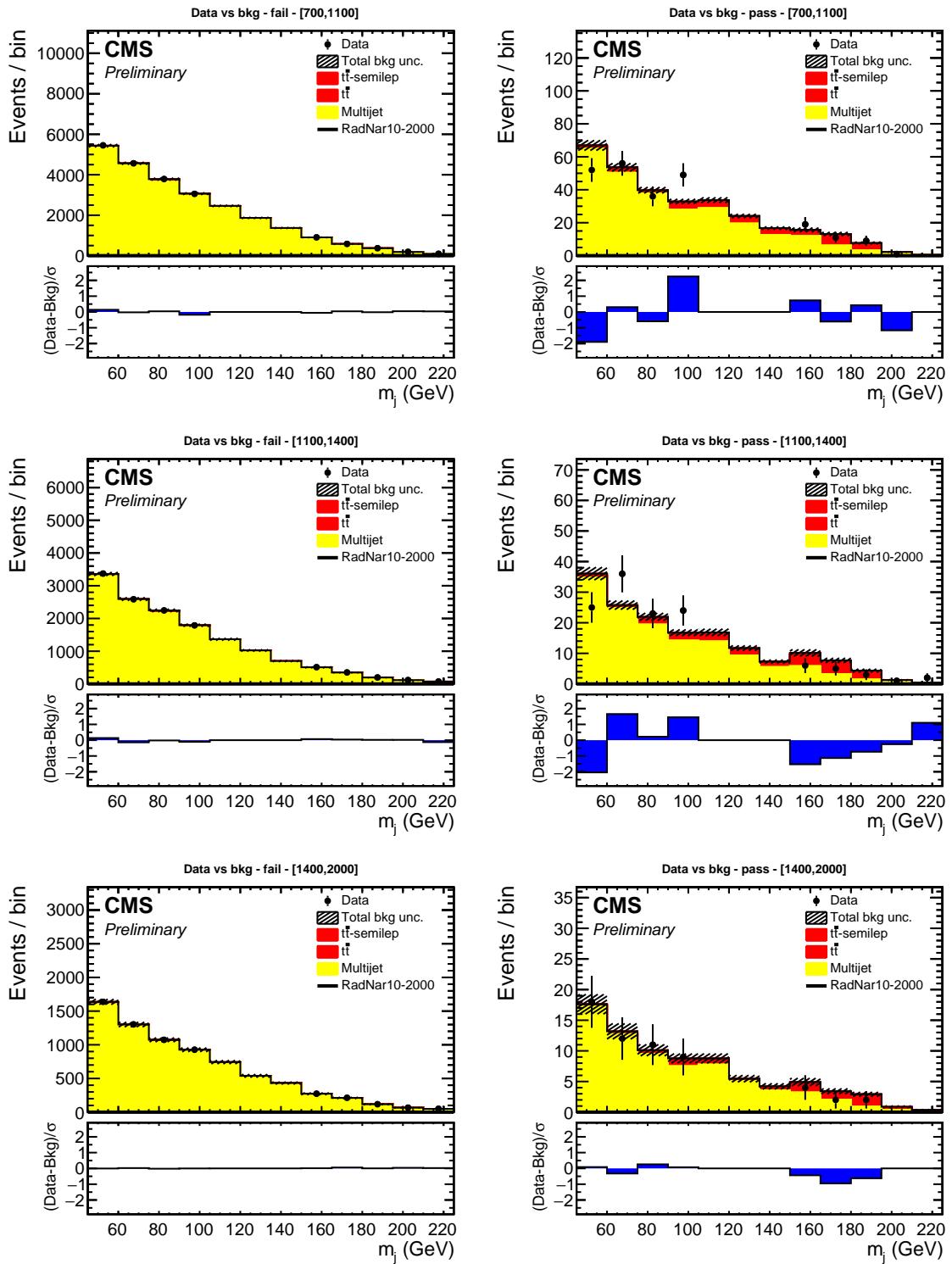


Figure A-59. Full Run 2 2+1 fits for M_j axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.

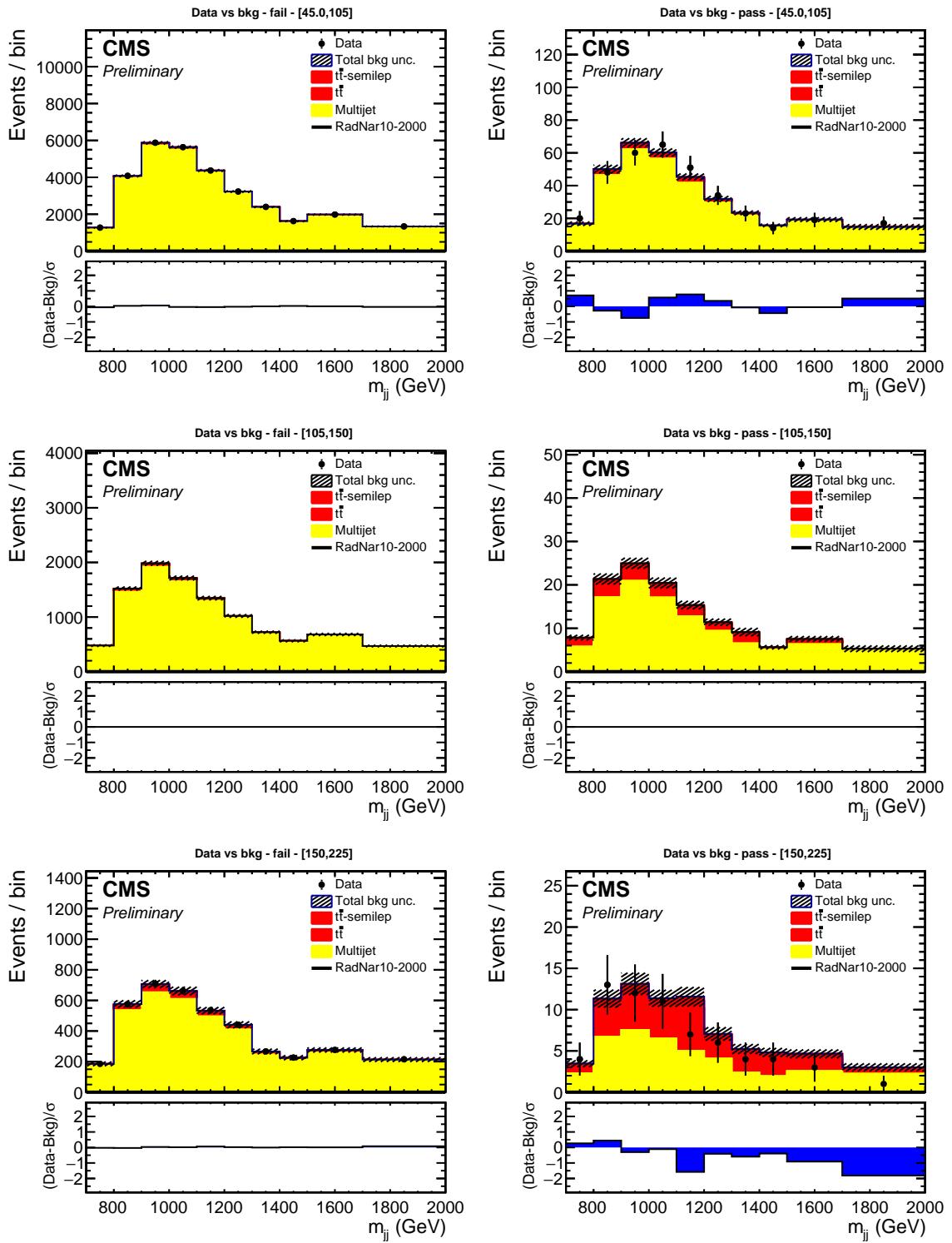


Figure A-60. Full Run 2 2+1 fits for M_{jj}^{red} axis including expected Radion 2000 GeV wide signal, normalized to the signal strength found by the fit.

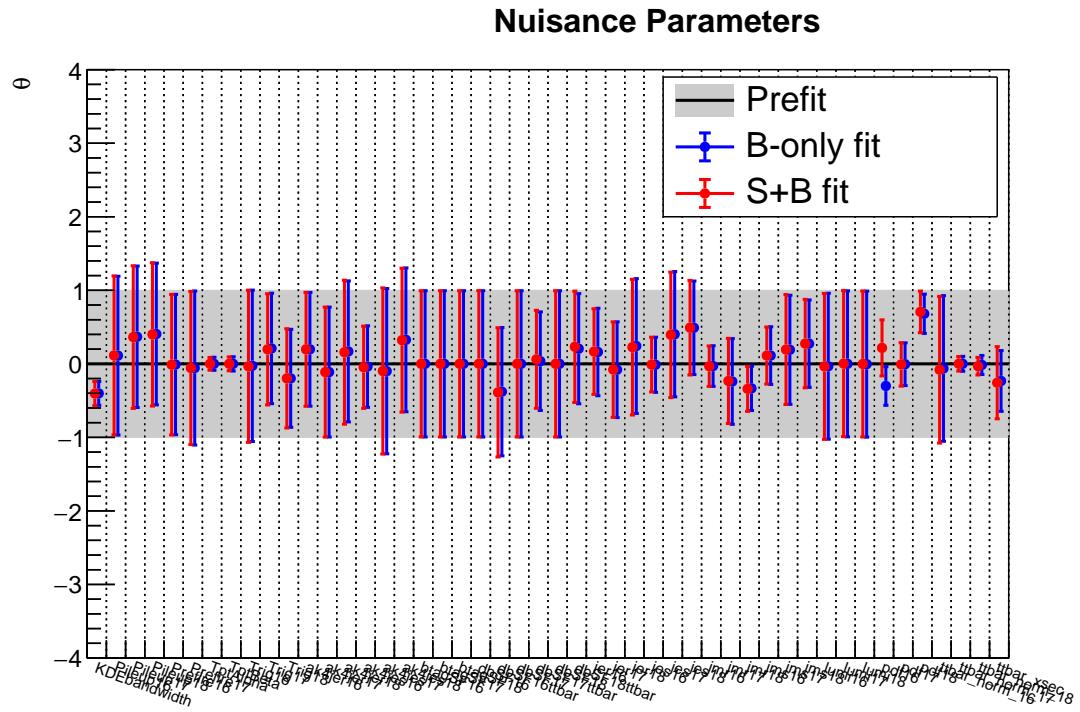


Figure A-61. Full Run 2 Nuissance Pulls Plot for 2 TeV wide Radion signal.

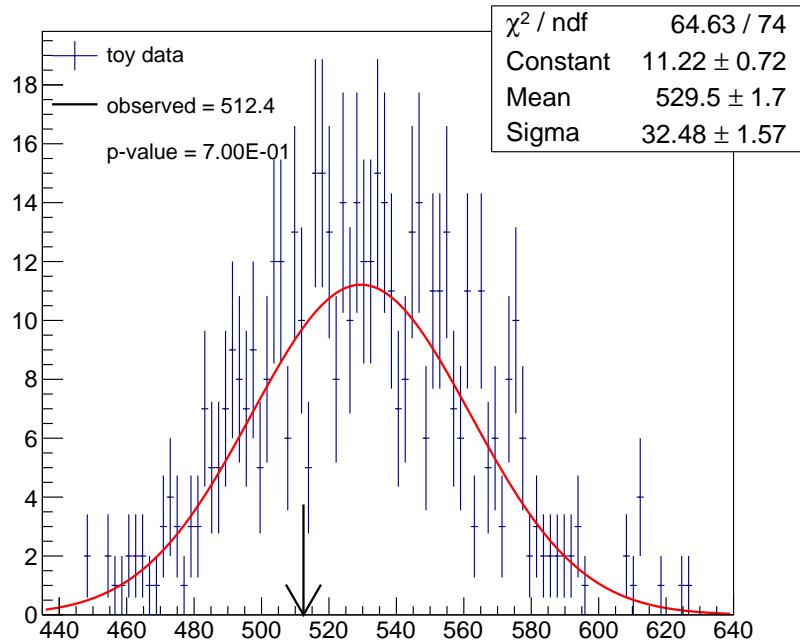


Figure A-62. Full Run 2 Goodness of Fit test plot with 500 toys for 2 TeV wide Radion signal.

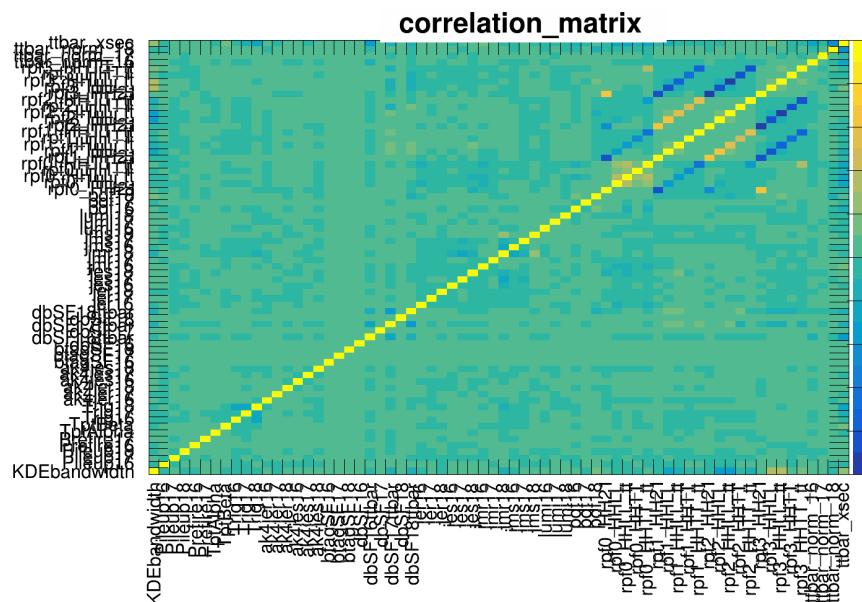


Figure A-63. Full Run 2 Correlation Matrix for 2 TeV wide Radion signal.

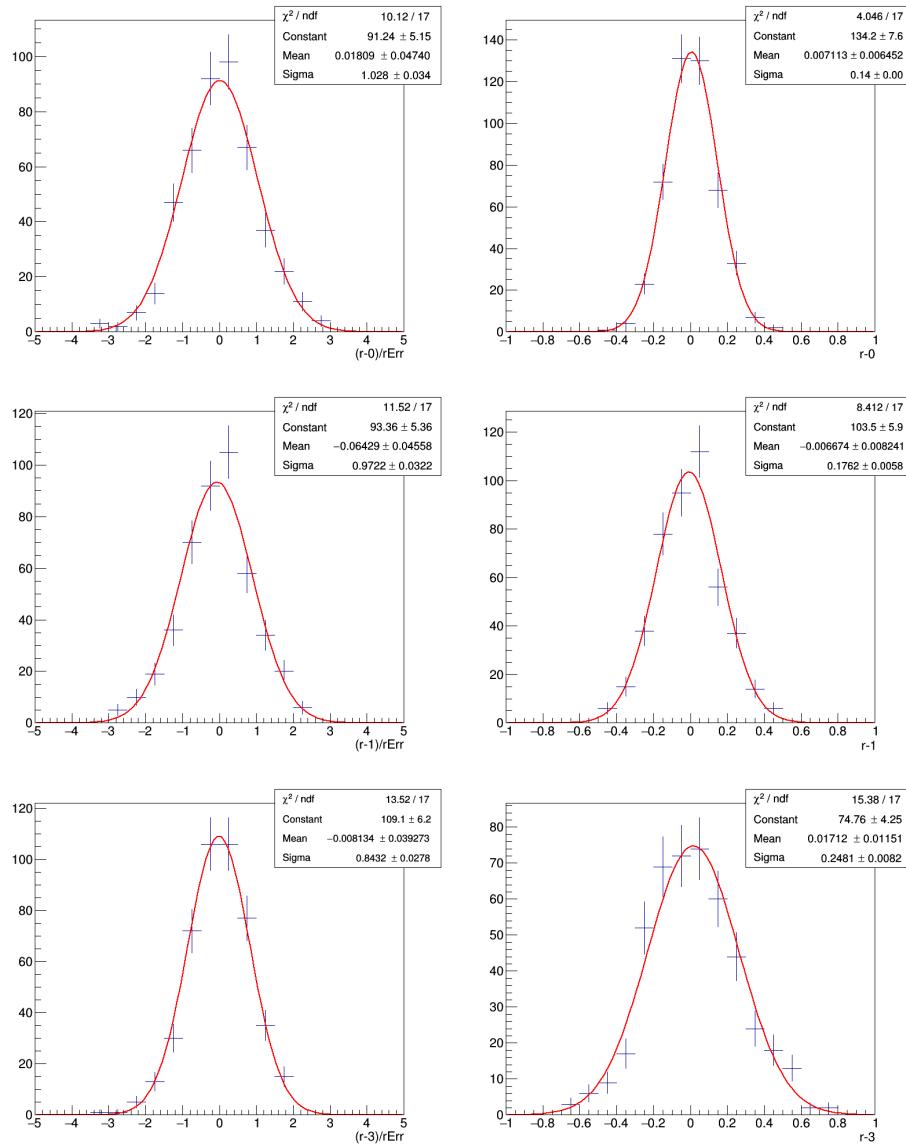


Figure A-64. Wide Signal (Left Column is Pull and Right is Strength) Injection for Full Run 2 at 1 TeV for $r = 0, 1, 3$ with 500 toys.

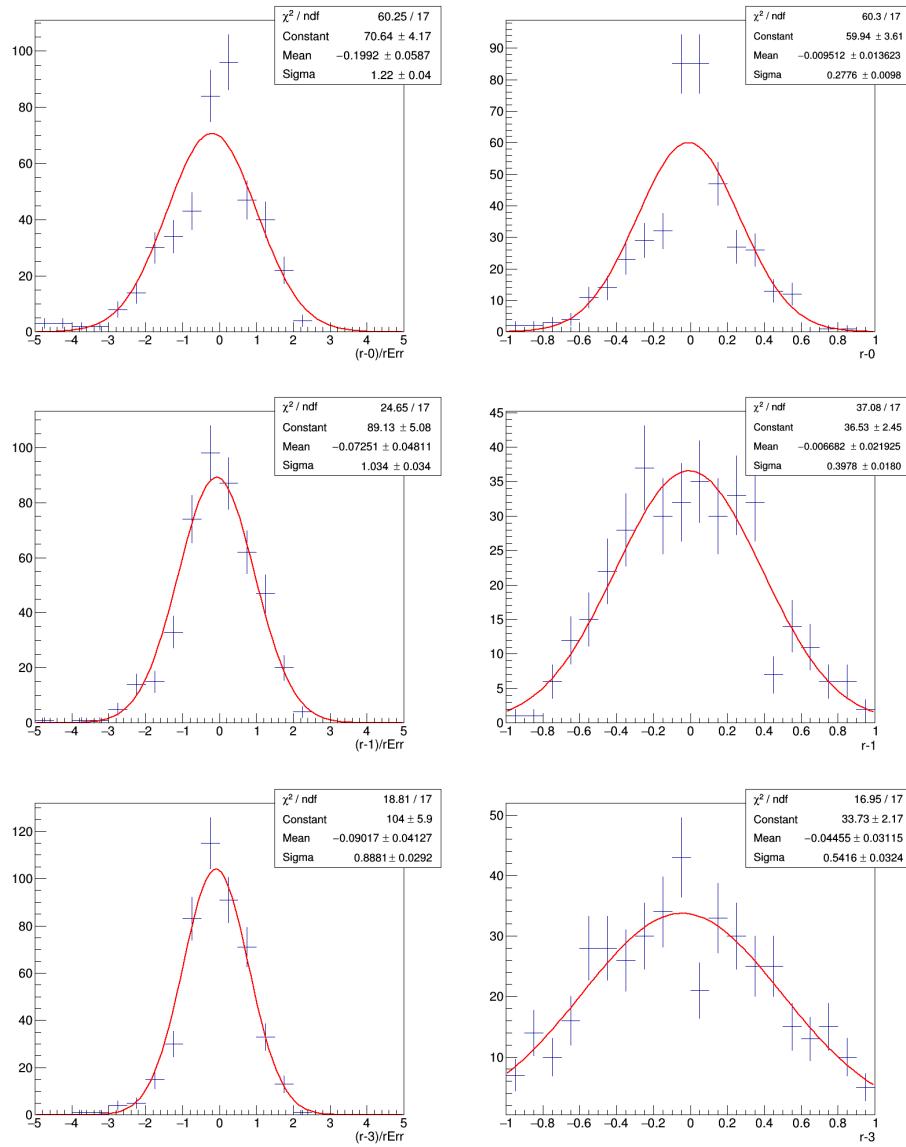


Figure A-65. Wide Signal (Left Column is Pull and Right is Strength) Injection for Full Run 2 at 2 TeV for $r = 0, 1, 3$ with 500 toys.

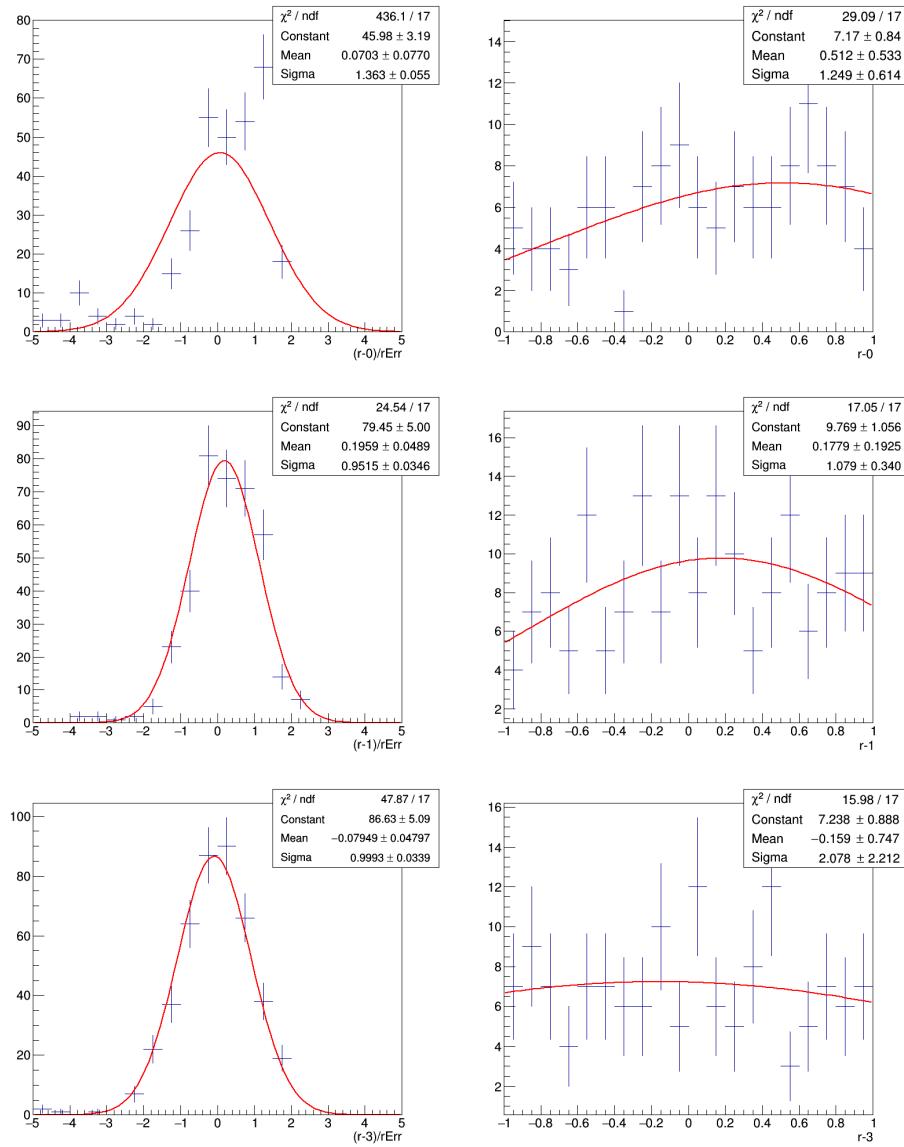


Figure A-66. Wide Signal (Left Column is Pull and Right is Strength) Injection for Full Run 2 at 3 TeV for $r = 0, 1, 3$ with 500 toys.