



Search for dark matter produced in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS detector using full Run2 data

A thesis submitted for the degree of
Doctor of Philosophy
in the Faculty of Science

by
Praveen Chandra Tiwari
under the guidance of
Dr. Jyothsna Rani Komaragiri

*Centre for High Energy Physics
Indian Institute of Science
Bangalore - 560012. India.*



To

My Teachers,

My Family and

My Friends

DECLARATION

I hereby declare that the work presented in this thesis entitled “**Search for dark matter produced in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS detector using full Run2 data**” is the outcome of studies and analyses performed by me under the supervision of Dr. Jyothsna Rani Komaragiri at the Centre for High Energy Physics, Indian Institute of Science, Bengaluru, India. The work presented in this thesis has not been submitted elsewhere for obtaining any degree or diploma in any university or institute. Due citations and acknowledgements have been made wherever the work of other investigations have been described.

Dated: 30 July 2021



Praveen Chandra Tiwari

Certified by:

Dr. Jyothsna Rani Komaragiri
Centre for High Energy Physics
Indian Institute of Science
Bengaluru 560012
India



Dr. Jyothsna Komaragiri
Centre for High Energy Physics
IISc Bengaluru - 560 012

ACKNOWLEDGEMENTS

This long and satisfying journey of PhD would not have been possible unless the contributions from innumerable people. With immense pleasure and gratitude, I am taking this opportunity to express my appreciation to them.

Focus towards the objective and advancement in the right direction are the elementary factors to be imbibed in a highly competitive academic environment of the Indian Institute of Science; certain personalities tend to make a lasting impression on a grad student's fresh and naive mind.

First of all, I would like to express my deepest heartfelt thanks to my Research Supervisor, Dr Jyothsna Rani Komaragiri, for her continuous support, guidance and constant motivation. She always had made sure that our work gets the topmost priority irrespective of all the logistical and other challenges faced and stood like a pillar to support the IISC CMS group.

This is an opportunity for me to thank all other faculty members of the department who have been encouraging and helpful at various times and on occasions.

As this thesis involves an experimental search for dark matter particles, I have been fortunate to be supported by the best people in the field.

With immense gratitude, I would like to thank my collaborator, Dr Raman Khurana, who thoroughly supported me throughout my PhD journey. His guidance has helped me tremendously learn the subject and acquire the necessary research skills. I shall always remain enormously indebted to him for providing me with an opportunity to excel in the subject. His enthusiastic encouragement and physics discussions made my PhD journey a memorable one. I thank him for all the enthu-

siasm he invoked in me during my research work. This thesis would not have been possible without the immense support I received from him in every phase.

During my doctoral work, I have been associated with some of the best minds in many fields. I would like to duly acknowledge my colleagues Lata, Deepak, Spandan, Zoheb and my friends Jitesh, Gaurav, Gagan, Mayur, Ayushi, Jagannath, Ibrahim, Nimisha, Debasmita, Sudipta, Amit, Priyanka Bajaj, Dr Priyanka Jayal, Sukriti, Rahool, Deepak, Priyanka Kushwah,

I am also thankful to Mr Keshava, Mr Sarvana and Ms Aruna, the office staff of the CHEP, who were most caring and patient in guiding me through the office procedures.

I want to express my deepest regards to my school teachers, Kavindra Digari , Pradeep Mehta, Vinaydeep Punetha, K S Samant, Jyothsna and all the others who showed me the right path and direction for pursuing physics.

I want to finish this acknowledgement on a high note by offering my sincere thanks to my mother, Mrs Daya Tiwari and my father, Mr Trilok Chndra Tiwari, who has stood by me through thick and thin.

I want to thank my brother Mr Bhuwan Tiwari and my sisters Mrs Sarita Joshi, Mrs Hansha Pant, Mrs Shobha Sethi, and Mrs Kavita Tiwari, who have always supported and motivated me to reach this stage. I salute you for all the selfless love, care, pain and sacrifice you did to shape my life.

I am forever grateful to my wife, Dr Shweta Pant, to consistently motivate me to do what I desire and cheer me up when I feel low. Her love has always made me feel secure and given me the courage to pursue my dreams.



My sincere regard goes to my father-in-law Mr Kamlesh Pant and mother-in-law Mrs Usha Pant for their love and moral support. I am also grateful to my other family members and friends who have supported me along the way. I would like to sincerely acknowledge the unconditional support that they have provided for reaching individual milestones.

I thank the Almighty for giving me the strength and patience to work through all these years so that today I can stand proudly with my head held high.

I offer my regards to all who supported me in any respect during the completion of the doctoral work. It is a human trait to miss some personalities who should be acknowledged. The absence of a name must not be perceived as negligence but should be attributed to my absentmindedness.



Preface

The composition of the universe has 5% visible matter (everything by which we are made, we see and we observe i.e. leptons, quarks and bosons), 27% of Dark Matter(DM) and rest 68% the Dark Energy. Although the contribution of the Dark Matter is very large as compared to the visible matter, the dark matter particles have not been detected so far. From Cosmology, there are very clear evidences that suggest the existence of dark matter, for example, the rotational curve of the galaxies, bullet cluster, gravitational lensing etc. Therefore, understanding the origin and nature of dark matter has been one of the most important goal of modern particle physics. In General, three complementary methods has been designed to search dark matter, Direct method(observing low-energy recoils of nuclei induced by the interaction with dark matter), Indirect method (searching decay/self annihilation of dark matter particles in outer space) and Collider searches (produce dark matter in a laboratory).

The dark matter production in the Collider is very challenging because of multiple reasons.

1. The strength of interaction of DM with ordinary particles is unknown to enable its production at a particle collider like Large Hadron Collider(LHC).
2. The mass of the DM candidate creates uncertainty if the centre-of-mass energy at LHC is enough to produce a dark matter candidate
3. The presence of DM can then only be revealed by an imbalance in the trans-

verse momentum in the detector measured as missing transverse momentum.

Several types of particle candidates for DM have been proposed in various physics models, all compatible with the properties of DM observed in astrophysical, cosmological experiments, and the observed relic density of the universe. One of the favoured hypotheses suggests that the DM is in the form of electrically neutral, stable, weakly interacting massive particle often called WIMP, with masses in a range between a few GeV to a few TeV. Such models open up the possibility of searching the DM production at LHC. In the collider search for DM, a pair of WIMP that recoil against visible SM particle(s) is looked at. Due to the neutral nature and negligibly small interaction cross-section with the SM particles, WIMPs become invisible to the particle detectors. The only way to ascertain the presence of the WIMP is inferred via measuring the imbalance in the total momentum in the plane transverse to the colliding beams, referred to as missing transverse momentum ($\vec{E}_T^{\text{miss}}$). This gives rise to a detector signature of $\vec{E}_T^{\text{miss}} + X$, where X is one or more than one SM particles produced in association with DM (photon, jet, pair heavy quarks, Higgs boson, Z/W boson, etc.)

In this thesis, the production of dark matter in association with the pair of bottom quark using the full Run2 data corresponding to the integrated luminosity of 137fb^{-1} has been analysed. The search is interpreted with simplified model which overcome the drawbacks of the traditional EFT models while maintaining a model-independent framework. The physics models for the scope of LHC are provided by ‘ATLAS+CMS Dark Matter Forum’, one of those models that meet the characteristics of generic benchmark mediator models is the 2HDM+a model. The 2HDM+a

model is a canonical 2HDM model with an additional pseudoscalar singlet interacting with dark matter particle.

For the analysis, two signal categories have been designed namely, signal region 1(SR1) and signal region 2(SR2) and then to model the background 6 control regions have been designed corresponding to each signal region. The presence of one (two) AK4 jet(s) in the detector, a high p_T^{miss} and not leptons is the typical experimental signature for the signal, hence the signal regions have these event selections along with pre-selection applied to the full analysis. The SRs are categorized on the basis of number of b-Jets in an event, event with one b-jet in out of two jets goes to SR1 and similarly an event with 2b-jets out of three jets goes to SR2 category. The control regions(CRs) are designed so that the major backgrounds can be modelled in a data driven way. In the CRs, the variable called Hadronic Recoil is defined as the sum of p_T^{miss} and lepton p_T , with the inclusion of lepton p_T the Hadronic Recoil essentially behaves as p_T^{miss} in signal region.

The events for the signal and muon channel control regions are selected with dedicated MET trigger with trigger threshold at around 120-140 GeV. In order to get rid of most of the multijet QCD background contamination and to be above the turnON region of the MET trigger events are selected with $p_T^{\text{miss}} > 250$ GeV. Single electron trigger are used for the electron channel control region. All the scale factors associated with the triggers, lepton IDs, lepton ISO, b-tagging along with electroweak reweighting, top p_T reweighting, NLO k-factors, and Pile reweights are applied to the analysis. Further, all the systematics uncertainties associated with the scale factors are calculated.

For the extraction of signal, a variable with best signal to background separation is taken and a binned profile likelihood fit method has been performed fitting simultaneously over all the control regions and the signal regions with the systematic uncertainties as nuisance parameters in the fit and limits on the mass of lighter pseudoscalar in 2HDM+a have been calculated.

To summarize, the present thesis explores 2HMD+a model in the context of $bb+p_T^{\text{miss}}$ analysis for the first time and put limits on the mass of the lighter and heavier pseudoscalar that produce the dark matter particle pair using the full Run2 data collected by CMS detector with centre-of-mass energy of $\sqrt{s} = 13\text{TeV}$.

CONTENTS

1	Introduction	1
2	Theory and Motivation	7
2.1	Standard Model	7
2.1.1	Fundamental Forces of Nature	9
2.1.2	Gauge of Standard Model	11
2.1.3	Limitations of Standard Model	14
2.2	Dark Matter search at LHC	24
3	Experimental Setup	33
3.1	Large Hadron Collider	36
3.1.1	LHC Parameters	38
3.2	THE CMS Detector	44
3.2.1	Inner Tracking System	48

CONTENTS

3.2.2	Electromagnetic Calorimeter	52
3.2.3	Hadronic Calorimeter	54
3.2.4	Solenoid Magnet	59
3.2.5	CMS Muon System	59
3.2.6	CMS Trigger System and DAQ	63
4	Data, Simulated samples and Trigger	79
4.1	Data and Simulated samples	79
4.1.1	Data	79
4.1.2	Background Simulation	81
4.2	Triggers	86
4.2.1	p_T^{miss} triggers	86
4.2.2	Single electron triggers	90
5	Simulation and Object Reconstruction	93
5.1	Event Reconstruction and Physics Object Selection	93
5.1.1	Reconstruction	93
5.1.2	Physics object selection	96
5.2	Monte Carlo Reweighting	109
5.2.1	Primary Vertex and Pileup Reweighting	109
5.2.2	L1 Prefiring in 2016 and 2017 data	110
6	Analysis Strategy and Event Selection	113
6.1	Event Selection	113

CONTENTS

6.1.1	Preselection	114
6.1.2	Signal region	115
6.1.3	Control regions	118
6.1.4	Signal Background Separation	124
7	Signal Extraction and Systematics Uncertainties	143
7.1	Signal extraction	143
7.2	Systematic uncertainties	157
7.2.1	Systematic Uncertainties on the background estimated using simulation	157
7.2.2	Systematic uncertainty on the Transfer Factor	162
8	Results	167
8.1	Results	167
8.1.1	Pulls and Impacts	168
8.1.2	Post-fit distributions	181
8.1.3	Limits	184
9	Summary and Outlook	187
9.1	Summary	187
9.2	Outlook	188
	Appendices	189
A	Plots for Detector Issues	191
A.1	Pixel DCDC failure	191

CONTENTS

A.2	EE Noise mitigation	191
A.3	L1 Prefiring in 2016 and 2017 data	196
B	Optimization	201
C	Plots of other variables for CRs	203
D	Systematic Uncertainty Plots	240
E	Contributions at CMS	339

LIST OF TABLES

2.1	Properties of Fundamental Forces	10
2.2	Phase space for which signal simulation datasets are produced. Same mass grid is used for all three years, 2016, 17 and 18.	29
2.3	Cross-section values for signal samples for bbDM analysis, the samples are produced using MadGraph5_aMC@NLO for event generation and PYTHIA8 for hadronization	31
3.1	The designed LHC parameters [1].	43
3.2	Input variable used for CSV and CSVv2 b-tagging algorithm. The symbol $\surd(x)$ implies that the variable used(not used)	71
4.1	Datasets used in the analysis for 2016 with luminosity.	80
4.2	Datasets used in the analysis for 2017 with luminosity	81
4.3	Datasets used in the analysis for 2018 with luminosity	81
4.4	Background Monte Carlo datasets used in the analysis for Run2	85
4.5	V+Jets Monte Carlo datasets(NLO) used in the analysis for 2016, 2017 and 2018	85
4.6	MET Trigger for all the three years	86
4.7	Single Electron trigger used in the analysis	90

LIST OF TABLES

5.1	Loose electron identification criteria for Electrons	100
5.2	Tight electron identification criteria for Electrons	100
5.3	Loose electron identification criteria for Photons	101
5.4	Loose Jet ID criteria for 2016	103
5.5	Tight Jet ID criteria for 2017 and 2018	103
5.6	MET Filter recommended by JetMET POG for Run II	107
6.1	Preselection applied to skim the events.	115
6.2	Selection applied on each of the two categories in the signal region, SR1 and SR2 (*MWP = Medium Working Point)	116
6.3	Event selection for 1b Control regions	120
6.4	Event selection for 2b Control regions	120
6.5	Background yield for SR (1b)	132
6.6	Background yield for SR (2b)	132
7.1	Contribution and fraction w.r.t total MC background of QCD background in each bin of $\cos \theta^*$ in signal region	156
7.2	QCD scales variations	160
7.3	Systematic uncertainties and their effect on the analysis. Uncertainties in red colour are uncorrelated throughout the three years	161
7.4	Source of experimental systematic uncertainties effecting the transfer factors (to mea- sure Z , W and $t\bar{t}$ background) along with their approximate size. Note that the W TF are used only for the 1b category and $t\bar{t}$ related TF are used only for the 2b category.	165

LIST OF FIGURES

1.1	The Particle discovery spectrum	3
1.2	Only about 5% of all the mass and energy in the universe is matter with which we are familiar here on Earth. Most ordinary matter consists of hydrogen and helium located in interstellar and intergalactic space. Only about one-half of 1% of the critical density of the universe is found in stars. Dark matter and dark energy, which have not yet been detected in earthbound laboratories, account for 95% of the contents of the universe. [2]	5
2.1	Standard Model Particle Physics	8
2.2	The interactions of particles in Standard Model. [3]	10
2.3	Color Confinement in QCD	13
2.4	The running coupling constant α_s of QCD [4]	14
2.5	Observed cross-section measurements for various physics processes as obtained by the CMS experiments at the LHC, with the theoretical predicted values. [4]	15
2.6	Observed cross-section measurements for various physics processes as obtained by the ATLAS experiments at the LHC, with the theoretical predicted values. [4]	16
2.7	The Higgs mass corrections from self interactions, gauge loops, and fermion loop	19

LIST OF FIGURES

2.8	Schematic showing the possible dark matter detection channels [5].	24
2.9	The Feynman diagram for the $bb+p_T^{\text{miss}}$ process within the 2HDM+a model	28
2.10	The left plots shows how the shape of generator p_T^{miss} variation with change in m_a . In the right plot, the variation of the generator p_T^{miss} w.r.t the mass of the dark matter particle	31
2.11	The comparison of cross-section values used for the signal samples for 2016, 2017 and 2018 datasets.	32
3.1	A diagram showing the tunnel of the LHC w.r.t French-Swiss border	34
3.2	CERN ACCELERATOR COMPLEX	34
3.3	Overview of HL-LHC Plan [6]	37
3.4	Computer generated diagram of an LHC dipole magnet [7].	39
3.5	Proton beams squeezed at the interaction point	41
3.6	Schematic of a hard scattering proton-proton collision [7].	42
3.7	Sectional view of the CMS detector [8].	45
3.8	Schematic of transverse slice of the CMS detector [9].	47
3.9	Schematic of the CMS tracker, each detector module is depicted by a line [10].	49
3.10	The geometrical layout of the pixel detector with η coverage [10].	50
3.11	The layout of the CMS ECAL [11].	53
3.12	The layout of the CMS HCAL [10].	56
3.13	Scintillator Trays.	57
3.14	Distribution of measured energy scaled to incident energy for pions at $\eta = 0$ at inci- dent energies of 10 GeV(left) and 200 GeV(right). The solid and dashed histograms are measurements without and with the HO.	58
3.15	Sketch of a cell showing drift lines and isochrones. The plates at the top and bottom of the cell are at ground potential. The voltages applied to the electrodes are +3600V for wires, +1800V for strips, and -1200V for cathodes [10]	61

3.16 Schematics of the CMS L1 trigger system. Data from the calorimeters are processed regionally (RCT) and then globally (GCT). Hits from the RPCs, CSCs, and DT are processed either via a pattern comparator or via a system of the segment- and track-finders and sent onwards to a global muon trigger (GMT). The information from the GCT and GMT is combined in a global trigger (GT), which makes the final trigger decision. This decision is sent to the tracker (TRK), ECAL, HCAL or muon systems (MU) via the trigger, timing and control (TTC) system. The data acquisition system (DAQ) reads data from various subsystems for offline storage [12].	64
3.17 Schematic of CMS trigger and DAQ system	66
3.18 A typical HLT path in CMS	67
3.19 Jet with a secondary vertex (SV) from the b or c hadron decay giving displaced tracks and large impact parameter.	69
3.20 Misidentification probability for c and light-flavour jets versus b jet identification efficiency for various b tagging algorithms in $t\bar{t}$ events [13].	73
3.21 Scatter plot of the output of DeepCSV algorithm offline(RECO)(y-axis) and online(HLT)(x-axis) for particle flow jets. The probability for b-tagged jet is probbb and probbb (plots in top row), similarly for c-tagged jets the probability is probcc and probcc (plots in middle row) and for light jets the probability is probudsg(plot in bottom row). The final discriminator value given for b-tag is probbb + probbb	75
3.22 Scatter plot of the output of DeepCSV algorithm offline(RECO)(y-axis) and online(HLT)(x-axis) for calorimetric jets. The probability for b-tagged jet is probbb and probbb (plots in top row), similarly for c-tagged jets the probability is probcc and probcc (plots in middle row) and for light jets the probability is probudsg(plot in bottom row). The final discriminator value given for b-tag is probbb + probbb	76
3.23 The timing study plot calculated for full HLT Menu without DeepCSV tagger (left) and with the DeepCSV tagger(right). The inclusion of HLT380_SixPFJet32_DoublePFJetTagDeepCSV_2p2 path only increase the total time taken by the full HLT menu by 0.53ms.	77

LIST OF FIGURES

4.1	The pie chart for SR1 (left) and SR2 (right) to show the % contribution of each background.	84
4.2	p_T^{miss} trigger efficiency scale factors for 2016 (left), 2017 (right), and 2018 (bottom) . . .	88
4.3	Trigger turnON for p_T^{miss} trigger for 2016 (top), 2017(middle) and 2018 (bottom). The distribution for data is shown in the left and simulation is shown on the right.	89
4.4	Single electron trigger efficiency scale factors for 2016 (top left), 2017 (top right), and 2018 (bottom)	91
5.1	Before and after PU reweighting plots in 2b category of $t\bar{t}(e\nu)$ for 2018	110
6.1	Figure showing the categories defined in the signal regions. The categories depend on the total number of jets and the number of b-tagged jets in an event.	117
6.2	Figure showing the the control regions and the position of the signal regions in the electron-muon space. The categories depend on the number and flavour of leptons, and also the number of jets in an event.	119
6.3	The signal to background separation of top six variables that were fed to the BDT framework. Out of 25 variables, $\cos(\theta)^*$ gives highest signal to background separation.	127
6.4	The bdt discriminator for 2b category of showing the separation of signal and background.	128
6.5	p_T^{miss} plot in 1b Category of Signal region for 2016, 2017 and 2018	129
6.6	p_T^{miss} and $\cos(\theta)^*$ plot in 2b Category of Signal region for 2016, 2017 and 2018. The signal in both set of the plots is scaled by 5. $\cos(\theta)^*$ provides better signal to background separation as compared to p_T^{miss}	130
6.7	Jet1 p_T plot in 1b and 2b Category of Signal region for 2016, 2017 and 2018	131
6.8	Recoil plot in 1b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	133
6.9	Jet1 p_T plot in 1b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	134

LIST OF FIGURES

6.10	$\cos(\theta)^*$ plot in 2b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	135
6.11	Jet1 p_T plot in 2b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018.	136
6.12	Recoil plot in 1b Category of $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	137
6.13	Jet1Pt plot in 1b Category of $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	138
6.14	$\cos(\theta)^*$ plot in 2b Category of Top(μ) and Top(e) control region for 2016, 2017 and 2018.	139
6.15	Jet1 p_T plot in 2b Category of Top(μ) and Top(e) control region for 2016, 2017 and 2018.	140
6.16	$\cos\theta^*$ plot in the QCD control region for 2016, 2017 and 2018	141
7.1	Distribution of transfer factors $W_{SR}/W_{\mu\nu}$ and $W_{SR}/W_{e\nu}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.	145
7.2	Distribution of transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.	146
7.3	Distribution of transfer factors $t\bar{t}_{SR}/t\bar{t}_{Top\mu CR}$ and $t\bar{t}_{SR}/t\bar{t}_{Top e CR}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.	147
7.4	Distribution of transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.	148
7.5	Uncertainties on the transfer factors $W_{SR}/W_{\mu\nu}$ and $W_{SR}/W_{e\nu}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018.	149
7.6	Uncertainties on the transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018.	150

LIST OF FIGURES

7.7	Uncertainties on the transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018.	151
7.8	Uncertainties on the transfer factors $t\bar{t}_{SR}/t\bar{t}_{Top\mu CR}$ and $t\bar{t}_{SR}/t\bar{t}_{Top e CR}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018.	152
7.9	Extrapolating QCD background from the QCD control region to signal region using an exponential function. The contribution after the $\Delta\phi > 0.5$ is very small compared to the total MC background in the signal region	156
8.1	Distribution of the pulls with data. Combination of 1b and 2b for full 2016 dataset is used for the fit.	169
8.2	Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for 2016 dataset. Combination of 1b and 2b is used for the fit.	170
8.3	Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2016.	171
8.4	Distribution of the pulls with data. Combination of 1b and 2b for full 2017 dataset is used for the fit.	172
8.5	Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for 2017 dataset. Combination of 1b and 2b is used for the fit.	173
8.6	Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2017.	174
8.7	Distribution of the pulls with data. Combination of 1b and 2b for full 2018 dataset is used for the fit.	175
8.8	Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for 2018 dataset. Combination of 1b and 2b is used for the fit.	176
8.9	Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2018.	177
8.10	Distribution of the pulls with data. Combination of 1b and 2b for full run2 dataset is used for the fit.	178

LIST OF FIGURES

8.11	Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for full run2 dataset. Combination of 1b and 2b is used for the fit.	179
8.12	Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for run2.	180
8.13	Summarizing MET for signal and Control region for 1b category for 2016, 2017 and 2018	182
8.14	Summarizing $\cos(\theta)^*$ for signal and Control region for 2b category for 2016, 2017 and 2018	183
8.15	95% CLs upper limit on the signal strength (ratio of 95% CLs cross-section of dark matter produced in association with pair of bottom quarks to the theoretical cross-section). The upper limits are shown as a function of m_a for $m_A = 600$ GeV, $\sin\theta=0.7$ and $\tan\beta=35$ for combination of 1b and 2b category for the full run-2 dataset.	185
1	Comparison of nBjet distribution for pre and post pixel scenario to check the effect on the region of interest for the analysis	191
4	E_T^{miss} Comparison plots in SR1 for all the eras for 2017 data taking, the era wise plots are all normalize to unity to check shape of the distribution with and without METv2 recipe applied	192
5	E_T^{miss} Comparison plots in SR1 for all the eras for 2017 data taking, the era wise plots are all without normalization to check difference in statistics of the E_T^{miss} distribution with and without METv2 recipe applied	193
6	E_T^{miss} Comparison plots in SR2 for all the eras for 2017 data taking, the era wise plots are all normalize to unity to check shape of the distribution with and without METv2 recipe applied	194
7	E_T^{miss} Comparison plots in SR2 for all the eras for 2017 data taking, the era wise plots are all without normalization to check difference in statistics of the E_T^{miss} distribution with and without METv2 recipe applied	195

LIST OF FIGURES

8	Comparison of E_T^{miss} distribution of total background in 1b and SR2, with normalization(in upper row) and without normalization(lower row) to check the effect of l1 prefire in 2017 MC	197
9	Recoil distribution of total background in 2b category of SR(including signal mass points) and muon CRs to check the effect of l1 prefire in 2017 MC	198
10	Recoil distribution of total background in electron CRs to check the effect of l1 prefire in 2017 MC	199
11	The left plot, Contribution of individual background in SR2. In the right plot the percentage change in the yield of individual background with and without prefiring. .	200
12	Comparison of upper limit plot for number of jets and jet p_T selection with CR only fit for 2017	202
13	CutFlow plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	204
14	CutFlow plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	205
15	CutFlow plot in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	206
16	CutFlow plot in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	207
17	E_T^{miss} plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	208
18	E_T^{miss} plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	209
19	E_T^{miss} plot in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	210
20	E_T^{miss} plot in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	211
21	Jet1 η in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	212
22	Jet1 η in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	213
23	Jet1 η in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	214

LIST OF FIGURES

24	Jet1 η in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	215
25	Jet1 ϕ in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	216
26	Jet1 ϕ in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	217
27	Jet1 ϕ in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	218
28	Jet1 ϕ in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	219
29	DeepCSV discriminator for Jet1 in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	220
30	DeepCSV discriminator for Jet1 in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	221
31	DeepCSV discriminator for Jet1 in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	222
32	DeepCSV discriminator for Jet1 in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	223
33	$\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	224
34	$\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	225
35	$\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	226
36	$\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	227
37	Z boson p_T in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	228
38	Z boson p_T in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	229
39	W boson p_T in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	230
40	W boson p_T in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	231
41	Z boson mass in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	232

LIST OF FIGURES

42	Z boson mass in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	233
43	W boson mass in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	234
44	W boson mass in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	235
45	Number of jets in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.	236
46	Number of jets in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.	237
47	Number of jets in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.	238
48	Number of jets in Top(μ) and Top(e) control region for 2016, 2017 and 2018.	239
49	PileUP uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	241
50	b-tag uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	242
51	fake b-tag uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	243
52	E_T^{miss} trigger uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	244
53	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	245
54	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	246
55	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	247
56	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	248
57	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	249

LIST OF FIGURES

58	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	250
59	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	251
60	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	252
61	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	253
62	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	254
63	JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018	255
64	PileUP uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	256
65	b-tag uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	257
66	fake b-tag uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	258
67	E_T^{miss} trigger uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	259
68	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	260
69	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	261
70	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	262
71	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	263

LIST OF FIGURES

72	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	264
73	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	265
74	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	266
75	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	267
76	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	268
77	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	269
78	JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018	270
79	PileUP uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	271
80	b-tag uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	272
81	Fake b-tag uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	273
82	Trigger uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	274
83	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	275
84	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	276
85	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	277

LIST OF FIGURES

86	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	278
87	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	279
88	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	280
89	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	281
90	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	282
91	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	283
92	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	284
93	JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018	285
94	muon ID and electron ID uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ respectively for 2016, 2017 and 2018	286
95	muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ respectively for 2016, 2017 and 2018	287
96	PileUP uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018	288
97	b-tag uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018	289
98	Fake b-tag uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018	290
99	Trigger uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018	291

LIST OF FIGURES

100	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	292
101	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	293
102	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	294
103	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	295
104	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	296
105	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	297
106	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	298
107	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	299
108	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	300
109	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	301
110	JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ for 2016, 2017 and 2018	302
111	muon ID and electron ID uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ respectively for 2016, 2017 and 2018	303
112	muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(ev)$ respectively for 2016, 2017 and 2018	304
113	PileUP uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	305

LIST OF FIGURES

114	b-tag uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	306
115	Fake b-tag uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	307
116	Trigger uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	308
117	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	309
118	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	310
119	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	311
120	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	312
121	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	313
122	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	314
123	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	315
124	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	316
125	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	317
126	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	318
127	JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	319

LIST OF FIGURES

128	muon ID and electron ID uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018	320
129	muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018	321
130	PileUP uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	322
131	b-tag uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	323
132	Fake b-tag uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	324
133	Trigger uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	325
134	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	326
135	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	327
136	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	328
137	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	329
138	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	330
139	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	331
140	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	332
141	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	333

LIST OF FIGURES

142	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	334
143	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	335
144	JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018	336
145	muon ID and electron ID uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018	337
146	muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018	338

CHAPTER 1

INTRODUCTION

From the vast structures as big as the universe to the tiniest amount of matter possible, the curiosity of knowing how the universe works has been consistent throughout the human history. That curiosity alone have led so far in the way of understanding the nature. From the atom to nucleus and then to the quarks, the goal of the particle physics has always been to unlock the mysteries of the world around us.

The story of curiosity is not new, it dates back to as early as the 5th century BCE, when the ancient Greek philosophers Democritus and Leucippus recorded the concept of the *atomos*, an indivisible building block of matter. This concept of undivided atom remained untouched for very long until the search for the smallest particle led J. J. Thomson to the discover one of first elementary particle “electron” [14] in 1897, in an experiment where THompson observed the bending of cathode rays

CHAPTER 1. INTRODUCTION

in the presence of a magnetic field. With the inclusion of electron, the new model of atom was developed called Plum Pudding Model [15]. This was soon followed by the discovery of nucleus by Ernest Rutherford in 1911 [16]. The Rutherford model of atom described the atom as a small, dense, positively charged core called a nucleus, in which nearly all the mass is concentrated, around which the negatively charge particles called electrons, circulate at some distance [17]. After the nucleus, the discovery of the proton by Rutherford in 1919 and then the discovery of neutron by J. Chadwick in 1932 [18] gave a new model of atom with proton, neutron in the nucleus and electron orbiting around the nucleus and proton, neutron and electrons are called the fundamental particles. In 1964, “Quark Model” was proposed by Gell-Mann and Zweig which stated that “quarks” are the fundamental particle and all the baryons and mesons consist quarks. The baryons like proton and neutron consists of three quarks and meson like pions are made up of quark-antiquark ($q\bar{q}$) pair. The first key evidence for the existence of quarks came from a series of inelastic electron-nucleon scattering experiments conducted between 1967 and 1973 at the Stanford Linear Accelerator Center [19]. The discovery of $W^\pm$, Z^0 boson followed soon after and the finally in 2012, the Higgs boson was discovered in Large Hadron Collider (LHC) at CERN [20]. All the particles hypothesized by SM have been found as shown in Figure 1.1

All these particle are put together in a an extensive theoretical framework known as the Standard Model of particle physics. It is the latest and greatest model for the fundamental particle developed over a century of research and which successfully describes most of the interactions of the elementary particles. SM of particle physics

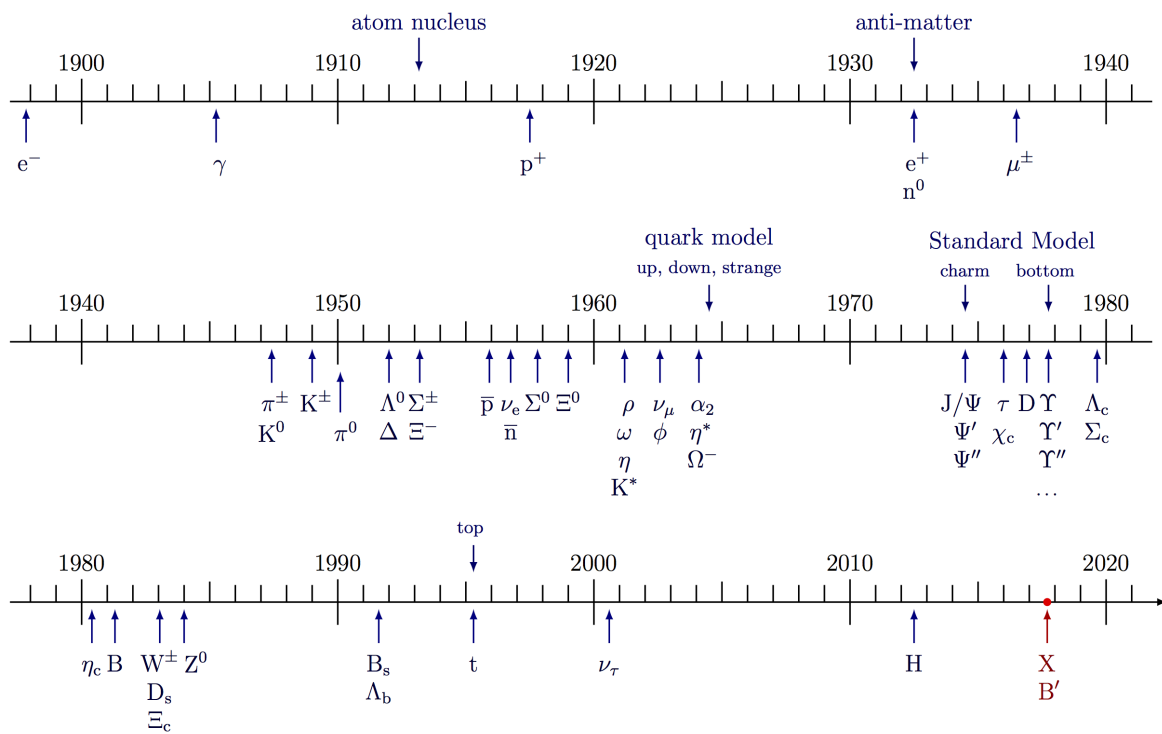


Figure 1.1: The Particle discovery spectrum

CHAPTER 1. INTRODUCTION

has been backed by numerous experiments over the time, the agreement between the prediction of SM and the experimental observation is quite remarkable [21]. For example, the existence of the top quark, Z, W, and Higgs bosons were predicted by SM long before these particles were experimentally discovered. In-fact, the SM has been the guiding theory for looking out for these particles and design the experiments to reach the energy required for the production and detection of these particles. Furthermore, various properties like the Higgs boson cross section and branching fractions have been very accurately predicted by the SM. As the famous mass-energy equivalence states that the mass can be converted into energy and visa-versa. Therefore to produce particles with higher mass we need very high energy in very small space, one way of doing this is to accelerate the particle to very high speeds and the collide them. This has been the basic idea behind the particle accelerators across the world and is an extremely successful method of probing the particles predicted by SM. It has provided great deal of information over the last few decades about fundamental particles. One of facility where this idea has been realized as one of the largest particle collider(Large Hadron Collider or LHC) is CERN. In 2012 at the LHC, Higgs boson was discovered by both the ATLAS [22] and CMS Collaborations [23] independently. Higgs boson being the final piece of the Standard Model of physics, has opened possibilities for very exciting new physics.

The SM will be discussed in upcoming chapter, however with the current understanding of the universe we know that only about 5% of the universe is visible matter, 27% accounts for the Dark matter and the remaining 68% constitutes Dark Energy. [2] So, the Standard Model has been a great theory in explaining the inter-

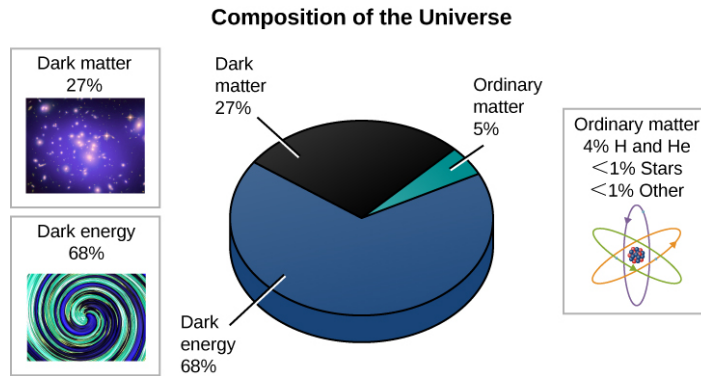


Figure 1.2: Only about 5% of all the mass and energy in the universe is matter with which we are familiar here on Earth. Most ordinary matter consists of hydrogen and helium located in interstellar and intergalactic space. Only about one-half of 1% of the critical density of the universe is found in stars. Dark matter and dark energy, which have not yet been detected in earthbound laboratories, account for 95% of the contents of the universe. [2]

action between elementary particles but it is still not a complete one, for example there is no place for gravity in it yet, the SM does not explain the existence of dark matter and dark energy and why there is an asymmetry in matter and antimatter in the universe. The Dark matter constituting 27% for the universe, backed by strong astrophysical evidence like the rotational curve [24], the bullet cluster [25] and by gravitational lensing [26]. However, the particle nature of DM is not covered by the reach of SM and remains completely unknown.

In the light of the recent developments and very direct evidence of dark matter, various beyond the SM theories prediction states that DM particles interact with SM particles although very weakly. There are three ways that broadly cover the search of DM particles, the first method is direct search where DM particle collides with SM particle and scatters off, the second method is where the DM particle decays

CHAPTER 1. INTRODUCTION

to SM particle and the third method is that where SM particles are collided in a particle collider like the LHC and then the DM particle will be produced and will recoil against the visible matter, since the DM particle does not interact with normal matter, it will go unnoticed in the detector. These events where there is a large imbalance in the transverse momentum become the signature events for the production of DM particle. In this thesis, one such signature where the DM produced in association with the bottom quarks has been analysed.

In this thesis, the search will be analysed in detail, starting from the theoretical motivation to the details of the CMS detector, the reconstruction of various physics objects, the analysis strategy, the event selection, background modeling and the treatment of the systematic and statistical uncertainties. Then the results will be interpreted with 2HDM+a model, which will be explored for the first time in the context of this search.

CHAPTER 2

THEORY AND MOTIVATION

2.1 Standard Model

A ‘Standard Model’ is a theoretical setup that predicts and correlates new data built from observation of nature. It describes all the elementary particles and their interaction with strong, weak and electromagnetic forces. The particles in the SM, with spin $-\frac{1}{2}$ are known as fermions which interact with the fundamental forces mediated by gauge bosons with spin-1. There are 12 fermions in the SM, six quarks that interact through the strong, weak and electromagnetic forces, and six leptons that interact through weak, electromagnetic, but not the strong force. Each of the fermions has an equivalent particle with the opposite electric charge as the fermion, and it is called an antiparticle. Both quarks and Leptons are further subdivided into three “generations” of particles, with heavier mass otherwise identical. There are two types

CHAPTER 2. THEORY AND MOTIVATION

Standard Model of Elementary Particles						interactions / force carriers (elementary bosons)	
three generations of matter (elementary fermions)			three generations of antimatter (elementary antifermions)				
I	II	III	I	II	III		
mass charge spin							
$\approx 2.2 \text{ MeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 1.28 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 173.1 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 2.2 \text{ MeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	$\approx 1.28 \text{ GeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	$\approx 173.1 \text{ GeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	0 0 1	$\approx 124.97 \text{ GeV}/c^2$ 0 0
u up	c charm	t top	$\bar{u}$ antiup	$\bar{c}$ anticharm	$\bar{t}$ antitop	g gluon	H higgs
d down	s strange	b bottom	$\bar{d}$ antidown	$\bar{s}$ antistrange	$\bar{b}$ antibottom	0 0 1	
e electron	μ muon	τ tau	e^+ positron	$\bar{\mu}$ antimuon	$\bar{\tau}$ antitau	0 0 1	
ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	$\bar{\nu}_e$ electron antineutrino	$\bar{\nu}_\mu$ muon antineutrino	$\bar{\nu}_\tau$ tau antineutrino	$\approx 91.19 \text{ GeV}/c^2$ 0 1	
$< 2.2 \text{ eV}/c^2$ 0 $\frac{1}{2}$	$< 0.17 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 18.2 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 2.2 \text{ eV}/c^2$ 0 $\frac{1}{2}$	$< 0.17 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 18.2 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	1 1 1	1 -1 1
W^+ W ⁺ boson	W^- W ⁻ boson						

Figure 2.1: Standard Model Particle Physics

of fermions in the SM, quarks and leptons. Figure 2.1 lists the mass, charge and spins of the leptons and quarks for each of the three generations. The leptons are electron(e), muon(μ) and tau(τ) with corresponding neutrinos electron neutrino(ν_e), muon neutrino(ν_μ) and tau neutrino(ν_τ) and corresponding antiparticles. The leptons have an electric charge of ± 1 , so they interact electromagnetically via virtual photons. The quarks, on the other hand, have two types of charge, electric and colour charge. So they can interact with gluons for strong interaction and $W^\pm/Z$ bosons for weak and photons for electromagnetic. The “down-type” quarks have an electric charge of $-\frac{1}{3}$, and the “up-type” have an electric charge of $\frac{2}{3}$ with colour

charges(or anti-colour charges) of red, blue or green. As a feature of QCD(Quantum chromodynamics) called confinement, only bound states of quarks with no overall colour charge are allowed, hence in nature, quarks are always found in composite states with no overall colour charge, and these composite particles are called Hadrons which include proton, neutron, pions etc. The SM constitutes five gauge bosons, photons(γ), $W^\pm$ boson, Z boson, gluons and the Higgs Boson, 4 out of these gauge bosons are responsible for three of the fundamental force of nature.

2.1.1 Fundamental Forces of Nature

With the current understanding of the universe, there are four kinds of forces in nature, namely Strong forces, Weak force, Electromagnetic forces and Gravitational forces. Out of these four, only three of them Strong, Weak and Electromagnetic forces are explained by the SM. Only Gravitational force is explained by the General Theory of Relativity; however, the particle equivalent of the gravitational force is yet not established. The interaction between the elementary particles and the force carriers in the SM is shown in Figure 2.2 with the basic properties of the forces, and their particle equivalent is given in the Table 2.1.

CHAPTER 2. THEORY AND MOTIVATION

Interaction	Force Carrier	Charge	Mass (GeV/c ²)	Range (m)	Coupling Constant (α)	Theory
Strong	gluons	0	0	10^{-15}	1	Quantum Chromodynamics
Weak	$W^{\pm}$	± 1	80.38 ± 0.012	10^{-18}	10^{-9}	Quantum
	Z	0	91.188 ± 0.002			Flavourdynamics
Electromagnetic	γ	0	0	∞	1/137	Quantum Electrodynamics
Gravitational	graviton	0	0	∞	10^{-38}	General Theory of Relativity

Table 2.1: Properties of Fundamental Forces

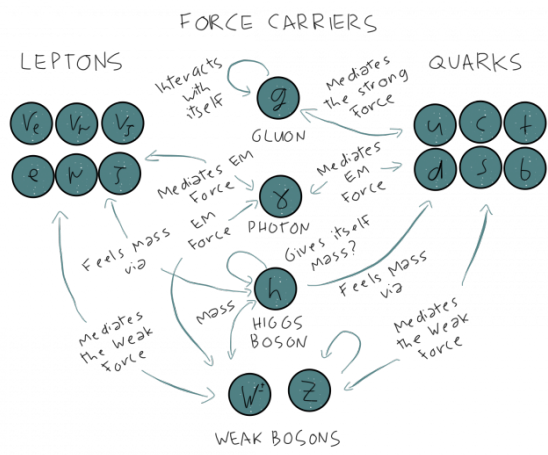


Figure 2.2: The interactions of particles in Standard Model. [3]

2.1.2 Gauge of Standard Model

The mathematical formulation of the Standard Model of particle physics takes us to the Lagrangian densities that are invariant under a global gauge transformation, and while locally, it varies with the transformation at each point. It is described with gauge group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$$

The mixing of the colour of the quarks and antiquarks is by $SU(3)_C$, the $SU(2)_L$ mixes the weak isospin, and the $U(1)_Y$ couples to the weak hypercharge Y . describe the transformation of particle states

The general form of lagrangian density for the standard model is given as

$$\mathcal{L}_{SM} = \mathcal{L}_{Dirac} + \mathcal{L}_{gauge} + \mathcal{L}_{\phi} + \mathcal{L}_{Yukawa} \quad (2.1)$$

where $\mathcal{L}_{Dirac}$ is for the fermions(spin- $\frac{1}{2}$), $\mathcal{L}_{gauge}$ refers to spin-1 mediators, the gauge bosons, $\mathcal{L}_{\phi}$ is for spin-0 scalar field(the Higgs doublet) and $\mathcal{L}_{Yukawa}$ is for the interaction between higgs and the fermions.

Quantum ElectroDynamics is an abelian gauge theory with the symmetry group $U(1)$. The gauge field for the QED is the electromagnetic field which mediates the interaction between the charged spin- $\frac{1}{2}$ fields. The QED Lagrangian for a spin- $\frac{1}{2}$ fermionic fields interacting with the electromagnetic field is given as

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.2)$$

CHAPTER 2. THEORY AND MOTIVATION

where, γ^μ are the Dirac matrices and $F_{\mu\nu}$ is the electromagnetic field tensor.

The symmetry group $U(1)$, leaves the Lagrangian invariant under global symmetry and hence according to Noether's theorem leads to the conservation of electric charge. To make the lagrangian invariant locally, a new gauge field “photon field” (massless with spin-1) is added as shown in the the equation 2.2.

In a similar way, weak interactions are characterized by $SU(2)$ gauge group as they need to have several fermionic flavours, different properties for left- and right-handed fields and massive gauge bosons $W^\pm$ and Z . When the electromagnetic interactions are also included to have a unified electroweak theory, we need an additional $U(1)$ group, then the symmetry group considered is $SU(2)_L \otimes U(1)_Y$, where L refers to left-handed fields [27]. This gives rise to three gauge fields, two for charged current interaction (W^+ , W^-) and one for neutral current interaction (Z^0).

Quantum chromodynamics, or QCD, is the quantum field theory that describes the strong interactions between quarks and gluons and represented by a non-abelian gauge symmetry group $SU(3)_C$. The local gauge invariance adds 8 gauge fields (gluon fields) that are the force carriers between colored quarks. The QCD lagrangian controls the dynamics of the quarks and gluons. The gauge invariant lagrangian is

$$\mathcal{L}_{QCD} = \bar{\psi}_i (i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} \quad (2.3)$$

where $\psi_i(x)$, is the quark field, in the fundamental representation of the $SU(3)$ gauge group. The QCD coupling constant is given as $g_s = \sqrt{4\pi\alpha_s}$, where α_s is the strength of the strong interaction, and it depends on the energy scales of the interactions.

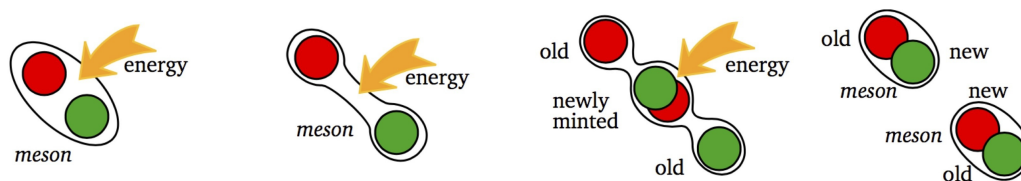


Figure 2.3: Color Confinement in QCD

This dependence of α_s to the energy of the interaction gives rise to the very unique and interesting properties of the strong interaction.

- **Color confinement:** Confinement is the phenomenon that quarks and gluons cannot be isolated, and therefore cannot be directly observed in normal conditions, i.e., in low energies where the coupling constant where α_s is large. If two quarks in a bound state are pulled apart, at a certain point, the colour force becomes so strong that it creates new quark-antiquark pair to conserve the energy and other quantum numbers as the energy of the color force field is converted into a mass of new quarks. This is the reason no single independent quark is found in nature. The confinement is better shown in Figure 2.3
- **Asymptotic Freedom:** It is a property of non-abelian gauge theories like QCD that makes interactions between particles to become weaker, asymptotically, when the length scale of interaction decreases and the energy scale increases. Unlike QED, the coupling constant in QCD depends on the scale of momentum transfer(q) and it is given as

$$\alpha_s(q^2) = \frac{12\pi}{\left[(11n - 2f) \ln \left(\frac{q^2}{\Lambda^2} \right) \right]} \quad (2.4)$$

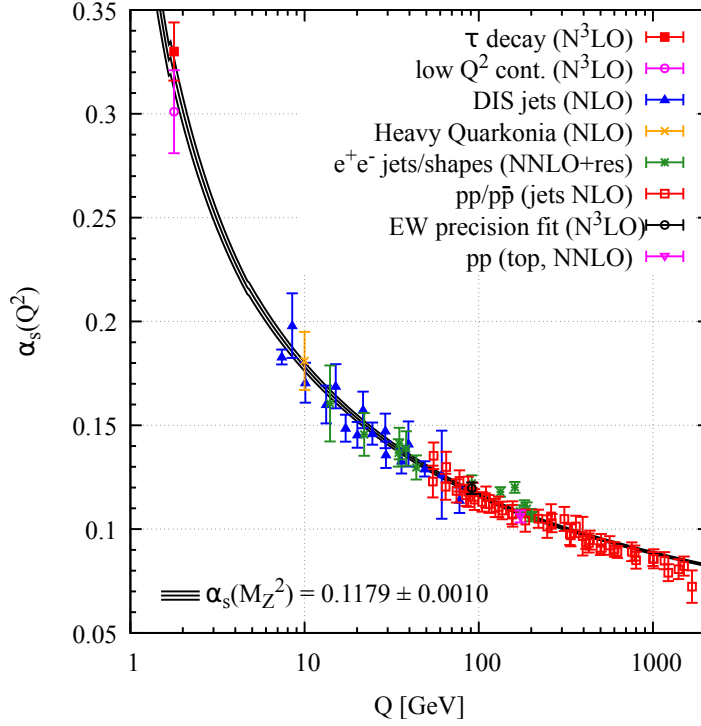


Figure 2.4: The running coupling constant α_s of QCD [4]

So, $\alpha_s \rightarrow 0$ as $q^2 \rightarrow \infty$, implying that at very high energies, the interaction asymptotically vanishes to zero as shown in Figure 2.4

2.1.3 Limitations of Standard Model

Standard Model has been a great theory so far as explained in previous sections. However, there are some questions which the Standard Model cannot explain, for example, according to SM, at the start of the universe (the big bang) matter and antimatter would have been equally produced and then annihilate as the universe cools. Yet, we do see that universe is composed of matter rather than antimatter. The gravity



[4]

dark energy.

2.5 and Figure 2.6.

new theory that describes the processes beyond the SM is required. An explanation

CHAPTER 2. THEORY AND MOTIVATION

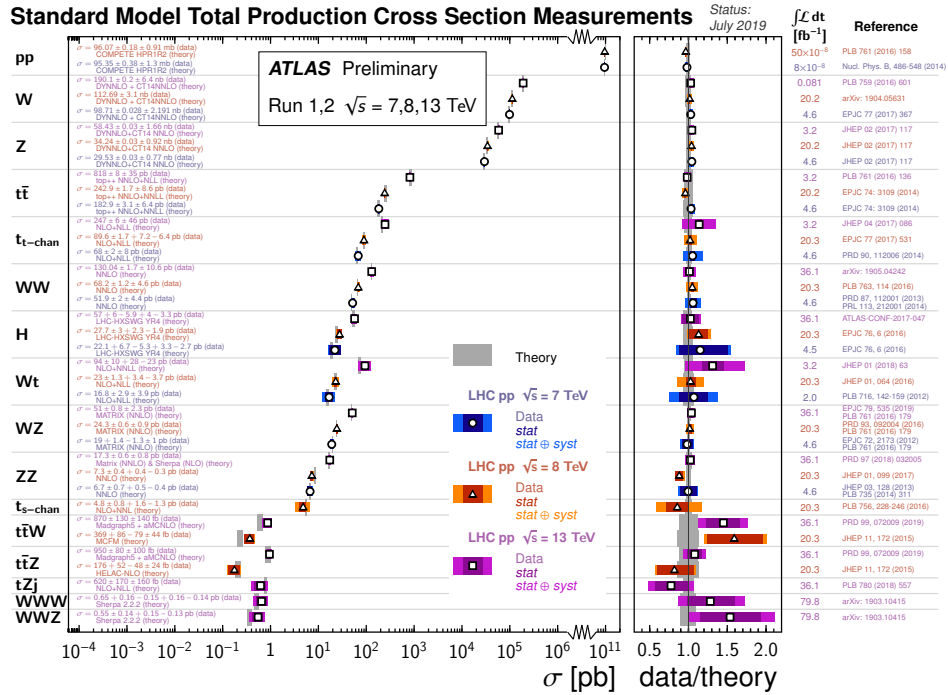


Figure 2.6: Observed cross-section measurements for various physics processes as obtained by the ATLAS experiments at the LHC, with the theoretical predicted values. [4]

2.1. STANDARD MODEL

of some of the drawbacks of SM is given in the following section.

CHAPTER 2. THEORY AND MOTIVATION

2.1.3.1 Matter-AntiMatter Asymmetry

Ideally, after the Big Bang, there should have been equal amounts of matter and anti-matter in the early universe. However, as of today, all that we see, from the smallest of the things to the largest of galaxies objects, is made almost entirely of matter. The antiparticles have the same mass as their particle counterparts; the electric charge is opposite; for example, the positron is the antiparticle of an electron. The theoretical framework for antimatter was given by Paul Dirac in 1928 and discovered in the form of anti-electrons(positrons e^+) by Carl Anderson in 1932. Andrei Sakharov in 1967, listed three necessary conditions for this asymmetry [30].

- Baryon number violation have to occur. Interactions violating baryon number would be necessary to have a state of thin universe as it is today.
- Charge Parity(CP)-violation, if the product of charge conjugation and parity is not violated, then the rate of particle production will be equal to that of antiparticle production.
- Departure from Thermal Equilibrium, if the early universe was in thermal equilibrium, then even Baryon number and CP-violating interactions could not produce a net asymmetry.

2.1.3.2 Neutrino Oscillations

Neutrinos, as according to the Standard Model, are chargeless and massless particles. However, in 1988, it was discovered by the Super-Kamiokande experiment

that the atmospheric muon neutrinos generated by the collision between the atmosphere on the earth and cosmic rays and were observed, in which the number of the upward going neutrinos was only half of the number of the downgoing neutrinos because the muon neutrinos were passing through the earth turn into tau neutrinos [31]. This phenomenon is called neutrino oscillation, and it is an indicator that the neutrinos are not massless particles.

2.1.3.3 The Hierarchy Problem

The Hierarchy Problem refers to the large difference in the scales between the mass of the electroweak gauge bosons/Higgs and the Planck mass ($M_{W,Z,H} \sim 100\text{GeV}$ and $M_{\text{Planck}} \sim 10^{19}\text{GeV}$) of the order of $\sim 10^{17}\text{GeV}$. As in QFT, scalar particles get large masses through large radiative corrections, and hence in the calculation of the Higgs mass, it receives contributions from huge quantum correction from self-interactions, gauge loops, and fermion loops as shown in Figure 2.7.

The loops shown in Figure 2.7 are quadratically divergent. Explicitly,

$$\delta m_H^2 = \frac{\Lambda^2}{32\pi^2} \left[6\lambda + \frac{1}{4}(9g^2 + 3g'^2) - y_t^2 \right] \quad (2.5)$$

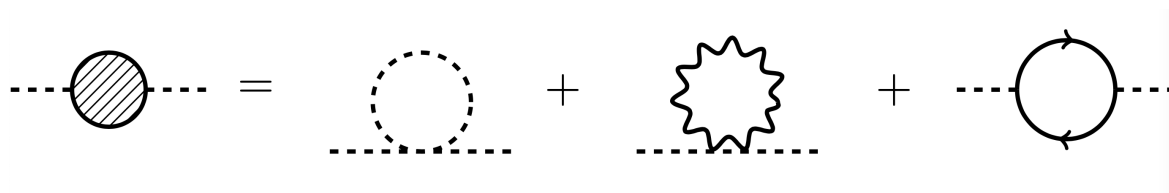


Figure 2.7: The Higgs mass corrections from self interactions, gauge loops, and fermion loop

CHAPTER 2. THEORY AND MOTIVATION

where Λ is the cutoff scale, so if $\Lambda \sim M_{\text{Planck}}$, then the quantum correction to the Higgs mass is much larger than the Higgs mass itself, $\delta m_H^2 \gg m_H^2$. Hence to balance such corrections, the standard model requires a consistent extension that takes care of the unnaturally large values for cancelling out the corrections.

2.1.3.4 Dark Matter and Dark Energy

As mentioned in the chapter 1, dark matter (DM) constitute about 27%, and dark energy(responsible for the accelerating expansion of the universe [32]) constitute 68% of the total mass-energy of the known universe. The Standard Model only explains the visible matter, but since the dark matter does not interact with photons(hence the name “Dark Matter”), there is no place for dark matter in the Standard Model so far. The evidence of the existence of the dark matter stands on very solid ground; few of the pieces of evidence are

- Galaxy rotation curves.
- Bullet Cluster.
- Velocity dispersions.
- Gravitational lensing.
- Structure formation.

However, one thing to note is that all the pieces of evidence for the existence of DM are gravitational in nature, and there has been no direct evidence. Neither experimentally nor by prediction of SM, we have seen a dark matter particle so far.

But in many beyond standard model theories, there are various plausible candidates for dark matter, and hence to have a dark matter candidate with all the constraints from various experiments, Marco Taoso, Gianfranco Bertone, and Antonio Masiero have made a 10 point test [33].

1. Does it match the appropriate relic density? In the early universe, the non-relativistic dark matter was in thermodynamic equilibrium with the plasma, and when the rate of interaction dropped below the rate of expansion of the universe, it froze out. So, the relic density today should be within the bounds of the limits set by CMB measurements.
2. Is it cold? The dark matter is required to be cold for the standard theory of structure formation to be consistent. Only a small part of the dark matter content of the universe can be allowed to be hot dark matter, provided that it is compatible with CMB data and structure formation
3. Is it neutral? Almost all of the dark matter candidates have to be electrically neutral as it does not interact electromagnetically.
4. Is it consistent with Big Bang Nucleosynthesis(BBN)? The BBN predicts the abundances of light elements produced in the very early universe; it is based on the set of coupled Boltzmann equations, which relate number densities of neutrons, protons, and light elements. The agreement between predictions of BBN (which is described by SM) and measurements [34] should be respected by the dark matter candidate.

CHAPTER 2. THEORY AND MOTIVATION

5. Does it leave stellar evolution unchanged? Since the standard model prediction about the total number of neutrinos from SN 1987A, their energy and their time distribution is in agreement with observed values, implying that the dark matter candidates should not change the stellar evolution above the constraints
6. Is it compatible with constraints on self-interactions? There are various observations that put constraints on the self-interaction cross-section of dark matter particles.
7. Is it consistent with direct DM searches? In the direct searches of DM, the DM particles are detected through the measurement of nuclear recoils produced by DM scattering. The DM candidate has to follow the limits on the mass and nature of DM particle that are put by Direct Detection Experiments
8. Is it compatible with gamma-ray constraints? Dark matter can be searched through the detection of products that are the result of the annihilation of the Dark matter particle. This is called Indirect search. When the DM candidate is in the mass range of GeV to TeV, then the result of this annihilation can give γ -rays. The result of this type of search can put several constraints on dark matter candidates.
9. Is it compatible with other astrophysical bounds? Apart from photons, neutrinos, positrons and antiprotons in cosmic-ray fluxes are also produced in the annihilation of DM. These put further constraints on the Dark matter candidates

10. Can it be probed experimentally? This is not a necessary requirement for a dark matter candidate, but this is the most important point from the viewpoint of the experiments.

One of the few candidates for dark matter is called Weakly interacting massive particles (WIMPs). The WIMPs are stable particle which arises in various extensions of Standard Model. The WIMPs typically range from a few GeV to a few TeV in the mass range, and they interact with normal matter in ways that are characteristics of weak interactions [35].

2.1.3.4.1 Experimental searches for dark matter particle For the detection of the nature of Dark Matter, there are three methods, and almost all the experiment relate to the searches of dark matter particle can be categorized with these three methods as shown in Figure 2.8.

- **Direct Detection:** To observe low-energy recoils (typically a few KeVs) of nuclei induced by interactions with particles of dark matter, which (in theory) pass through the Earth
- **Indirect detection:** Dark matter can annihilate into photons, neutrinos etc., so in the indirect detection, the search is for the products of the self-annihilation or decay of dark matter particles in outer space.
- **Collider searches:** A method complementary to direct and indirect detection of dark matter to produce DM in a laboratory, i.e. in the particle colliders. For example, at CMS, physicists look for generic DM signature of non-interacting

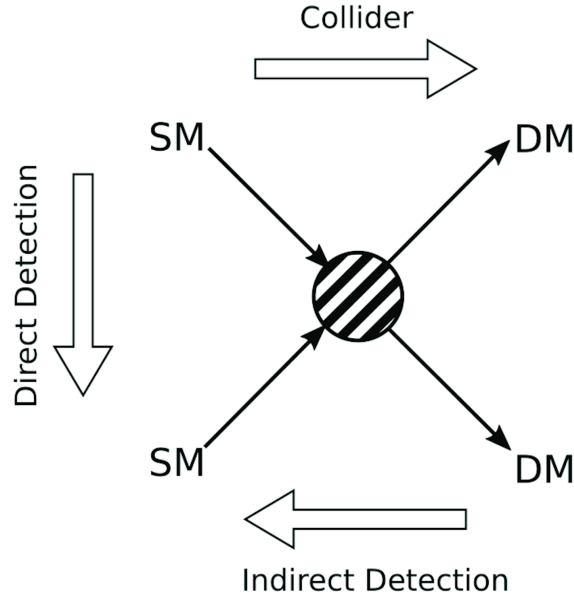


Figure 2.8: Schematic showing the possible dark matter detection channels [5].

particles, since dark Matter particles produced at the Large Hadron Collider(LHC) would escape the detector unnoticed must recoil against the SM particles to become “visible”. Therefore, general topologies to search for the production of Dark Matter at the LHC are ‘mono- X’ signatures (a visible object recoiling against missing energy), where X can be photons, W/Z and jets

2.2 Dark Matter search at LHC

The dark matter (DM) constitute about 26% [36] of the total mass-energy of the known universe. All the evidence for the existence of DM are gravitational in nature, and there has been no direct detection of DM candidates in experiments. The search for DM at the Large Hadron Collider (LHC) is very challenging because

2.2. DARK MATTER SEARCH AT LHC

of multiple reasons, namely, (i) it is unknown whether DM interacts sufficiently strongly with the ordinary matter to enable its production at the LHC; (ii) even if the production mechanism is possible, the unknown mass of the DM candidate creates uncertainty if the centre-of-mass energy at LHC is enough to produce dark matter candidate or not; and (iii) if DM candidate is produced at LHC against these odds, it will not be directly visible in the particle detector like other SM particles. The presence of DM can then only be revealed by an imbalance in the transverse momentum in the detector measured as missing transverse momentum.

Several theories beyond-the-SM (BSM) suggest the particle nature of DM [37], [38]. Several types of particle candidates for DM are proposed, all compatible with the properties of DM observed in astrophysical, cosmological experiments, and the observed relic density of the universe [39]. One of the most favoured hypotheses is that the bulk of the DM is in the form of electrically neutral, stable, weakly interacting massive particle (WIMP) [40], with masses in a range between few GeV to a few TeV. Thus open the possibility of DM production at the high energy colliders [41]. The collider search for DM mainly looks for a pair production of WIMP that recoils against one or more visible SM particle(s). Due to the neutral nature and negligibly small interaction cross-section with the SM particles, WIMPs do not leave a trajectory in the tracker of any noticeable energy deposit in the calorimeter, hence invisible to the particle detectors. The presence of the WIMP in the hadron collider experiment can, therefore, be inferred via measuring the imbalance in the total momentum in the plane transverse to the colliding beams, referred to as missing transverse momentum ($\vec{p}_T^{\text{miss}}$). This gives rise to a detector signature of $p_T^{\text{miss}} + X$, where X

CHAPTER 2. THEORY AND MOTIVATION

is one or more than one SM particle produced in association with DM (photon, jet, pair heavy quarks, Higgs boson, Z/W boson, etc.), commonly dubbed as “mono-X” signature. In this note, we summarize the search for dark matter produced in association with a pair of bottom quarks ($b\bar{b} + p_T^{\text{miss}}$) final state using the pp collision data collected during Run 2 corresponding to an integrated luminosity of 137 fb^{-1} . The search for dark matter produced in association with heavy flavour quarks is complementary to other mono-X searches at LHC. A number of searches have been carried out at both ATLAS and CMS including the mono-top [42], $p_T^{\text{miss}} + t\bar{t}$ [43], and $p_T^{\text{miss}} + b\bar{b}$ [44]. In this analysis, we inspect the simplified model, which overcomes the drawbacks of the traditional EFT models while maintaining a model-independent framework. The signal models for Run-2 DM searches at the LHC have been discussed and agreed upon in the ‘ATLAS+CMS Dark Matter Forum’ [45]. One of the models that meet the characteristics of generic benchmark mediator models is the 2HDM+a model [46, 47]. Figure 2.9 shows the Feynman diagram for the production of $b\bar{b} + p_T^{\text{miss}}$ process within the 2HDM+a framework. The 2HDM+a model is a canonical 2HDM model with additional pseudoscalar singlet interacting with dark matter particles. The model has 5 Higgs bosons as in 2HDM and a CP odd light pseudoscalar mediator (a). The nomenclature used is SM Higgs boson (h), Charged Higgs($H^\pm$), Heavier Scalar(H), Heavier Pseudoscalar (A) and a CP odd light pseudoscalar mediator (a). The mass of respective boson is a parameter of the model

2.2. DARK MATTER SEARCH AT LHC

with a constraint $m_H = m_{H^\pm}$. The model has 14 parameters:

$$\left\{ \begin{array}{c} v, M_h, M_A, M_H, M_{H^\pm}, M_a, m_\chi \\ \cos(\beta - \alpha), \tan \beta, \sin \theta, \\ y_\chi, \lambda_3, \lambda_{P_1}, \lambda_{P_2} \end{array} \right\} \quad (2.6)$$

, where, v is the EW VEV ≈ 246 GeV, m_χ is mass of the dark matter candidate, y_χ is dark-sector Yukawa coupling, θ is the mixing angle of the two CP-odd weak spin-0 eigenstates, α denotes the mixing angle between the two CP-even weak spin-0 eigenstates, $\tan\beta$ is ratio of the vacuum expectation values (VEVs) of the two Higgs doublets, $\lambda_3, \lambda_{P_1}, \lambda_{P_2}$ are the quartic couplings, and the mass of the SM-like CP-even mass $m_h = 125$ GeV. In the decoupling limit $\sin(\beta - \alpha) = 1$, $\lambda_3 = \lambda_{P_1} = \lambda_{P_2} = 3$ are fixed for boundedness of Higgs potential, electroweak precision constraints give $m_A = m_H = m_{H^\pm}$ and dark matter coupling y_χ is set to 1 resulting in five free parameters. The free parameter includes the mass of lighter pseudoscalar m_a , the mass of heavy pseudoscalar m_A , $\tan\beta$, $\sin\theta$, and m_χ . In this model, the coupling between the heavier pseudoscalar A and bottom quark is proportional to $\tan\beta$; hence with high $\tan\beta$, we get a higher cross-section for a process having bottom quark in the hard processes. The heavier pseudoscalar A couples to dark matter particle with strength $\sin\theta$, whereas lighter pseudoscalar couples to dark matter particle with strength $\cos\theta$. The coupling of a and A can be shown as follows.

$$-y_\chi (A \sin\theta + a \cos\theta) \bar{\chi} \gamma^5 \chi \quad (2.7)$$

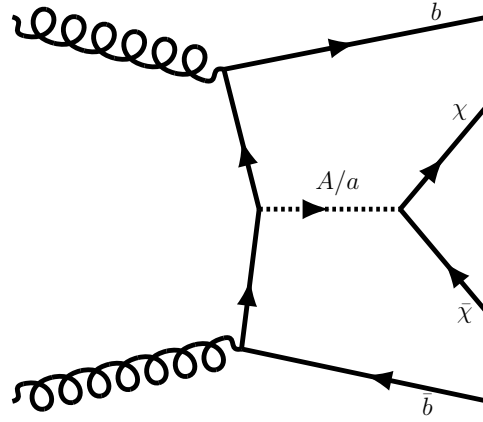


Figure 2.9: The Feynman diagram for the $bb + p_T^{\text{miss}}$ process within the 2HDM+a model

The recommended range of these parameter as per the white paper is as follows:

- $\sim 0.2 \leq \tan\beta \leq 50$
- $0 \leq \sin\theta \leq 1$
- $100 \leq m_a \leq 1000$, lower limit is 100 GeV
- $600 \leq m_A \leq 1200$, lower limit is 580 GeV

For the purpose of production of signal samples, the coupling values are chosen to be in a phase space where analysis is sensitive:

- $\tan\beta = 35$
- $\sin\theta = 0.7$
- $m_\chi = 1 \text{ GeV}$

2.2. DARK MATTER SEARCH AT LHC

Heavier Pseudoscalar m_A GeV	Lighter Pseudoscalar m_a GeV														
600	10	50	100	150	200	250	300	350	400	450	500				
1200	10	50	100	150	200	250	300	350	400	450	500	700	750	1000	

Table 2.2: Phase space for which signal simulation datasets are produced. Same mass grid is used for all three years, 2016, 17 and 18.

These parameter do not change the kinematical quantities significantly but significantly change the production cross-section. Therefore with the choice of parameters for which signal simulation datasets are produced, a full scan of the whole parameter space is possible by rescaling cross-section without producing more samples. the grid in Table 2.2 for masses of m_a and m_A has been used.

As depicted in Figure 2.9 the final state consists of two b-jets and high missing transverse momentum. Therefore, the signal is extracted using two analysis categories based on the number of b-jets. Signal regions are referred to as SR1 and SR2 when there are exactly 1 and 2 b-jets, respectively. SR1 is mainly to catch the signal where one of the b-jet is either not reconstructed or lost due to detector acceptance etc. And contribute to the final search.

Figure 2.10 shows the distribution of true missing transverse momentum (gen-MET) for various values of m_a . It can be noticed that the p_T^{miss} spectrum becomes harder as the mass of light pseudoscalar m_a is increased from $m_a = 100$ GeV to 500 GeV. However, due to the coupling strengths involved in the production of these physics processes, the dependence on the m_A is minimal, as the main contribution comes from the diagram when a is coupling to the dark matter candidate. The dependence of the distribution of gen p_T^{miss} on the dark matter is minimal; however,

CHAPTER 2. THEORY AND MOTIVATION

the distribution does change for the off-shell cases, e.g. $m_a < 2 \times m_{\text{DM}}$.

Further, it can be seen that the bulk of events are in the low p_T^{miss} events, and due to trigger selection used in this search, described in section 4.2, the events with p_T^{miss} lower than 200 GeV are not of much importance. Due to this reason, signal sample production requires p_T^{miss} to be higher than 150 GeV, which is 100 GeV lower than the reco level selection to account for the resolution and turn on effects. Table 2.3 lists the signal samples used in the analysis along with respective cross-sections for 2016 samples. The cross-section values used for 2016 and 2017/18 datasets are also visualised in Figure 2.11. At low m_a the cross-section value is independent of m_A ; however, as the mass value of m_a increases, the change in the cross-section is visible, and change is greater at high m_a values. This is expected due to two mediators in the production mechanism and giving rise to such behaviour.

We have noticed that for the 2016 signal sample production, there was a feature due to which the signal samples were produced without the required threshold of $\text{gen-}p_T^{\text{miss}} > 150 \text{ GeV}$. This resulted in very low signal selection efficiency and, therefore, relatively large statistical uncertainties. Therefore the cross-section used for 2016 and 2017/18 signal samples are different for the moment. A set of private signal samples are being produced for the grid needed for the 2016 setup. The signal samples will therefore be updated once high stats datasets are available and validated.

2.2. DARK MATTER SEARCH AT LHC

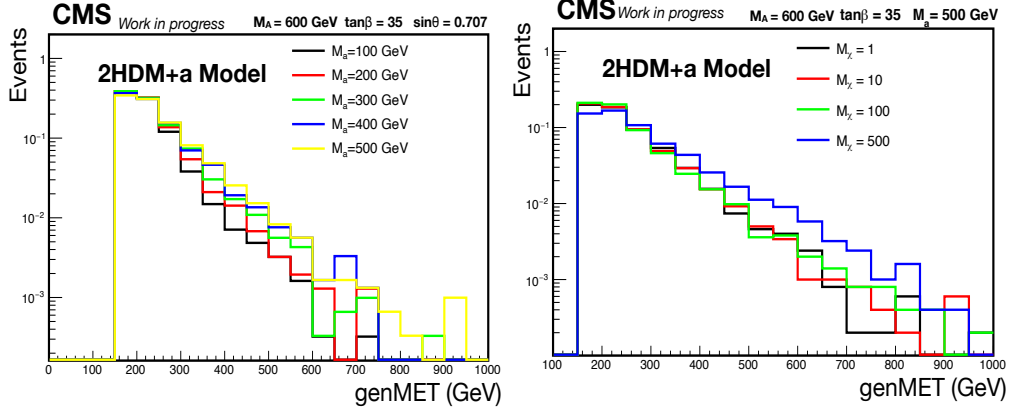


Figure 2.10: The left plots shows how the shape of generator p_T^{miss} variation with change in m_a . In the right plot, the variation of the generator p_T^{miss} w.r.t the mass of the dark matter particle

	Crossection(pb)	
ma (GeV)	for mA = 600 GeV	for mA = 1200 GeV
10	0.91120	0.90530
50	0.62520	0.62480
100	0.42320	0.40960
150	0.28280	0.26940
200	0.18800	0.17330
250	0.12550	0.11630
300	0.08584	0.07677
350	0.06183	0.05271
400	0.04522	0.03663
450	0.03421	0.02660
500	0.02579	0.01897
700	-	0.00575
750	-	0.00438
1000	-	0.00142

Table 2.3: Cross-section values for signal samples for bbDM analysis, the samples are produced using MadGraph5_aMC@NLO for event generation and PYTHIA8 for hadronization

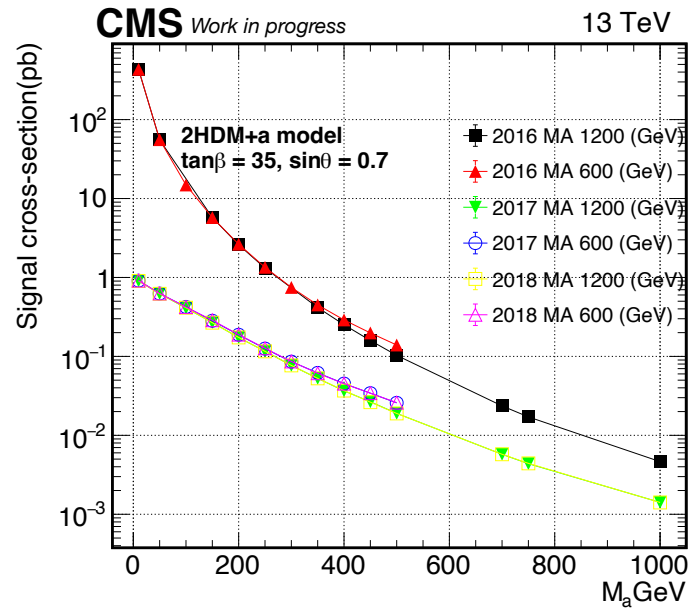


Figure 2.11: The comparison of cross-section values used for the signal samples for 2016, 2017 and 2018 datasets.

CHAPTER 3

EXPERIMENTAL SETUP

As mentioned in the previous sections, the collider search for Dark Matter requires a powerful particle accelerator that collides high energy particles to produce the DM particles. One such particle collider is known as the Large Hadron Collider(LHC). The LHC is the largest and highest-energy particle collider in the world, which has been build in the same housing where a Large Electron-Positron (LEP) was built with a 27km tunnel at the CERN accelerator complex, beneath the France-Switzerland border near Geneva (Figure 3.1).The accelerator complex is shown in Figure 3.2, is a set of machines that accelerate particles to higher energies in increasing order. Each of the machine in the set boosts the energy of a beam of particles, before injecting it into the consecutive machine in the sequence. The Large Hadron Collider (LHC) is the final step of the beam before the beams are collided with an acceleration up to the energy of 6.5 TeV per beam.A simple hydrogen gas cylinder

CHAPTER 3. EXPERIMENTAL SETUP

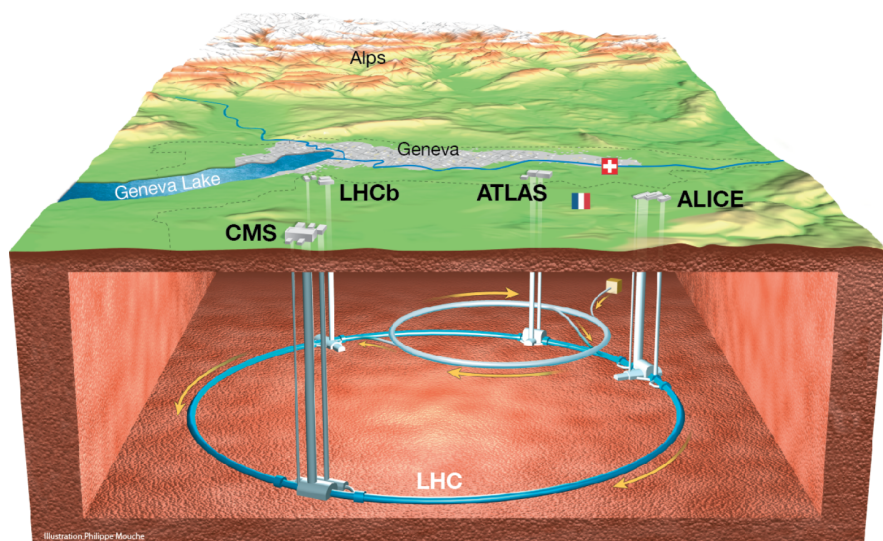


Figure 3.1: A diagram showing the tunnel of the LHC w.r.t French-Swiss border

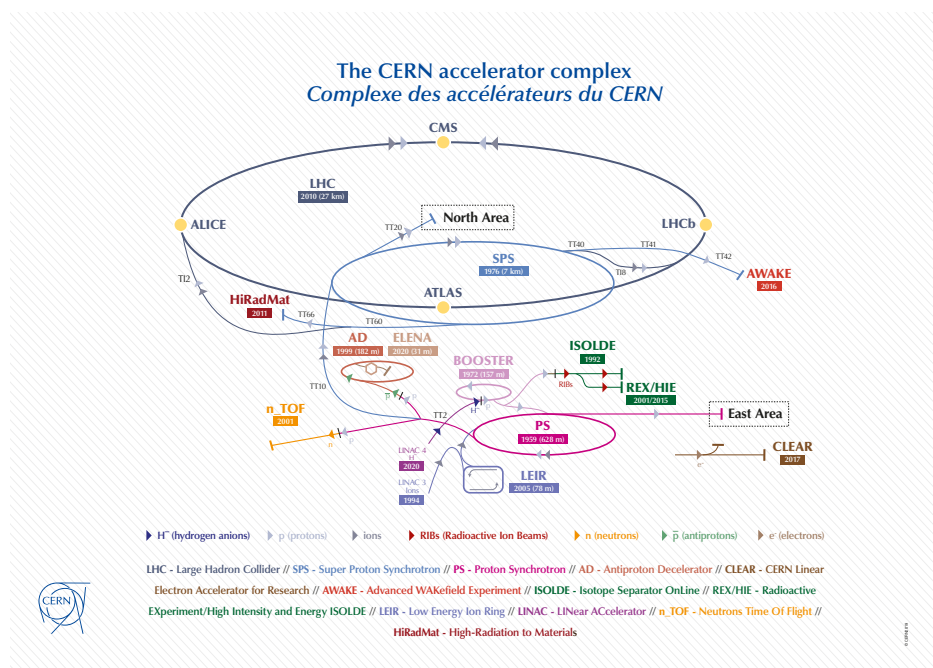


Figure 3.2: CERN ACCELERATOR COMPLEX

works as the proton source; an electric field is applied to get the protons from the hydrogen. The first accelerator in the chain is a linear accelerator called LINAC that accelerates the protons to 50 MeV. Then the beam is injected into the Proton Synchrotron Booster (PSB) to increase the proton energy to 1.4 GeV, which is then followed by the Proton Synchrotron (PS), which further increases the beam energy to 25 GeV. The next step for the protons beams is the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV. Finally, the protons are moved to the two beam pipes of the LHC. Both the beam are circulated but in opposite directions; one circulates clockwise while the other beam circulates anticlockwise. The time taken to fill each LHC ring is 4 minutes and 20 seconds, then it takes another 20 minutes to reach the maximum energy of 6.5 TeV for the protons. Under normal operating conditions, the proton beams circulate for many hours inside the LHC beam pipes and are brought into collision inside the centre of four detectors, namely ATLAS, CMS, ALICE and LHCb, where the total centre-of-mass energy at the point of collision is equal to 13 TeV. The protons (or heavy ions) in the beams are accelerated in bunches; each bunch contains about 10^{11} protons. In general, 2800 proton bunch circulate inside LHC ring with a time difference of 25ns, and the LHC delivers 40 million bunch crossings per second [48] ($\Rightarrow$ 40MHz).

At CERN, there are four main experiments as mentioned above, LHCb (Large Hadron Collider beauty), ALICE (A Large Ion Collider Experiment), CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC ApparatuS). Out of them, two are the general-purpose detector designed with the discovery of Higgs Boson in mind, ATLAS and CMS. ALICE is specifically made to analyze quark-gluons plasma created by heavy-

CHAPTER 3. EXPERIMENTAL SETUP

ion collision, while LHCb is built explicitly to study B mesons.

3.1 Large Hadron Collider

The LHC stands for Large Hadron Collider. It is called the Large Hadron Collider because of its size, the famous 27km circumference, the colliding particles are protons (one of the particles of a more extensive set of particles called Hadrons), and of course because it is a particle collider. The LHC is designed to accelerate two circular protons beams in opposite direction up to a COM energy $\sqrt{s} = 14\text{GeV}$ with a peak instantaneous luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. Besides proton beams, proton-lead, lead-lead and Xenon-Xenon collision also take place in LHC with COM energy of $\sqrt{s} = 8.16, 5.02$ and 5.44TeV respectively. The first beam in LHC was injected on 10th September 2008; however, because of some faulty electrical connection between the superconducting magnets, it was stopped. Finally, on 20th November 2009 after a year the first successful proton-proton (pp) collisions at $\sqrt{s} = 900\text{GeV}$ took place. In March of 2010, the energy of proton beams was increased to 3.5 TeV leading to a $\sqrt{s} = 7\text{TeV}$. Next, data collection started in April of 2012 and ended in February of 2013, which was with $\sqrt{s} = 8\text{TeV}$, and was called as “Run I” of LHC. After LS1 (called as long shutdown 1) for 2 years, the data collection for “Run II” started with each proton beam of energy 6.5 TeV giving $\sqrt{s} = 13\text{TeV}$. Run II ended in March of 2018, finishing up to 150fb^{-1} of data, and LHC entered the LS2 period. After the current LS2, “Run III” for LHC is scheduled at $\sqrt{s} = 14\text{TeV}$ between 2021-2024 with an expected integrated luminosity of 300fb^{-1} . After the start of LS3, the accelerator

3.1. LARGE HADRON COLLIDER



Figure 3.3: Overview of HL-LHC Plan [6]

infrastructure will be improved to accommodate very high values of instantaneous luminosity ranging between 5.0 to $7.5 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ resulting in 3000 to 4000fb^{-1} of data for the full accelerator lifetime. This next stage of LHC with very high luminosity is called “HL-LHC”. Figure 3.3 illustrates the detailed timeline of LHC run operations. Before the injection of the proton beams into the LHC machine, protons are accelerated by a series of accelerators and each accelerator stage raises the energy of the protons plus transfers them to the next stage which has been mention in the introduction of this chapter. The details about the LHC parameters and performance of the LHC will be discussed in the following section.

CHAPTER 3. EXPERIMENTAL SETUP

3.1.1 LHC Parameters

The LHC is a collider with a two-ring superconducting proton accelerator and is designed to explore the TeV energy scale. The design of LHC is based on fundamental principles of physics used in an exquisite way. A few of the main components of LHC are explained in the following section:

3.1.1.1 Magnetic system in LHC

The LHC is made of eight arcs and eight ‘insertions’; the arcs are 2.45km long with eight 545m long straight sections with a total of 154 dipoles ‘bending’ magnets in each arc. An insertion consists of two (one at each end) transition regions which are also called ‘dispersion suppressors’ plus a long straight section. Each arc contains 23 arc cells with a regular lattice structure. Each arc cell is 106.9 m long and has a main dipole magnet(Figure 3.4) plus quadrupole magnet and few other magnets. 1232 superconducting dipoles in the LHC ring. which operate reliably at the nominal magnetic field of 8.33 T which is 10^5 times that of the Earth’s, corresponding to the centre-of-mass energy of 14 TeV. The LHC magnets are made up of niobium-titanium (NbTi) alloy and the superconducting magnet system operates in superfluid helium(≈ 96 tonnes of it in cryogenic system) at 1.9 K. Further, the cables of the magnets are of a very special design and conduct current without resistance in their superconducting state

3.1. LARGE HADRON COLLIDER

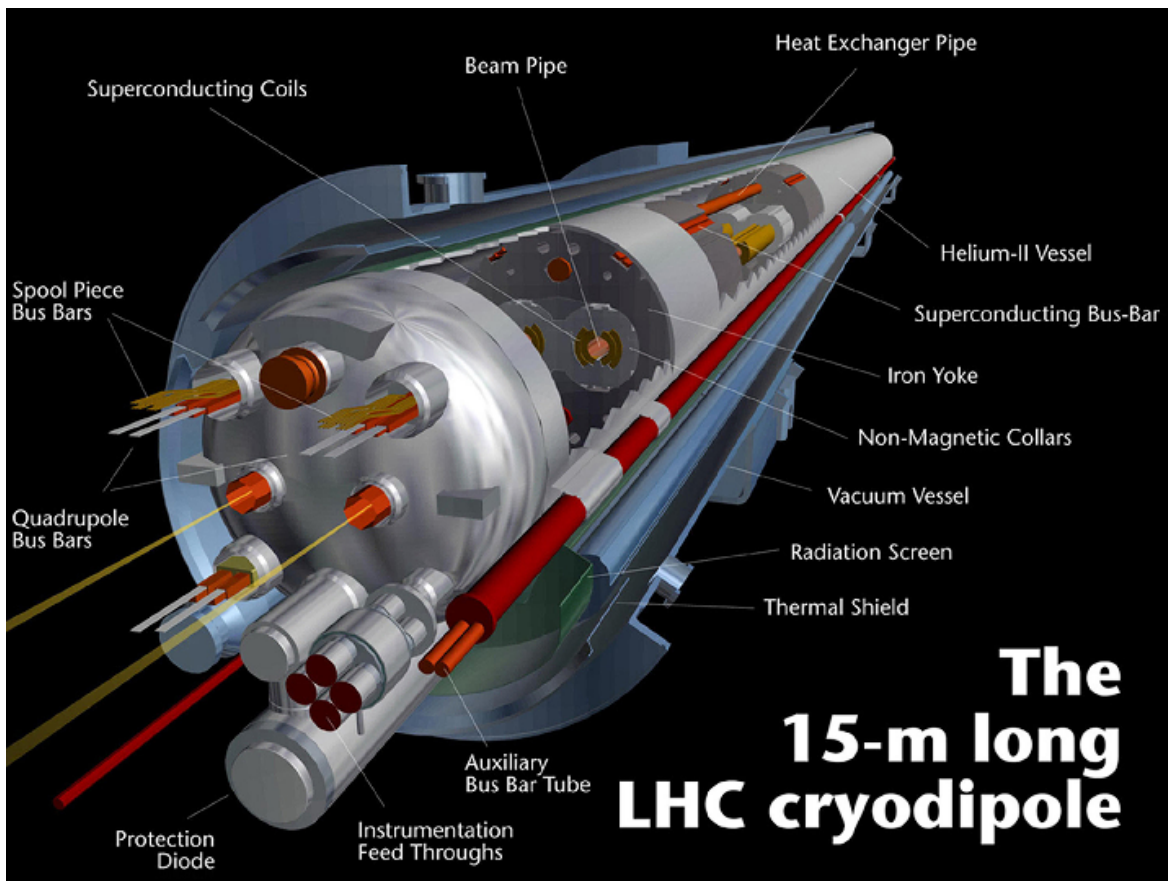


Figure 3.4: Computer generated diagram of an LHC dipole magnet [7].

CHAPTER 3. EXPERIMENTAL SETUP

3.1.1.2 Collisions

The protons (or lead nuclei) inside the LHC ring are accelerated to speeds close to the speed of light and then collided. With such high energy collisions in the centre of detectors, the information about the smallest components of matter can be extracted. In these collisions, new particles can be created, and which will provide information about physics beyond the SM of physics. As mentioned, each beam of a proton consists of bunches that contain more or less 10^{11} protons. The collision happens between the two bunches at the centre point of the detectors. At the centre points of the detectors, the beams are squeezed down as much as possible to increase the chances of a collision. However, the cross-section of a proton is tiny; therefore, even after squeezing the bunches down to 64 microns at the interaction point, only very few head-on collisions per bunch crossing occur. A pictorial representation of the collision of two beams is shown in Figure 3.5 The bunches cross every 25ns, with around 600 million collisions happening per second. Most of the protons do not collide at all, and hence the beams are kept circulating for hours to maximize the collisions from each bunch. A rate for the collision can be calculated from the luminosity and the cross-section of the bunches colliding; the total p-p cross-section at 7 TeV is approximately 110 mbarn which contains inelastic, single diffractive and elastic cross-section, however only the inelastic scattering is seen by the detector which is 60 mbarn hence with the luminosity of 10^{34} the inelastic event rate becomes $10^{34} * 60 * 10^{-27} = 600$ million per second and with 25ns time difference between bunches, the event rate comes out to be 19 collisions per bunch crossing.

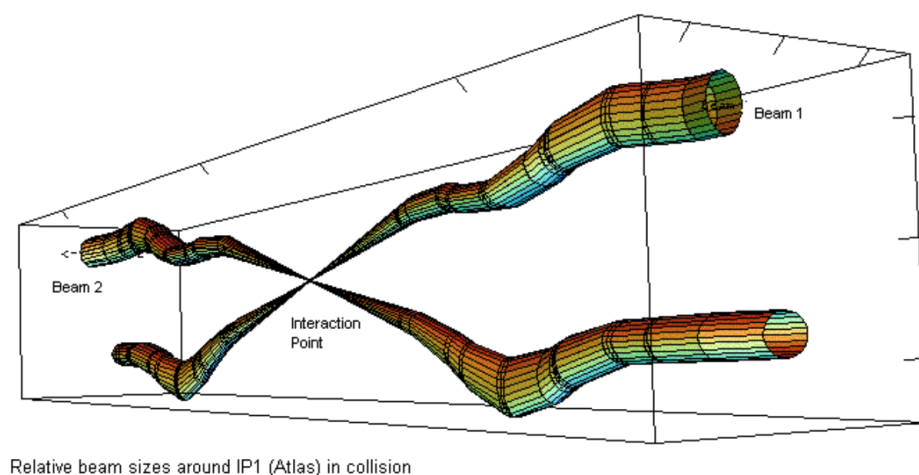


Figure 3.5: Proton beams squeezed at the interaction point

From the physics point of view, when the collision of the two protons occurs, there are various physics processes that go on. When the proton collides head-on, the collision is called the hard scattering process. Figure 3.6 shows the hard scattering process. The hard scattering is the process with very high momentum transfer due to parton interaction. The interactions between all the other protons which were not involved in the hard scattering give rise to “**Underlying Events**”(UE). The proton can radiate a gauge boson before and after the hard scattering, this is known as “**Initial State Radiation**”(ISR) and “**Final state Radiation**”(FSR) respectively. One of the most important by-products of the collision is the “**Pileup**”; in every collision, the bunches cross one another, more than one proton-proton collision takes place, which is known as pileup. In general, 23 collisions occur between different pairs of protons in every bunch crossing on average, which leads to low p_T events called Pileup events. The number of pileup vertices (interaction points) depends on the

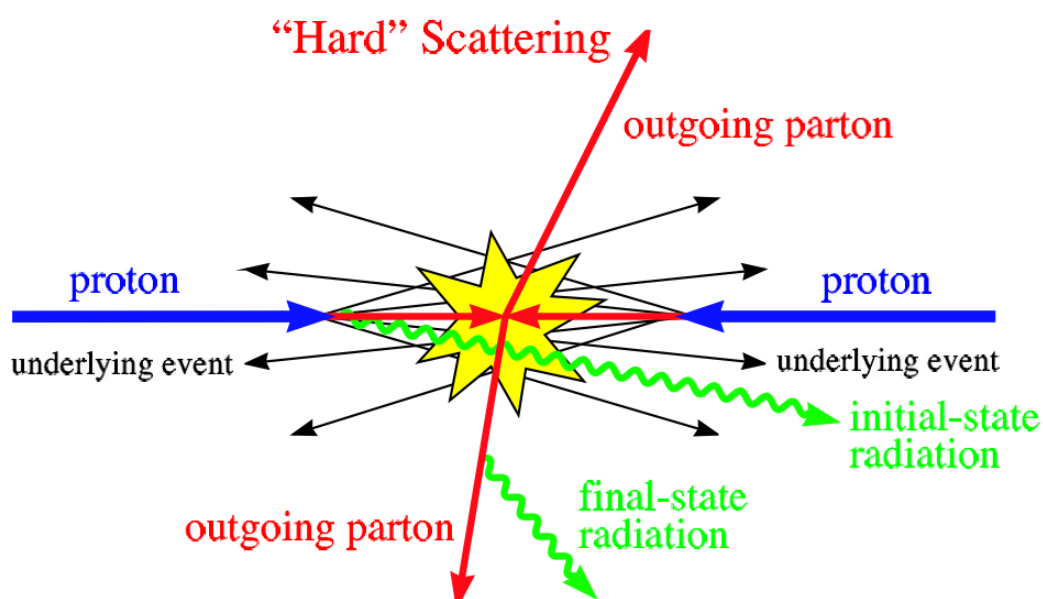


Figure 3.6: Schematic of a hard scattering proton-proton collision [7].

3.1. LARGE HADRON COLLIDER

Circumference	26659 m
Dipole operating temperature	1,9 K
Number of arcs (2450 m long)	8
Number of lattice cells per arc	23
Number of straight sections (545 m long)	8
Number of magnets (dipoles, quadrupoles ... dodecapoles)	9300
Number of RF cavities	8 per beam
Momentum at collision	7 TeV/c
Momentum at injection	450 GeV/c
Peak magnetic dipole field	8.33 T
Main Dipole Length	14.3 m
Number de Ni-Ti filaments in each strand	6500
Minimum distance between bunches	~ 7 m
Bunch spacing	25 ns
Design Luminosity	$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
No. of bunches per proton beam	2808
No. of protons per bunch (at start)	$1.15 \cdot 10^{11}$
Number of turns per second	11245
Beam lifetime	10 h
Average crossing rate	31.6 MHz

Table 3.1: The designed LHC parameters [1].

integrated luminosity; it increases with the increase in the Integrated Luminosity. For Run III, the number of pileup vertices has been expected to be from 140 to 200. A list of few of the numbers used in LHC are given in Table 3.1

3.2 The CMS Detector

The CMS detector is an acronym for **Compact Muon Solenoid** detector. It is a multi-purpose detector covering almost 4π and is one of the two general-purpose detectors being operated at the LHC at CERN. It is located at the IP5 of the LHC, however 100m underground.

As discussed in section 3.1.1.2, about 19 inelastic collision occur in every collision. And with this, around 1000 charged particles will be ejected out from the interaction region after each bunch crossing in 25ns. The problem of detecting the events of importance further increases due to the pileup and the detector's response time. However, CMS is built for dealing with such problems by having the following features [10]:

- Good muon identification and momentum resolution over a wide range of momenta and angles
- Good charged-particle momentum resolution and reconstruction efficiency.
- Efficient triggering and offline tagging of particles
- Good electromagnetic energy resolution.
- Good missing-transverse-energy and dijet-mass resolution

All the above-mentioned points can be addressed by looking at the design of the CMS detector in Figure 3.7 and will be discussed further in this section. The detector has a onion like structure, where each layer is a detector in itself for stopping,

3.2. THE CMS DETECTOR

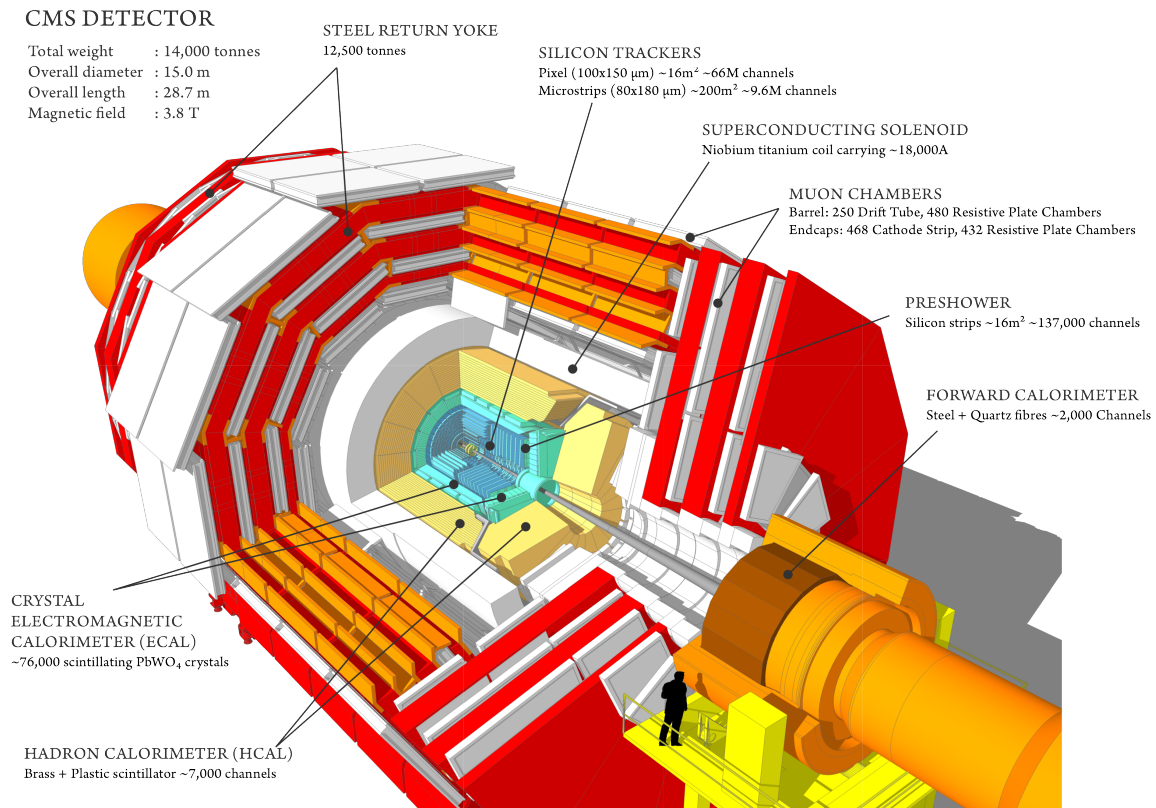


Figure 3.7: Sectional view of the CMS detector [8].

CHAPTER 3. EXPERIMENTAL SETUP

tracking or measuring various particle coming out from p-p or heavy-ion collisions. CMS detector is cylindrical in shape and hence each layer takes a cylindrical shape as well. The components of the detector that are parallel to the beamline are called barrel region and the two ends closing the detector are called endcap regions. At the centre of the CMS detector, a magnetic field goes up to 3.8T produced by the superconducting solenoid. Then comes the silicon tracker, which determines the tracks of charged particles that are coming from the collisions. After the tracker, calorimeters come into the picture that gives the details about the energy of the particles. There are two types of calorimeter used in CMS, electromagnetic and hadron calorimeter. The electromagnetic calorimeter is used to get the energy of particles like the photon, electrons, pions etc., whereas the hadron calorimeter is specifically for the hadrons, i.e. proton and neutrons. Immediately after the hadron calorimeter, there is a solenoid with an iron yoke which is used to track the particles. After the solenoid, the outermost layer of the CMS detector is specifically designed to track muons, where the drift tubes(DT), resistive plate chambers(RPC) are mounted. The path of various particles inside the CMS detector is shown in Figure 3.8

To better understand the workings of the CMS detector, the first step is to understand the geometry of the CMS detector, and it starts with coordinate a system. In CMS, the origin is located at the center of the detector(the interaction point). The z -axis is along the beam line and the $x - y$ plane is in the transverse direction to the z -axis. The positive y -axis points upwards and the positive x -axis points the center of LHC ring. With r, θ & ϕ , where ϕ , the azimuthal angle, θ is the angle from the beam axis and r is the distance from the interaction point. The ϕ is 0 along the

3.2. THE CMS DETECTOR

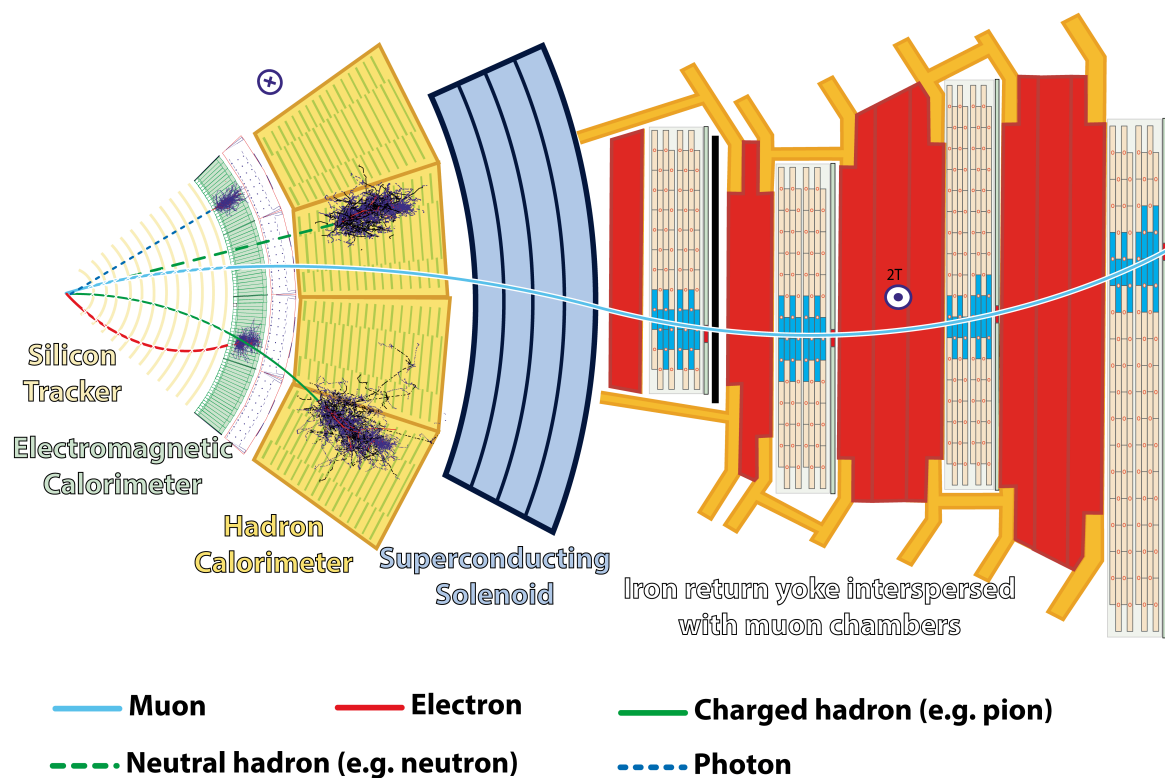


Figure 3.8: Schematic of transverse slice of the CMS detector [9].

CHAPTER 3. EXPERIMENTAL SETUP

positive x -axis, the θ is not used as such in the CMS, because it is not a Lorentz Invariant quantity, so to have a Lorentz Invariant quantity from θ , variable called rapidity is defined with a momentum $p = (E, p_x, p_y, p_z)$ as:

$$y = \frac{1}{2} \log \frac{E + p_z}{E - p_z} \quad (3.1)$$

for $\beta \rightarrow 1$ (massless) limit, the rapidity is given as:

$$\eta = \frac{1}{2} \log \left(\frac{1 + \cos(\theta)}{1 - \cos(\theta)} \right) \quad (3.2)$$

In this limit the rapidity variable gets a new name pseudo-rapidity, and along with ϕ a distance ΔR is calculated as :

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} \quad (3.3)$$

The value of η ranges from 0 to 5, where 0 implies particle in $x - y$ plane, and five implying along the beamline. Now the nomenclature and variables are defined, in the next section, each individual detector system in the CMS detector is discussed.

3.2.1 Inner Tracking System

One of the most important parts of reconstructing the physics objects at CMS is to measure the tracks made by the charged particles; the inner tracking system of CMS has been designed for this purpose. It not only provides precise and efficient tracking but also give very precise secondary vertex reconstruction. The length of

3.2. THE CMS DETECTOR

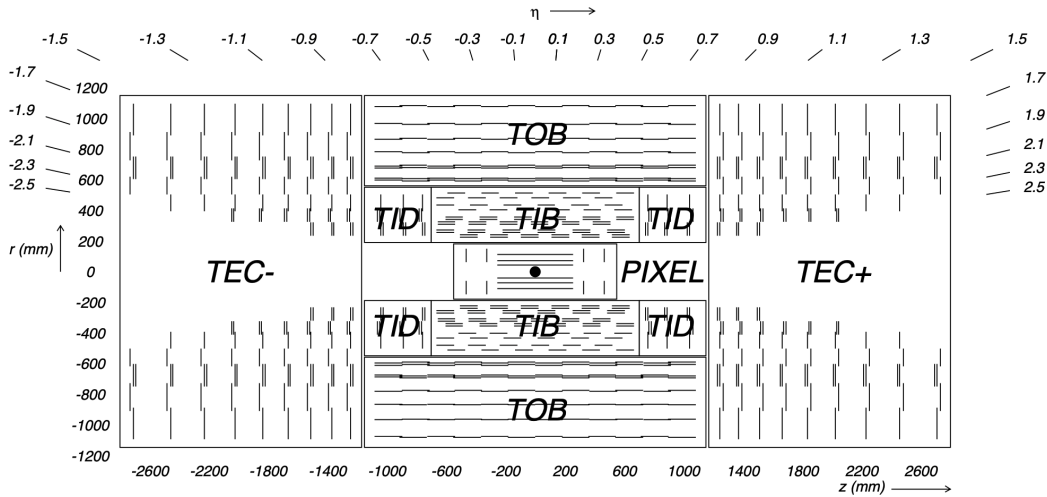


Figure 3.9: Schematic of the CMS tracker, each detector module is depicted by a line [10].

the inner tracking system is 5.8m, and the diameter is 2.5 m. The magnetic field of 3.8T on the full volume to the tracker homogenously is provided by the CMS solenoid. The tracker receives the highest level of radiation compared to the rest of the detector, further with the LHC luminosity of $10^{34} \text{cm}^{-2} \text{s}^{-1}$ and high pileup, the detector technology featuring high granularity and fast response is used in the tracker. Therefore silicon detector technology is used in the entire tracker system. There are two parts of the tracker, a pixel detector with 3 barrel layers and a 10 barrel layer silicon strip tracker extending up to 1.1 m and these detectors are enclosed by 2 and 9 disks in each endcap region. In total, 1440 pixel detectors and 15148 silicon strip detectors are mounted in the inner tracker system of CMS. A schematics of the tracker is shown in Figure 3.9

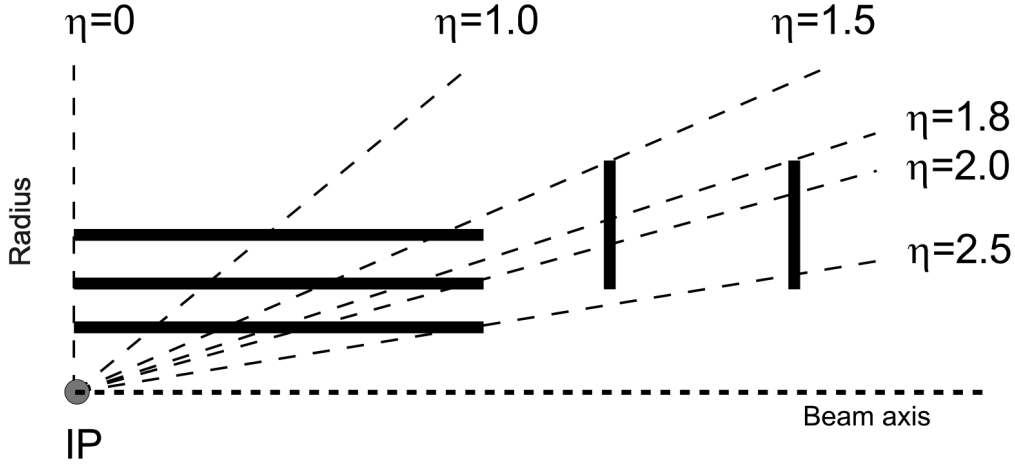


Figure 3.10: The geometrical layout of the pixel detector with η coverage [10].

3.2.1.1 Pixel Detector

The pixel detector is the closest detector to the collision point; it receives the highest amount of radiation due to the pp collisions. With pixel detector, precise tracking point (z -axis and $r - \phi$ plane) and impact parameter resolution is achieved, which is essential for the reconstruction of secondary vertices from b and τ decays, and forming seed tracks for the outer track reconstruction and HLT.

Consisting of three barrel regions (BPix) and two endcaps (FPix) it covers the pseudorapidity η from -2.5 to 2.5 . The geometrical layout of the pixel detector is shown in Figure 3.10. As mentioned that the pixel detector receives highest amount the radiations resulting in fake rates and degradation of resolution, to overcome this the pixel detector upgraded before the start of 2017 data taking, the upgrade was called Phase1 Pixel Upgrade [49] which included 1 more barrel layer(total 4) and 1

more end-cap disk, along with new readout chip, reduction of material, and more efficient cooling and powering systems [49]. In the next phase i.e. Phase2, with HL-LHC(High Luminosity LHC), the tracker will be completely upgraded to a new to cope up with very high pileup as well as very high levels of radiation.

3.2.1.2 Silicon Strip Detector

In the silicon strip detectors, long strips of the silicon are used instead of the smaller pixel used in pixel detector; the sensor elements are single-sided p-on-n-type silicon micro-strip sensors. The strip tracker is divided into four subsystems as shown in Figure 3.9, Tracker Inner Barrel(TIB) is the innermost barrel region with four layers then comes the Tracker Inner Disks(TID), inner endcap region with three layers after this region comes Tracker Outer Barrel(TOB) which has 6 barrel layers and finally the Tracker endcap (TEC) which has nine endcap layers. The silicon strip tracker provides three-dimensional information for the separation of the particle tracks with a single-sided module and a stereo module, with strips tilted by 100 mrad mounted back to back, a configuration with which two inner layers of TIB, TID and TOB, 1st, 2nd and 5th ring of TEC are equipped. With this setup, at least eight measurement points can be obtained for high momentum tracks with $|\eta| < 2.5$. The transverse momentum resolution of the tracker varies in the range from 0.7% – 1.5% at 1 and 100GeV, respectively. The tracks reconstructed by the CMS tracker are used in particle identification, vertex reconstruction, charge identification, the momentum of the charged particle.

3.2.2 Electromagnetic Calorimeter

The calorimeter is a device that measures the energy of the particles, a charged particle interaction with the calorimeter material and produces showers of secondary (mostly electrons and via Bremsstrahlung) particles of lower energies till the energy of the primary particle goes below the critical energy. There are certain particles that interact very less with the calorimeter, for example, neutrinos; the energy carried by them cannot be measured directly; it is calculated with the use of the momentum balance of the full interaction and is called Missing Transverse Energy (MET).

The CMS electromagnetic calorimeter (ECAL) is a hermetic homogeneous calorimeter that is made up of lead tungstate (PbWO_4) crystal mounted in both barrel and endcap region. With the use of PbWO_4 , a high-density crystal gives the calorimeter all important characteristics in the LHC environment, i.e. it is fast, has fine granularity and is radiation-resistant. Further, to add the capability to detect two photons from one of the postulated decay of Higgs boson, a pre-shower detector is placed in front of the endcap crystals. The properties of PbWO_4 crystal-like, high density (8.28g/cm^3), short radiation length of 0.89cm and small Molière radius of 2.2cm are best suited for a CMS electromagnetic calorimeter, giving it high granularity and high energy resolution. The total thickness of the ECAL is kept greater than 25 radiation lengths. The PbWO_4 crystal emits the scintillating blue light ($\lambda \sim 420\text{nm}$) after ionizing, which is received by the photodetectors. The total amount of scintillation and the energy deposited by photons and electrons in the crystal are proportional, and with this proportionality, the energy of the particles is measured. The layout of

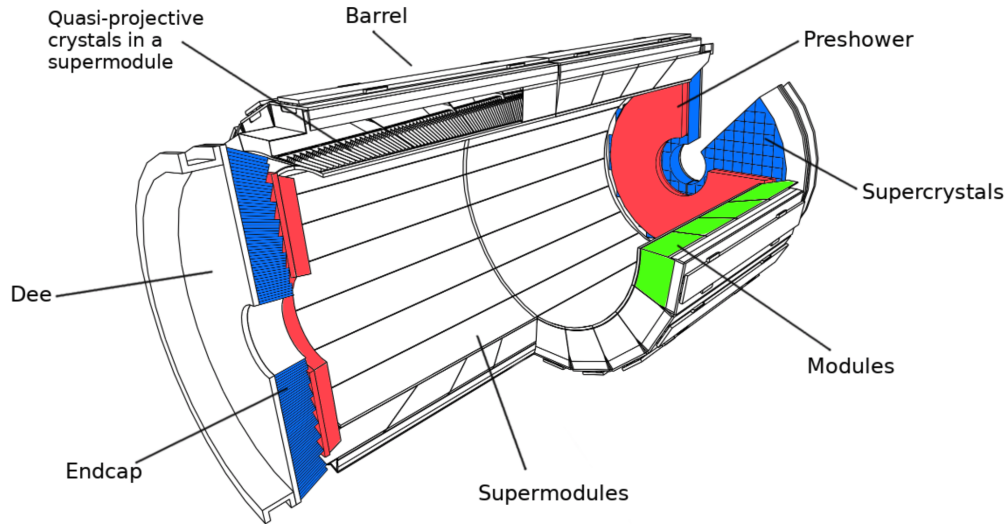


Figure 3.11: The layout of the CMS ECAL [11].

CMS ECAL is shown in Figure 3.11

3.2.2.0.1 ECAL Barrel: The barrel region of the ECAL called EB which covers $|\eta| < 1.479$. The EB has 360-fold granularity in ϕ and $(2 * 85)$ in η , with a total number of crystals to be 61200, extending from inner radius 1.29 m to 1.77 m. In EB, silicon Avalanche PhotoDiodes (APD) are used to receive the light coming out of the PbWO_4 crystals. Each APD is operated at a voltage of 340-430 V and at a gain of 50. The cross-section of crystal in the $\eta - \phi$ plane is 0.0174×0.0174 , with the total volume of all EB crystals at 8.14m^3 and weight at 67.4T.

3.2.2.0.2 ECAL Endcap: The ECAL endcap also called EE covers the pseudorapidity range $1.479 < |\eta| < 3.0$. With the 1.6 cm shift toward the interaction point in the 3.8T magnetic field, the longitudinal distance between endcap envelop and the

CHAPTER 3. EXPERIMENTAL SETUP

interaction point is 3.154m. In EE, identically shaped crystals are grouped together in the stack of 5×5 crystals which are then called supercrystals, or SCs. The two individual halves of each disk in EE is called *Dee*, and each *Dee* contains 3662 crystals arranged in two ways, 138 SCs on the inner circumference and 18 partial SCs on the outer circumference.

3.2.2.0.3 ECAL Preshower: The preshower is a sampling calorimeter (a type of calorimeter that has a layer of passive absorber with an active detector layer). In CMS preshower, lead radiators are the passive absorbers that initiate the electromagnetic showers, and silicon strip sensors are the active detector layer to measure the energy deposited and the transverse shower profiles.

The preshower detector with a total thickness of 20cm ($3 * X_0$ long, where X_0 is radiation length) was mainly designed to identify neutral pions (π^0) in the endcap regions in $1.653 < |\eta| < 2.6$ range, however, it also improves the determination of the position of photons and electrons with high granularity.

3.2.3 Hadronic Calorimeter

The CMS Hadron Calorimeter (HCAL) has been designed to measure the energy of the particles that are called “hadrons”, the particles which are made up of quarks and gluons, the most common example includes protons, neutrons, pions and kaons etc. It further provides a measurement of the presence of non-interacting and/or uncharged particles such as neutrinos indirectly. The CMS hadron calorimeter must have the following properties [50]:

- Good hermeticity
- Good transverse granularity
- Moderate energy resolution
- Sufficient depth for hadron shower containment

Further, the choice of PbWO_4 crystals for ECAL influences the design and performance of HCAL; hence the optimization of the HCAL and ECAL is necessary. For this reason, a lateral granularity $\Delta\eta * \Delta\phi = 0.087 * 0.087$ for $|\eta| < 2.0$ is chosen to match that of the ECAL which is very sufficient for di-jet separation mass resolution. The HCAL has two separate regions. The barrel(HB) and endcap (HE) region of the HCAL covers the pseudorapidity range ($|\eta| < 3.0$) and a forward calorimeter that covers the rest of the pseudorapidity range, i.e. $3.0 < |\eta| < 5.0$. A schematic of CMS HCAL is shown in Figure 3.12

3.2.3.1 Barrel HCAL (HB)

The HB covers the pseudorapidity range of $|\eta| < 1.3$; it is a sampling calorimeter with two half-barrel sections(HB+ and HB-).

Absorber: The HB contains 36 identical azimuthal wedges which make up two half barrels, HB+ and HB-. The wedges are made from flat brass absorber plates(70% Cu, 30% Zn) which are aligned parallel to beam axis. However,for structural strength, the innermost plate and the outermost plate are made of stainless steel.

Scintillator: A sampling calorimeter requires a active medium to capture the photons, in case of CMS HCAL, a tile and wavelength shifting fibre concept is used in

CHAPTER 3. EXPERIMENTAL SETUP

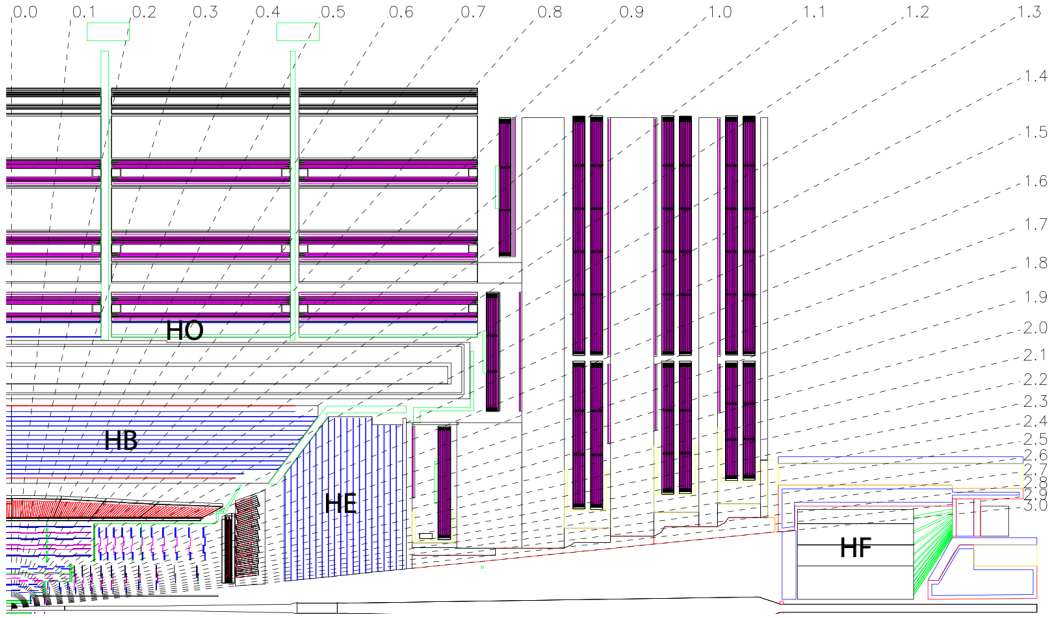


Figure 3.12: The layout of the CMS HCAL [10].

an active medium. The HCAL consists of around 70000 tiles which are grouped in a single mechanical scintillator tray unit as shown in Figure 3.13. A scintillator tray is made of individual scintillators with edges painted with white color and they are wrapped in Tyvek 1073D (made of pure polyethylene fibres) material and attached to a 0.5mm thick plastic substrate with plastic rivets. Each layer of HB contains 108 scintillator trays, however, in case of damage, individual scintillator trays can be replaced without disassembly of the absorber.

3.2. THE CMS DETECTOR

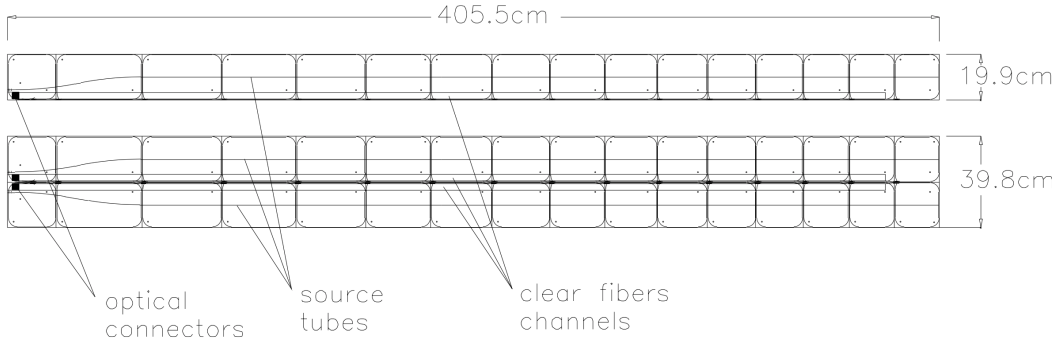


Figure 3.13: Scintillator Trays.

3.2.3.2 Endcap HCAL (HE):

The endcap region of CMS HCAL covers $1.3 < |\eta| < 3$ range, with is 13.2% of the solid angle, it also contains about 34% of the particles produced in the final state.

absorber:

The absorber has been designed to minimize the cracks between HB and HE. The brass plates that are 79mm thick with 9mm gaps to accommodate the scintillators are bolted in a staggered geometry with no projective “dead” material. The total length of the calorimeter, including electromagnetic crystals, is about ten interaction lengths (λ_I). The outer layers of HE have a cutout region for the installation of the photodetectors and front-end electronics. Further, the HE contains 20916 plastic tiles and has granularity of $0.17 \times 0.17 (\Delta\eta \times \Delta\phi)$ for $|\eta| \leq 1.6$.

Scintillator trays: In the scintillator tray attached to HE, Wavelength shifting (WLS) fibres are used to capture scintillating light and is designed in a way that minimizes the dead zones. The narrow edges of the scintillator are painted and grouped together into a frame to form a tray with 1368 trays in total.

CHAPTER 3. EXPERIMENTAL SETUP

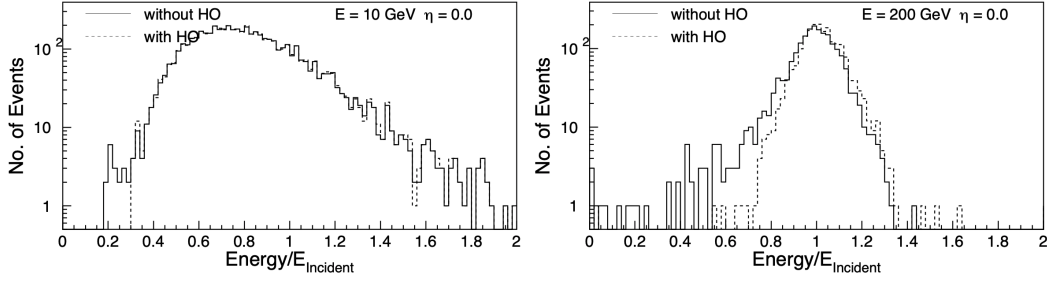


Figure 3.14: Distribution of measured energy scaled to incident energy for pions at $\eta = 0$ at incident energies of 10 GeV(left) and 200 GeV(right). The solid and dashed histograms are measurements without and with the HO.

3.2.3.3 Outer HCAL (HO)

The combined stopping power of EB plus HB is insufficient for containment for hadron showers in the central pseudorapidity region, therefore to get the adequate sampling depth for the $|\eta| < 1.3$ range, the HCAL is also extended outside the solenoid, which is called outer calorimeter (HO). The solenoid coil is utilized by HO as an additional absorber with $1.4/\sin\theta$ interaction length generally used to detect late starting showers and further to measure shower energy after HB. In HO, Silicon Photomultipliers (SiPM) are used instead of HPDs due to the insensitivity to magnetic fields in the return yoke. The inclusion of the HO in the HCAL improves the energy resolution as shown in Figure 3.14 [51].

3.2.3.4 HCAL Forward (HF)

The HCAL forward calorimeter receives a very high particle flux, 760 GeV per p-p interaction, compared to only 100 GeV for the rest of the detector. Due to this rea-

son, stainless steel and quartz fibres have been used as an absorber and active material, respectively, in HF. The signal is generated when Cherenkov light is generated by charged shower particle above the Cherenkov threshold and is captured by the quartz fibres is measured by multi-anode Photomultiplier tubes (PMT), which are shielded behind 40 cm thick steel.

3.2.4 Solenoid Magnet

The superconducting magnet of CMS has a 6m internal diameter and is 12.5m in length, designed to reach the magnetic field of 4T, which is about 10^5 times the magnetic field of the earth. In practice, the magnetic field of the magnet is kept at 3.8 T to increase the life of the magnet system as decided by the CMS collaboration. To return the magnetic flux, 10000 tonnes of the iron yoke, which consist of 5 wheels and two endcaps, each of which has three disks, is used, and the diameter of the iron yoke is 14.6m with 21.6m length. The high magnetic field of 3.8 T bends the charged particles in the tracker giving curved tracks which are used to calculate the momentum of those particles(provides good momentum measurement up to $|\eta| < 2.4$), the magnetic field also helps in reducing the pileup as it bends the low $p_T(\leq 0.75 \text{ GeV})$ particle very strongly so that they do not even reach the ECAL.

3.2.5 CMS Muon System

Good muon detection is a very powerful tool for the discovery of various physics processes with high background and high luminosity. There are many BSM searches

CHAPTER 3. EXPERIMENTAL SETUP

where muon comes in a final state. So getting a good handle on the identification and isolation of muons is very important and is achieved by using multiple sub-detectors combined at CMS. In general, the information from all the subdetectors is taken into account in order to measure muon properties, but the ones specific to muons are Drift Tubes (DT), Cathode Strip Chamber (CSC) and Resistive Plate Chamber (RPC). The full muon system has the capacity to determine the charge and reconstruct the momentum of muons over the entire energy range of the LHC. Muons are minimum ionizing particles(MIP), and hence they escape the detector by losing a very small fraction of energy in the form of ionization loss which is captured by three muon subdetectors.

The **DTs** [52] are mounted in barrel part of the detector, they cover pseudorapidity up to $|\eta| < 1.2$. DT has four stations, namely, MB1, MB2, MB3 and MB4, which are alternated with the segmented yoke. The inner three cylinders have 60 DTs each, and the outer ones have 70 DTs. In the $|\eta| < 1.2$ region, the low muon occupancy rate and magnetic field contained in the yoke allow easy handling of DTs. Drift tubes are made up of drift cells, which are a gas-filled tube surrounded by parallel planes of aluminium and I-shapes aluminium beam cathode as shown in Figure 3.15. The drift cells operates at a gain of 10^5 , which allows them to work within a plateau of efficiency with a wide threshold range. The gas used in the drift cell is Ar/CO₂(85%/15%) gas mixture as it has good quenching properties and saturated drift velocity. The efficiency and resolution of individual drift cell are about 99.8% and 180 μ m respectively.

The **CSCs** [53] are multiwire proportional chambers that consist of 6 anode wire

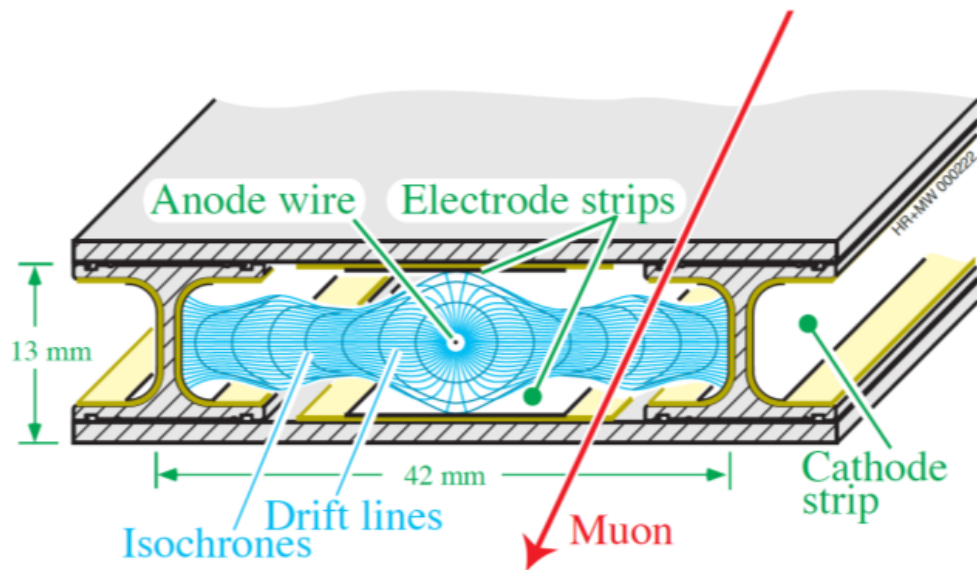


Figure 3.15: Sketch of a cell showing drift lines and isochrones. The plates at the top and bottom of the cell are at ground potential. The voltages applied to the electrodes are +3600V for wires, +1800V for strips, and -1200V for cathodes [10]

CHAPTER 3. EXPERIMENTAL SETUP

planes interleaved with seven cathode panels, and the CSC subdetector is composed of rings made up of trapezoidal chambers mounted on eight disks divided in two endcaps covering $0.9 < |\eta| < 2.4$. In total, there are 468 trapezoidal chambers in the CSC subdetector. Each chamber has six detecting layers, composed of an anode wire between two cathode strips made up of copper. CSCs chambers have four stations, namely ME1, ME2, ME3, and ME4, where ME implies Muon Endcap. The names are designated in a more detailed way as $ME \pm S/R$, where the $\pm$ sign indicates which endcap, R is the ring number, and S indicates the “station”. The CSCs are used to measure the azimuthal angle (ϕ) of the muon tracks precisely. The CSCs provide a spatial resolution of $80 \mu\text{m}$ with a gas mixture of 10% CF_4 , 40% Ar and 50% CO_2 . Further, the CF_4 is also used to prevent polymerization on the wires of CSC. The RPCs [54] are gaseous particle detector which utilizes a uniform and constant electric field produced by two parallel plates. An RPC that combines spatial resolution and time resolution, which is comparable to that of scintillators. However, an RPC can tag the time of an ionizing event very fast (1-2 ns) compared to 25ns, which is the time between 2 consecutive LHC bunch crossings (BX). The fastness of RPC in tagging time of the ionizing event make it a fast muon trigger device, which unambiguously identifies the BX to which the muon track is associated.

The CMS RPC cover both barrel as well as endcap region. In total 1056 RPC detectors are mounted in CMS covering $|\eta| < 1.9$. In the barrel region, five wheels are placed (480 RPCs), whereas eight disks are mounted in the endcap region (576 RPCs), four on the positive side of z -axis and four on the negative side of z -axis. The gas used for the gaps in RPC is 95.2% of tetrafluoroethane Freon ($\text{C}_2\text{H}_2\text{F}_4$), 4.5%

of isobutane (C_2H_{10}) and 0.3% of sulphur hexafluoride (SF_6).

3.2.6 CMS Trigger System and DAQ

The proton-proton and heavy-ion collision in LHC occur with high rates of interaction. For pp-collision, two bunch collide every 25 ns, which corresponds to 40 MHz frequency (equivalent to $\approx 10^9$ interactions per second). One of the main points about the pp interaction in the collision is that not all interactions are useful; only a few of them contain interesting physics, and most of the events come from soft interaction and are called minimum bias events. Secondly, it is very difficult to store and process such a large amount of data which is not feasible for the CERN main computing farm, which is limited. Therefore to filter out all such events, the CMS trigger system is used from where the event selection process starts; basically, the event will be stored when it will pass some criteria which are used to select an event of interest. The rate reduction is made in two steps, one is called Level-1 (L1) Trigger, and the other one is called High-Level Trigger (HLT).

3.2.6.1 Level-1 Trigger

The L1 trigger is a hardware trigger system with a fixed latency. The L1 trigger system has to decide whether the event should be accepted or rejected using the information from the subdetectors within $3\mu\text{s}$. Figure 3.16 shows the schematic of L1 trigger. The trigger information from the calorimeters is taken in two stages; stage 1 is called regional calorimeter trigger (RCT), and then comes stage 2, called global calorimeter trigger (GCT). For the L1 muon trigger, all three muon detec-

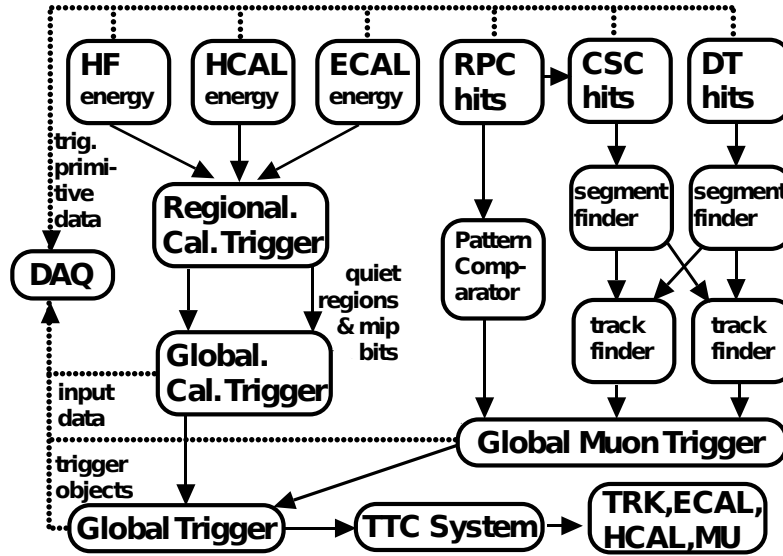


Figure 3.16: Schematics of the CMS L1 trigger system. Data from the calorimeters are processed regionally (RCT) and then globally (GCT). Hits from the RPCs, CSCs, and DT are processed either via a pattern comparator or via a system of the segment- and track-finders and sent onwards to a global muon trigger (GMT). The information from the GCT and GMT is combined in a global trigger (GT), which makes the final trigger decision. This decision is sent to the tracker (TRK), ECAL, HCAL or muon systems (MU) via the trigger, timing and control (TTC) system. The data acquisition system (DAQ) reads data from various subsystems for offline storage [12].

tors participate; the information from RPCs, CSCs and DTs (track segments from the hit information) is used for the pattern recognition algorithm, which identifies the muon candidates and momentum of the muon from the bending in the magnetic field. This information is provided to CSCTF (CSC track finder) and DTTF (DT track finder), and the hit information from RPCs is directly sent to the pattern comparator trigger (PACT). So with the information from CSCTF, DTTF and PACT is sent to GMT (Global Muon Trigger). In the final step, the information from GMT and GCT is combined to form a Global Trigger (GT). A menu of L1 triggers, which is essentially a list of objects with specific selection requirements (electron, muons, jets, photons or taus), is implemented at the final step.

3.2.6.2 CMS High Level Trigger (HLT)

The event that passes the selections criteria via the LT trigger is then fed as an input to the High-Level Trigger. The HLT is a software-based triggering system with a CPU farm having arrays of multicore computers. The HLT reduces the event rate from 100kHz (L1 Trigger output) to 1kHz with a latency of 100ms. The event selection procedure at the HLT is very similar to offline processing. The menu of triggers similar to the L1 trigger is made for HLT as well, and when an event passes the HLT, it gets stored permanently on the disc by data acquisition (DAQ) system for further analysis as shown in Figure 3.17.

At HLT, various workflows are written for various physics analysis, such workflows are called HLT paths, and full HLT menu consists of large list of these paths similar to L1 trigger menu. A simple structure of a typical HLT path is shown in

CHAPTER 3. EXPERIMENTAL SETUP

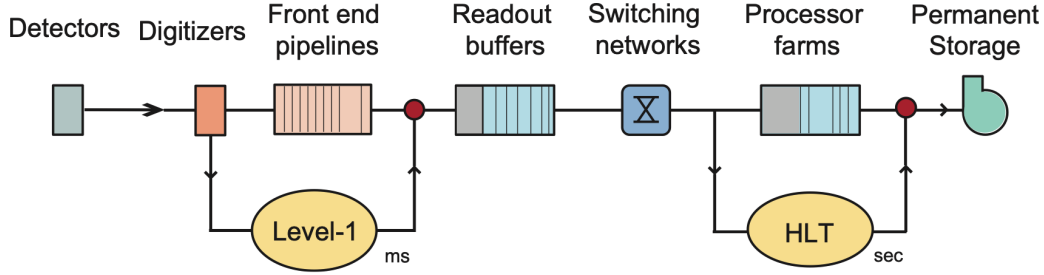


Figure 3.17: Schematic of CMS trigger and DAQ system

Figure 3.18 The **muons** in HLT are reconstructed in various steps; the first step is to get the input from the L1 trigger, and then starts the “Level-2 reconstruction”, based on Kalman-filtering. The output of the Level-2 trigger is fed to the Level-3 reconstruction step. In this step, the track parameters are constrained to interaction region in the silicon tracker in $\eta\phi$ plane, multiple trajectories are developed using Kalman-filtering, and then best-reconstructed trajectory (on the basis of χ^2 test) is chosen.

In the case of **electrons and photons**, the ECAL clusters corresponding to the L1 seed are reconstructed. The clusters are grouped in ϕ direction to get the Bremsstrahlung radiation; they are called Super-Clusters(SC). A threshold on the E_T of SCs is provided by the paths present in the HLT menu. Along with the SCs, ECAL isolation, energy reconstruction in HCAL, reconstruction of pixel hits globally for electrons, track reconstruction with Kalman-filtering and tracker isolation is performed to provide the electrons and photons for offline reconstruction.

The **jets** are reconstructed using iterative cone algorithm, generally anti- K_T algorithm with cone radius of $R = 0.4$ (AK4jets), however for fat jets, anti- K_T with

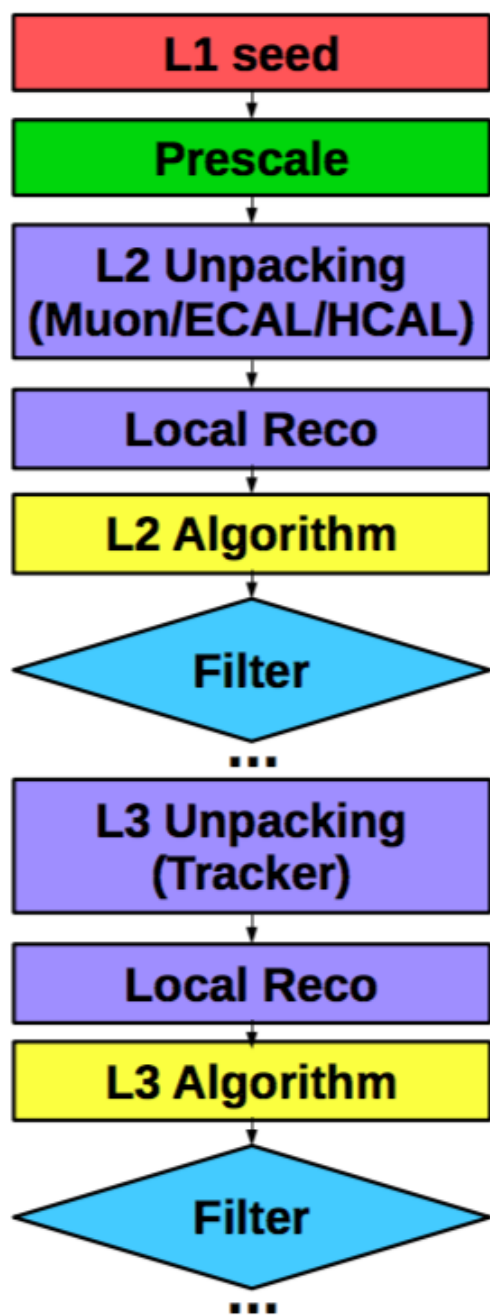


Figure 3.18: A typical HLT path in CMS

CHAPTER 3. EXPERIMENTAL SETUP

$R=0.8$ (AK8) or Cambridge-Aachen with $R = 1.5$ (CA15) is used depending upon the requirement in the HLT trigger paths. The main input for the jet reconstruction is the calorimeter towers constructed from HCAL cells and corresponding ECAL crystals; in general, the calorimeter towers must have $p_T > 0.5\text{GeV}$ and at least one tower having $p_T > 1\text{GeV}$

3.2.6.3 b-tagging at HLT

After the reconstruction of a jet, the next most important object reconstruction done at HLT is to identify whether a jet is originated from b-quark or not. The process of identifying the original parton flavour of the jet is Flavour tagging. The algorithms used for the flavour tagging use the properties of the hadrons that are present in the jets as a result of hadronization of b or c-quark. Few of the different properties of a b-quark are as follows:

- b-quarks hadronize into B-hadrons (B^+ , B^- , B^0 , B_s , Λ_B)
- b-quarks/B-hadrons have sizable lifetime
 - $c\tau \sim 500\mu\text{m} \rightarrow \beta\gamma c\tau \sim 5\text{mm} @ 50\text{ GeV}$
- The b quark and in turn B hadrons have a large mass ($\sim 5\text{ GeV}$)
- B hadron decay often contains leptons(20%)

With the large lifetime of hadrons containing b-quark results in large displacement gives rise to displaced tracks and with these displaced track a secondary vertex(SV) can be reconstructed as shown in Figure 3.19. The distance between primary vertex

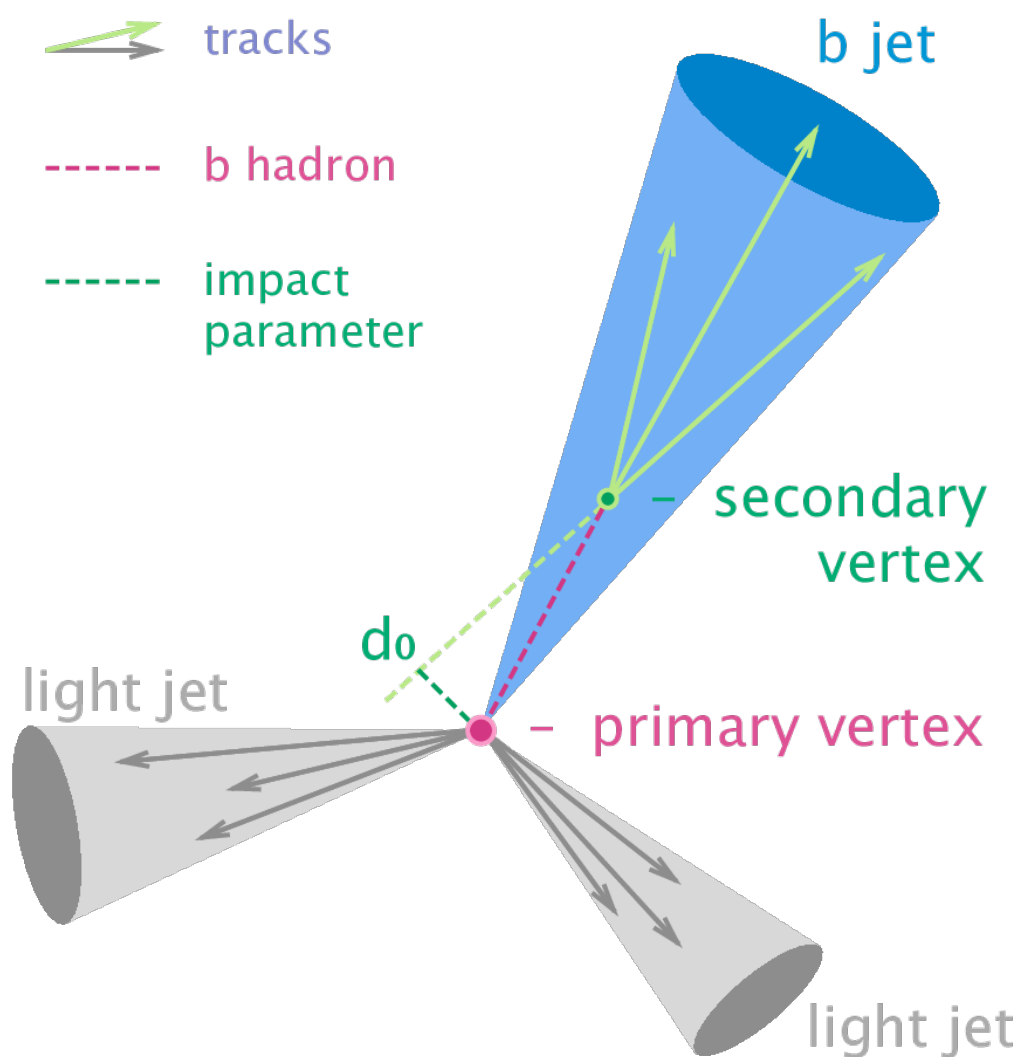


Figure 3.19: Jet with a secondary vertex (SV) from the b or c hadron decay giving displaced tracks and large impact parameter.

CHAPTER 3. EXPERIMENTAL SETUP

and tracks (point of closest approach) is called impact parameter (IP), the IP can be defined both 2 (transverse plane w.r.t beam line) and in 3 dimensions. Using all the information of secondary vertex, impact parameter, displaced tracks, and/or soft-lepton variables(the presence of electrons/muons) a jet identification algorithm has been designed.

During the Run1, jet probability (JP) [55] and combined secondary vertex(CSV) taggers were used. In the JP tagging method, the impact parameter significance of the tracks corresponding to the jet is used to obtain a likelihood that the jet originates from the primary vertex. A version of the jet probability algorithm, known as JBP, was also used where more weight is given to four tracks(with $\Delta R < 0.4$) with the highest impact parameter in the jet probability calculation. In Combined Secondary Vertex (CSV) [55], in this method, the information of the secondary vertex and the track-based lifetime is used in addition to other variables like IP to get an algorithm to provide a discriminator for heavy flavour and light flavoured jets. There are several other algorithms using this information to discriminate the heavy flavoured jets, for example, Simple Secondary Vertex (SSV).

However, for Run2, an improved version of CSV, called CSVv2 [13] is used; it is based on the CSV algorithm with the use of the multivariate technique. Due to the use of multivariate technique, the algorithm is trained on inclusive multijet events. Table 3.2 shows the variables used in CSV and CSVv2

3.2.6.3.1 Implementing DeepCSV at HLT During the 2016 data taking, the CSVv2 tagging algorithm was being used at the HLT level for the identification of the b-jets.

3.2. THE CMS DETECTOR

Input variable	CSV	CSVv2
Number of SV	x	✓
Corrected SV mass	✓	✓
SV energy ratio	✓	✓
SV 2D flight distance significance	✓	✓
$\Delta R(\text{SV}, \text{jet})$	x	✓
3D IP significance of the first four tracks	✓	✓
Number of tracks from SV	✓	✓
Track η_{rel}	x	✓
Track $p_{T,\text{rel}}$	x	✓
$\Delta R(\text{track}, \text{jet})$	x	✓
Track $p_{T,\text{rel}}$ ratio	x	✓
Track distance	x	✓
Track decay length	x	✓
Summed tracks E_T ratio	x	✓
$\Delta R(\text{summed tracks}, \text{jet})$	x	✓
First track 2D IP significance above c threshold	x	✓
Number of selected tracks	x	✓
Jet η	x	✓
Jet p_T	x	✓

Table 3.2: Input variable used for CSV and CSVv2 b-tagging algorithm. The symbol ✓(x) implies that the variable used(not used)

CHAPTER 3. EXPERIMENTAL SETUP

However, heavy-flavour jet identification can be further improved by using deep machine learning [56]. So, CSVv2 updated to a new version (at the offline level) in which, a deep neural network with more nodes per layer and more hidden layers is used. For the training of the DeepCSV algorithm, a mixture of multijet and $t\bar{t}$ events is used with the KERAS deep learning library which is interfaced with the TensorFlow in the backend. The output of deepCSV is used as a single powerful discriminator; the output is provided for b, c and light jets separately. Figure 3.20 shows the performance of the DeepCSV tagger as compared with other taggers in the form of a ROC curve.

To apply the DeepCSV at the HLT level, the first step is to choose an HLT path from the HLT menu, which has been used with CSVv2 algorithm:

- HLT_{PFHT380_SixPFJet32_DoublePFBTagCSV_2p2} :
 - This path contains minimum of six jets with $p_T > 32$ GeV that are identified by using particle flow algorithm.
 - Minimum jetHT (scalar sum of all jet momentum) of 380 GeV.
 - At-least two jets should pass the CSV discriminator cut of 0.75.

and then use this path as the basis and then replace the CSVv2 algorithm with DeepCSV in the HLT path.

- HLT_{PFHT380_SixPFJet32_DoublePFBTagDeepCSV_2p2}
 - This path have exactly same selection as the path with csv discriminator,

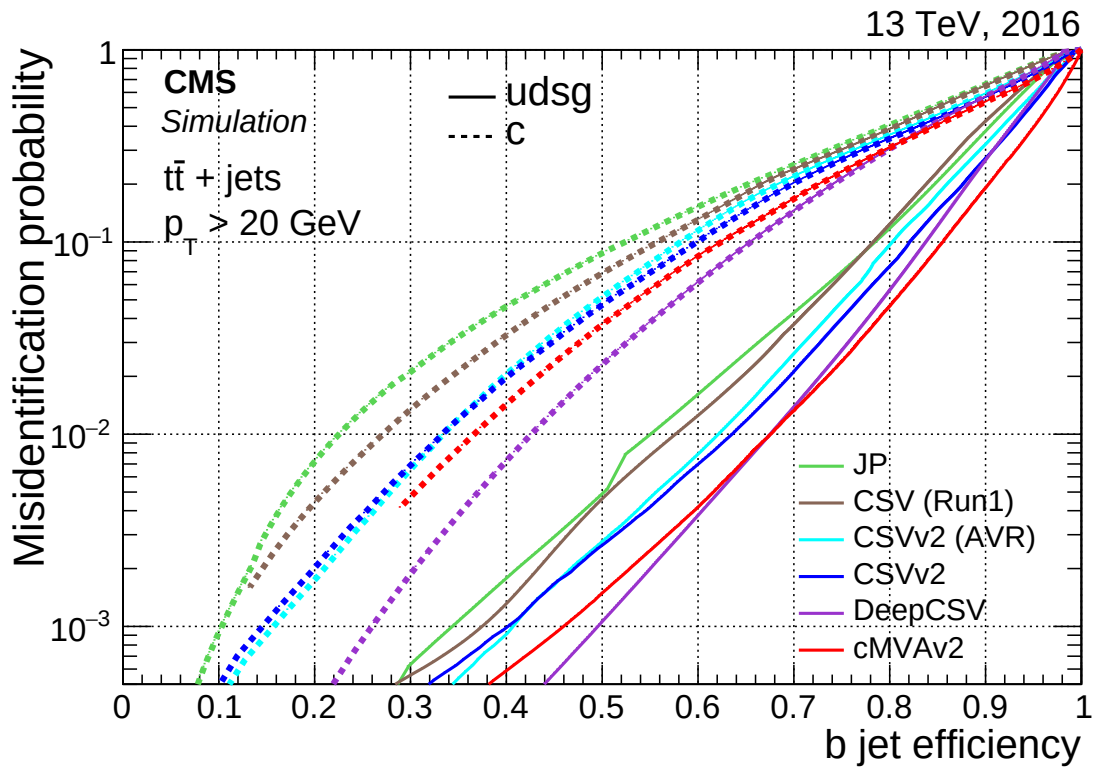


Figure 3.20: Misidentification probability for c and light-flavour jets versus b jet identification efficiency for various b tagging algorithms in $t\bar{t}$ events [13].

CHAPTER 3. EXPERIMENTAL SETUP

except out of all jets, at-least two jets should pass the deepCSV discriminator cut of 0.5

- The basic sequence Flow of the DeepCSV module is
 - CaloJet reconstruction
 - CaloJet Filter
 - Btag sequence for CaloJet
 - PFJet Reconstruction
 - PFJet Filter
 - Btag sequence for PFJet

The DeepCSV module used for offline reconstruction only works for particle flow jets, however, for online implementation, a template of DeepCSV was made in order to use DeepCSV for calo jet as well. The template was validated by plotting the scatter plots in Figure 3.21, where in Y-axis has the probability value from offline reconstruction and X-axis contains the probability obtained by the template made for HLT for particle flow jets. A similar exercise had been carried out for the calo-jets as shown in 3.22. No deviations from the straight line in the scatter plots were seen, which confirms the correct implementation of offline DeepCSV to the implementation at HLT. Once the DeepCSV module was available for both calo and pf jets, the DeepCSV path was constructed using ConfDB database. After the path has been made for the DeepCSV, the timing, rate and efficiency of the path are calculated:

3.2. THE CMS DETECTOR

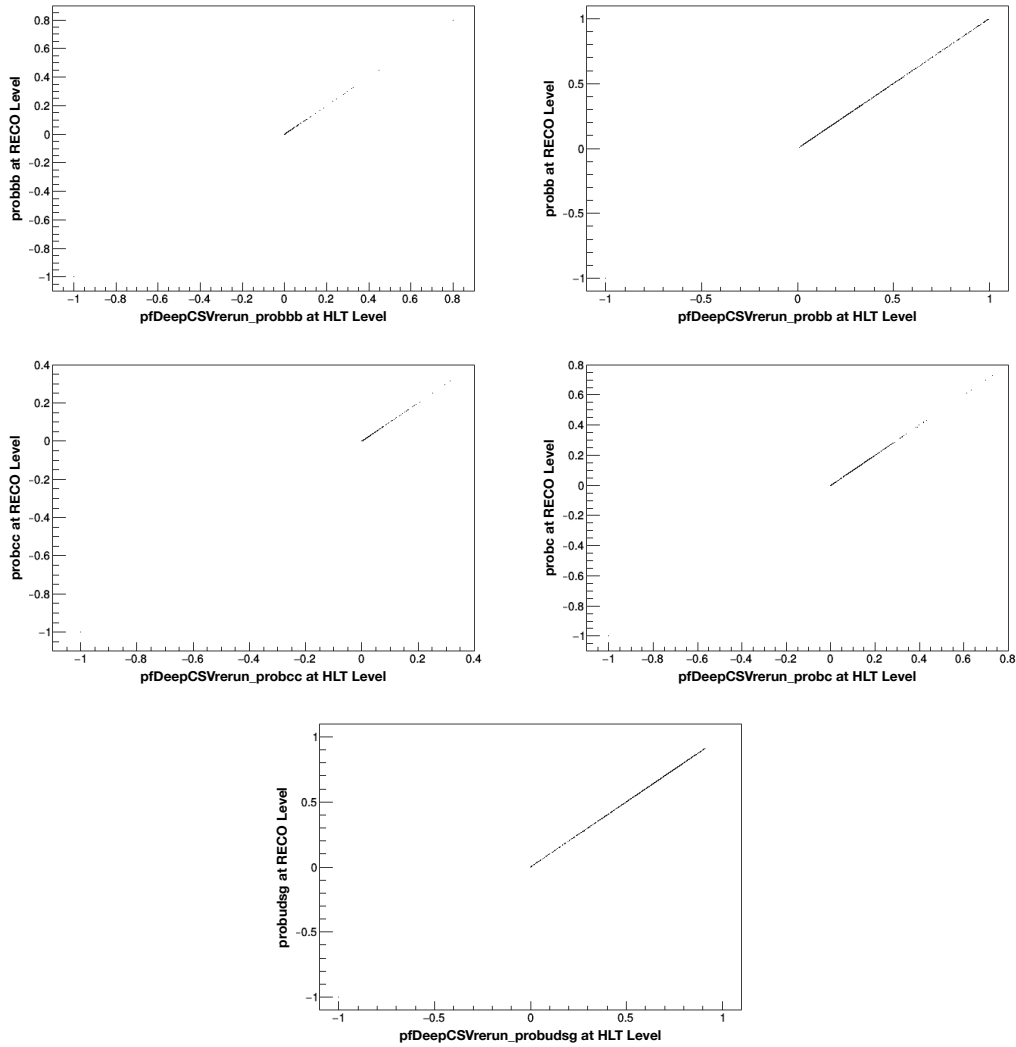


Figure 3.21: Scatter plot of the output of DeepCSV algorithm offline(RECO)(y-axis) and online(HLT)(x-axis) for particle flow jets. The probability for b-tagged jet is probbb and probbbb(plots in top row), similarly for c-tagged jets the probability is probcc and probccc(plots in middle row) and for light jets the probability is probudsg(plot in bottom row). The final discriminator value given for b-tag is $\text{probbb} + \text{probbb}$

CHAPTER 3. EXPERIMENTAL SETUP

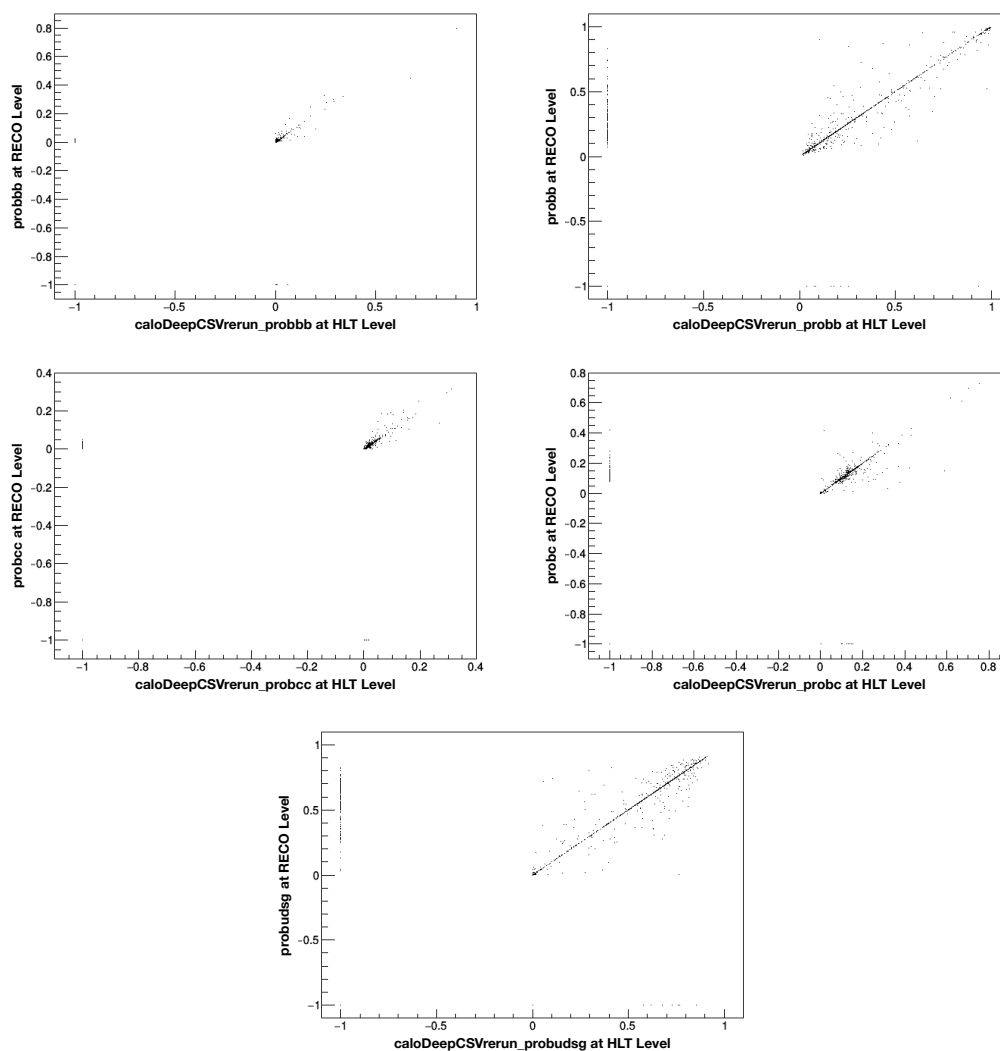


Figure 3.22: Scatter plot of the output of DeepCSV algorithm offline(RECO)(y-axis) and online(HLT)(x-axis) for calorimetric jets. The probability for b-tagged jet is probbb and probbbb(plots in top row), similarly for c-tagged jets the probability is probcc and probcc(plots in middle row) and for light jets the probability is probudsg(plot in bottom row). The final discriminator value given for b-tag is $\text{probbb} + \text{probbb}$

3.2. THE CMS DETECTOR

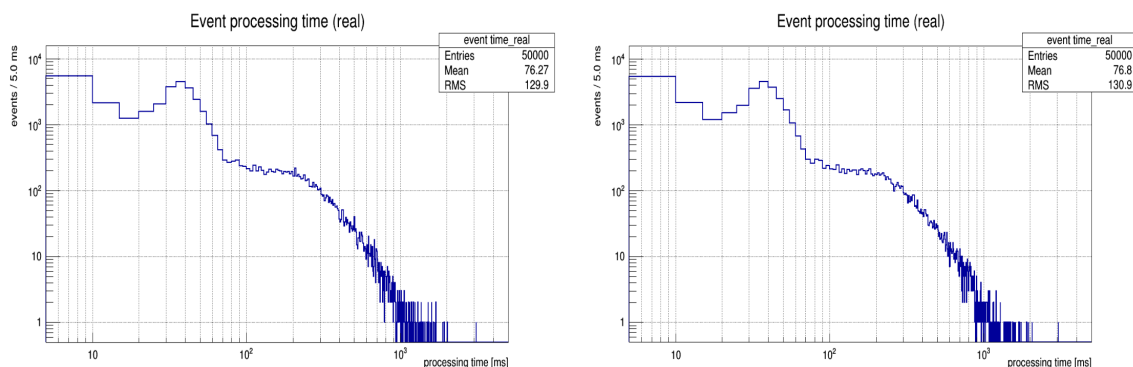


Figure 3.23: The timing study plot calculated for full HLT Menu without DeepCSV tagger (left) and with the DeepCSV tagger(right). The inclusion of HLTPFHT380_SixPFJet32_DoublePFBTagDeepCSV_2p2 path only increase the total time taken by the full HLT menu by 0.53ms.

- The path should not take more time which can cause problems in proper functioning of the HLT, as time is of at-most importance at HLT due the fixed bunch crossing time fo 25ns. Hence, monitoring and testing of the performance of the CPU time for the various HLT path is very important for the implementation of new triggers in HLT menu. The test for the timing measurement is done in specific benchmarked machines available in CMS server. The timing comparison of the HLT path with and without DeepCSV implementation has been shown in Figure 3.23
- The path should not have higher rate as the main purpose of the HLT is to reduce the rate of events so that only important and useful events can be saved for further analysis. To calculate the trigger rates, HLT developers have created a software that reports the rates of a selected list of triggers, primary

CHAPTER 3. EXPERIMENTAL SETUP

datasets, and streams, updated them and then compares them to the predicted rates. The rate of the DeepCSV path with

pileUp range = [54,58]

Luminosity = 2.0E+34

Using QCD & $t\bar{t}$ 83X MC Samples

=== RATES for **DeepCSV** ===

HLT_PFHT380_SixJet32_DoubleBTagDeepCSV_p051_v1 : 24.7 ± 1.2

=== RATES for **CSV** ===

HLT_PFHT380_SixJet32_DoubleBTagCSV_p075_v1 : 24.4 ± 2.2

=====

The HLT path was deployed online in mid of 2017 data taking.

In the following chapters the details about the use of dataset, simulated samples, reconstruction of various physics objects and there selection in the context of the $bb+p_T^{\text{miss}}$ analysis, and a analysis strategy will be discussed.

CHAPTER 4

DATA, SIMULATED SAMPLES AND TRIGGER

4.1 Data and Simulated samples

4.1.1 Data

The proton-proton collision data with $\sqrt{s} = 13$ TeV and 25 ns minimum bunch spacing collected in 2016, 2017, and 2018 have been used in this analysis. For signal and muon controls regions, the MET dataset is used, however for the electron control regions, SingleElectron (2016 and 2017) and EGamma (2018) datasets are used. Table [4.1](#), [4.2](#), and [4.3](#) summarize the datasets with an example for the MET dataset used in the analysis and the approximate luminosity as reported from the golden JSON using the `brilcalc` and the `normtag`.

- **Run2016B** to **Run2016H** eras of the CMS data-taking corresponding to an in-

CHAPTER 4. DATA, SIMULATED SAMPLES AND TRIGGER

Dataset	Luminosity(fb^{-1})
/MET/Run2016B-17Jul2018.ver2-v1	≈ 8.66
/MET/Run2016C-17Jul2018-v1	≈ 7.58
/MET/Run2016D-17Jul2018-v1	≈ 3.10
/MET/Run2016E-17Jul2018-v1	≈ 3.94
/MET/Run2016F-17Jul2018-v1	≈ 4.26
/MET/Run2016G-17Jul2018-v1	≈ 2.57
/MET/Run2016H-17Jul2018-v1	≈ 5.71
Total	35.82

Table 4.1: Datasets used in the analysis for 2016 with luminosity.

egrated luminosity of 35.9 fb^{-1} with good runs (lumis) selected using json file:

`Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt`.

The dataset has been processed using `94X_dataRun2_v10` global tag.

- **Run2017B to Run2017F** eras of the CMS data-taking corresponding to an integrated luminosity of 41.5 fb^{-1} with good runs (lumis) selected using json file:

`Cert_294927-306462_13TeV_PromptReco_Collisions17_JSON.txt`

The dataset has been processed using `94X_dataRun2_v11` global tag.

- **Run2018A to Run2018D** eras of the CMS data-taking corresponding to an integrated luminosity of 59.64 fb^{-1} with good (lumis) runs using json file:

`Cert_314472-325175_13TeV_PromptReco_Collisions18_JSON.txt`.

The dataset has been processed using `102X_dataRun2_Prompt_v15` global tag.

4.1. DATA AND SIMULATED SAMPLES

Dataset	Luminosity(fb^{-1})
/MET/Run2017B-31Mar2018-v1	≈ 4.79
/MET/Run2017C-31Mar2018-v1	≈ 9.75
/MET/Run2017D-31Mar2018-v1	≈ 4.30
/MET/Run2017E-31Mar2018-v1	≈ 9.36
/MET/Run2017F-31Mar2018-v1	≈ 13.30
Total	41.5

Table 4.2: Datasets used in the analysis for 2017 with luminosity

Dataset	Luminosity(fb^{-1})
/MET/Run2018A-17Sep2018-v1	≈ 13.9
/MET/Run2018B-17Sep2018-v1	≈ 7.06
/MET/Run2018C-17Sep2018-v1	≈ 6.88
/MET/Run2018D-PromptReco-v2	≈ 31.8
Total	59.64

Table 4.3: Datasets used in the analysis for 2018 with luminosity

4.1.2 Background Simulation

There are several standard model processes with much higher production cross-section produced in pp collisions which can mimic the experimental signature in the detector, i.e. one or two b-jets in the presence of high p_T^{miss} . These processes are the background for the signal searched for. The major background, which constitutes about 80% of the total background, include $Z(\nu\nu) + \text{jets}$, $t\bar{t}$. The number below is reported for the signal category with two bjets (one bjet).

$$Z(\nu\nu) + \text{jets}$$

is the main irreducible background for this analysis, where the Z boson decays to

CHAPTER 4. DATA, SIMULATED SAMPLES AND TRIGGER

two neutrinos, and in the final state, we see E_T^{miss} and jets/b-jets. Figure 4.1 shows the % contribution of each of the backgrounds.

$$Z(\nu\nu) + \text{jets}$$

constitute around 60% (70%) of the total background. Leading order (LO) simulation samples (in QCD using the MadGraph generator) in bins of hadronic transverse energy sum (H_T) are used to measure the pre-fit background contribution in signal region and control regions as described later in this note 5.1.2.

Another major background is $t\bar{t}$ events, where semi-leptonic decays of the $t\bar{t}$ can mimic the signal if lepton from the decay of W boson is lost. $t\bar{t}$ constitute around 17% (5%) of the total background. In addition, there is a small contribution from the hadronic decays of the single top events. The MC samples for $t\bar{t}$ are produced with the aMC@NLO and Powheg generators at Next-to-leading order (NLO), and Single-top events are generated with the Powheg generator at NLO.

W+jets process poses as a reducible background to the signal when lepton coming from the decay of the W boson is lost due to detector acceptance, reconstruction or identification criteria. This is the third-largest background and contributes less than 10% (21%) of the total background. The MC samples produced to have this background are at leading order (LO) in QCD using the MadGraph generator exclusively, i.e., in bins of hadronic transverse energy sum (H_T).

Pair of bosons (dibosons, VV) and QCD multijet collectively 6% (5.2%) of total backgrounds and referenced as others in the remaining part of the note. The MC

4.1. DATA AND SIMULATED SAMPLES

samples for VV are at leading order (LO) in QCD using the MadGraph generator. QCD events become the background of this analysis mostly due to the mismeasurement of the high p_T jets or if they go undetected by the detector. Although the QCD event with a final state and selection as in this analysis are very few, due to high cross-sections associated with QCD samples, we consider the MC samples for this background, which are produced at leading order (LO) in QCD using the MadGraph generator in bins of hadronic transverse energy sum (H_T).

The contribution of the major backgrounds $Z(\nu\nu) + \text{jets}$ and $t\bar{t}$ in 2b, and $Z(\nu\nu) + \text{jets}$ and $W(\ell\nu) + \text{jets}$ in 1b are estimated during the signal extraction process using the simultaneous fit of SR and dedicated control regions which are enriched in respective background. However, the remaining backgrounds are estimated using the simulation.

$Z(\ell\ell) + \text{jets}$ has similar physics as $Z(\nu\nu) + \text{jets}$. Therefore $Z(\ell\ell) + \text{jets}$ process is used to estimate the contribution of $Z(\nu\nu) + \text{jets}$. The MC samples produced to have this background are at leading order (LO) in QCD using the MadGraph generator exclusively, i.e., in bins of hadronic transverse energy sum (H_T).

For 2016: Monte Carlo samples were produced within `RunIISummer16miniAODv3` campaign and processed in `CMSSW_9_4_13` using the global tag `mcRun2_asymptotic_v3` conditions to match the 2016 data conditions. The list of sample and cross-section value for each of these samples are summarised in the Table 4.4.

For 2017: Monte Carlo samples were produced within `RunIIFall17miniAODv2` campaign and processed in `CMSSW_9_4_13` using the global tag `mc2017_realistic_v17` conditions to match the 2017 data conditions. The list of sample and cross-section

CHAPTER 4. DATA, SIMULATED SAMPLES AND TRIGGER

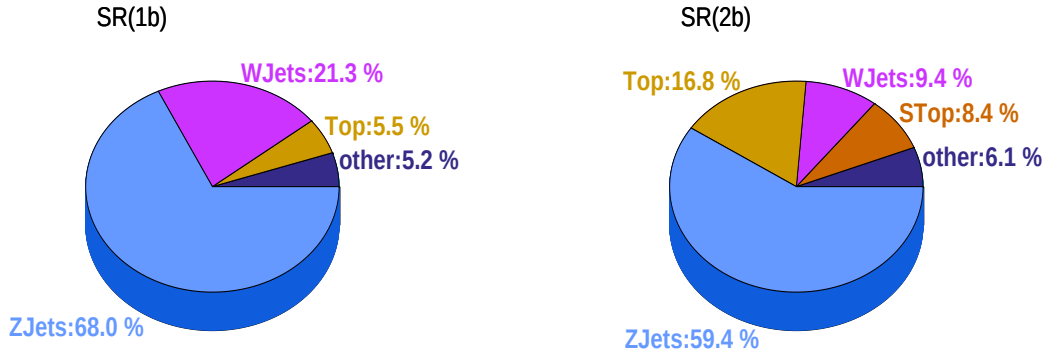


Figure 4.1: The pie chart for SR1 (left) and SR2 (right) to show the % contribution of each background.

value for each of these samples are summarised in the Table 4.4.

For 2018: Monte Carlo samples were produced within RunII Autumn18 miniAOD campaign and processed in CMSSW_10_2_18 using the global tag `upgrade2018-realistic_v20` conditions to match the 2018 data conditions. The list of sample and cross-section value for each of these samples are summarised in the Table 4.4.

4.1. DATA AND SIMULATED SAMPLES

Dataset Name	Run2 xsec(pb)	Dataset Name	Run2 xsec(pb)
DYJetsToLL_M-50_HT-100to200	161	ST_t-channel_antitop_4f_inclusiveDecays	80.95
DYJetsToLL_M-50_HT-200to400	48.58	ST_t-channel_top_4f_inclusiveDecays	136.02
DYJetsToLL_M-50_HT-400to600	6.982	ST_tW_antitop_5f_inclusiveDecays	38.06
DYJetsToLL_M-50_HT-600to800	1.747	ST_tW_top_5f_inclusiveDecays	38.09
DYJetsToLL_M-50_HT-800to1200	0.8052	TTTo2L2Nu	72.14
DYJetsToLL_M-50_HT-1200to2500	0.1927	TTToHadronic	314
DYJetsToLL_M-50_HT-2500toInf	0.003478	TTToSemiLeptonic	300.95
GJets_HT-40To100	18620	QCD_HT500to700	29990
GJets_HT-100To200	8622	QCD_HT700to1000	6351
GJets_HT-200To400	2193	QCD_HT1000to1500	1094
GJets_HT-400To600	258.5	QCD_HT1500to2000	98.99
GJets_HT-600ToInf	85.21	QCD_HT2000toInf	20.23
ST_s-channel_4f_leptonDecays	10.32	-	-

Table 4.4: Background Monte Carlo datasets used in the analysis for Run2

2016		2017		2018	
NLO Sample	Xsec(pb)	NLO Sample	Xsec(pb)	NLO Sample	Xsec(pb)
DYJetsToLL_Pt-100To250	81.42	DY1JetsToLL_M-50_LHEZpT_150-250	9.5	DYJetsToLL_Pt-0To50	106200
DYJetsToLL_Pt-250To400	2.991	DY1JetsToLL_M-50_LHEZpT_250-400	1.097	DYJetsToLL_Pt-50To100	409.8
DYJetsToLL_Pt-400To650	0.3882	DY1JetsToLL_M-50_LHEZpT_400-inf	0.1207	DYJetsToLL_Pt-100To250	97.26
DYJetsToLL_Pt-50To100	354.9	DY1JetsToLL_M-50_LHEZpT_50-150	315.1	DYJetsToLL_Pt-250To400	3.76
DYJetsToLL_Pt-650ToInf	0.03737	DY2JetsToLL_M-50_LHEZpT_150-250	15.73	DYJetsToLL_Pt-400To650	0.52
		DY2JetsToLL_M-50_LHEZpT_250-400	2.74	DYJetsToLL_Pt-650ToInf	0.05
		DY2JetsToLL_M-50_LHEZpT_400-inf	0.4492		
		DY2JetsToLL_M-50_LHEZpT_50-150	169		
DYJetsToNuNu_PtZ-100To250	164.73	Z1JetsToNuNu_M-50_LHEZpT_150-250	18.04	Z1JetsToNuNu_M-50_LHEZpT_150-250	17.97
DYJetsToNuNu_PtZ-250To400	6.219	Z1JetsToNuNu_M-50_LHEZpT_250-400	2.051	Z1JetsToNuNu_M-50_LHEZpT_250-400	2.053
DYJetsToNuNu_PtZ-400To650	0.8355	Z1JetsToNuNu_M-50_LHEZpT_400-inf	0.2251	Z1JetsToNuNu_M-50_LHEZpT_400-inf	0.224
DYJetsToNuNu_PtZ-50To100	709.5	Z1JetsToNuNu_M-50_LHEZpT_50-150	598.9	Z1JetsToNuNu_M-50_LHEZpT_50-150	598
DYJetsToNuNu_PtZ-650ToInf	0.07812	Z2JetsToNuNu_M-50_LHEZpT_400-inf	0.8472	Z2JetsToNuNu_M-50_LHEZpT_150-250	29.66
		Z2JetsToNuNu_M-50_LHEZpT_150-250	29.6	Z2JetsToNuNu_M-50_LHEZpT_250-400	5.214
		Z2JetsToNuNu_M-50_LHEZpT_250-400	5.174	Z2JetsToNuNu_M-50_LHEZpT_50-150	326
		Z2JetsToNuNu_M-50_LHEZpT_50-150	326.3		
WJetsToLNu_Pt-100To250	626.2	W1JetsToLNu_LHEWpT_100-150	284.06	WJetsToLNu_Pt-100To250	771.2
WJetsToLNu_Pt-250To400	21.83	W1JetsToLNu_LHEWpT_150-250	71.71	WJetsToLNu_Pt-250To400	28.06
WJetsToLNu_Pt-400To600	2.635	W1JetsToLNu_LHEWpT_250-400	8.06	WJetsToLNu_Pt-400To600	3.59
WJetsToLNu_Pt-600ToInf	0.4102	W1JetsToLNu_LHEWpT_400-inf	0.89	WJetsToLNu_Pt-50To100	3578
WJetsToLNu_TuneCUETP8M1	60380	W2JetsToLNu_LHEWpT_100-150	281.63	WJetsToLNu_Pt-600ToInf	0.55
		W2JetsToLNu_LHEWpT_150-250	108.59		
		W2JetsToLNu_LHEWpT_250-400	18.03		
		W2JetsToLNu_LHEWpT_400-inf	3.0		

Table 4.5: V+Jets Monte Carlo datasets(NLO) used in the analysis for 2016, 2017 and 2018

4.2 Triggers

The very first step of the event selection is to collect the signal-like events using a set of High Level Triggers (HLT). Different triggers are used for signal and control regions. The events in the signal region and control regions with 1 or 2 muons are required to pass the p_T^{miss} trigger whereas the events in the control regions with 1 or 2 electrons are required to pass the single electron triggers.

4.2.1 p_T^{miss} triggers

Table 4.6 summaries the p_T^{miss} triggers used in the analysis. In the trigger algorithm, objects identified as particle flow muons [57] are removed from the event before the p_T^{miss} objects are calculated. In this manner, using the same triggers, we can select di-muon (mainly $Z \rightarrow \mu^+ \mu^-$) and single muon (mainly $W \rightarrow \mu \nu$) events.

For Run2 the performance of the p_T^{miss} triggers is measured using and single muon ($W \rightarrow \mu \nu$ decays) and double muon ($Z \rightarrow \mu^+ \mu^-$ decays) events. In particular, the single muon events are selected from the SingleMuon dataset using single muon trigger and one muon is required to be loose and have p_T larger than 20 GeV. A central jet with p_T larger than 100 GeV is required in addition to the lepton

	2016	2017	2018
p_T^{miss} Trigger	HLT_PFMETNoMu120_PFMHT NoMu120_IDTight_v OR HLT_PFMET170_HBHECleaned_v	HLT_PFMETNoMu120_PFMHT NoMu120_IDTight_PFHT60_v OR HLT_PFMETNoMu120_PFMHT NoMu120_IDTight_v OR HLT_PFMETNoMu140_PFMHT NoMu140_IDTight_v	HLT_PFMETNoMu120_PFMHT NoMu120_IDTight_PFHT60_v OR HLT_PFMETNoMu120_PFMHT NoMu120_IDTight_v OR HLT_PFMETNoMu140_PFMHT NoMu140_IDTight_v

Table 4.6: MET Trigger for all the three years

and the trigger efficiency is computed with respect to the hadronic recoil.

The scale factors are measured as a function of hadronic recoil using $W \rightarrow \mu\nu$ and $Z \rightarrow \mu\mu$ control samples. The control sample for $W \rightarrow \mu\nu$ is extracted using following selection:

- HLT IsoMu27(24)
- Exactly one loose(tight) μ with $p_T > 20$ and $|\eta| < 2.4$
- Leading Ak4Jet $p_T > 50\text{GeV}$
- veto loose electron with $p_T > 10$ and $|\eta| < 2.5$

However the control sample for $Z \rightarrow \mu\mu$ is formed using similar selection except that in $Z \rightarrow \mu\mu$ region two muons are required, the second muon is required to pass `looseID`. We can name the selection used to construct the control samples as *preselection*. The events in the control sample just constructed are required to pass the set of p_T^{miss} triggers used in the analysis. The efficiency in the respective region is then measure as a function of the hadronic recoil using the definition:

$$\text{Efficiency} = \frac{\text{logical OR of } p_T^{\text{miss}} \text{ triggers \&\& } preselection}{preselection} \quad (4.1)$$

The efficiency of the p_T^{miss} is measured in $Z \rightarrow \mu\mu$ and $W \rightarrow \mu\nu$ events using data and simulation events, as shown in Figure 4.3. The ratio of efficiency measured in data and simulation is shown in Figure 4.2. It shows that is good agreement in the scale factors measured in the $W \rightarrow \mu\nu$ and $Z \rightarrow \mu\mu$ regions. The difference

CHAPTER 4. DATA, SIMULATED SAMPLES AND TRIGGER

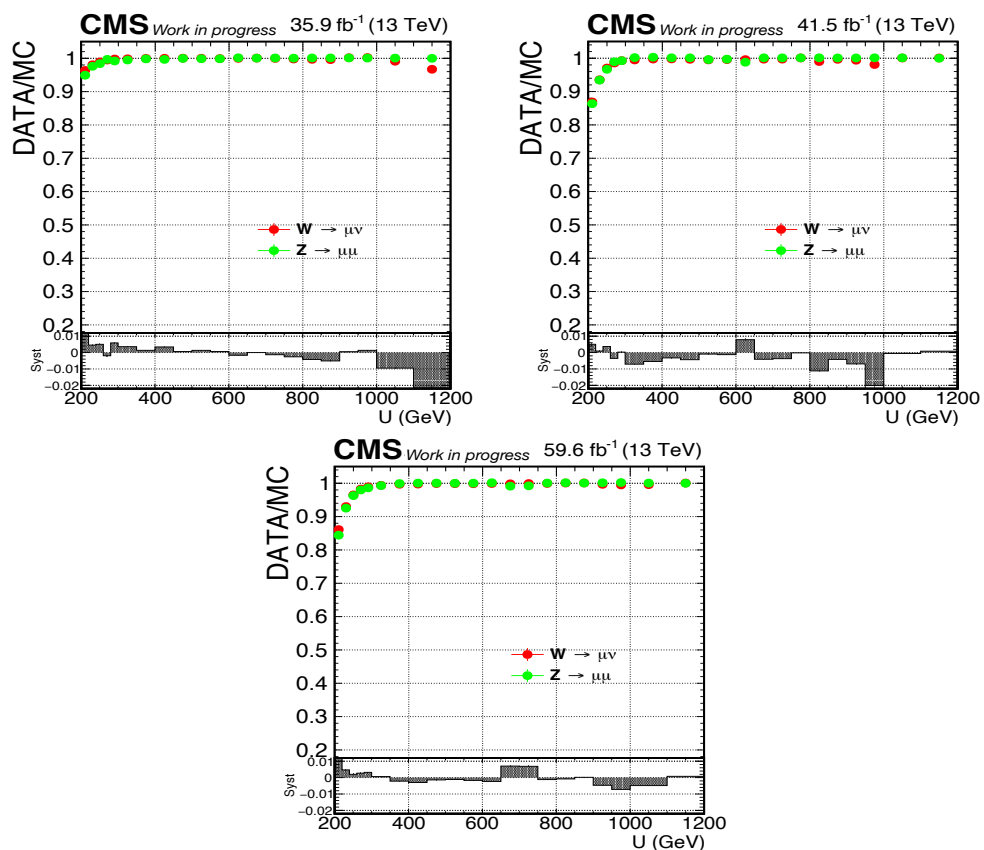


Figure 4.2: p_T^{miss} trigger efficiency scale factors for 2016 (left), 2017 (right), and 2018 (bottom)

in the measurement due to different selection used here is taken as the systematic uncertainty for the p_T^{miss} triggers.

4.2. TRIGGERS

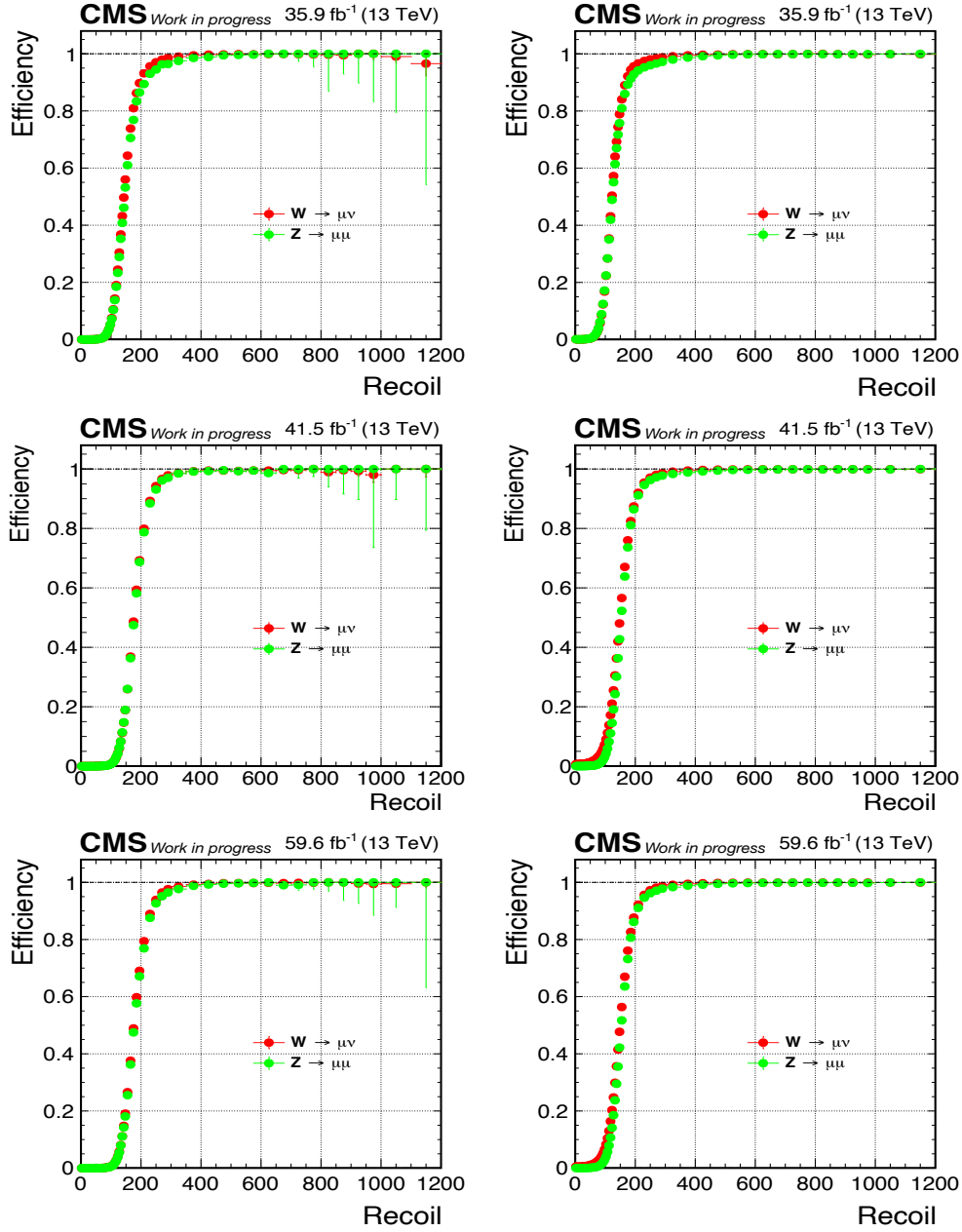


Figure 4.3: Trigger turnON for p_T^{miss} trigger for 2016 (top), 2017(middle) and 2018 (bottom). The distribution for data is shown in the left and simulation is shown on the right.

4.2.2 Single electron triggers

The events in the regions with 1 or 2 electrons use the single electron triggers as the first step of the analysis selection. Table 4.7 summarize the single electron triggers used in the analysis for each year. For each data taking period 2016, 2017 and 2018 the efficiency scale factors for single-electron trigger measured using tag & probe method [58] is either provided by the EGamma POG (2016) or blessed and recommended by the EGamma POG (2017 and 2018). [59]. Figure 4.4 shows the trigger scale factors used for each year. For 2017, `HLT_Ele32_WPTight_Gsf_v*` only existed after run 302026, start of Run2017D. However, `HLT_Ele32_WPTight_Gsf_L1DoubleEG` trigger path is seeded by the normal single L1 EG seeds and the double EG seeds hence is a strict super-set of `HLT_Ele32_WPTight_Gsf_v*`. So to turn `HLT_Ele32_WPTight_Gsf_L1DoubleEG` trigger into a `HLT_Ele32_WPTight_Gsf_v` trigger, we require the trigger object passing `hltEle32L1DoubleEGWPTightGsfTrackIsoFilter` to also pass `hltEGL1SingleEGOrFilter`, because `hltEle32L1DoubleEGWPTightGsfTrackIsoFilter` is the last filter of `HLT_Ele32_WPTight_Gsf_L1DoubleEG` and after making a trigger object `hltEGL1SingleEGOrFilter` filter as well, we directly get the trigger object that passes `HLT_Ele32_WPTight_Gsf_v`. Single electron triggers are also used to select dielectron events to mitigate any inefficiencies of the dielectron triggers.

Single Electron triggers	2016	2017	2018
	<code>HLT_Ele27_WPTight_Gsf</code>	<code>HLT_Ele32_WPTight_Gsf_L1DoubleEG</code>	<code>HLT_Ele32_WPTight_Gsf</code>

Table 4.7: Single Electron trigger used in the analysis

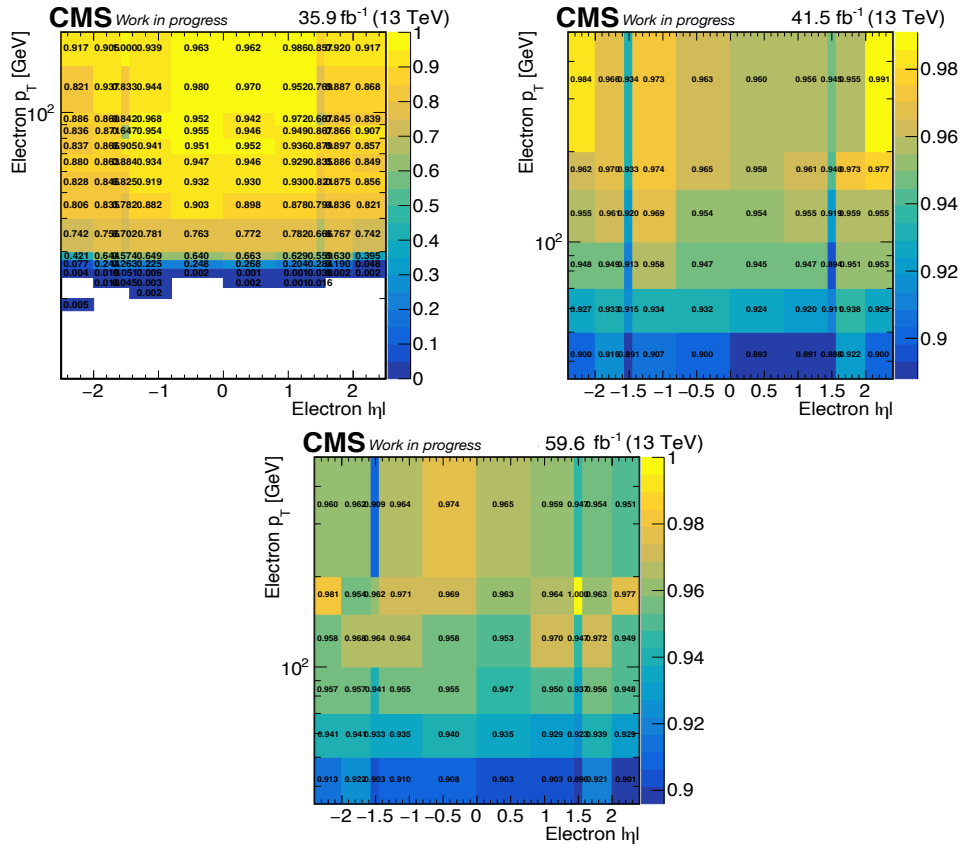


Figure 4.4: Single electron trigger efficiency scale factors for 2016 (top left), 2017 (top right), and 2018 (bottom)

CHAPTER 5

SIMULATION AND OBJECT RECONSTRUCTION

5.1 Event Reconstruction and Physics Object Selection

5.1.1 Reconstruction

In CMS, a global event reconstruction is performed by the particle-flow algorithm [60, 61], which reconstructs and identifies individual particles produced in each collision with an optimized combination of information from the various elements of the CMS detector.

Electrons are reconstructed in the tracker acceptance region $|\eta| < 2.5$ by matching the energy deposits in the ECAL with tracks reconstructed in the inner tracker. The electron identification relies on a cut-based algorithm, which defines a set of

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

selections based on observables sensitive to the shape and energy deposit along the electron trajectory, the direction and momentum of the track in the inner tracker, and the compatibility with the primary vertex of the event. In addition, specific cuts are applied to remove electrons produced by photon conversions.

Muons are reconstructed within the acceptance of the CMS muon systems $|\eta| < 2.4$ by the information from both the muon spectrometer and the silicon tracker. Muon candidates are identified by selection criteria based on the compatibility of the track reconstructed using the silicon tracker only and the combination of the hits in both the tracker and muon spectrometer. Additional requirements are based on the compatibility of the trajectory with the primary vertex and the number of hits observed in the tracker and muon systems. Further, muons and electrons are required to be isolated. The isolation is defined as the scalar sum of the p_T of all the particle-flow candidates (excluding the lepton itself) in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the lepton direction, after removing the contribution from pileup. The cone radius parameter is chosen to be 0.3 for electrons and 0.4 for muons.

Photons are reconstructed as ECAL energy clusters in the detector. The energy of the candidate photons is determined from the sum of signals in the ECAL crystals corrected to recover the energy deposited by bremsstrahlung and photon conversions. Photons are identified with cut-based criteria to distinguish real photons from jets and electrons, based on information including the isolation and transverse shape of the ECAL deposit and the ratio of the energy in the HCAL towers behind the supercluster to the electromagnetic energy in the supercluster.

Hadronically decaying tau leptons are reconstructed from the particle-flow candi-

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

dates as composite objects using the hadron plus strips (HPS) algorithm [62] combining one or three charged pions with up to two neutral pions, the latter also reconstructed by the particle-flow algorithm from the photons arising from the $\pi^0 \rightarrow \gamma\gamma$ decay.

Jets are reconstructed from the clustering of the particle-flow candidates, both charged and neutral, utilizing the anti-kT algorithm [63] with a clustering parameter of 0.4. Additional pp interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified to be originating from pileup vertices are discarded (Charged Hadron Subtraction, CHS), and an offset correction is applied to correct for remaining contributions. Jet energy corrections, extracted from data in multijet, g+jets, and Z+jets events, are applied to both data and simulation to correct for differences in the detector response relative to simulated events [64]. All the jets are further identified using a standard criterion based on the fractions of energy associated with all the particle-flow candidates associated with the jet.

A b-tagging algorithm (DeepCSV) is used to identify jets that originate from the hadronization of bottom quarks. The algorithm combines information from the tracks associated with the primary and secondary (displaced) vertices [65, 66] into a deep neural network. Standard b-tagging working points, namely loose, medium, and tight, are provided by CMS BTV POG corresponding to a particular value of mistag rate.

The loose working point corresponds to 10% of mistag rate; the medium working

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

point corresponds to 1% of mistag rate, and a tight working point corresponds to 0.1% of mistag rate.

The missing transverse momentum vector $\vec{p}_T^{\text{miss}}$ is defined as the projection on the plane perpendicular to the beams of the negative vector sum of the momenta of all particles reconstructed by the particle-flow algorithm, and its magnitude is referred to as p_T^{miss} . Corrections for the p_T^{miss} response and resolution are derived in g+jets and Z+jets events and applied to simulated events [67]

5.1.2 Physics object selection

All the major physics objects, i.e. electron, muon, photon, jets, and p_T^{miss} have been used in this analysis to identify the signal-like event, to suppress backgrounds and to define control regions in data for background estimation. The basic workflow for identifying the physics objects and reconstructing an event is discussed in section 5.1.1, the details of selection requirements imposed on them are described in the following subsections.

5.1.2.1 Muons

The usage of light leptons (electron and muon) is twofold in the analysis: (i) to veto the additional leptons (ii) to reconstruct the W/Z boson. Generally, a tighter identification criterion is used to reconstruct the W/Z boson, and a looser selection is used when the lepton is vetoed. Muon objects are required to pass the $|\eta| < 2.4$ and $p_T > 10$ GeV (30 GeV) for loose (tight) muons. In order to identify genuine muons in the event we use following variables [68]

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

- Loose ID for Muon Identification
- Muon objects are reconstructed as a global or tracker muon
- Muon objects are reconstructed as a particle flow muon
- The relative isolation less than 0.25 (loose Isolation)

The isolation value is computed as the p_T sum of all particle-flow charged hadrons associated to the primary vertex, particle-flow neutral hadrons and particle-flow photons in a ΔR cone ($\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$) of 0.4 around the muon. The isolation value is corrected for pileup by applying the $\Delta\beta$ correction and is taken relative to the p_T of the muon as given in the following equation:

$$\text{Rel. Iso} = \frac{\sum p_T^{\text{charged hadrons}} + \max\left(\sum p_T^{\text{neutral hadrons}} + \sum p_T^{\text{photons}} - 0.5 \cdot \sum p_T^{\text{PU charged hadrons}}, 0\right)}{p_T^\mu} \quad (5.1)$$

Beyond charged and neutral hadrons and photons, an additional term is introduced to subtract the pileup contribution, called $\Delta\beta$ correction. Since charged hadrons coming from other vertices can be easily subtracted from the event and thus do not enter in the first term, the same procedure cannot be done for the neutral fraction. It is verified that considering the contribution of PU neutral deposits, about half of the charged ones is a good approximation, and this term is subtracted from the neutral component of the primary vertex, having the care to check that the total neutral contribution cannot be lower than zero. The cone radius parameter is

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

set to 0.4 for muons. For muons, a “tight” cut at 0.15 has been applied, following POGs suggestions.

Muons used in the control samples are required to pass tighter identification [69] criteria compared to those listed above for veto in the Signal region. These are as follows:

- Muon reconstructed as a Global Muon
- Muon reconstructed as Particle-Flow muon
- Normalized χ_2 of the global track less than 10
- At least one muon-chamber hit included in the global-muon track fit
- Muon segments in at least two muon stations
- Muon’s tracker track has transverse impact parameter $d_{xy} < 2$ mm
- The longitudinal distance of the tracker track w.r.t the primary vertex is $d_z < 5$ mm
- Number of pixel hits > 0
- Hits on at least 5 tracker layers

The efficiency of these selection requirements (both loose and tight) is measured separately in data and simulation using the tag-n-probe technique, which is centrally computed by the muon POG, and accordingly, scale factors provided by CMS MUON POG [70–72] are applied to Monte Carlo events.

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

5.1.2.2 Electrons

In order to identify genuine electrons in the event, the electron candidates are required to have

- $p_T > 10$ (30) GeV
- $|\eta| < 2.5$ (excluding with $\eta \in [1.4442, 1.566]$ to avoid gap)
- loose (tight) identification and isolation

The electron identification used in this analysis is based on the cut-based ID recommended by the EGamma POG [73]. These are listed in Table 5.1. The cut-based ID definitions already contain the isolation cuts; therefore, no additional Isolation cut is required. In the signal region, events with well-reconstructed electrons are rejected to suppress electroweak backgrounds. Further, electrons are also used in defining certain control regions for estimating the key backgrounds in a data-driven manner in the analysis. Electrons used in these control samples are subjected to tighter identification and isolation requirements compared to selection used for signal regions. These are given in Table 5.2. The selection for electrons in the electron channel control region is tighter as compared to the loose electron selection for the signal region not only in terms of electron identification but also by impact parameter cuts for electrons. The impact parameter cuts include $|d_0|$ (in cm) < 0.05 (0.10) in barrel (endcap) and $|dz|$ (in cm) < 0.10 (0.20) in barrel (endcap). The data to simulation scale factors for Loose and Tight electron ID are used, which are recommended by the EGAMMA POG [74].

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

Variable	Selection (Barrel)	Selection (Endcap)
Full 5x5 $\sigma_{i\eta i\eta}$	< 0.0112	< 0.0425
$ \Delta\eta_{seed} $	< 0.00377	< 0.00674
$ \Delta\phi_{in} $	< 0.0884	< 0.169
H/E	$< 0.05 + 1.16/E_{SC} + 0.0324\rho/E_{SC}$	$< 0.0441 + 2.54/E_{SC} + 0.183 * \rho/E_{SC}$
Relative isolation (ρ correction)	$< 0.112 + 0.506/p_T$	$0.108 + 0.963/p_T$
1/E - 1/p	< 0.193	< 0.111
Expected Missing Inner Hits	≤ 1	≤ 1
Pass Conversion Veto	Yes	Yes

Table 5.1: Loose electron identification criteria for Electrons

Variable	Selection (Barrel)	Selection (Endcap)
Full 5x5 $\sigma_{i\eta i\eta}$	< 0.0104	< 0.0353
$ \Delta\eta_{seed} $	< 0.00255	< 0.00501
$ \Delta\phi_{in} $	< 0.022	< 0.0236
H/E	$< 0.026 + 1.15/E_{SC} + 0.0324\rho/E_{SC}$	$< 0.0188 + 2.06/E_{SC} + 0.183 * \rho/E_{SC}$
Relative isolation (ρ correction)	$< 0.0287 + 0.506/p_T$	$0.0445 + 0.963/p_T$
1/E - 1/p	< 0.159	< 0.0197
Expected Missing Inner Hits	≤ 1	≤ 1
Pass Conversion Veto	Yes	Yes

Table 5.2: Tight electron identification criteria for Electrons

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

5.1.2.3 Photons

A photon veto is applied in the analysis for the signal as well as the control regions to suppress photon induced backgrounds ($Z(\nu\nu) + \gamma\text{+jet}$, $W(\ell\nu) + \gamma\text{+jet}$). Events having one or more photons with

- $p_T > 15 \text{ GeV}$
- $|\eta| < 2.5$
- loose identification and isolation

are rejected. The identification criteria used for photon is cut-based photon Identification as recommended by EGamma POG [73] for Run2 analyses. The selection cuts used for the loose identification with efficiency of 90.08% in Barrel and 90.65% in endcap is given Table 5.3.

The isolation value is computed in a ΔR cone of 0.3 and is corrected for pileup using the ρ -correction as in the case of electrons. Photons are further cleaned against loose leptons and jets with $\Delta R = 0.4$.

	Selection (Barrel)	Selection (Endcap)
HoverE	< 0.04596	< 0.0590
$\sigma_{i\eta i\eta}$	< 0.0106	< 0.0272
Rho corrected PF charged hadron isolation	< 1.694	< 2.089
Rho corrected PF neutral hadron isolation	$< 24.032 + 0.01512 * p_T + 2.259e - 05 * p_T^2$	$< 19.722 + 0.0117 * p_T + 2.3e - 05 * p_T^2$
Rho corrected PF photon isolation	$< 2.876 + 0.004017 * p_T$	$< 4.162 + 0.0037 * p_T$

Table 5.3: Loose electron identification criteria for Photons

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

5.1.2.4 AK4 Jets

The AK4 jets considered in the analysis are required to have p_T larger than 30 GeV and $|\eta| < 2.5$. The jets are subjected to the loose jet identification requirements for 2016 analysis [75], the selection cuts are given in Table 5.4, however for 2017 and 2018 tight jet identification are used as recommended by JETMET POG [76, 77], the selection cuts are given in Table 5.5. AK4 jets are cleaned using ΔR cut against electron and muon in the analysis by rejecting a jet if $\Delta R(\text{jet}, \text{lep}) < 0.4$.

5.1.2.4.1 HEM 15/16 Failure During a large part of data-taking in 2018, one of the HCAL submodules stopped functioning, which resulted in no record of HCAL deposits in the region of $-3.0 < \eta < -1.3$ and $-1.57 < \phi < -0.87$ for the affected portion of the data set. The main consequence of the missing HCAL deposits for this analysis is that the transverse momentum of the jets in the region is typically under-measured, which results in over-measured E_T^{miss} . To mitigate the effect of the HEM 15/16 failure on this analysis, the events containing the jets (which are not cleaned w.r.t leptons with ΔR) in the region of detector where $\eta \in (-3.0, -1.3)$ and $\phi \in (-1.57, -0.87)$ have been vetoed.

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

	For $-2.7 \leq \eta \leq 2.7$	For $2.7 < \eta \leq 3.0$	For $ \eta > 3.0$
PF Jet ID	Loose		
Neutral Hadron Fraction	<0.99	<0.98	-
Neutral EM Fraction	<0.99	>0.01	<0.90
Number of Constituents	1	-	-
Number of Neutral Particles	-	>2	>10
And for $-2.4 \leq \eta \leq 2.4$ in addition apply			
Charged Hadron Fraction	>0	-	-
Charged Multiplicity	>0	-	-
Charged EM Fraction	<0.99	-	-

Table 5.4: Loose Jet ID criteria for 2016

	For $ \eta \leq 2.7$		For $2.7 < \eta \leq 3.0$	For $ \eta > 3.0$
	$ \eta \leq 2.6$	$2.6 \leq \eta \leq 2.7$ (Only for 2018)		
PF Jet ID	Tight			
Neutral Hadron Fraction	<0.90		-	>0.02
Neutral EM Fraction	<0.90	<0.99	>0.02 & <0.99	<0.90
Number of Constituents	>1		-	-
Number of Neutral Particles	-		>2	>10
And for $-2.4 \leq \eta \leq 2.4$ in addition apply				
Charged Hadron Fraction	>0		-	-
Charged Multiplicity	>0		-	-
Charged EM Fraction	-		-	-

Table 5.5: Tight Jet ID criteria for 2017 and 2018

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

5.1.2.5 Taus

The presence of hadronically-decaying tau only acts as a veto for the events both in the signal and in the control regions to suppress electroweak backgrounds. The selection criteria for taus is

- $p_T > 20 \text{ GeV}$
- $|\eta| < 2.3$

However, for tau hadronic decay, the reconstruction associates reconstructed hadronic tau decay to a decay mode.

- `decayModeFinding`

Which has been applied in this analysis. Furthermore, for tau identification, various discriminators have been provided by the CMS TAU POG [78], which can directly be accessed from the MiniAOD; in this analysis, we have used

- `disc_againstMuonTight3`
- `disc_againstElectronTightMVA6`

Further, The loose isolation for taus has been applied by using `disc_byLooseIsolation-MVArun2017v2DBoldDMwLT2017`.

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

5.1.2.6 b-tagged Jets

In this analysis, b-jets are used, which pass a medium working point of the deepCSV b-tagging algorithm. The choice of working point has been optimized based on the signal significance. Based on the number of medium b-tagged jets in the event, we have two signal categories (??), SR1 with one bjet and SR2 second with two b jets. In identifying b-jets, we lower the p_T threshold on the jets to 30 GeV and require that they be in the tracker region ($|\eta| < 2.4$). It is known that b-tagging efficiency is not the same in data and MC, so in order to take into account this, the BTV POG [79] provides collections of b-tagging scale factors for b-jets and mistagged light jets, measured for different physics processes, for the supported tagging algorithms and the three standard working points as discussed above. A weight is calculated on a per-event basis as a the function of the b-tagging status of the jets in the event [80]. The probability of a given configuration of jets in MC simulation and data is defined as:

$$\begin{aligned} P(MC) &= \prod_{i=\text{tagged}} \epsilon_i \prod_{j=\text{not tagged}} (1 - \epsilon_j) \\ P(DATA) &= \prod_{i=\text{tagged}} SF_i \epsilon_i \prod_{j=\text{not tagged}} (1 - SF_j \epsilon_j) \end{aligned} \tag{5.2}$$

where ϵ_i is the MC b-tagging efficiency, and SF_i and ϵ_i are the functions of the jet flavor, jet p_T , and jet η . The event weight is now calculated as

$$w = \frac{P(DATA)}{P(MC)} \tag{5.3}$$

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

After a few months of the smooth operation of DCDC converters, some converters started to fail once the luminosity of the accelerator was increased at the beginning of October 2017. However, we did not get any sizeable impact of the 2017 pixel issues on the MET. After correcting for the EE noise (the so-called E_T^{miss} v2 recipe), we find a decent agreement between data and MC both in regions enriched in fake E_T^{miss} and genuine MET. One of the main objects which might get affected by the pixel failure is b-tagging, in our analysis, we require one or two bjets in our signal and control region, hence we checked the change in the nBjet distribution for the region of our interest in a high bjet control region i.e. $t\bar{t}$. For the second bin, the difference is less than 3% which corresponds to 1b category of signal and control region and for the third bin, the difference is less than 8% which corresponds to 2b category of signal and control regions as shown in figure 1

5.1.2.7 Missing Transverse Energy

The use of missing transverse energy (p_T^{miss}) is central to the analyses presented here. All the corrections to the momenta of jets reconstructed in the event, i.e., the jet energy correction (JEC) is propagated to the p_T^{miss} , and the Type 1 corrected p_T^{miss} is used. In the control regions, hadronic recoil is used as a proxy to the p_T^{miss} in the signal region, where the hadronic recoil can be computed by adding the p_T of the lepton(s) to the p_T^{miss} computation.

$$\vec{U} = -(\vec{p}_T^{\text{miss}} + \sum \vec{p}_T^{\text{lep}}) \quad (5.4)$$

5.1. EVENT RECONSTRUCTION AND PHYSICS OBJECT SELECTION

2016	2017 and 2018
Flag_HBHENoiseFilter	Flag_HBHENoiseFilter
Flag_globalTightHalo2016Filter	Flag_globalSuperTightHalo2016Filter
Flag_eeBadScFilter	Flag_eeBadScFilter
Flag_goodVertices	Flag_goodVertices
Flag_EcalDeadCellTriggerPrimitiveFilter	Flag_EcalDeadCellTriggerPrimitiveFilter
Flag_BadPFMuonFilter	Flag_BadPFMuonFilter
Flag_HBHENoiseIsoFilter	Flag_HBHENoiseIsoFilter
	ecalBadCalibReducedMINIAODFilter

Table 5.6: MET Filter recommended by JetMET POG for Run II

In the case of $Z \rightarrow \nu\nu$ decays, $\vec{U} = \vec{p}_T^{\text{miss}}$. In this sense, the recoil in the control regions can be used to mimic the p_T^{miss} in $Z \rightarrow \nu\nu$ events that pass the signal selection.

Events in data can end up with a large p_T^{miss} due to detector noise, mis-reconstruction, and beam backgrounds. To tag these events several p_T^{miss} filters, as listed in Table 5.6, recommended by the JetMET POG [81] and have been used in the analysis for data.

For 2017 data, EE noise at high $|\eta|$ increased the fake p_T^{miss} tail significantly; hence a tool has been developed to implement a recipe to account for the change in the p_T^{miss} , it has been described in a TWiki [MissingETUncertaintyPrescription](#). After correcting for the EE noise (the so-called E_T^{miss} v2 recipe), we find a decent agreement between data and MC both in regions enriched in fake E_T^{miss} and genuine MET.

CHAPTER 5. SIMULATION AND OBJECT RECONSTRUCTION

During 2017 data taking, an enhancement of the instrumental p_T^{miss} tail in data compared to MC was observed both at the trigger level (L1 and HLT) and offline. The feature showed a strong dependency on the run era, and it was most pronounced in 2017F [82]. It was found that EE noise at high $|\eta|$ increased the fake p_T^{miss} tail in 2017 data. A technique [named recipe “v2”] was developed [83] to mitigate that effect in which we fully drop jets and unclustered PF candidates with $2.65 < |\eta| < 3.139$ and $p_T^{\text{raw}} < 50$ GeV in the calculation of p_T^{miss} . In the analysis presented here, we have compared the effect of this noise mitigation to the p_T^{miss} distribution in data. The analysis is blinded in the signal region. Therefore, we checked the distribution without any requirement on the number of b-jets, the number of jets and the number of leptons in the event 6.1.2.

Figure 5 and 7 in appendix shows the comparison of p_T^{miss} in SR1 and SR2-like phase space for each of the year of data taking in 2017. The plots compare the number of events selected with and without mitigation. The observed p_T^{miss} distribution after mitigation is softer w.r.t without mitigation and give rise to an overall reduction in statistics by 10%, i.e. the events which earlier have $p_T^{\text{miss}} > 250$ GeV now have $p_T^{\text{miss}} < 250$ GeV. Hence those events do not appear in the post-mitigation distribution. Figure 4 and 6 shows the comparison of p_T^{miss} in SR1 and SR2-like phase space for all the eras for 2017 data taking after the distribution is normalized to unity compare the shape of the p_T^{miss} distribution with and without mitigation.

5.2 Monte Carlo Reweighting

5.2.1 Primary Vertex and Pileup Reweighting

Several primary vertices are typically reconstructed in an event due to pileup. A loose set of selection criteria is applied to select well-reconstructed primary vertices, and the vertex with the highest sum of p_T of constituent tracks is chosen as the vertex corresponding to the hard interaction. The simulation of pileup conditions in the Monte Carlo samples does not match with the one observed in the data. The simulation samples are therefore corrected based on the number of vertices simulated and observed in the data. The number of vertices in data is measured using a minimum bias cross-section of 69.2 mb, and in the simulation, the number is taken from the *getTrueNumInteractions* variable in the miniAOD simulation. The ratio of these two templates then served as the event weights and was used to correct the number of pileup interactions in the event.

Figure 5.1 shows the effect of pileup reweighting in the $t\bar{t}(e\nu)$ control region for 2b category for 2018. The overall agreement in the peak region improved after reweighting is applied.

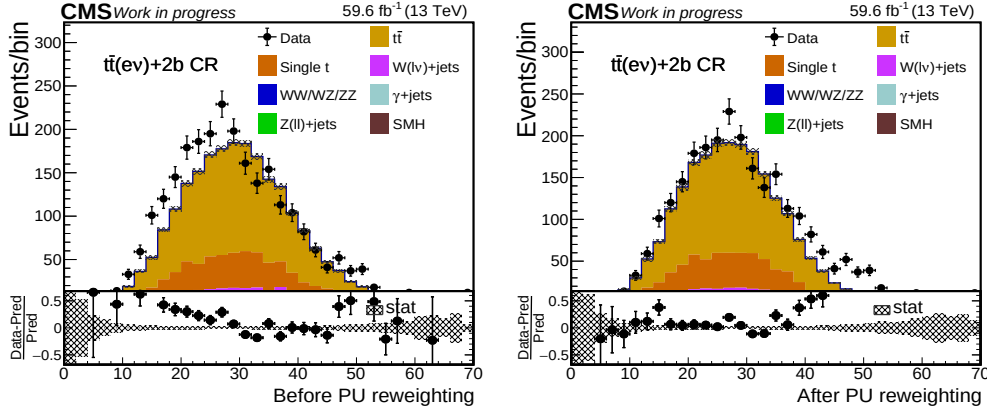


Figure 5.1: Before and after PU reweighting plots in 2b category of $t\bar{t}(ev)$ for 2018

5.2.2 L1 Prefiring in 2016 and 2017 data

In 2016 and 2017 data taking, the gradual timing shift of ECAL was not properly propagated to L1 trigger primitives (TP), which resulted in a significant fraction of high eta TP being mistakenly associated with the previous bunch crossing. Since Level 1 rules forbid two consecutive bunch crossings to fire, a consequence of this (in addition to not finding the TP in the bx 0) is that events can be self vetoed if a significant amount of ECAL energy is found in the region of $2 < |\eta| < 3$ [84]. This effect was not described by the MC; hence a recipe was developed to compute the probability for an event not to prefire, which can then be applied to the MC.

For the effect of the L1 prefire, the analysis is repeated by removing events with jets with $p_T > 100$ and $2.25 < |\eta| < 3.0$ and the effect on the final observable is found to be more than 1%/. The overall change in the yield is of the order of 2%-3% in the signal region as shown in Figure 11. Therefore for 2016 and 2017 simulation

5.2. MONTE CARLO REWEIGHTING

samples, a per-event prefiring weight is obtained and applied as the product of the non-prefiring probability of all objects.

This is achieved by running over all offline photons and jets found in the event and assign them a prefiring probability, the efficiency, for example, for jets is given as

$$\varepsilon_i^{pref} = \frac{N(\text{jets matched to L1IsoEG in bx-1})}{N(\text{all jets})} \quad (5.5)$$

The final event weight is obtained as the product of the non-prefiring probability of all objects (measured using unpreferable events), namely:

$$\omega = 1 - P(\text{prefiring}) = \prod_{i=\text{photons, jets}} (1 - \varepsilon_i^{pref}(\eta, p_T^{(EM)})) \quad (5.6)$$

The upward and downward values of the event weight(systematics) are obtained by shifting all prefiring probabilities plus or minus their uncertainties. The uncertainty is taken as the maximum between 20 percents of the object prefiring probability and the statistical uncertainty associated with the considered bin in the prefiring map.

CHAPTER 6

ANALYSIS STRATEGY AND EVENT SELECTION

6.1 Event Selection

The event selection for the signal and control regions are separated into two steps, (i) preselection (which is common for all the SR and CRs) and (ii) full selection (exclusive for each SR and CR). This section describes the details of preselection and full selection applied to construct the signal and control regions, which are later used to measure the major background using a simultaneous fit. The analysis considers two signal regions based on the number of b-tagged jets in the event, named as 1b and 2b categories (or SR1 SR2 interchangeably) as shown in Figure 6.1. Similarly, the CRs are constructed separately for 1b and 2b categories and are described below.

6.1.1 Preselection

The presence of one (two) AK4 jet(s) in the detector and a high p_T^{miss} is the typical experimental signature for the signal being probed in this physics analysis. The events are preselected using the set of selection criteria summarised in Table 6.1. The preselection of an event starts with the trigger selection. One of the trigger listed in the table 4.6 and 4.7 is required to pass in order to select the event. This ensures the events needed for SR, single electron, double electron, single muon, double muon control regions are saved for further analysis. An event is kept for further processing if it has either p_T^{miss} or hadronic recoil greater than 250 GeV. The hadronic recoil is calculated for one lepton and two lepton cases and is used to preselect the events. Further, to clean the data against fake p_T^{miss} tail events, the recommended p_T^{miss} filters are applied on data events only. The leading jet is required to have $p_T > 100$ GeV and within $|\eta| < 2.5$ and required to pass the identification criteria described in Table 5.5 and 5.4. Loose Jet ID is used for 2016 analysis, and Tight ID is used for 2017 and 2018 analysis. The subleading and remaining jets are required to have $p_T > 30$ GeV and identification criteria. All the jets considered in the events should not overlap with any other loose ID passed lepton, or hadronic tau or photon.

6.1. EVENT SELECTION

	Preselection
trigger	as described in Table 4.6 and 4.7
MET Filter	as described in Table 5.6
p_T^{miss} or Hadronic Recoil	$> 250 \text{ GeV}$
Jet	$p_T > 30 \text{ GeV} \ \& \ \eta < 2.5$
Leading Jet	$p_T > 100 \text{ GeV}$ and ID as described in Table 5.5 5.4
γ	veto

Table 6.1: Preselection applied to skim the events.

6.1.2 Signal region

Events in SR 2b category are required to have a maximum of three AK4 jets, out of which two of them are required to pass the medium WP b-tagging (deep csv algorithm). The leading jet is required to pass a $p_T > 100 \text{ GeV}$, and the subleading jet must have $p_T > 30 \text{ GeV}$ to enter the 2b signal category. Events that contain at least one reconstructed and identified isolated electron, muon, tau-lepton or photon candidates with $p_T > 10 \text{ GeV}$ (electrons and muons), $p_T > 15 \text{ GeV}$ (γ), $p_T > 18 \text{ GeV}$ (τ) are rejected, in order to minimize contaminations from background associated to leptons and MET, mostly $W \rightarrow l\nu$ with leptons reconstructed in the detector acceptance, and $t\bar{t}$ leptonic and semileptonic decays. The signal region is defined on top of the preselection cuts by requiring a clean final state with b-tagged jet(s) recoiling against the missing transverse momentum. Further, to remove events with abnormal MET because of reconstruction failures of particle flow algorithm, $\Delta(\text{PFE}_T^{\text{miss}}, \text{Calo}E_T^{\text{miss}}) = (\text{PFE}_T^{\text{miss}} - \text{Calo}E_T^{\text{miss}})/\text{Recoil} < 0.5$ is applied. The multijet background arising due to mis-measured jets or fake MET events are rejected using

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

	Category 1: SR1	Category 2: SR2
Trigger	Table 4.6	
E_T^{miss}	> 250	
Jet1	$p_T > 100 \text{ GeV}$	
Jet2	(if present) $p_T > 30 \text{ GeV}$	$p_T > 30 \text{ GeV}$
Jet3	veto	(if present) $p_T > 30 \text{ GeV}$
n b-tagged jets	1 (deepCSV MWP*)	2 (deepCSV MWP*)
Lepton (e/μ)	veto	
Hadronic τ	veto	
$\min(\Delta\phi) [\text{jet}, p_T^{\text{miss}}]$	> 0.5	
$(\text{PFE}_T^{\text{miss}} - \text{Calo}E_T^{\text{miss}})/\text{MET}$	< 0.5	

Table 6.2: Selection applied on each of the two categories in the signal region, SR1 and SR2

(*MWP = Medium Working Point)

$\min(\Delta\phi)(\text{jet}, p_T^{\text{miss}}) > 0.5$. Events with additional b-tagged jets are rejected to reduce the contribution of $t\bar{t}$.

Events in SR 1b category are required to have a maximum of two AK4 jets, and the leading jet is to pass the medium b-tagging selection and $p_T > 100 \text{ GeV}$. Additional lepton, photon, and hadronic τ vetos are places similar to the 2b category. The rest of the event selection remains the same in order to reject the same set of backgrounds. The event selection for 1b and 2b categories are summarized in Table 6.2.

Figure 6.6 and 6.7 shows the recoil, $\cos\theta^*$ and transverse momentum of the leading b-jet in the event for Signal Region of 2b category. Figure 6.5 and 6.7 shows the

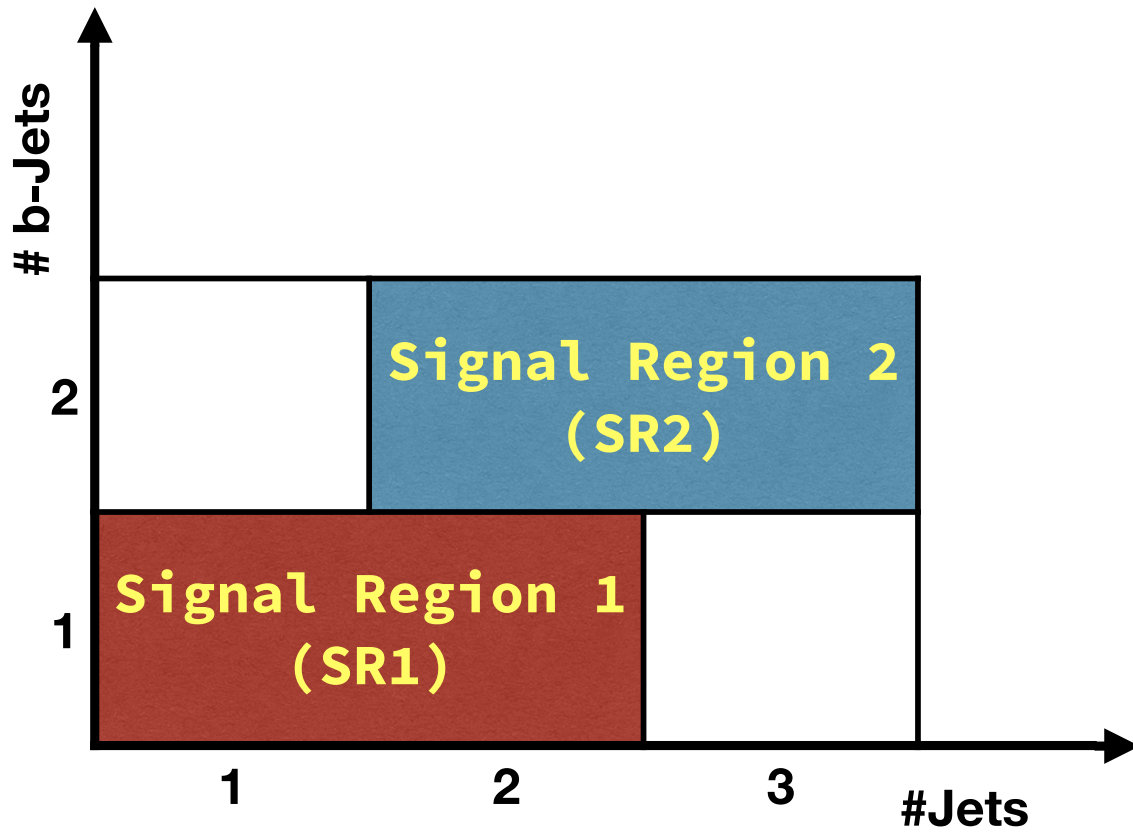


Figure 6.1: Figure showing the categories defined in the signal regions. The categories depend on the total number of jets and the number of b-tagged jets in an event.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

recoil and transverse momentum of the leading b-jet in the event for Signal Region of 1b category.

6.1.3 Control regions

Control regions are defined to estimate the major backgrounds, i.e., $Z + \text{jets}$, and $t\bar{t}$ for 2b category and $Z + \text{jets}$, and $W + \text{jets}$ for 1b category. The dominant $Z(\nu\nu) + \text{jets}$ is estimated by using the $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events. For the $t\bar{t}$ and $W + \text{jets}$ respective control is formed. The CRs are kept orthogonal to the SRs by using vetos in SR, mainly the number of jets, b-jets and leptons. Figure 6.2 shows the selection of the number of jets, b-jets and leptons in order to construct the respective control regions.

Four control regions, $Z(ee) + \text{jets}$, $Z(\mu\mu) + \text{jets}$, $t\bar{t}(e)$ and $t\bar{t}(\mu)$ are defined for 2b category to estimate the $Z(\nu\nu) + \text{jets}$ and $t\bar{t}$ in the SR via simultaneous fit of SR and CRs. However, since the major backgrounds in the 1b category are $Z(\nu\nu) + \text{jets}$ and $W + \text{jets}$, the control regions for $(Z(ee) + \text{jets})$, $(Z(\mu\mu) + \text{jets})$, $W(\mu\nu) + \text{jets}$ and $W(e\nu) + \text{jets}$ are constructed to constrain the normalization of the most important backgrounds. The detailed selection for the control regions is given in Table 6.3 and 6.4

The lepton veto is removed from the selection, allowing for the presence of one or two isolated leptons. In this case, the E_T^{miss} is computed by removing the momenta of the selected leptons (denoted as hadronic recoil).

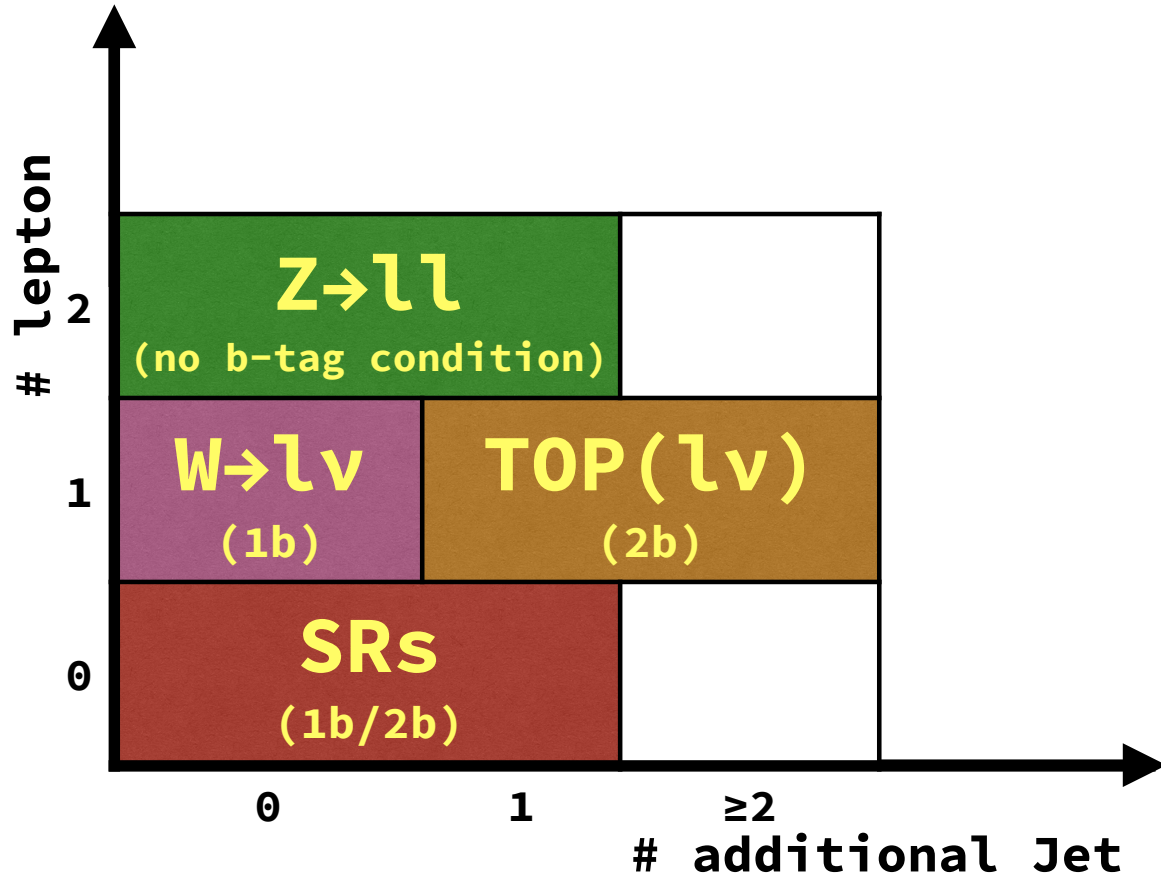


Figure 6.2: Figure showing the the control regions and the position of the signal regions in the electron-muon space. The categories depend on the number and flavour of leptons, and also the number of jets in an event.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

	Z($\mu\mu$)	Z(ee)	W($\mu\nu$)	W($\nu\nu$)
# Electron	0	2	0	1
# Muon	2	0	1	0
Lepton 1	$p_T > 30 \text{ GeV}$			
	Tight ID	Tight ID	Tight ID	Tight ID
	Iso < 0.15		Iso < 0.15	
Lepton 2	$p_T > 10 \text{ GeV}$	$p_T > 10 \text{ GeV}$	Veto	
	Loose ID	Loose ID		
	Iso < 0.25			
Recoil	$> 250 \text{ GeV}$			
	No Cut		$> 100 \text{ GeV}$	
Mass	$70 \text{ GeV} < \text{Reco Z Mass} < 110 \text{ GeV}$		$W \text{ trans Mass} < 160 \text{ GeV}$	
Additional jet	0 or 1		0	
$\frac{ (\text{PF}p_T^{\text{miss}} - \text{Calop}_T^{\text{miss}}) }{\text{Recoil}}$	< 0.5			

Table 6.3: Event selection for 1b Control regions

	Z($\mu\mu$)	Z(ee)	Top (μ)	Top (e)
# Electron	0	2	0	1
# Muon	2	0	1	0
Lepton 1	$p_T > 30 \text{ GeV}$			
	Tight ID	Tight ID	Tight ID	Tight ID
	Iso < 0.15		Iso < 0.15	
Lepton 2	$p_T > 10 \text{ GeV}$	$p_T > 10 \text{ GeV}$	Veto	
	Loose ID	Loose ID		
	Iso < 0.25			
Recoil	$> 250 \text{ GeV}$			
Mass	$70 \text{ GeV} < \text{Reco Z Mass} < 110 \text{ GeV}$		$W \text{ trans Mass} < 160 \text{ GeV}$	
Additional jet	0 or 1		≥ 1	
$\frac{ (\text{PF}p_T^{\text{miss}} - \text{Calo}p_T^{\text{miss}}) }{\text{Recoil}}$	< 0.5			

Table 6.4: Event selection for 2b Control regions

6.1.3.1 Z Control region

The main background of the analysis $Z \rightarrow \nu\nu$ is modelled using the $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ control regions. Z CRs are constructed using the selection similar to the signal region with a few exception listed in the table 6.3 and table 6.4 for 1b and 2b categories respectively. The change in the selection w.r.t the respective signal region is to make sure that the given control region is enriched in the $Z \rightarrow e^+e^-$ or $Z \rightarrow \mu^+\mu^-$ events. The $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ regions require two isolated electrons or muons with opposite charge; the dilepton invariant mass is required to lie in the $70 < m_{ll} < 110\text{GeV}$ window. The Z boson decay into a pair of muons/electrons and a pair of neutrinos are very similar from a theoretical point of view and hence becomes a natural process to be used for a good estimate of the major background. One of the major changes w.r.t the SR is removing the b-tagging on both the jets for 2b and one jet for 1b category to ensure high statistics in the control region.

Figure 6.10 and 6.11 shows the recoil, $\cos\theta^*$ and transverse momentum of the leading b-jet in the event for $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ of 2b category. Figure 6.8 and 6.9 shows the recoil and transverse momentum of the leading b-jet in the event for $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ of 1b category.

The distributions on the left are for the electron channel and muons on the right. The observables are shown for three different years, 2016 (top), 2017 (middle) and 2018 (bottom). From these figures, it can be noted that the prefit background estimate and shape of the distribution is well modelled by the simulation.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

6.1.3.2 W control region

The W control region formed only for the 1b category and is based on the same concept behind the Z control region. On top of the signal region selections, exactly one lepton (electron or muon) is required, and other additional leptons are vetoed. The hadronic recoil vector is recomputed by removing the lepton momentum. The transverse mass selection ($0 < m_T^W < 160$ GeV) is used to reject the multijet background. To avoid the effects of mis-modelling in the low E_T^{miss} region in this CR, $E_T^{\text{miss}} > 100$ GeV selection has been used. Further, in the W control region, no additional jet in the event is allowed.

Figure 6.12 and 6.13 shows the recoil and transverse momentum of the leading b-jet in the event for $W \rightarrow e\nu + \text{jets}$ and $W \rightarrow \mu\nu + \text{jets}$ of 1b category.

The distributions on the left are for the electron channel and muons on the right. The observables are shown for three different years, 2016 (top), 2017 (middle) and 2018 (bottom). From these figures, it can be noted that the prefit background estimate and shape of the distribution is well modelled by the simulation.

6.1.3.3 Top control region

$t\bar{t}$ control region includes events that satisfy the condition of presence of one lepton and any number of additional jets. This region is useful to obtain an almost-pure $t\bar{t}$ sample. The control region selection is required to be as close as possible to the signal region. Hence the same cuts applied to the missing transverse momentum in the signal region ($E_T^{\text{miss}} > 250$ GeV) are applied to the hadronic recoil evaluated in the

control regions (hadronic recoil $> 250\text{GeV}$). The same b-tagging selection is applied as in the SR, and additional leptons, τ -leptons, and photons are vetoed as defined in Section 5. Furthermore, control regions are split depending on the flavour of the leptons involved (e or μ) in the same way as in the signal regions.

Figure 6.14 and 6.15 shows the recoil, $\cos\theta^*$ and transverse momentum of the leading b-jet in the event for $t\bar{t}(e)$ and $t\bar{t}(\mu)$ of 2b category. The distributions on the left are for the electron channel and muons on the right. The observables are shown for three different years, 2016 (top), 2017 (middle) and 2018 (bottom). From these figures, it can be noted that the prefit background estimate and shape of the distribution is well modelled by the simulation.

Figure 6.14 and 6.15 shows the recoil, $\cos\theta^*$ and transverse momentum of the leading b-jet in the event for $t\bar{t}(e)$ and $t\bar{t}(\mu)$ of 2b category. The distributions on the left are for the electron channel and muons on the right. The observables are shown for three different years, 2016 (top), 2017 (middle) and 2018 (bottom). From these figures, it can be noted that the prefit background estimate and shape of the distribution is well modelled by the simulation.

6.1.3.4 QCD control region

To estimate the QCD background in the signal region a dedicated control region is defined with same selection as in signal regions as shown in Table 6.2 with inverted $\min\Delta\phi(jet, p_T^{\text{miss}})$ cut for both 1b and 2b categories. The data MC agreement in the QCD control region for 2b category is shown in Figure 6.16, the data MC agreement

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

comes around 60% to 90% in all the three years.

However, the contribution of QCD background in signal region comes out to be very small (only present in only two bins out of four) and therefore it becomes very difficult to extract the QCD contribution in signal region by just using the QCD control regions with per bin transfer factor method, which is explained in detail in section 7.1.

6.1.4 Signal Background Separation

In order to separate the signal from various background processes(discussed in section 4.1.2) and to increase the sensitivity of the analysis a boosted decision tree(BDT) was trained using 25 variables and tested on an independent dataset.

1. E_T^{miss} : Missing Transverse Energy
2. $\phi(E_T^{\text{miss}})$: azimuthal angle of E_T^{miss} in the event.
3. E_T^{miss} significance: defined as the log-likelihood ratio of the observed E_T^{miss} to true E_T^{miss} .
4. JetHT: This is the scalar sum of all the jets in a selected event.
5. Jet1 p_T : Transverse momentum of leading jet in the event.
6. Jet1 η : pseudorapidity of leading jet in the event.
7. Jet1 ϕ : azimuthal angle of leading jet in the event.

8. Jet1deepCSV: A discriminator value coming from the output of b-tagging algorithm for leading jet. The value of the discriminator is greater than the medium working point as described in subsection [5.1.2.6](#)
9. Jet2 p_T : Transverse momentum of sub-leading jet in the event.
10. Jet2 η : pseudorapidity of sub-leading jet in the event.
11. Jet2 ϕ : azimuthal angle of sub-leading jet in the event.
12. Jet2deepCSV: A discriminator value coming from the output of b-tagging algorithm for sub-leading jet.
13. $\Delta R(\text{Jet1}, \text{Jet2})$: This is separation between two leading jets in η and ϕ plane, which is denoted by ΔR .
14. $\min(\Delta\phi(\text{jet}, MET))$: minimum ϕ separation between jets and E_T^{miss}
15. $\Delta\phi(\text{Jet1}, \text{Jet2})$: ϕ separation of leading and sub-leading jets.
16. $\Delta\eta(\text{Jet1}, \text{Jet2})$: η separation of leading and sub-leading jets.
17. dijet mass: invariant mass of two jets
18. dijet p_T : Transverse momentum of di-jet object
19. dijet η : pseudorapidity of di-jet object
20. dijet ϕ : azimuthal angle of di-jet object
21. Jet1 p_T / E_T^{miss} : ratio of transverse momentum of leading jet to E_T^{miss}

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

- 22. $\text{Jet1 } p_T / \text{Jet2 } p_T$: ratio of transverse momentum of leading jet to the transverse momentum of sub-leading jet
- 23. δ^+, δ^- : linear combinations of $\min(\Delta\phi(\text{jets}, p_T^{\text{miss}}))$ and $\Delta\phi$ between the b-jets
- 24. $\cos(\theta)^*$: using difference of pseudorapidity of the two leading and sub-leading jets defined as $\left| \tanh \frac{\Delta\eta(\text{Jet1}, \text{Jet2})}{2} \right|$ This variable was introduced in [85] and used in [86] in the context of $t\bar{t} + E_T^{\text{miss}}$ analysis.

With these variables, the BDT provides ranking with respect to the power of separation for signal and background of a variable, and according to these rankings, the $\cos(\theta)^*$ variable performs best and has the highest separation of signal and background. Figure 6.3, the signal to background separation plot for the above-listed variables is shown.

The full output of the BDT is shown in Figure 6.4, and it can be seen that it is able to separate the signal and background very well. This full bdt exercise has been done for 2b category control and signal regions. However, for most of the variables mentioned above, at least two jets are required and hence cannot be used for 1b signal and control regions because of their definition discussed in subsection 6.1.2 and 6.1.3.

Even though the area under the curve is much better for the BDT w.r.t any single variable of the Receiver Operating Characteristics curve, the background estimation method relies on the control region and statistics in the control region; the overall sensitivity gain with the use of BDT is negligible, and hence the BDT itself could not be used, but it provided the best out of the 25 variables that were used for the

6.1. EVENT SELECTION

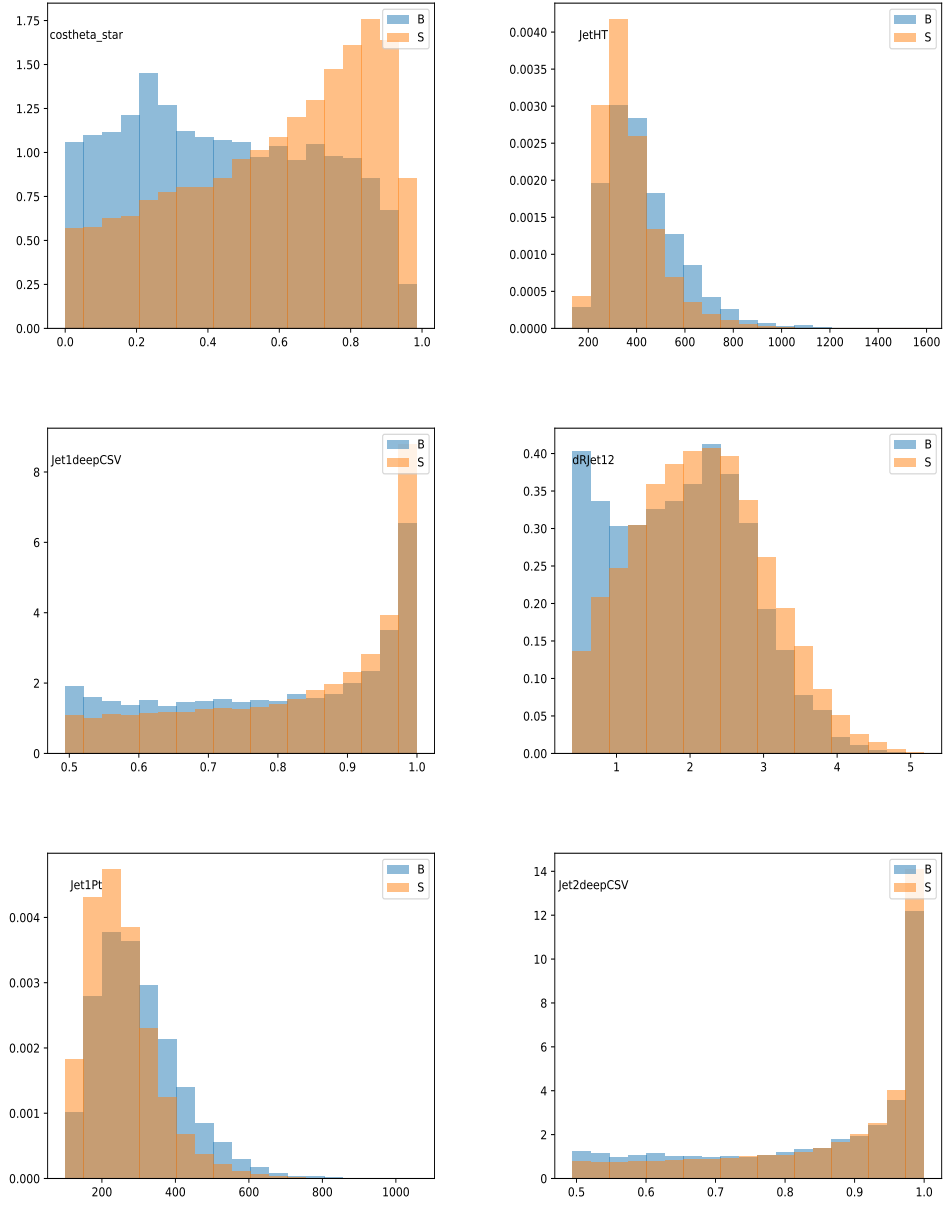


Figure 6.3: The signal to background separation of top six variables that were fed to the BDT framework. Out of 25 variables, $\cos(\theta)^*$ gives highest signal to background separation.

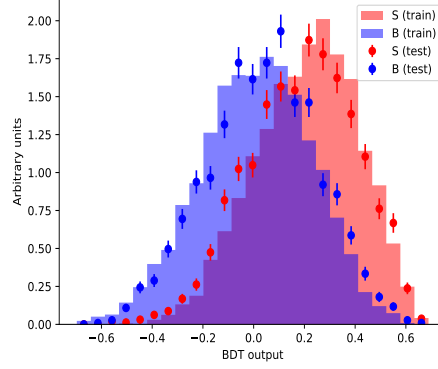


Figure 6.4: The bdt discriminator for 2b category of showing the separation of signal and background.

training the BDT. So, $\cos(\theta)^*$ (shown in Figure 6.3) is used for the signal extraction. One more point to note that most of the variables fed to BDT require at least two jets in an event, however for 1b category both in signal and control regions, we have at most two jets, and hence the BDT cannot be used in this category. Due to this reason, for 1b category E_T^{miss} remains the main observable, but for 2b category, $\cos(\theta)^*$ is used. In section 8.1, the results with $\cos(\theta)^*$ in 2b and E_T^{miss} in 1b are discussed. In section 8.1, the results with $\cos(\theta)^*$ in 2b and E_T^{miss} in 1b are discussed.

6.1. EVENT SELECTION

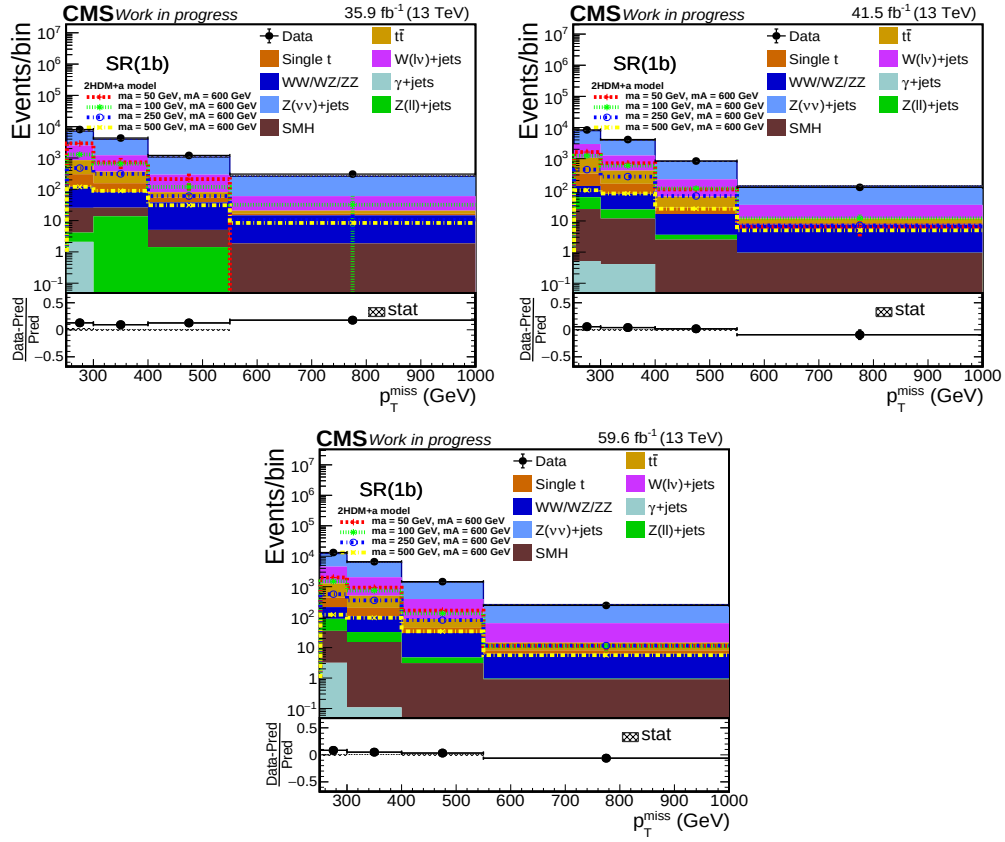


Figure 6.5: p_T^{miss} plot in 1b Category of Signal region for 2016, 2017 and 2018

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

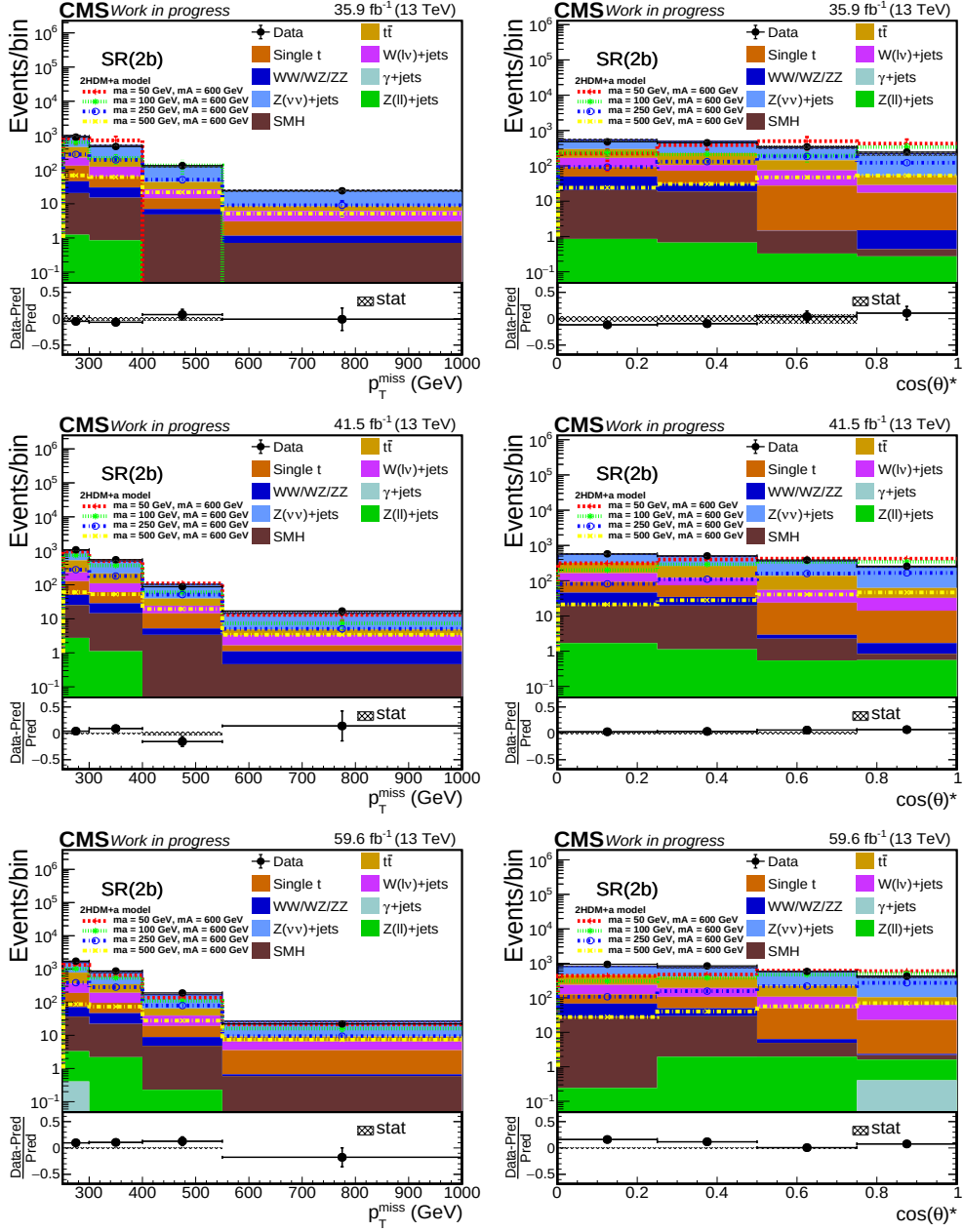


Figure 6.6: p_T^{miss} and $\cos(\theta)^*$ plot in 2b Category of Signal region for 2016, 2017 and 2018. The signal in both set of the plots is scaled by 5. $\cos(\theta)^*$ provides better signal to background separation as compared to p_T^{miss}

6.1. EVENT SELECTION

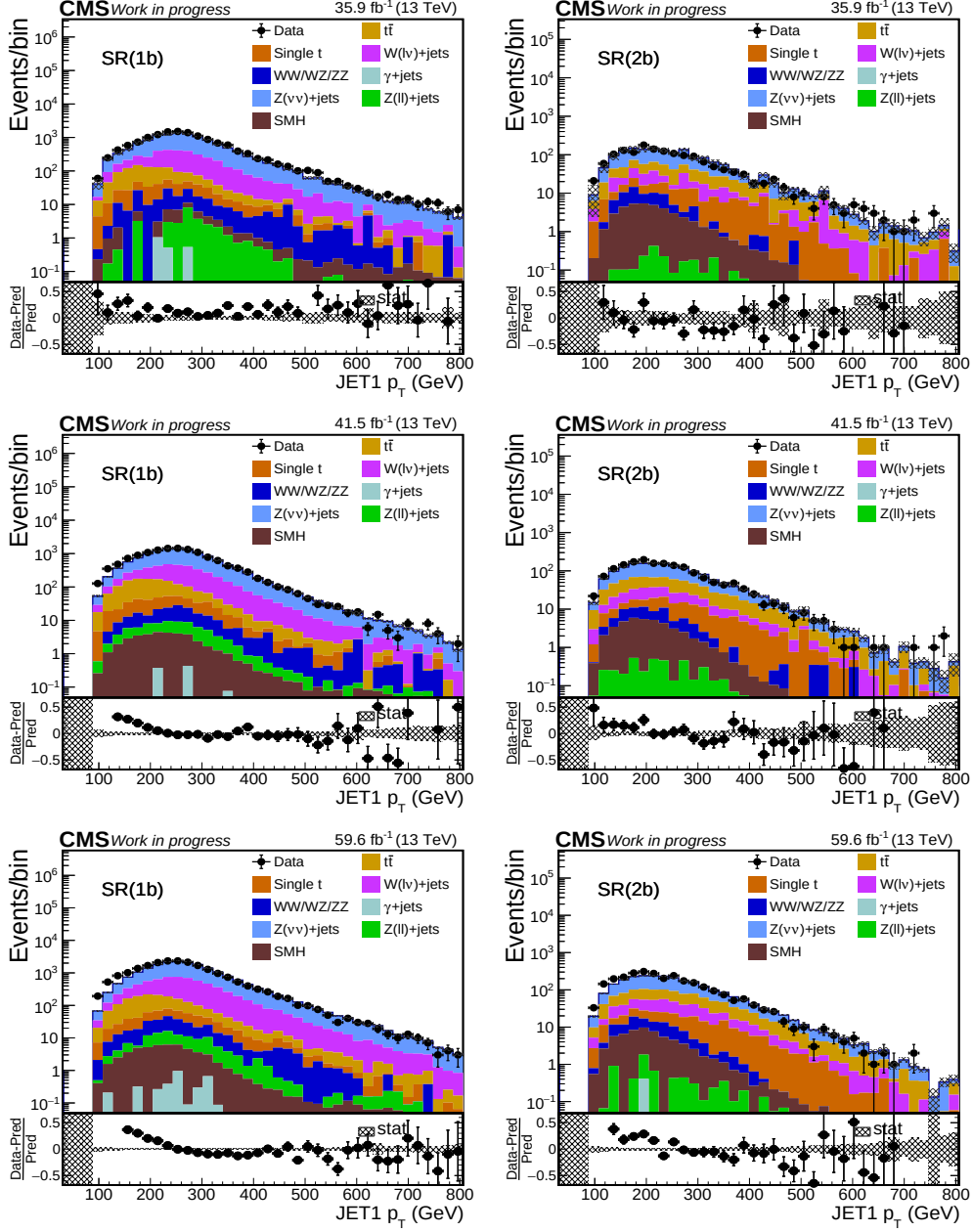


Figure 6.7: Jet1 p_T plot in 1b and 2b Category of Signal region for 2016, 2017 and 2018

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

Table 6.5: Background yield for SR (1b)

	2016	2017	2018
$t\bar{t}$	790.63 $\pm$ 9.18	1026.72 $\pm$ 12.31	1192.53 $\pm$ 7.59
Single t	276.52 $\pm$ 10.45	291.01 $\pm$ 11.74	338.55 $\pm$ 5.76
W(l ν)+jets	2779.79 $\pm$ 77.73	2915.21 $\pm$ 17.60	5185.18 $\pm$ 30.59
WW/WZ/ZZ	195.87 $\pm$ 17.96	145.72 $\pm$ 11.03	231.46 $\pm$ 8.24
Z($\nu\nu$)+jets	8407.12 $\pm$ 144.90	7851.18 $\pm$ 31.49	12279.17 $\pm$ 20.48
Z(l l)+jets	17.20 $\pm$ 10.18	46.36 $\pm$ 2.29	71.44 $\pm$ 4.57
SM H	39.25 $\pm$ 0.44	38.63 $\pm$ 0.33	51.45 $\pm$ 0.19
Total Bkg	12508.66 $\pm$ 166.31	12315.83 $\pm$ 41.45	19353.08 $\pm$ 39.19
data_obs	13993.00 $\pm$ 118.25	13455.00 $\pm$ 116.00	21562.00 $\pm$ 146.83
$m_a = 50\text{GeV}, m_A = 600\text{GeV}$	3869.61 $\pm$ 667.03	2450.17 $\pm$ 62.89	3105.46 $\pm$ 60.36
$m_a = 100\text{GeV}, m_A = 600\text{GeV}$	2102.19 $\pm$ 244.14	1958.10 $\pm$ 47.57	2387.48 $\pm$ 45.49
$m_a = 250\text{GeV}, m_A = 600\text{GeV}$	858.42 $\pm$ 46.87	780.99 $\pm$ 14.59	1015.18 $\pm$ 15.37
$m_a = 500\text{GeV}, m_A = 600\text{GeV}$	248.70 $\pm$ 8.09	202.91 $\pm$ 3.44	255.45 $\pm$ 3.81

Table 6.6: Background yield for SR (2b)

	2016	2017	2018
$t\bar{t}$	347.81 $\pm$ 6.01	414.83 $\pm$ 7.52	592.11 $\pm$ 5.36
Single t	126.78 $\pm$ 7.04	120.08 $\pm$ 7.50	180.47 $\pm$ 4.38
W(l ν)+jets	167.08 $\pm$ 19.56	165.76 $\pm$ 4.48	294.26 $\pm$ 6.91
WW/WZ/ZZ	42.44 $\pm$ 4.88	41.63 $\pm$ 3.48	62.61 $\pm$ 3.88
Z($\nu\nu$)+jets	869.66 $\pm$ 47.71	789.13 $\pm$ 9.58	1219.49 $\pm$ 6.37
Z(l l)+jets	2.09 $\pm$ 0.35	3.93 $\pm$ 0.58	5.38 $\pm$ 1.33
SM H	39.21 $\pm$ 0.43	39.79 $\pm$ 0.32	58.54 $\pm$ 0.20
Total Bkg	1595.11 $\pm$ 52.62	1575.16 $\pm$ 15.40	2413.27 $\pm$ 12.38
data_obs	1523.00 $\pm$ 39.03	1729.00 $\pm$ 41.58	2792.00 $\pm$ 52.84
$m_a = 50\text{GeV}, m_A = 600\text{GeV}$	1536.51 $\pm$ 402.16	1558.17 $\pm$ 47.95	2115.98 $\pm$ 48.50
$m_a = 100\text{GeV}, m_A = 600\text{GeV}$	984.89 $\pm$ 164.53	1180.87 $\pm$ 35.08	1670.08 $\pm$ 37.59
$m_a = 250\text{GeV}, m_A = 600\text{GeV}$	531.16 $\pm$ 36.38	517.46 $\pm$ 11.39	754.44 $\pm$ 12.98
$m_a = 500\text{GeV}, m_A = 600\text{GeV}$	155.09 $\pm$ 6.28	136.39 $\pm$ 2.69	194.86 $\pm$ 3.25

6.1. EVENT SELECTION

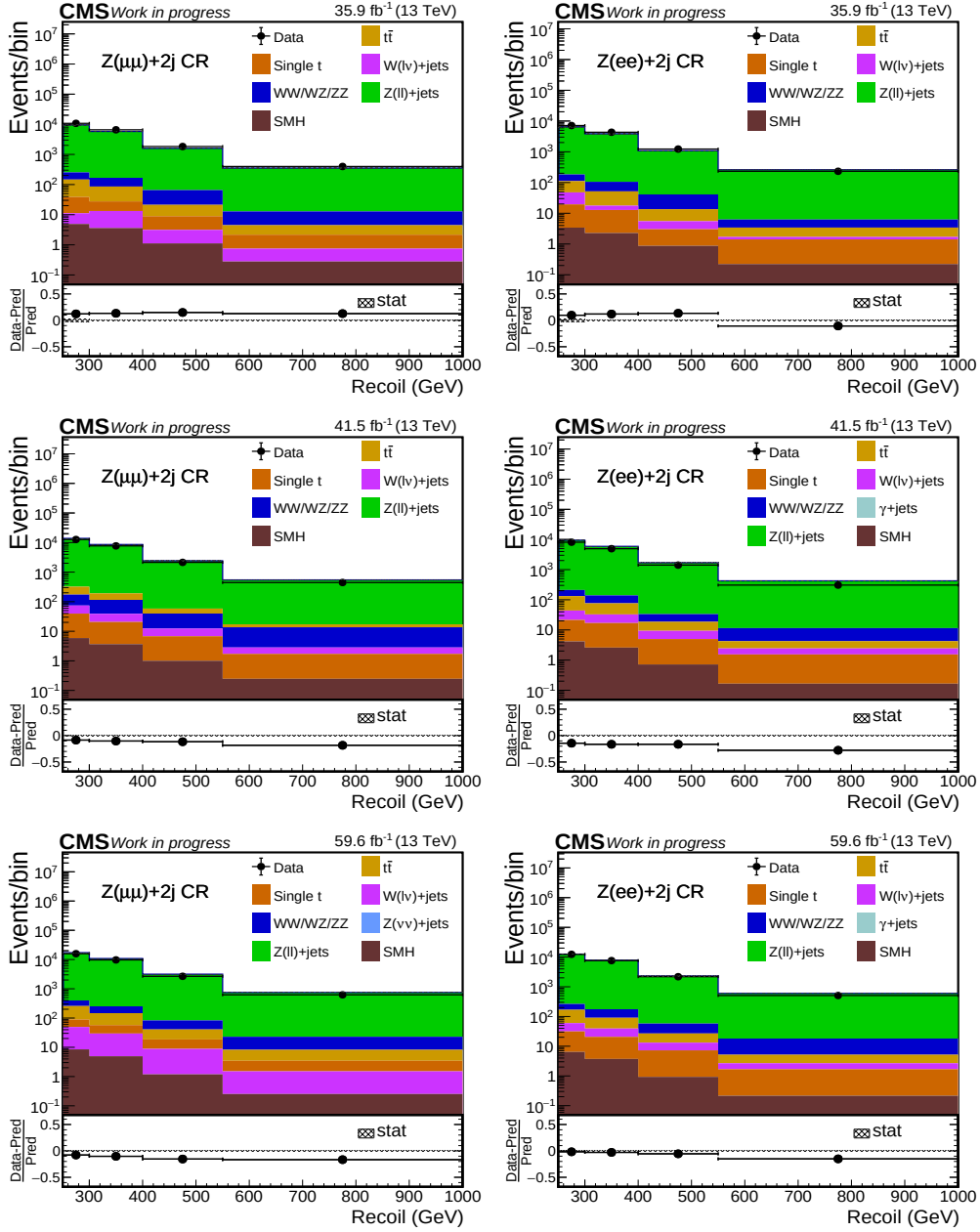


Figure 6.8: Recoil plot in 1b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

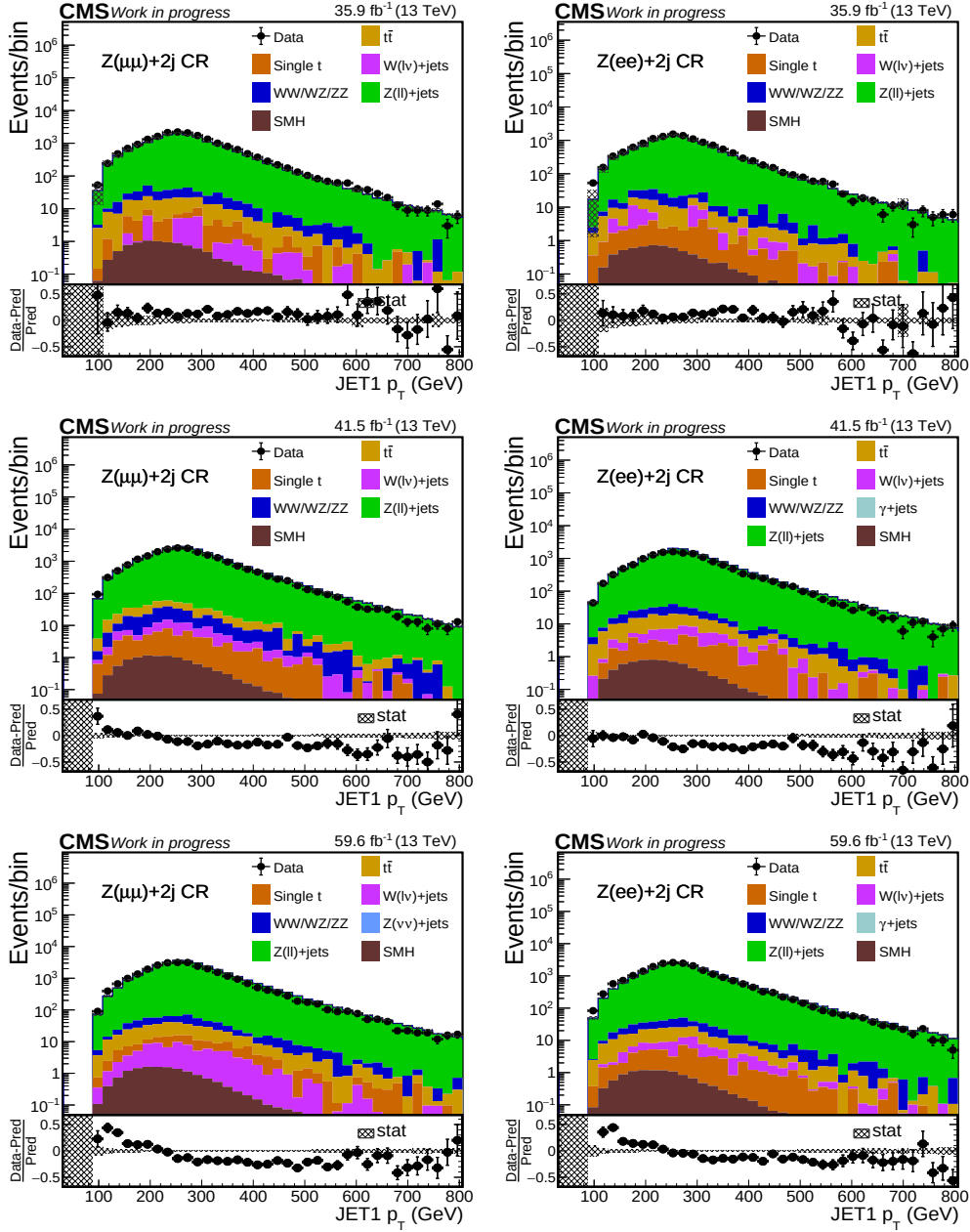


Figure 6.9: Jet1 p_T plot in 1b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

6.1. EVENT SELECTION

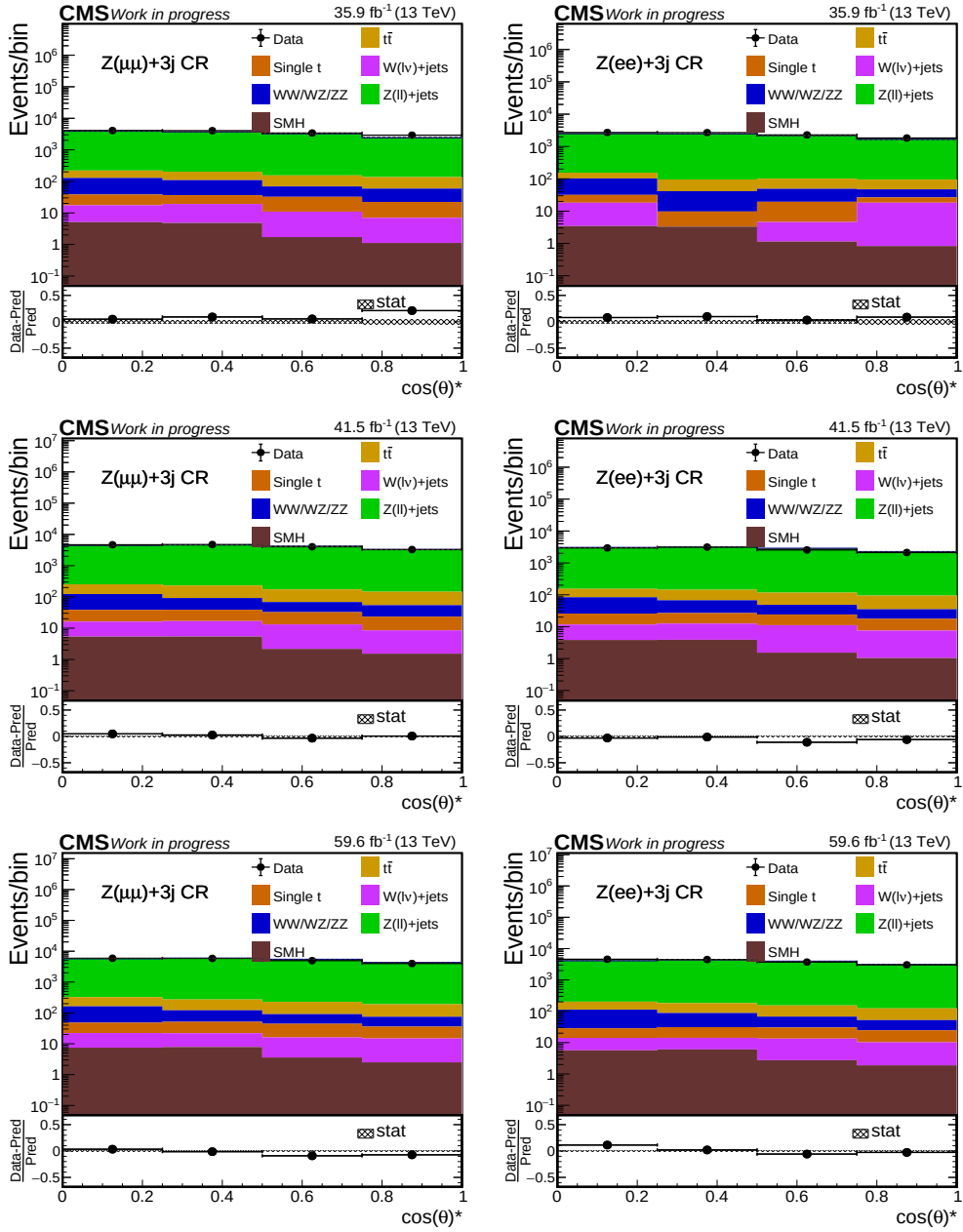


Figure 6.10: $\cos(\theta)^*$ plot in 2b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

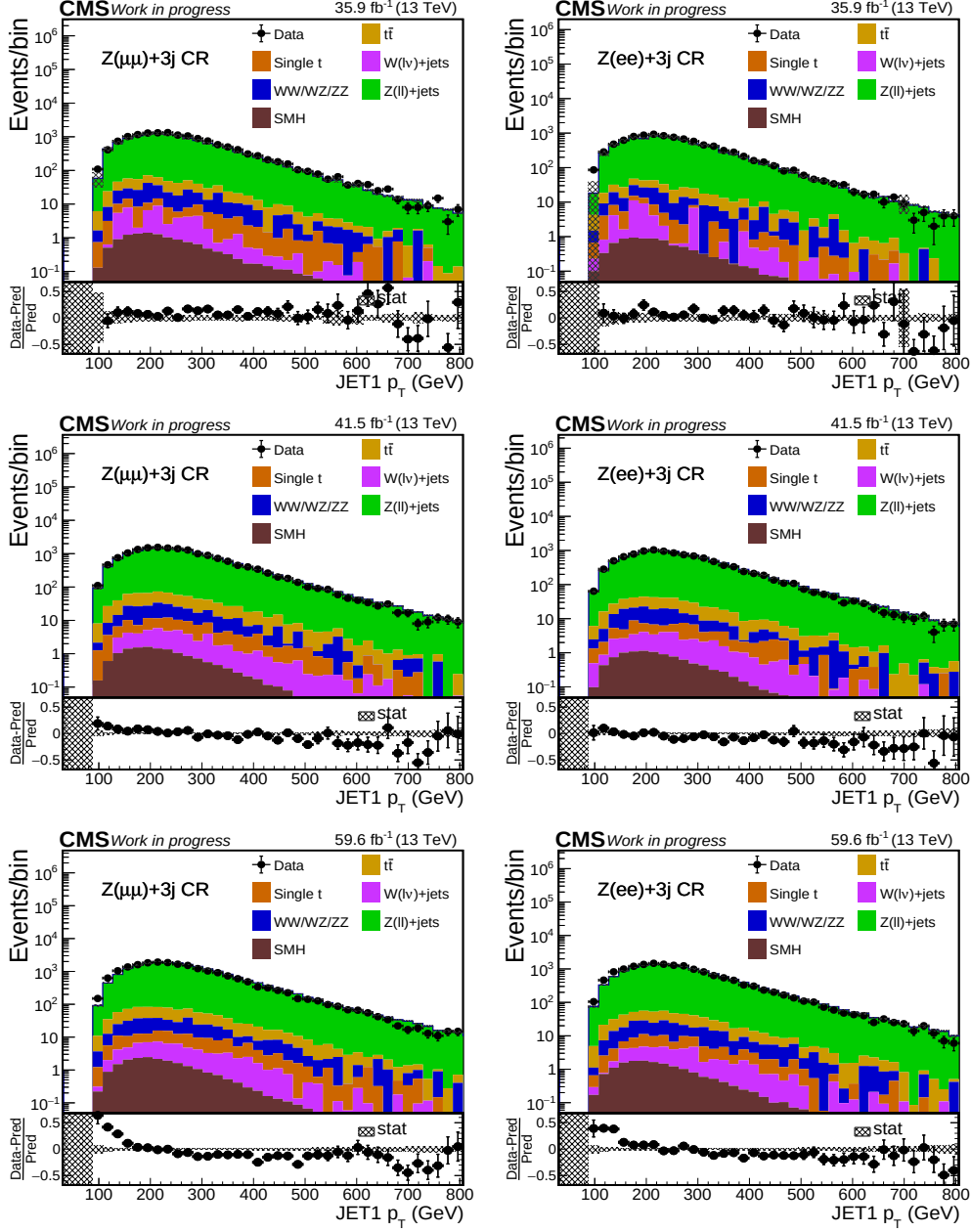


Figure 6.11: $\text{Jet1 } p_T$ plot in 2b Category of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018.

6.1. EVENT SELECTION

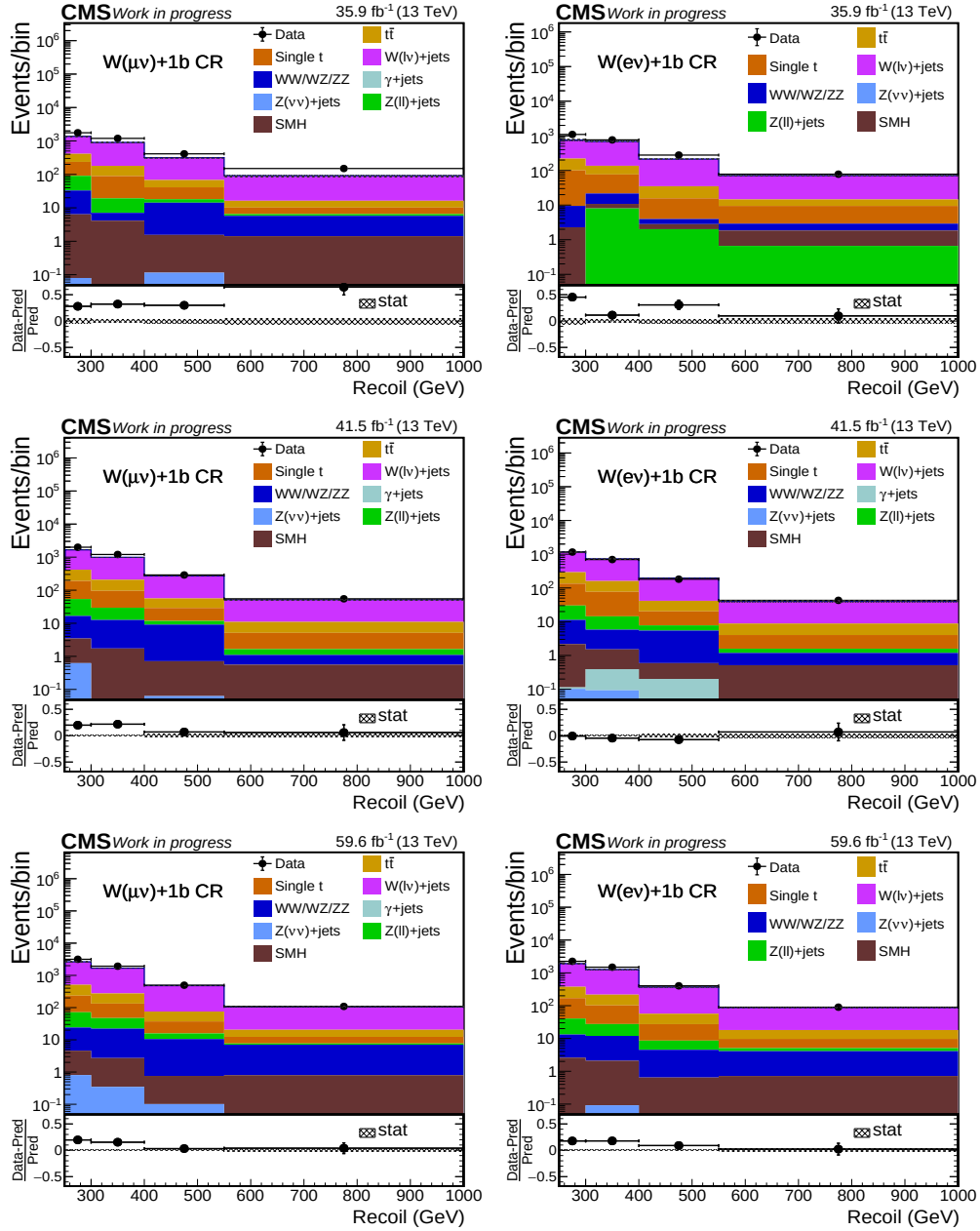


Figure 6.12: Recoil plot in 1b Category of $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

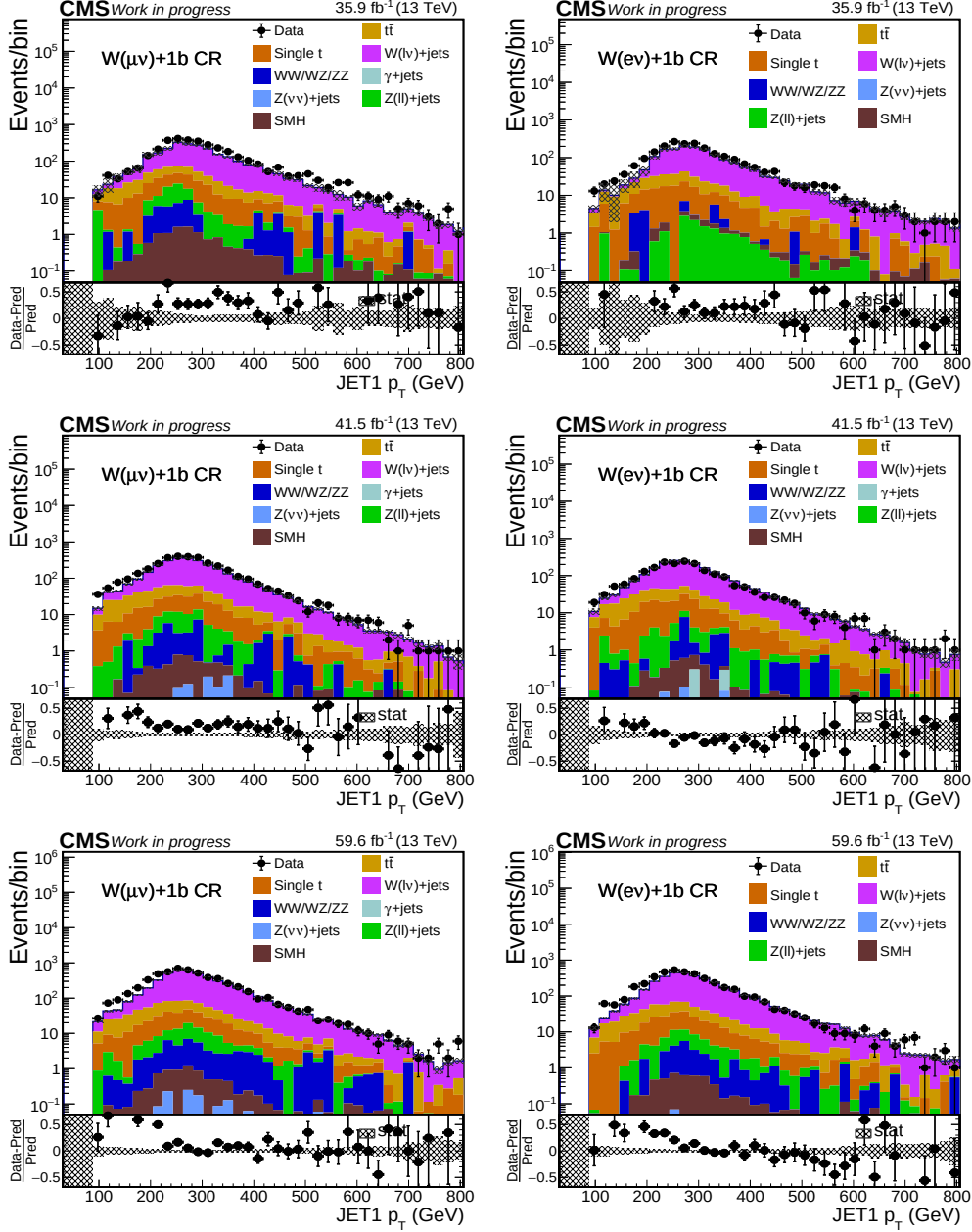


Figure 6.13: Jet1Pt plot in 1b Category of $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

6.1. EVENT SELECTION

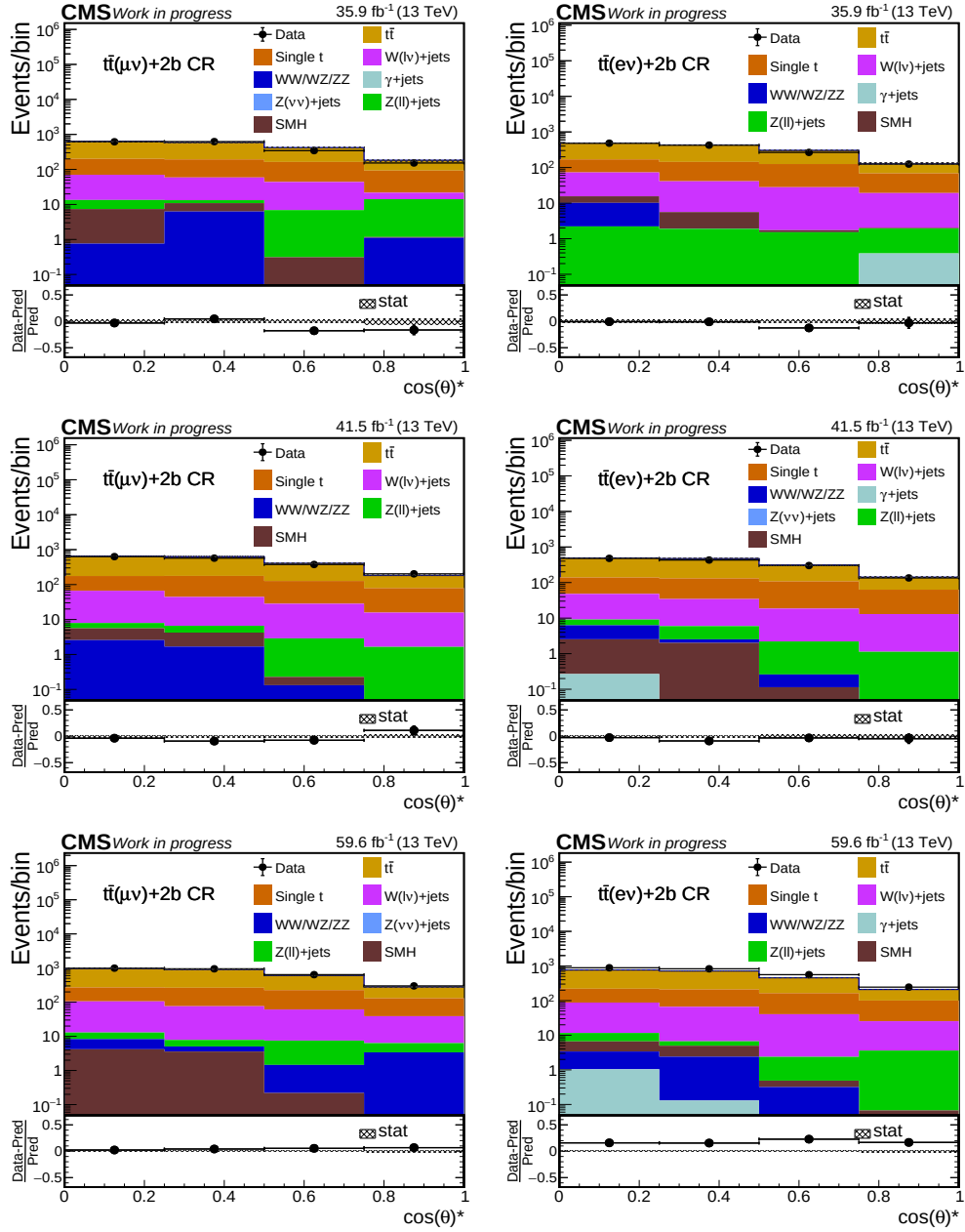


Figure 6.14: $\cos(\theta)^*$ plot in 2b Category of Top(μ) and Top(e) control region for 2016, 2017 and 2018.

CHAPTER 6. ANALYSIS STRATEGY AND EVENT SELECTION

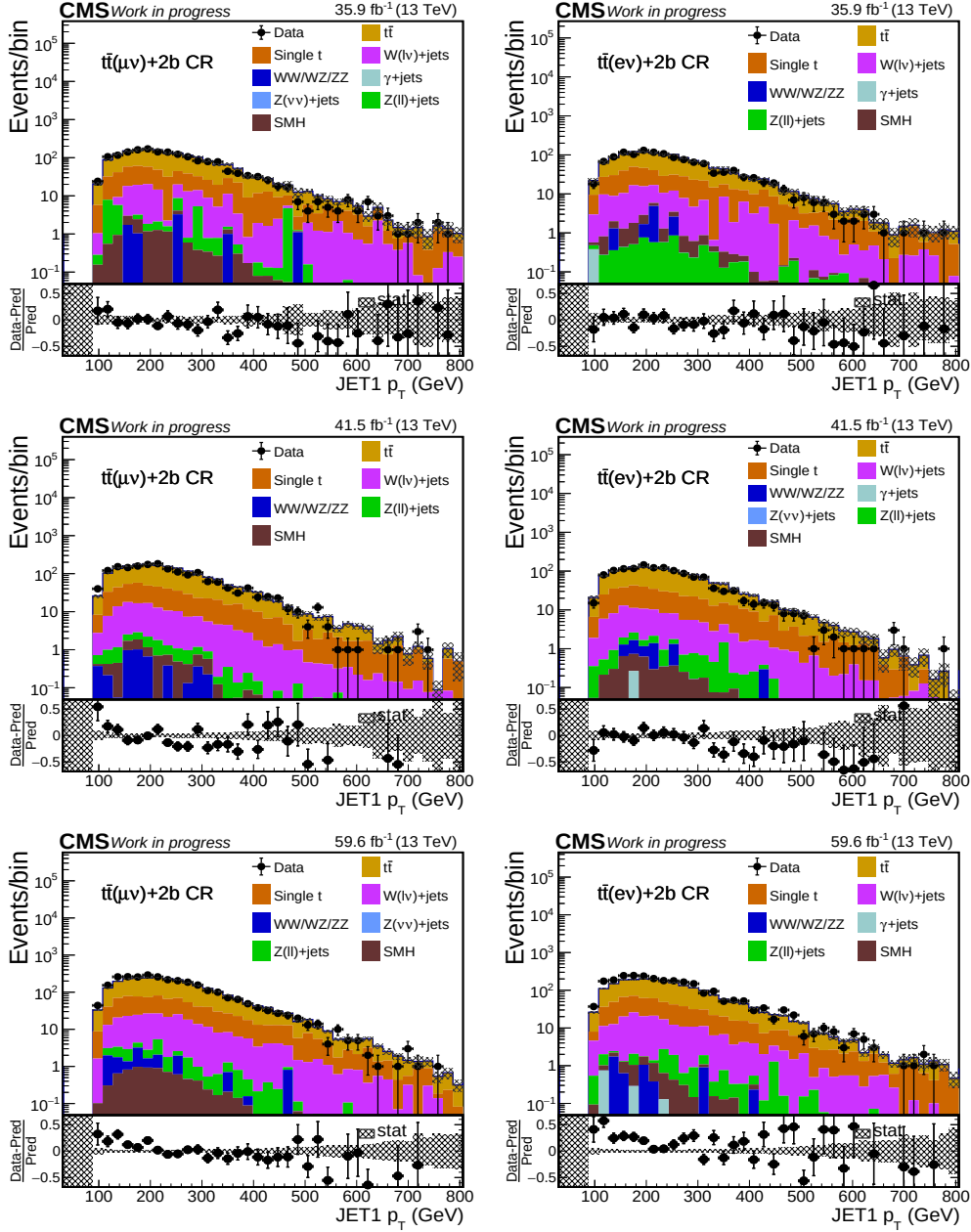


Figure 6.15: Jet1 p_T plot in 2b Category of Top(μ) and Top(e) control region for 2016, 2017 and 2018.

6.1. EVENT SELECTION

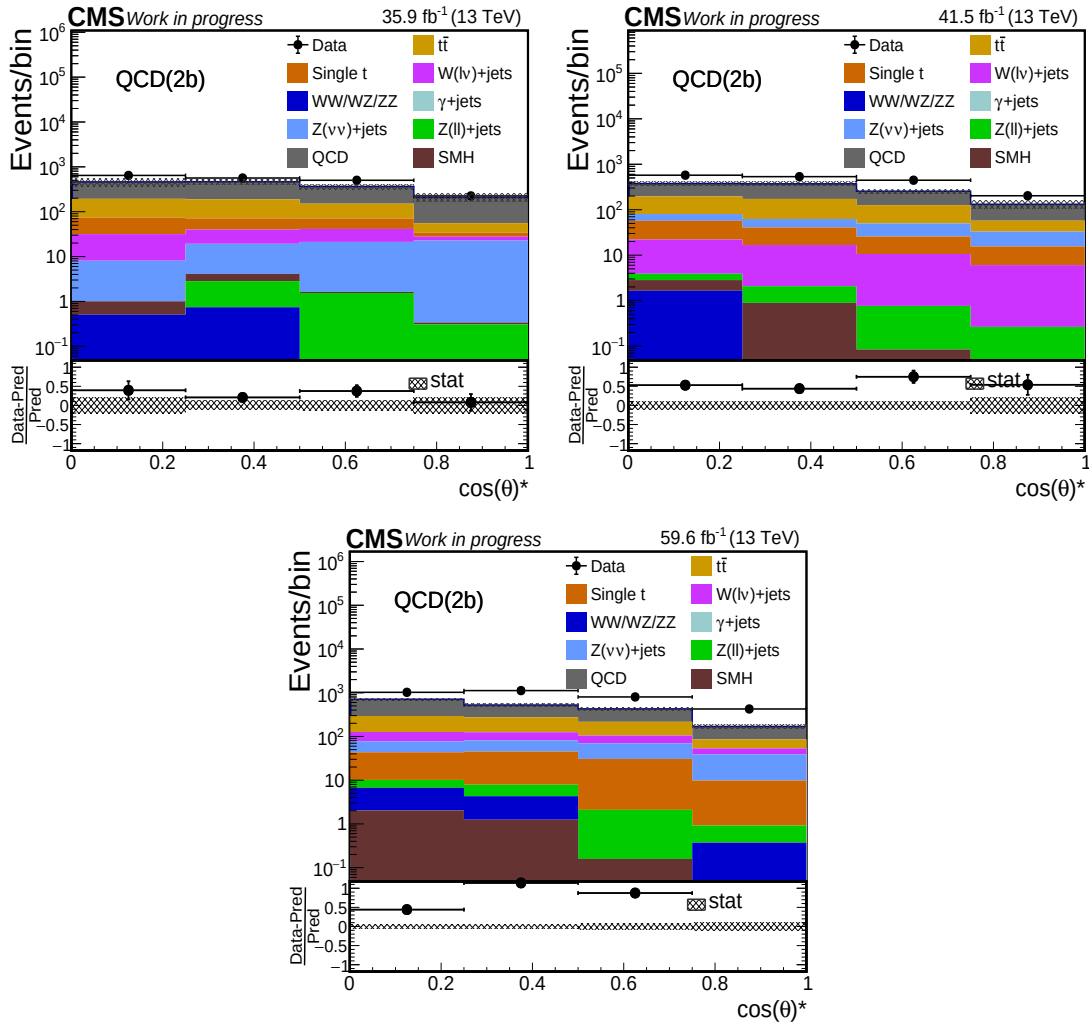


Figure 6.16: $\cos\theta^*$ plot in the QCD control region for 2016, 2017 and 2018

CHAPTER 7

SIGNAL EXTRACTION AND SYSTEMATICS

UNCERTAINTIES

7.1 Signal extraction

The major backgrounds $t\bar{t}$ and $Z(\nu\nu) + \text{jets}$ in 2b category and $W + \text{jets}$ and $Z(\nu\nu) + \text{jets}$ in 1b category are estimated using a simultaneous fit of signal and control regions. For each control region, the recoil variable is computed by removing either the muon(s) or electron(s) from the p_T^{miss} calculation. The fit is performed by tying each bin of p_T^{miss} in the signal region to the same bin of the recoil in the control region via transfer factors, treating the systematic uncertainties as nuisance parameters in the fit. The binning used for the recoil variable in 1b category and $\cos\theta^*$ in 2b categories are as follows:

- 1b: 250, 300, 400, 550, 1000

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

- 2b: 0.0, 0.25, 0.50, 0.75, 1.00

The binning is chosen to ensure maximum sensitivity for individual categories. The binned transfer factors are defined as the ratio of number of events for a given background in the signal region to the number of events of respective background in the corresponding control region. For example, they can be represented as:

$$R_i^Z = \frac{N_i^{Z \rightarrow \nu\nu}}{N_i^{Z \rightarrow ll}}, \quad R_i^W = \frac{N_i^{W \rightarrow l\nu}}{N_i^{W+jets}}, \quad R_i^t = \frac{N_i^{tt}}{N_i^{tt}} \quad (7.1)$$

The distribution of the TF for each year and each categories are shown in figure 7.1, 7.2, 7.3 and 7.4. The trend and size of the TF looks in good agreement when compare with different years and different lepton flavor. The statistical and systematic uncertainty on the transfer factors are described in 7.2.2. The likelihood for 1b and 2b are constructed using a similar procedure. The full likelihood can then be expressed as the product of two individual likelihoods.

$$\mathcal{L} = \mathcal{L}_{1b} \times \mathcal{L}_{2b}$$

7.1. SIGNAL EXTRACTION

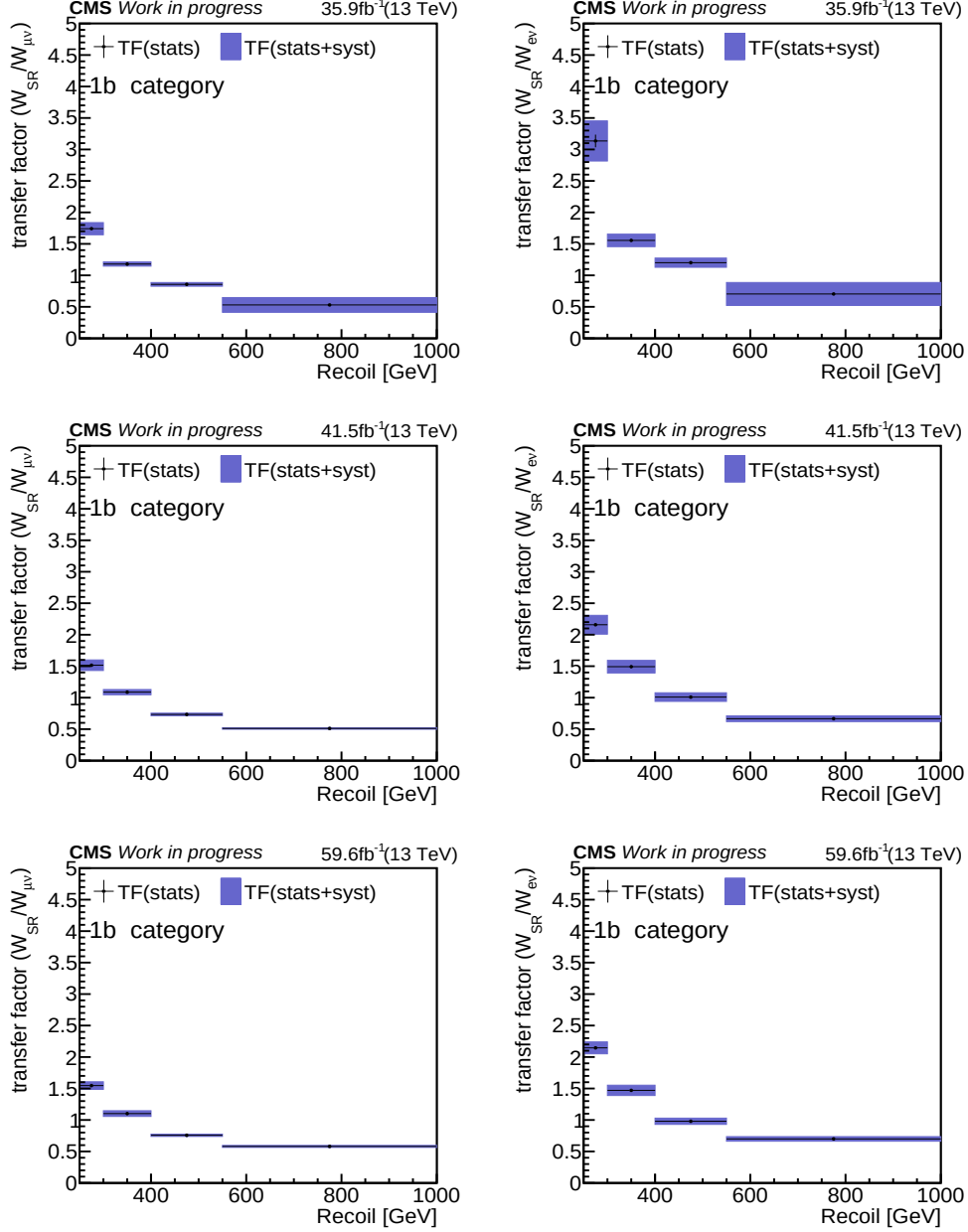


Figure 7.1: Distribution of transfer factors $W_{SR}/W_{\mu\nu}$ and $W_{SR}/W_{e\nu}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

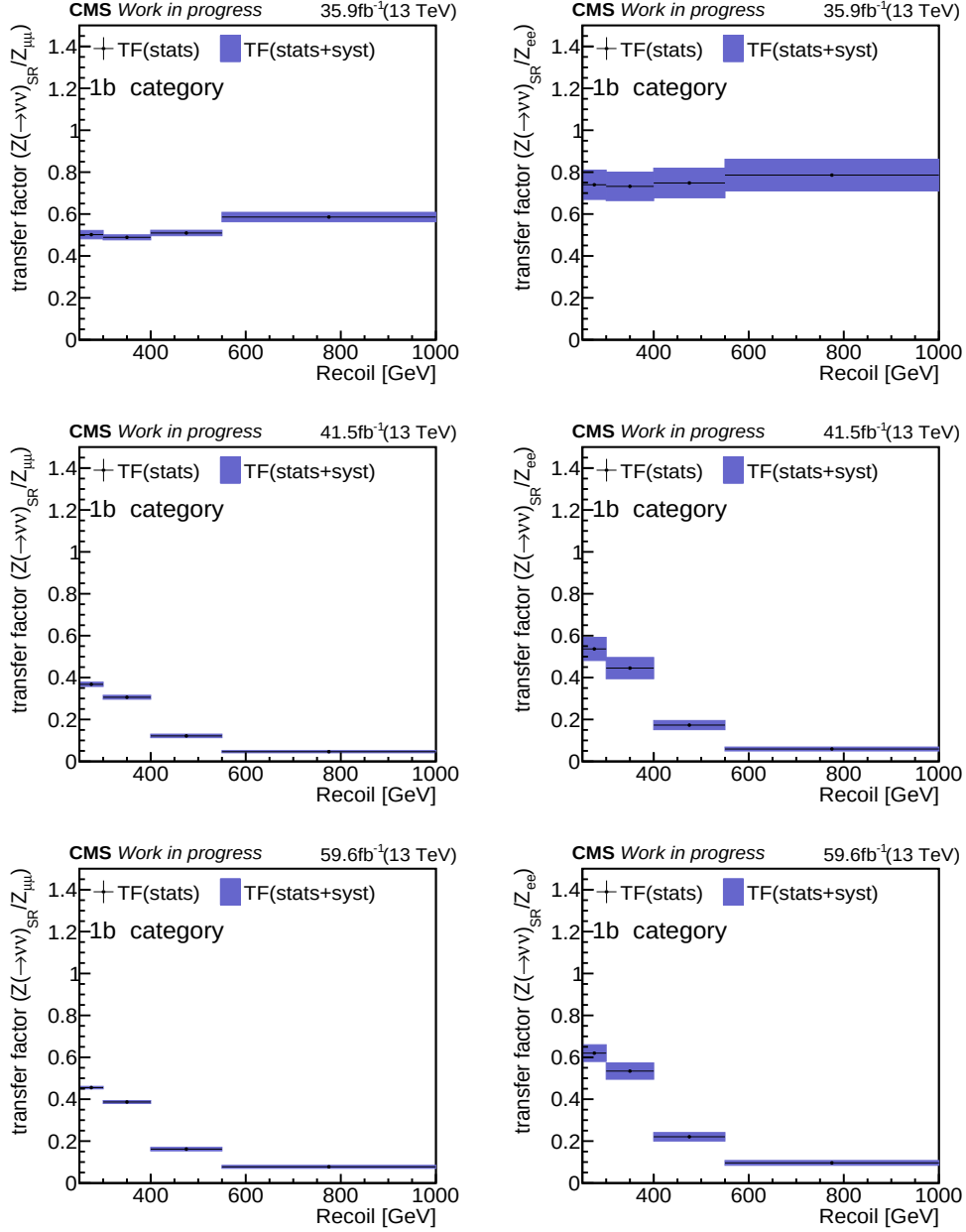


Figure 7.2: Distribution of transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.

7.1. SIGNAL EXTRACTION

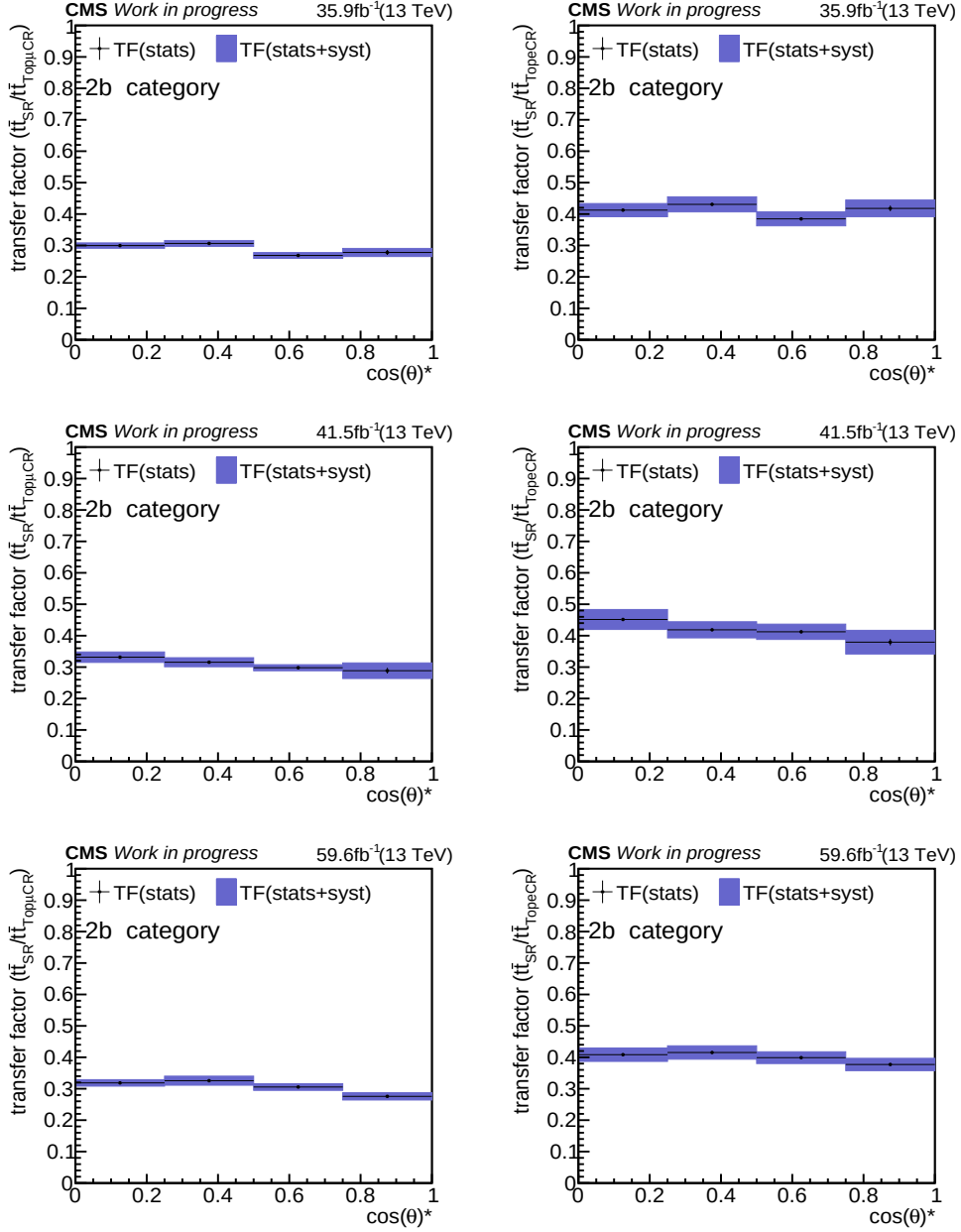


Figure 7.3: Distribution of transfer factors $t\bar{t}_{SR}/t\bar{t}_{Top\mu CR}$ and $t\bar{t}_{SR}/t\bar{t}_{Top e CR}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

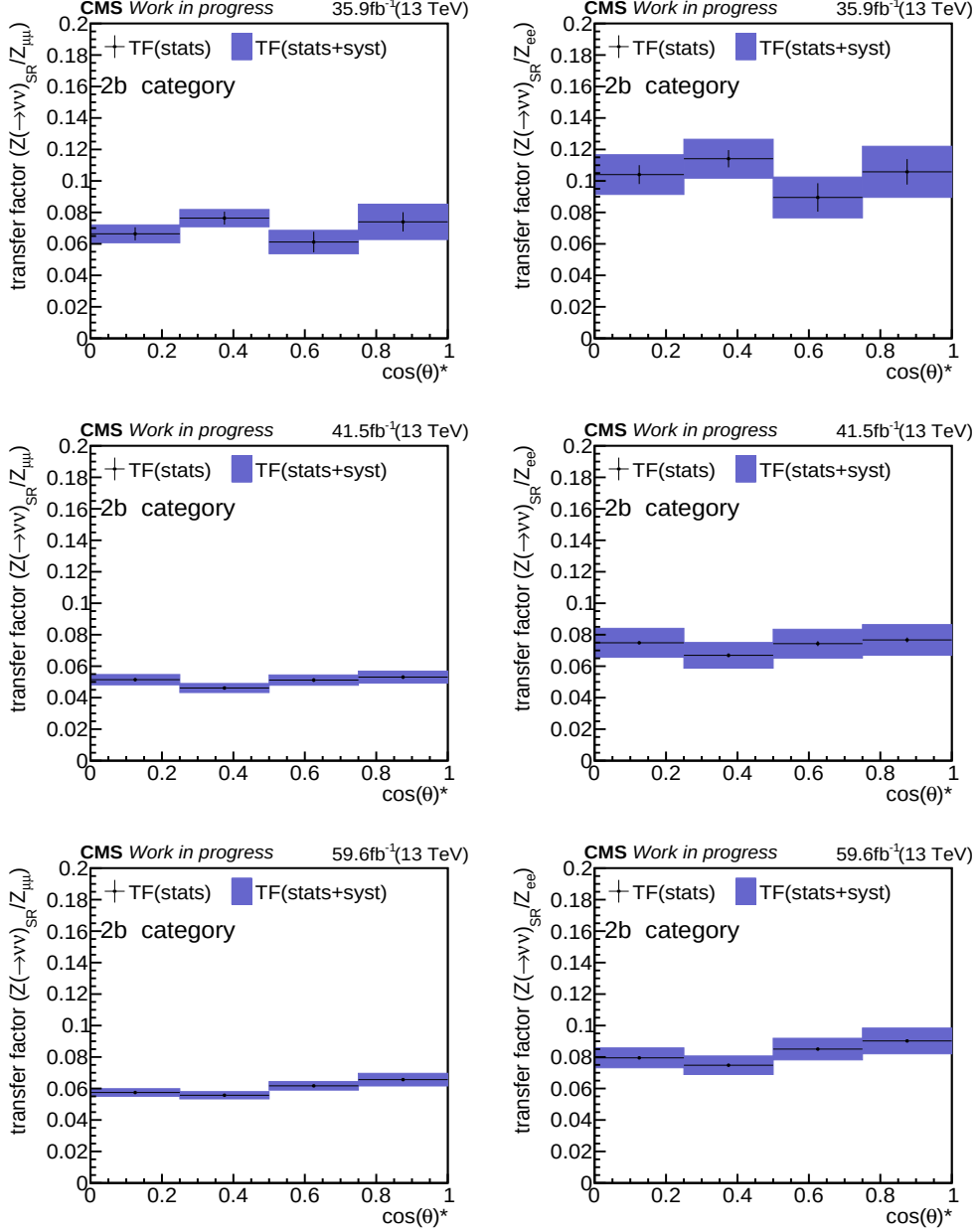


Figure 7.4: Distribution of transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018. The statistical uncertainty on the TF and nuisances associated with them used during the signal extraction fit are also shown.

7.1. SIGNAL EXTRACTION

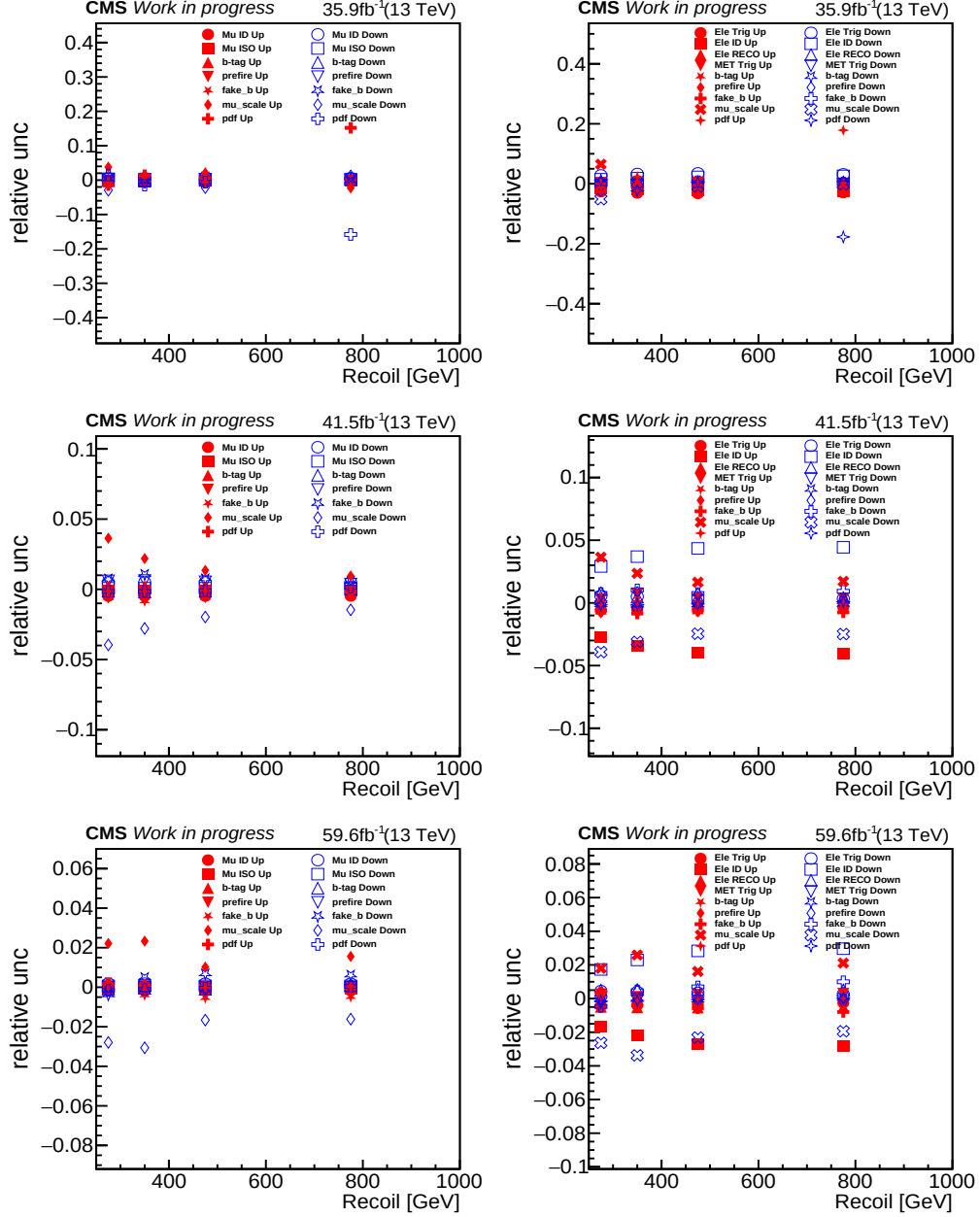


Figure 7.5: Uncertainties on the transfer factors $W_{SR}/W_{\mu\nu}$ and W_{SR}/W_{ev} for 1b category as a function of Recoil variable for 2016, 2017 and 2018.

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

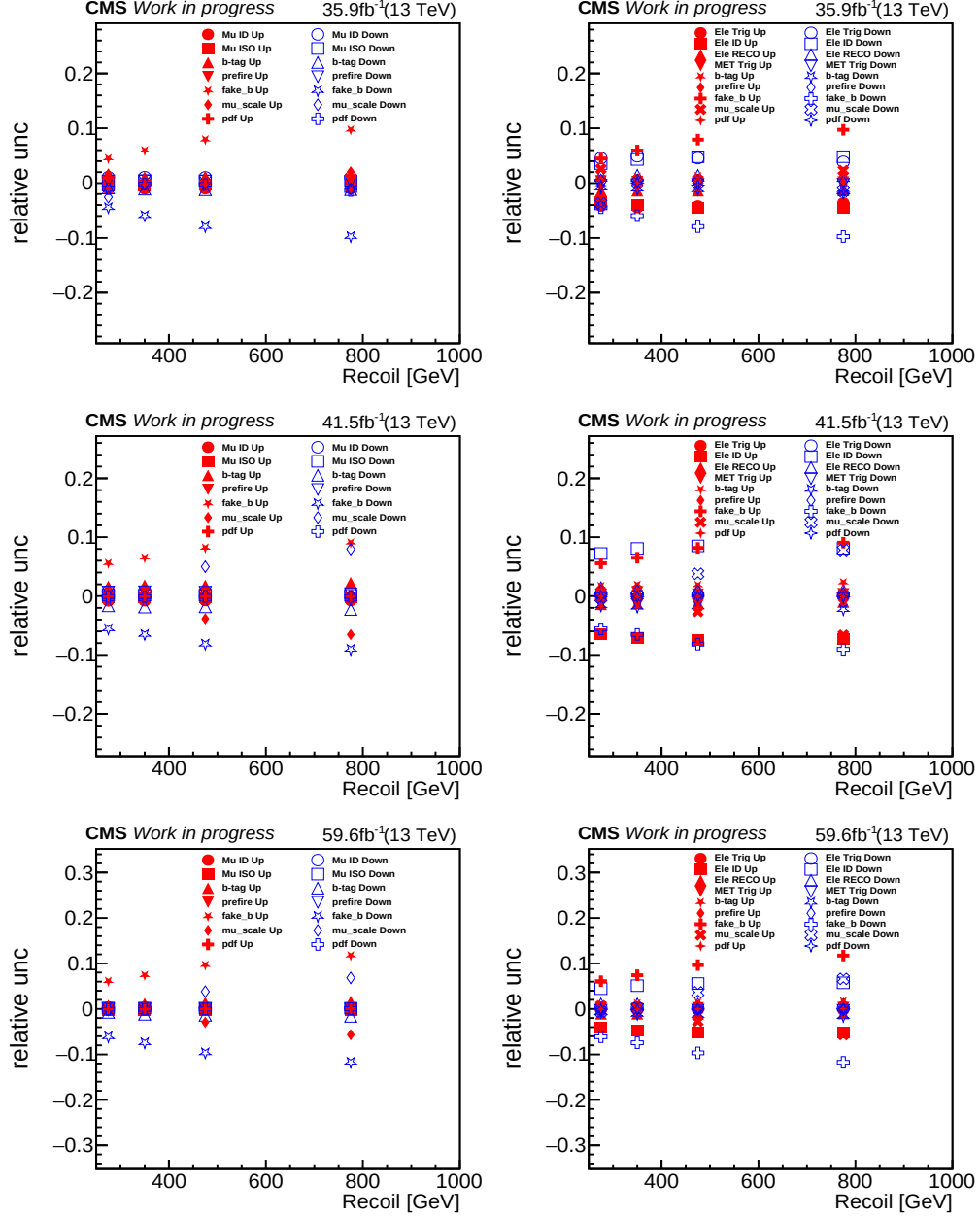


Figure 7.6: Uncertainties on the transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 1b category as a function of Recoil variable for 2016, 2017 and 2018.

7.1. SIGNAL EXTRACTION

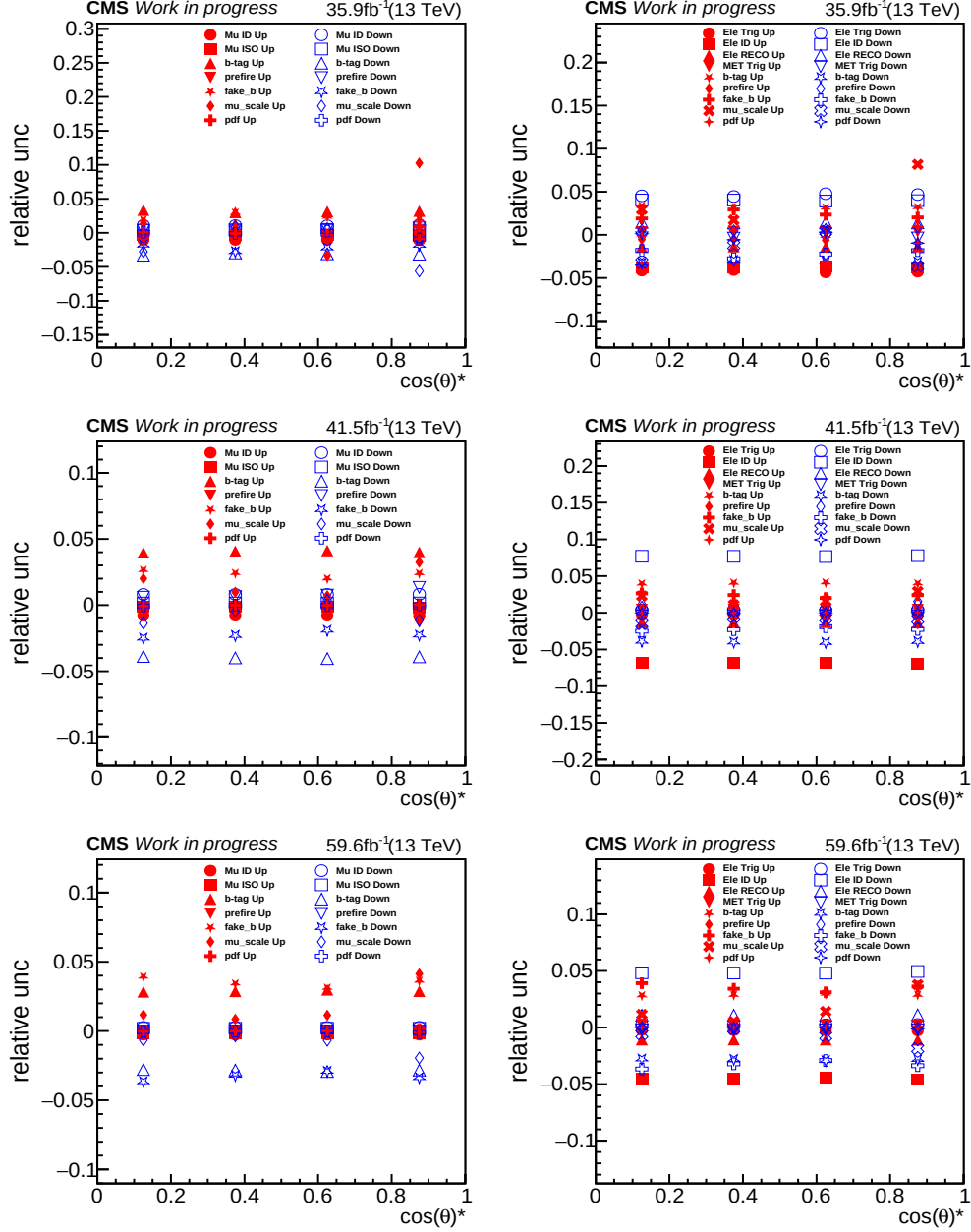


Figure 7.7: Uncertainties on the transfer factors $Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$ and $Z(\rightarrow \nu\nu)_{SR}/Z_{ee}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018.

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

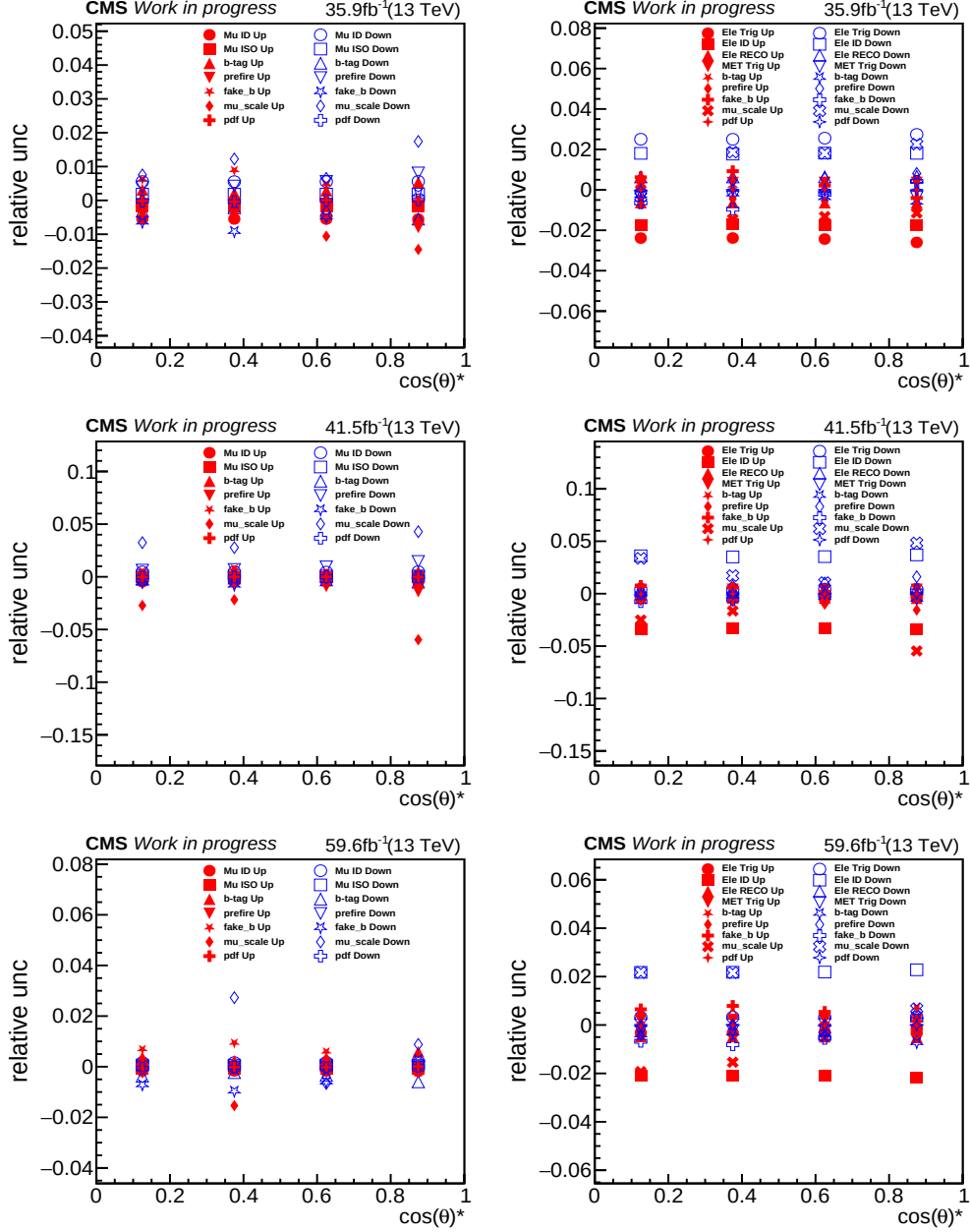


Figure 7.8: Uncertainties on the transfer factors $t\bar{t}_{SR}/t\bar{t}_{Top\mu CR}$ and $t\bar{t}_{SR}/t\bar{t}_{Top e CR}$ for 2b category as a function of $\cos(\theta)^*$ variable for 2016, 2017 and 2018.

The likelihood for one category say 2b can be written as follows:

$$\begin{aligned}
\mathcal{L}_{2b}(\mu^{Z \rightarrow \nu\nu}, \mu^{W+jets}, \mu^{tt}, \boldsymbol{\theta}) &= \prod_{i=1 \text{ to nbins}} \text{Poisson} \left(d_i^Z | B_i^Z(\boldsymbol{\theta}) + \frac{\mu_i^{Z \rightarrow \nu\nu}}{R_i^Z(\boldsymbol{\theta})} \right) \\
&\times \prod_{i=1 \text{ to nbins}} \text{Poisson} \left(d_i^W | B_i^W(\boldsymbol{\theta}) + \frac{\mu_i^{W+jets}}{R_i^W(\boldsymbol{\theta})} \right) \\
&\times \prod_{i=1 \text{ to nbins}} \text{Poisson} \left(d_i^{tt} | B_i^{tt}(\boldsymbol{\theta}) + \frac{\mu_i^{tt}}{R_i^{tt}(\boldsymbol{\theta})} \right)
\end{aligned} \tag{7.2}$$

where, $\mu^{Z \rightarrow \nu\nu}$, μ^{W+jets} , μ^{tt} are the background yields of $Z \rightarrow \nu\nu$, $W+jets$ and $t\bar{t}$ processes in the signal region. These are to be estimated during the fit and kept freely floating. $\mu_i^{Z \rightarrow \nu\nu}$, μ_i^{W+jets} , and μ_i^{tt} are the background yield of respective process in the i^{th} bin of the signal region. The systematic uncertainties ($\boldsymbol{\theta}$) enter the likelihood as a perturbation on the transfer factors R_i^Z , R_i^W , and R_i^{tt} . The per bin transfer factors R_i^Z , R_i^W , and R_i^{tt} are defined as the ratio of the yield of process in the SR to the yield in the respective control region used to model it.

The observed number of events in each region is denoted by d_i^{region} . B_i^{region} denotes the number of expected background events in each region, excluding the background being estimated by the fit (e.g. $Z \rightarrow \mu\mu$ in the $Z\mu\mu$ control region).

The $Z \rightarrow \nu\nu$ background prediction in the signal region is connected to the yields of $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ events in the dilepton control samples. The associated trans-

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

fer factors account for the differences in the branching ratio of Z bosons to charged leptons relative to neutrinos and the impact of lepton acceptance and selection efficiencies. In the case of dielectron events, the transfer factor also takes into account the difference in the trigger efficiencies as the trigger used in the signal region and dielectron region are different.

As described in the event selection, the single electron and single muon control regions with no additional jets are used only for 1b category. These events are W+jets enriched. The W+jets background prediction in the signal region is connected to the yields of $W \rightarrow \mu\nu$ and $W \rightarrow e\nu$ events in the single lepton control samples. The associated transfer factors account for the differences in the lepton acceptance and selection efficiencies. In the case of electron events, the transfer factor also takes into account the difference in the trigger efficiencies as the trigger used in the signal region and dielectron region are different.

The events with > 0 additional jets are $t\bar{t}$ enriched and used for 2b category to estimate the $t\bar{t}$ background in the signal region. Similar to W+jets, the $t\bar{t}$ background prediction in the signal region is connected to the yields of $t\bar{t} \rightarrow \mu\nu+2b+\text{Jets}$ and $t\bar{t} \rightarrow e\nu+2b+\text{Jets}$ events in the semi-leptonic $t\bar{t}$ control samples. The associated transfer factors account for the differences in the impact of lepton acceptance, and selection efficiencies. In the case of single electron events, the transfer factor also takes into account the difference in the trigger efficiencies as the trigger used in the signal region and single electron region are different.

The nuisances affecting the transfer factors are kept Gaussian constrained in the fit. The relative statistical uncertainty on the transfer factor measured from the sim-

ulation is also treated as Gaussian constrained nuisance in the fit.

For the estimation of QCD background, some contribution from QCD is required in each of the bin in signal region is required to extract transfer factor corresponding to QCD contribution. Unfortunately, the QCD contribution is very small and is not present in all the bins in signal region(section 6.1.3.4). To quantify the contribution, the QCD background defined as the difference of data and total MC background i.e. $QCD = DATA - Total\ MC\ background$ is plotted against $\min(\Delta\phi)$ cut i.e. for each entry of $\min(\Delta\phi)$ in the y-axis will correspond to the number of events passing the $\min(\Delta\phi)$ cut in QCD control region. The QCD contribution exponentially decreases with the increase in the $\min(\Delta\phi)$ cut as show in Figure 7.9.

For fitting the contribution curve is fitter with an exponential function with two parameters, and has been used as equation of the form $y = a. \exp(b.x)$. The curve is fitted using the range of $\min(\Delta\phi)$ values going from 0 to 0.5 till the QCD control region. The value of the QCD contribution in the signal region (region with $\min(\Delta\phi) > 0.5$ in the Figure 7.9) are calculated using the fitted curve and the total error is calculated by varying each parameter individually by 1σ and then added both the errors in quadrature.

The value of contribution in each bin of signal region is given in the Table 7.1. As from the value from the table it is clear that the total contribution in each of the bin is $< 0.5\%$ and hence no uncertainty on the QCD MC background is used in the final fit for this analysis.

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

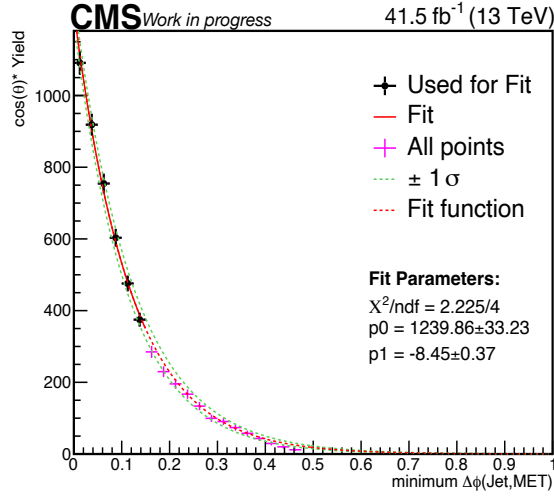


Figure 7.9: Extrapolating QCD background from the QCD control region to signal region using an exponential function. The contribution after the $\Delta\phi > 0.5$ is very small compared to the total MC background in the signal region

	QCD Contribution	Fraction w.r.t Total MC in SR(%)
Bin1	0.1135 ± 0.0535	0.022
Bin2	1.3437 ± 0.7013	0.304
Bin3	0.0347 ± 0.0184	0.010
Bin4	0.0496 ± 0.0330	0.022

Table 7.1: Contribution and fraction w.r.t total MC background of QCD background in each bin of $\cos \theta^*$ in signal region

7.2 Systematic uncertainties

The background and signal predictions are affected by systematic uncertainties taken into account in the signal extraction fit. This section includes the description of the sources of systematic uncertainties used for the MC simulation processes and the transfer factors.

7.2.1 Systematic Uncertainties on the background estimated using simulation

The contribution of diboson, single top, QCD and SM Higgs boson backgrounds are taken from the simulation in both 1b and 2b categories. Therefore following set of uncertainties affect these processes. Most of these sources are treated as shape unless specified. Whenever the shape of the nuisance is used, average % over the four bins is calculated and reported in the table for the shape systematic uncertainty to give an approximate size of the nuisance used in the prefit.

In addition to the above, $t\bar{t}$ in the 1b category and W+jets in the 2b category are also estimated using simulation and hence are affected by these nuisances.

7.2.1.1 Luminosity

All non-data driven backgrounds that are mentioned above are assigned with uncertainty in the luminosity measurement of 2.5% for 2016 and 2018, and 2.3% (log-normal) for 2017 data taking as recommended by the Lumi POG [87–89]. This uncertainty affects diboson, single top, QCD and SM Higgs boson in both 1b and 2b

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

categories. In addition, it affects the $t\bar{t}$ in the 1b category and W +jets in the 2b category.

7.2.1.2 The QCD multijet background

The contribution of QCD in signal region comes very low ($< 0.5\%$), therefore, no uncertainty on the QCD MC background is used in the final fit. The details of the extraction of the QCD contribution is discussed in section [7.1](#)

7.2.1.3 MET Trigger

As described earlier in section [4.2](#) the trigger efficiency scale factors are measured in 1, and 2 lepton phase space and the difference in the scale factors taken as the systematic uncertainty. A shape-based systematic uncertainty as a function of the recoil variable is applied to all the simulation samples in the signal and control regions.

7.2.1.4 Electron Trigger

The electron trigger efficiency scale factors are provided by the EGM POG for all three years, along with the systematic uncertainty associated with them. The uncertainty assigned on average is of the order of 1% [[59](#)].

7.2.1.5 Jet Energy Calibration

11 source shape uncertainty is used as recommended by JME POG [[90](#)]. The shape uncertainty is applied to all the simulation-based processes in all signal and control

7.2. SYSTEMATIC UNCERTAINTIES

regions. For each of the sources, the up and down variation is extracted as a function of Recoil in 1b and $\cos\theta^*$ in 2b category, respectively. The average effect is of the order of 2-3%.

7.2.1.6 Pile-Up

The minimum bias cross-section to calculate the pileup re-weight scale factors are derived using 69.2 mb [91]. The uncertainty on the minimum bias cross-section of 4.6% [91] is used to re-calculate the pileup re-weight by adding and subtracting 4.6% from the central value of the cross-section. The variation is propagated to the observable being fit here by using alternate pileup re-weight scale factors and used as shape uncertainty.

7.2.1.7 Cross-section uncertainty

The processes which are measured using simulation samples are assigned with respective uncertainty in the cross-section. They mainly include single top, diboson and standard model Higgs boson background. An uncertainty of 20% is assigned to single t and diboson background yield as reported by CMS in ref [92] and [93, 94]. The systematic uncertainty for each background process is correlated between SR and CRs and across the years.

7.2.1.8 QCD renormalization and factorization scale

The systematic uncertainty in the SM Higgs boson is due to the variation in the renormalization and factorization scale. This is done for both signal and SM h back-

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

muR	muF
1.0	1.0
1.0	2.0
1.0	0.5
2.0	1.0
2.0	2.0
0.5	1.0
0.5	0.5

Table 7.2: QCD scales variations

ground. The recoil/ $\cos\theta^*$ variable is reweighed using the weights for the case when renormalization and factorization scale are multiplied or divided by a factor of 2; all the combinations are given in Table 7.2. The envelope of the distributions is treated as the shape uncertainty. For the signal, the only shape of the observable is allowed to change for the and effect on the normalization is neglected (i.e. the normalization of up and down templates are forced to be the same as the central).

7.2.1.9 PDF variation

The alternate set of PDF weights variation (set of 101 weights) is used to derive the pdf systematic variation. The nuisance (at the moment) affect the SM h and signal process in a correlated way. The recoil/ $\cos\theta^*$ variable is reweighed using the alternate set of PDF weights, and the envelope of the distributions are treated as shape uncertainty. As done for the QCD scale uncertainty, for the signal process, only shape of the observable is allowed to change for the and effect on the normalization is neglected (i.e. the normalization of up and down templates are forced to be the same as the central)

7.2. SYSTEMATIC UNCERTAINTIES

A summary of all systematics is listed in table 7.3 and all the plots showing up and down systematics for all the analysis regions are shown in Appendix D.

	Region(CR/SR)	Processes	Size
Pile up reweighting	All	all except bkg_{fit}	shape(3%)
b tag efficiency	All	signal, all except bkg_{fit}	shape(3-8%)
fake b	All	all except bkg_{fit}	shape(1-7%)
electron Identification	Electron	all except bkg_{fit}	shape(5-7%)
electron Reconstruction efficiency	Electron	all except bkg_{fit}	shape(3-5%)
muon Identification	Muon	all except bkg_{fit}	shape(<1.%)
muon isolation	Muon	all except bkg_{fit}	shape(<1.%)
renormalization and factorization scale	All	all except bkg_{fit}	shape(16-42%)
pdf	All	all except bkg_{fit}	shape(25-50%)
electron trigger	Electron	all except bkg_{fit}	shape(1.%)
MET trigger	Muon & Signal	all except bkg_{fit}	shape(1%)
JECAbsolute	All	all except bkg_{fit}	shape(2.%)
JECAbsolute_YEAR	All	all except bkg_{fit}	shape(1.%)
JECBBEC1	All	all except bkg_{fit}	shape(1.%)
JECBBEC1_YEAR	All	all except bkg_{fit}	shape(0.5%)
JECEC2	All	all except bkg_{fit}	shape(0.09%)
JECEC2_YEAR	All	all except bkg_{fit}	shape(0.5%)
JECFlavorQCD	All	all except bkg_{fit}	shape(1-3%)
JECHF	All	all except bkg_{fit}	shape(0.3%)
JECHF_YEAR	All	all except bkg_{fit}	shape(0.2%)
JECRelativeBal	All	all except bkg_{fit}	shape(0.7%)
JECRelativeSample_YEAR	All	all except bkg_{fit}	shape(0.6%)
tau veto	All	all except bkg_{fit}	shape
Luminosity	All	all except bkg_{fit}	2.5%
*cross-section VV	All	only DIBOSON	lnN 20%
*cross-section Higgs boson	All	only SMH	lnN 20%
*cross-section Single Top	All	only SingleTop	lnN 10%

Table 7.3: Systematic uncertainties and their effect on the analysis. Uncertainties in red colour are uncorrelated throughout the three years

7.2.2 Systematic uncertainty on the Transfer Factor

7.2.2.1 MET Trigger

The MET trigger, used to trigger the events in the signal and muon control region, is $>95\%$ efficient above 250 GeV and $>97\%$ efficient above 350 GeV. The efficiency is measured in data and MC using W and Z (muons) events. The difference in the measurement in the single muon and dimuon region is assigned as the systematic uncertainty on the MET trigger efficiency. Recoil dependent (shape) uncertainty is applied to all the transfer factors. The uncertainty on the TF is calculated by varying the SR template used to measure the TF up and down by 1σ from the central value with all the bins correlated and remeasuring the TF for up and down, respectively. The relative % change in the TF in each bin is then treated as the shape uncertainty on the TF and treated as gaussian constrained. This nuisance affect the 1 electron and 2 electron control regions as they don't use the MET based triggers.

7.2.2.2 Electron Trigger

The electron triggers are used for the event triggering in the W and Z control regions. The scale factors to account for the difference in the data and simulation are used as recommended by the EGM POG [58]. The uncertainty on the scale factors is propagated to the per bin transfer factors using the same procedure as described for MET Trigger. This uncertainty affects the 1 electron and 2 electron control regions as they use different triggers w.r.t the signal region.

7.2. SYSTEMATIC UNCERTAINTIES

7.2.2.3 Muon reconstruction and identification efficiency

One of the differences between signal region and $W(\mu\nu)(Z(\mu^+\mu^-))$ control regions is the number of reconstructed and identified muons. The signal region requires exactly zero muons; however, $W(\mu\nu)(Z(\mu^+\mu^-))$ is required to have exactly one (two) muon(s). Therefore the systematic uncertainties due to muon reconstruction and muon identification are taken into account as recommended by the MUON POG [68–72] and propagated to the transfer factors involving $W(\mu\nu)(Z(\mu^+\mu^-))$ control regions. In order to propagate the uncertainty due to muon reco and ID the control region templates are varied by $\pm 1 \sigma$ with all bins correlated, and two more sets of TF are extracted. The % change in the TF is then treated as the nuisance, which is gaussian constrained during the signal extraction fit.

7.2.2.4 Electron reconstruction and identification efficiency

Similar to the muon part explained above, the difference in SR and $W(ev)(Z(e^+e^-))$ CR is the number of reconstructed and identified electrons. SR requires exactly zero electrons, and $W(ev)(Z(e^+e^-))$ CR is required to have exactly one (two) electron(s). Therefore the systematic uncertainties due to electron reconstruction and electron identification are taken into account as recommended by the EGM POG [73, 74] and propagated to the transfer factors involving $W(ev), t\bar{t}(e) Z(e^+e^-)$ control regions. The uncertainty on the scale factors is propagated to the per bin transfer factors using the same procedure as described for MET Trigger.

CHAPTER 7. SIGNAL EXTRACTION AND SYSTEMATICS UNCERTAINTIES

7.2.2.5 b-tagging

As described in the event selection that for $Z \rightarrow ll + \text{jets}$ control regions, the b-tagging requirement has been removed on the jets to increase the statistics in these regions. Therefore, an additional nuisance is added on the TF to take into this account this change in the selection w.r.t the signal region. The % change in the TF is calculated by scaling the b-tagging scale factor by $\pm 1 \sigma$ and deriving the TF for up and down cases, respectively (as done for other nuisances).

Table 7.4 lists all the sources of nuisances affecting the TFs and their average % value. All these nuisances are kept gaussian constraint along with statistical uncertainty on the TF, which is estimated in a similar way by scaling the numerator and denominator templates by $\pm 1 \sigma$ of the statistical uncertainty.

7.2.2.6 Pre-firing

For 2016 and 2017, the effect of pre-firing is propagated to all the TFs. The uncertainty on the TF is obtained by varying the SR and CR templates simultaneously by $\pm 1 \sigma$ correlated for four bins.

The distribution of the uncertainties on the transfer factors is shown in Figure 7.5, 7.6, 7.7 and 7.8

7.2. SYSTEMATIC UNCERTAINTIES

Source	Process	Size
Electron reconstruction efficiency	$W_{SR}/W_{e\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{ee}, t\bar{t}_{SR}/t\bar{t}_{\text{TopeCR}}$	1%
Electron ID efficiency	$W_{SR}/W_{e\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{ee}, t\bar{t}_{SR}/t\bar{t}_{\text{TopeCR}}$	3%
Electron Trigger efficiency	$W_{SR}/W_{e\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{ee}, t\bar{t}_{SR}/t\bar{t}_{\text{TopeCR}}$	2%
Muon tracking efficiency (2016 only)	$W_{SR}/W_{\mu\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}, t\bar{t}_{SR}/t\bar{t}_{\text{Top}\mu\text{CR}}$	< 1%
Muon ID efficiency	$W_{SR}/W_{\mu\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}, t\bar{t}_{SR}/t\bar{t}_{\text{Top}\mu\text{CR}}$	1%
Muon isolation efficiency	$W_{SR}/W_{\mu\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}, t\bar{t}_{SR}/t\bar{t}_{\text{Top}\mu\text{CR}}$	1%
b-tagging	$Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$	1%
Pre-firing	$W_{SR}/W_{e\nu}, Z(\rightarrow \nu\nu)_{SR}/Z_{ee}, t\bar{t}_{SR}/t\bar{t}_{\text{TopeCR}}$	0.05%
renormalization and factorization scale	$Z(\rightarrow \nu\nu)_{SR}/Z_{ee}, Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$	< 2%
pdf	$Z(\rightarrow \nu\nu)_{SR}/Z_{ee}, Z(\rightarrow \nu\nu)_{SR}/Z_{\mu\mu}$	<1%

Table 7.4: Source of experimental systematic uncertainties effecting the transfer factors (to measure Z , W and $t\bar{t}$ background) along with their approximate size. Note that the W TF are used only for the 1b category and $t\bar{t}$ related TF are used only for the 2b category.

CHAPTER 8

RESULTS

8.1 Results

In the analysis, the signal is extracted from a simultaneous fit of SR and CRs in 1b and 2b categories. In the 1b category sensitive observable is p_T^{miss} and $\cos(\theta)^*$ (define in 6.1.4) is used for 2b category. The upper limits are calculated at 95% confidence level (CL) using a modified frequentist method(CLs) computed with an asymptotic approximation [95]. A binned profile likelihood ratio is used as the test statistic in which systematic uncertainties are modelled as nuisance parameters. During the simultaneous fit, the major backgrounds $Z(\nu\nu) + \text{jets}$ and $W + \text{jets}$ in 1b and $Z(\nu\nu) + \text{jets}$ and $t\bar{t}$ in 2b categories are estimated on fly using binned transfer factors measured from simulation but allowed to vary by the gaussian constrained systematic uncertainties which act as small perturbation to change the size of the

CHAPTER 8. RESULTS

transfer factor in each of the bins. In the following section,

- checks performed for the final fit model, pulls and impact distributions.
- prefit and postfit comparison for the 1b and 2b categories for individual years.
- 95% CLs upper limit on the signal strength, μ .

8.1.1 Pulls and Impacts

Figure 8.1, 8.2 and 8.3 shows the pulls with data, control region only (i.e. signal region is masked) and the distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2016.

Figure 8.4, 8.5 and 8.6 shows the pulls with data, control region only (i.e. signal region is masked) and the distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2017.

Figure 8.7, 8.8 and 8.9 shows the pulls with data, control region only (i.e. signal region is masked) and the distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2018.

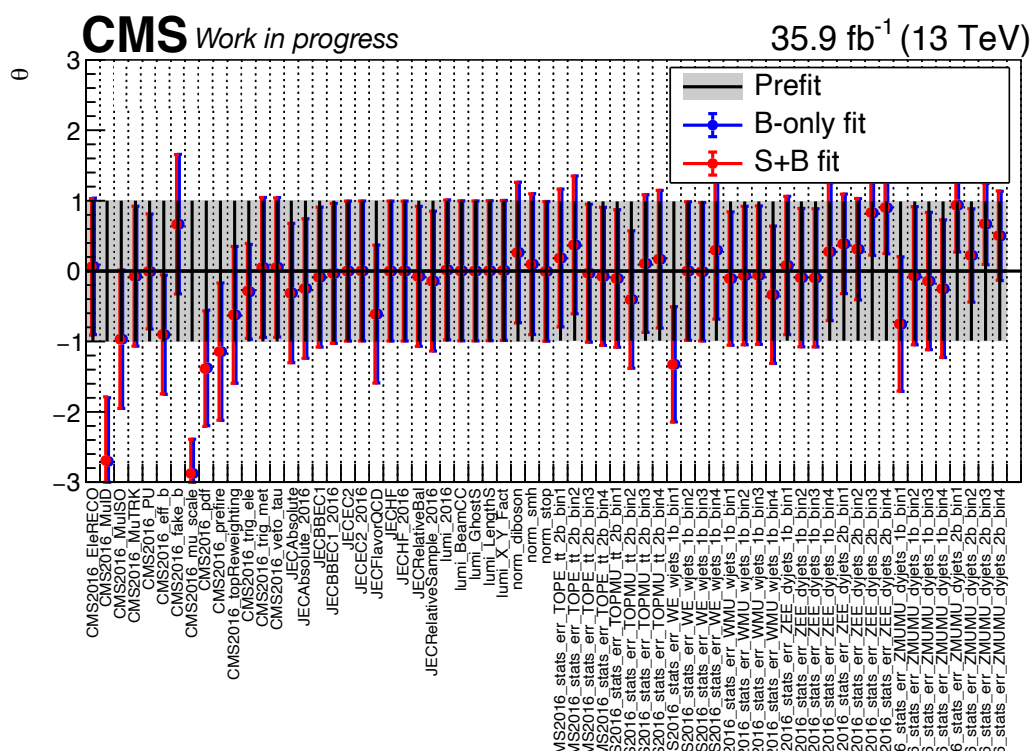


Figure 8.1: Distribution of the pulls with data. Combination of 1b and 2b for full 2016 dataset is used for the fit.

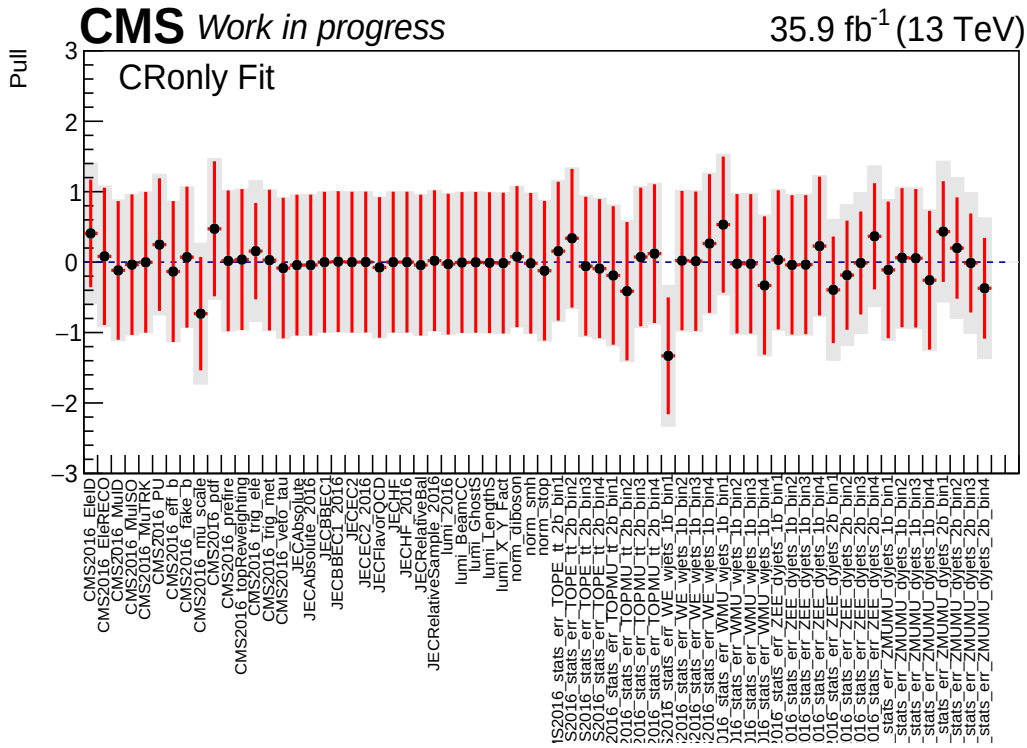


Figure 8.2: Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for 2016 dataset. Combination of 1b and 2b is used for the fit.

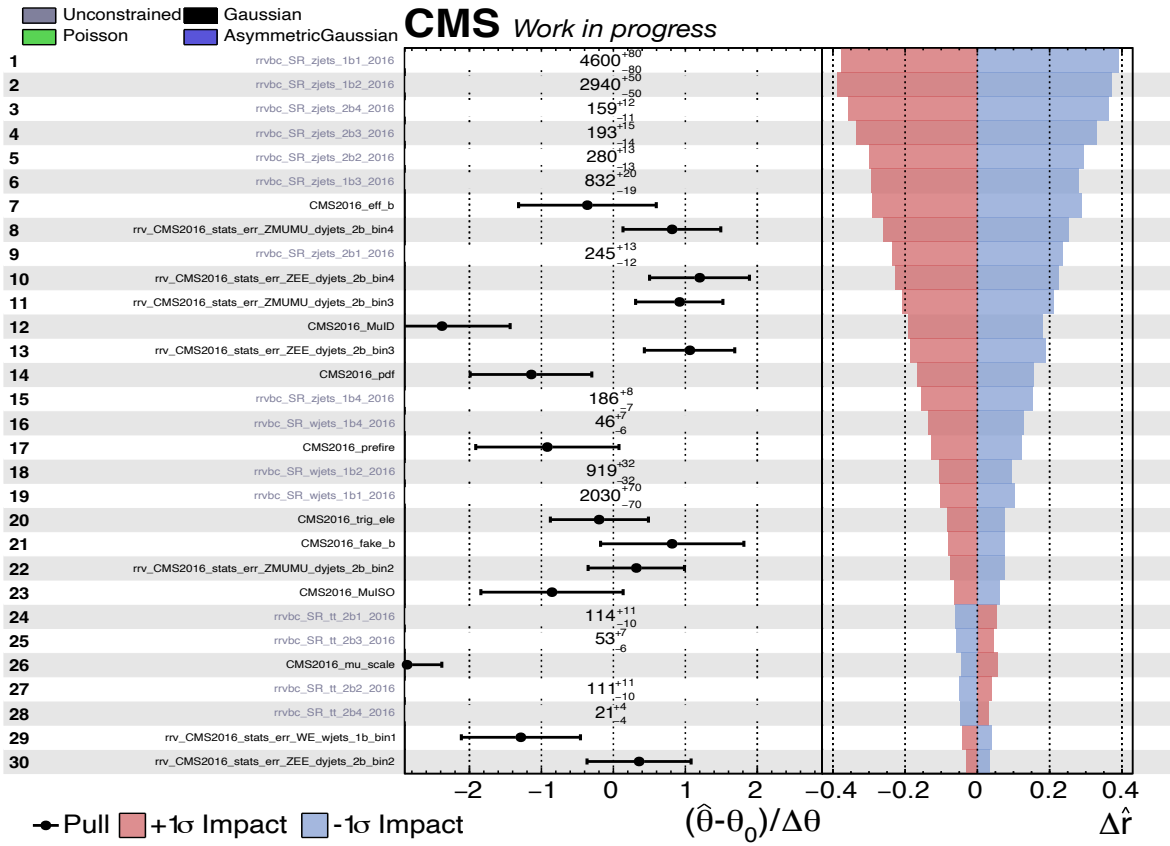
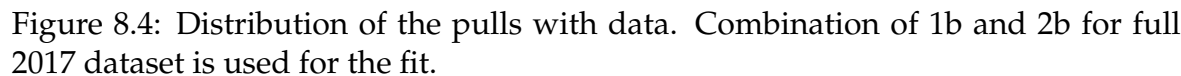


Figure 8.3: Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2016.



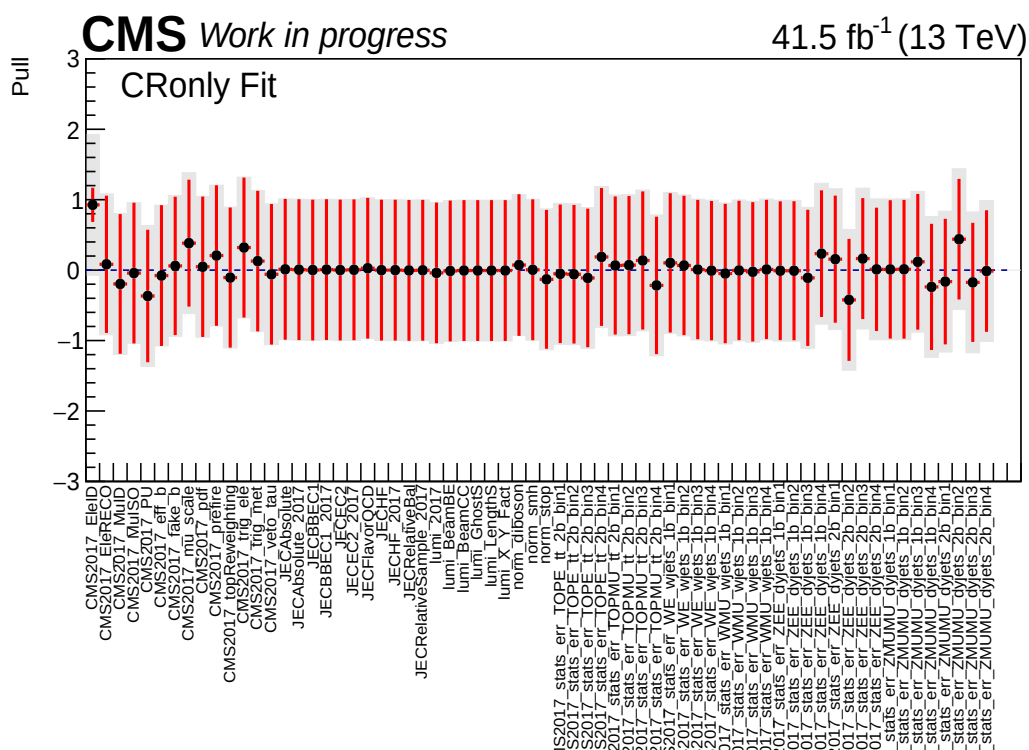


Figure 8.5: Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for 2017 dataset. Combination of 1b and 2b is used for the fit.

CHAPTER 8. RESULTS

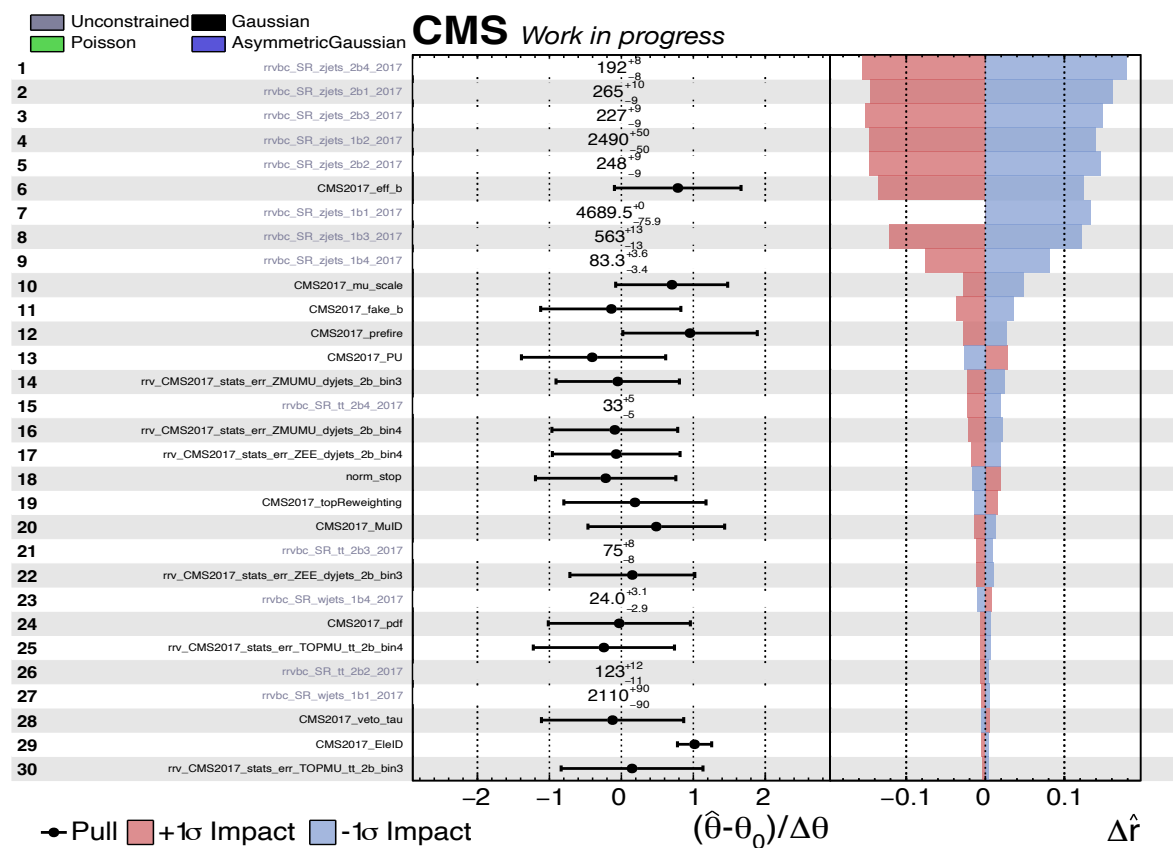
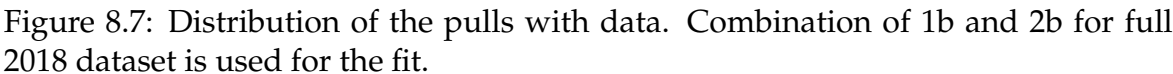


Figure 8.6: Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2017.



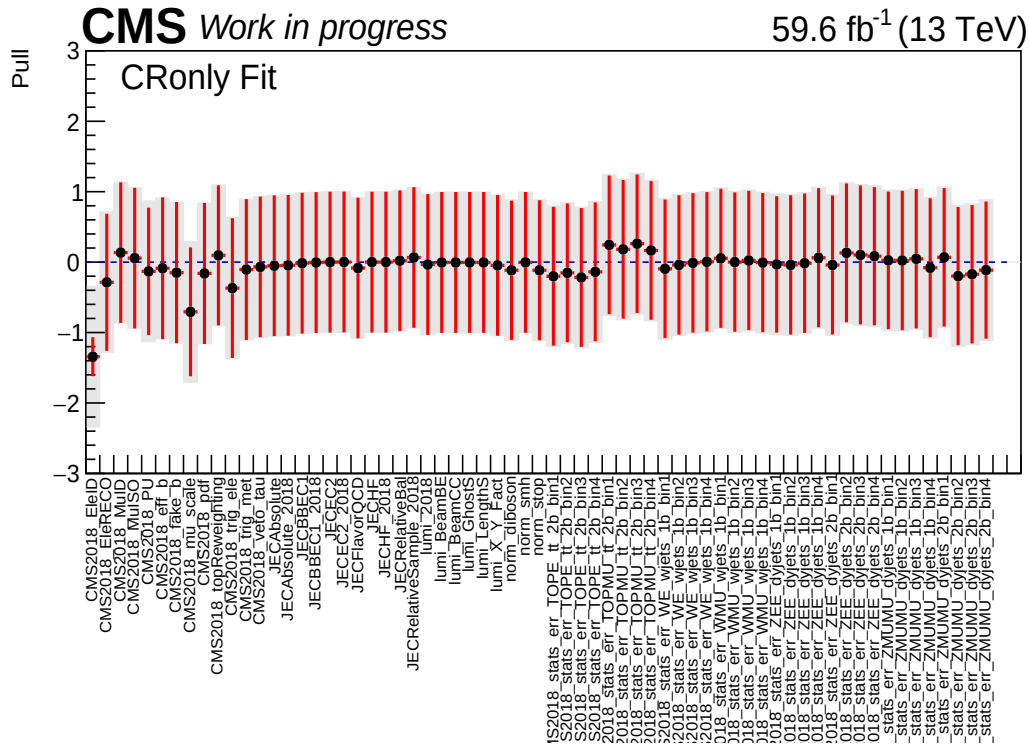


Figure 8.8: Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for 2018 dataset. Combination of 1b and 2b is used for the fit.

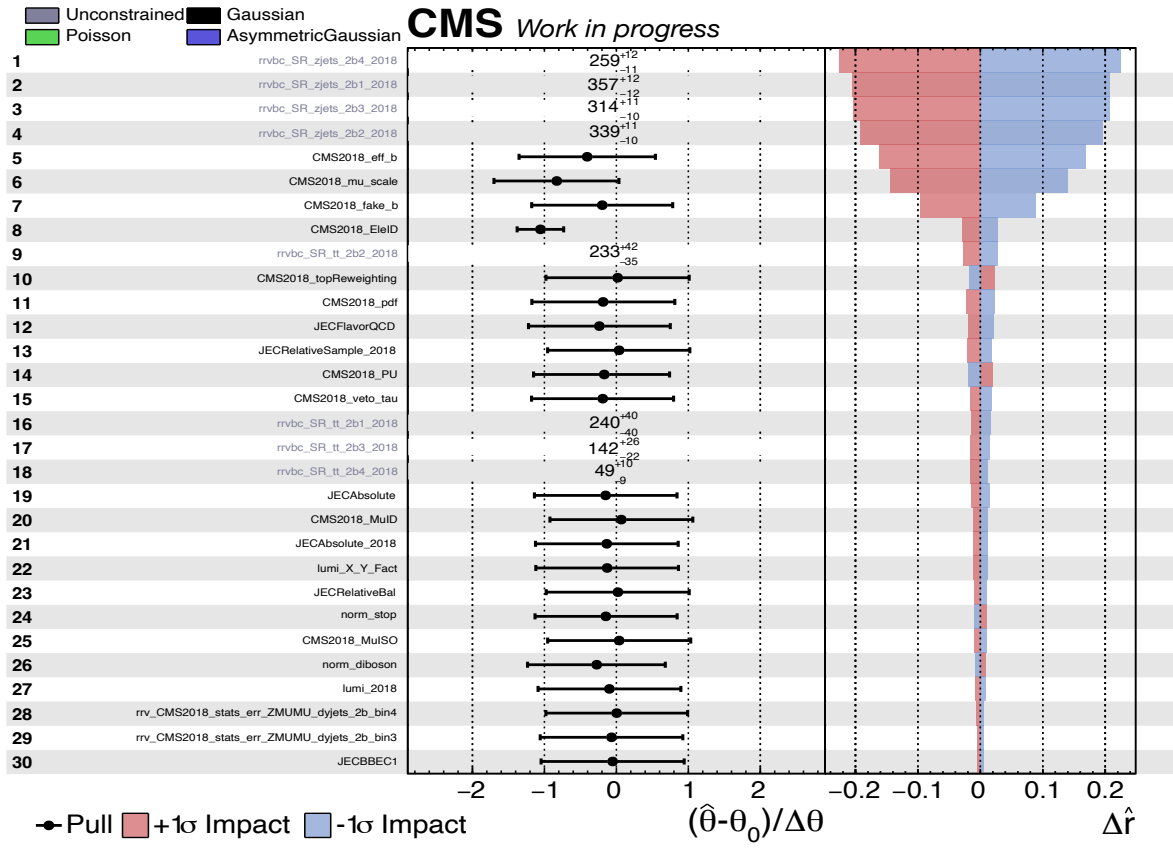


Figure 8.9: Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for 2018.

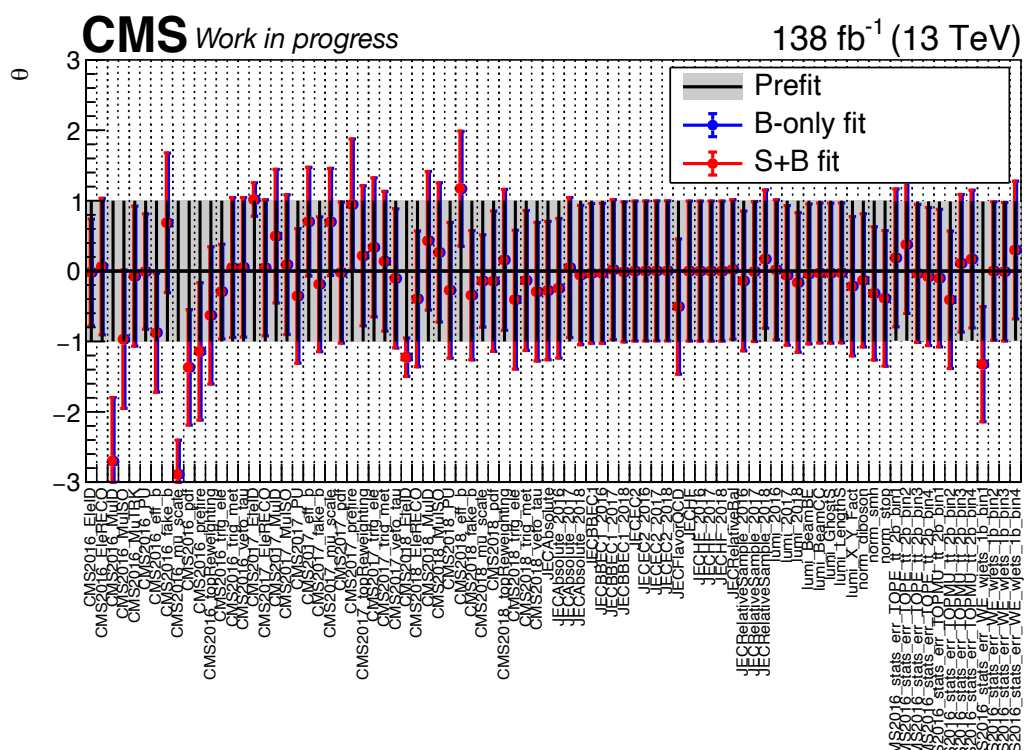


Figure 8.10: Distribution of the pulls with data. Combination of 1b and 2b for full run2 dataset is used for the fit.

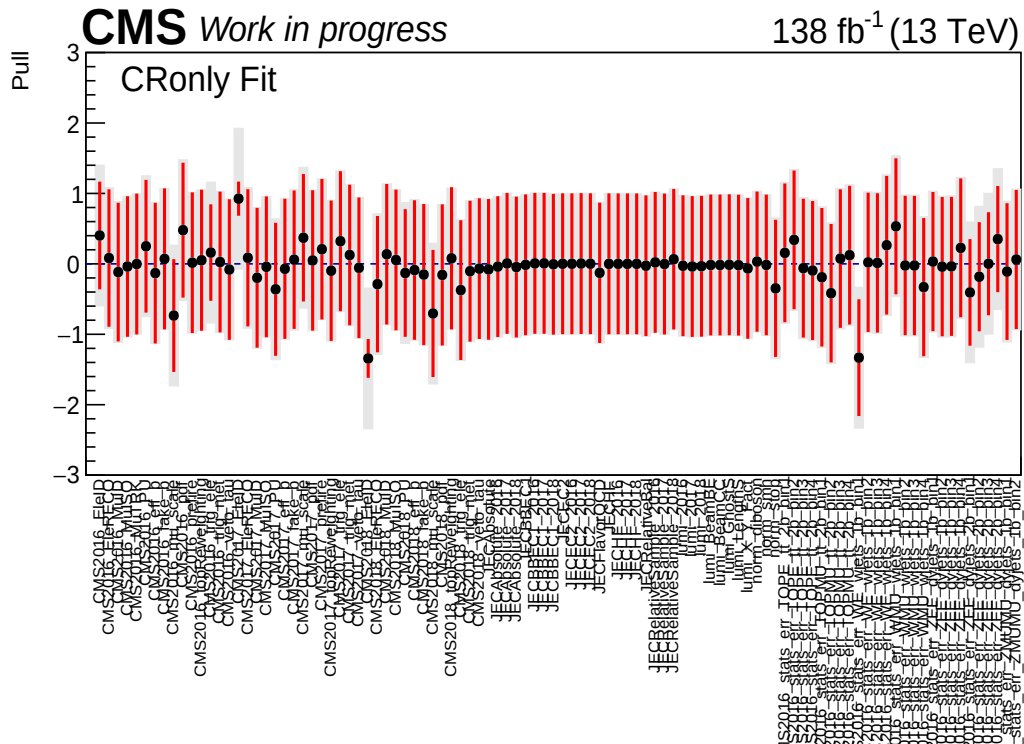


Figure 8.11: Distribution of the pulls when signal region is masked, masked region include 1b and 2b SR for full run2 dataset. Combination of 1b and 2b is used for the fit.

CHAPTER 8. RESULTS

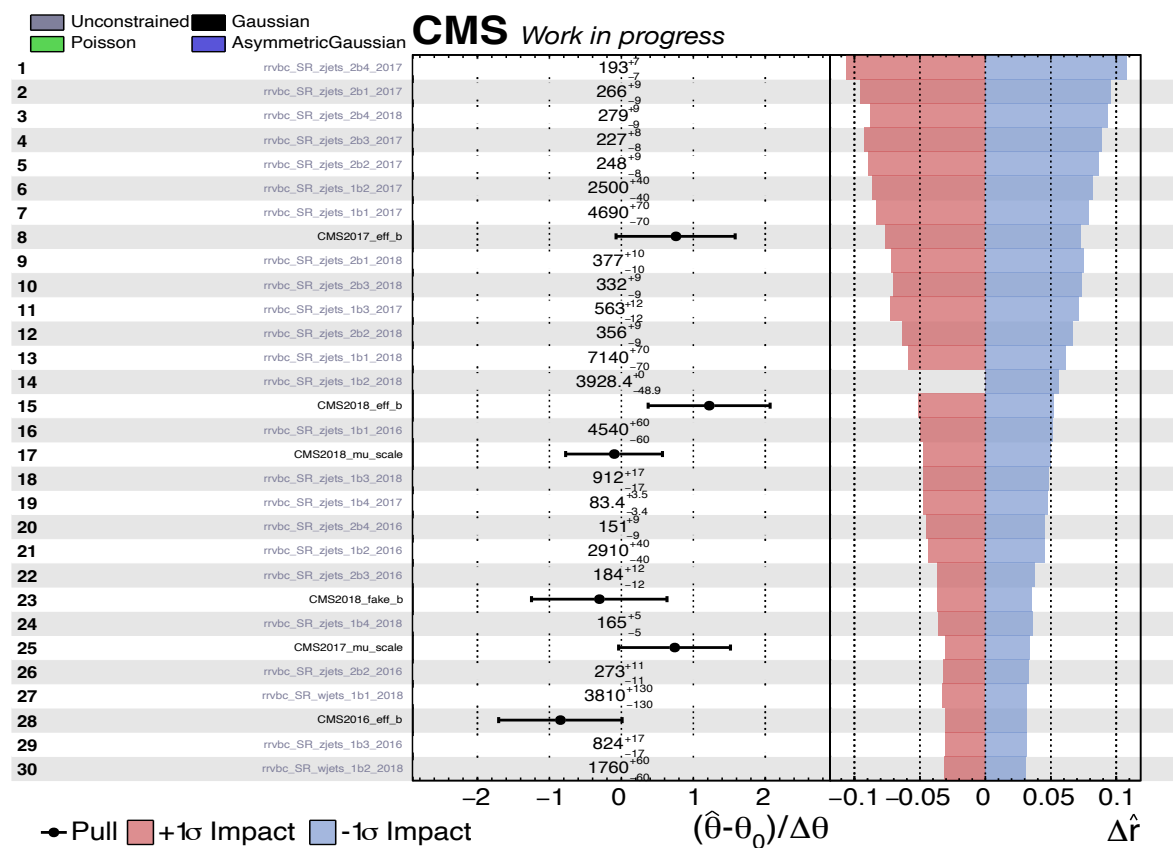


Figure 8.12: Distribution of the impact of first 30 nuisances used in the combination of 1b and 2b category for run2.

8.1.2 Post-fit distributions

The post fit plots for all the CR for 1b and 2b categories are shown in the Figure [8.13](#) and [8.14](#). The postfit distributions are extracted by using a software package developed by CMS (called *combine* for calculating impacts, pulls and limits) while masking the signal regions. The overall agreement between data and simulation is improved after the fit.

CHAPTER 8. RESULTS

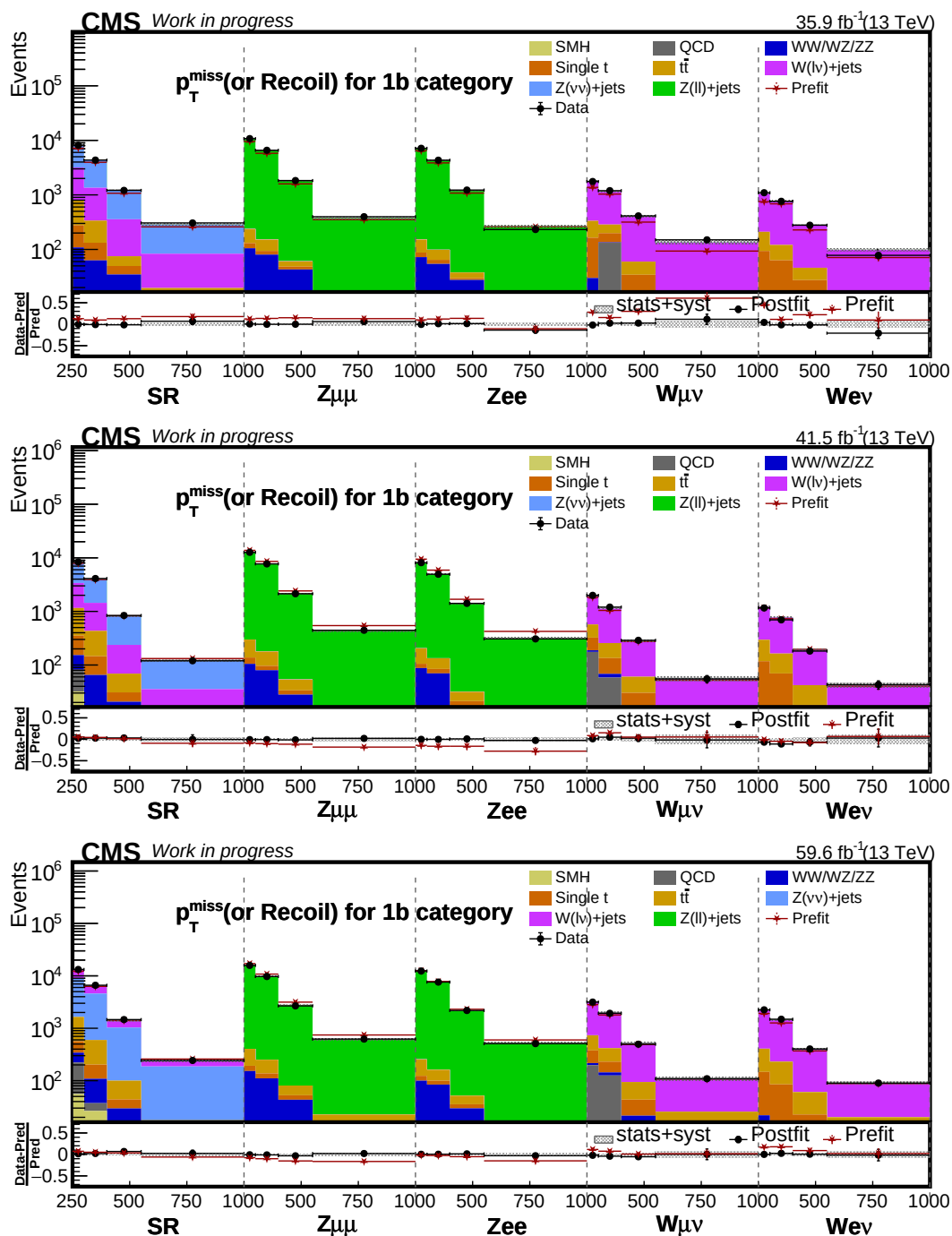


Figure 8.13: Summarizing MET for signal and Control region for 1b category for 2016, 2017 and 2018

8.1. RESULTS

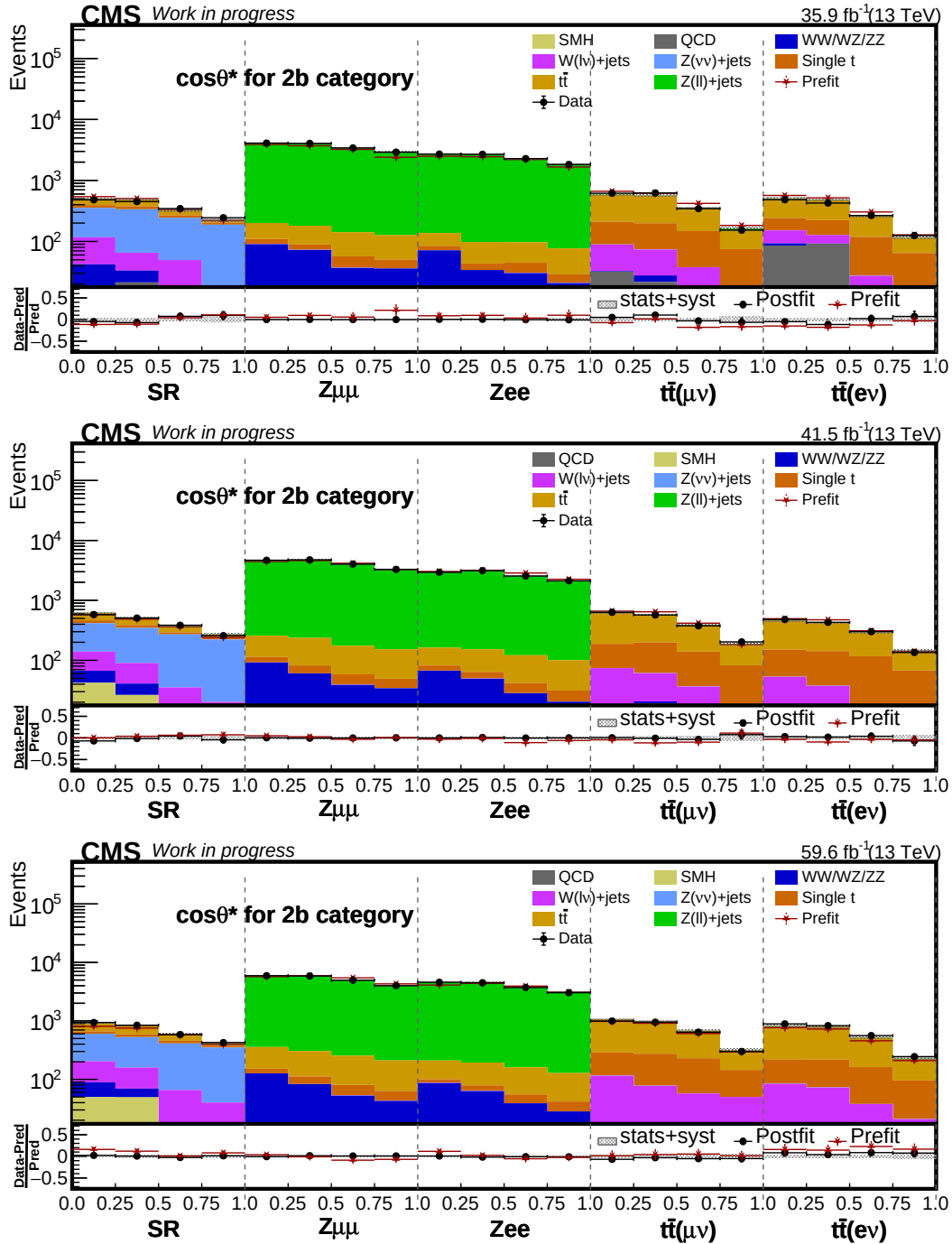


Figure 8.14: Summarizing $\cos(\theta)^*$ for signal and Control region for 2b category for 2016, 2017 and 2018

8.1.3 Limits

The combined limit for full Run 2 analysis is shown in the Figure [8.15](#). With full Run 2 limit, the observed exclusion plot ranges from $m_a = 10$ GeV to 300 GeV, which covers most of the phase space that has been scanned for this analysis.

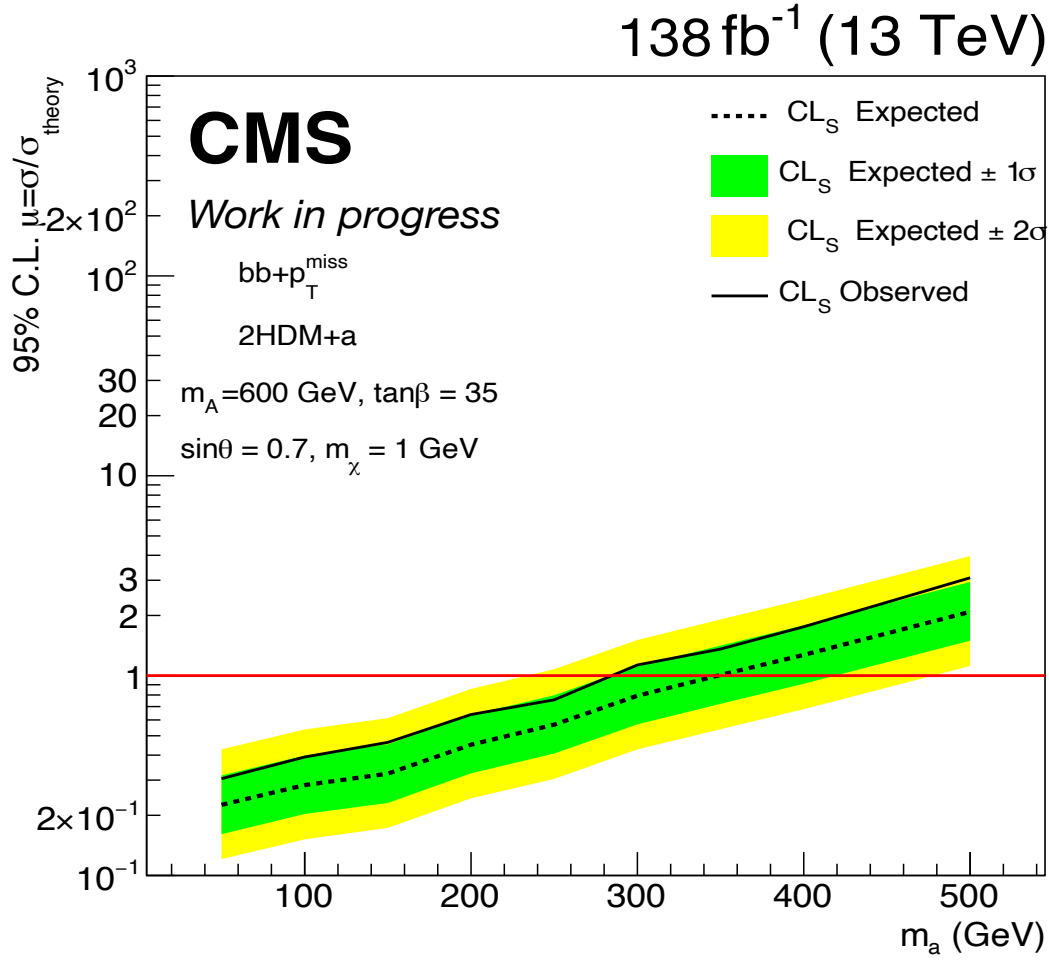


Figure 8.15: 95% CLs upper limit on the signal strength (ratio of 95% CLs cross-section of dark matter produced in association with pair of bottom quarks to the theoretical cross-section). The upper limits are shown as a function of m_a for $m_A = 600 \text{ GeV}$, $\sin\theta=0.7$ and $\tan\beta=35$ for combination of 1b and 2b category for the full run-2 dataset.

CHAPTER 9

SUMMARY AND OUTLOOK

9.1 Summary

A search for dark matter produced in association with a pair of bottom quarks is performed using full run 2 data corresponding to an integrated luminosity of 137 fb^{-1} . The analysis aims to capture the signal using two independent categories, 1b and 2b, to maximize the physics reach. A simultaneous fit of SR and CRs is performed while tying the regions via binned transfer factors, which are allowed to vary within gaussian constrained systematic uncertainties. The results are interpreted in the framework of a simplified model that is an extension of type-2 2HDM+a model where there is additional lighter pseudoscalar a that along with a heavier pseudoscalar A of the model couple to the dark matter and standard model particles. Present search performs the best in the high $\tan\beta$ phase space where the production

CHAPTER 9. SUMMARY AND OUTLOOK

cross-section for this process is enhanced. Therefore, this provides a complimentary search to constrain the 2HDM+a model parameter phase space in the region where other MET+X analyses are not very sensitive. The observed exclusion range using full run-2 dataset is from $m_a = 10$ GeV to 300 GeV, for $\tan\beta=35$, $\sin\theta=0.7$ and $m_A=600$ GeV.

9.2 Outlook

The future scope of the present analysis is listed below:

- The next straightforward step for this analysis would be to analyse the data that will be recorded in LHC Run3.
- In addition to the interpretation with the current model (2HDM+a) which has two pseudo-scalar that decay to dark matter particle, the analysis can be further interpreted with second-generation leptoquarks where pair produced leptoquarks can give MET and two jets as one of the final state.
- There are some other BSM models for e.g. R-parity violating (RPV) supersymmetry where bottom squark pair production can also give MET+ b-Jets as a final state.
- However, it will require some optimisation to interpret the analysis with different models as well as to perform the analysis with Run3 data. One of the main challenge for Run3 would be to have MET triggers with high threshold.

Appendices

A Plots for Detector Issues

A.1 Pixel DCDC failure

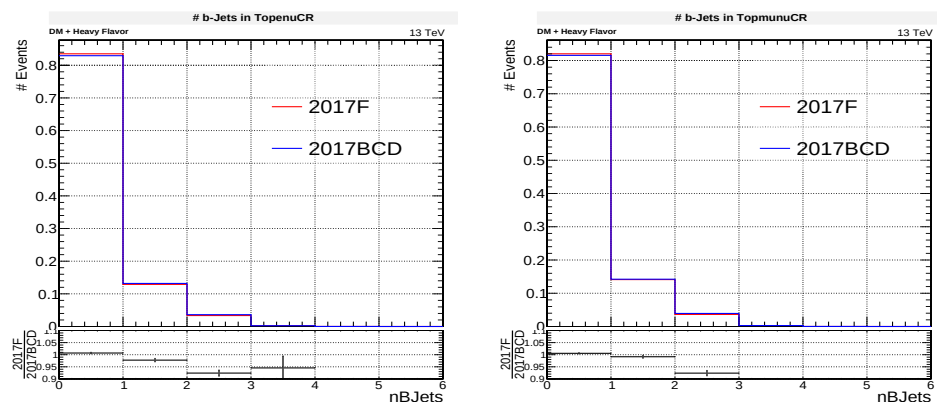
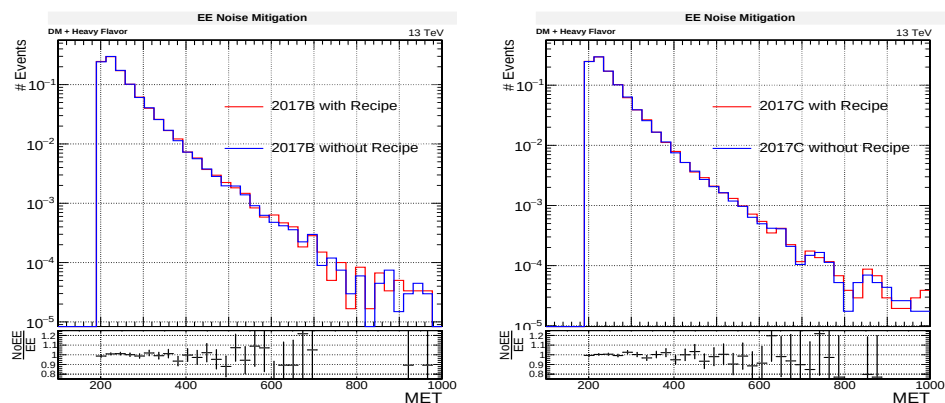


Figure 1: Comparison of nBjet distribution for pre and post pixel scenario to check the effect on the region of interest for the analysis

A.2 EE Noise mitigation



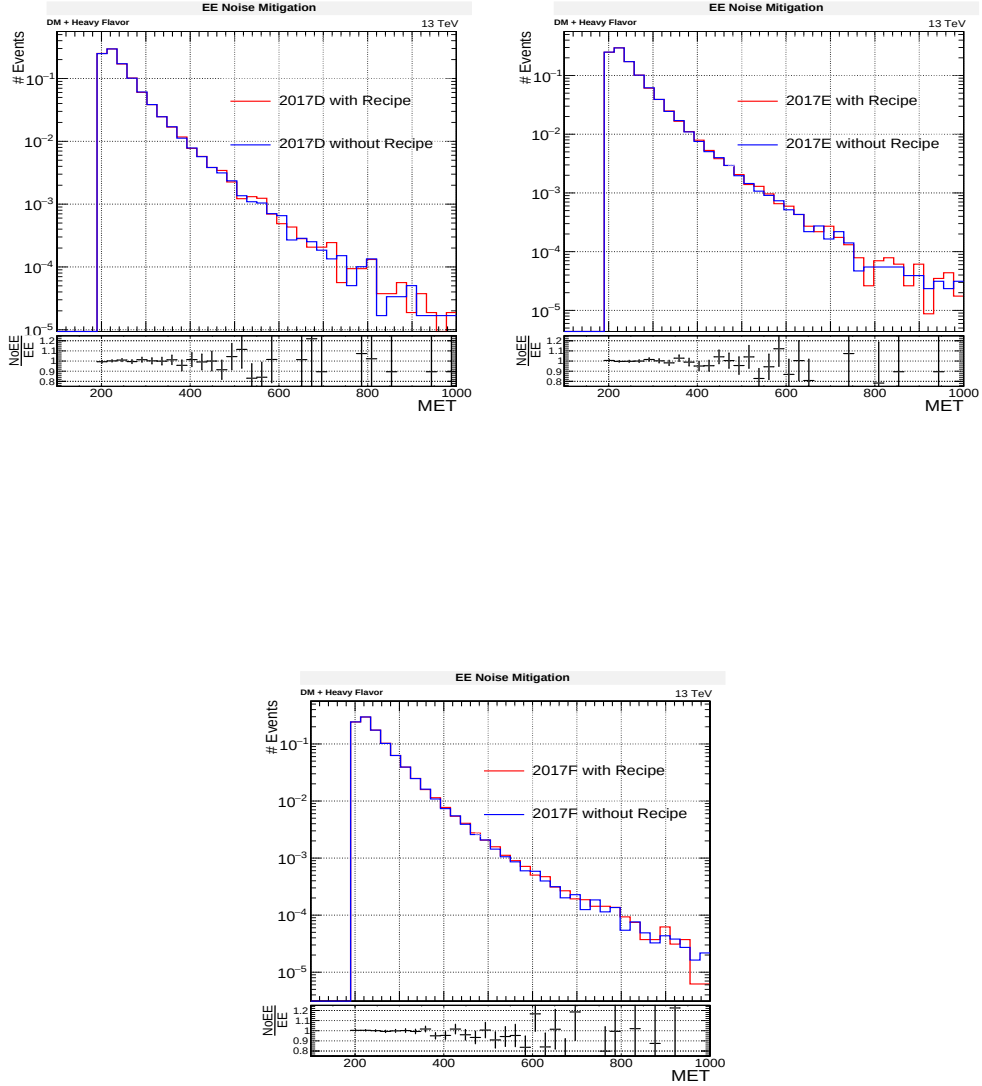


Figure 4: E_T^{miss} Comparison plots in SR1 for all the eras for 2017 data taking, the era wise plots are all normalize to unity to check shape of the distribution with and without METv2 recipe applied

A. PLOTS FOR DETECTOR ISSUES

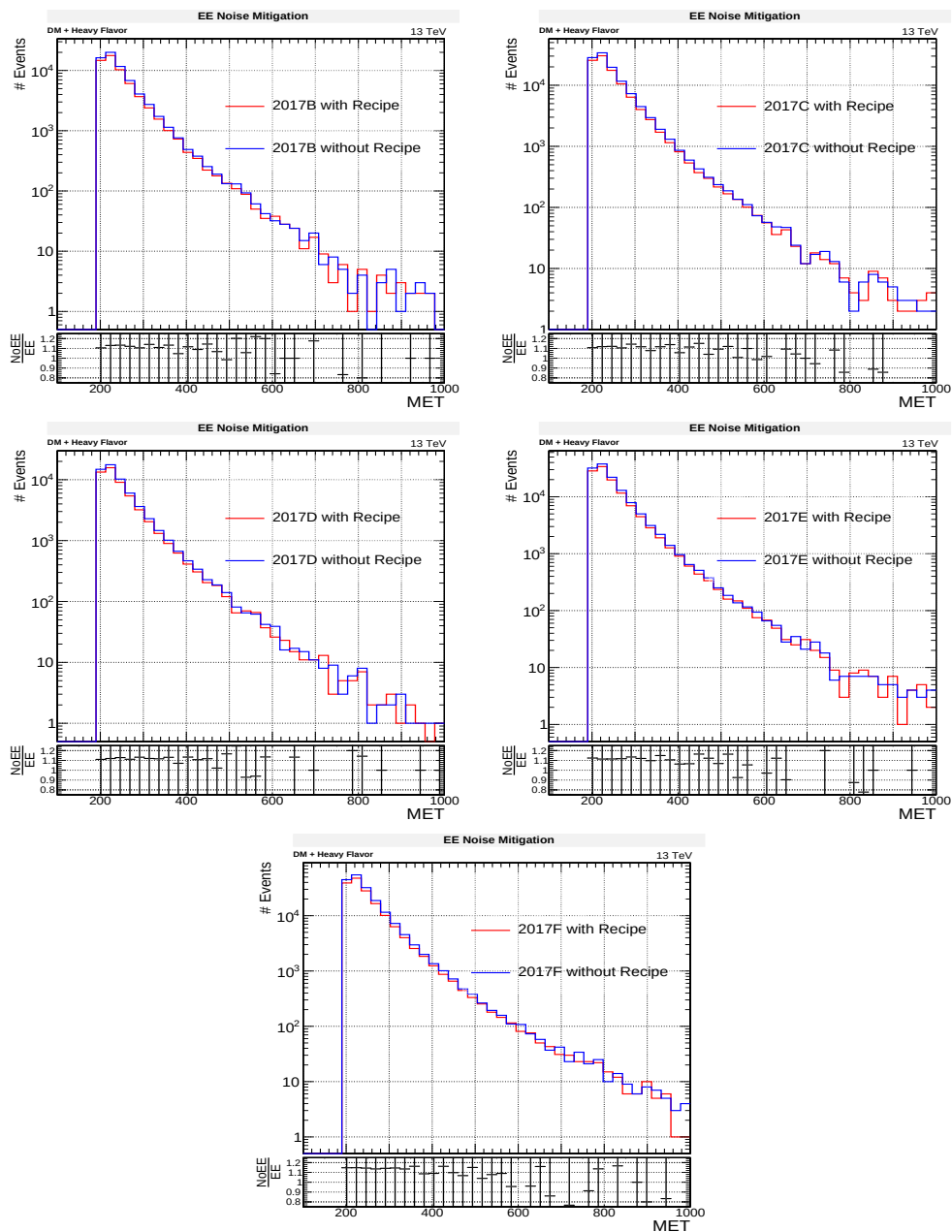


Figure 5: E_T^{miss} Comparison plots in SR1 for all the eras for 2017 data taking, the era wise plots are all without normalization to check difference in statistics of the E_T^{miss} distribution with and without METv2 recipe applied

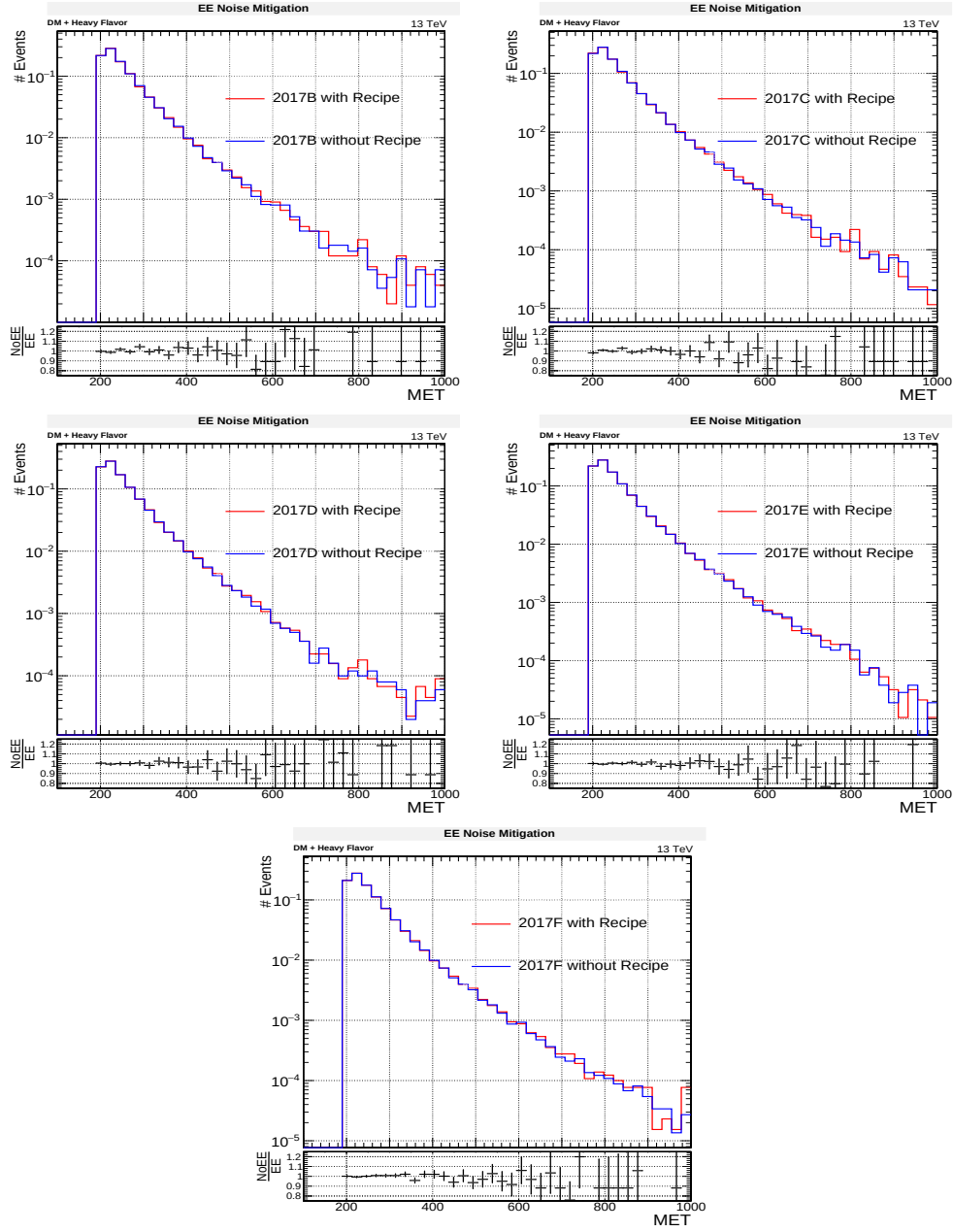


Figure 6: E_T^{miss} Comparison plots in SR2 for all the eras for 2017 data taking, the era wise plots are all normalize to unity to check shape of the distribution with and without METv2 recipe applied

A. PLOTS FOR DETECTOR ISSUES

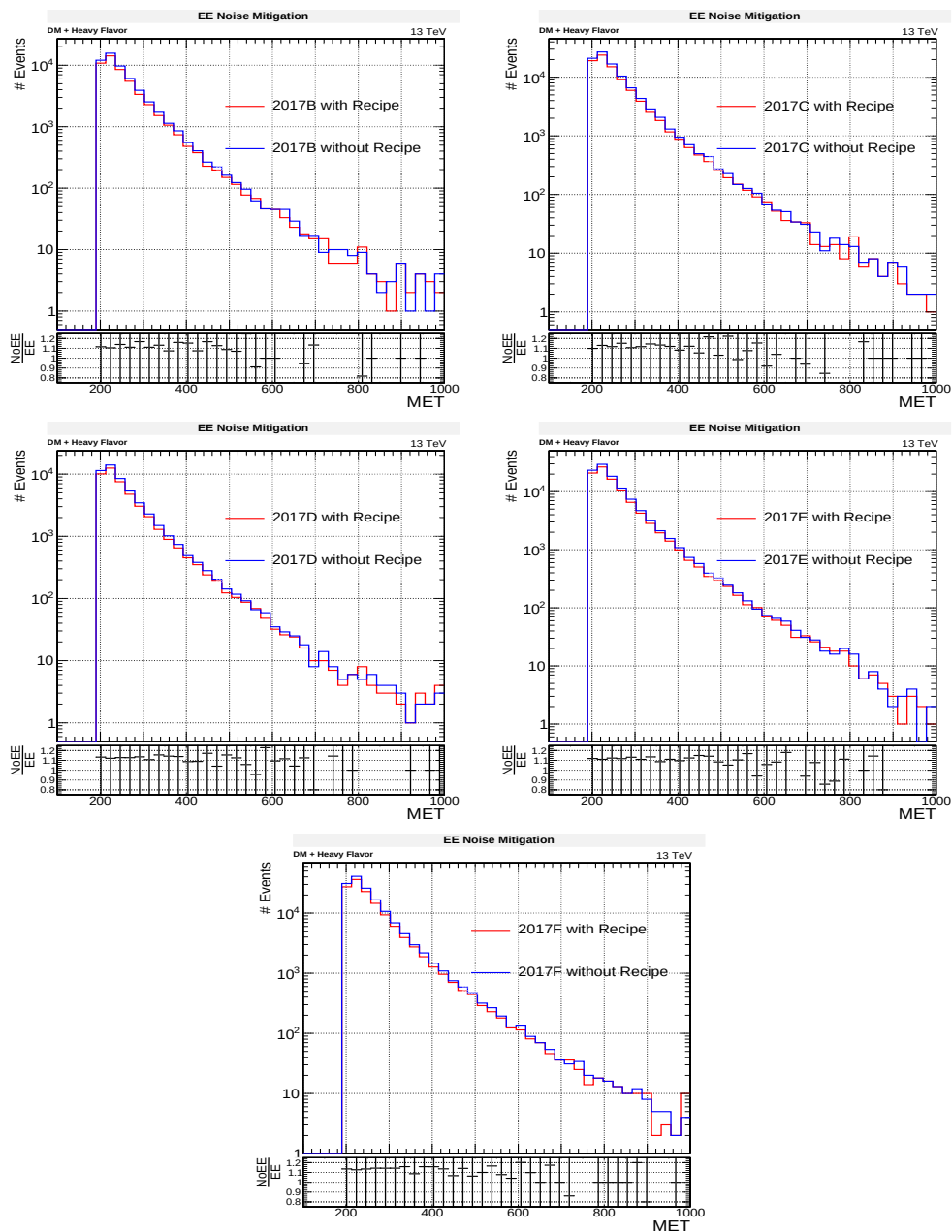


Figure 7: E_T^{miss} Comparison plots in SR2 for all the eras for 2017 data taking, the era wise plots are all without normalization to check difference in statistics of the E_T^{miss} distribution with and without METv2 recipe applied

A.3 L1 Prefiring in 2016 and 2017 data

To clearly see the effect of the uncertainties and the application of prefiring weights we plot the uncertainties as well plot without the application of prefiring weight in a single plot. By looking at the plot it becomes evident that the uncertainty associated with prefiring weight is around 0.7%, and the effect of prefiring weight on the total background in SR2 is 3.36%

In the signal region, the individual background contribution is as shown in the pie chart, the highest contributor is Z+Jets background. The effect in the SR will be the cumulative affect of the change in individual background

A. PLOTS FOR DETECTOR ISSUES

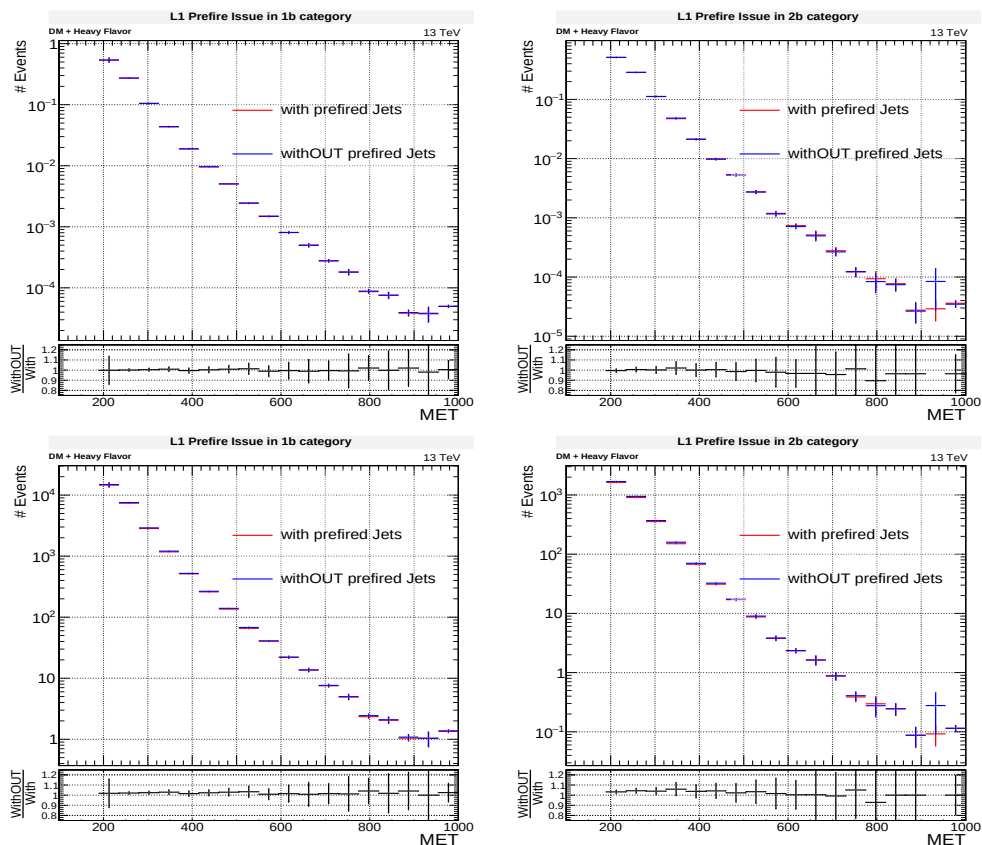


Figure 8: Comparison of E_T^{miss} distribution of total background in 1b and SR2, with normalization(in upper row) and without normalization(lower row) to check the effect of l1 prefire in 2017 MC

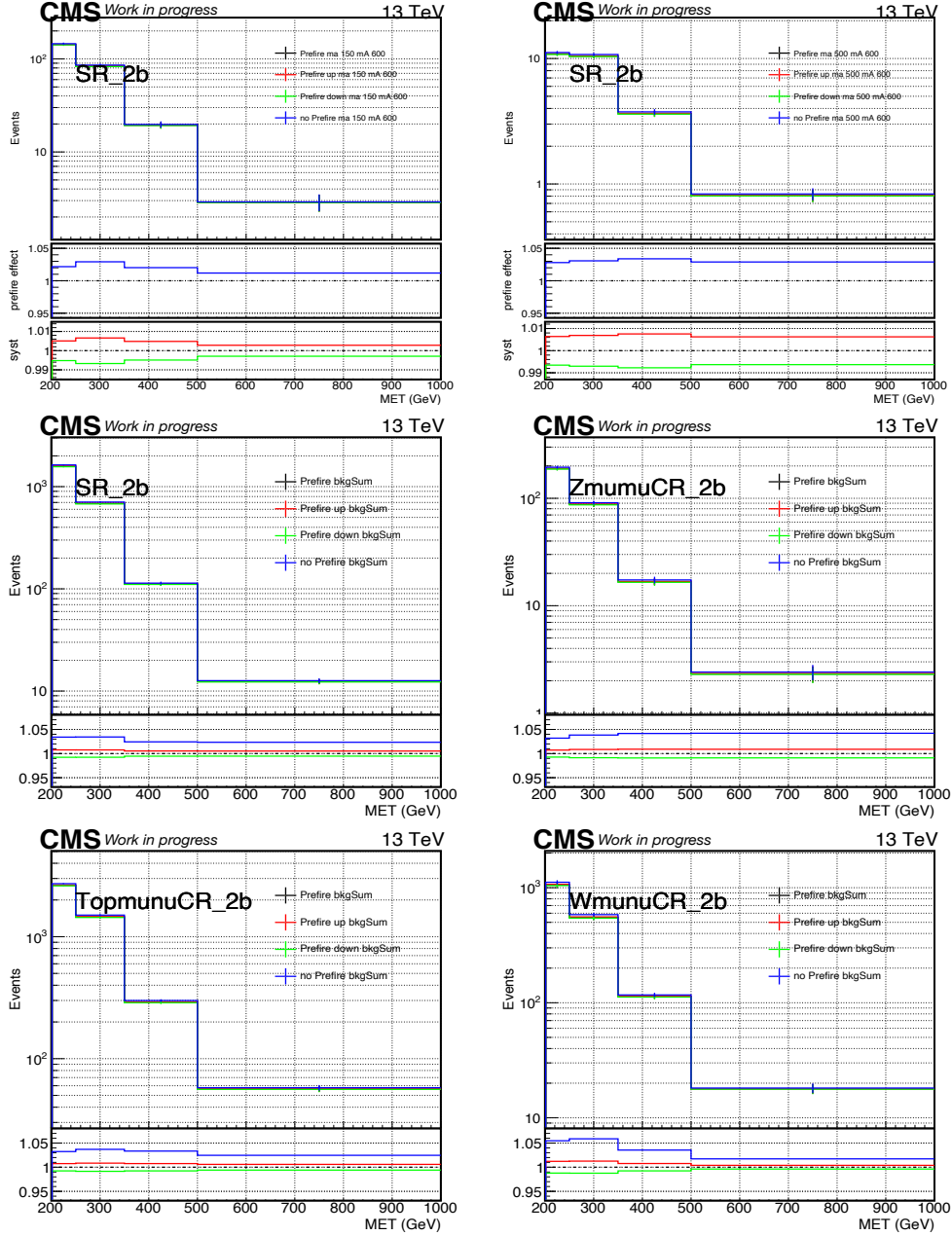


Figure 9: Recoil distribution of total background in 2b category of SR(including signal mass points) and muon CRs to check the effect of l1 prefire in 2017 MC

A. PLOTS FOR DETECTOR ISSUES

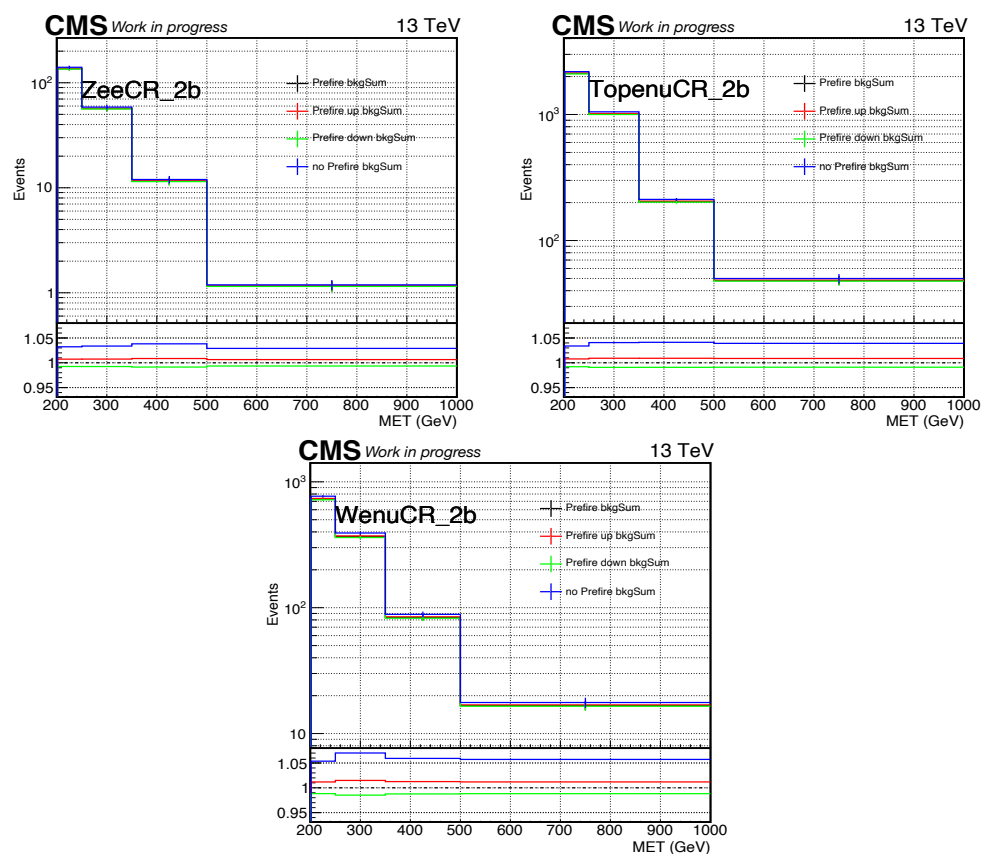


Figure 10: Recoil distribution of total background in electron CRs to check the effect of l1 prefire in 2017 MC

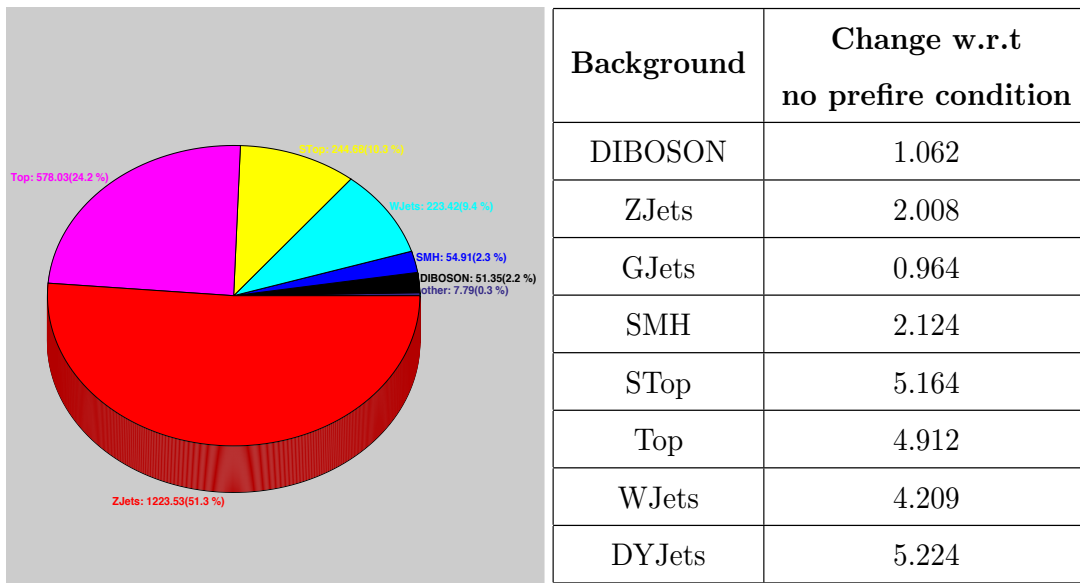


Figure 11: The left plot, Contribution of individual background in SR2. In the right plot the percentage change in the yield of individual background with and without prefiring.

B Optimization

The E_T^{miss} cut that has been applied in the analysis so far has been 200 GeV, However it has been noted out that at $E_T^{\text{miss}} \approx 200$ GeV, the data/MC scale factors comes out to be around 0.85, which is far away from 1. Hence, to get the most efficient region of the turnOn region for the MET we have shifted to $E_T^{\text{miss}} > 250$ GeV, by the shifting the $E_T^{\text{miss}} > 250$ GeV the need of additional treatment of the uncertainties around 200 GeV MET vanishes.

The jet selection in the analysis requires maximum of 3 jets in the signal regions with 2b category. This is to have orthogonality with the $t\bar{t}$ +DM analyses. Furthermore to optimise the jet selection, the upper limit(CR only fit) has been calculated (12) to see how much difference the change in selection w.r.t number of jets.

And for the p_T of the leading jet, we move from 50 GeV to 100 GeV, which has also been studied using the upper limit calculation w.r.t leading jet p_T selection (12).

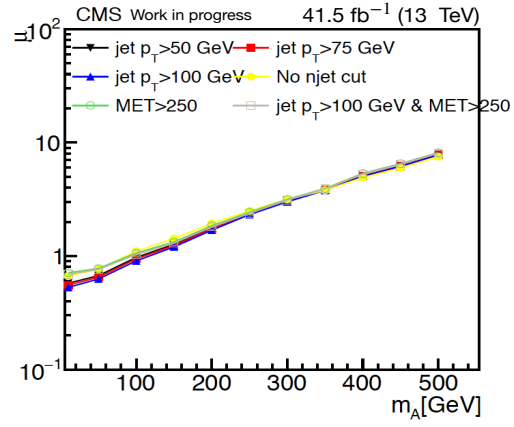


Figure 12: Comparison of upper limit plot for number of jets and jet p_T selection with CR only fit for 2017

C Plots of other variables for CRS

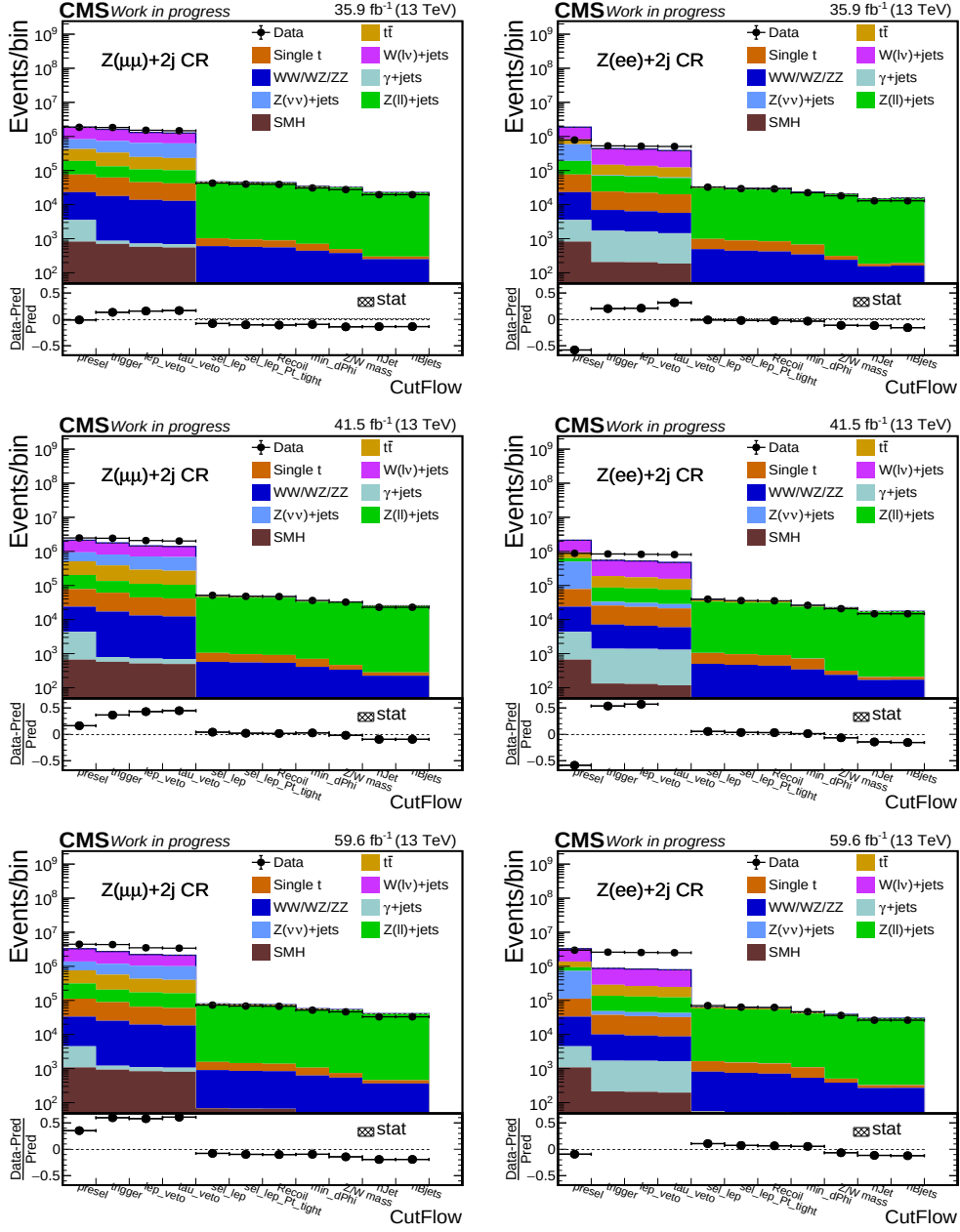


Figure 13: CutFlow plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

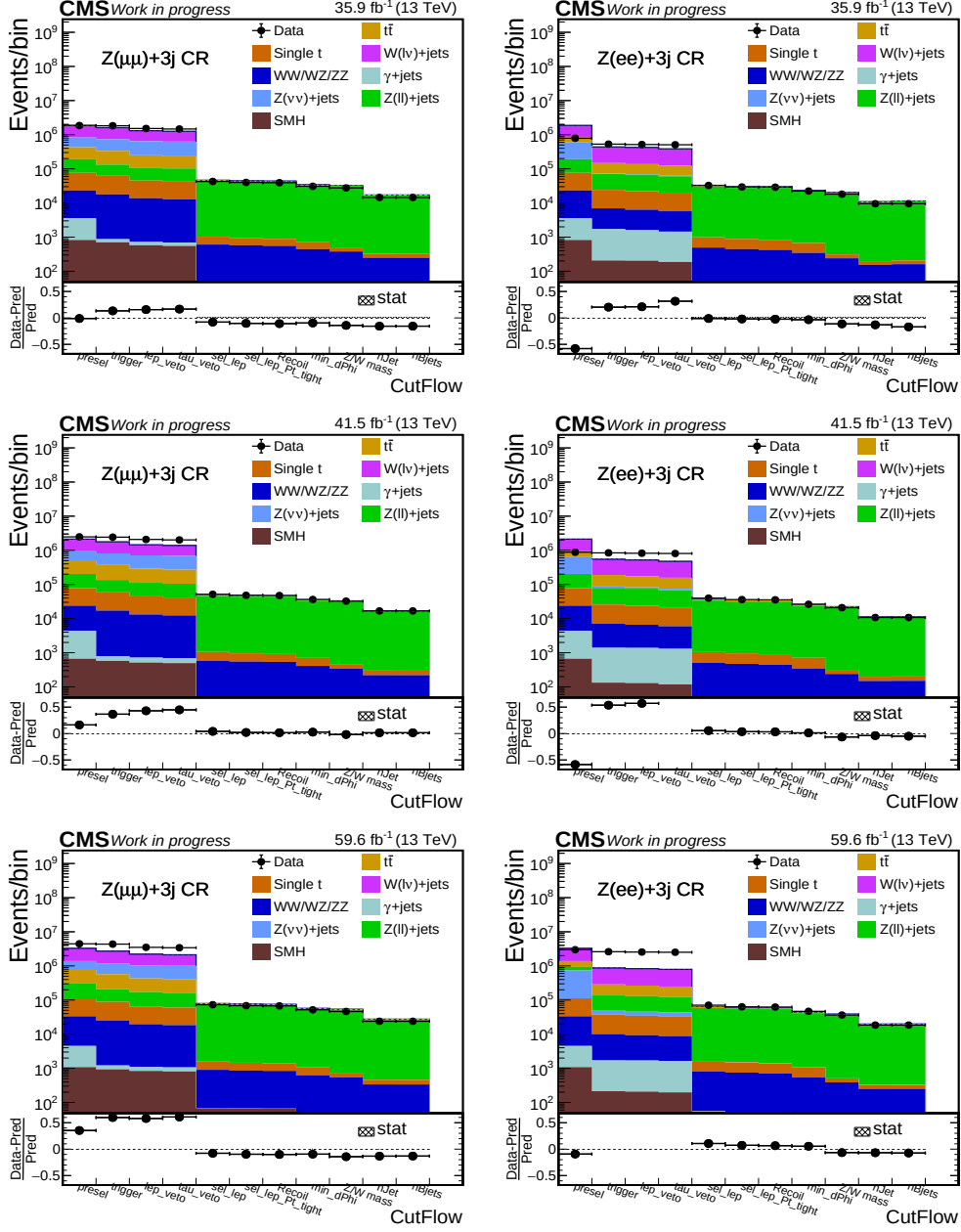


Figure 14: CutFlow plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

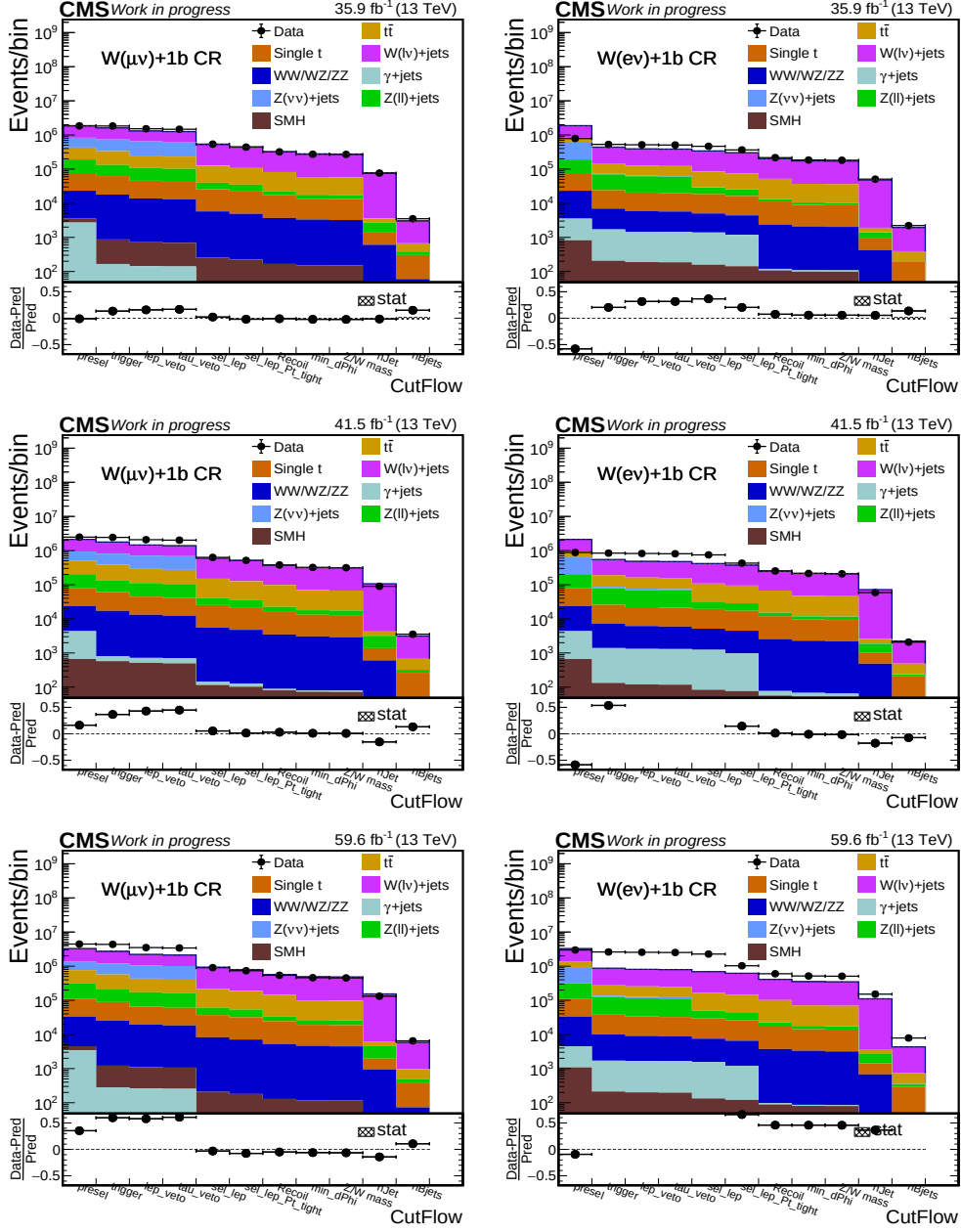


Figure 15: CutFlow plot in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

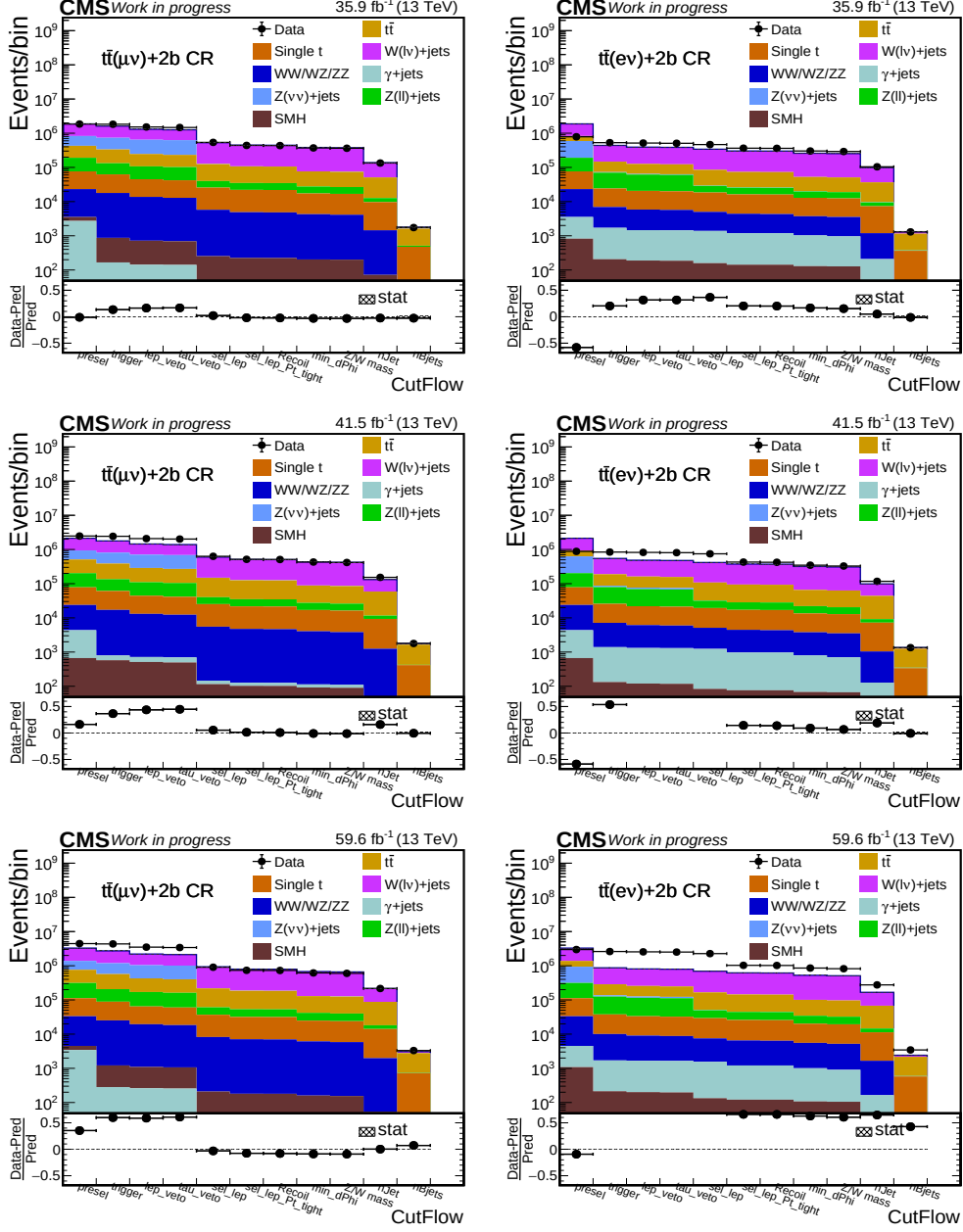


Figure 16: CutFlow plot in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

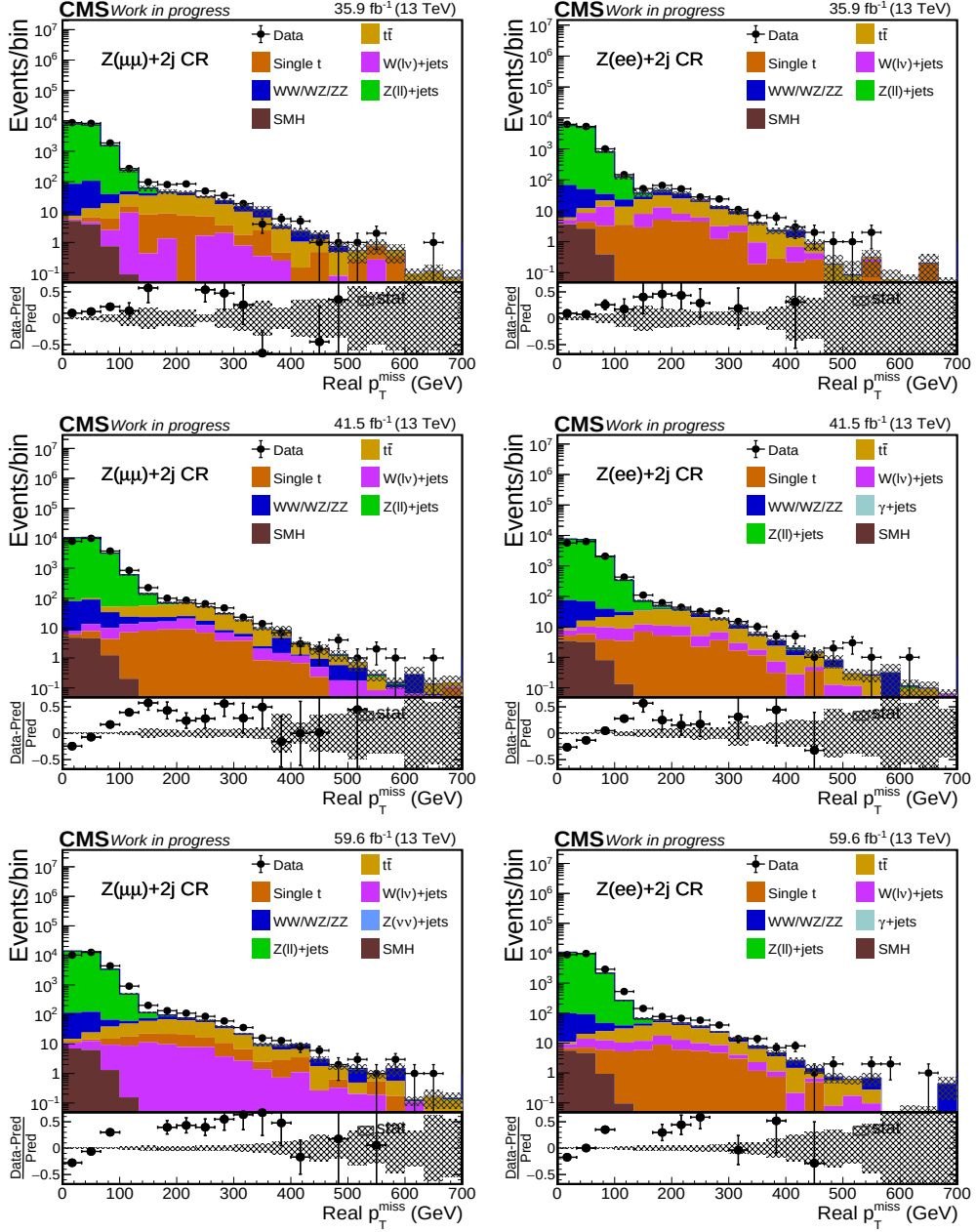


Figure 17: E_T^{miss} plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

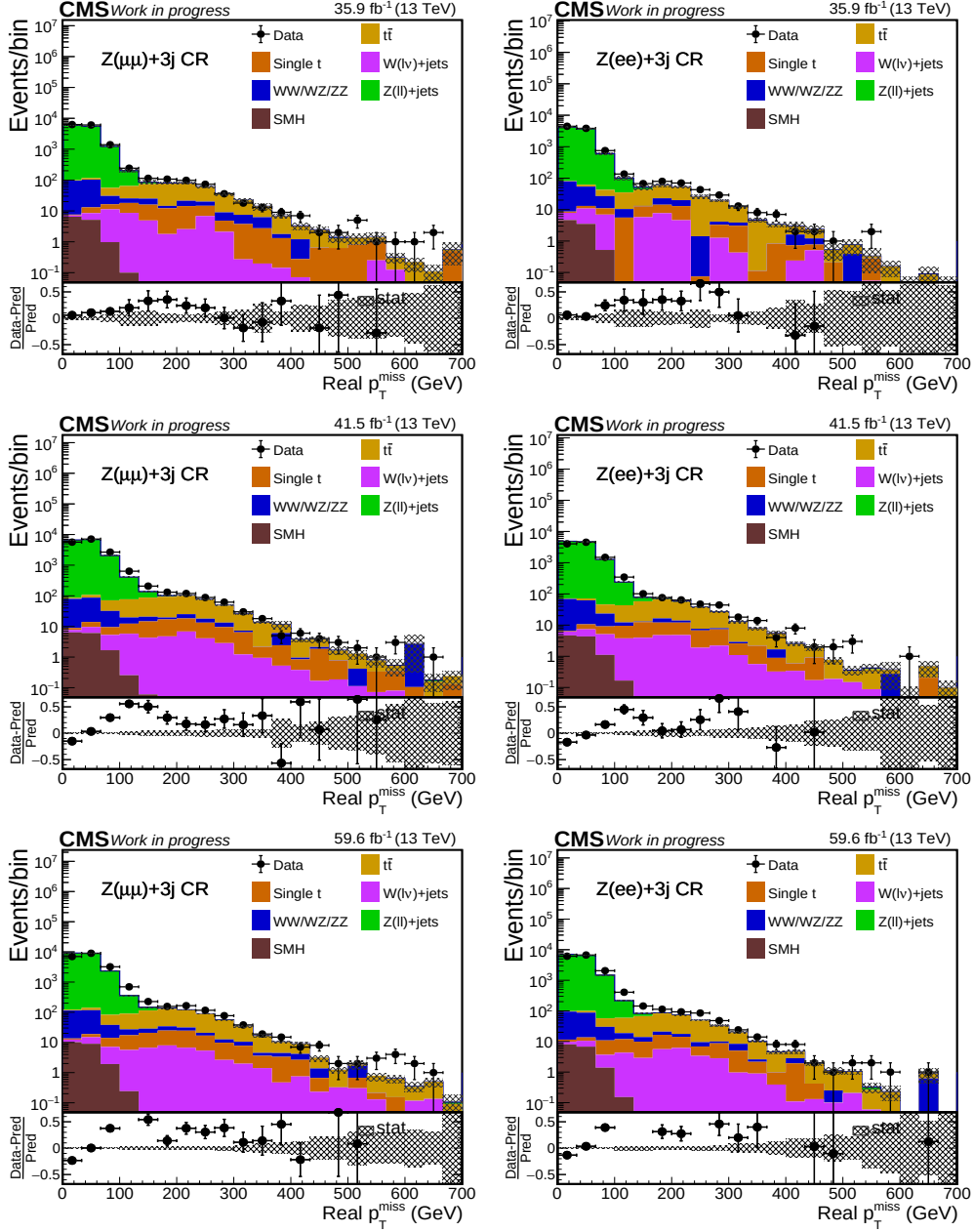


Figure 18: E_T^{miss} plot in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

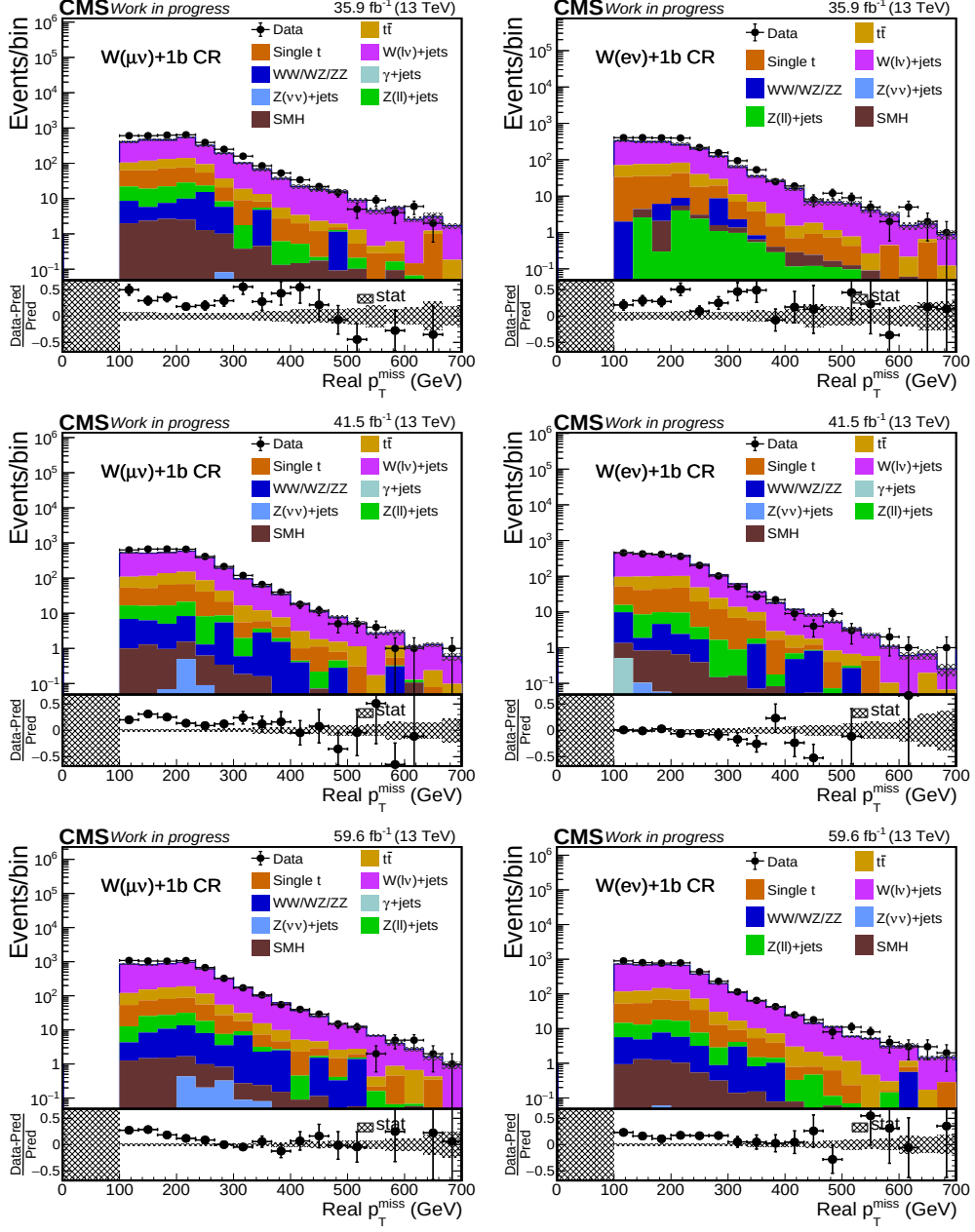


Figure 19: E_T^{miss} plot in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

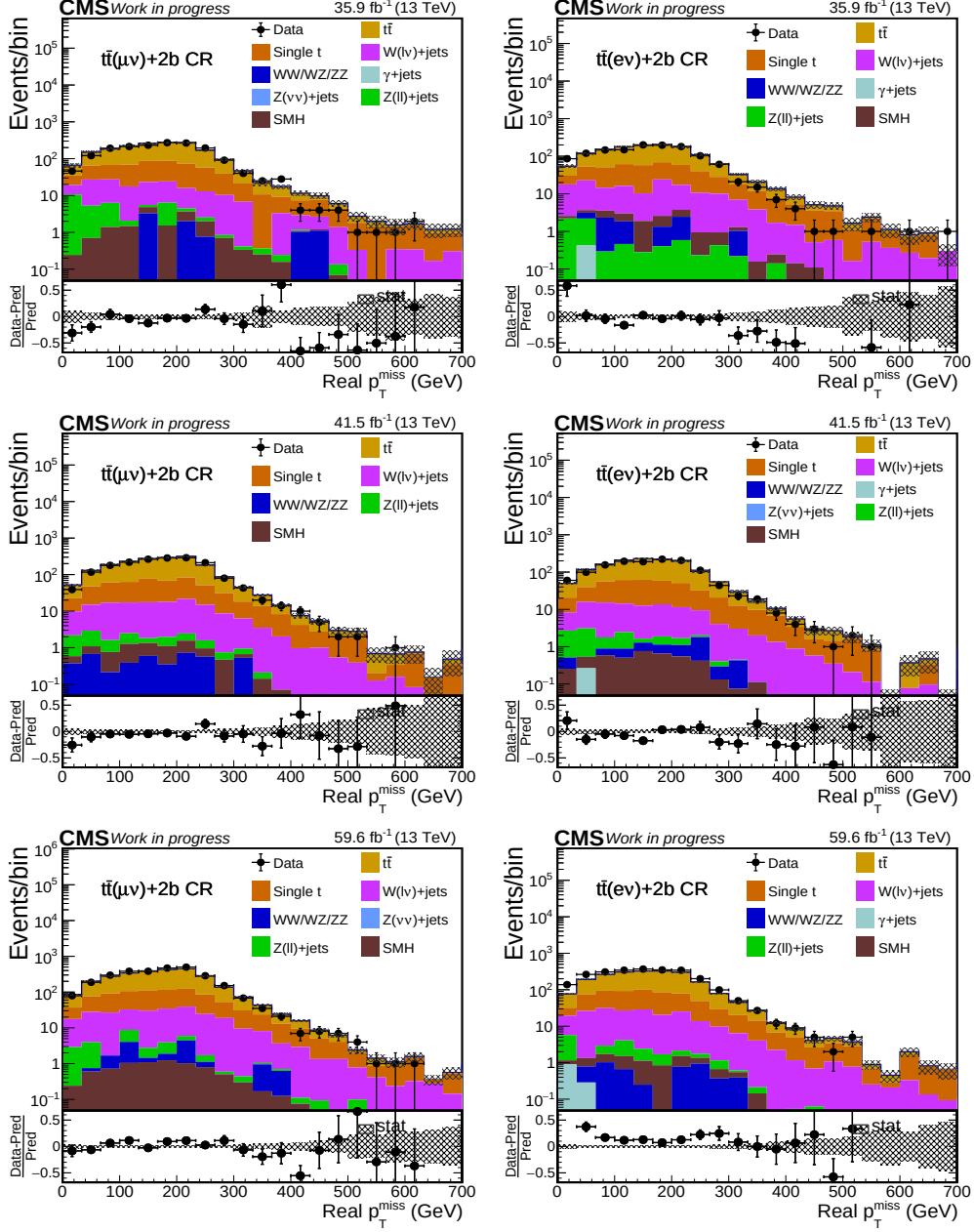


Figure 20: E_T^{miss} plot in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

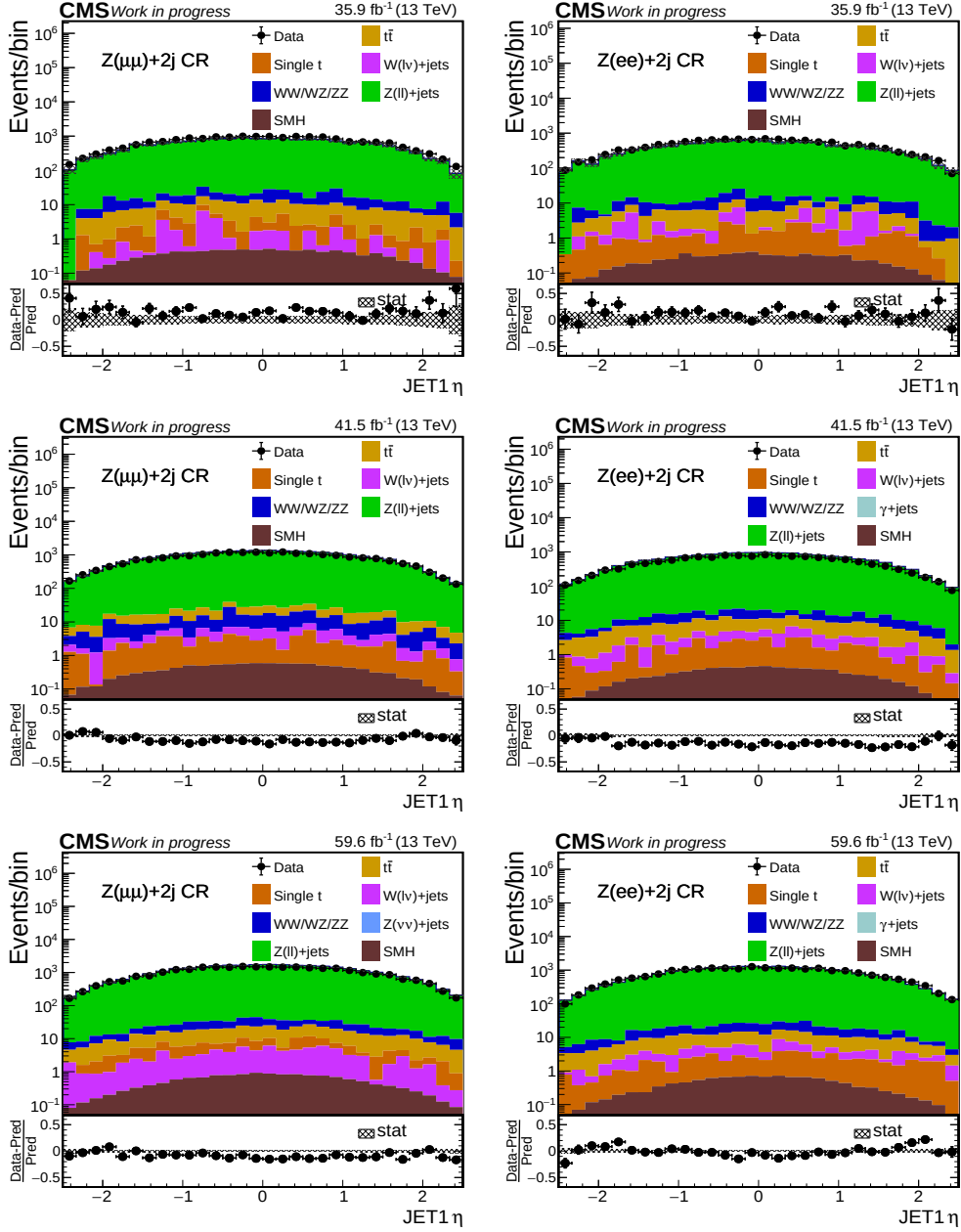


Figure 21: Jet1 η in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

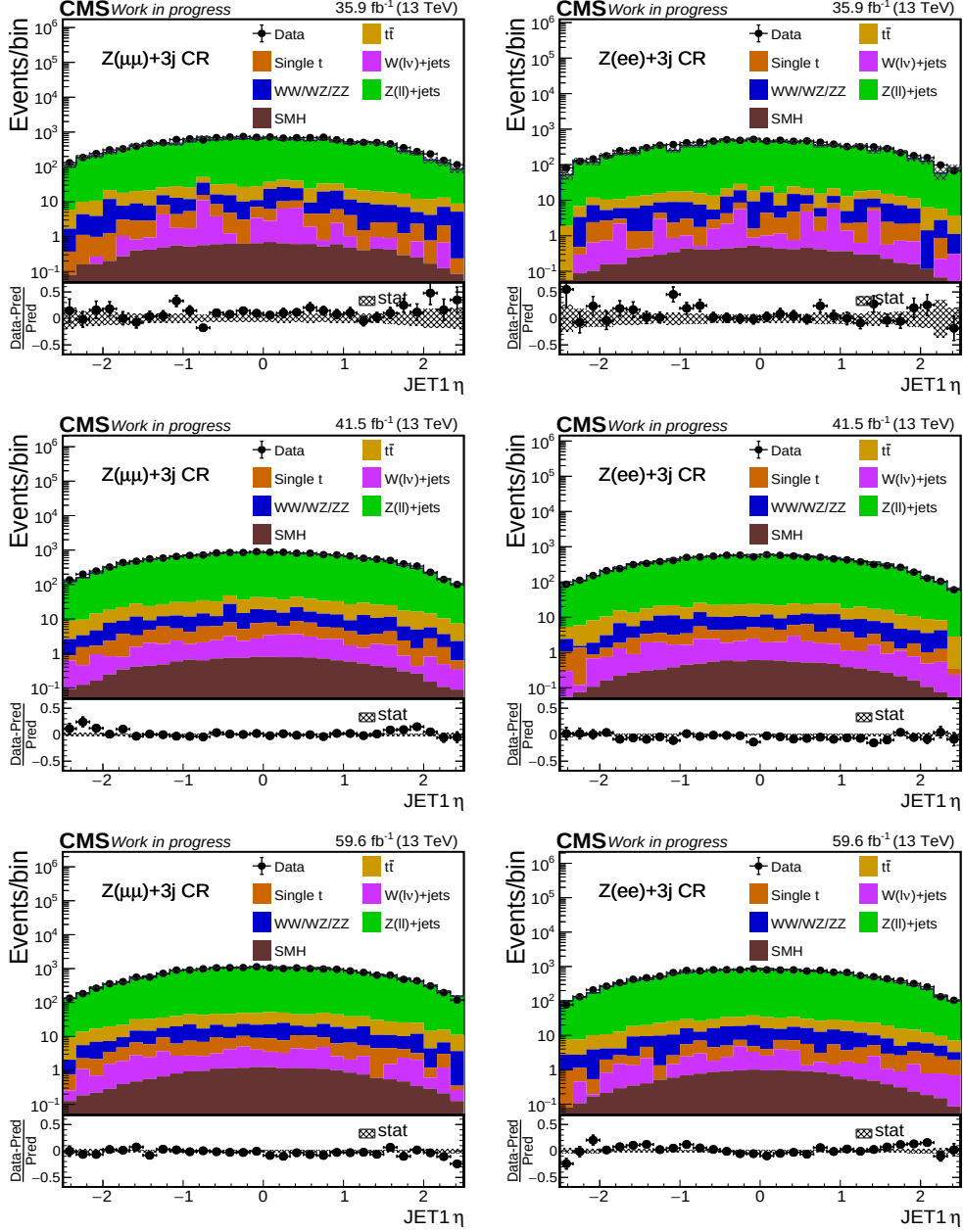


Figure 22: Jet1 η in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

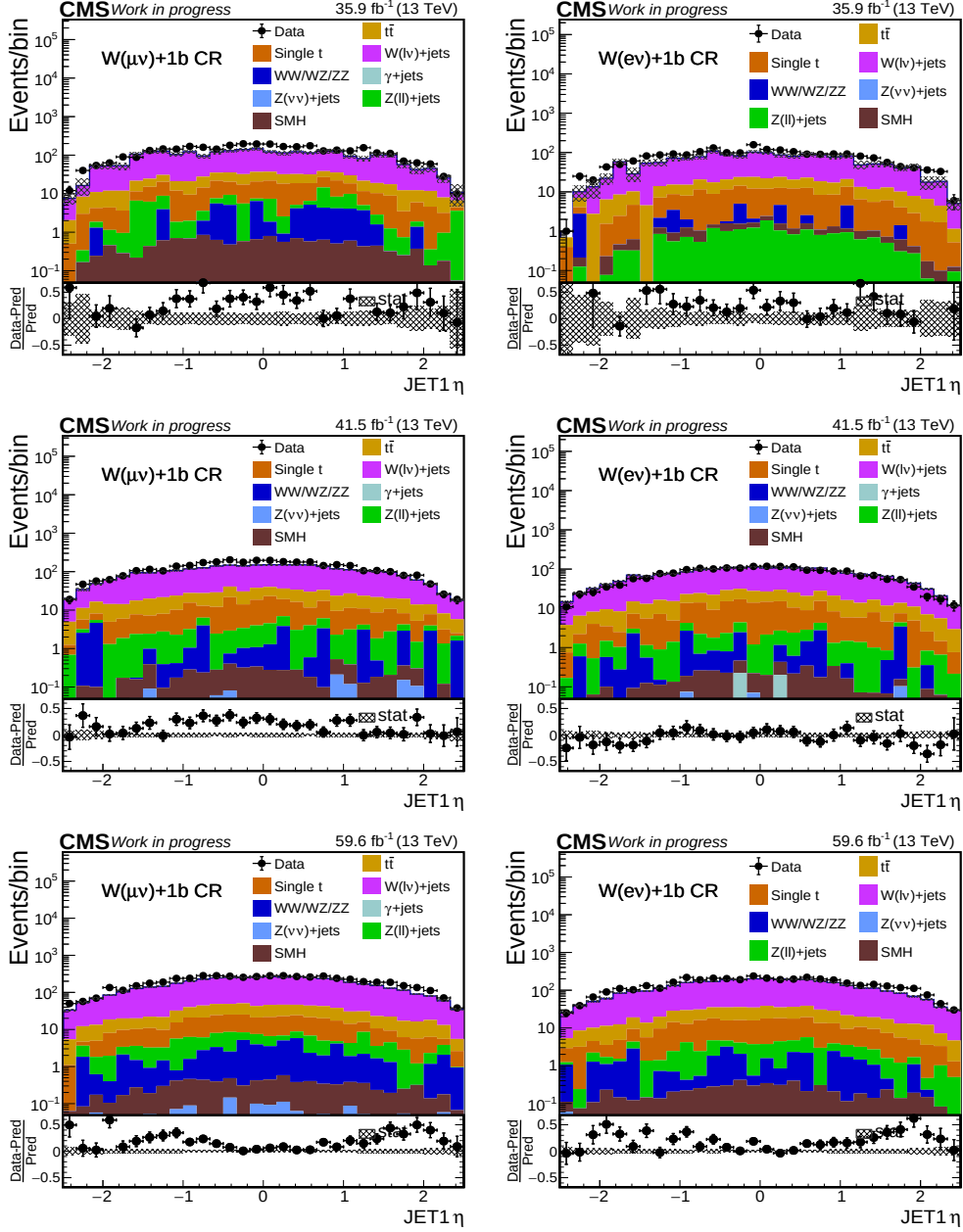


Figure 23: Jet1 η in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

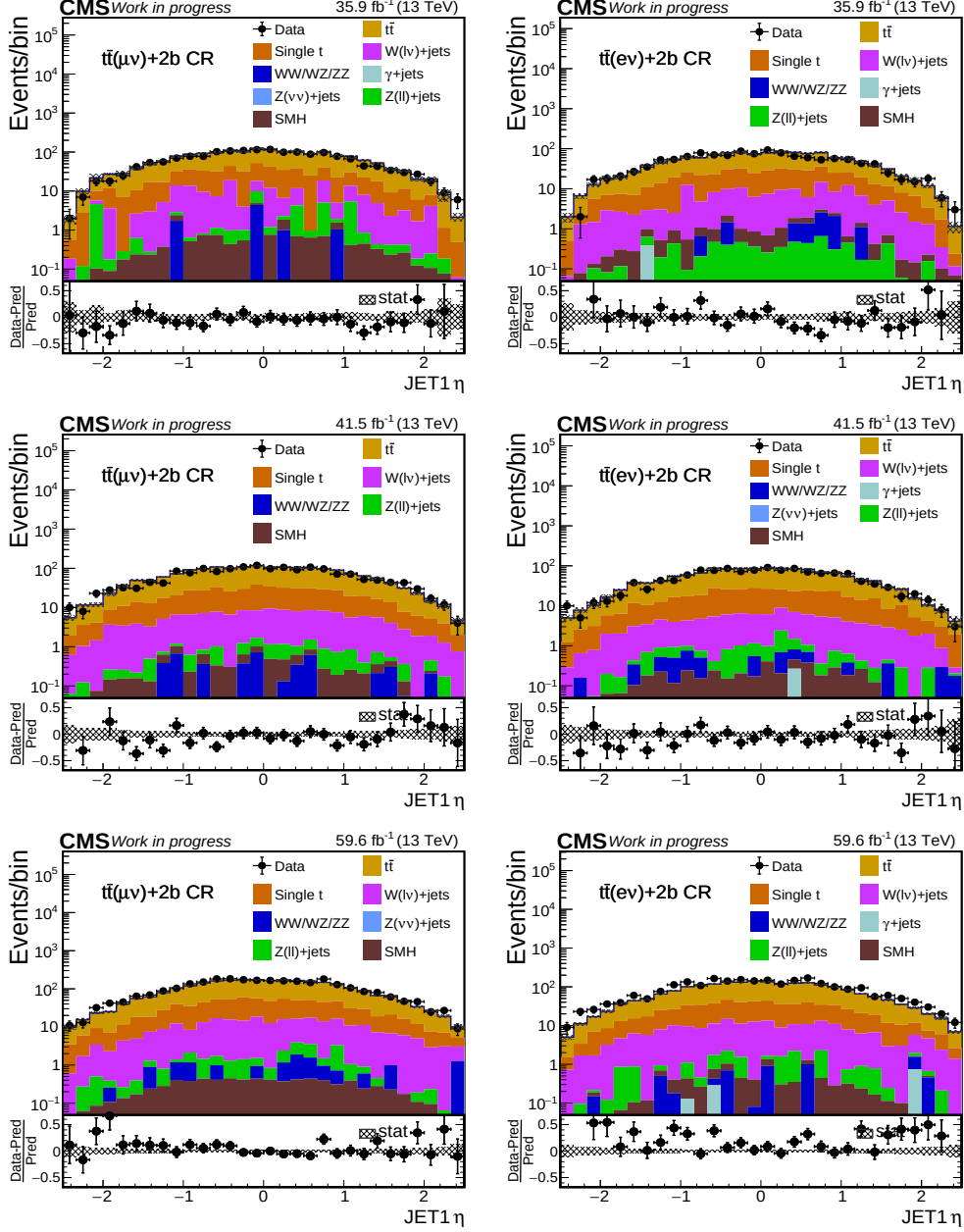


Figure 24: Jet1 η in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

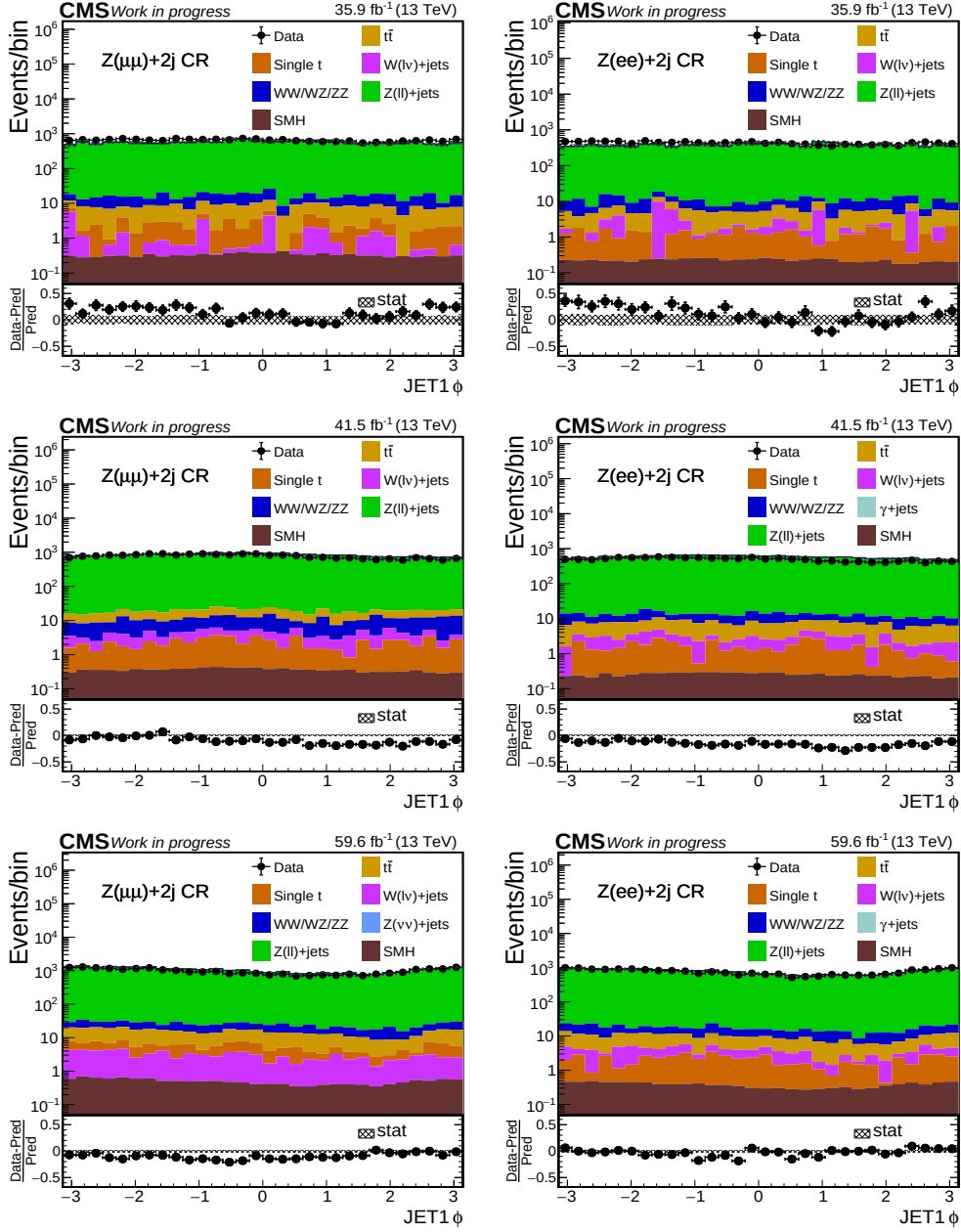


Figure 25: Jet1 ϕ in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

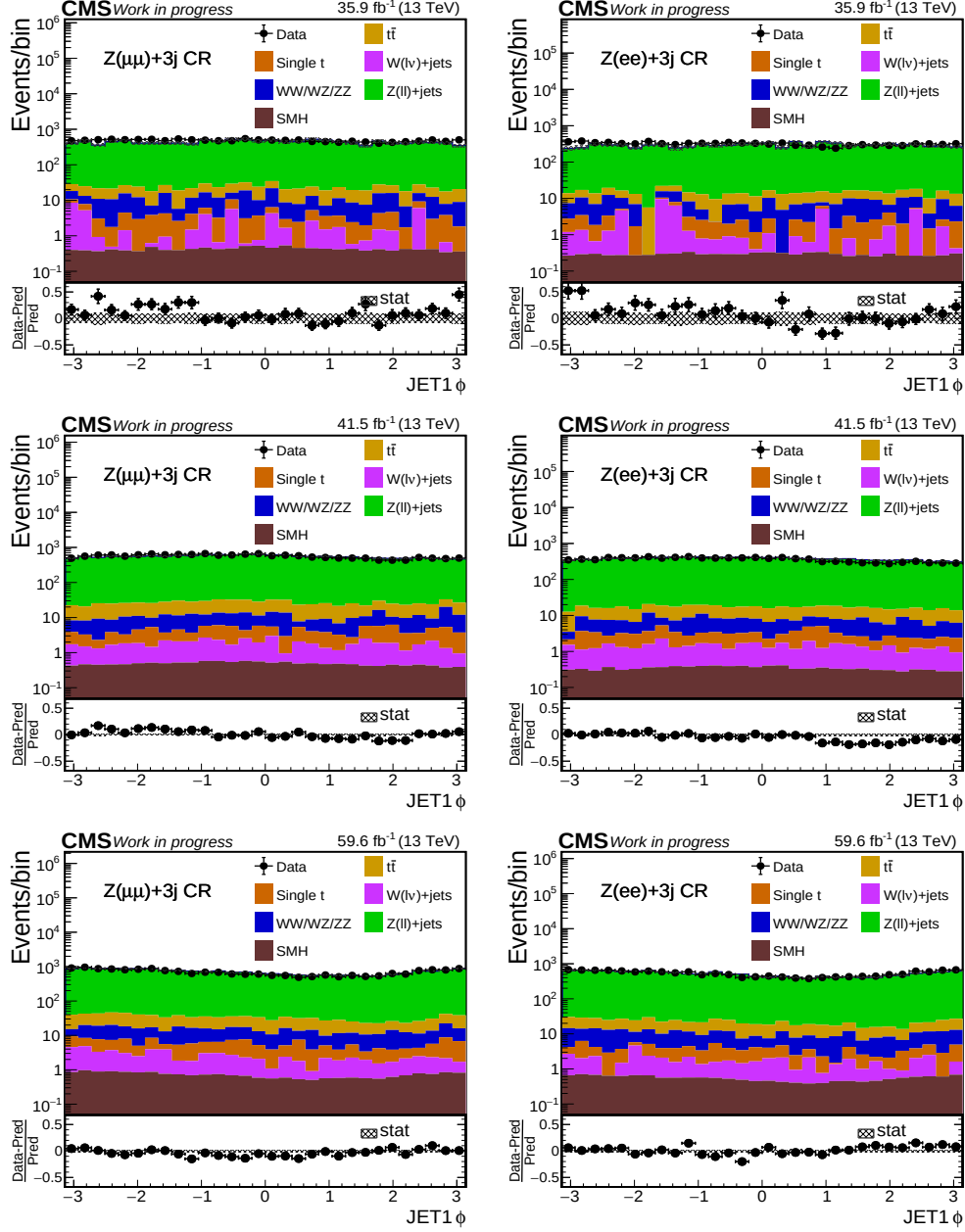


Figure 26: $\text{Jet1 } \phi$ in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

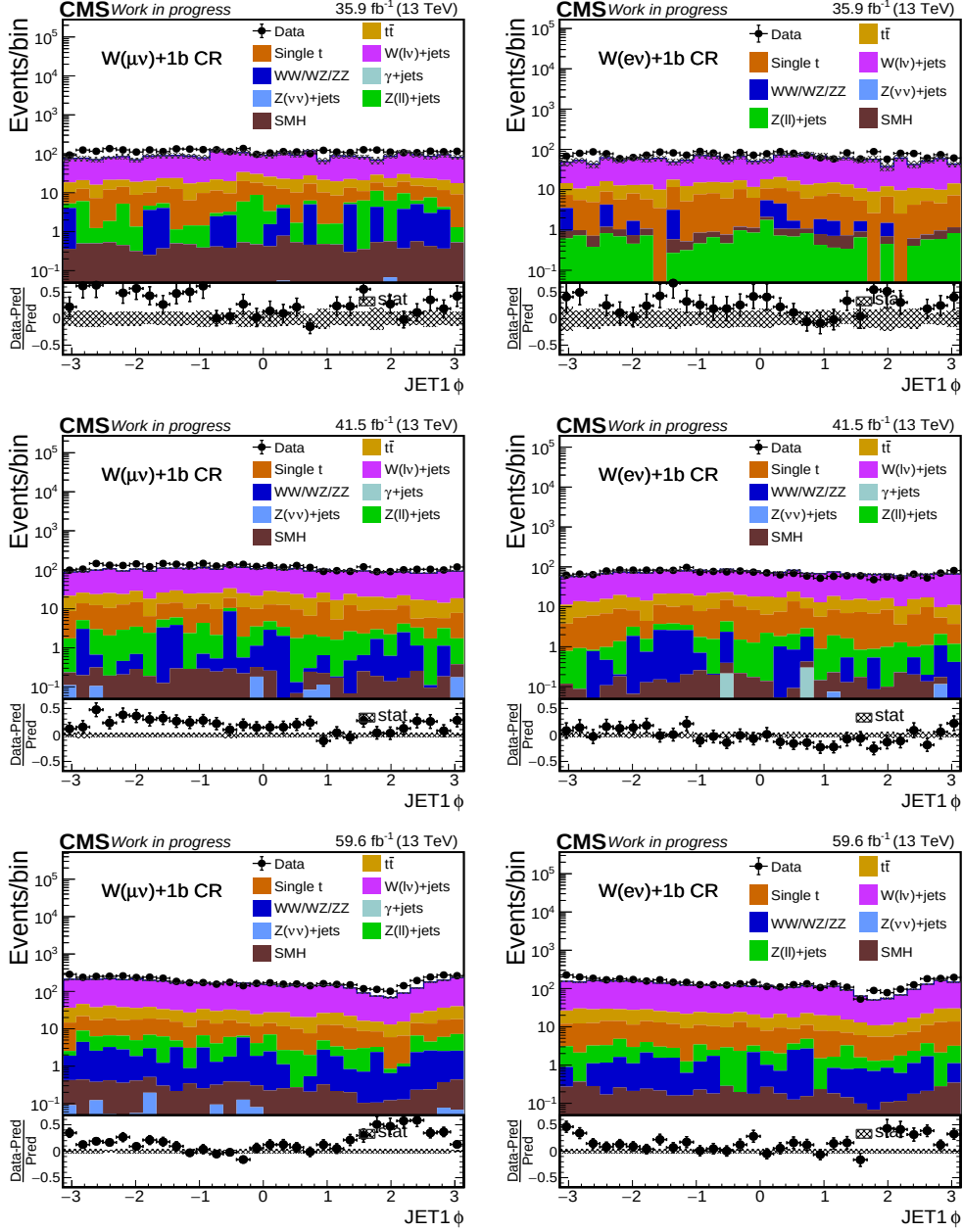


Figure 27: $\text{Jet1 } \phi$ in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

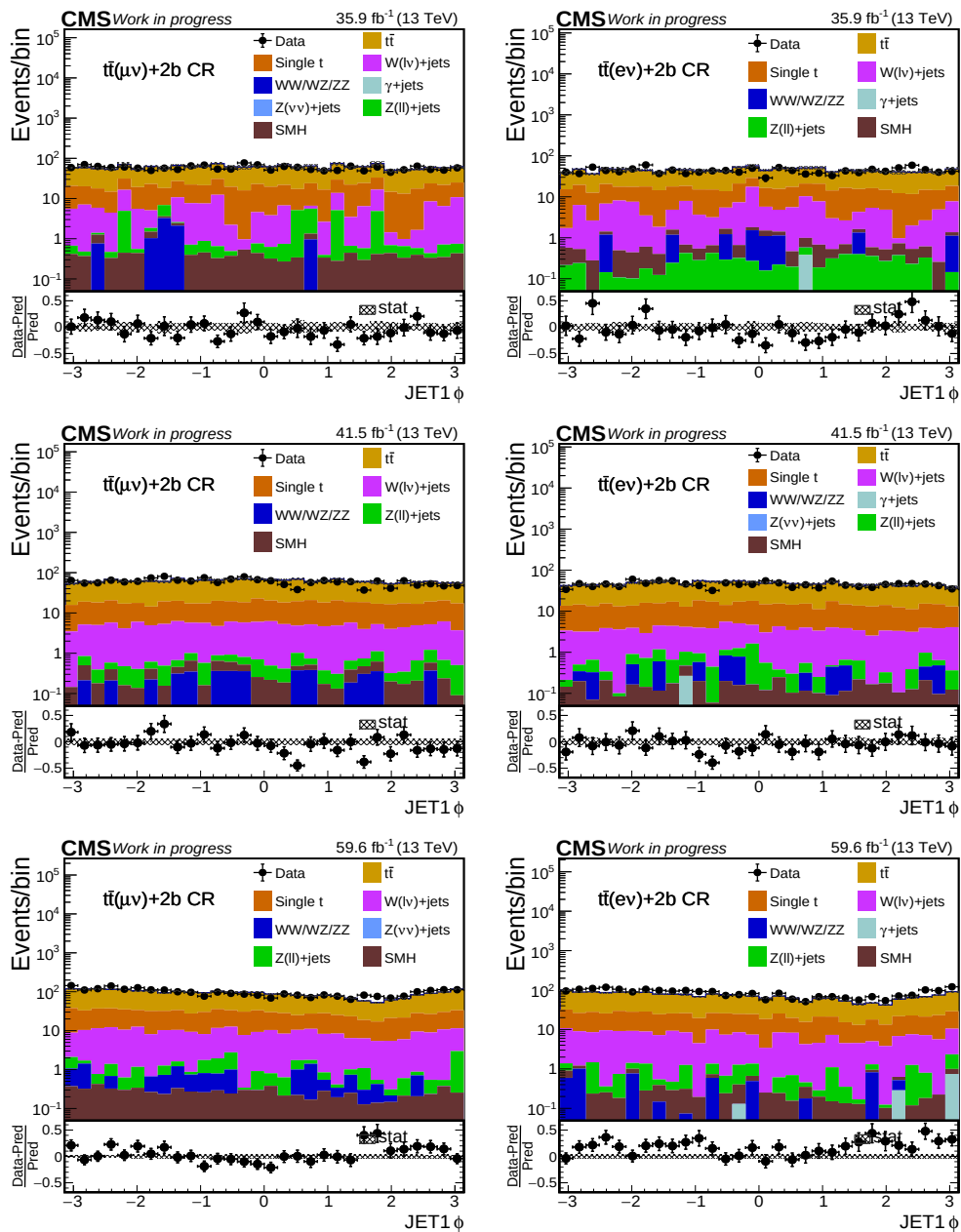


Figure 28: $\text{Jet1 } \phi$ in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

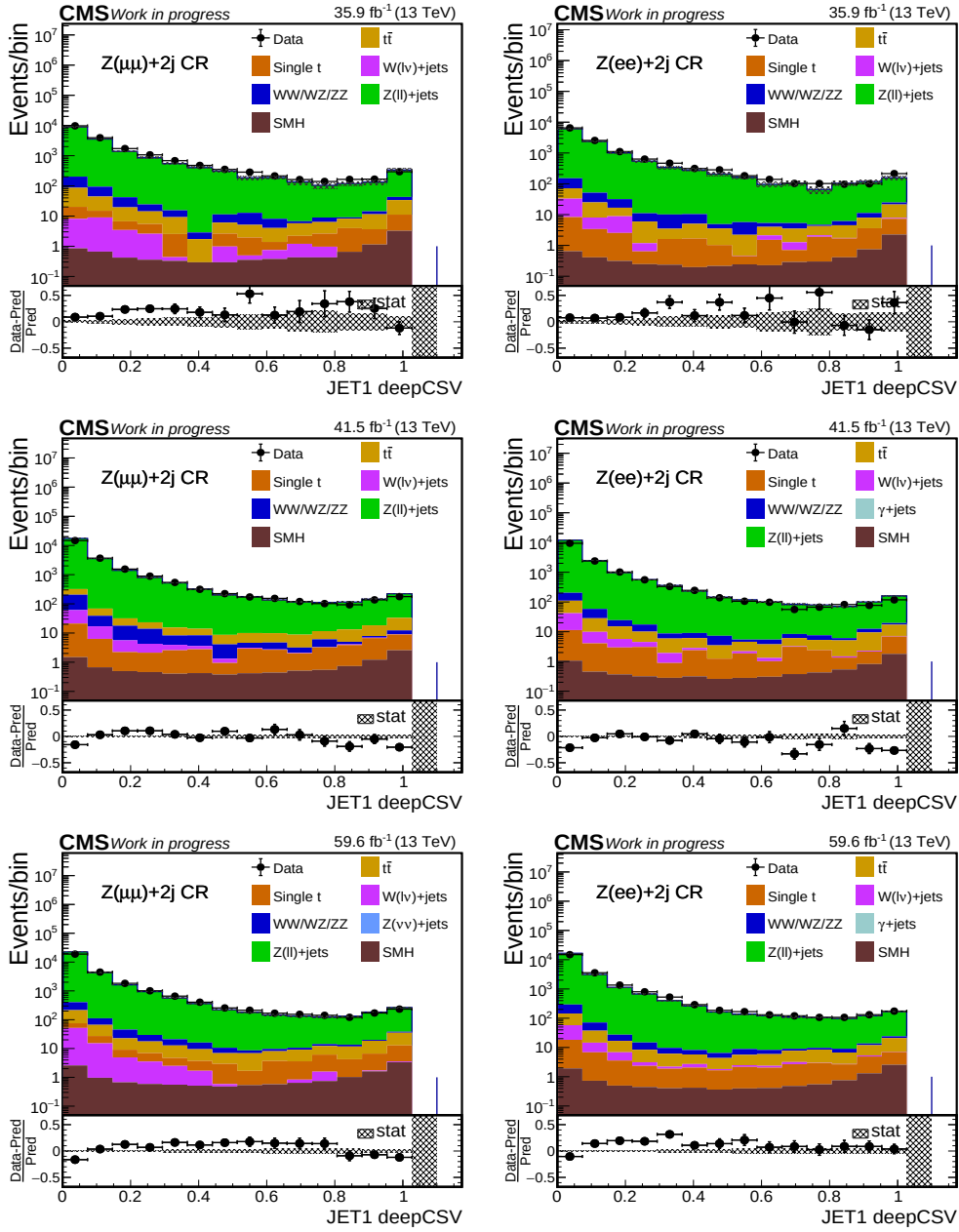


Figure 29: DeepCSV discriminator for Jet1 in Z→μμ and Z→ee control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

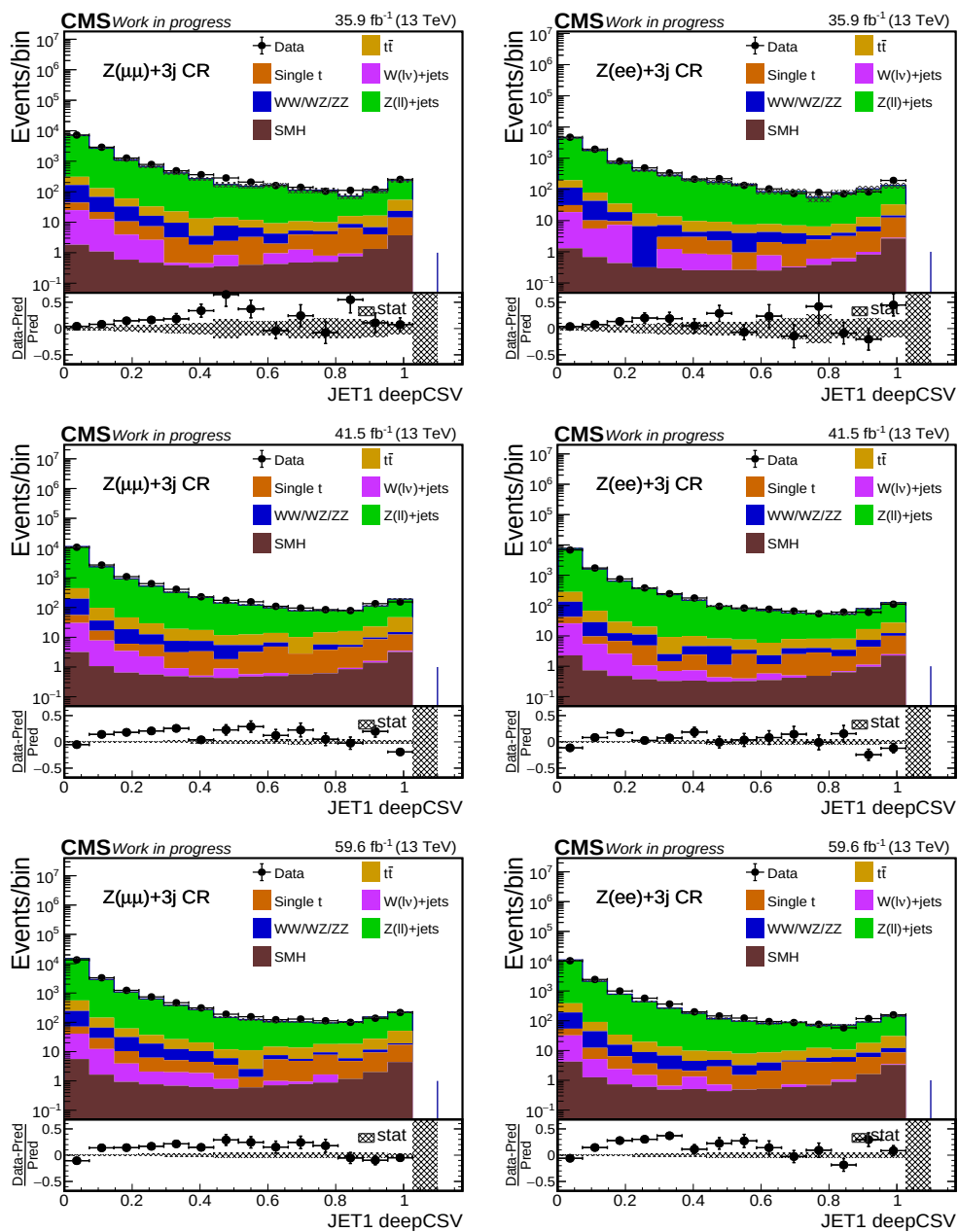


Figure 30: DeepCSV discriminator for Jet1 in Z→μμ and Z→ee control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

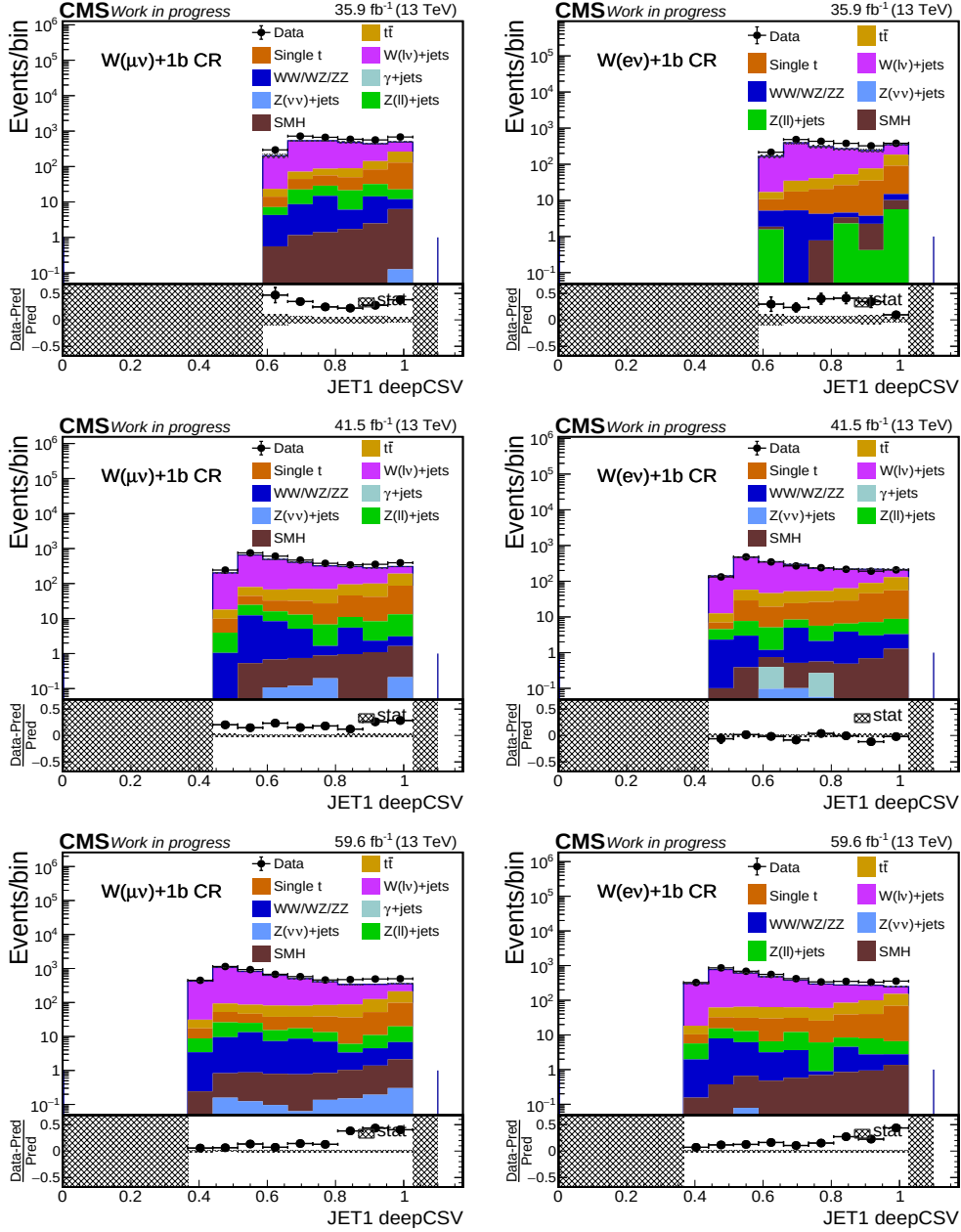


Figure 31: DeepCSV discriminator for Jet1 in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

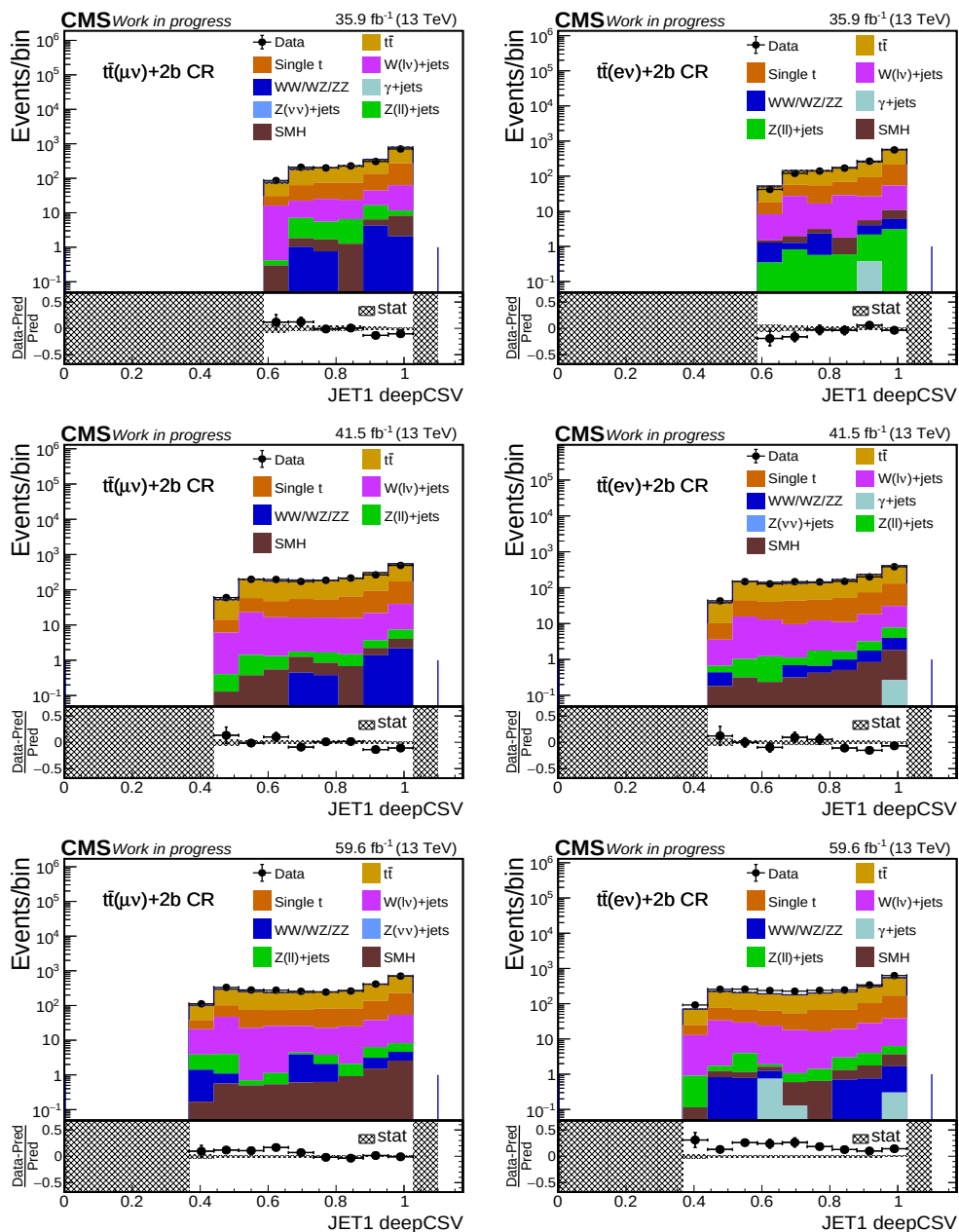


Figure 32: DeepCSV discriminator for Jet1 in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

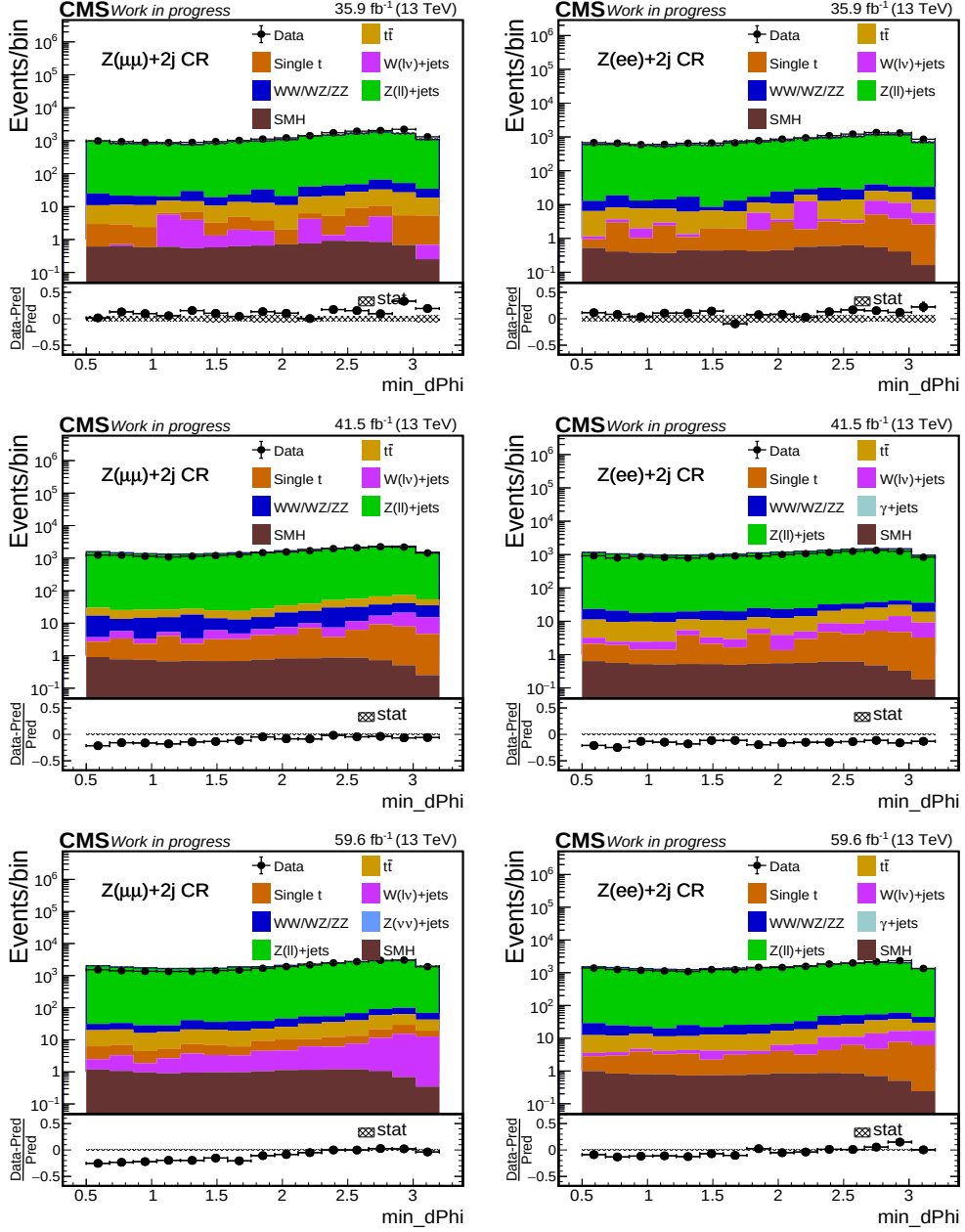


Figure 33: $\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in Z $\rightarrow$ $\mu\mu$ and Z $\rightarrow$ ee control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

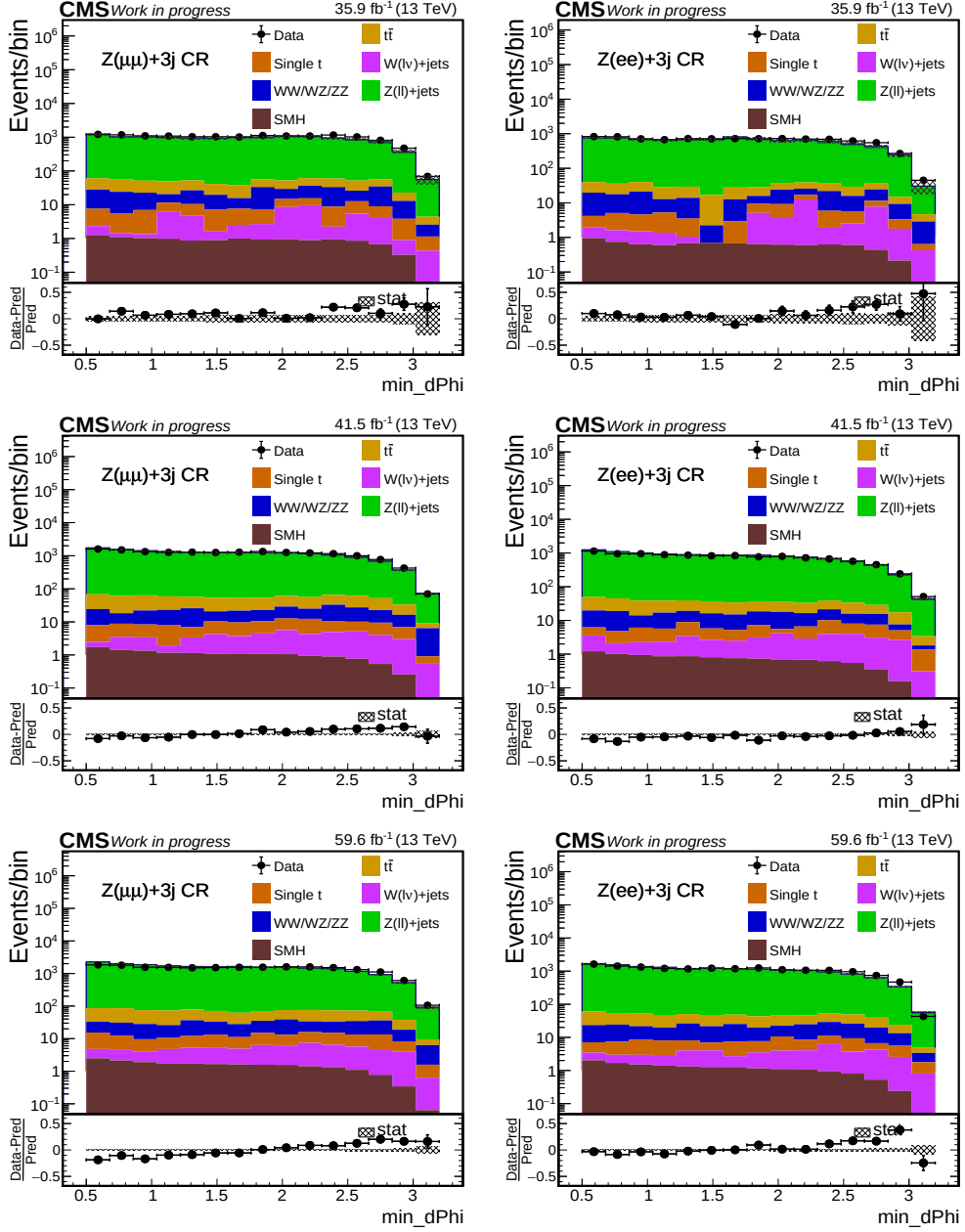


Figure 34: $\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in Z $\rightarrow$ $\mu\mu$ and Z $\rightarrow$ ee control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

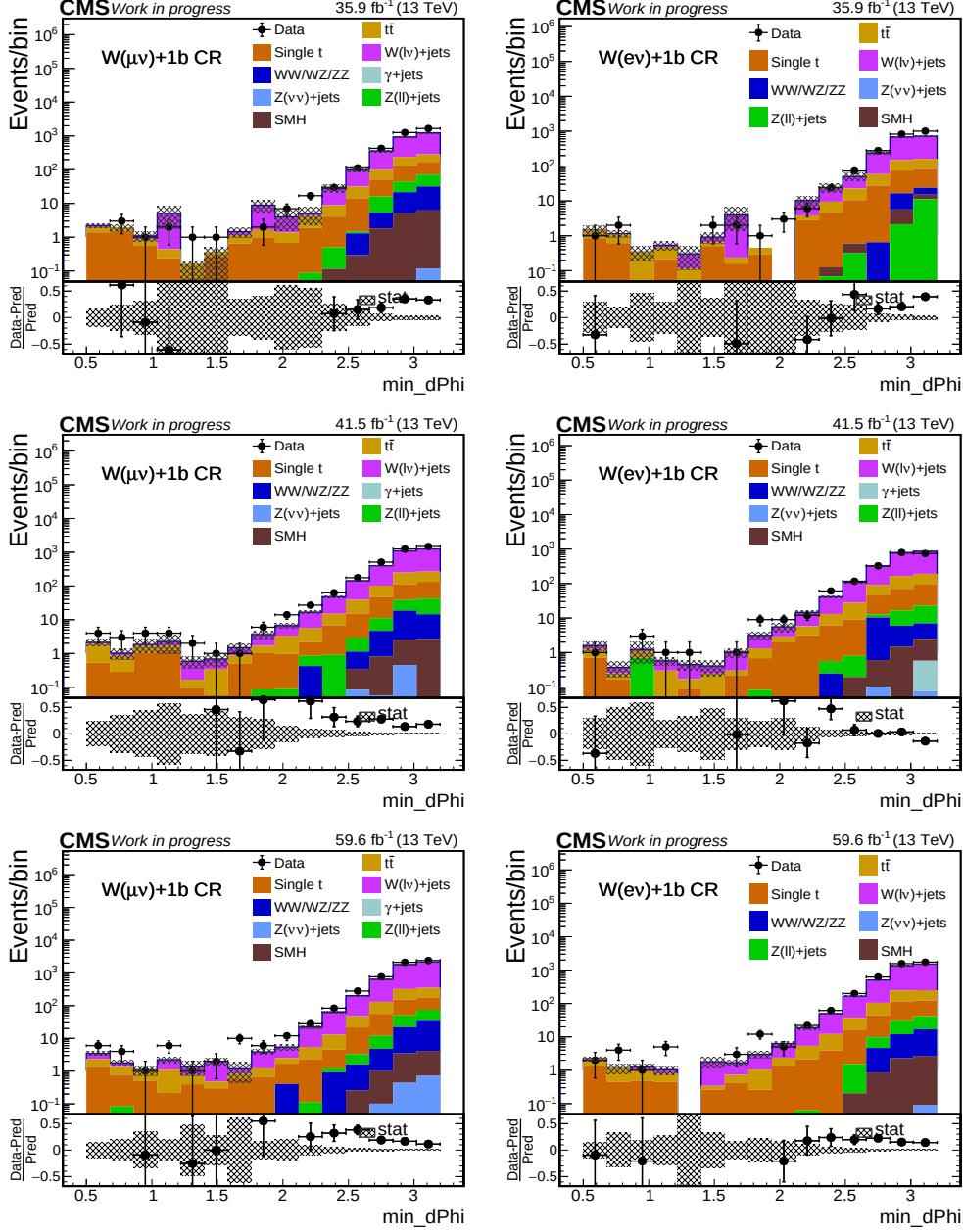


Figure 35: $\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

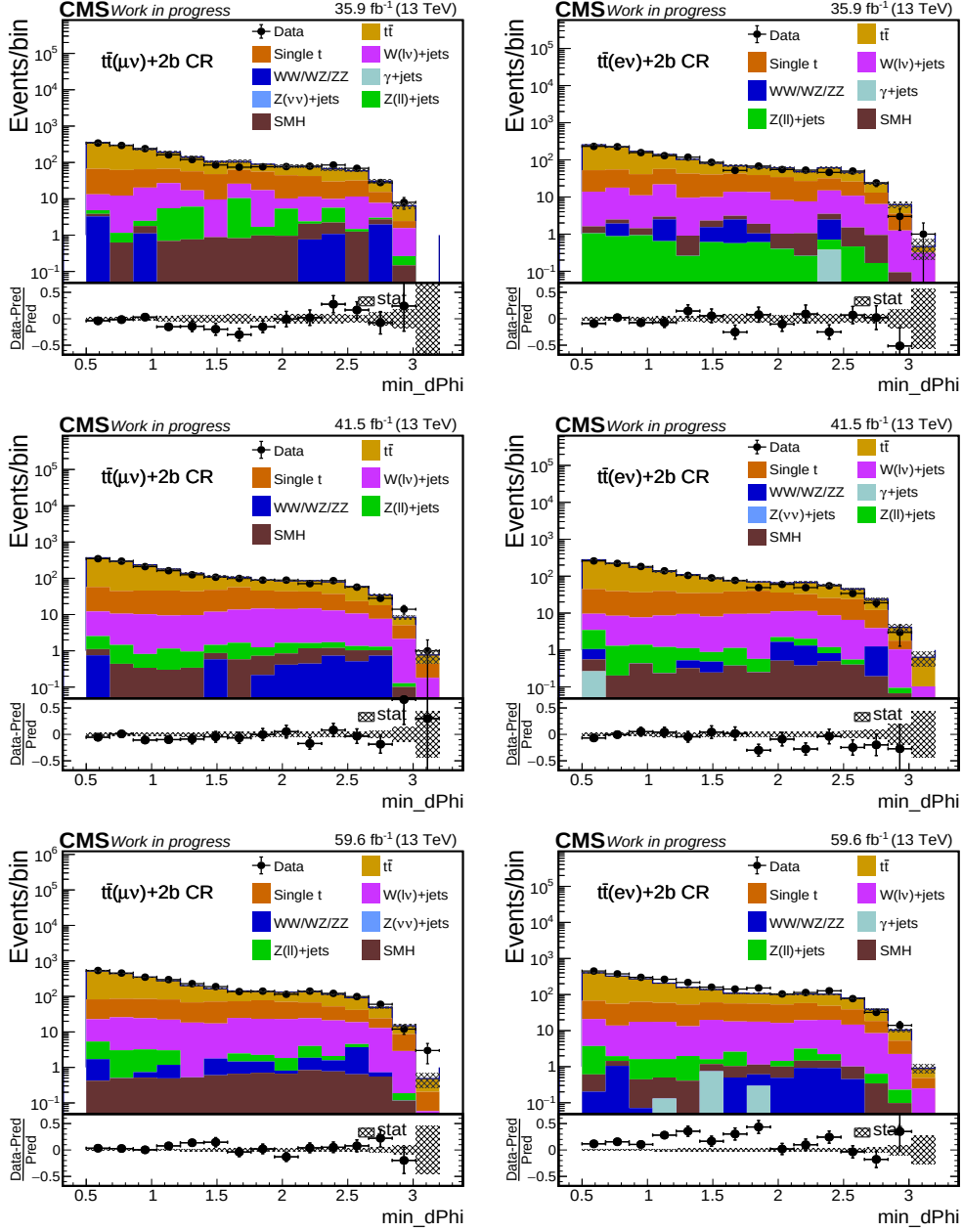


Figure 36: $\min(\Delta\phi(\text{jet}, E_T^{\text{miss}}))$ in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

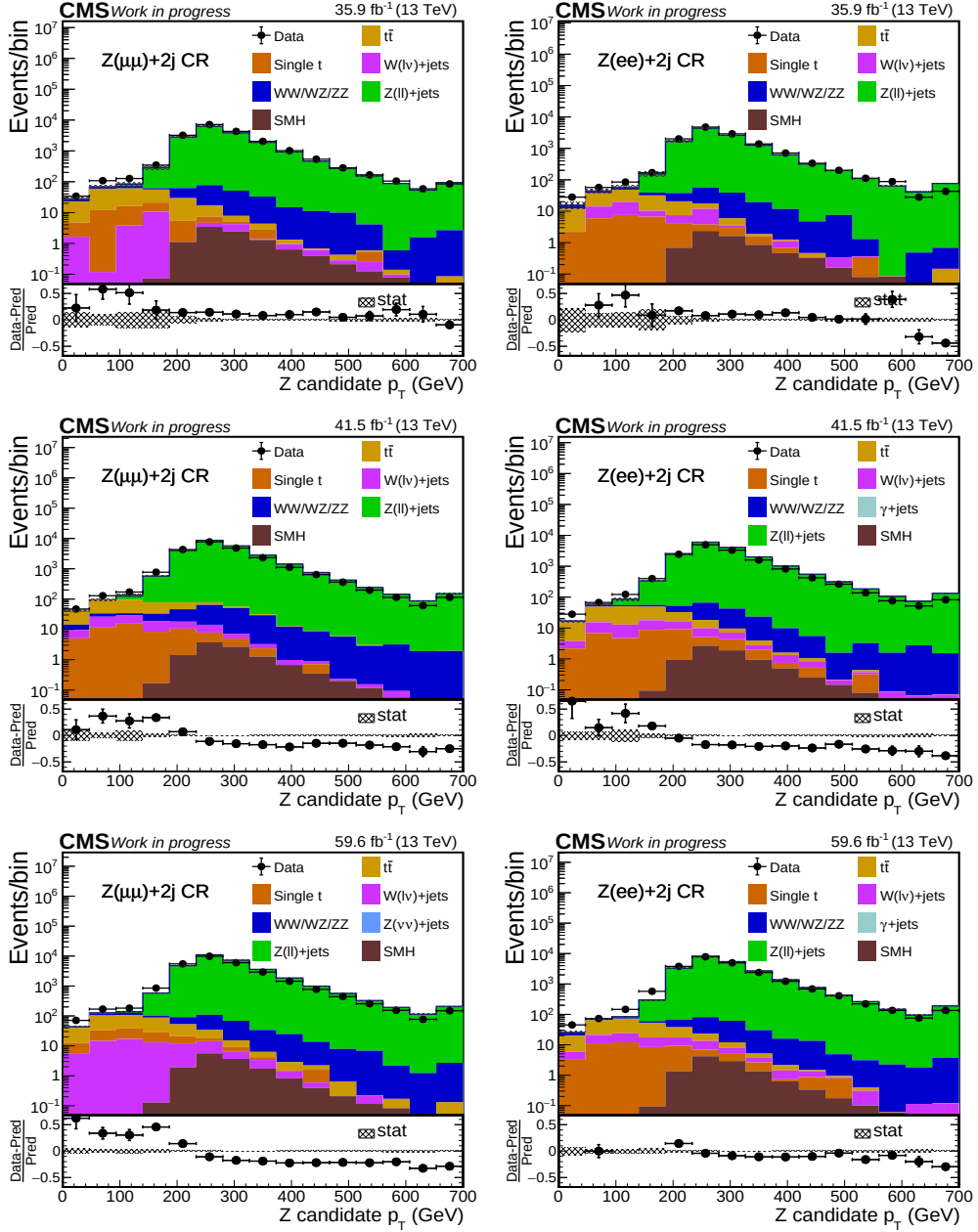


Figure 37: Z boson p_T in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

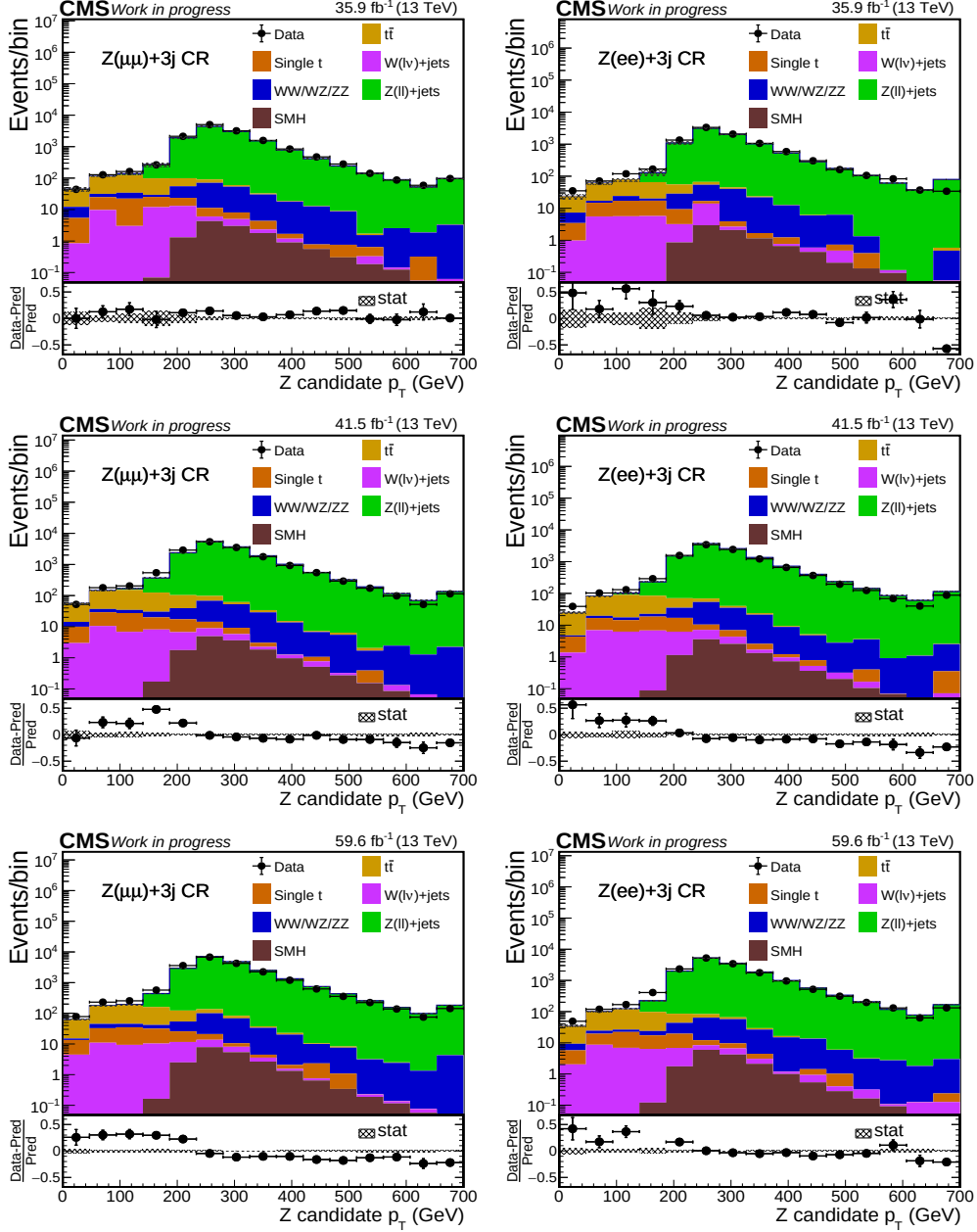


Figure 38: Z boson p_T in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

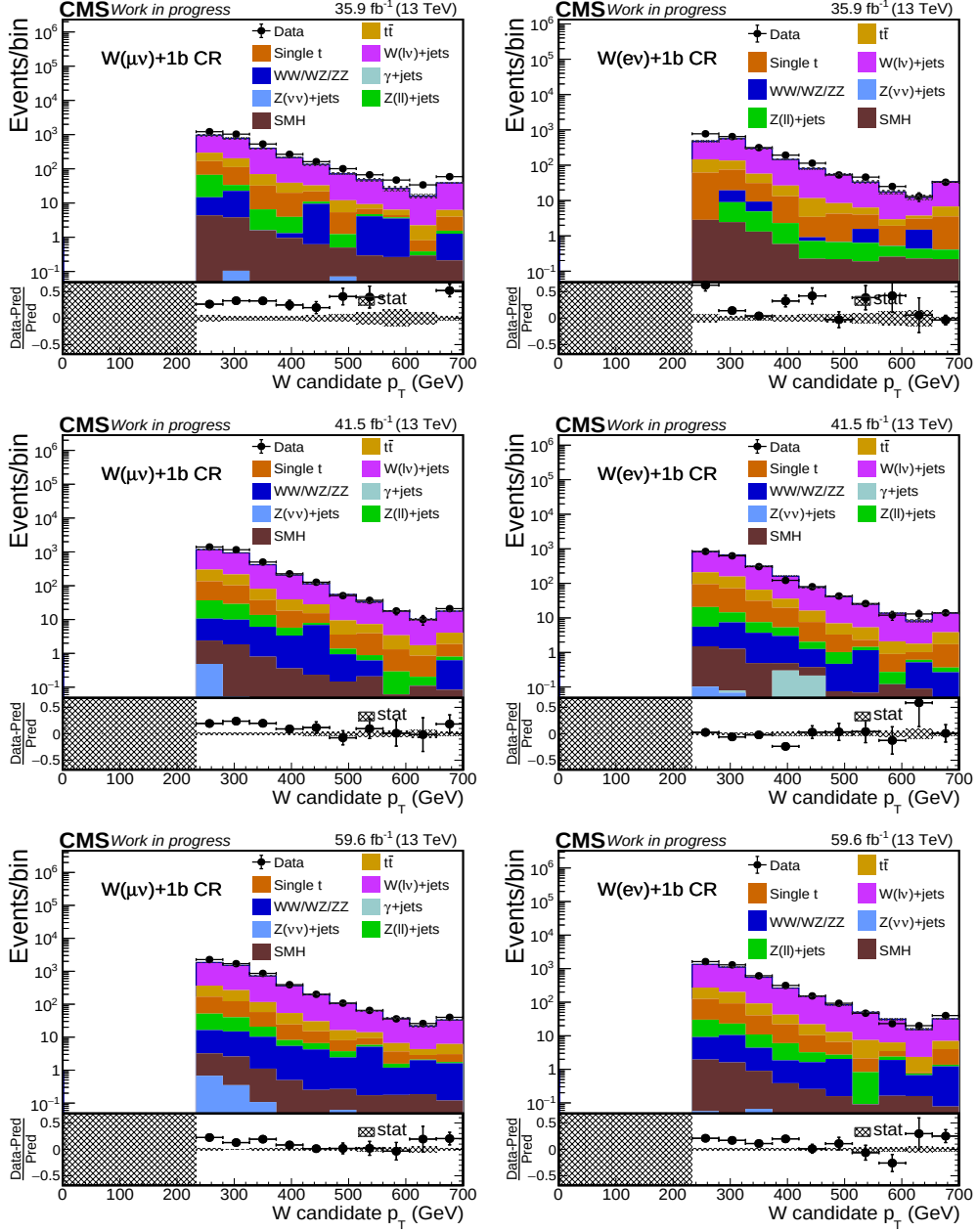


Figure 39: W boson p_T in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

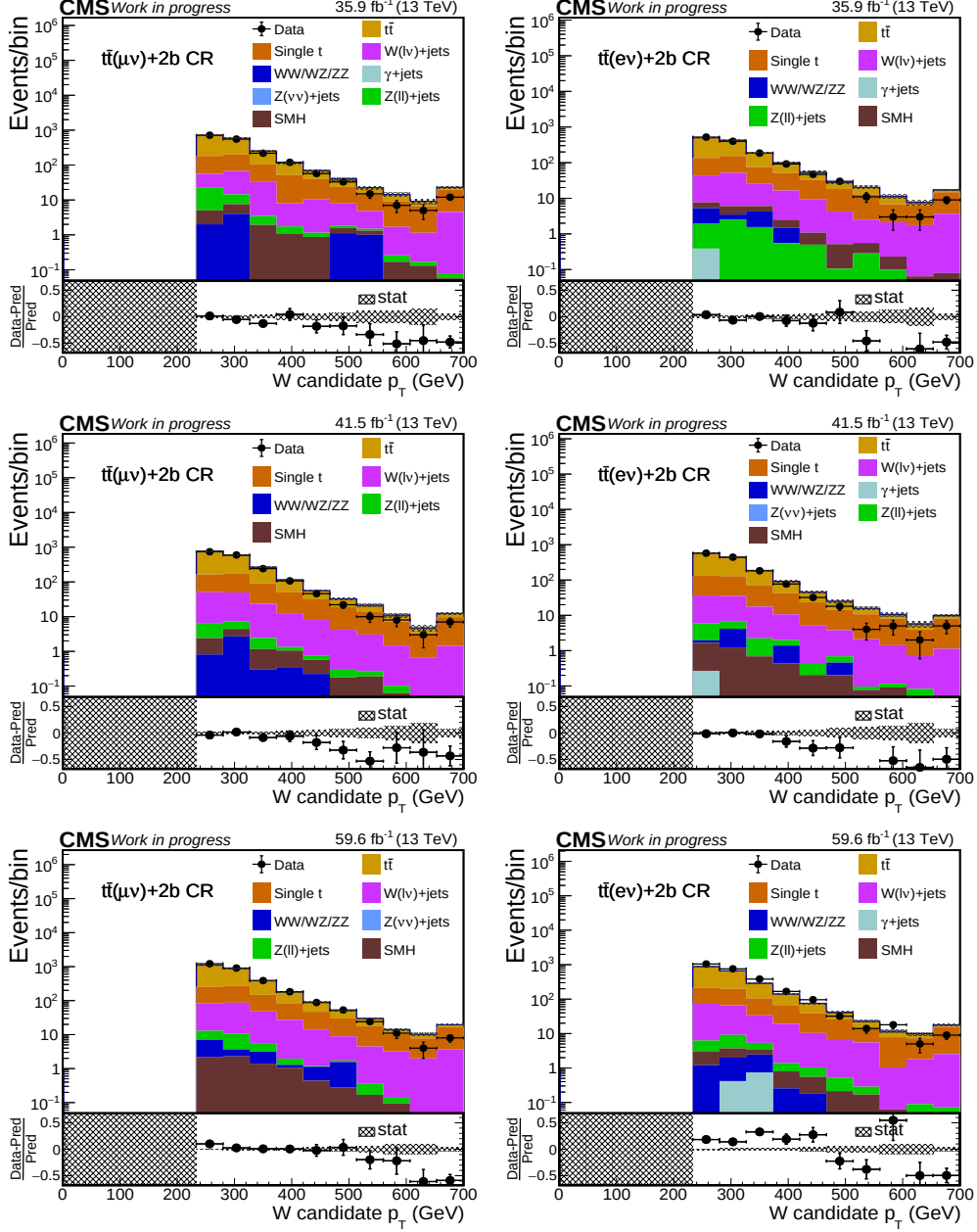


Figure 40: W boson p_T in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

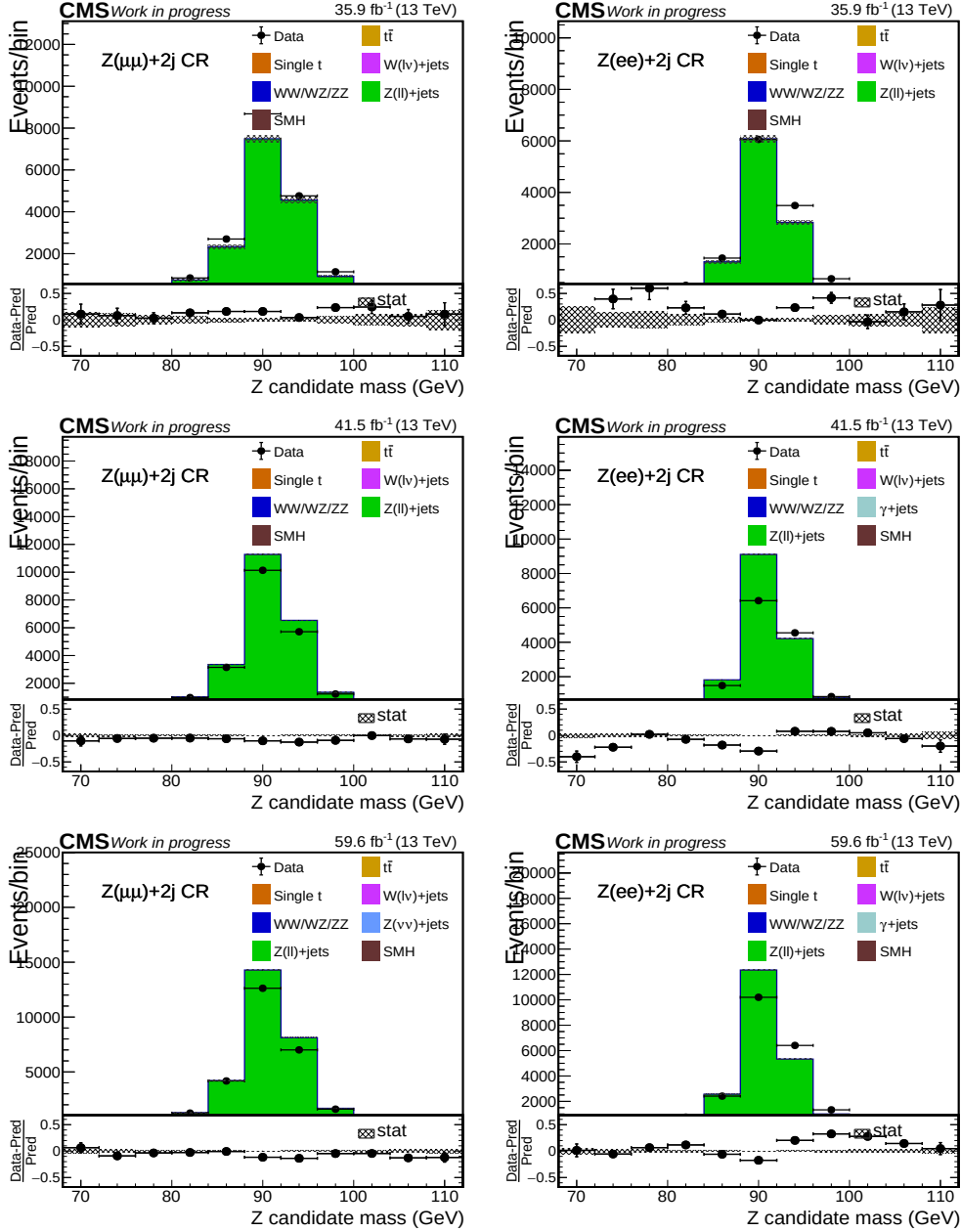


Figure 41: Z boson mass in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

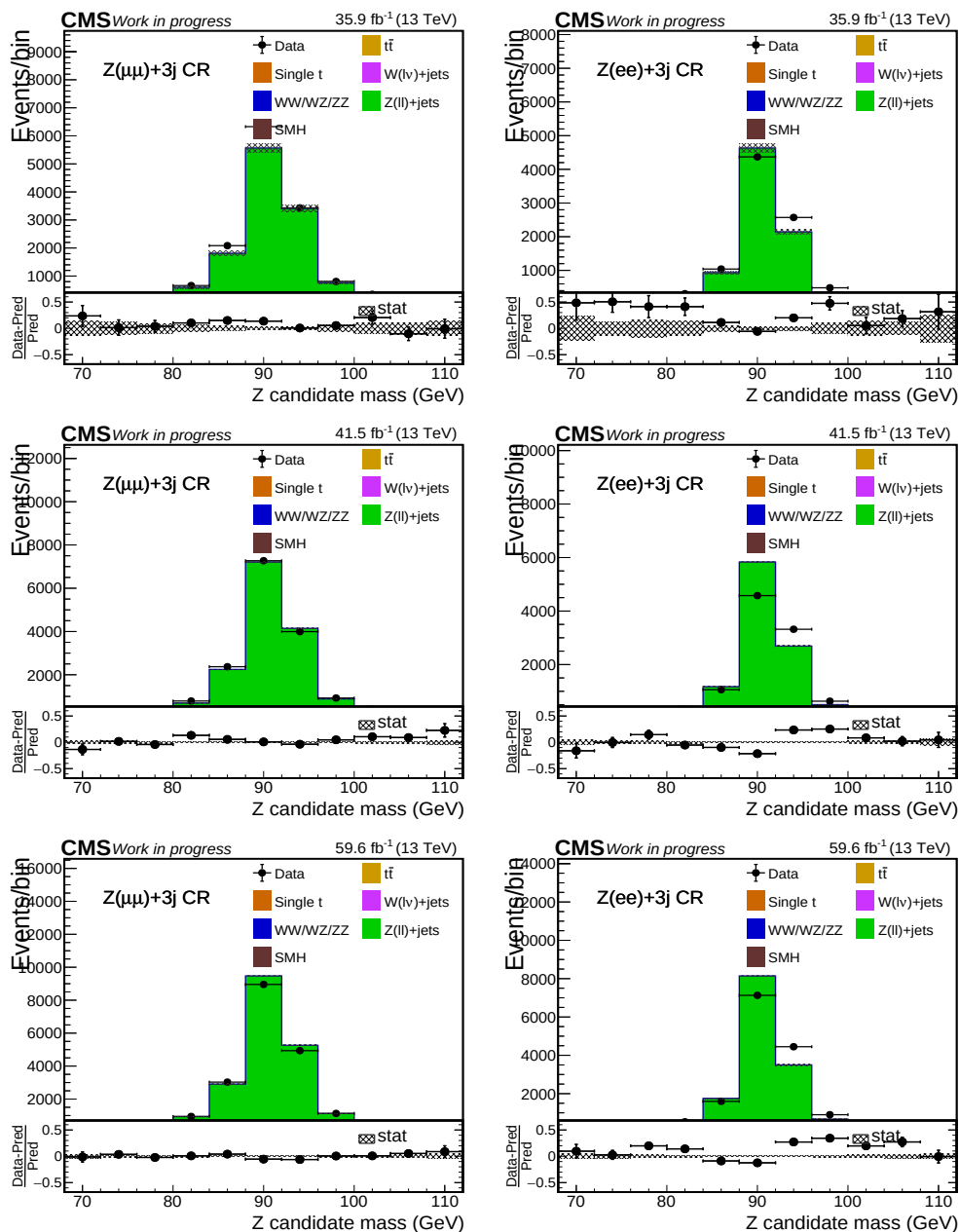


Figure 42: Z boson mass in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

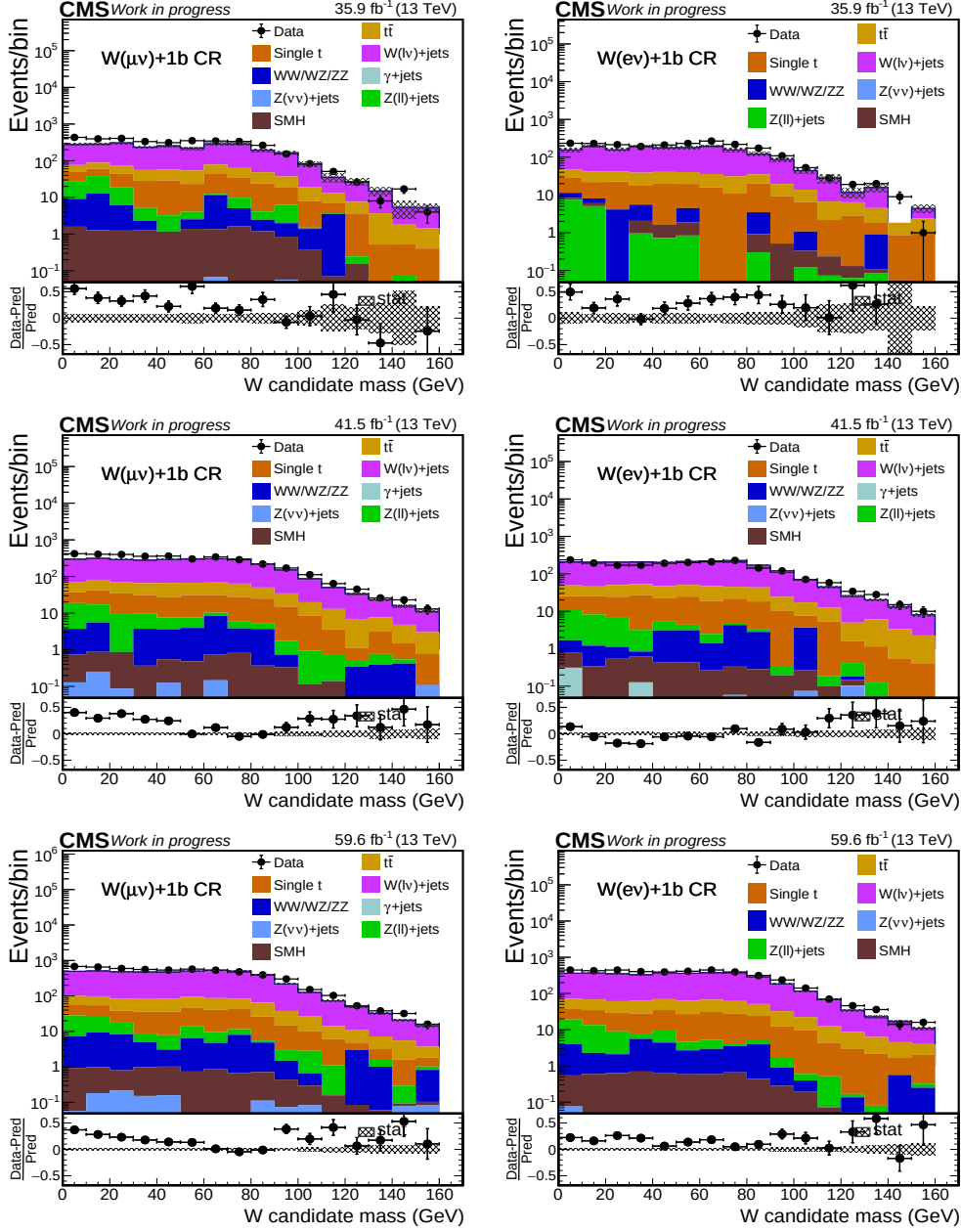


Figure 43: W boson mass in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

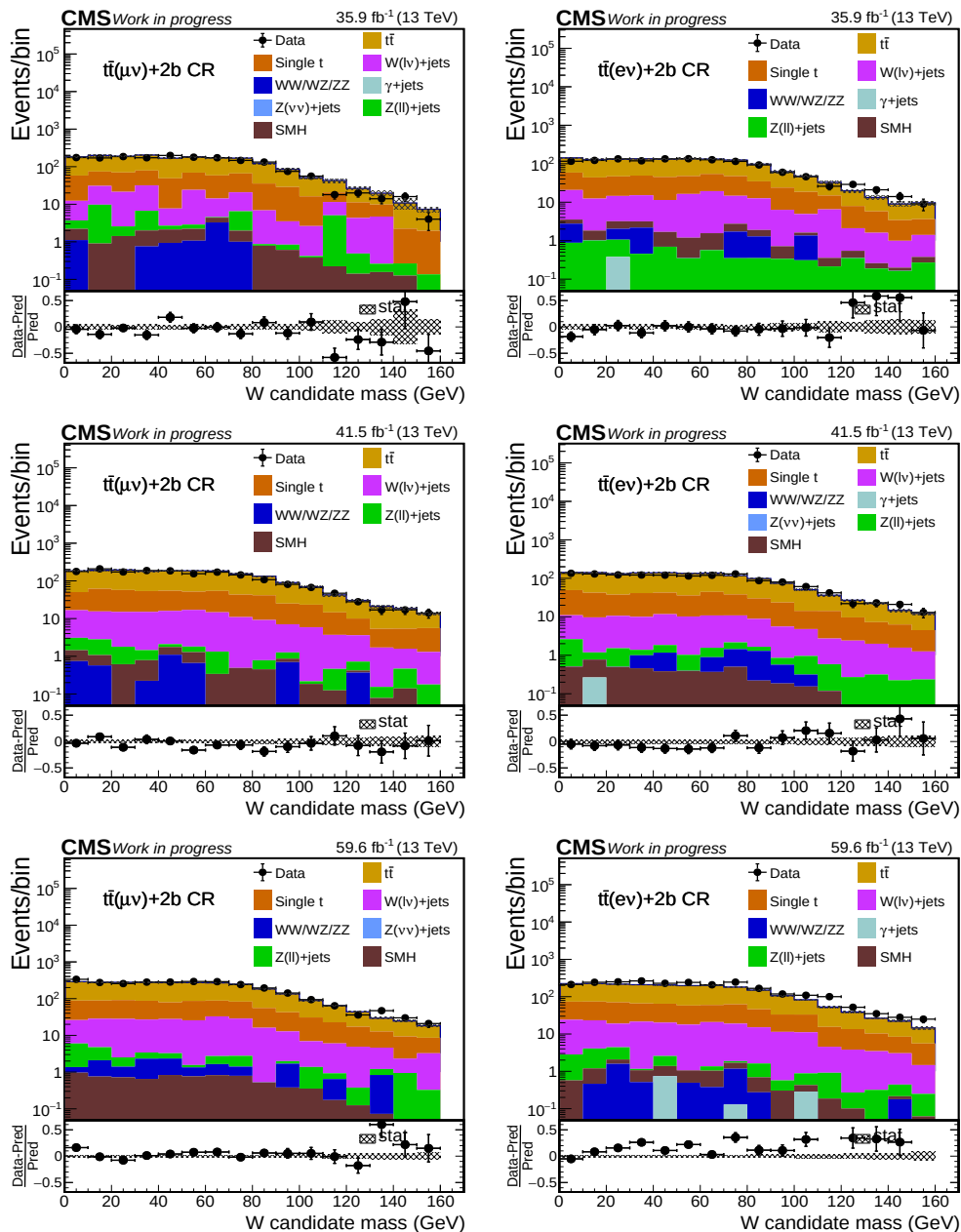


Figure 44: W boson mass in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

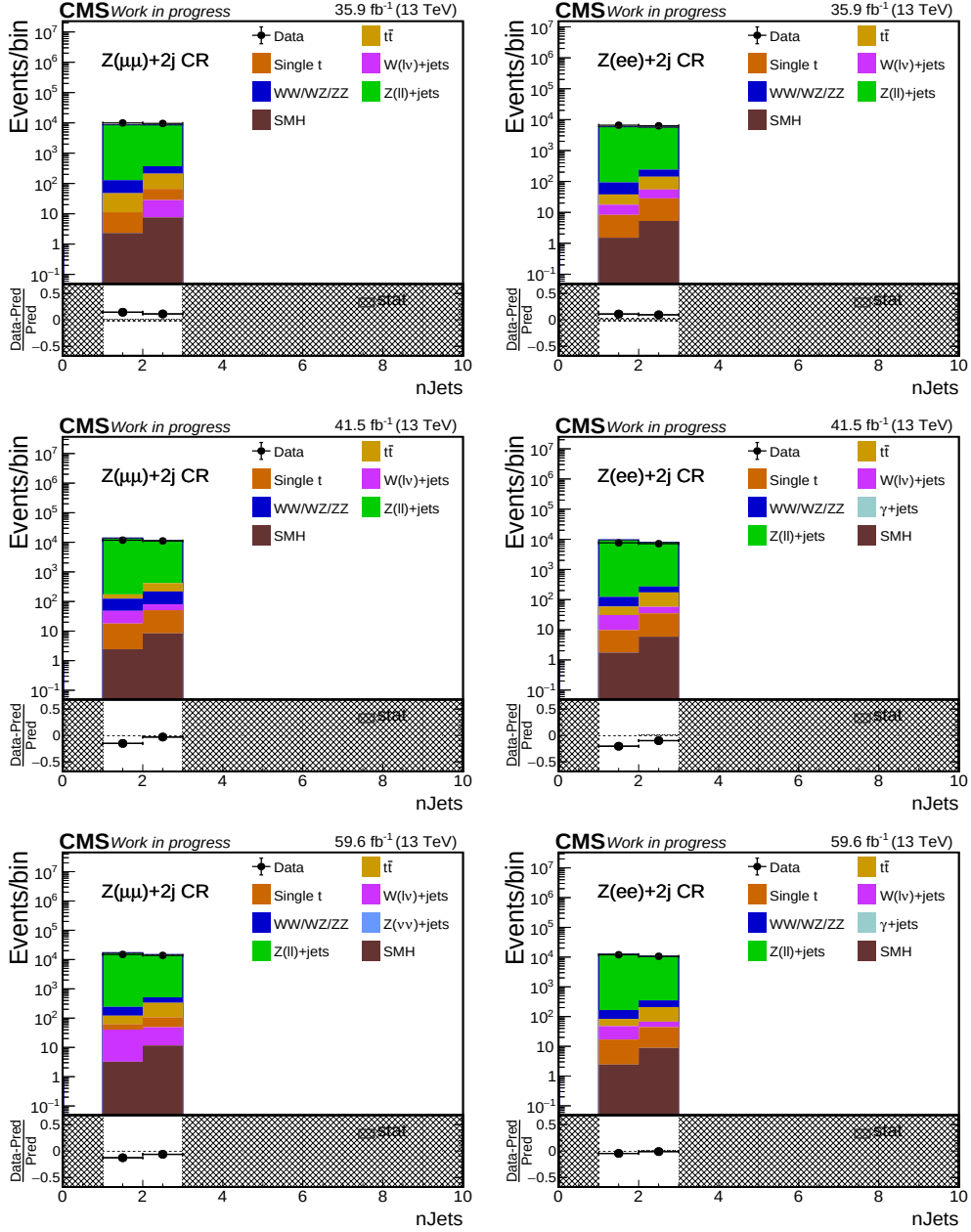


Figure 45: Number of jets in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+2j.

C. PLOTS OF OTHER VARIABLES FOR CRS

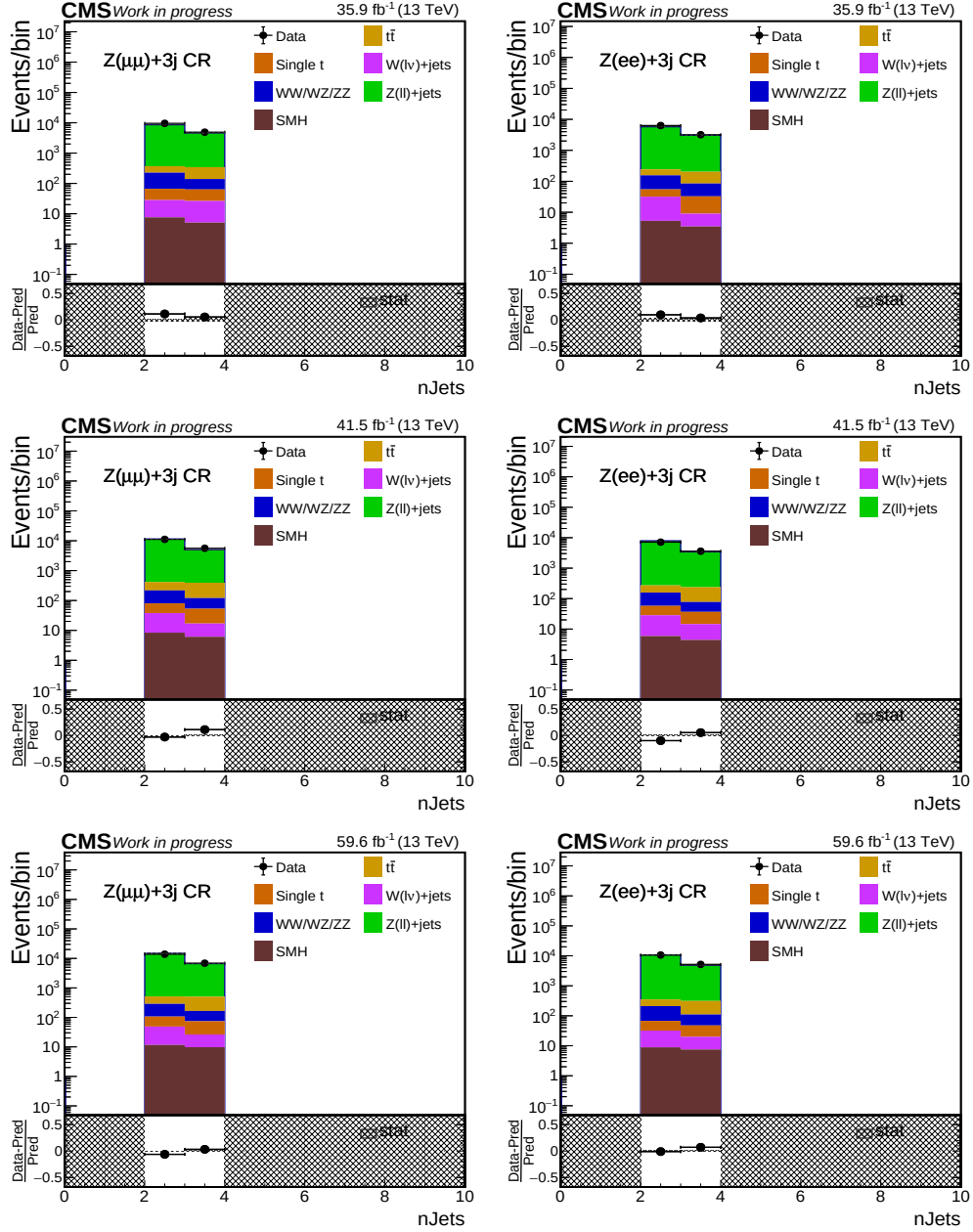


Figure 46: Number of jets in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ control region for 2016, 2017 and 2018. For Z CRs, the b-tagging requirement is relaxed hence they become Z+3j.

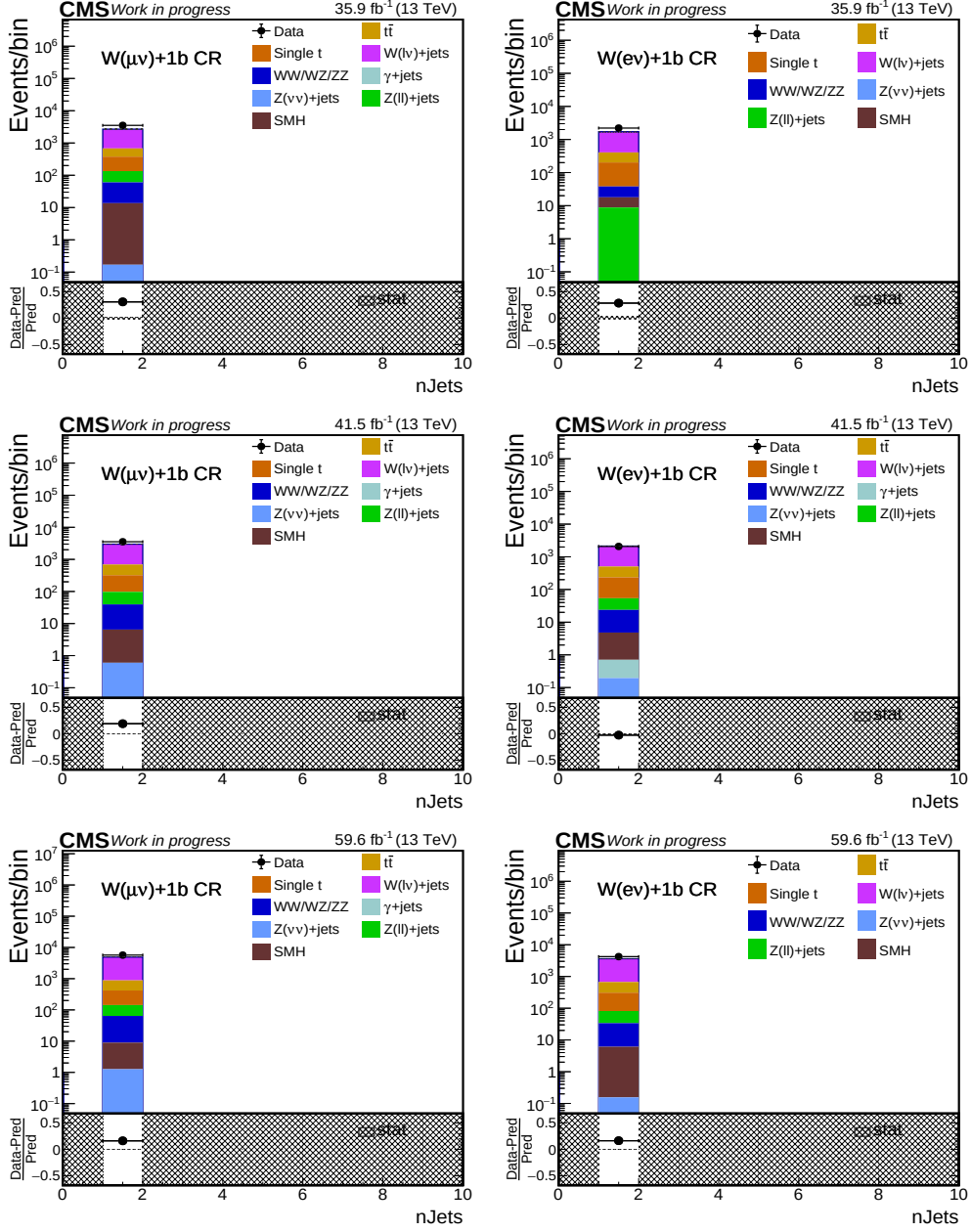


Figure 47: Number of jets in $W \rightarrow l\mu$ control region for 2016, 2017 and 2018.

C. PLOTS OF OTHER VARIABLES FOR CRS

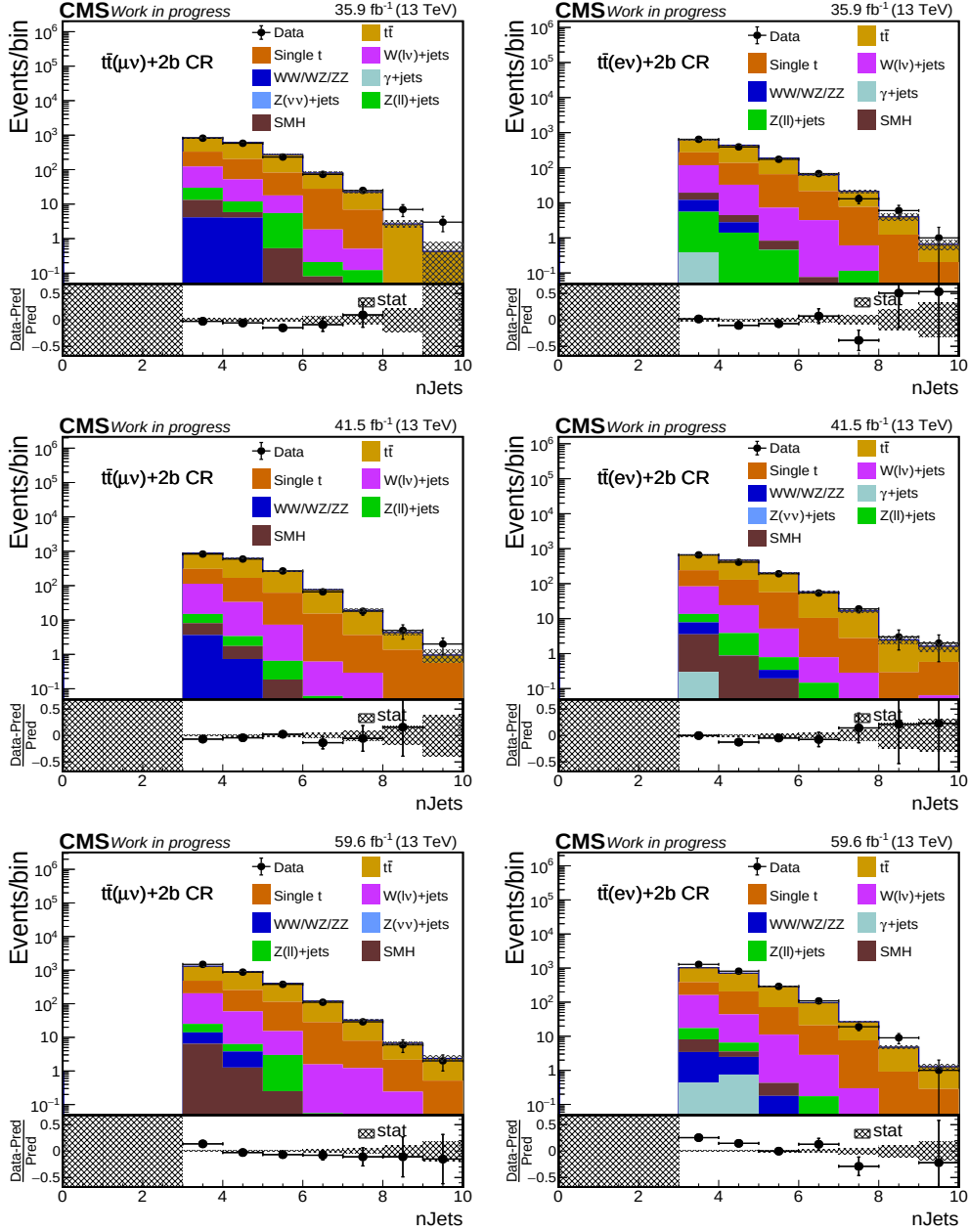


Figure 48: Number of jets in Top(μ) and Top(e) control region for 2016, 2017 and 2018.

D Systematic Uncertainty Plots

D. SYSTEMATIC UNCERTAINTY PLOTS

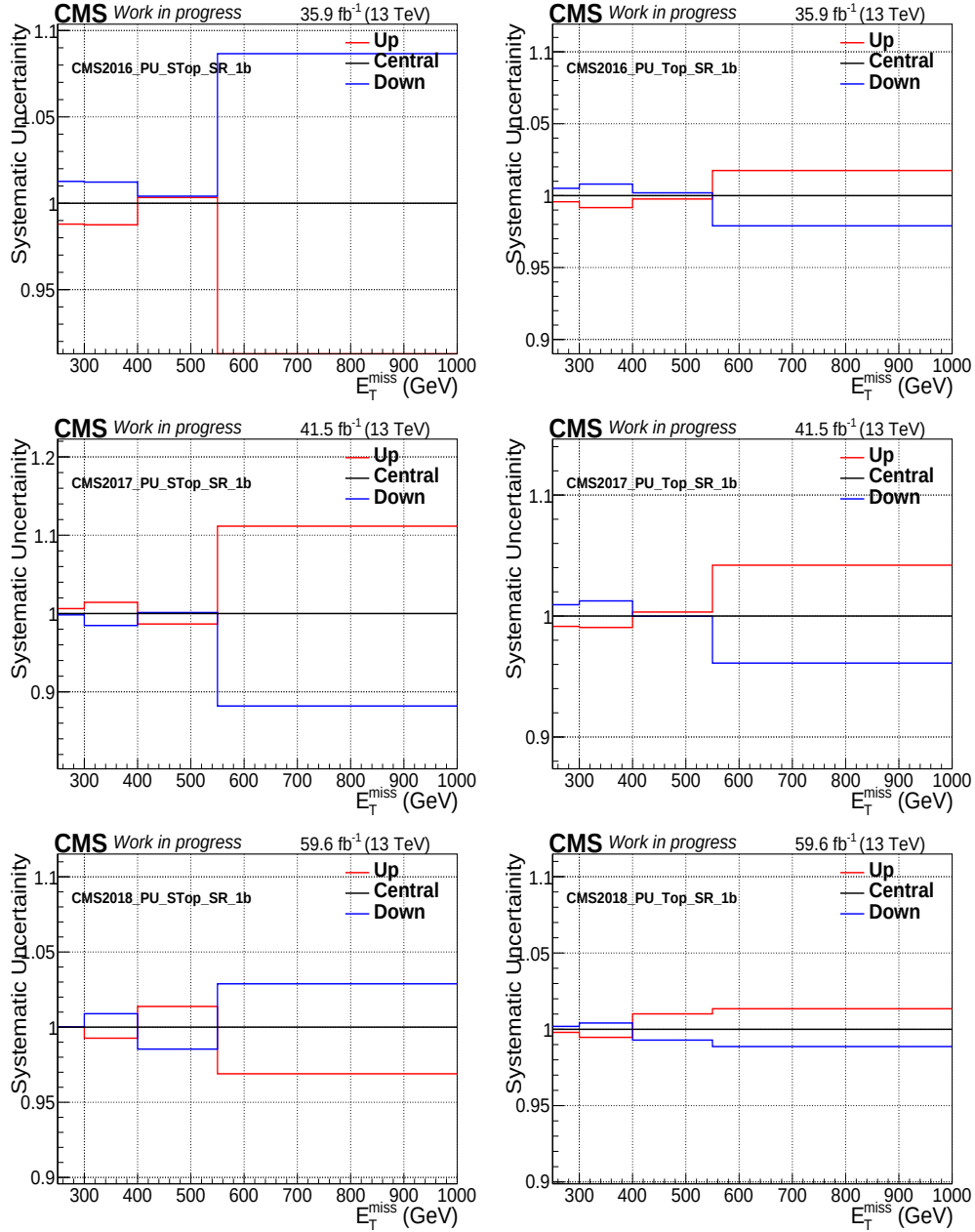


Figure 49: PileUP uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

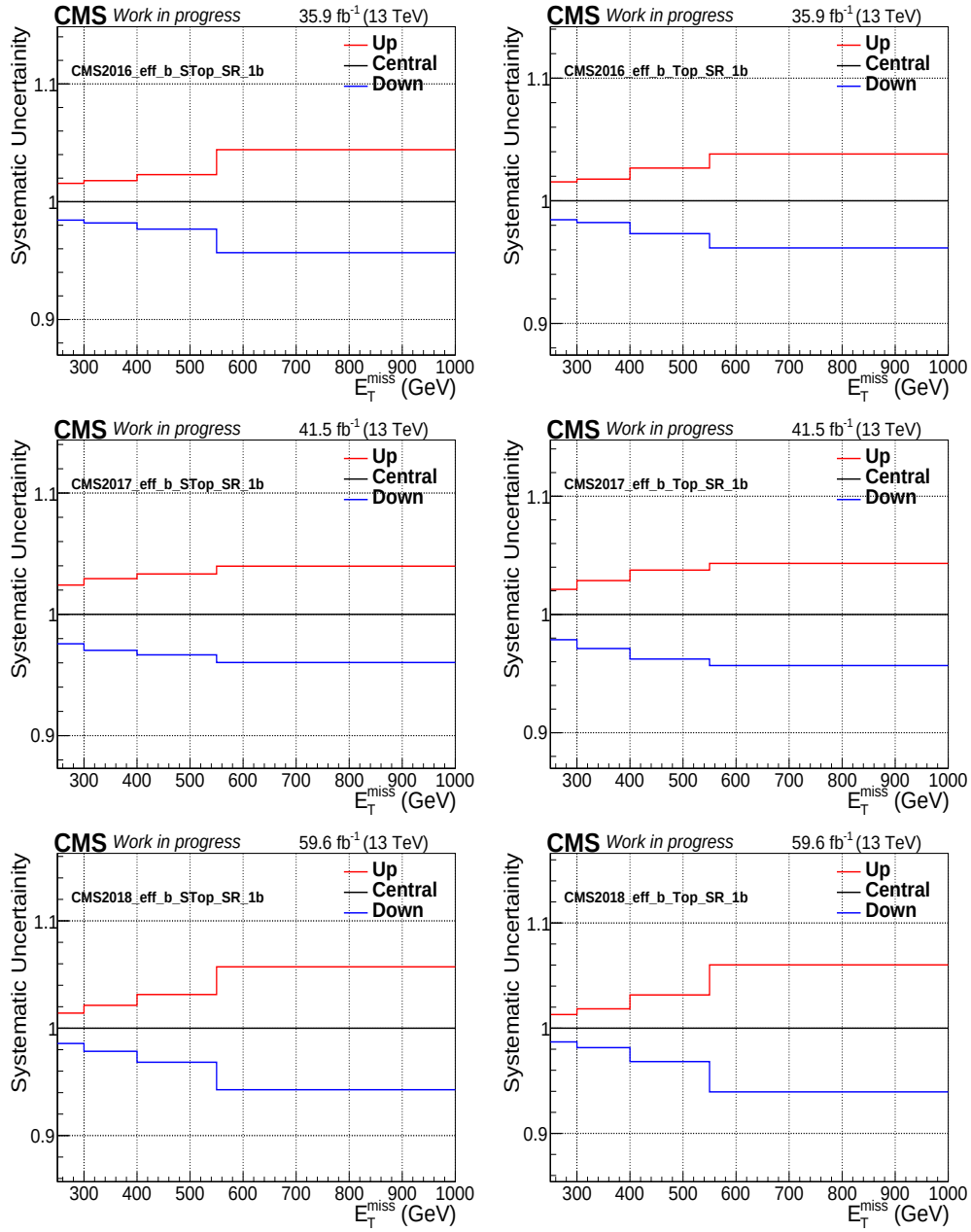


Figure 50: b-tag uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

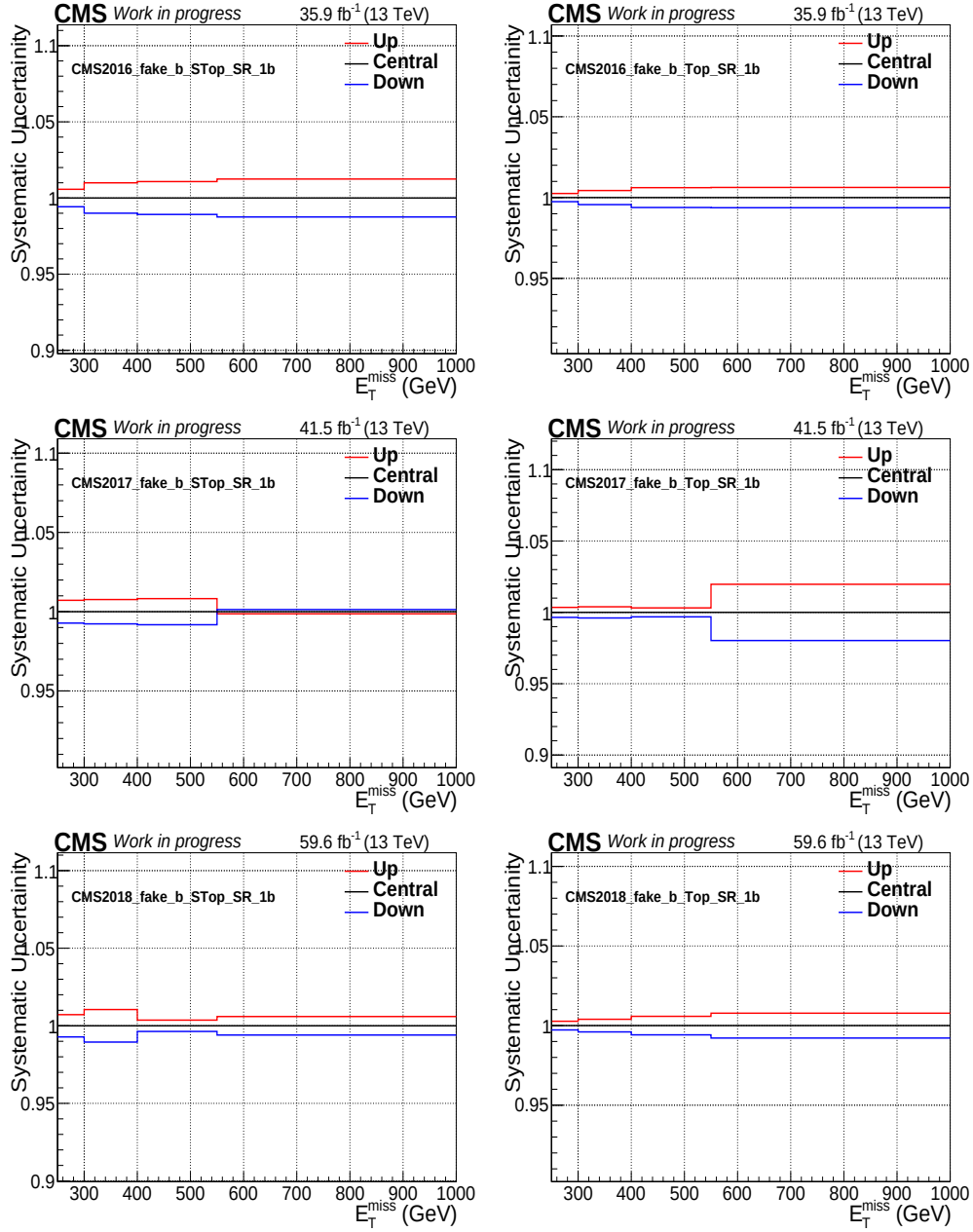


Figure 51: fake b-tag uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

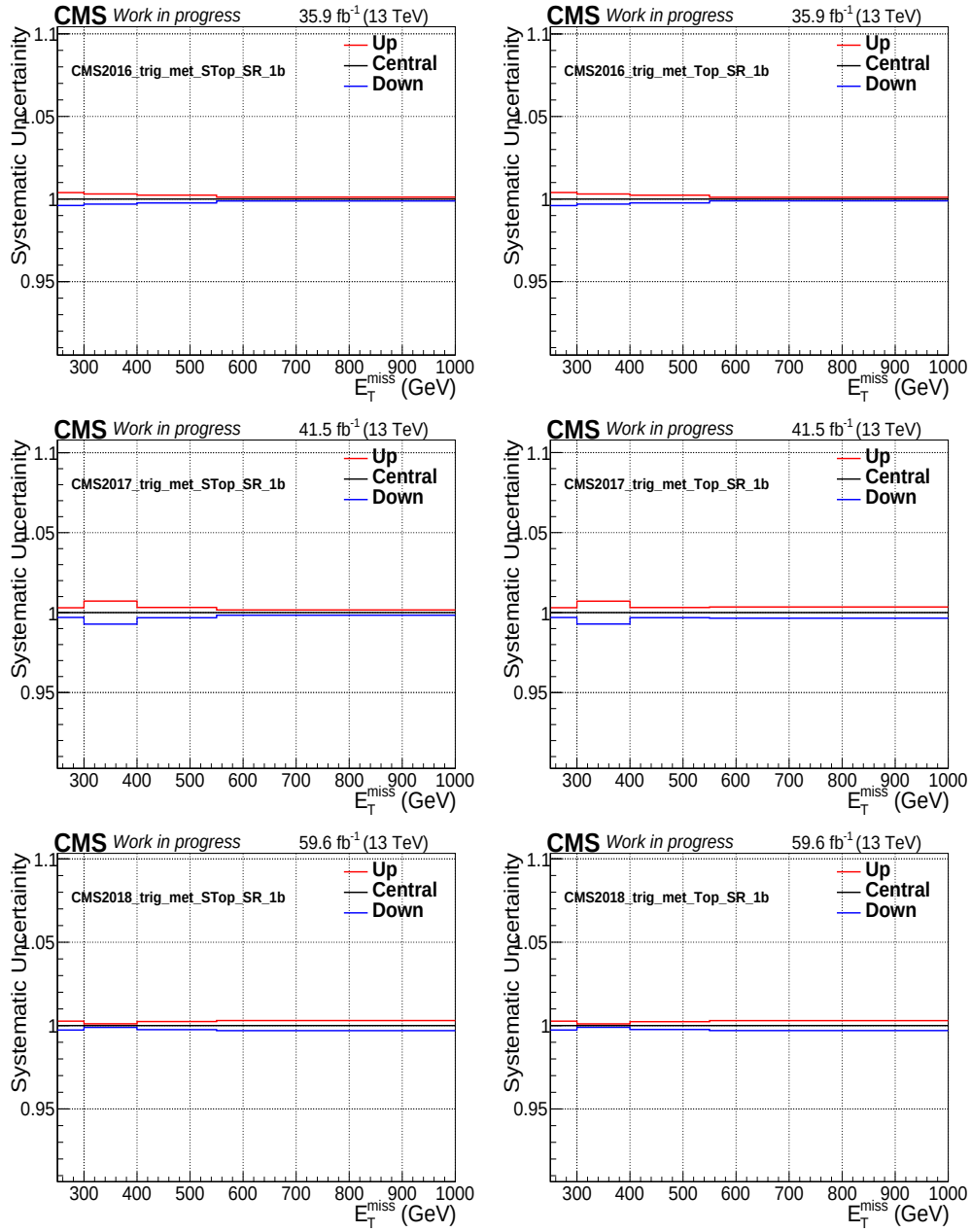


Figure 52: E_T^{miss} trigger uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

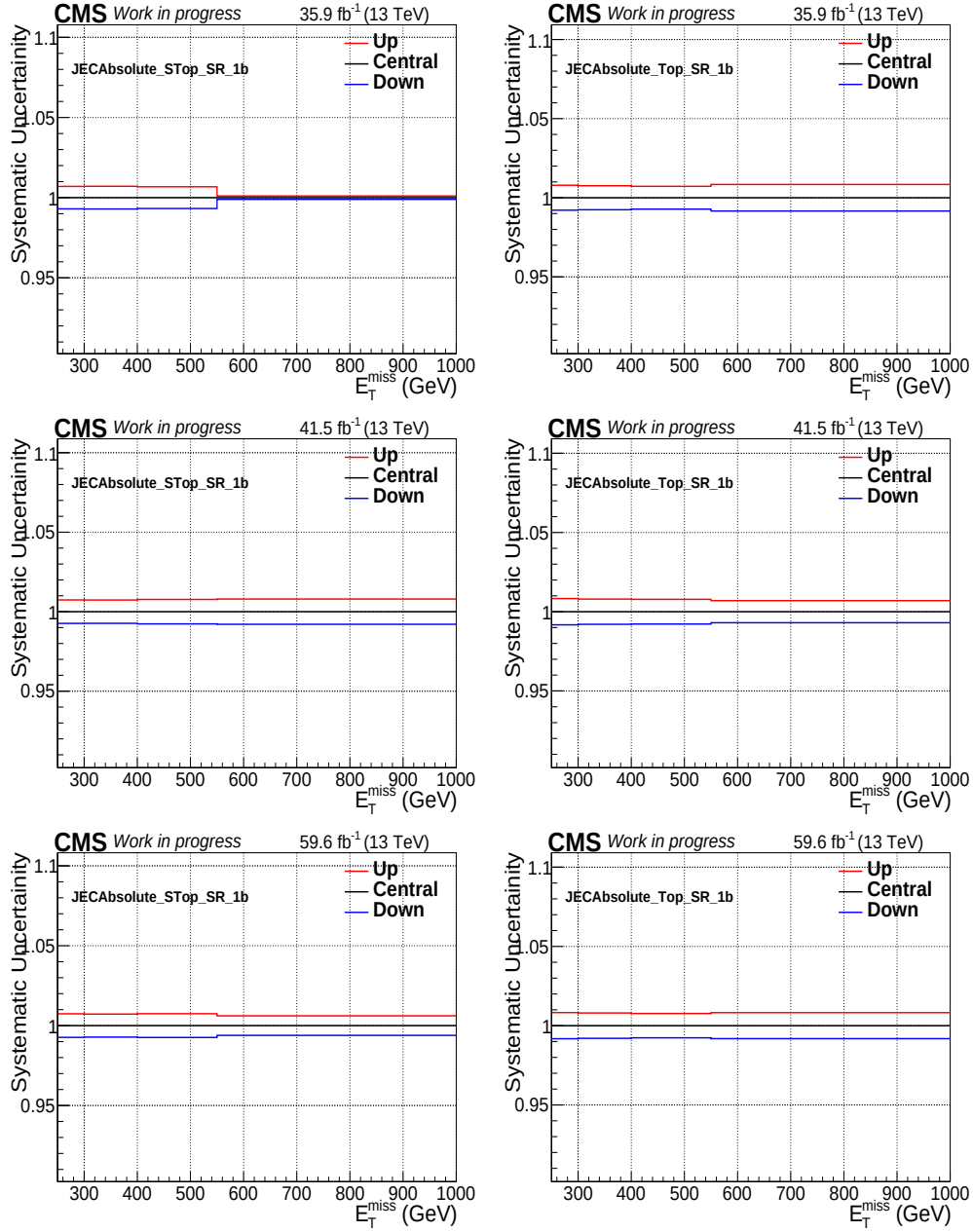


Figure 53: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

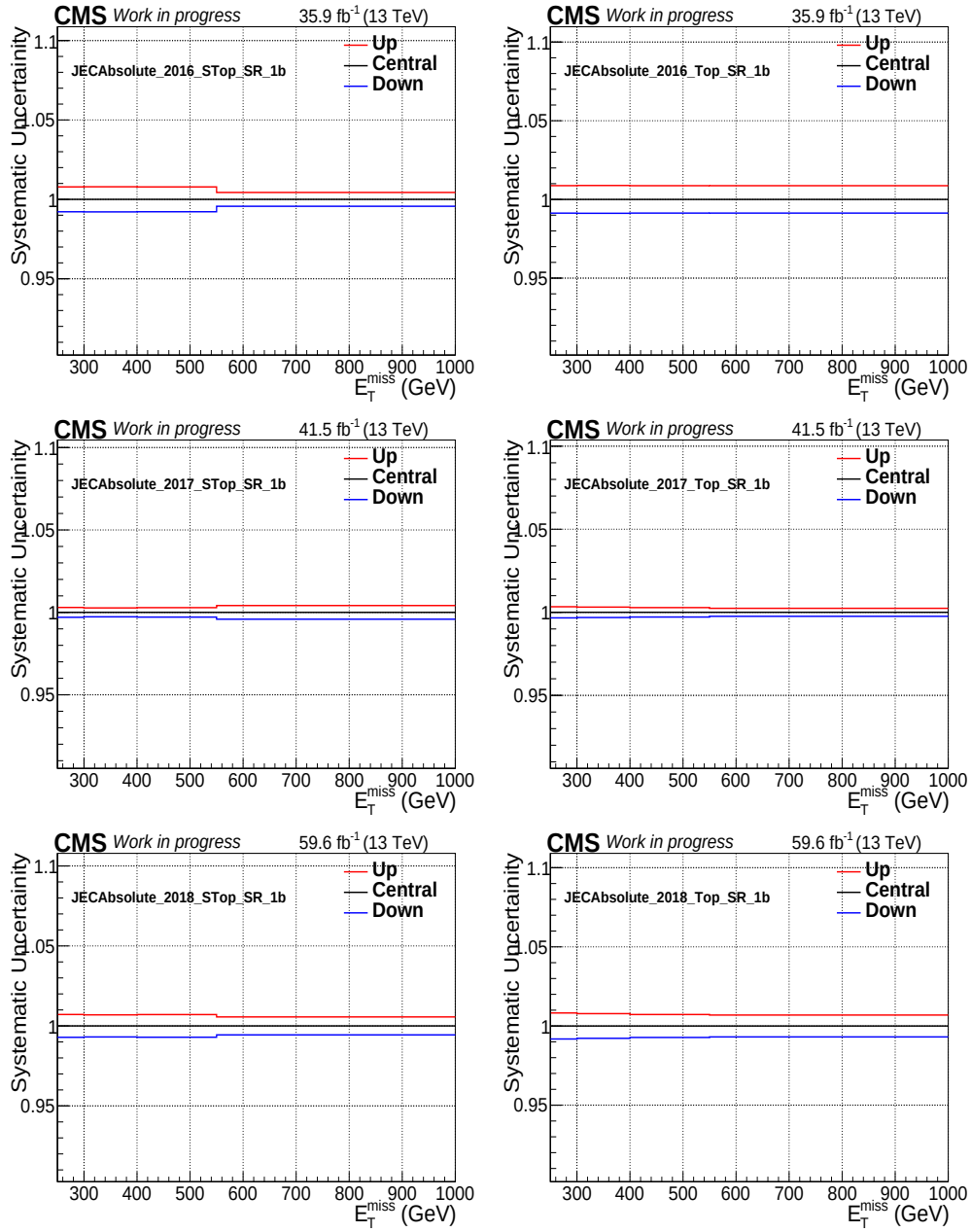


Figure 54: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

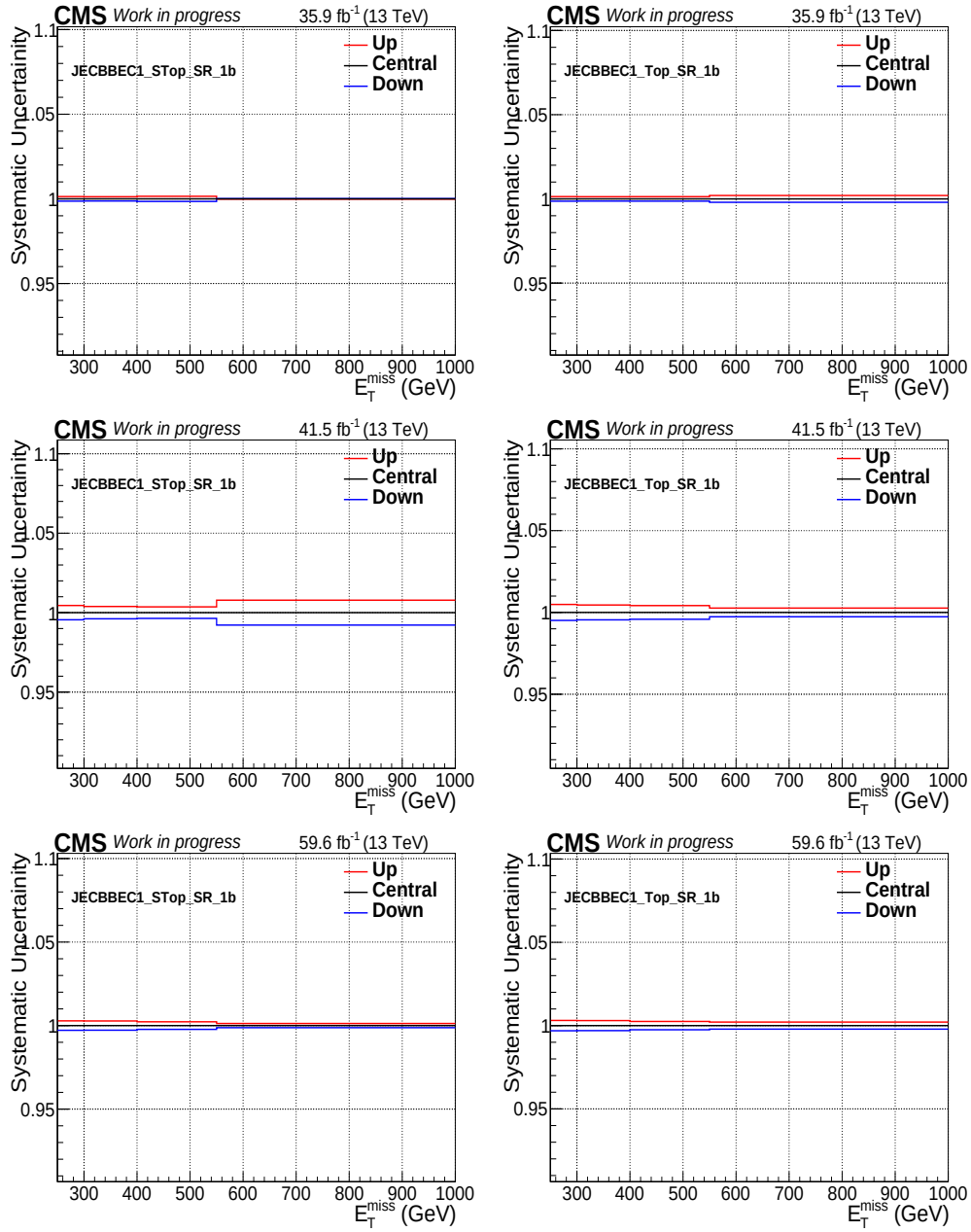


Figure 55: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

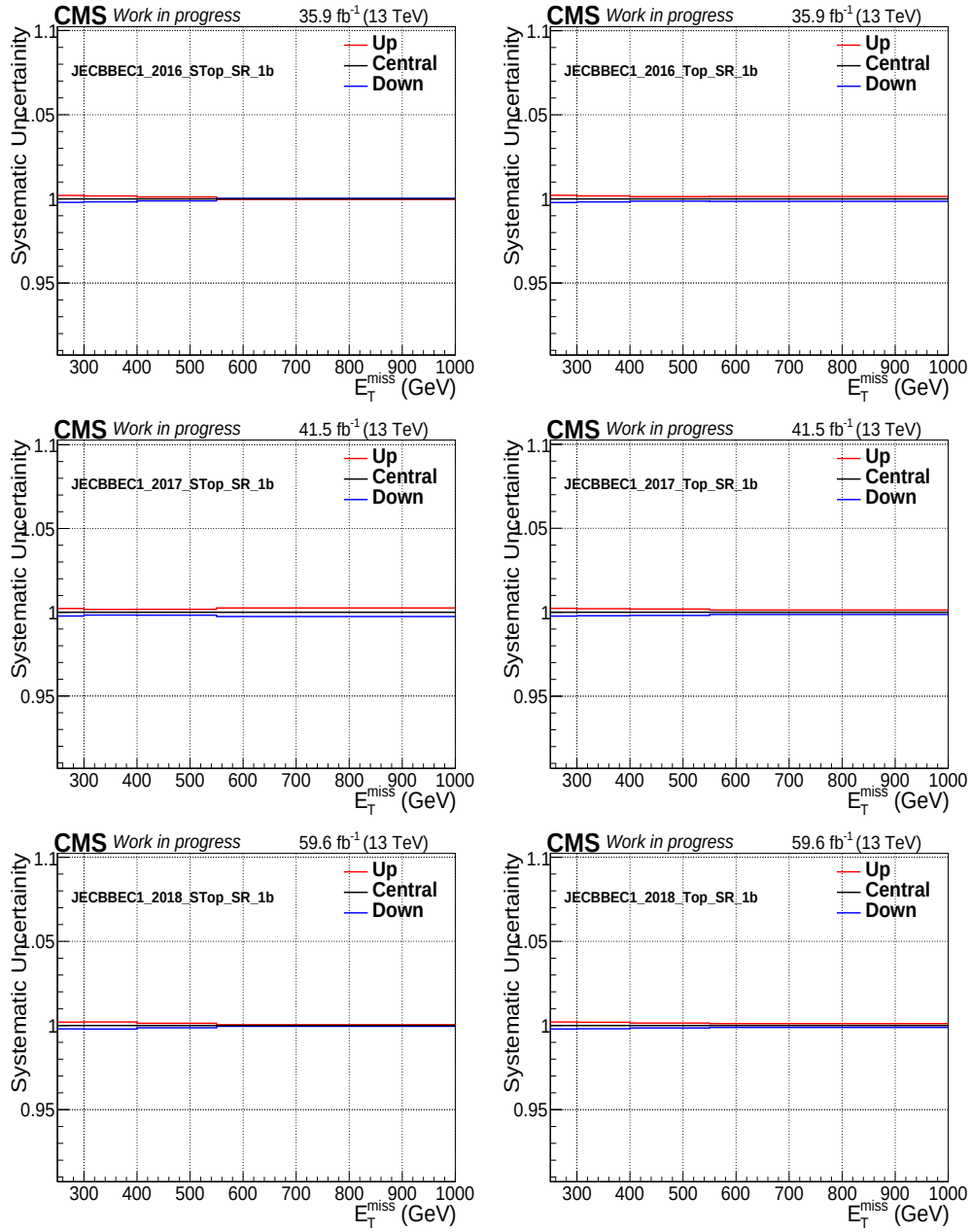


Figure 56: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

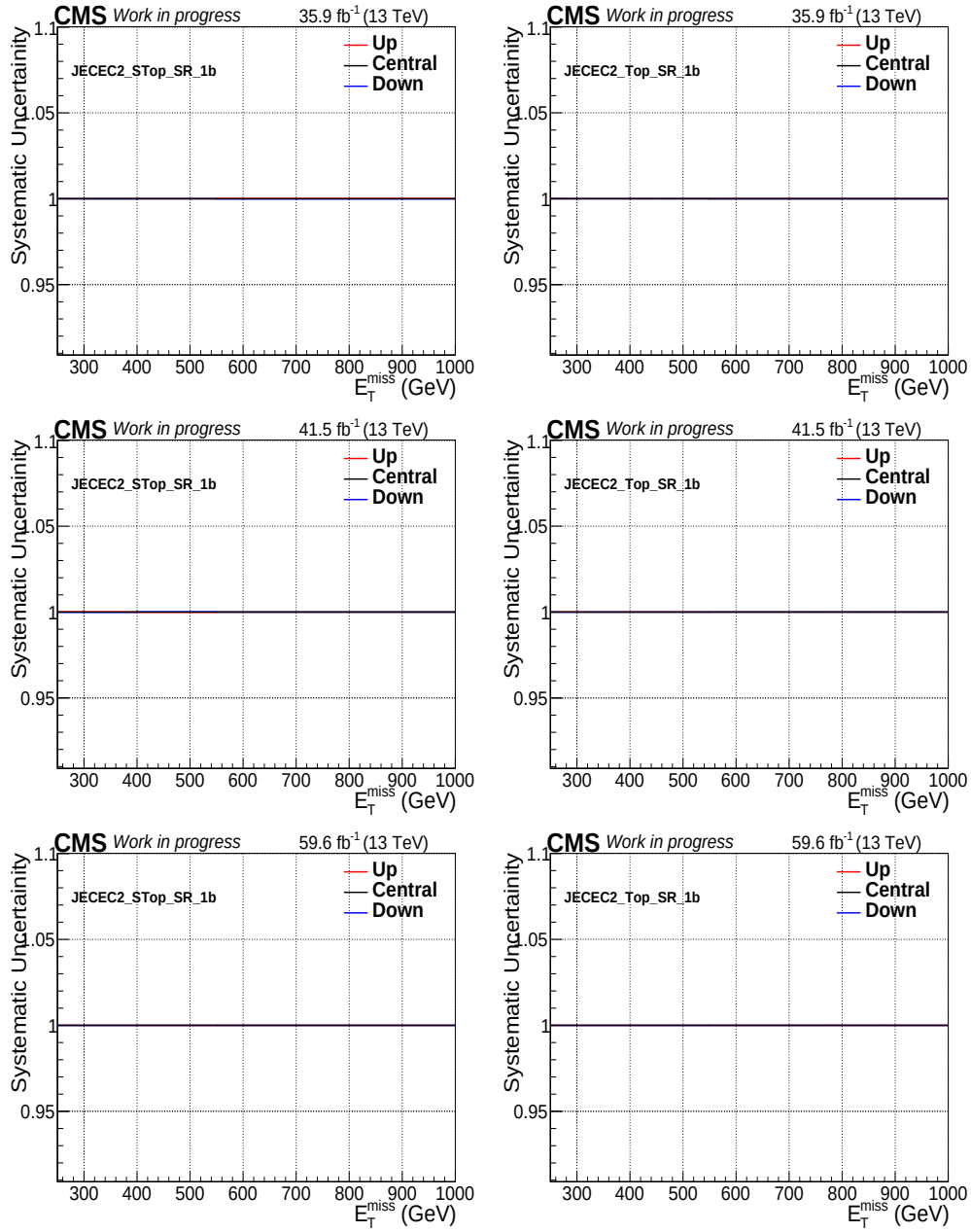


Figure 57: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

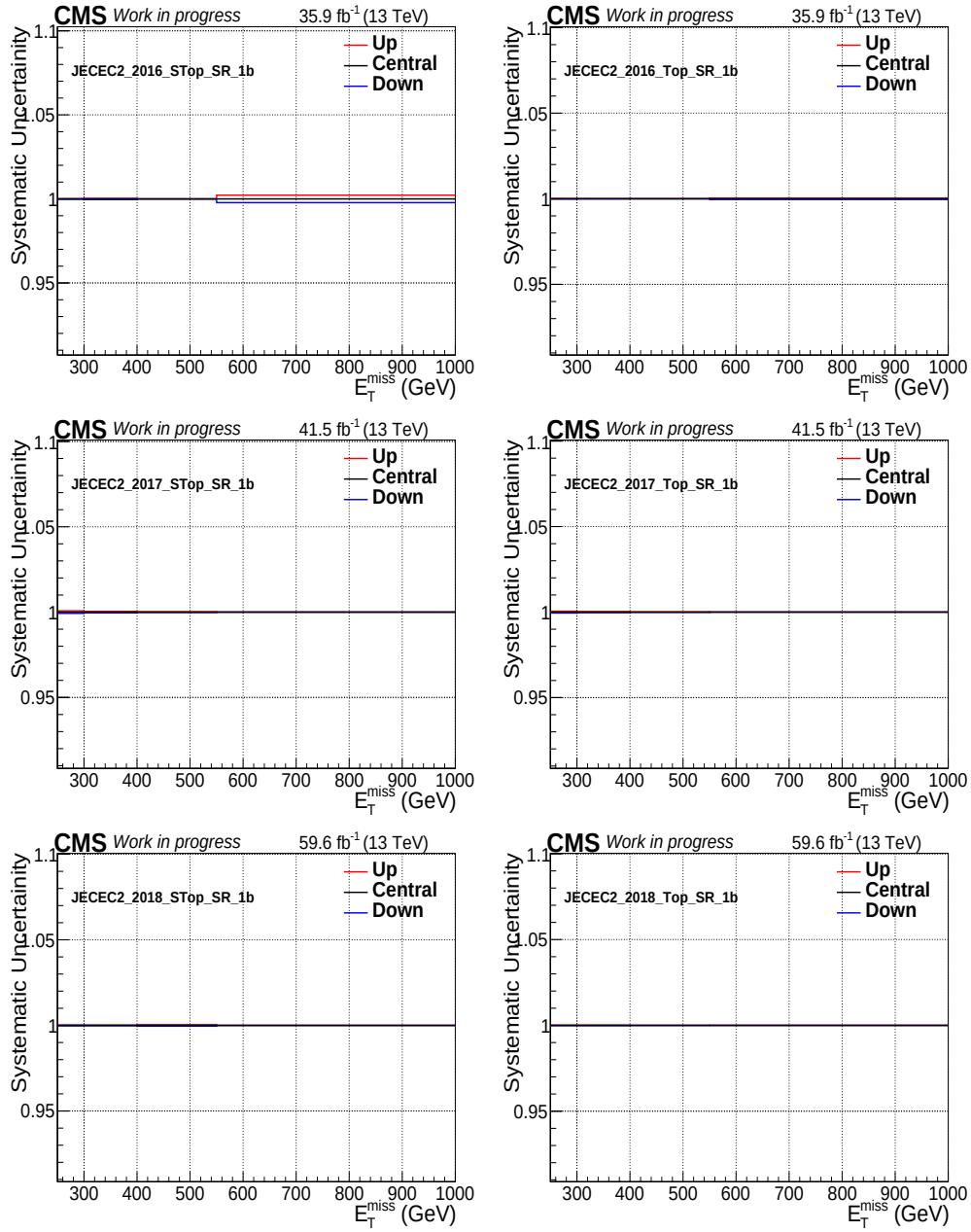


Figure 58: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

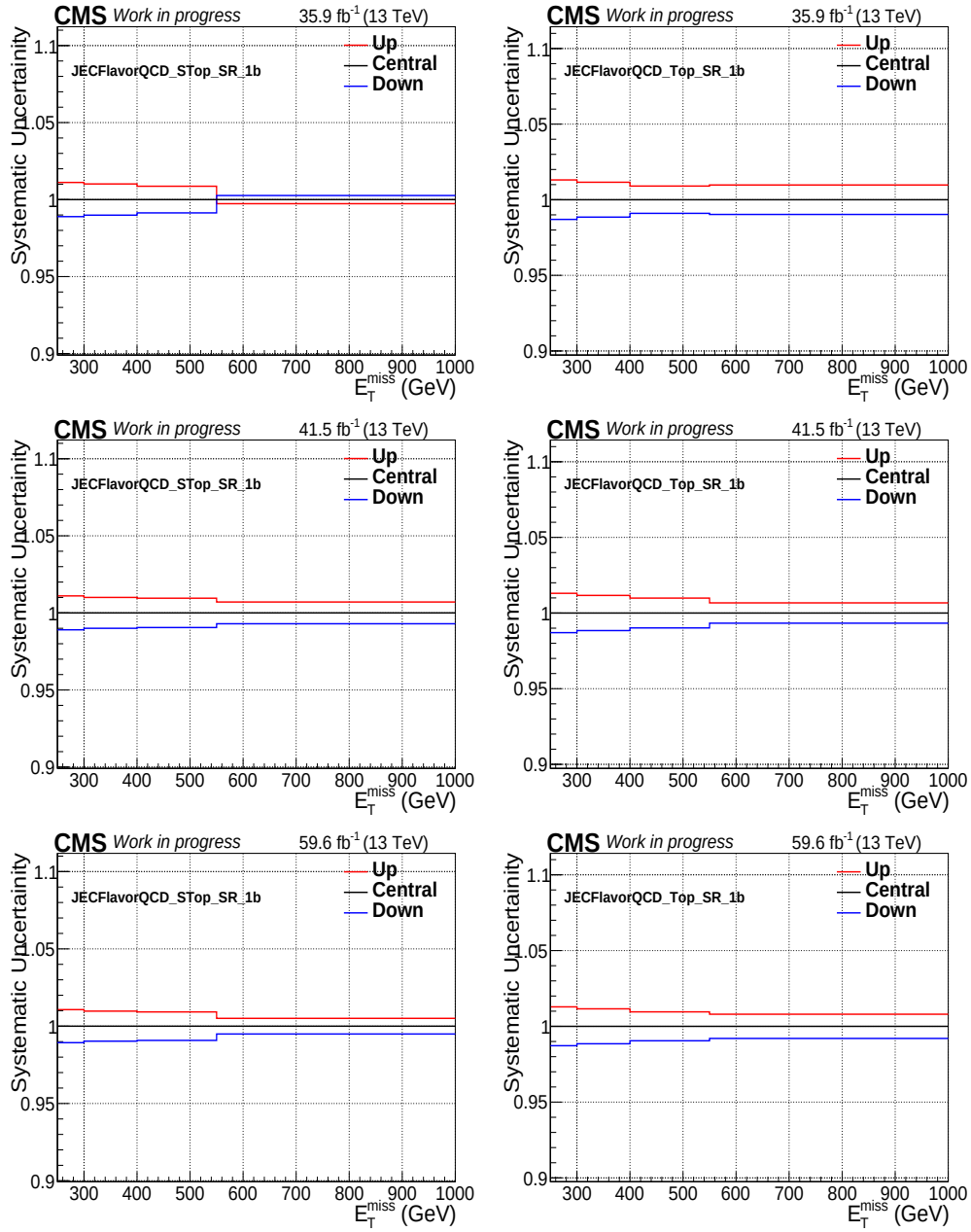


Figure 59: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

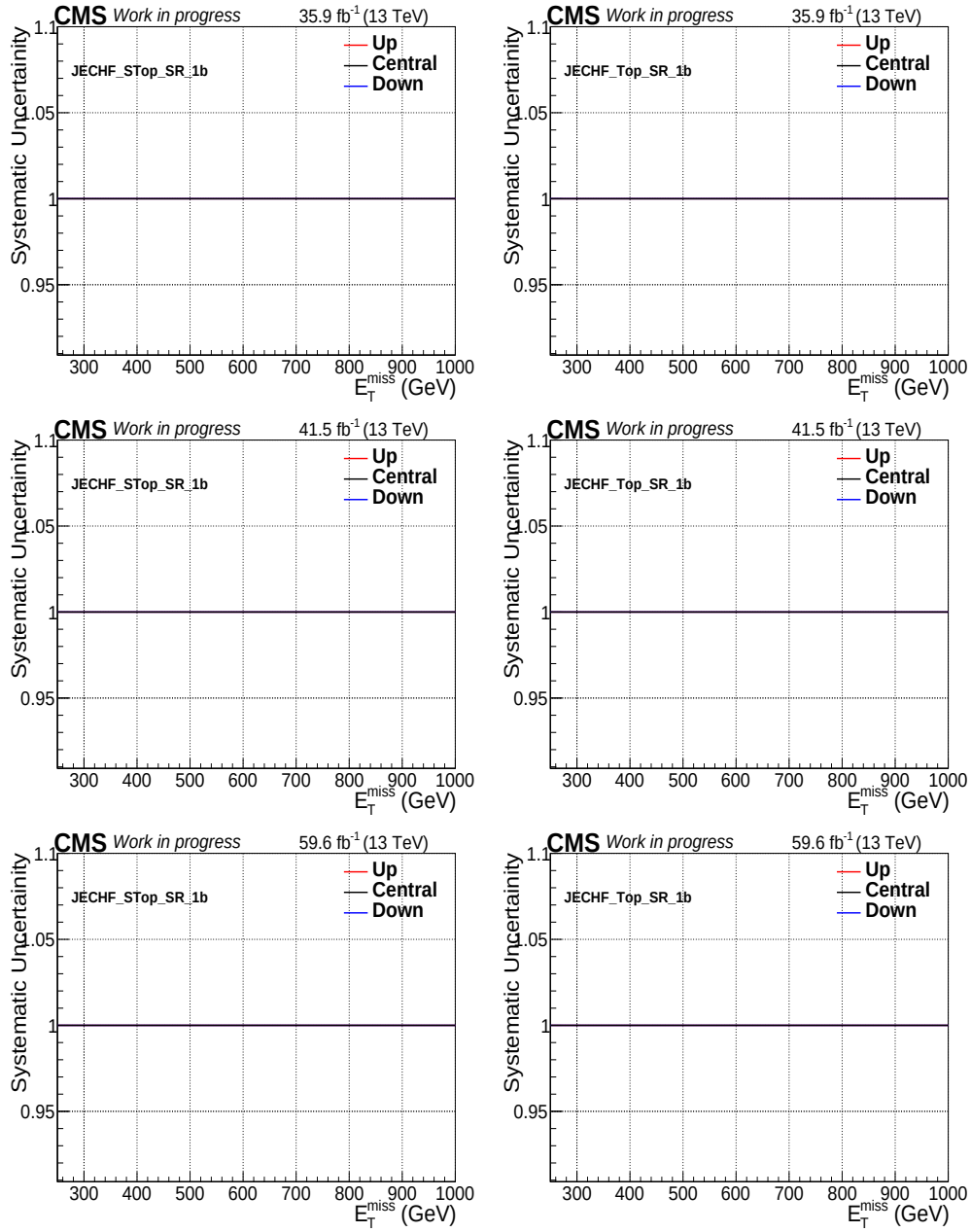


Figure 60: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

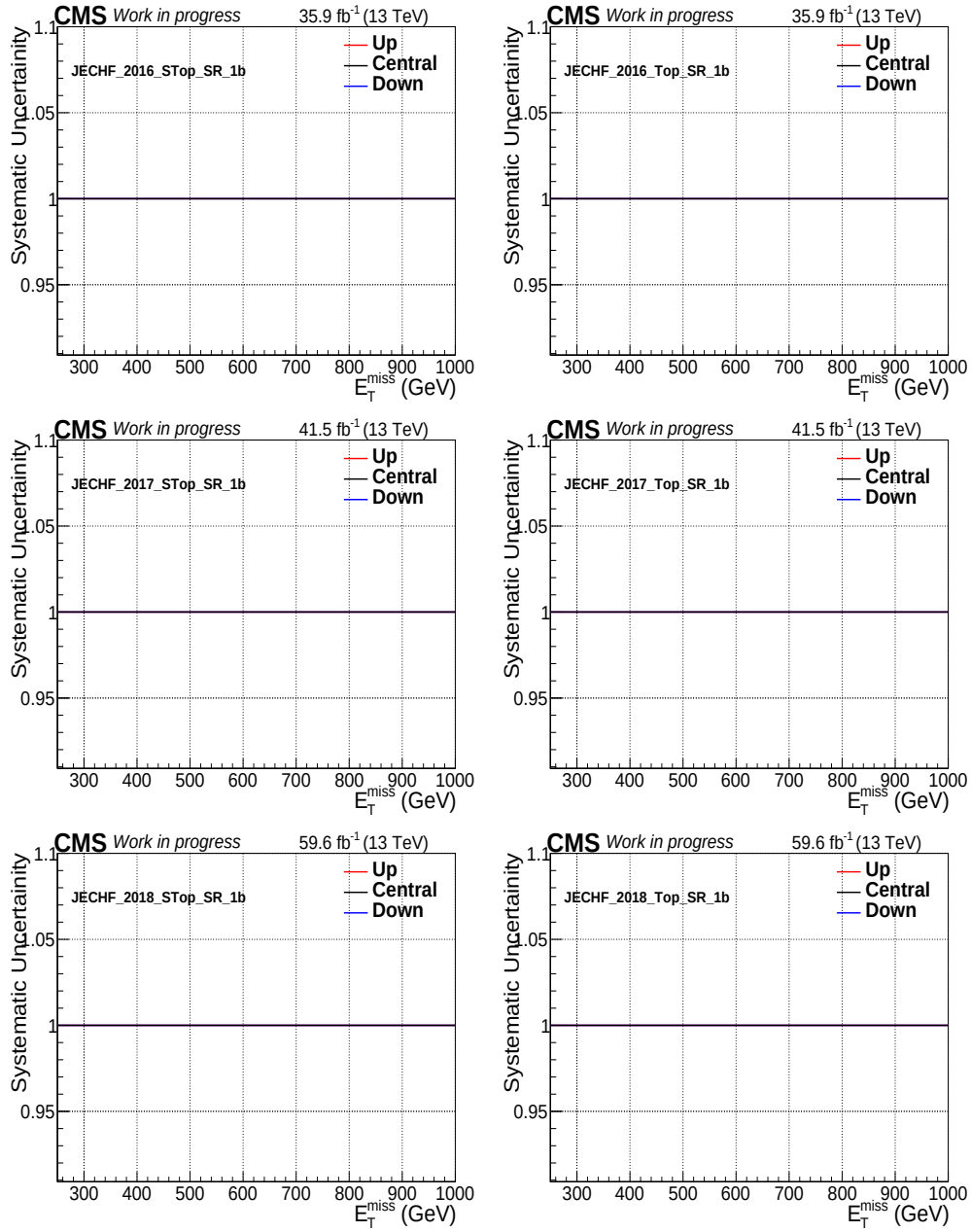


Figure 61: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

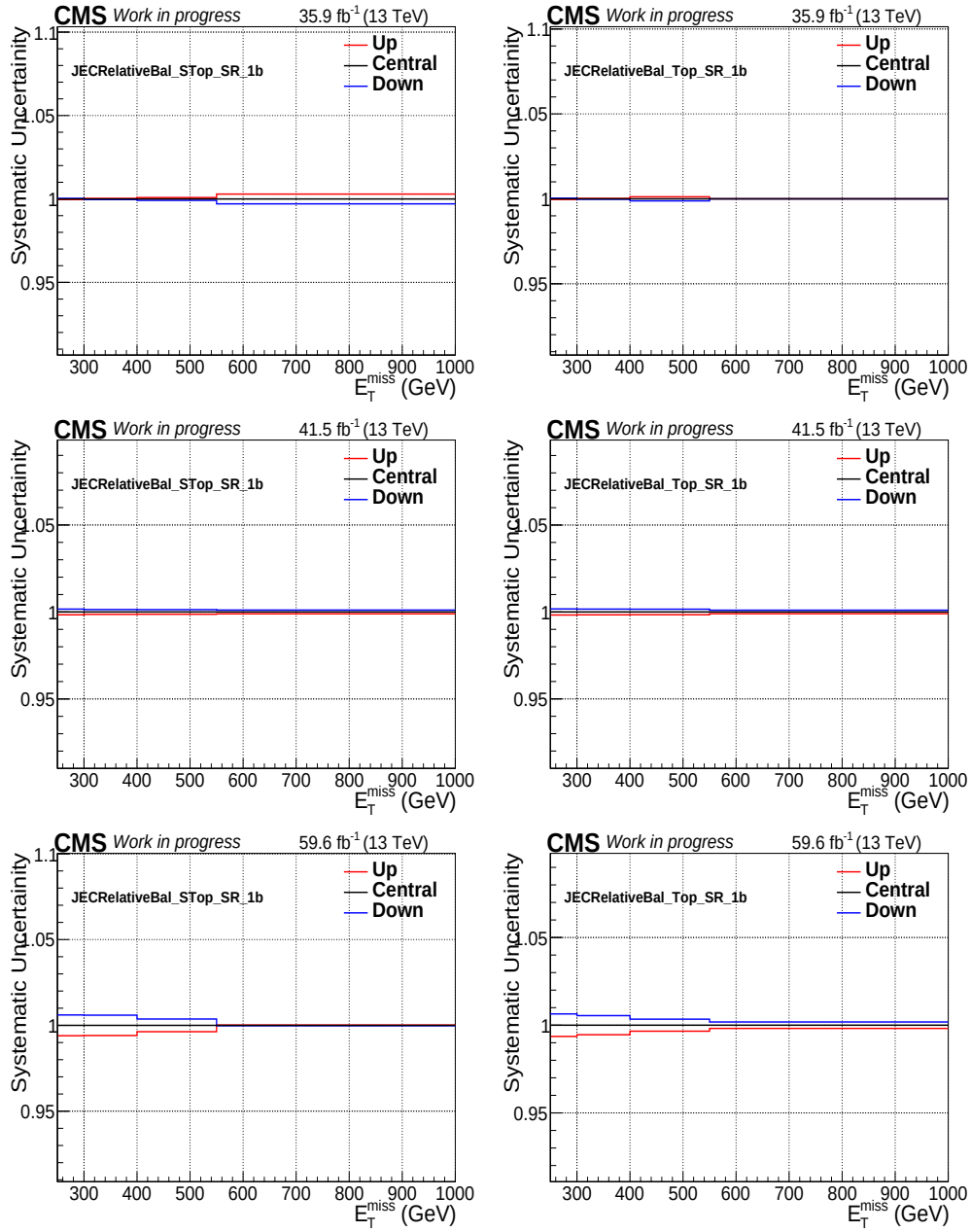


Figure 62: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

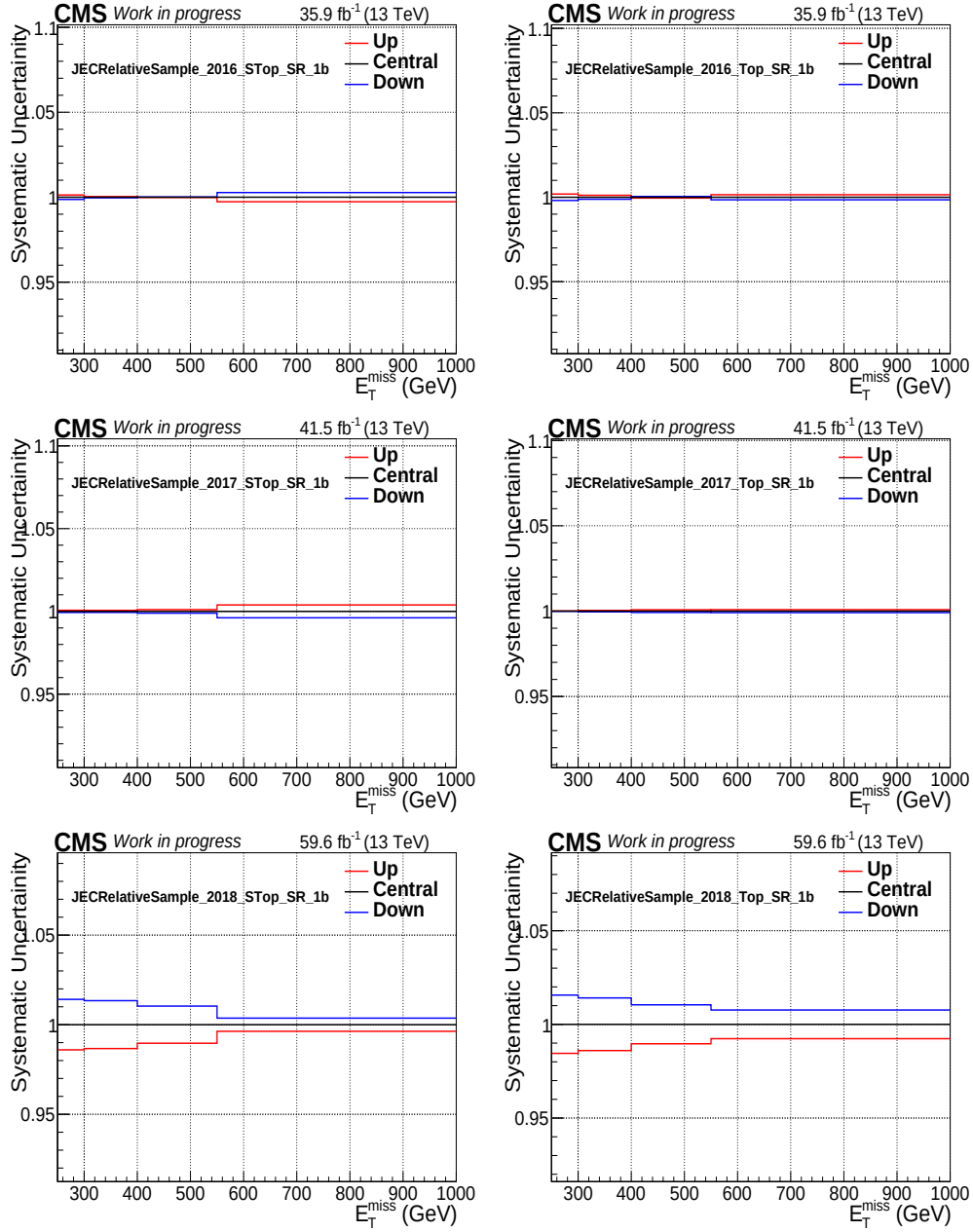


Figure 63: JEC uncertainty for single t and $t\bar{t}$ background 1b Category of Signal region for 2016, 2017 and 2018

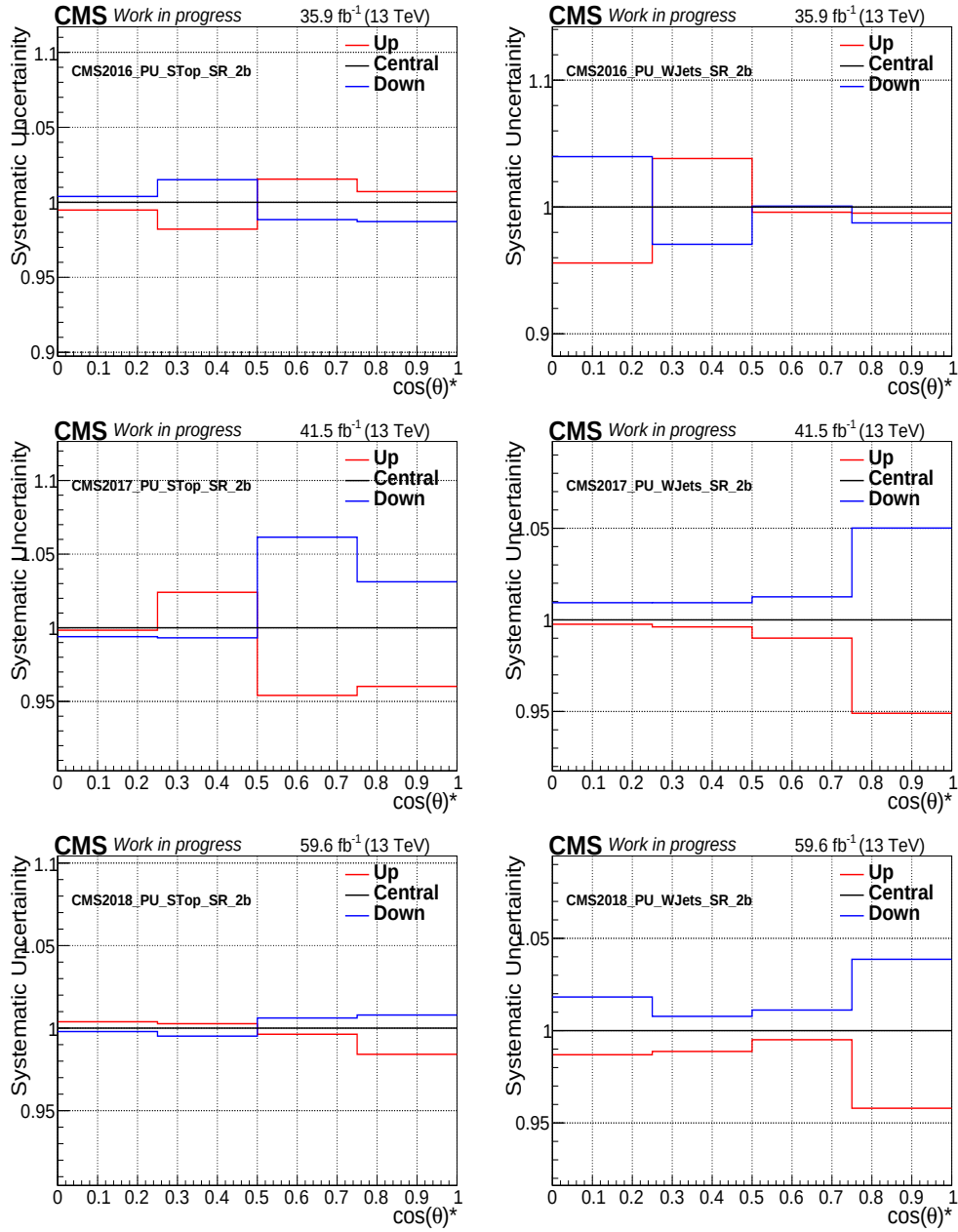


Figure 64: PileUP uncertainty for single t and Wjets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

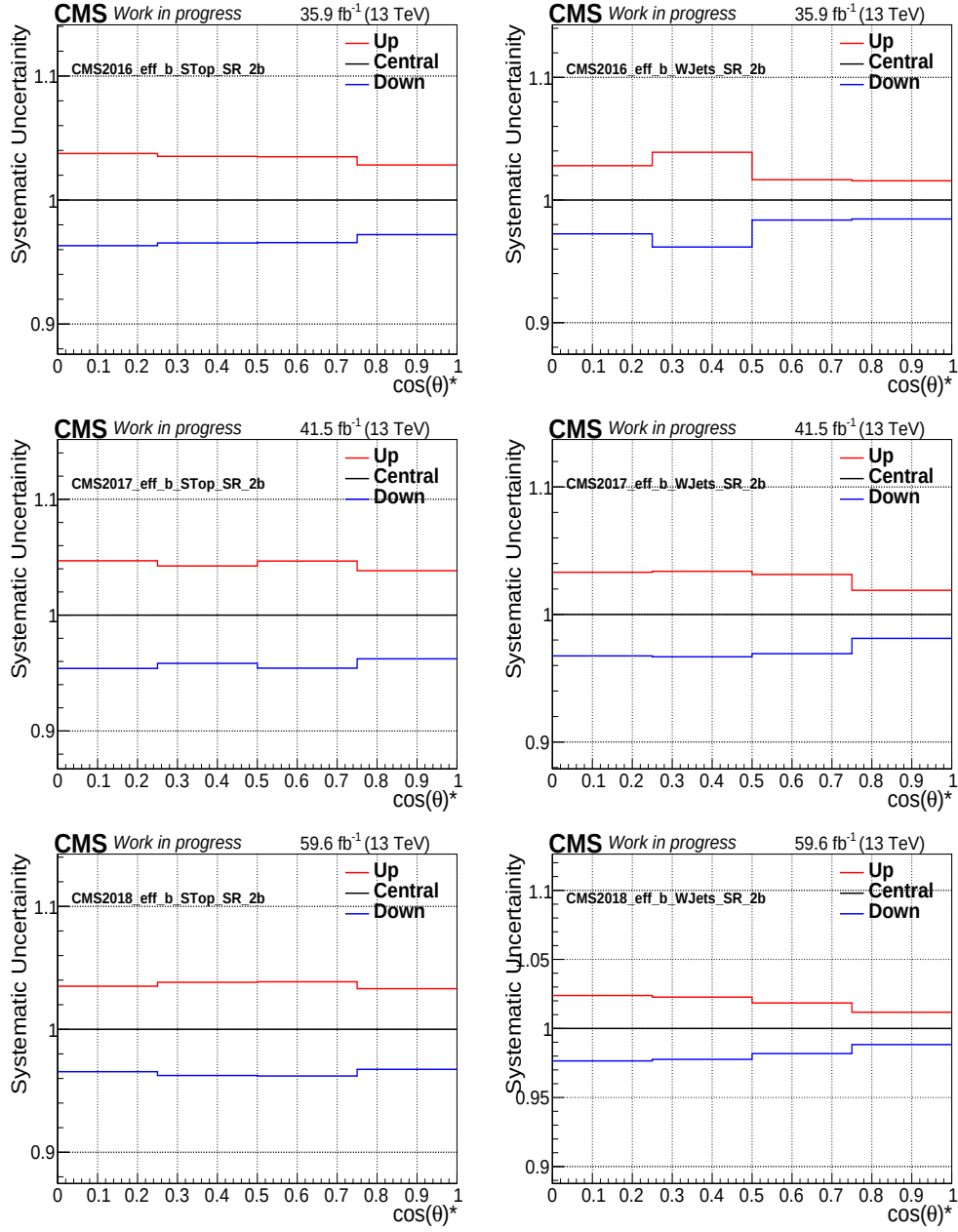


Figure 65: b-tag uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

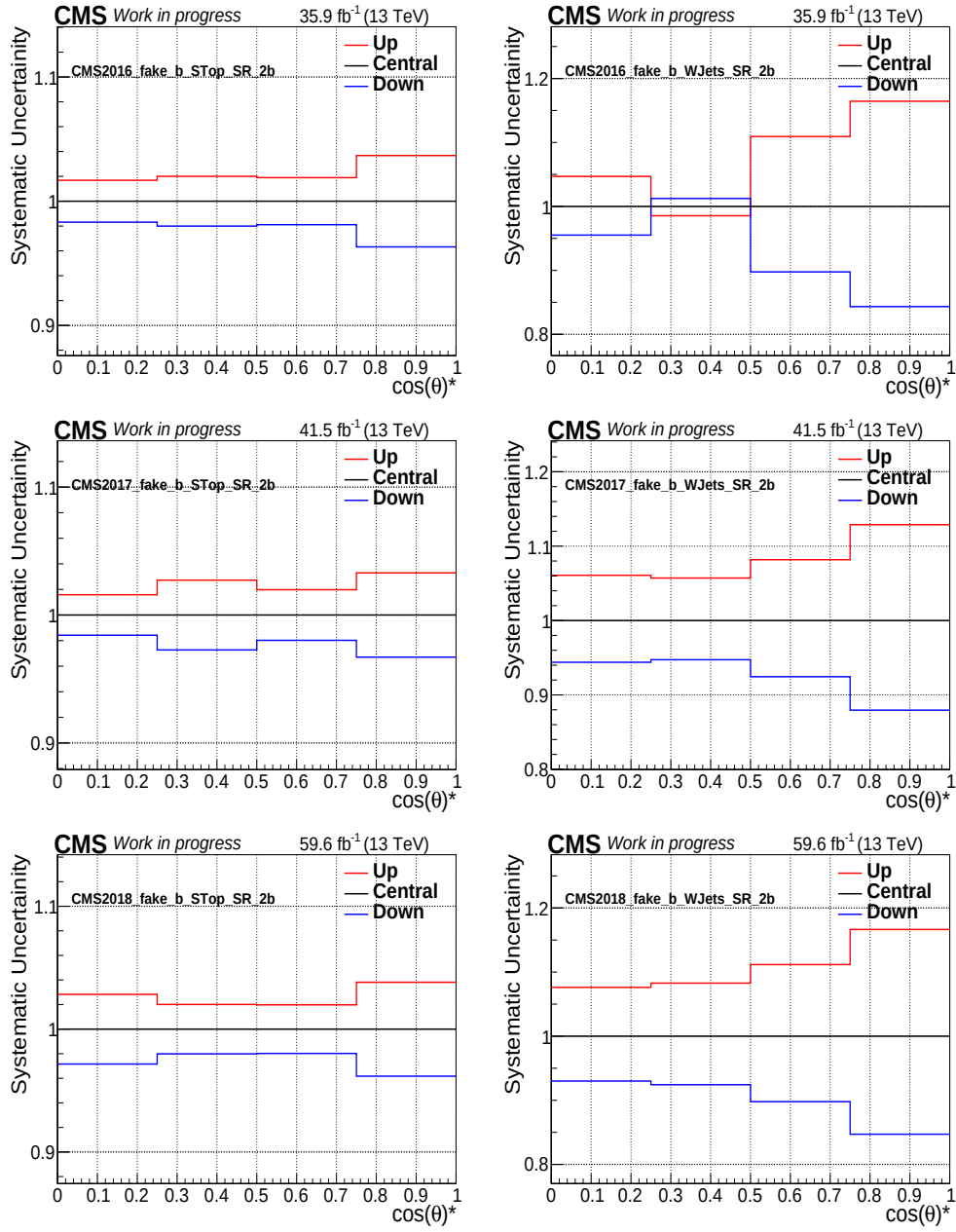


Figure 66: fake b-tag uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

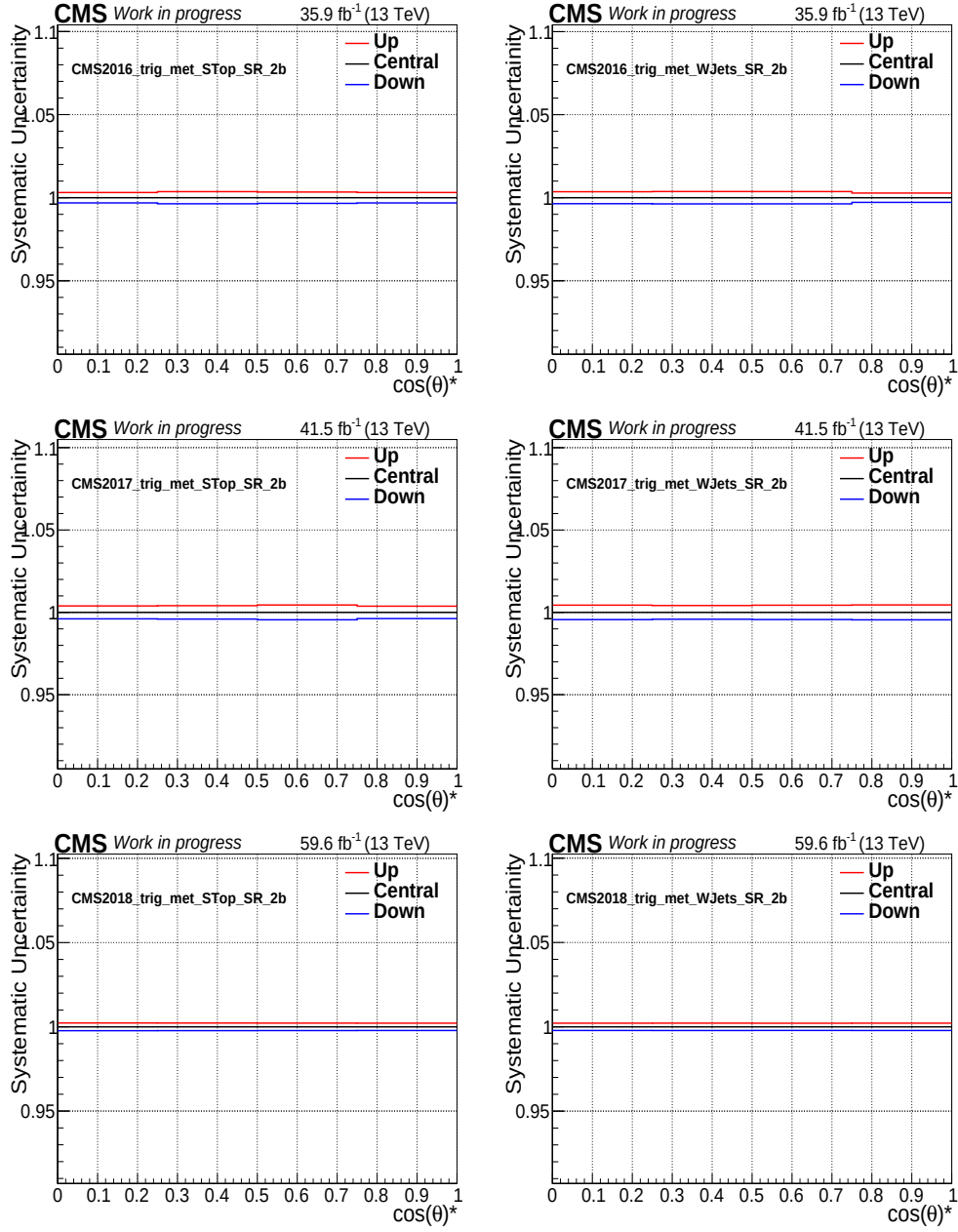


Figure 67: E_T^{miss} trigger uncertainty for single t and Wjets background 2b Category of Signal region for 2016, 2017 and 2018

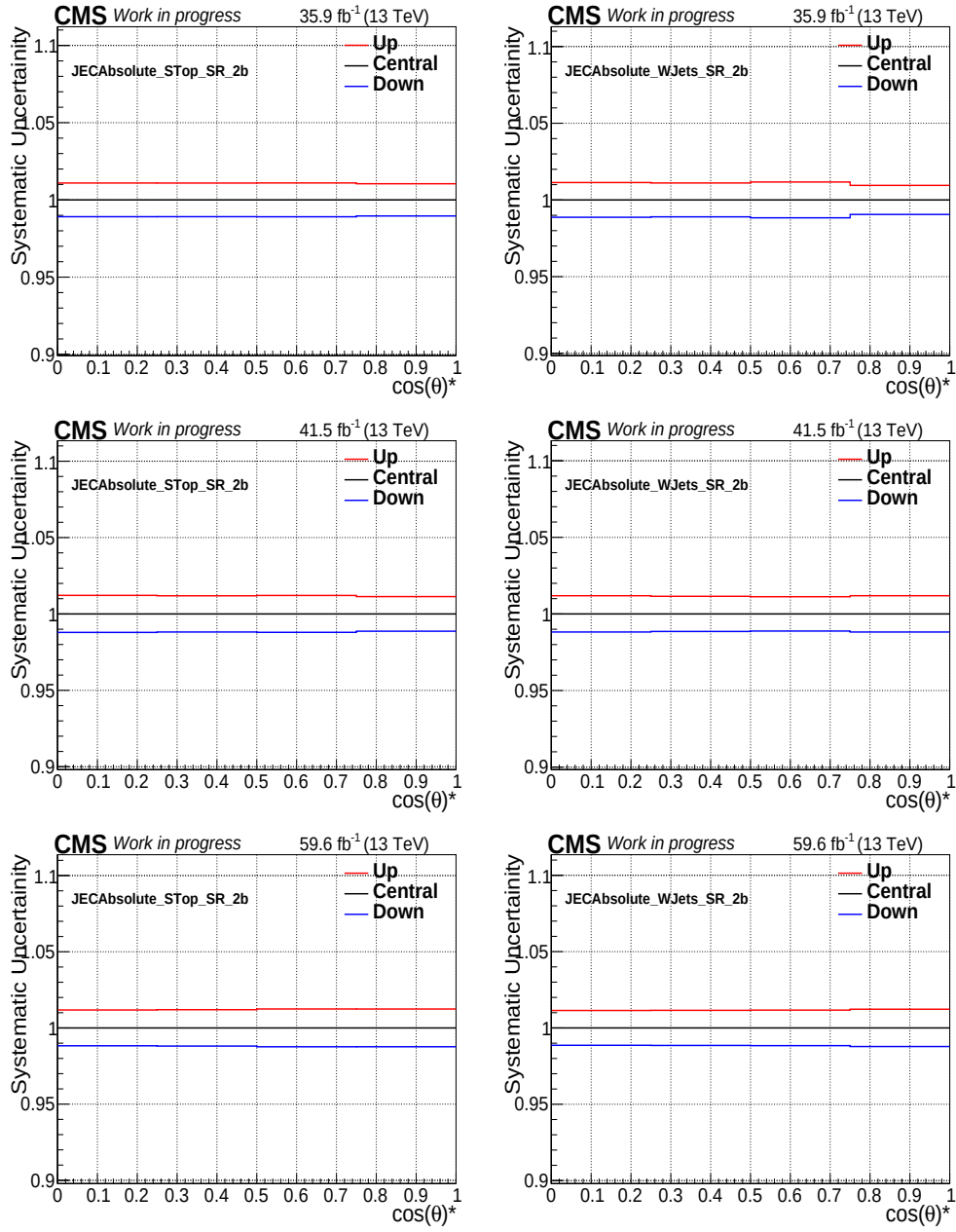


Figure 68: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

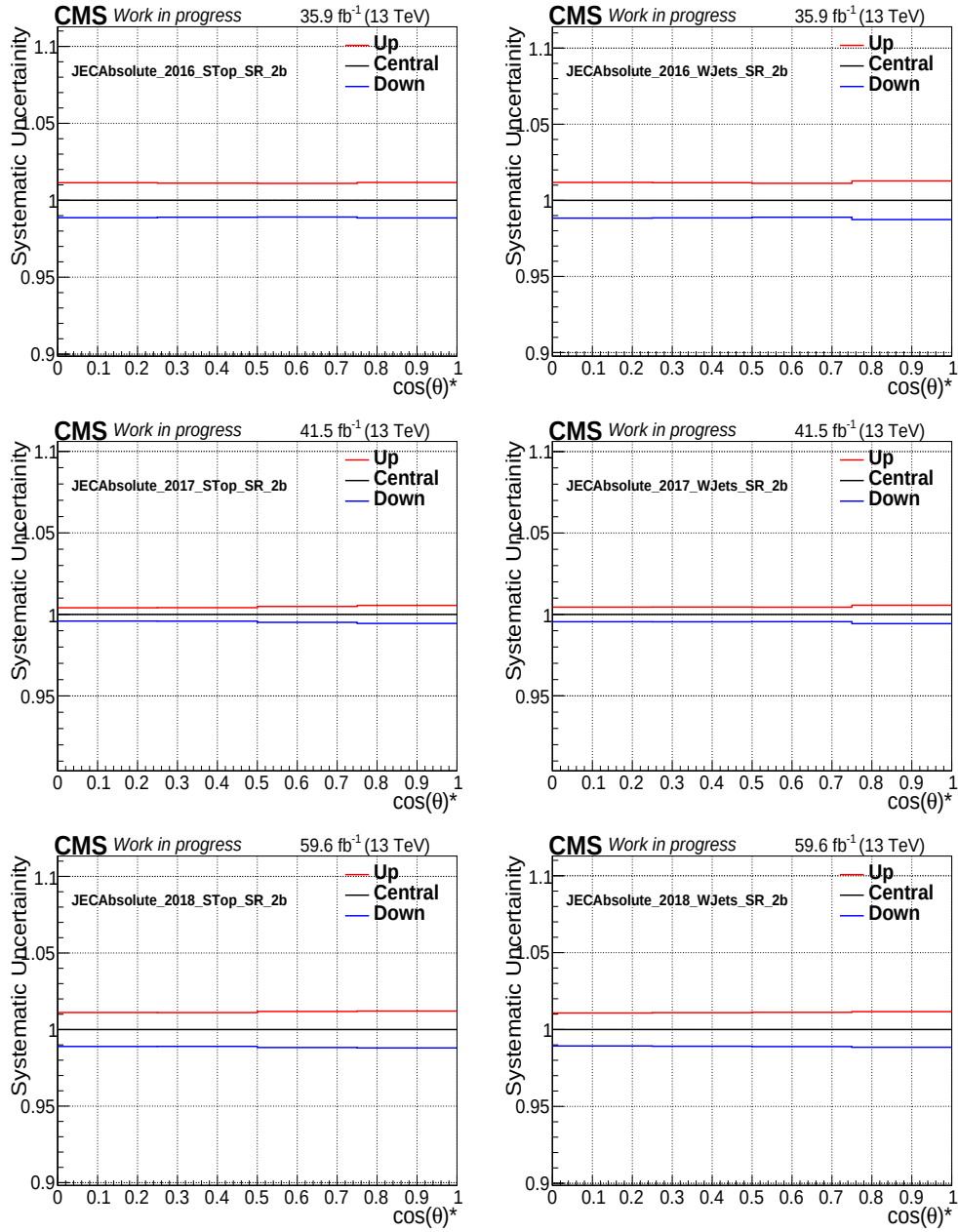


Figure 69: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

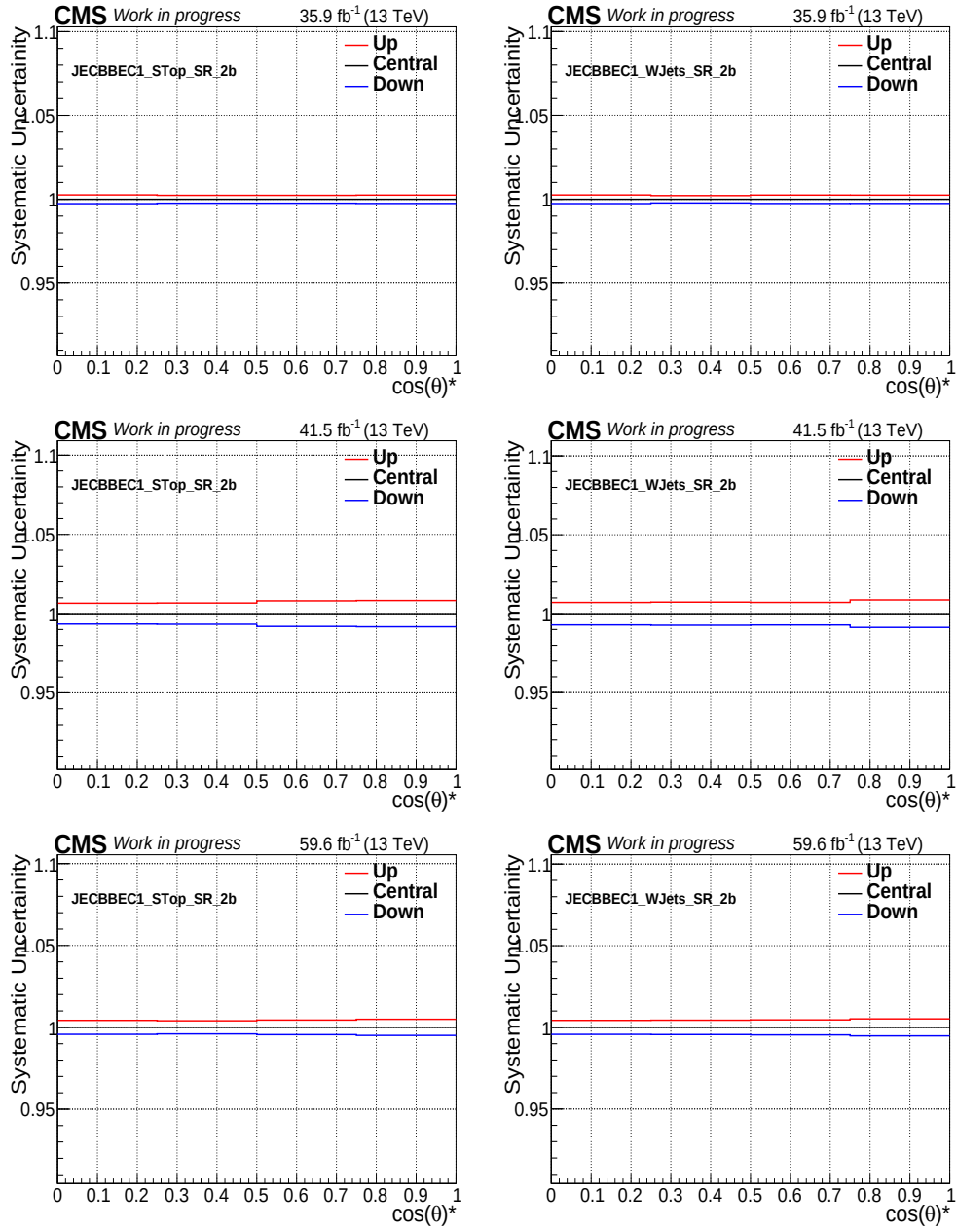


Figure 70: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

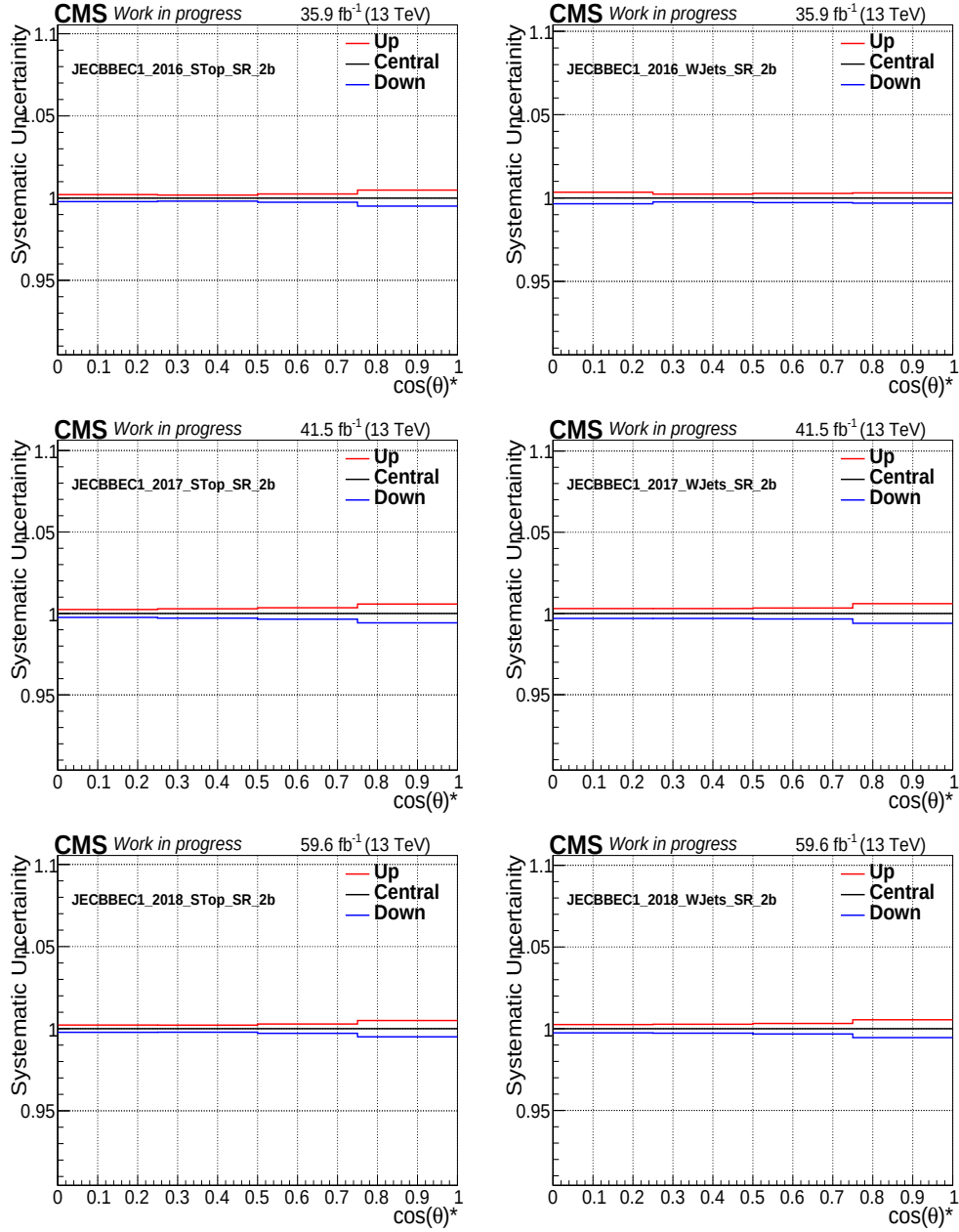


Figure 71: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

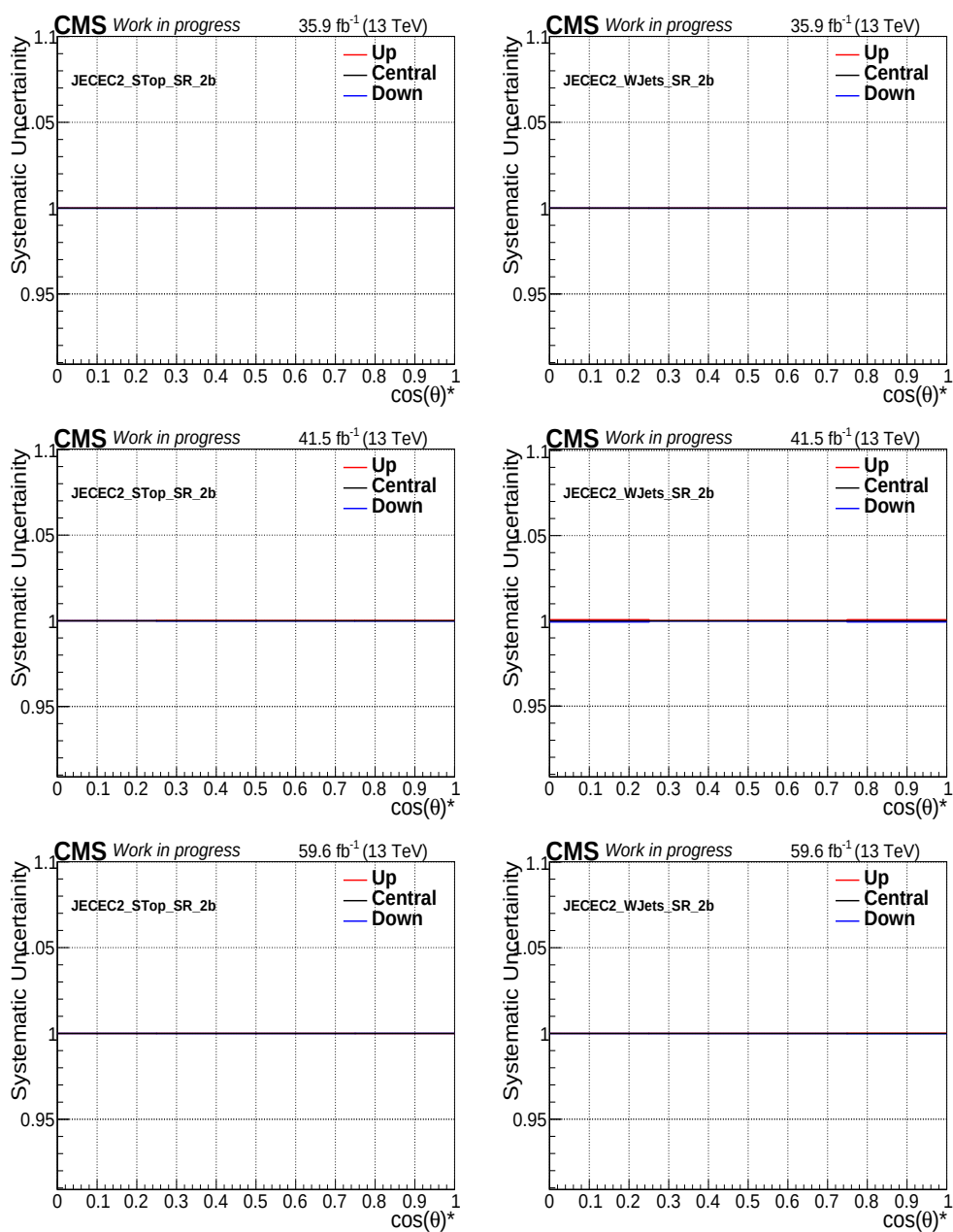


Figure 72: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

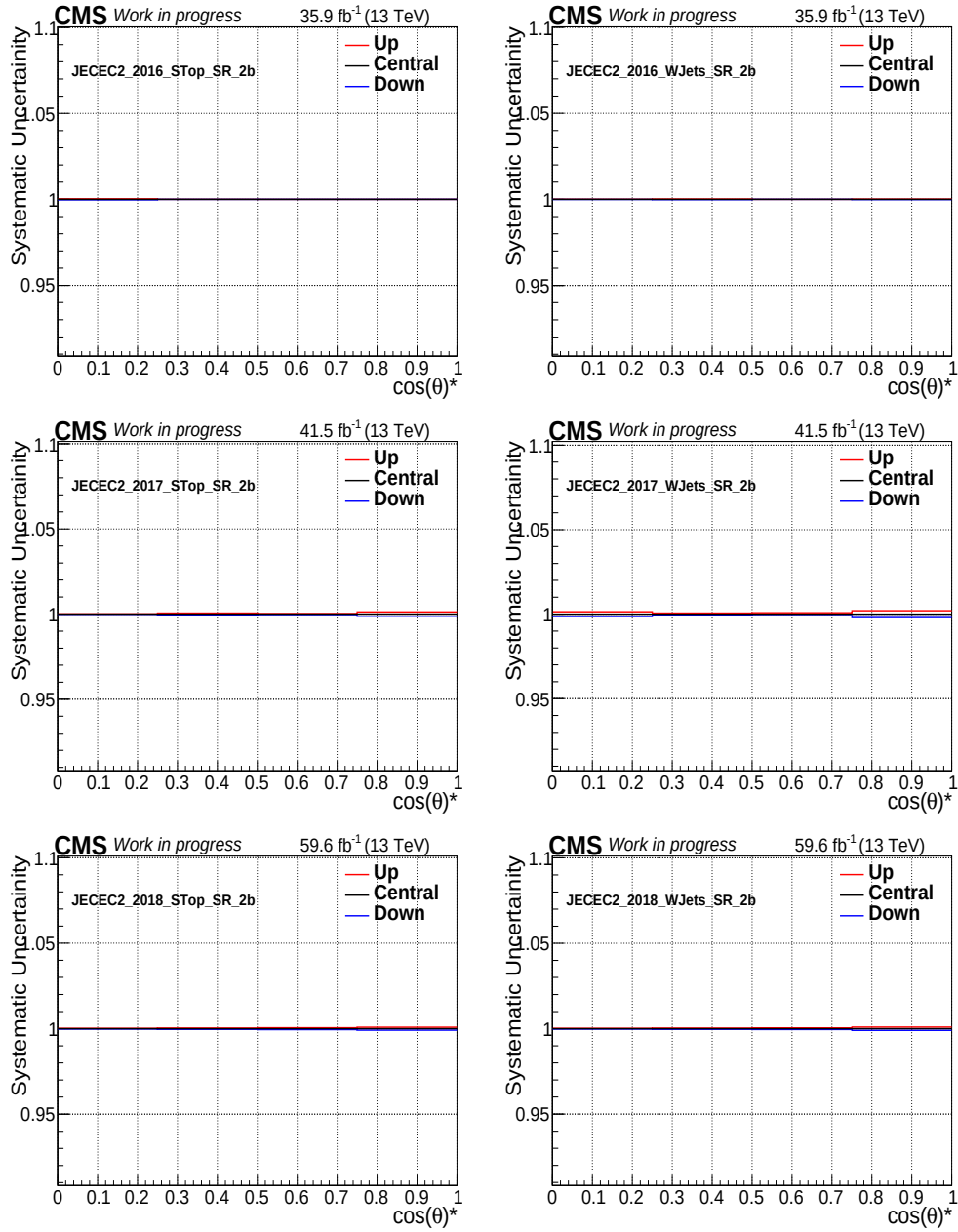


Figure 73: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

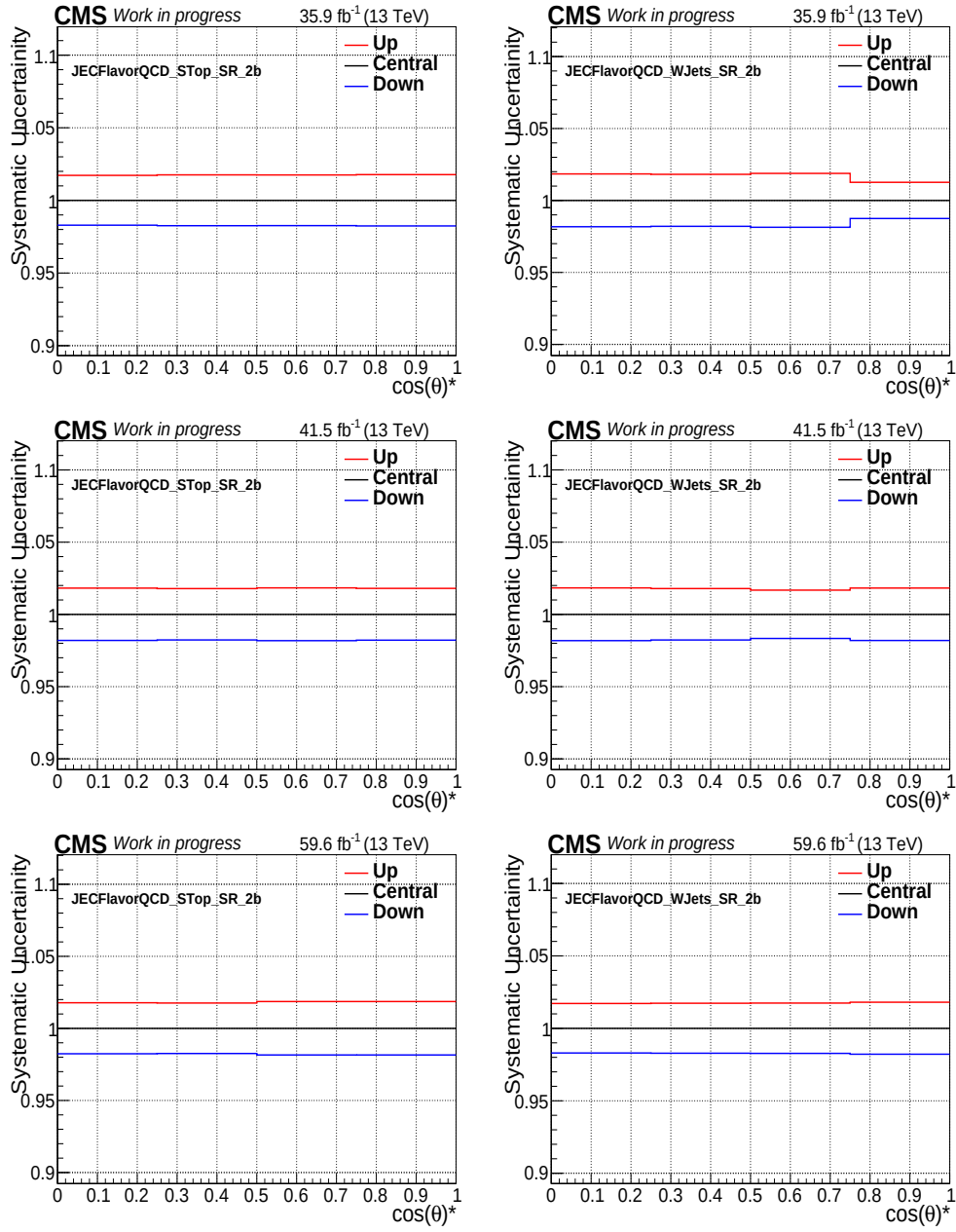


Figure 74: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

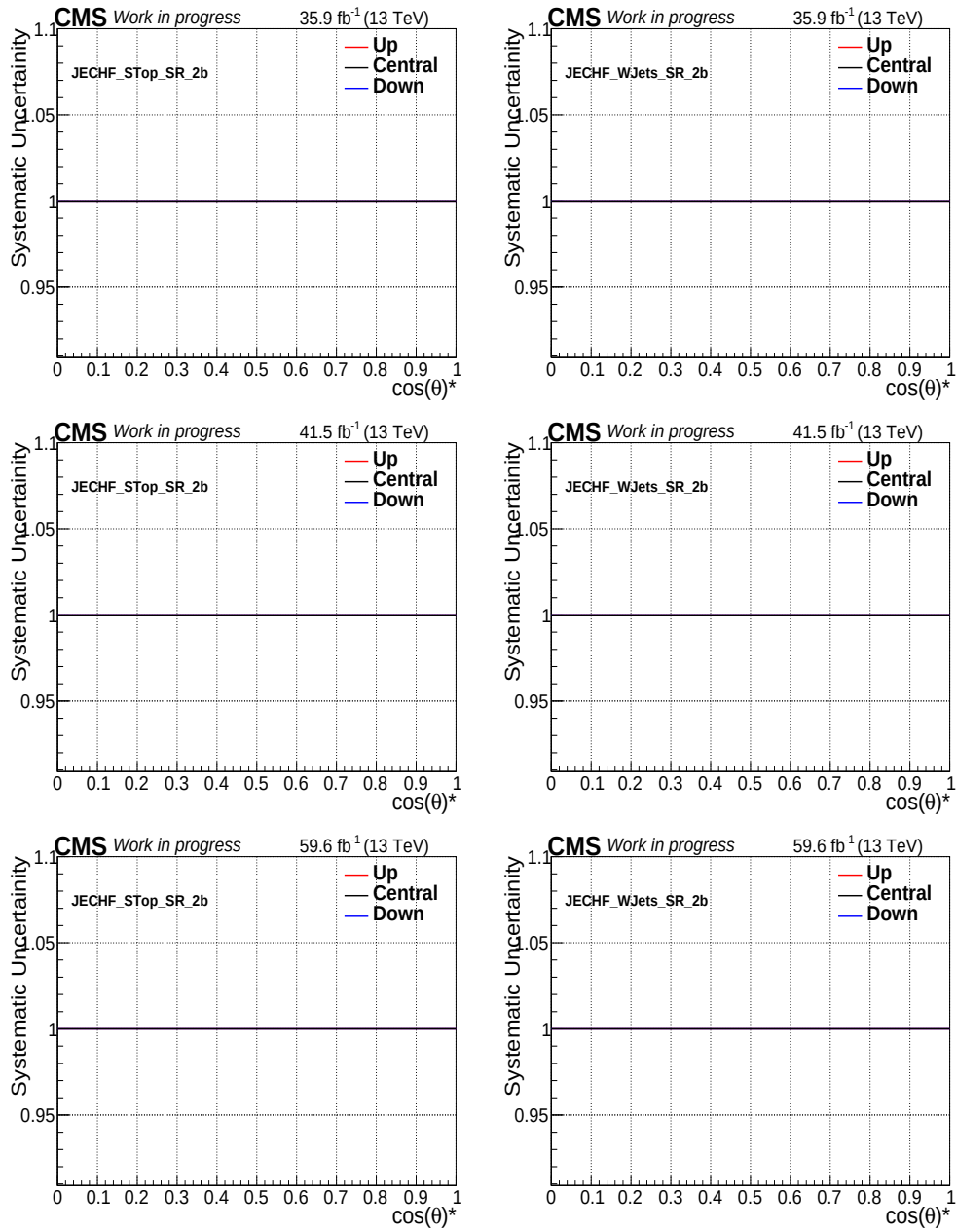


Figure 75: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

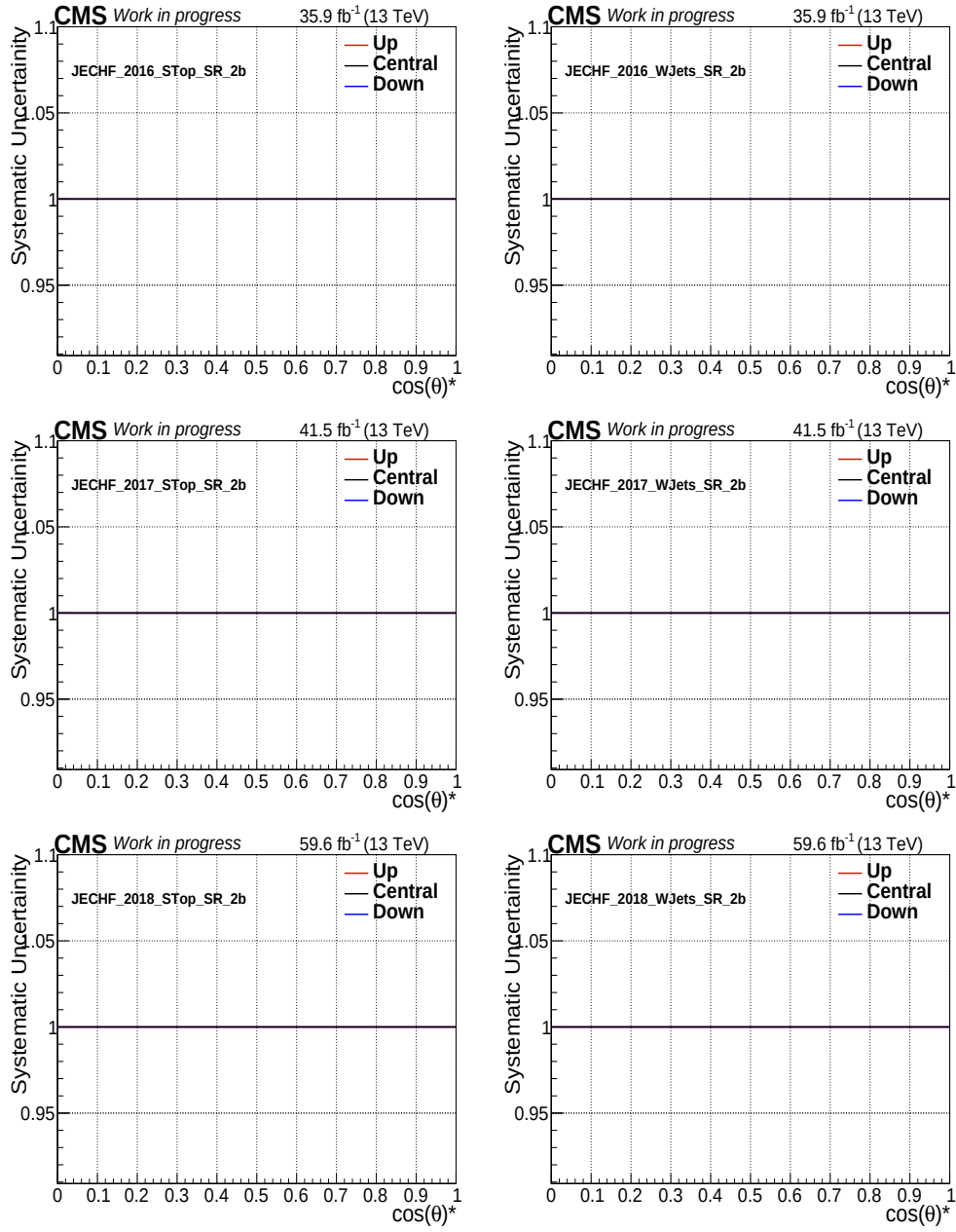


Figure 76: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

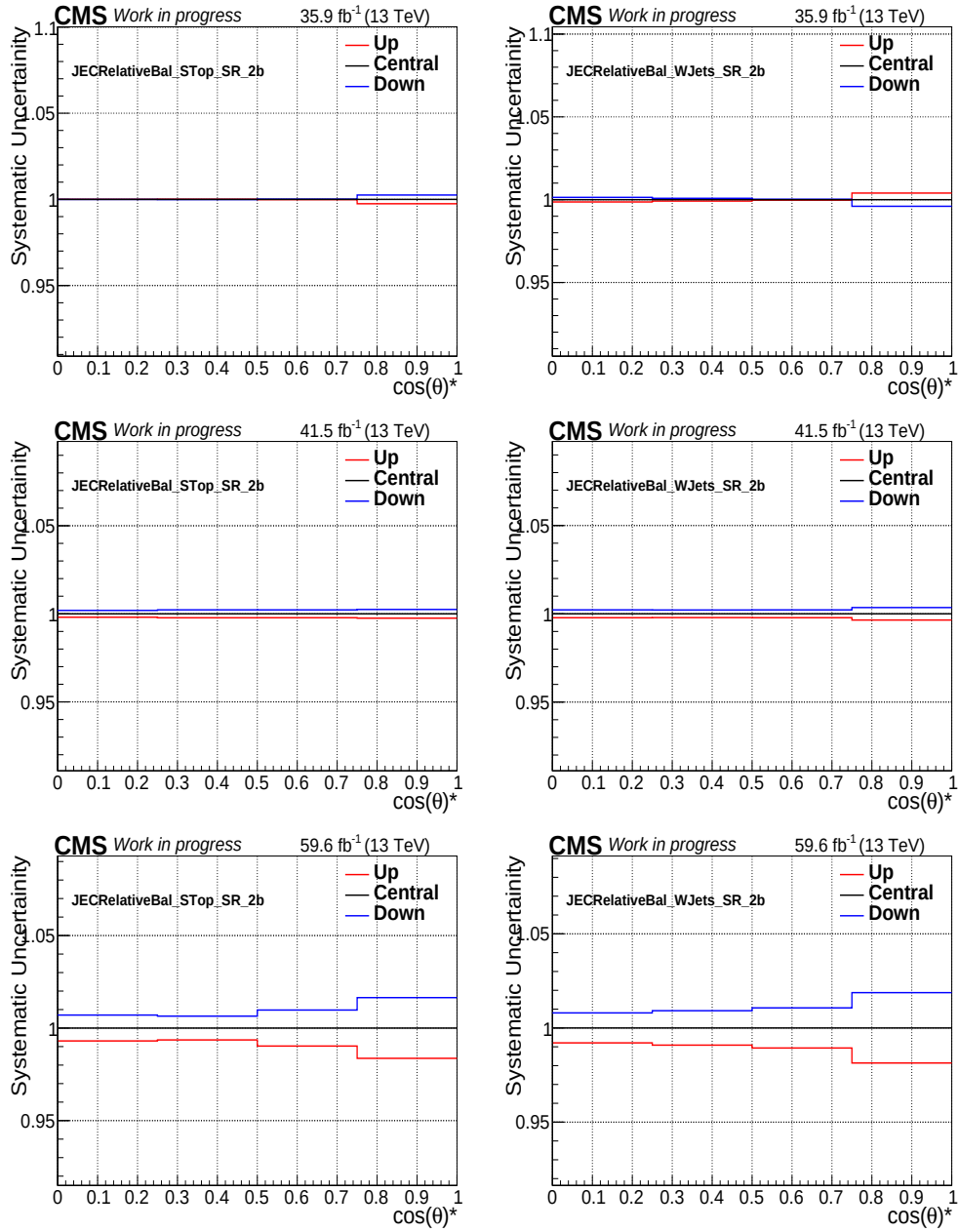


Figure 77: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

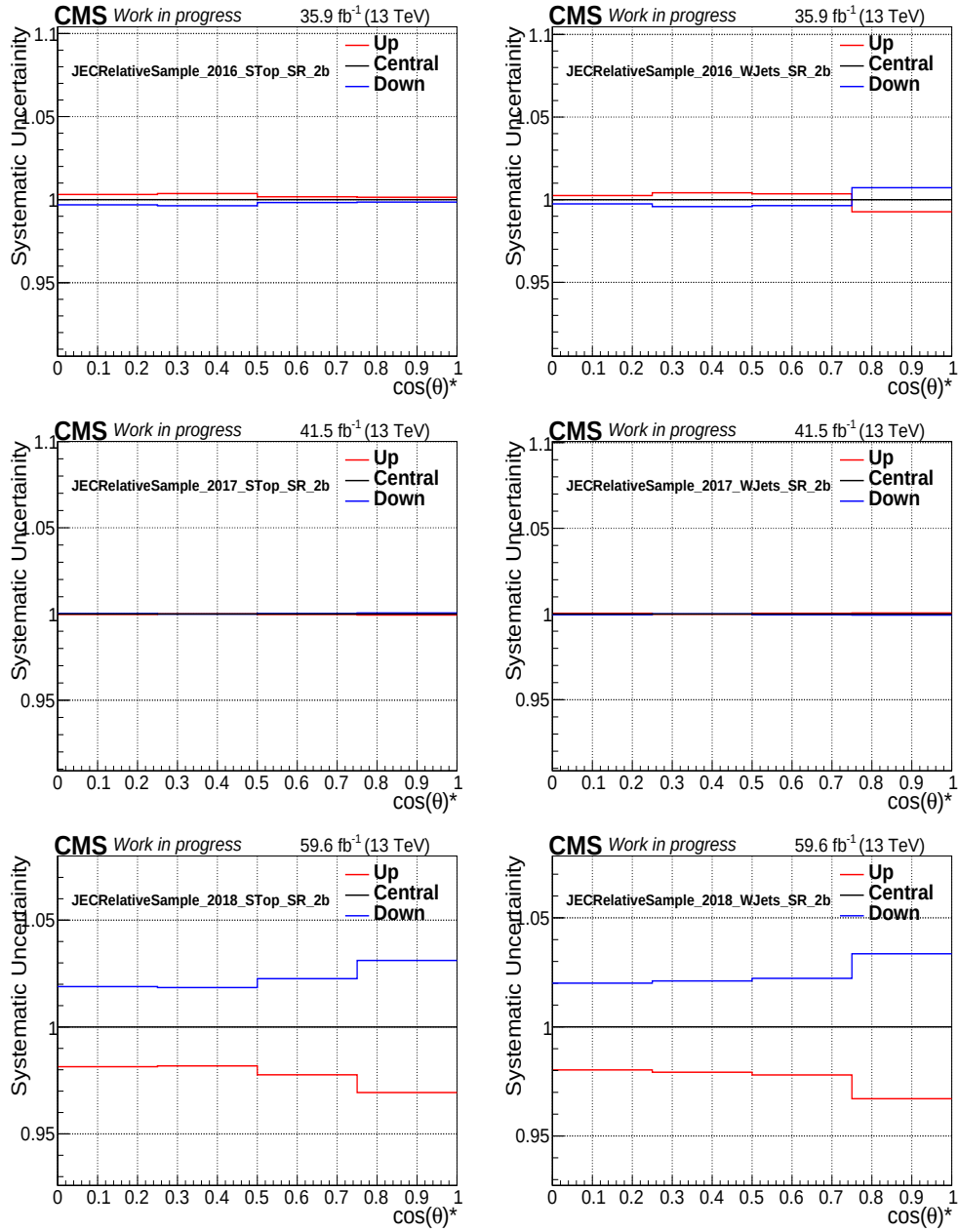


Figure 78: JEC uncertainty for single t and WJets background 2b Category of Signal region for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

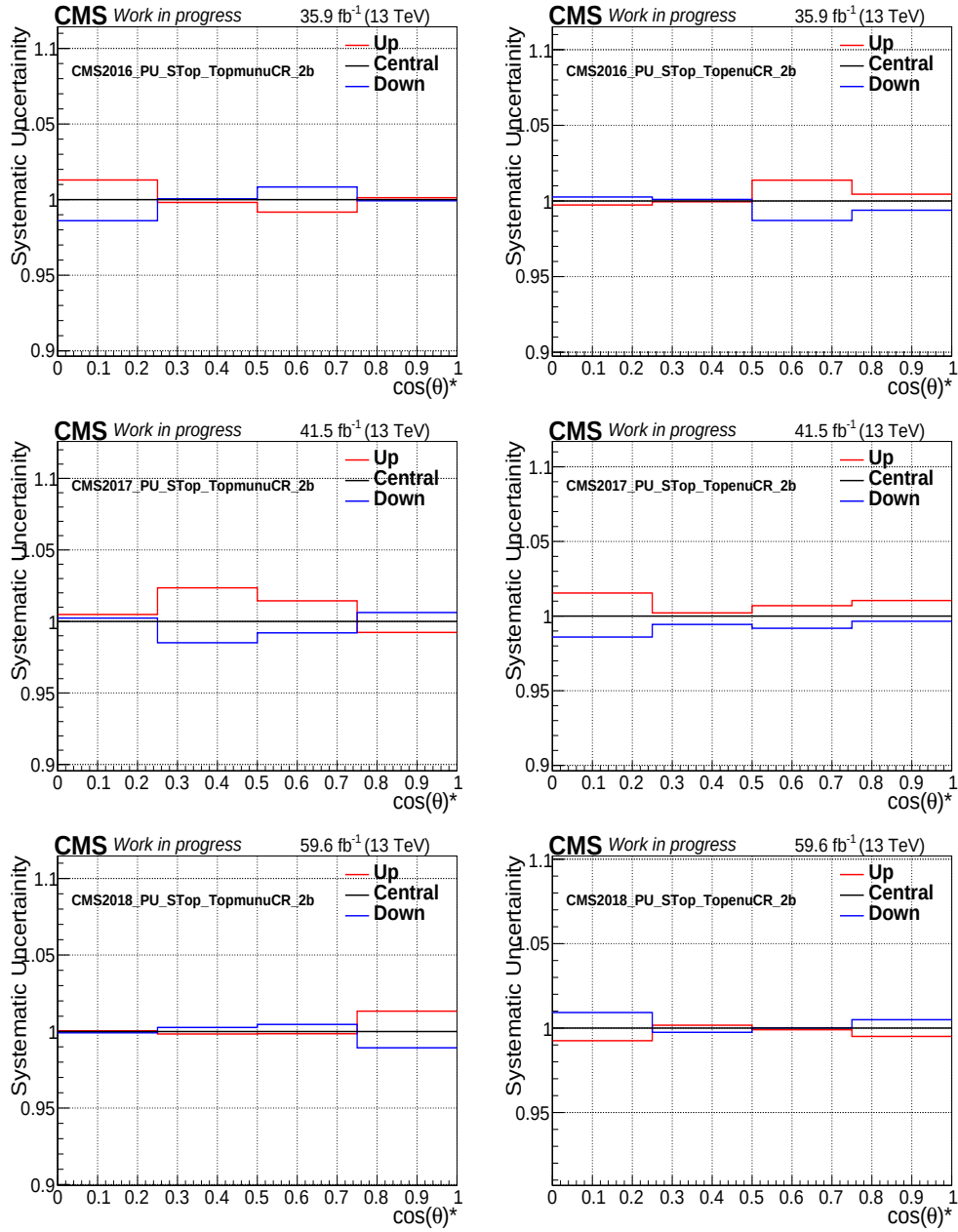


Figure 79: PileUP uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

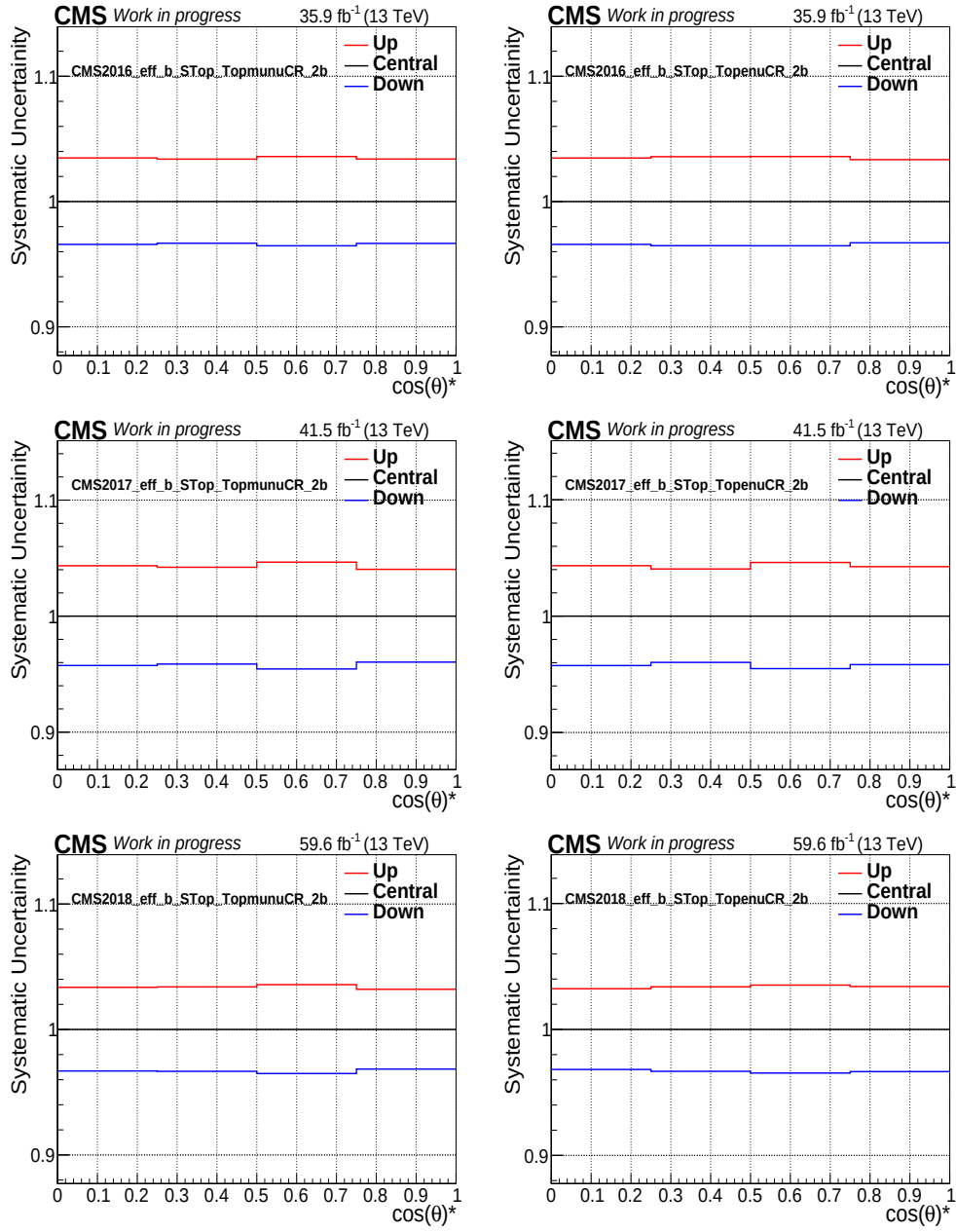


Figure 80: b-tag uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

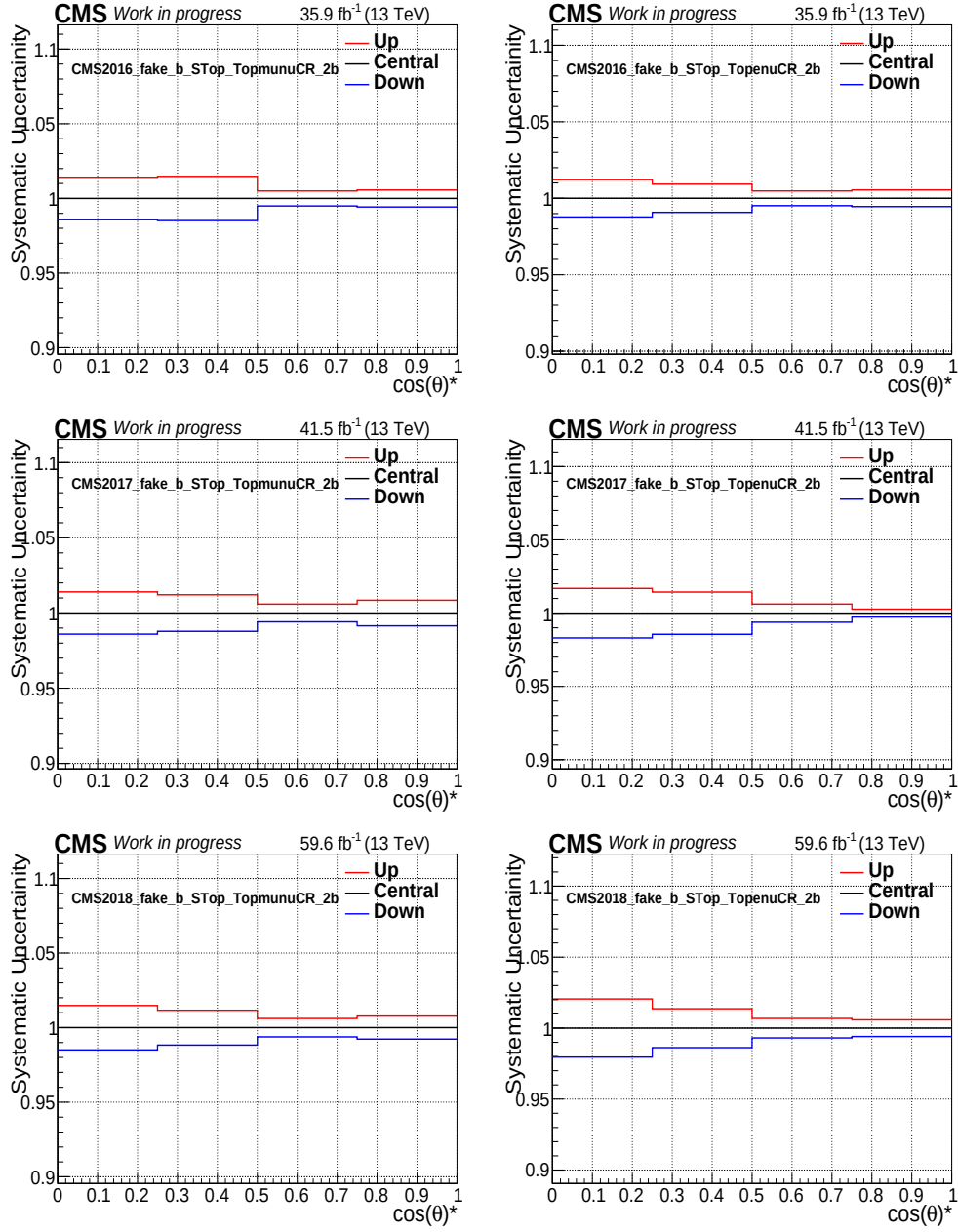


Figure 81: Fake b-tag uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

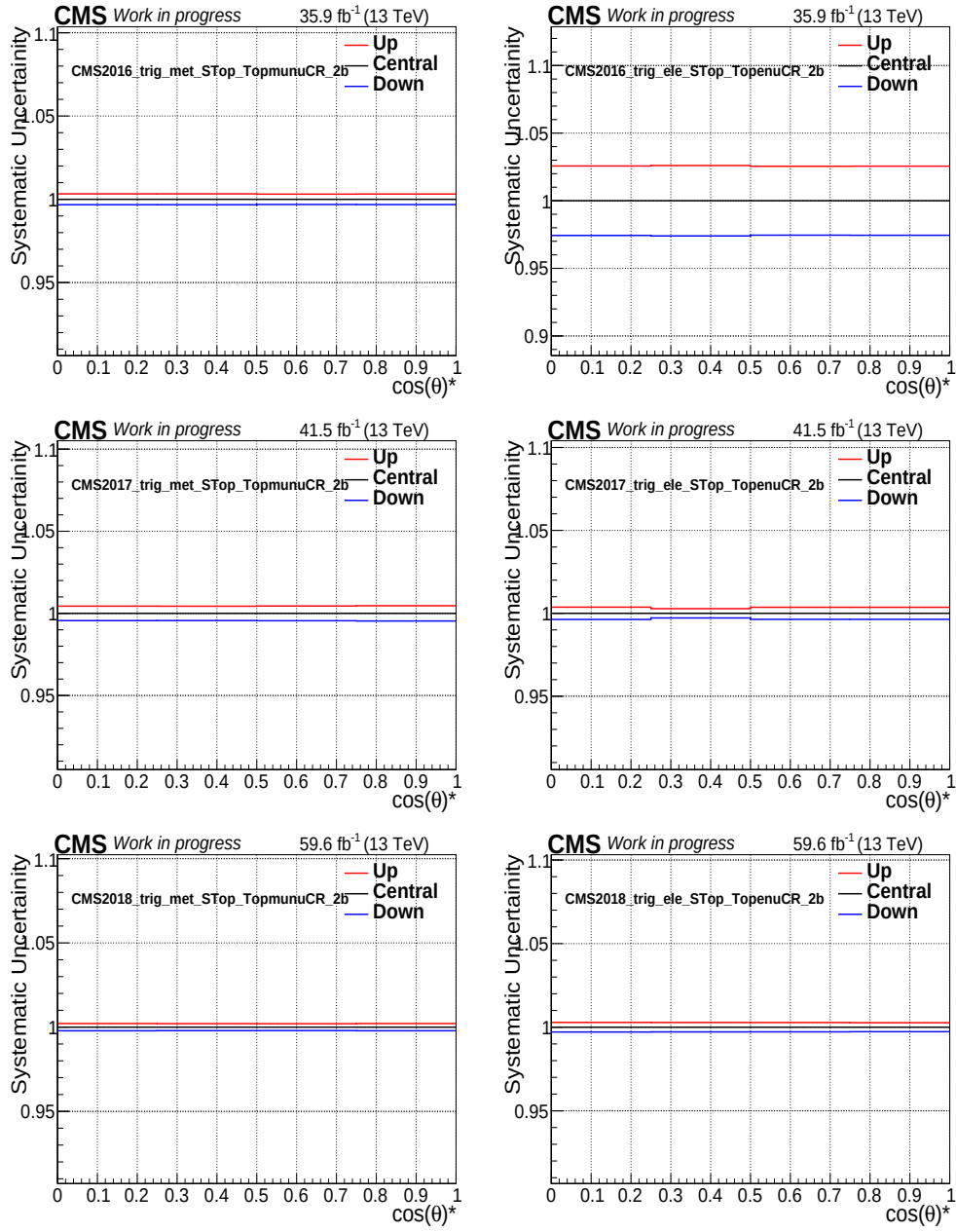


Figure 82: Trigger uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

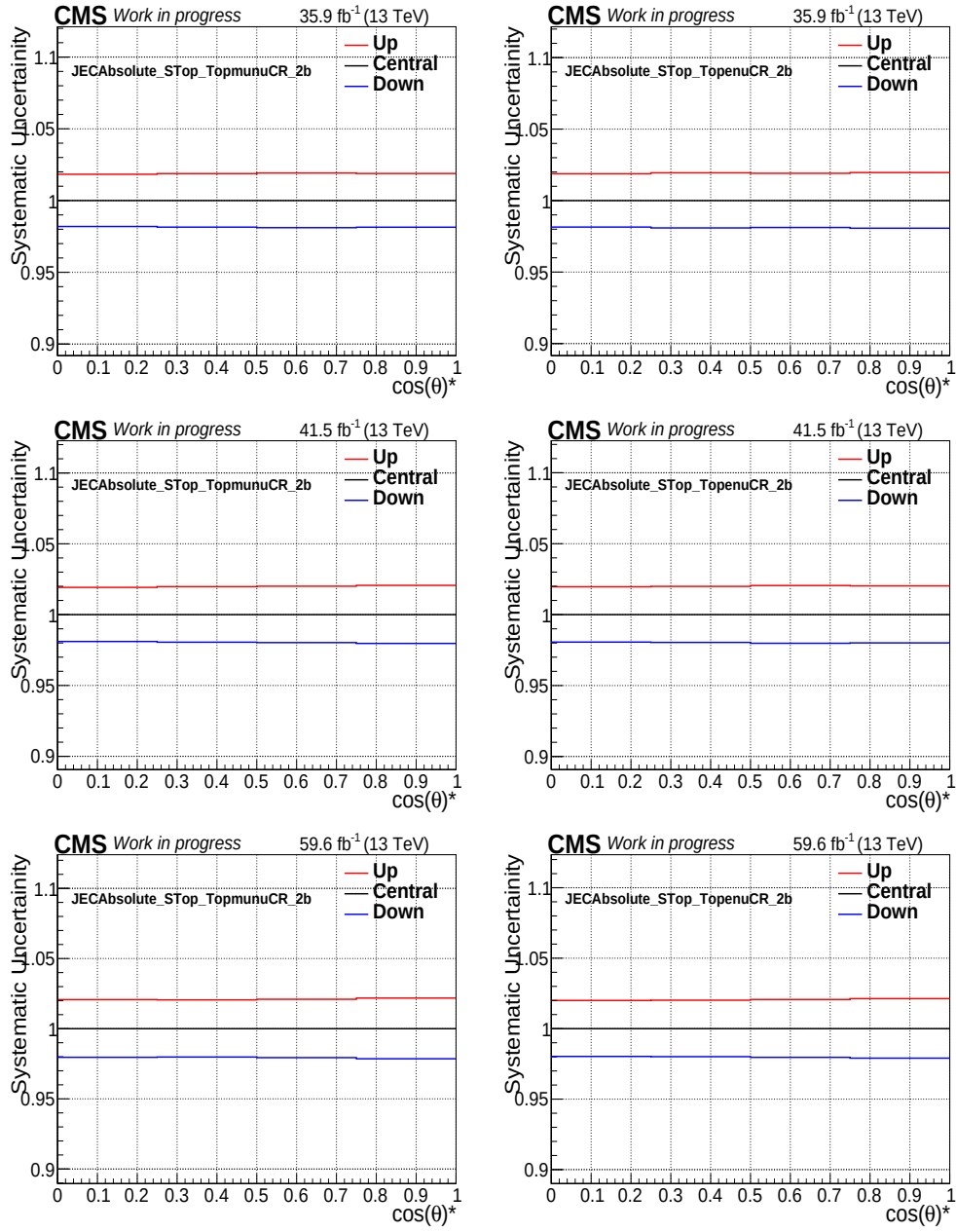


Figure 83: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

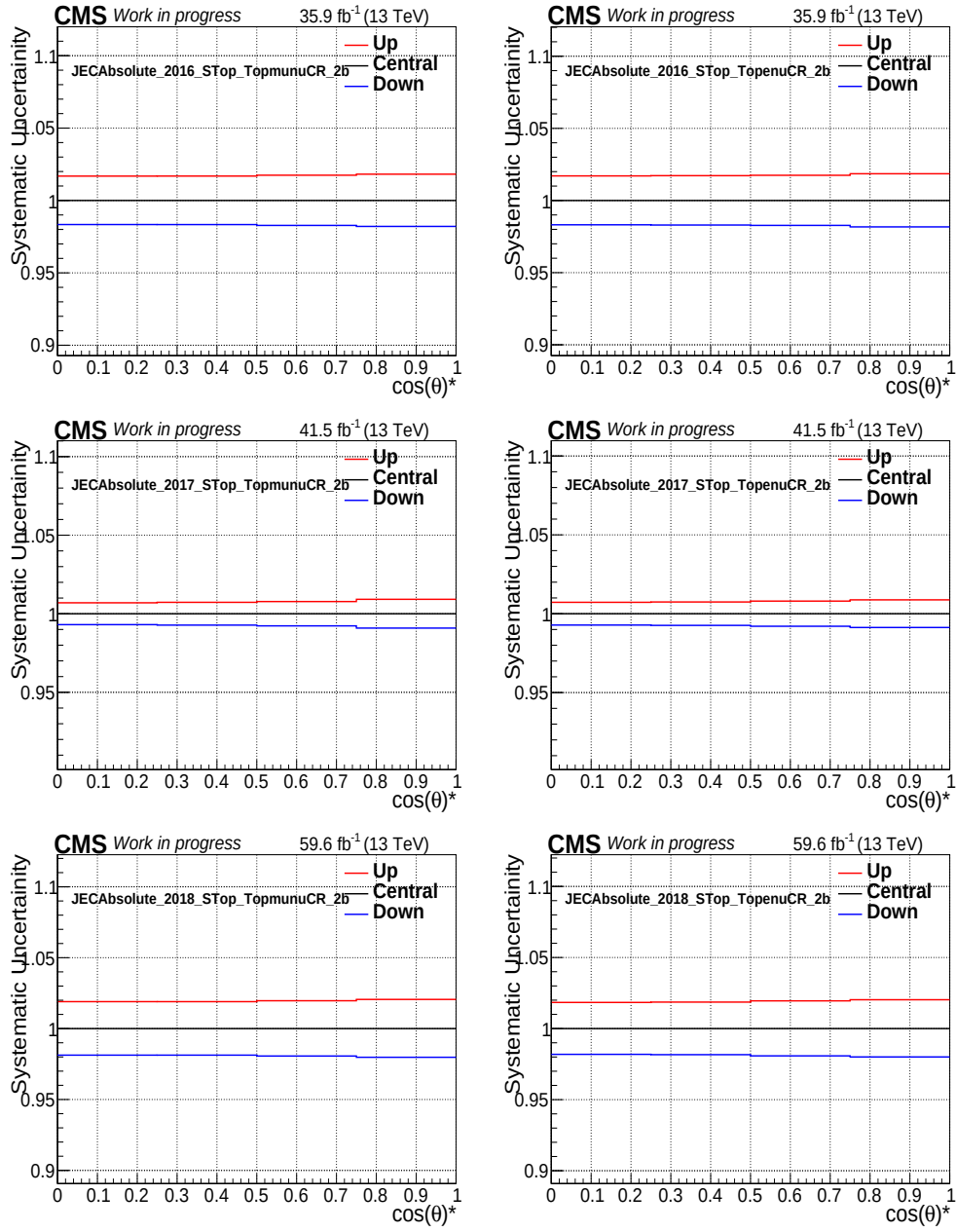


Figure 84: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

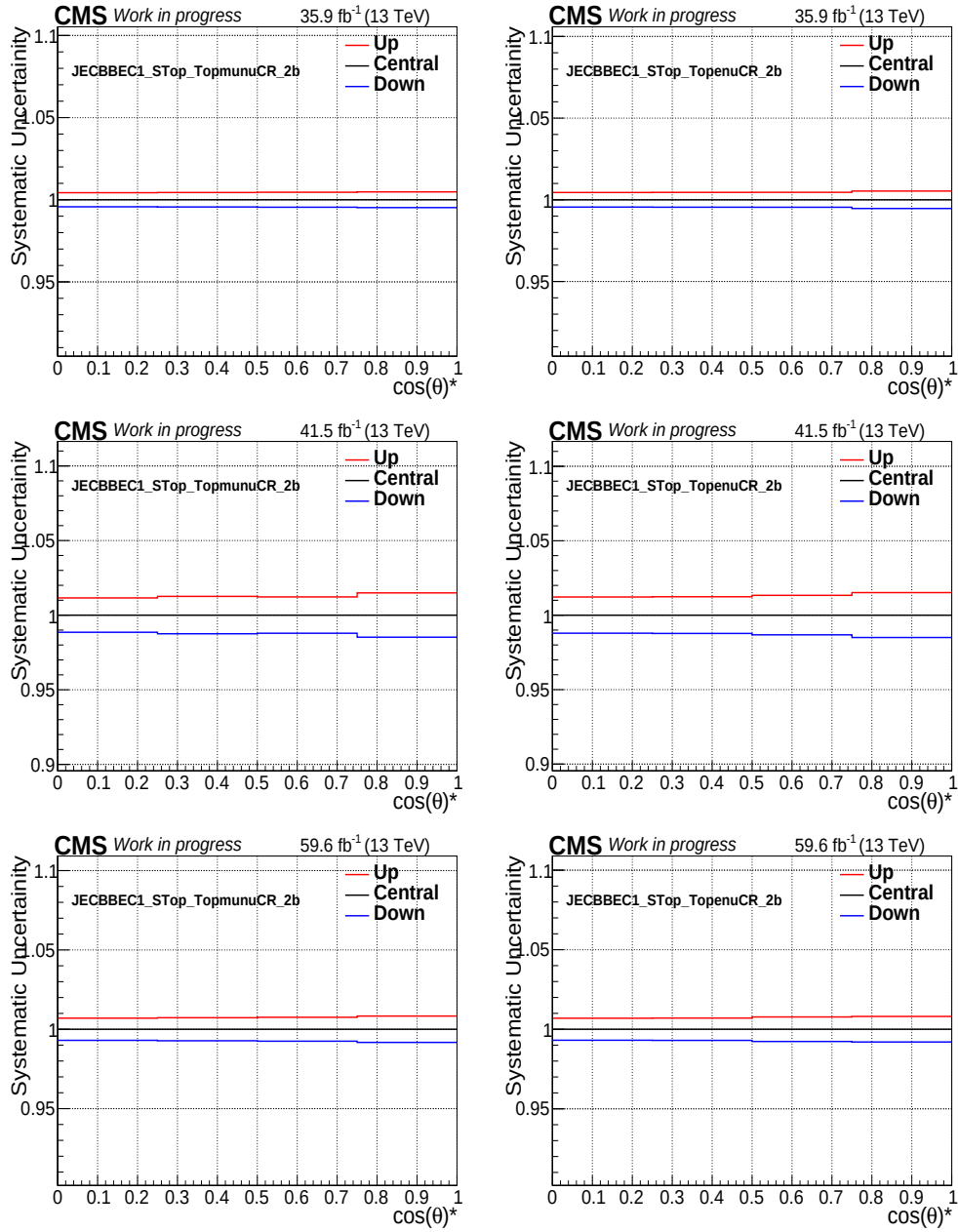


Figure 85: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

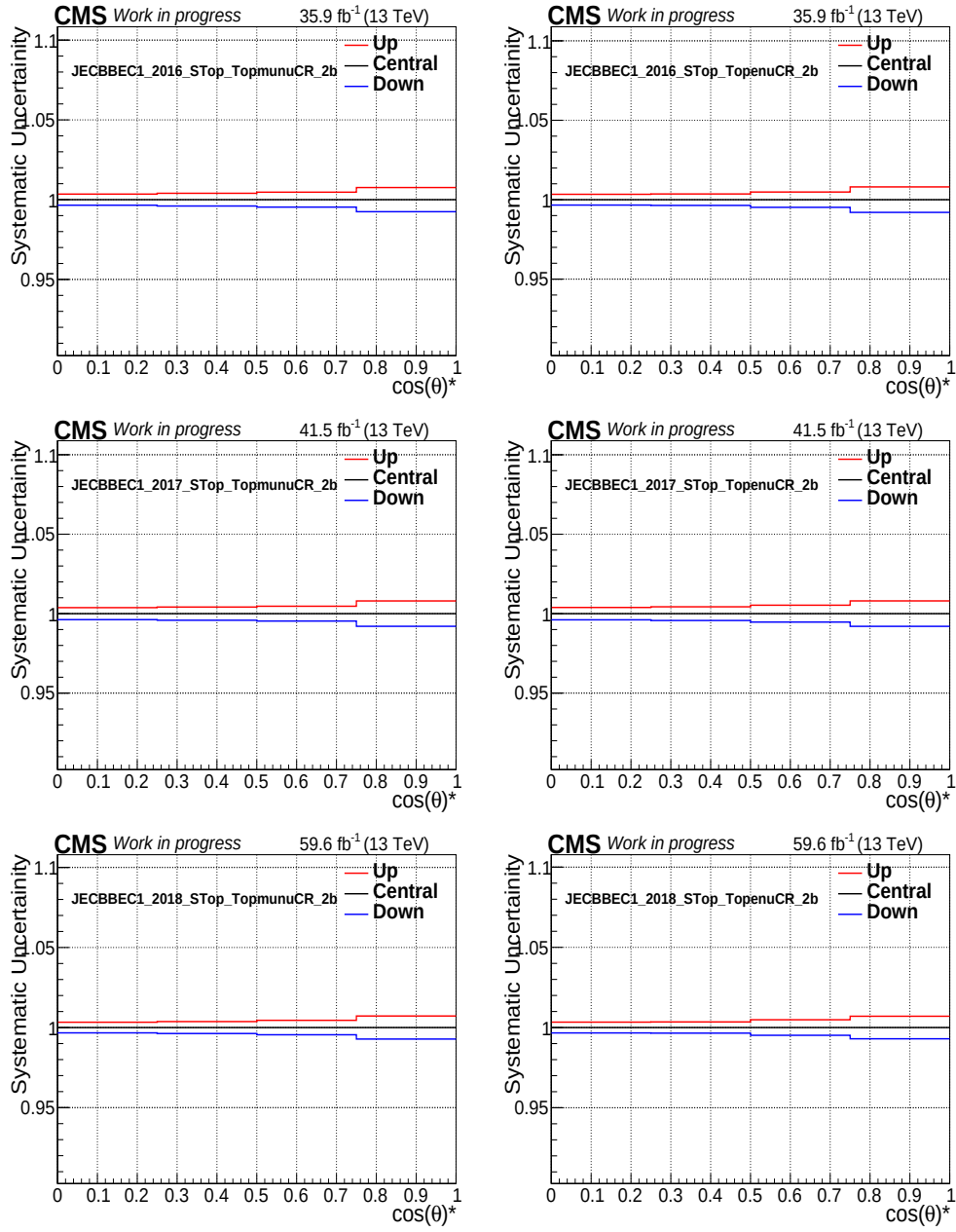


Figure 86: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

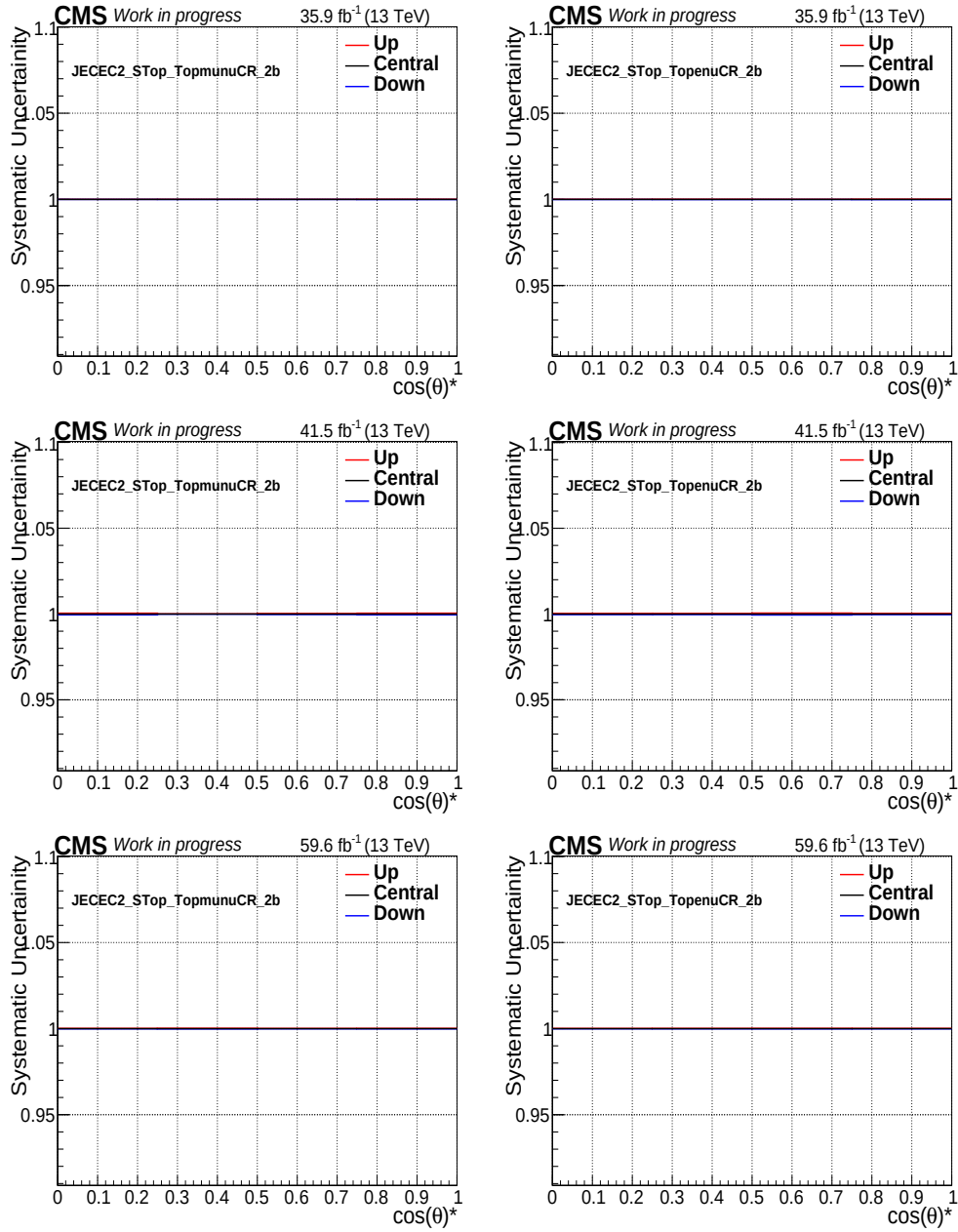


Figure 87: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

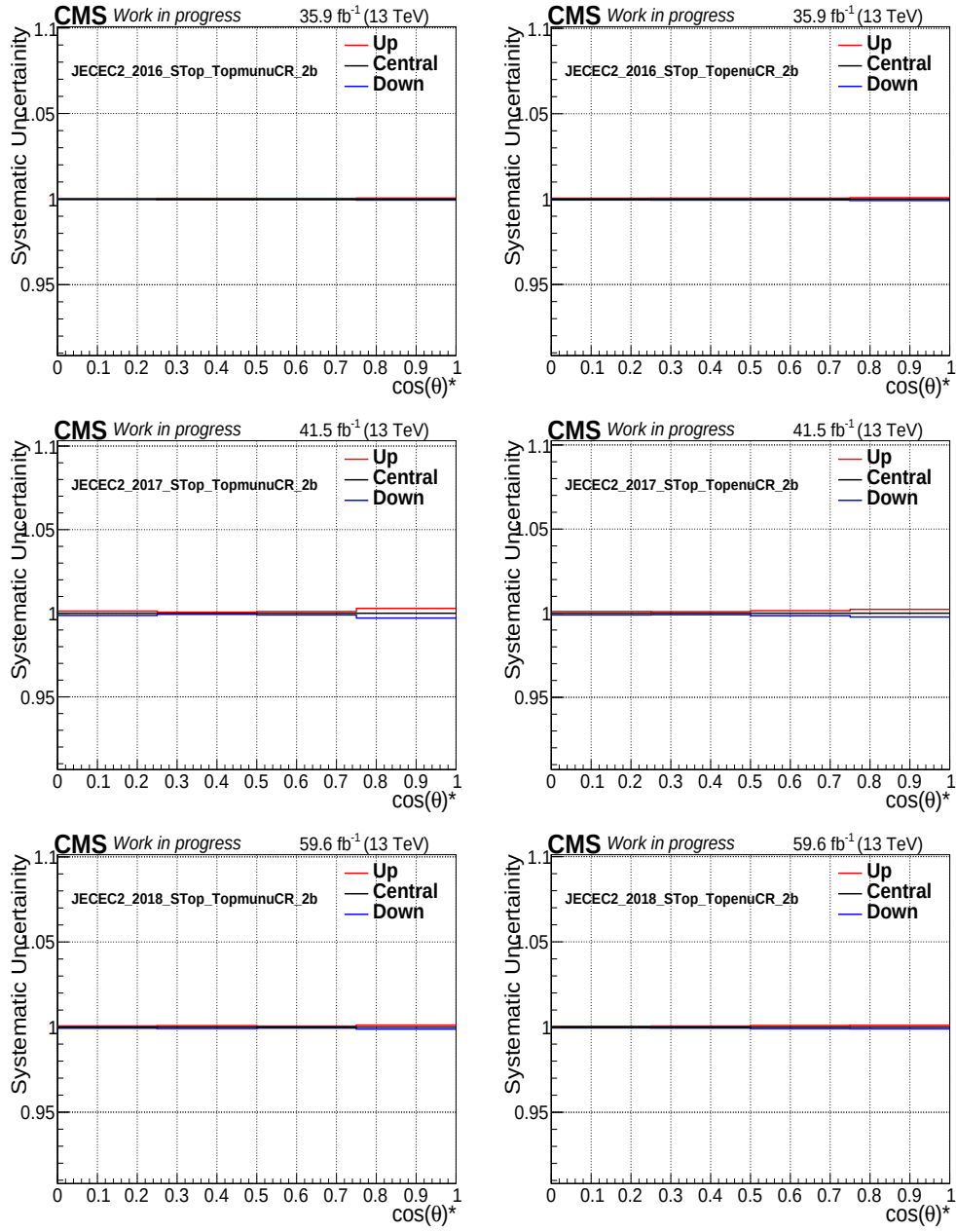


Figure 88: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

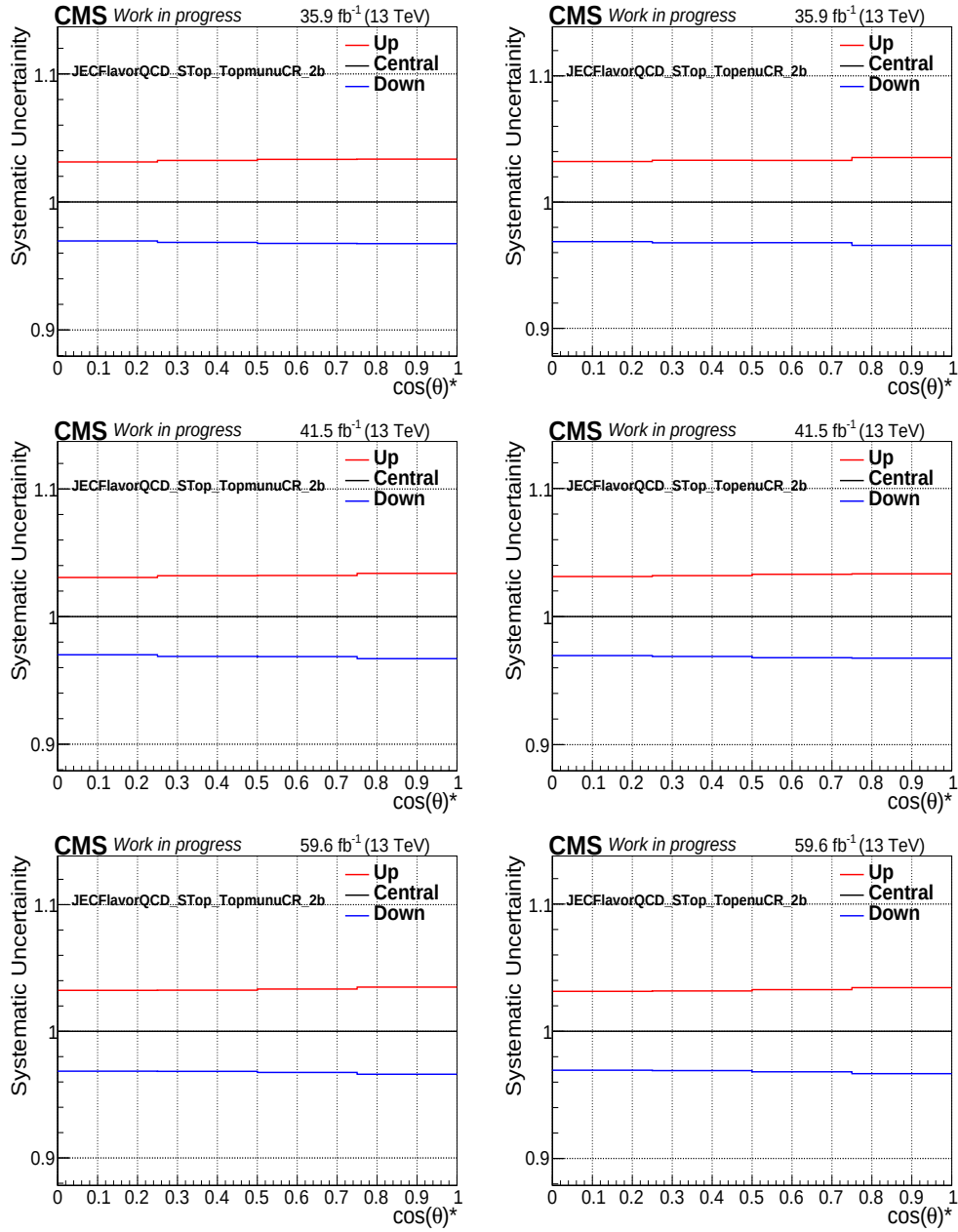


Figure 89: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

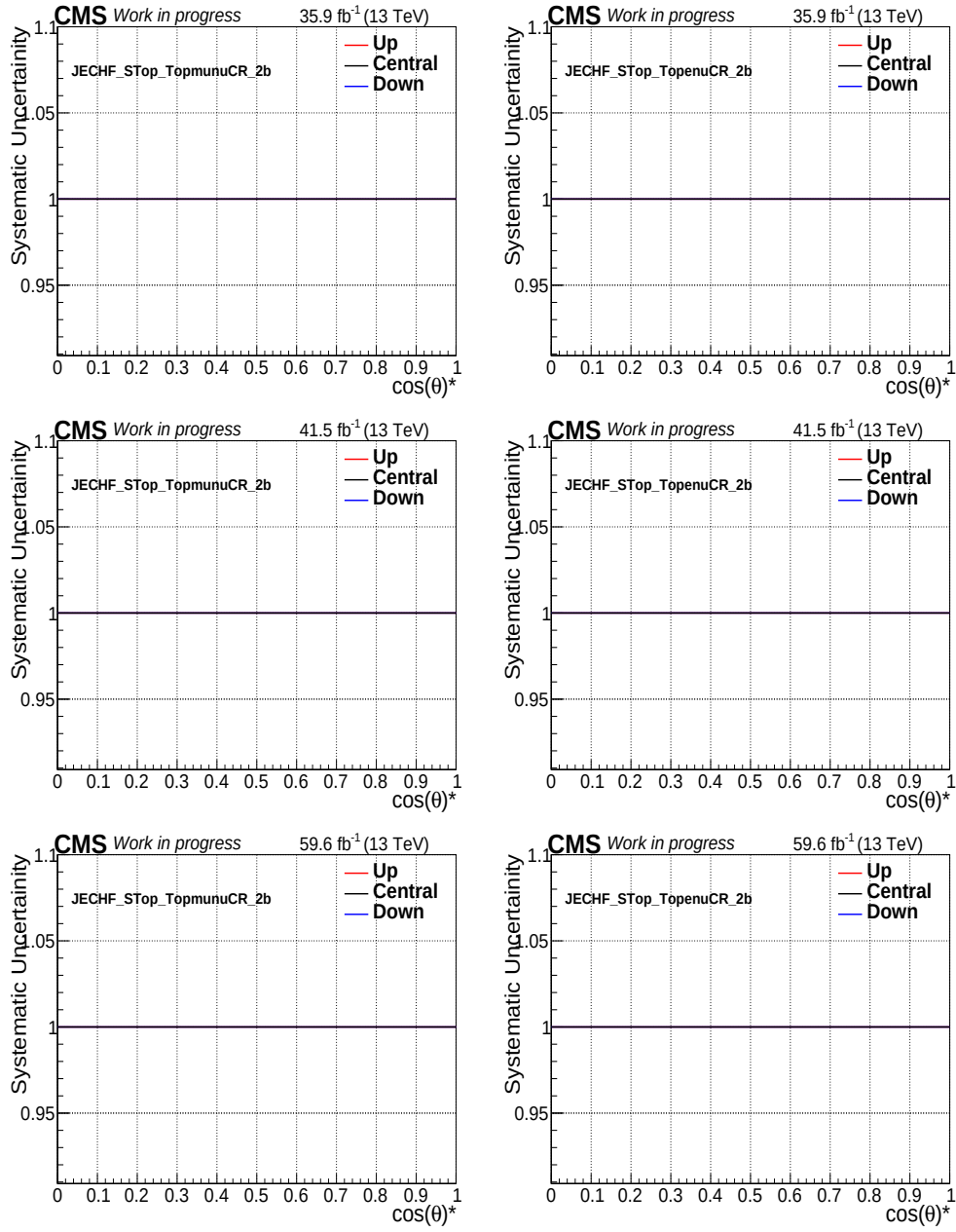


Figure 90: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

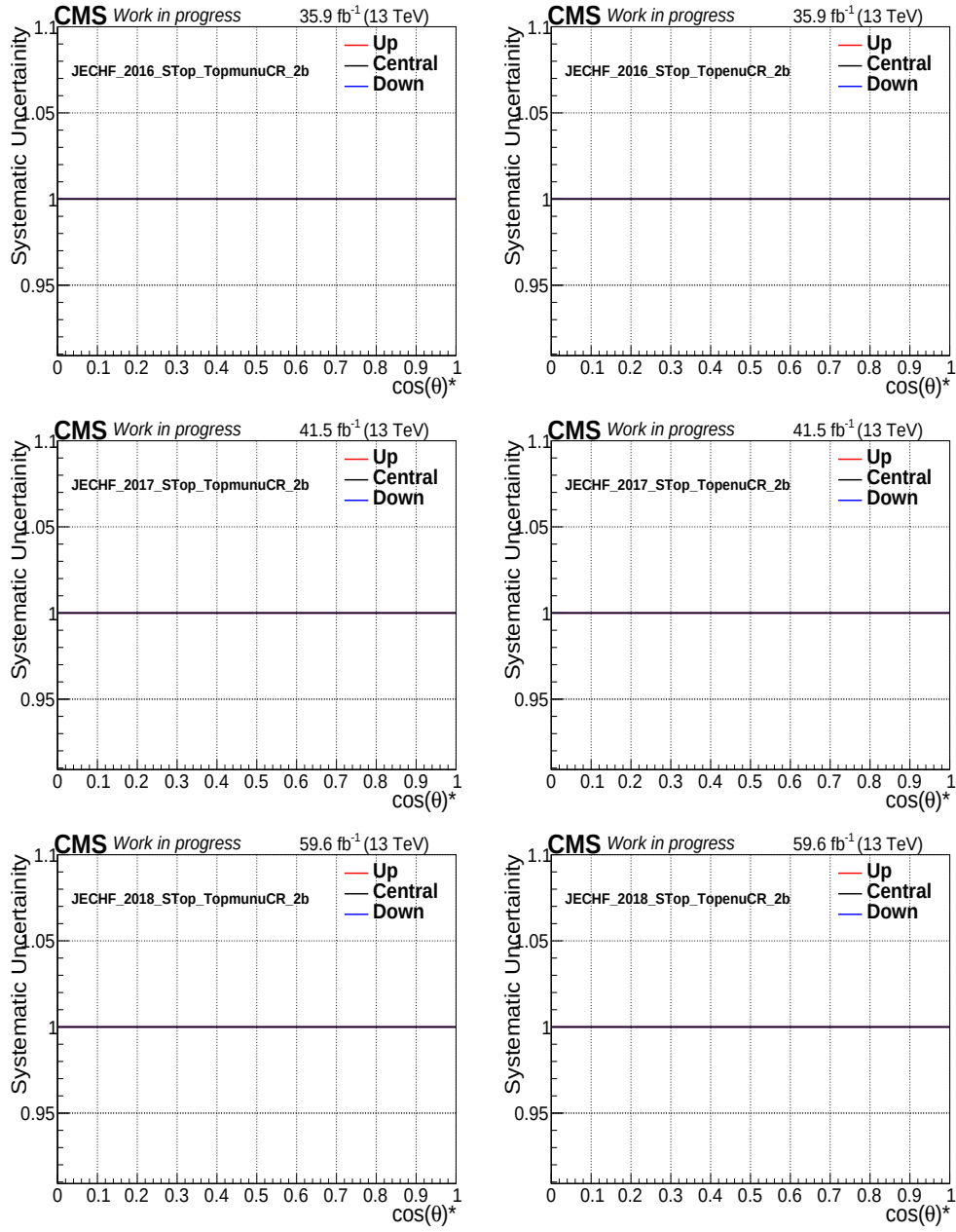


Figure 91: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

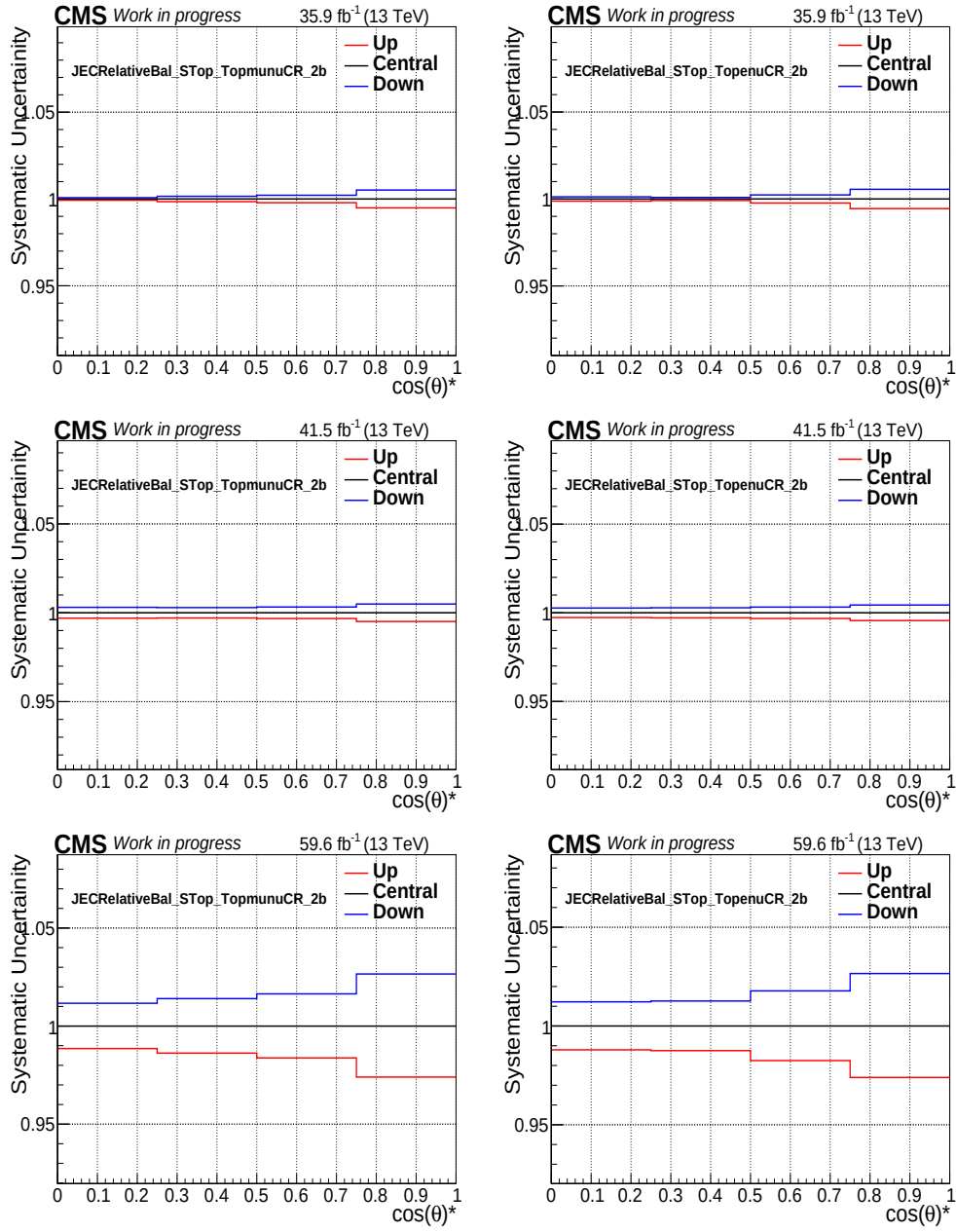


Figure 92: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

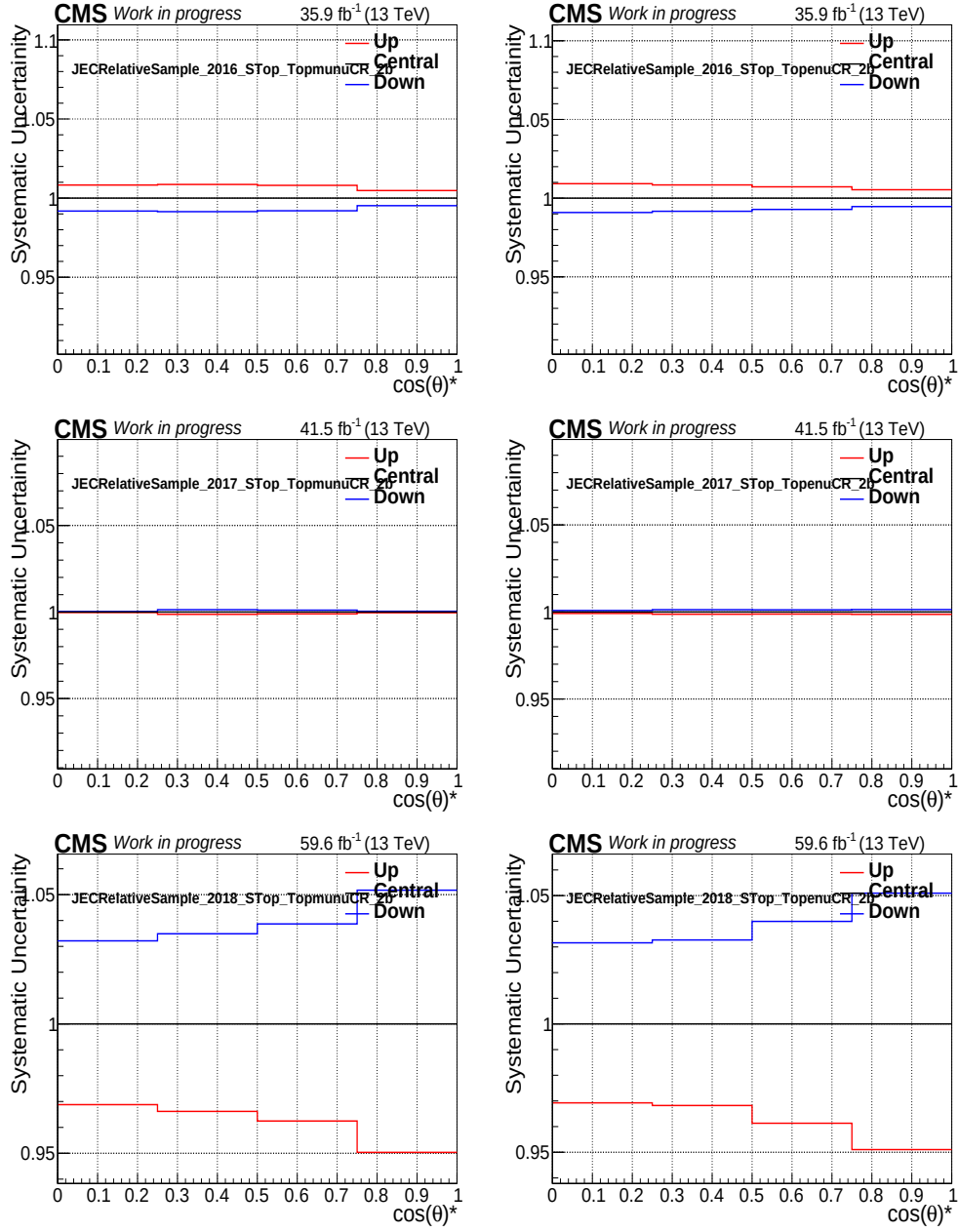


Figure 93: JEC uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ for 2016, 2017 and 2018

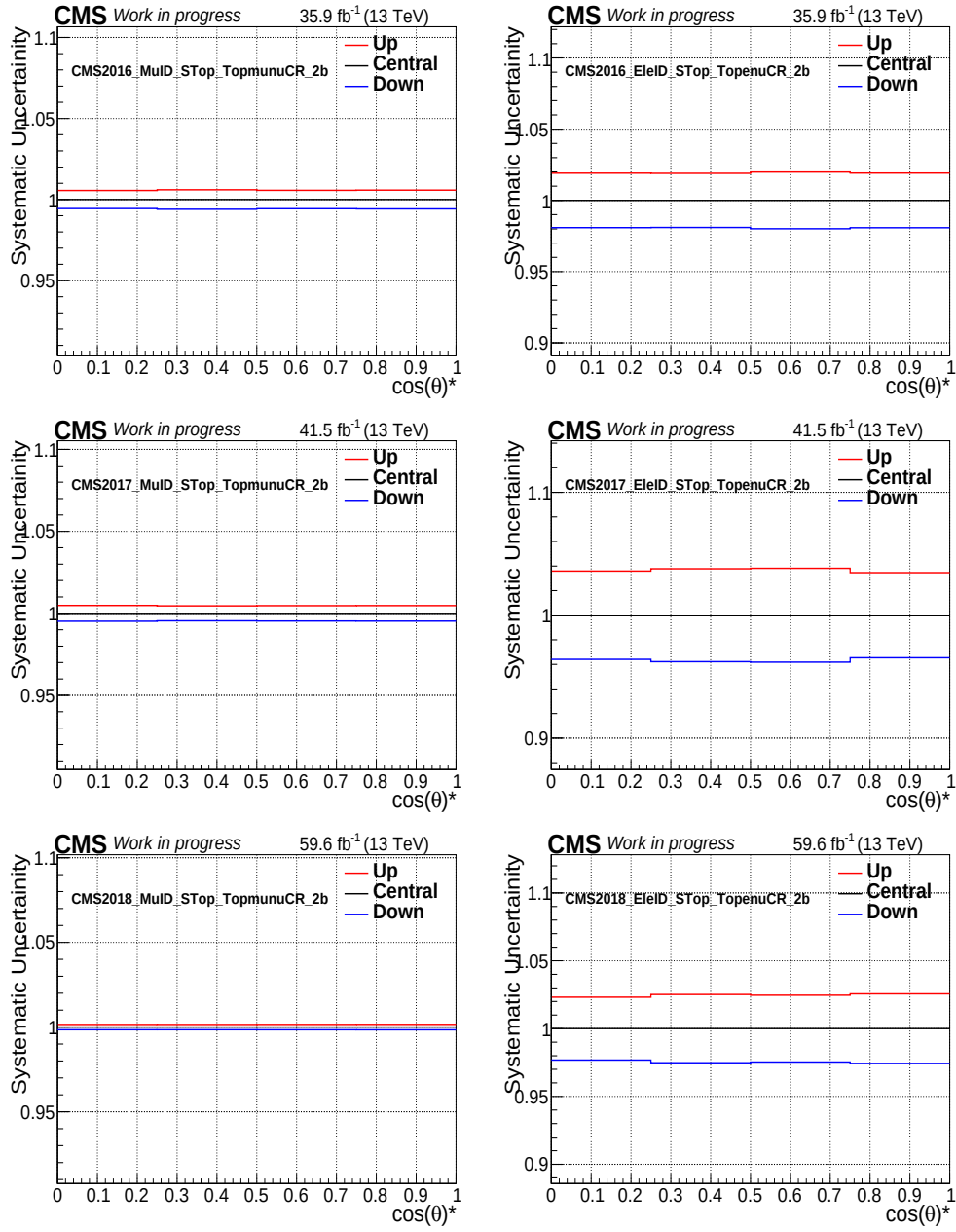


Figure 94: muon ID and electron ID uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ respectively for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

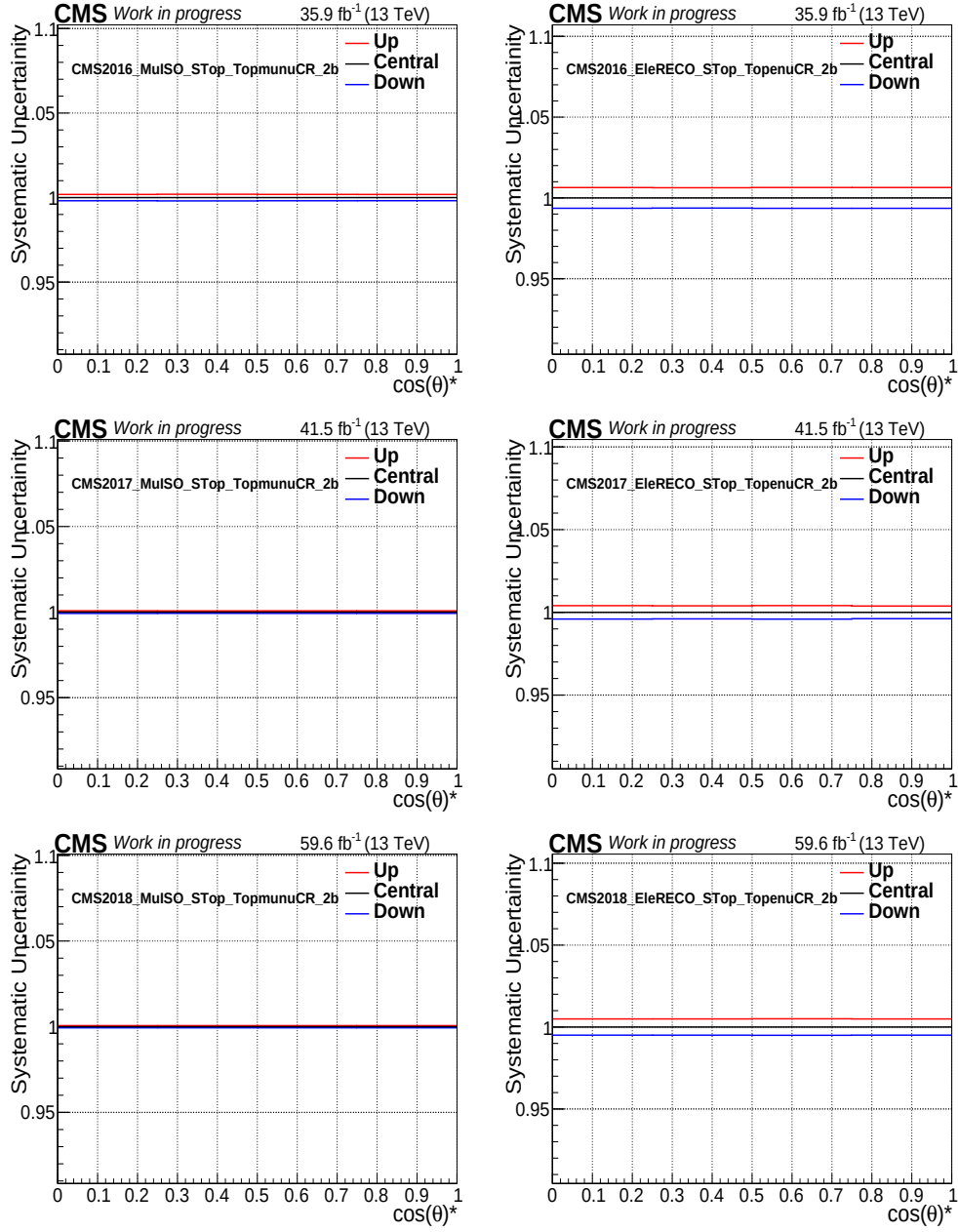


Figure 95: muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 2b Category of $t\bar{t}(\mu)$ and $t\bar{t}(e)$ respectively for 2016, 2017 and 2018

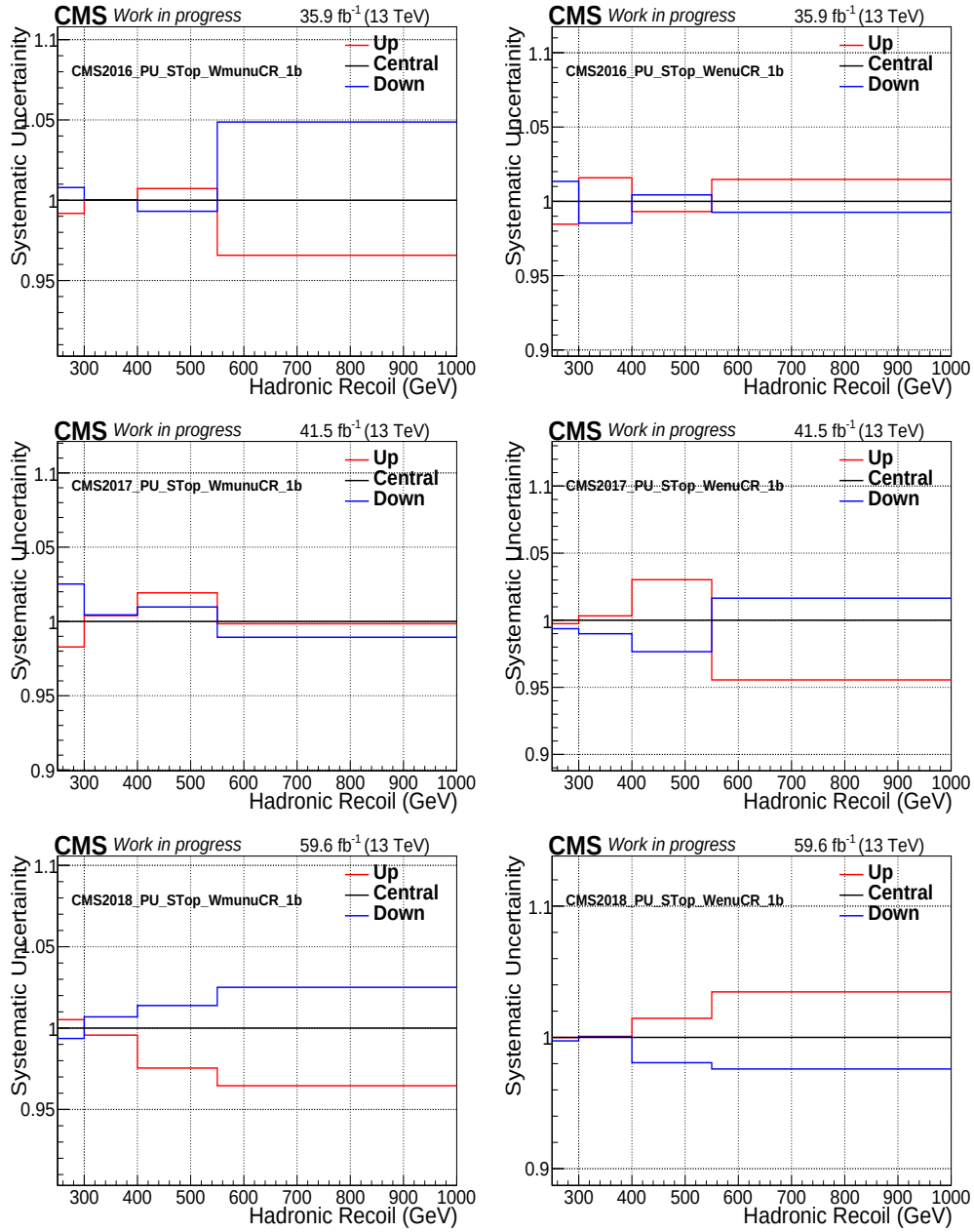


Figure 96: PileUP uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

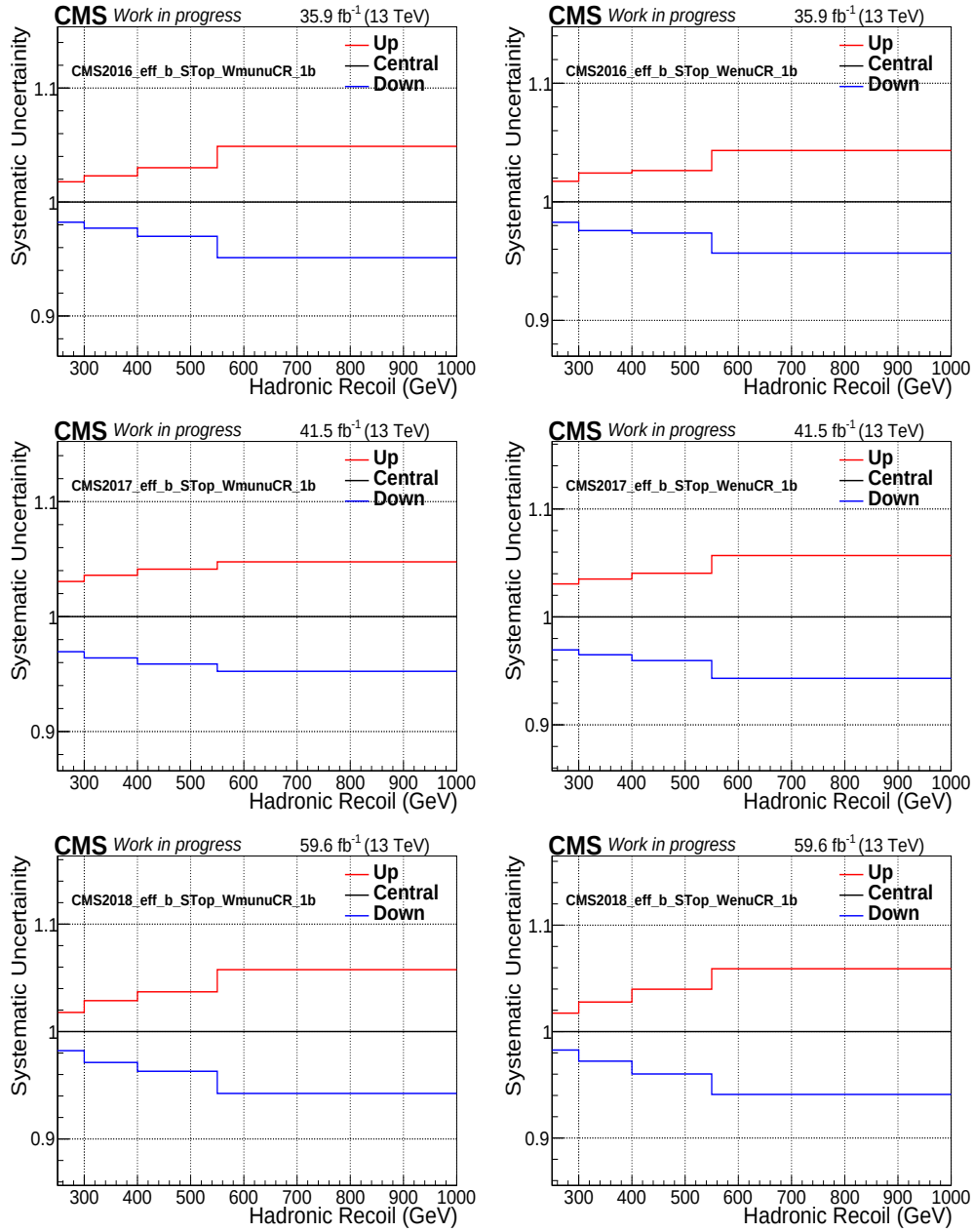


Figure 97: b-tag uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

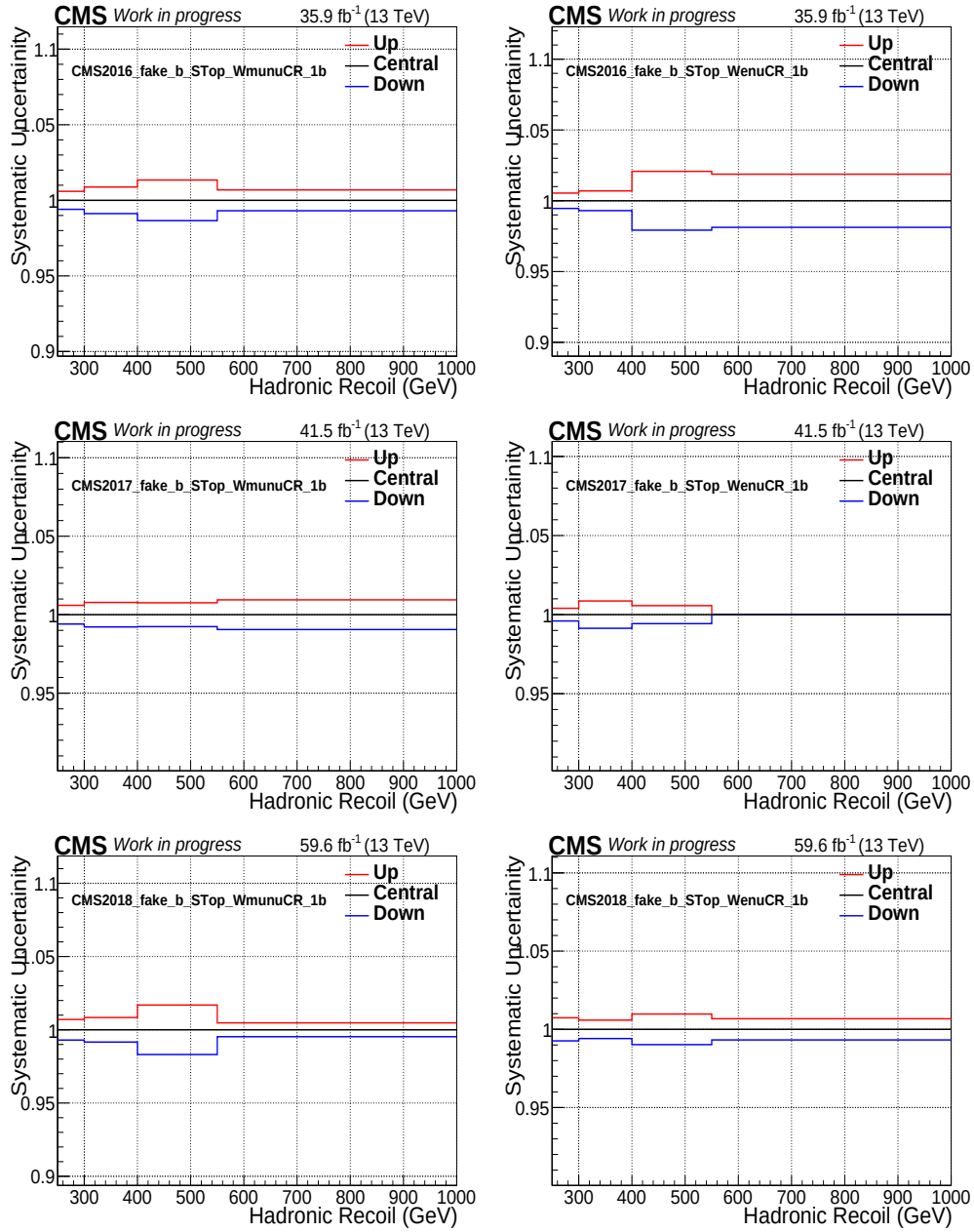


Figure 98: Fake b-tag uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

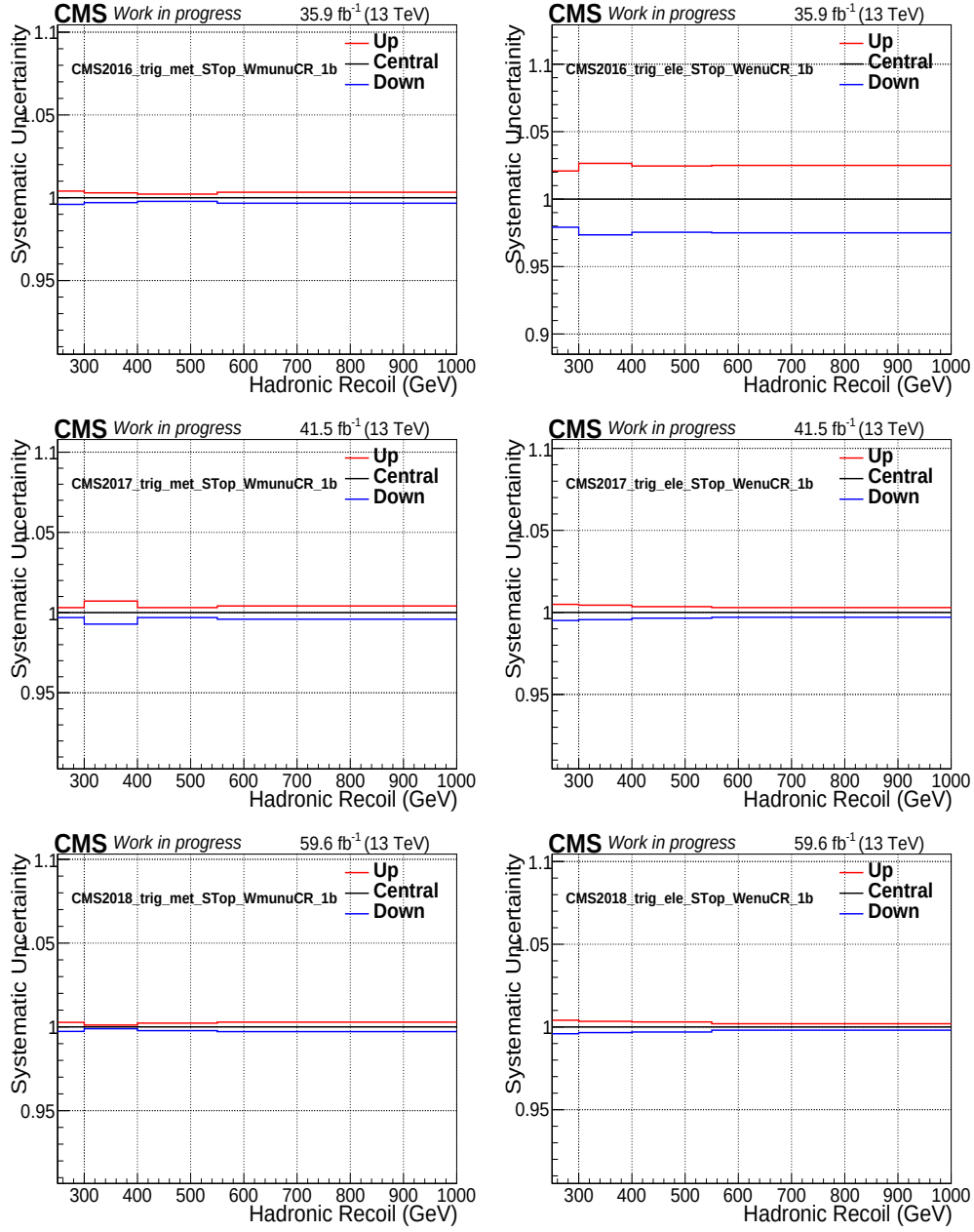


Figure 99: Trigger uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

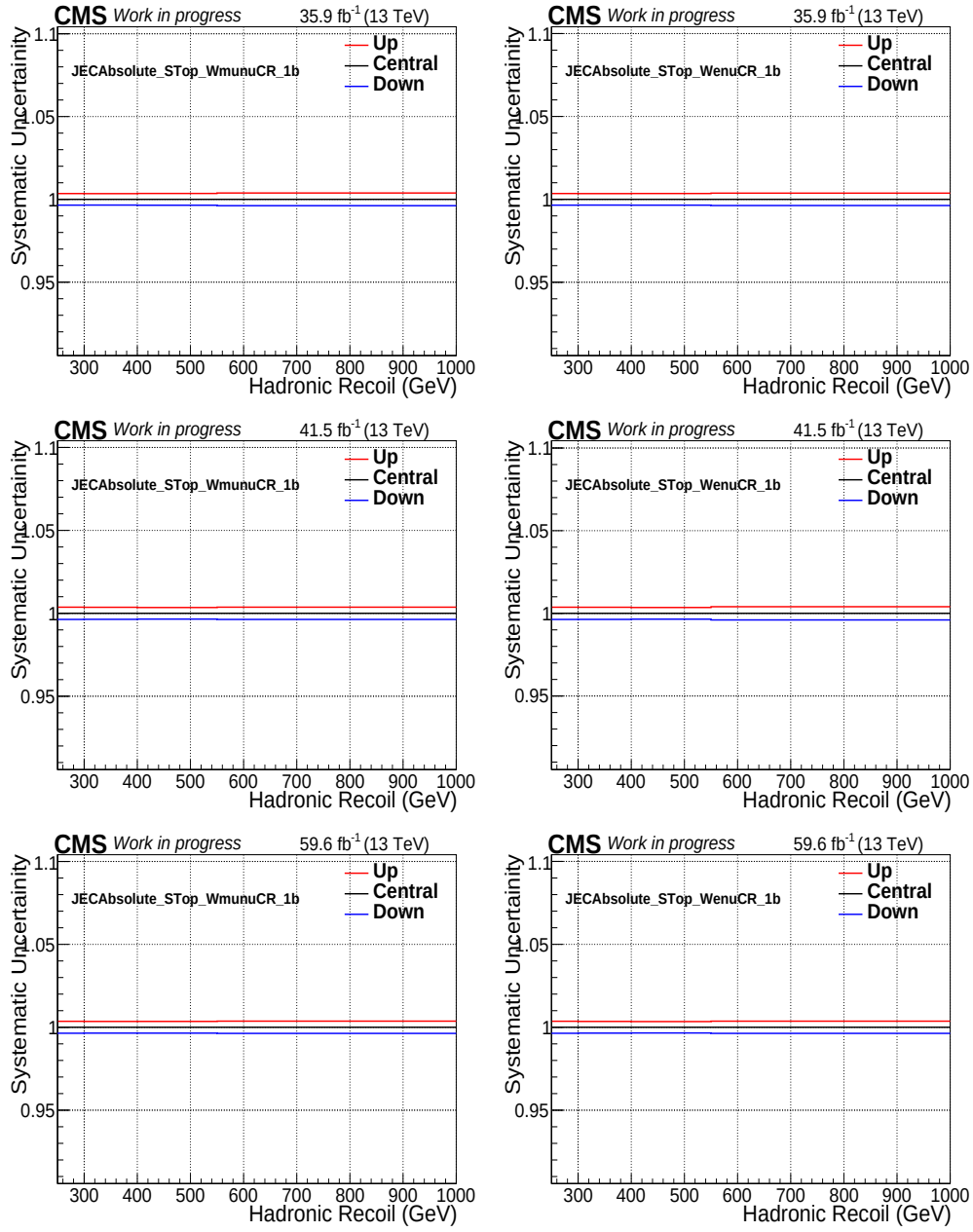


Figure 100: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

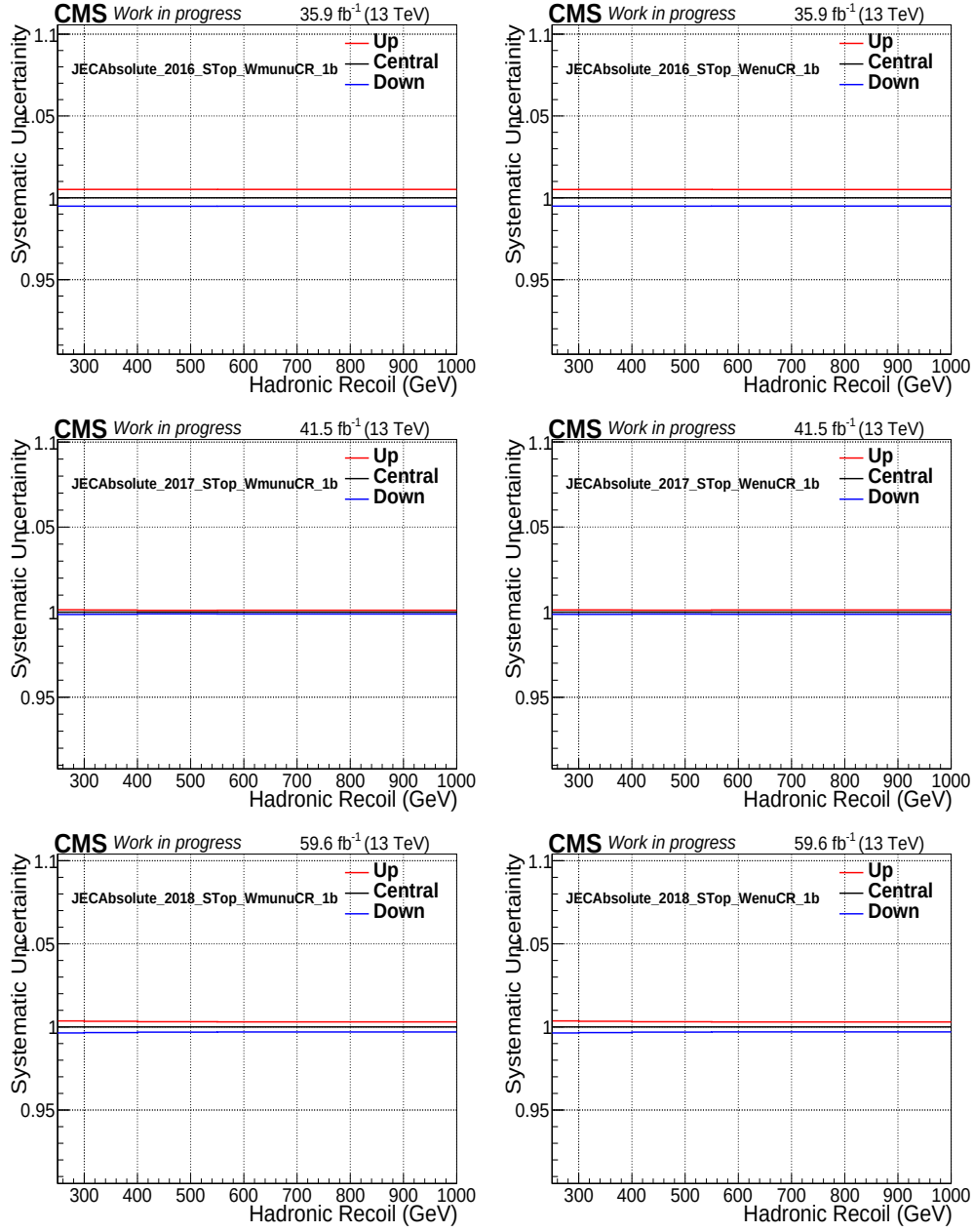


Figure 101: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

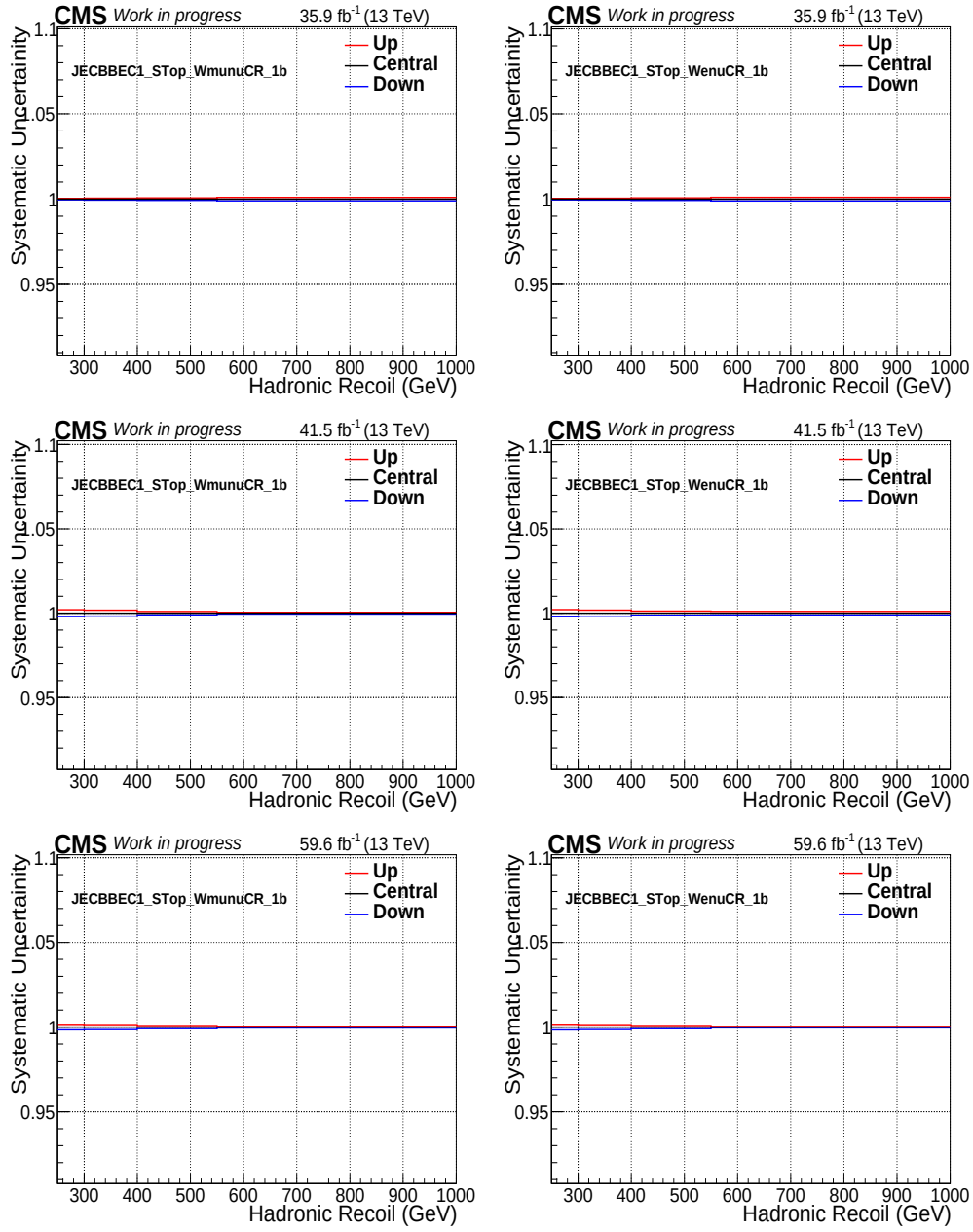


Figure 102: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

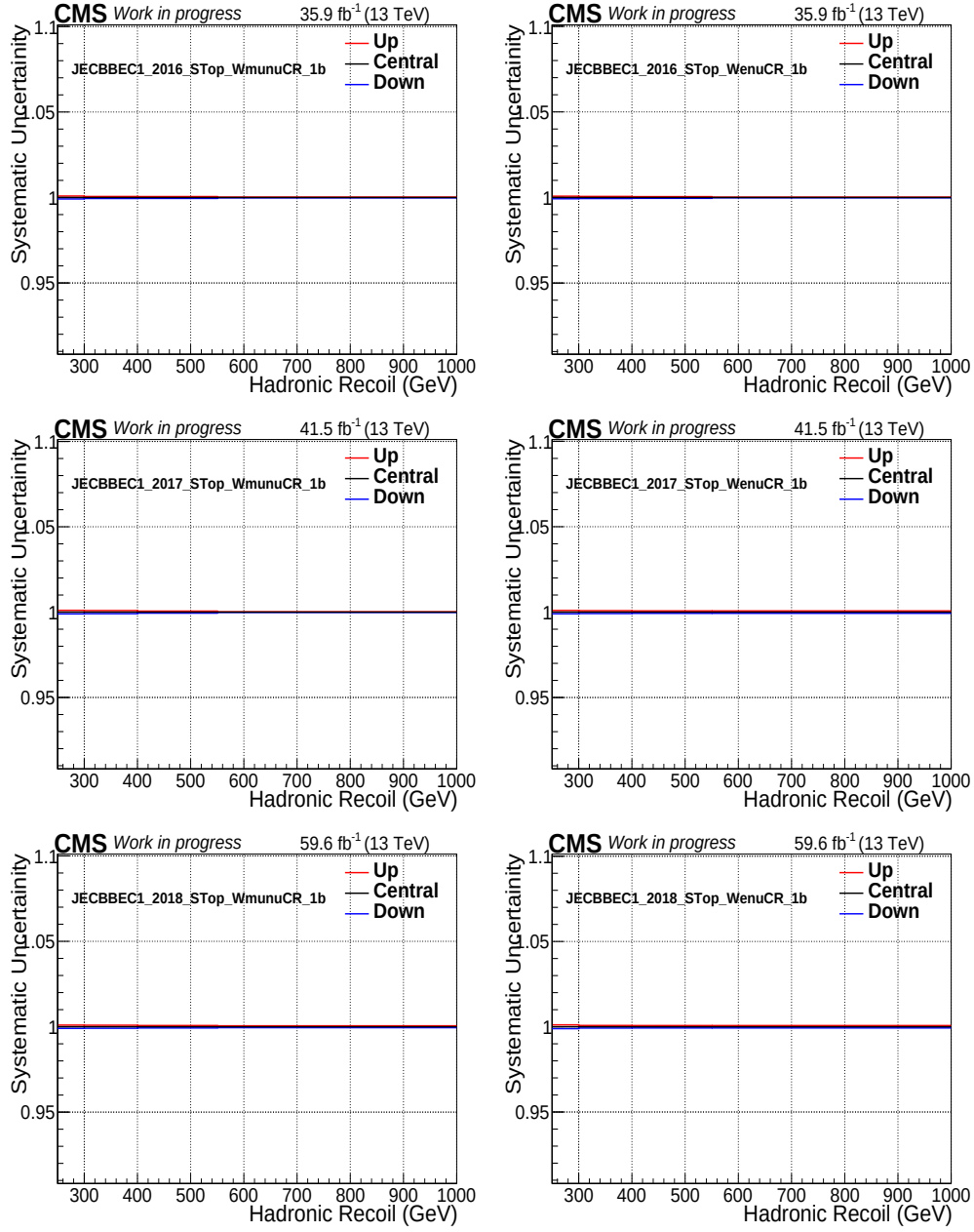


Figure 103: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

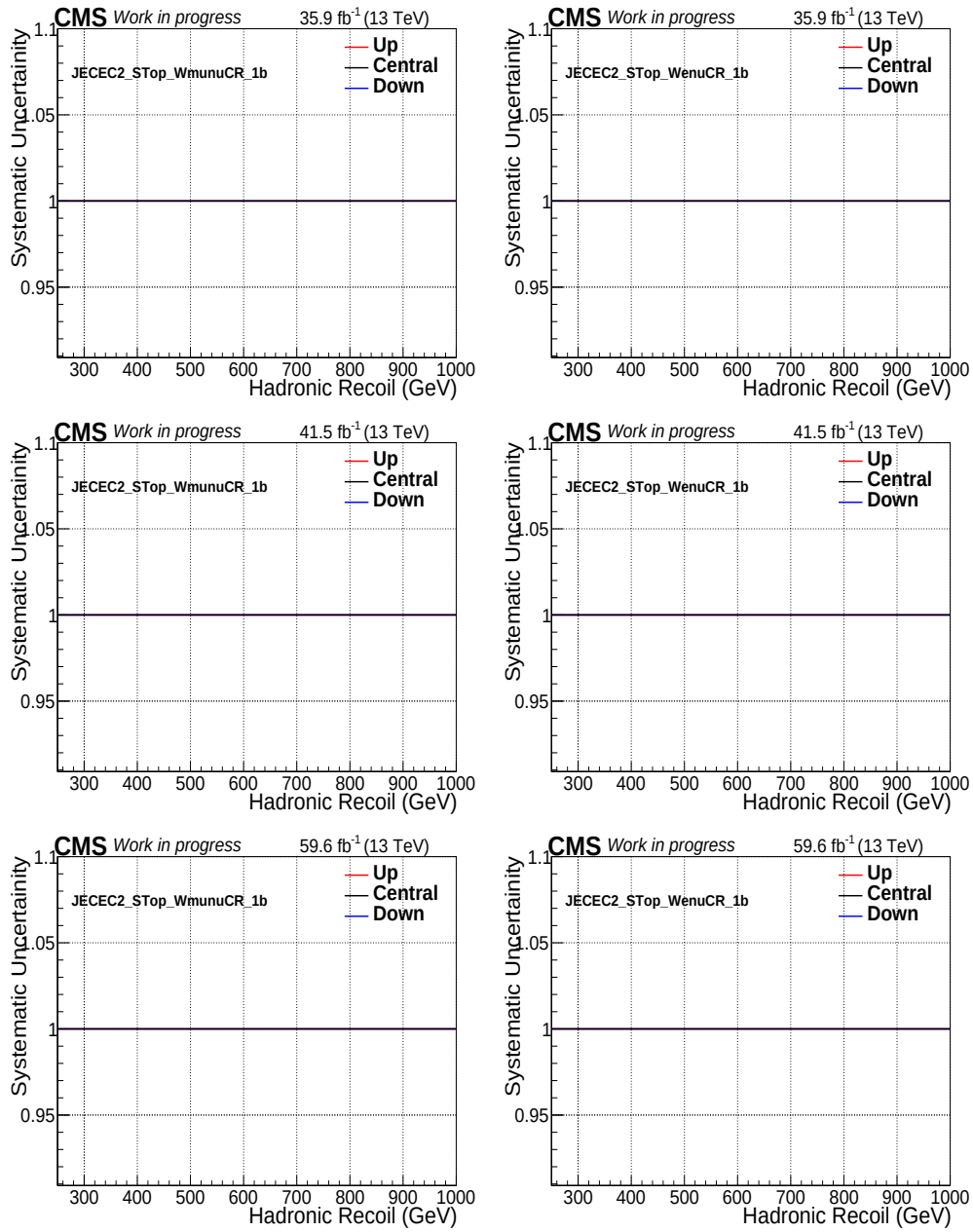


Figure 104: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

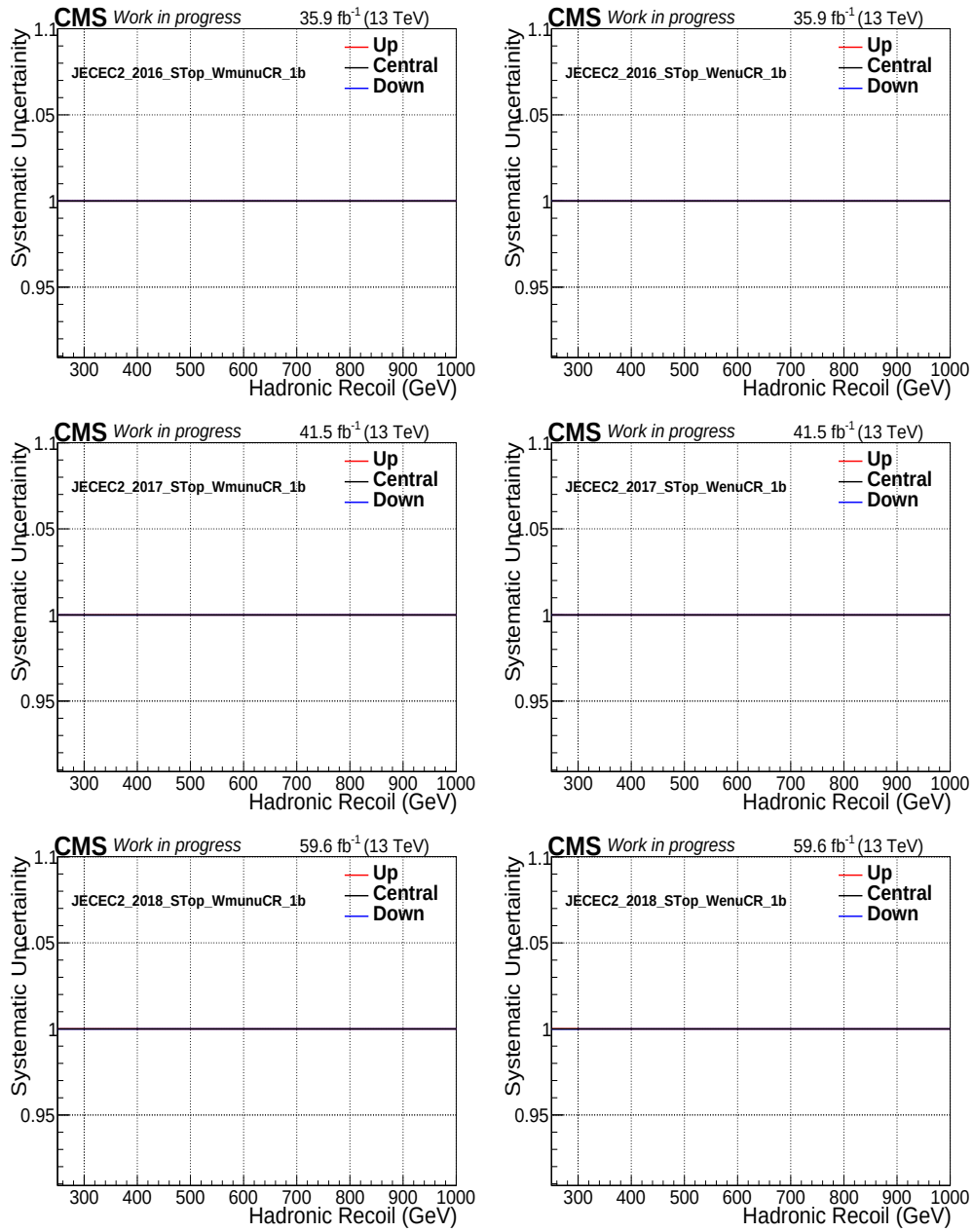


Figure 105: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

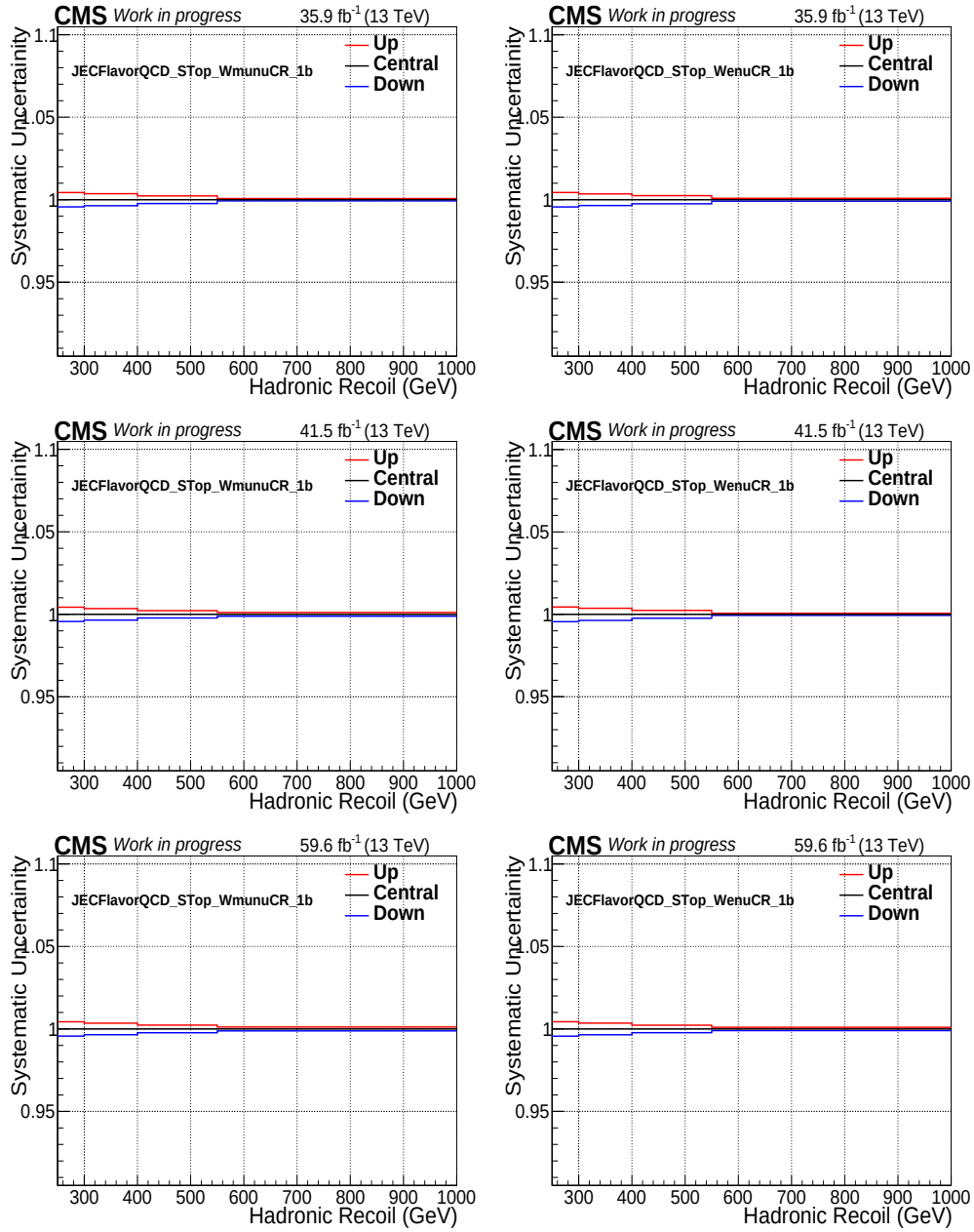


Figure 106: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

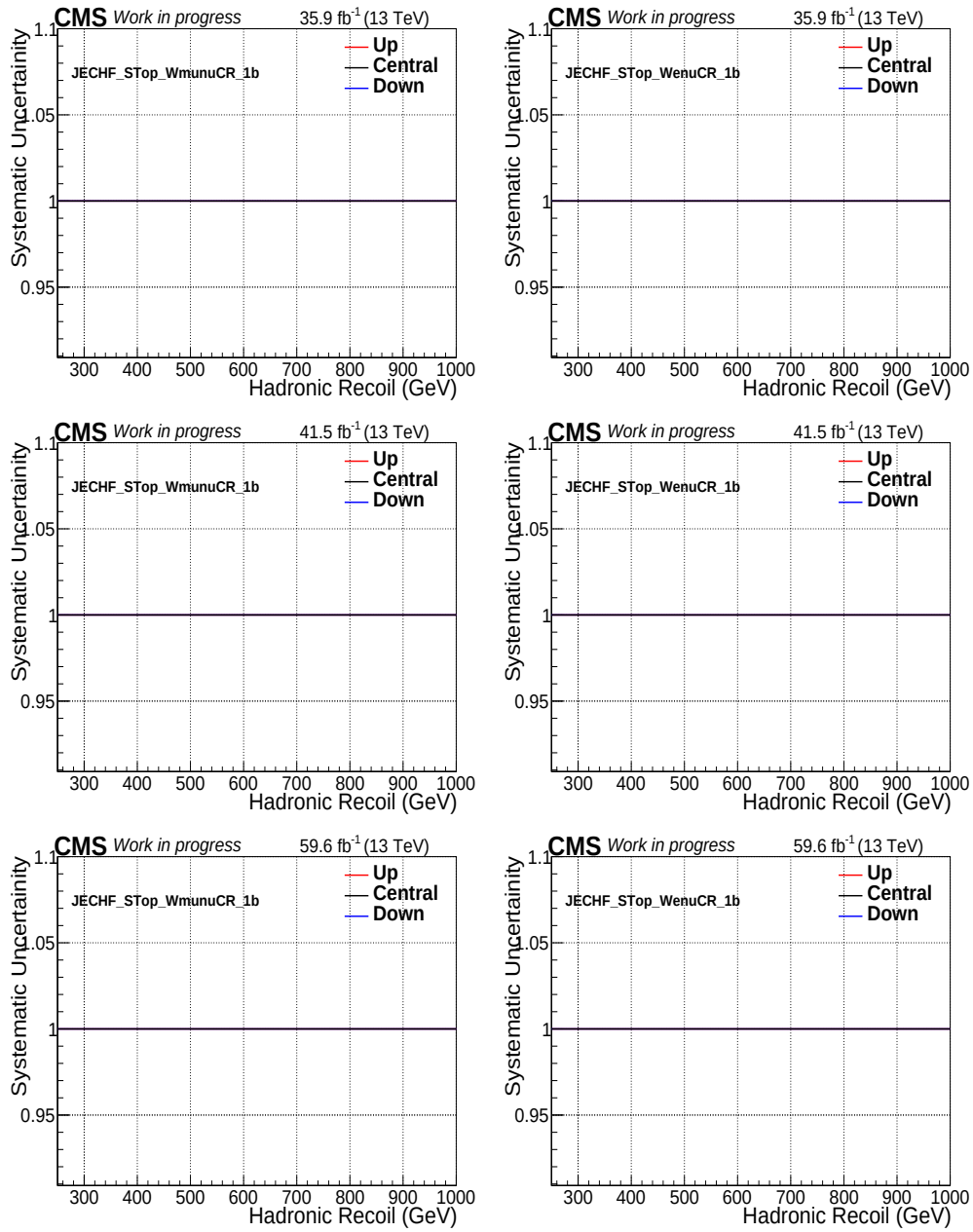


Figure 107: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

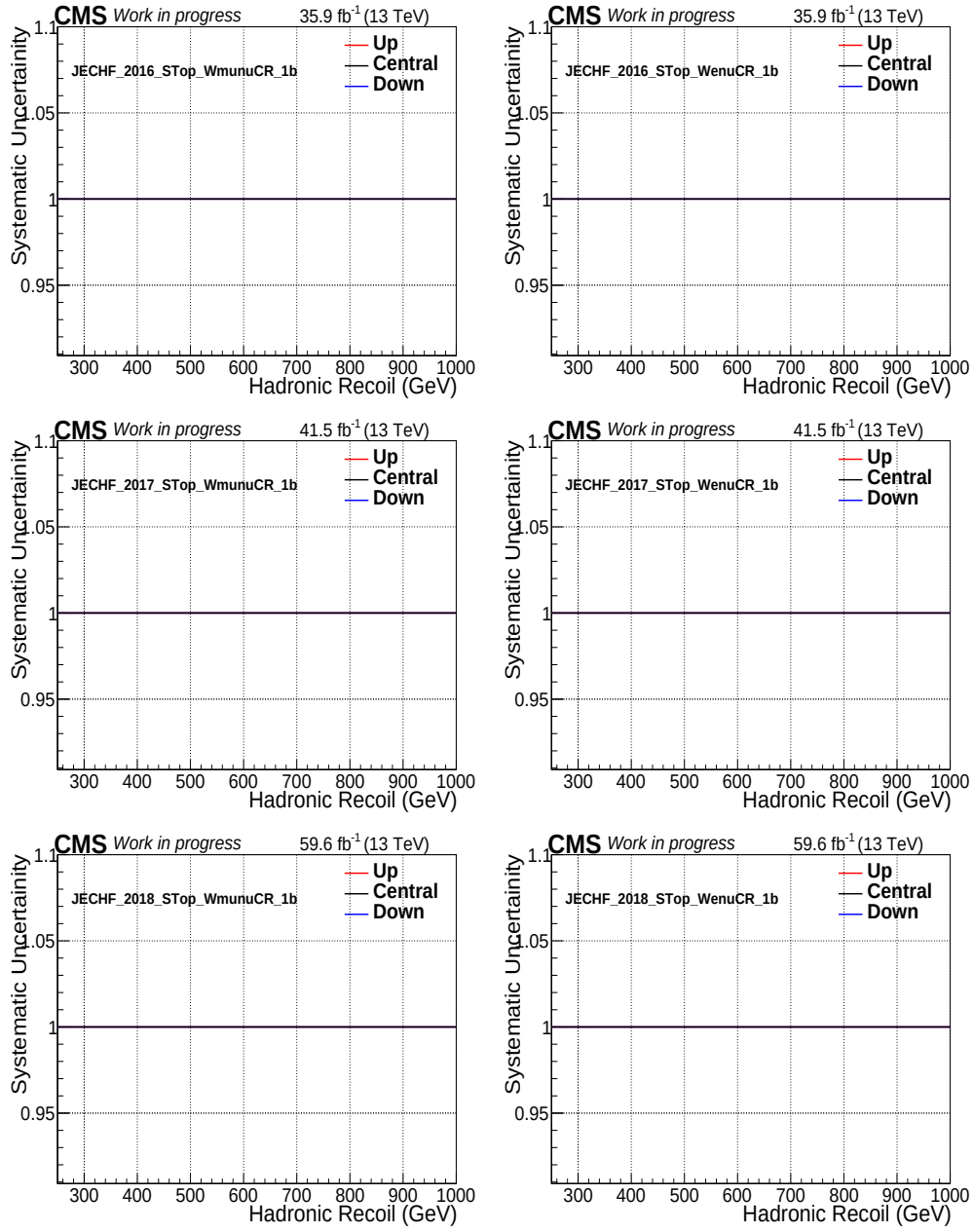


Figure 108: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

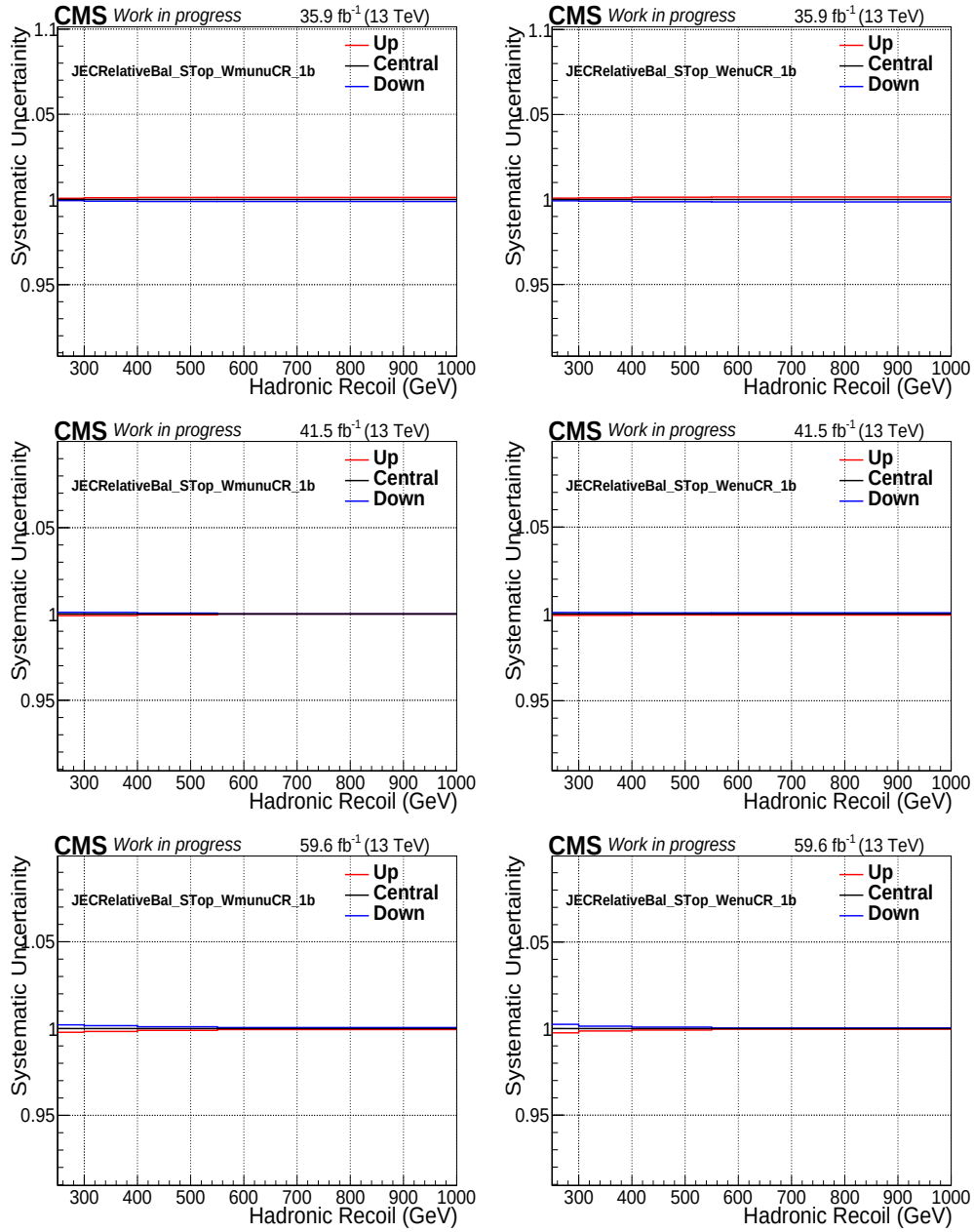


Figure 109: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

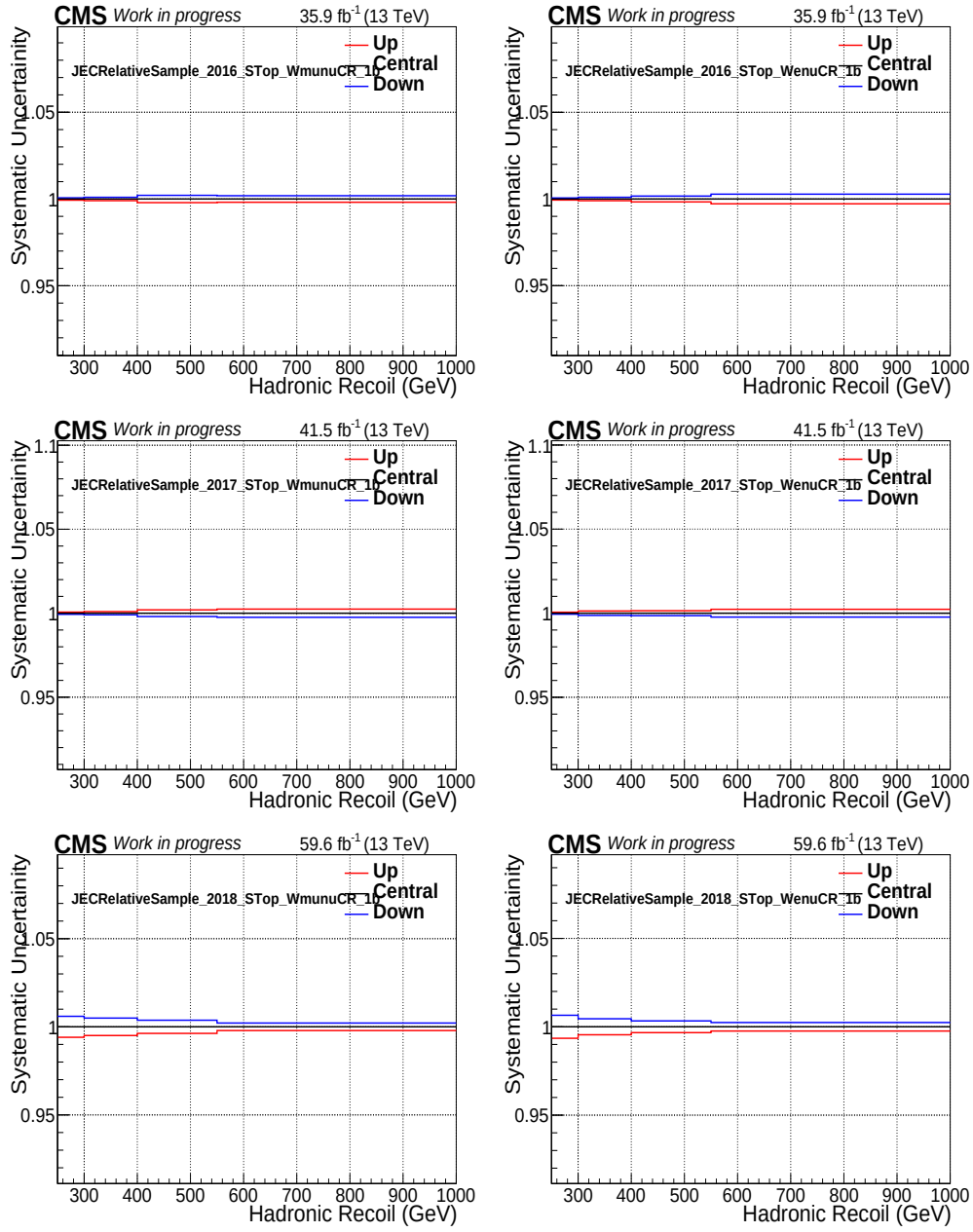


Figure 110: JEC uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

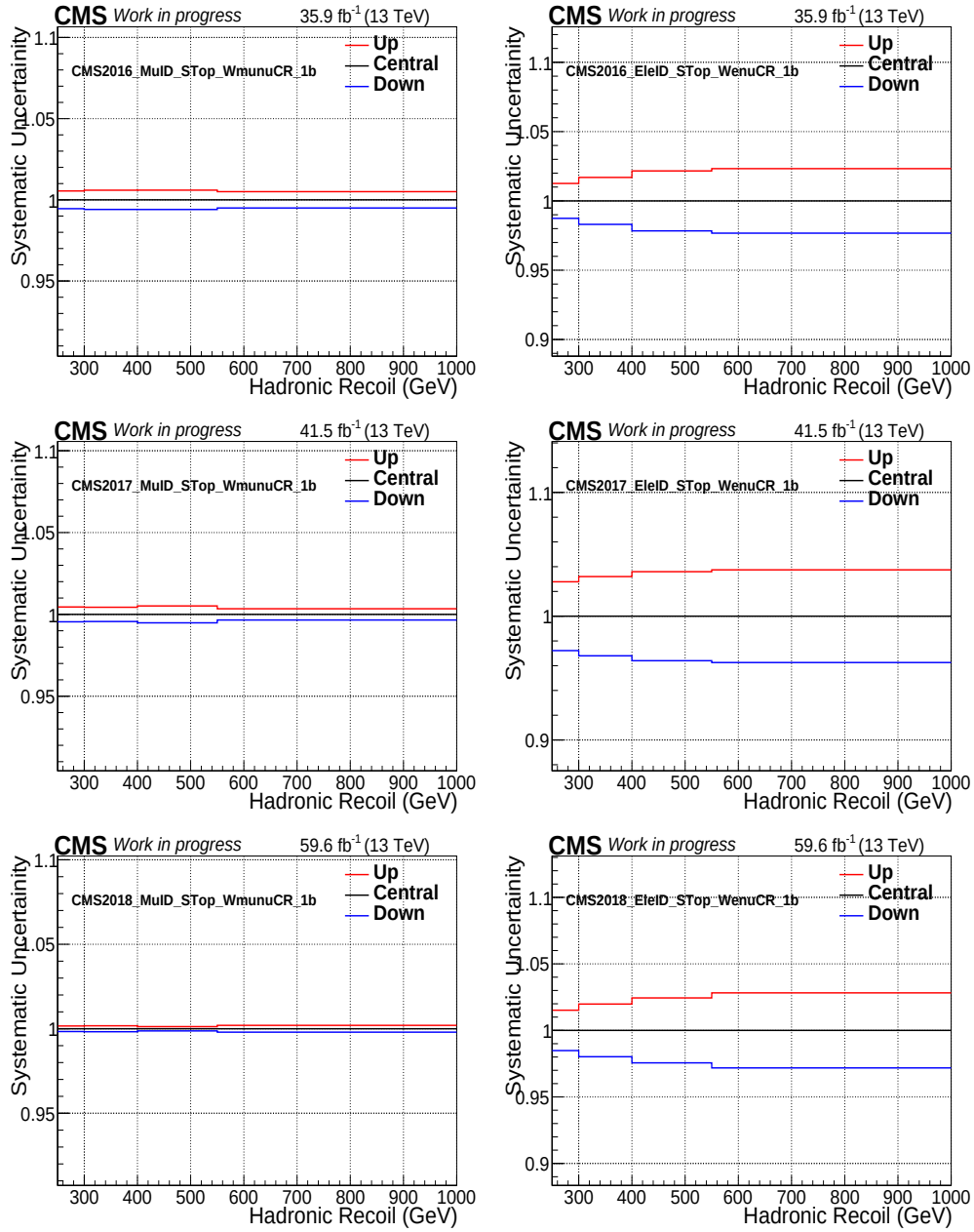


Figure 111: muon ID and electron ID uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ respectively for 2016, 2017 and 2018

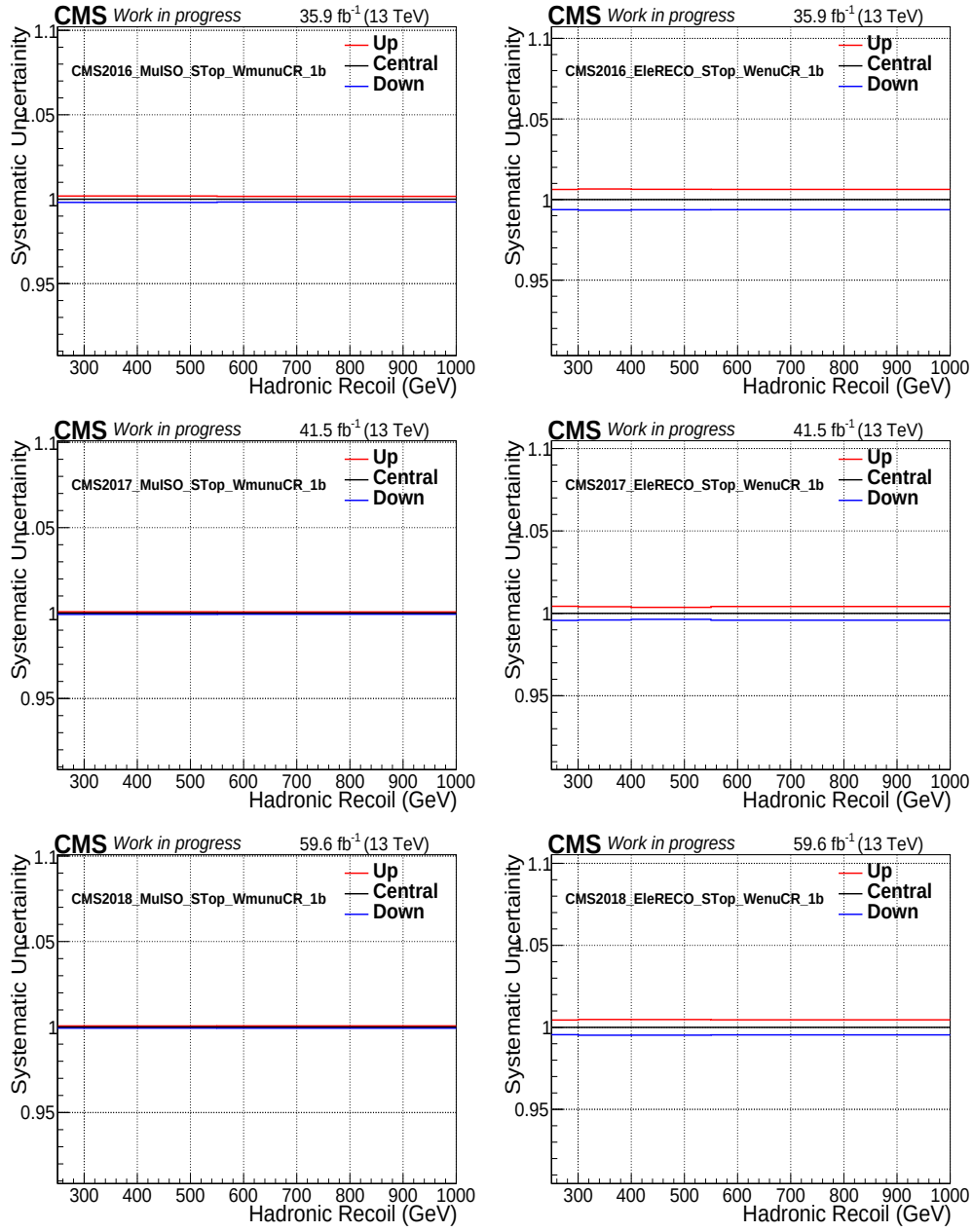


Figure 112: muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 1b Category of $W(\mu\nu)$ and $W(e\nu)$ respectively for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

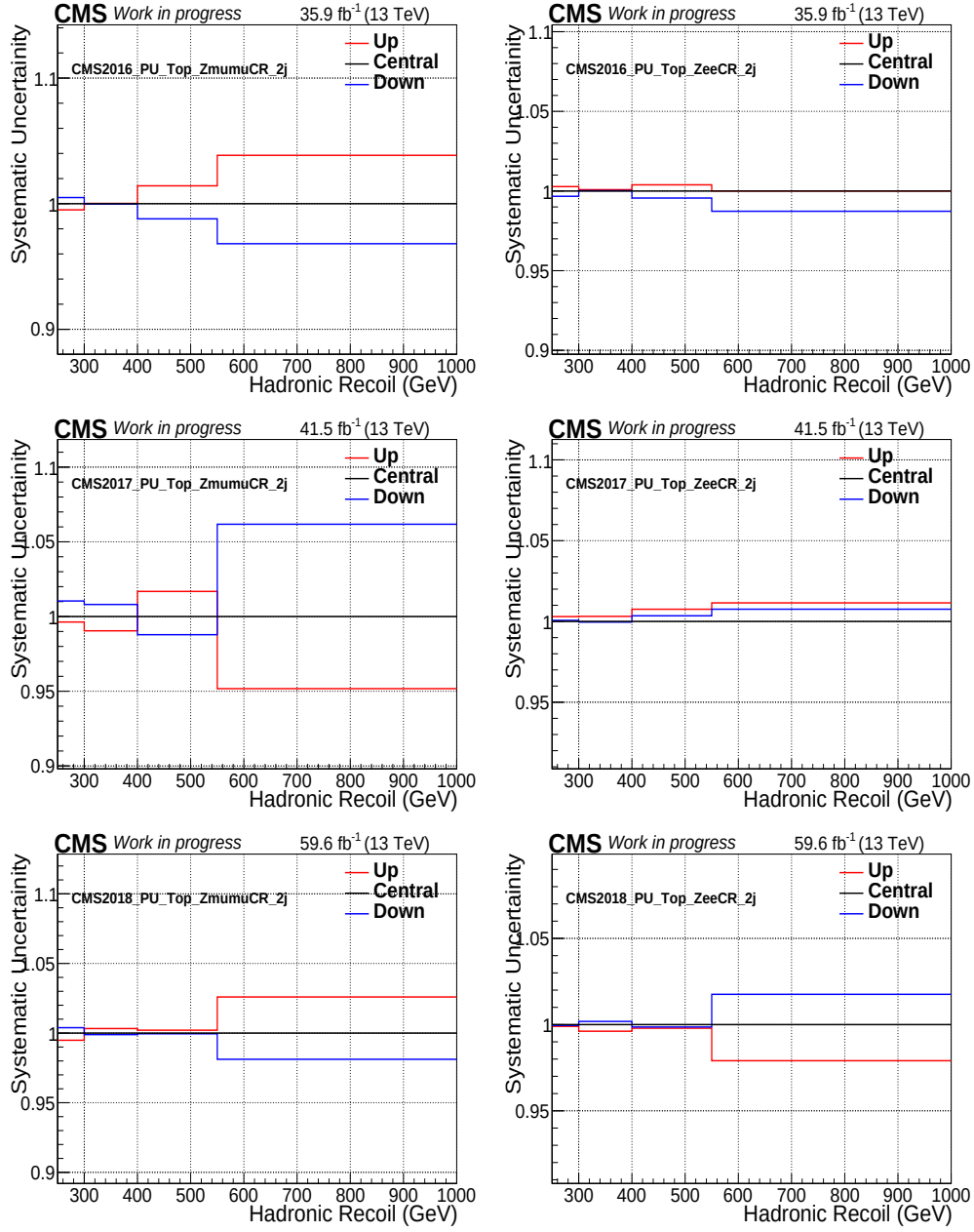


Figure 113: PileUP uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

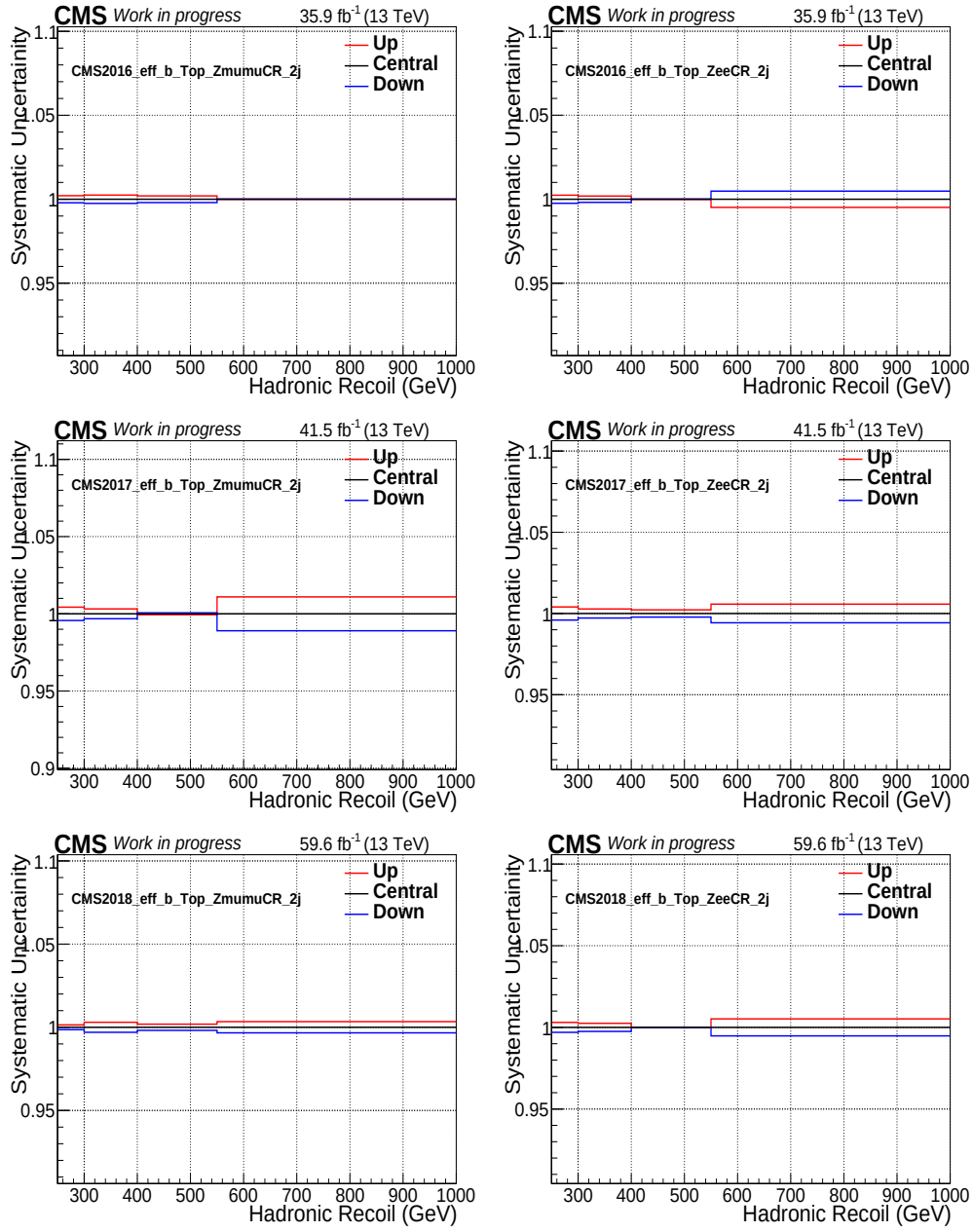


Figure 114: b-tag uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

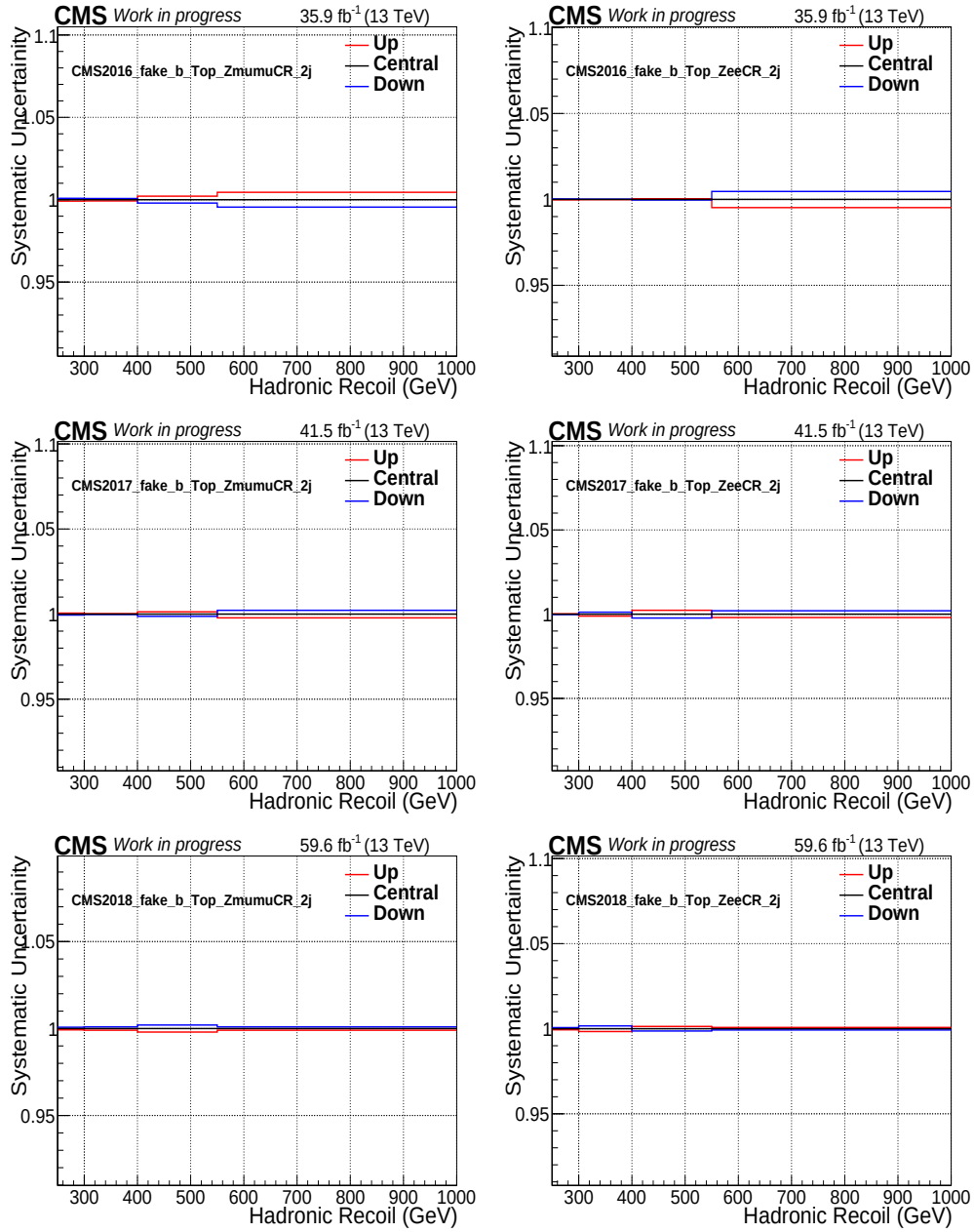


Figure 115: Fake b-tag uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

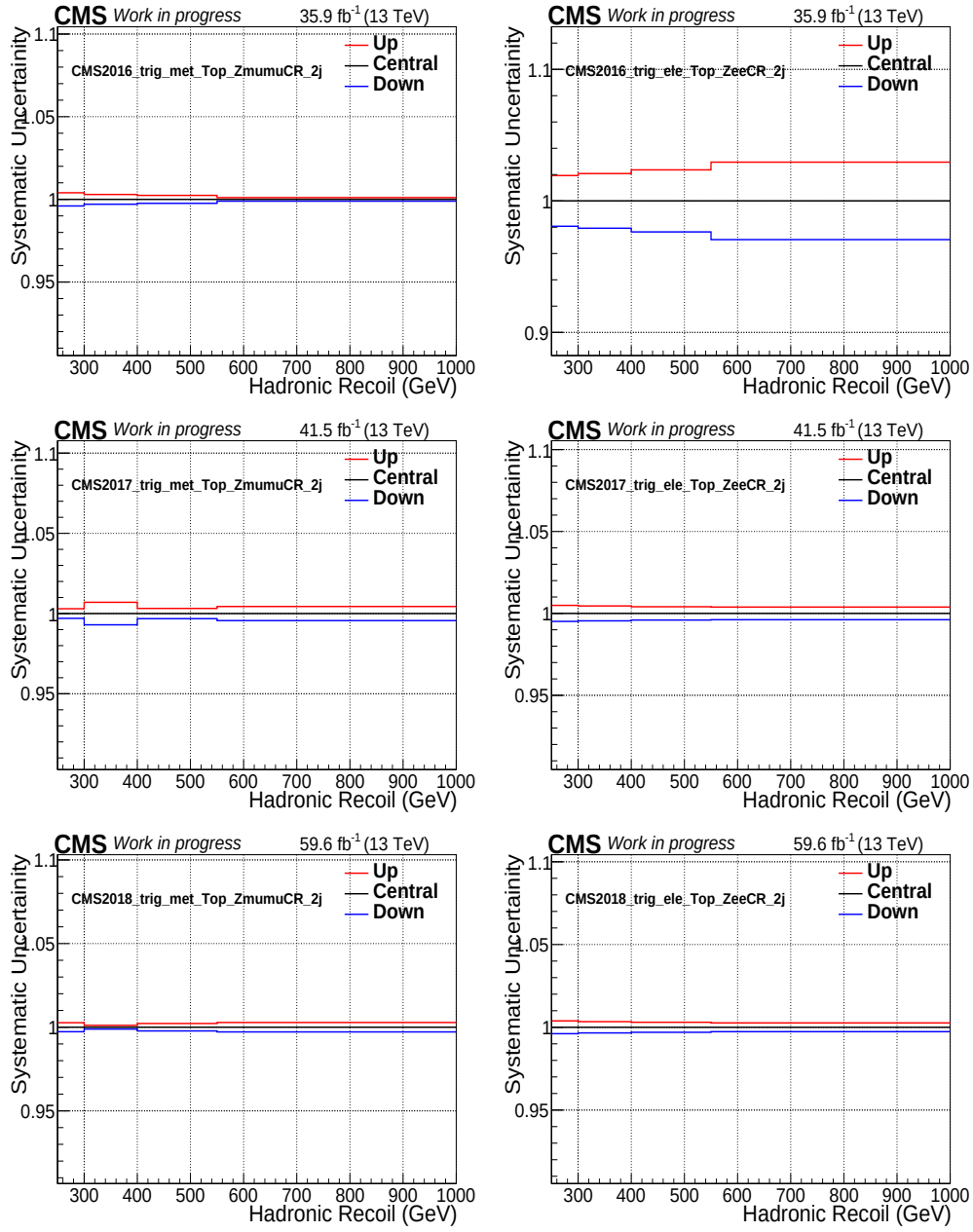


Figure 116: Trigger uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

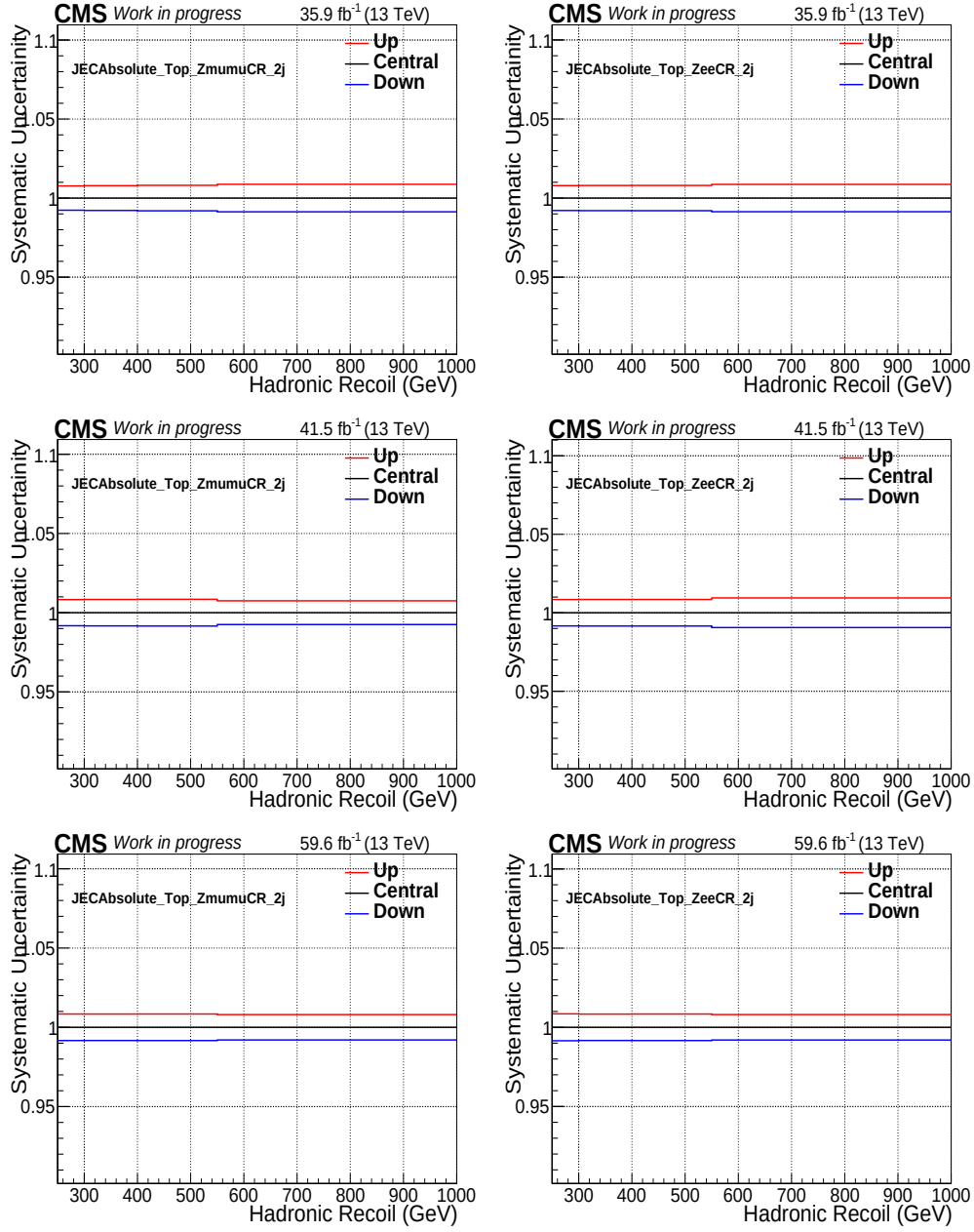


Figure 117: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

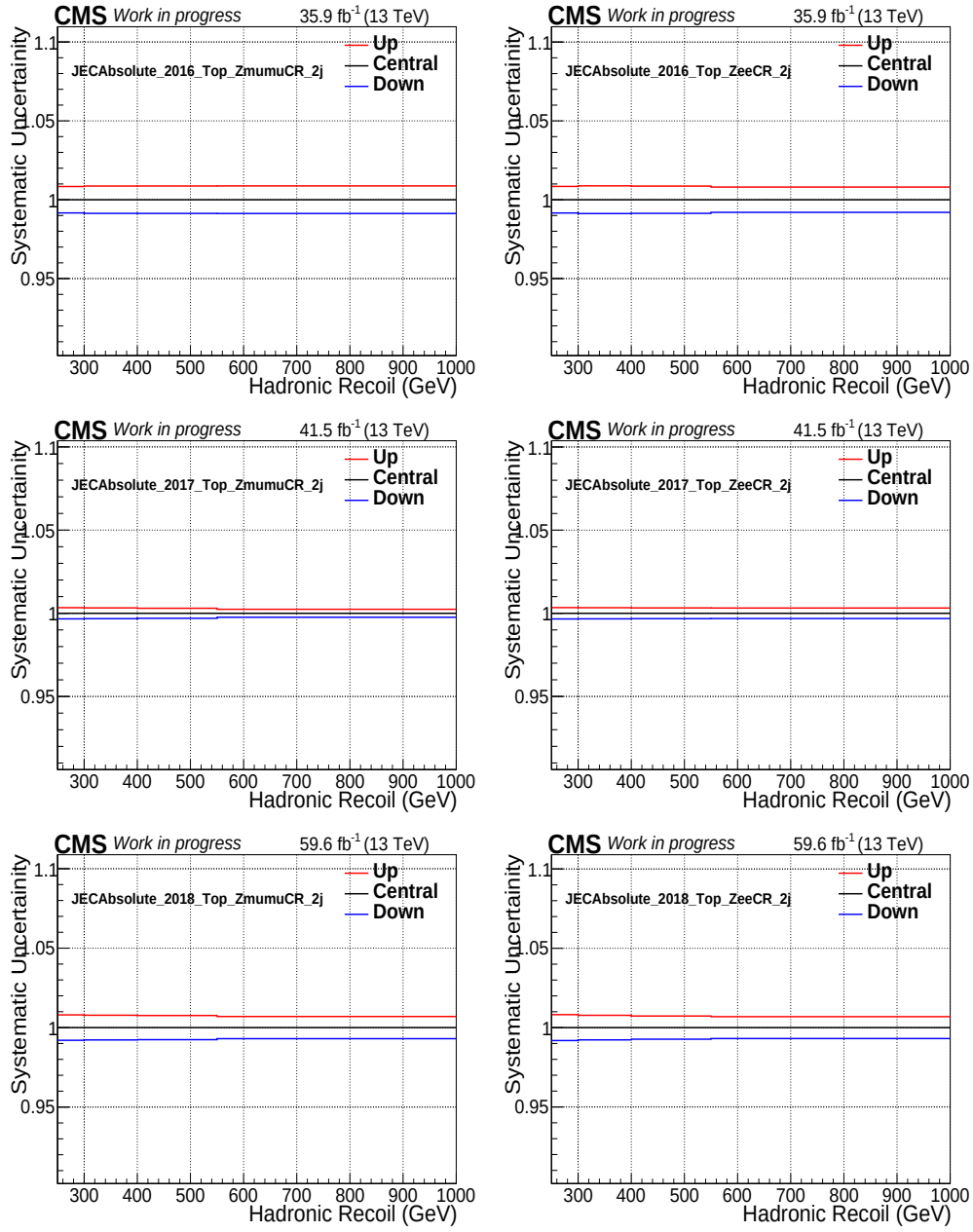


Figure 118: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

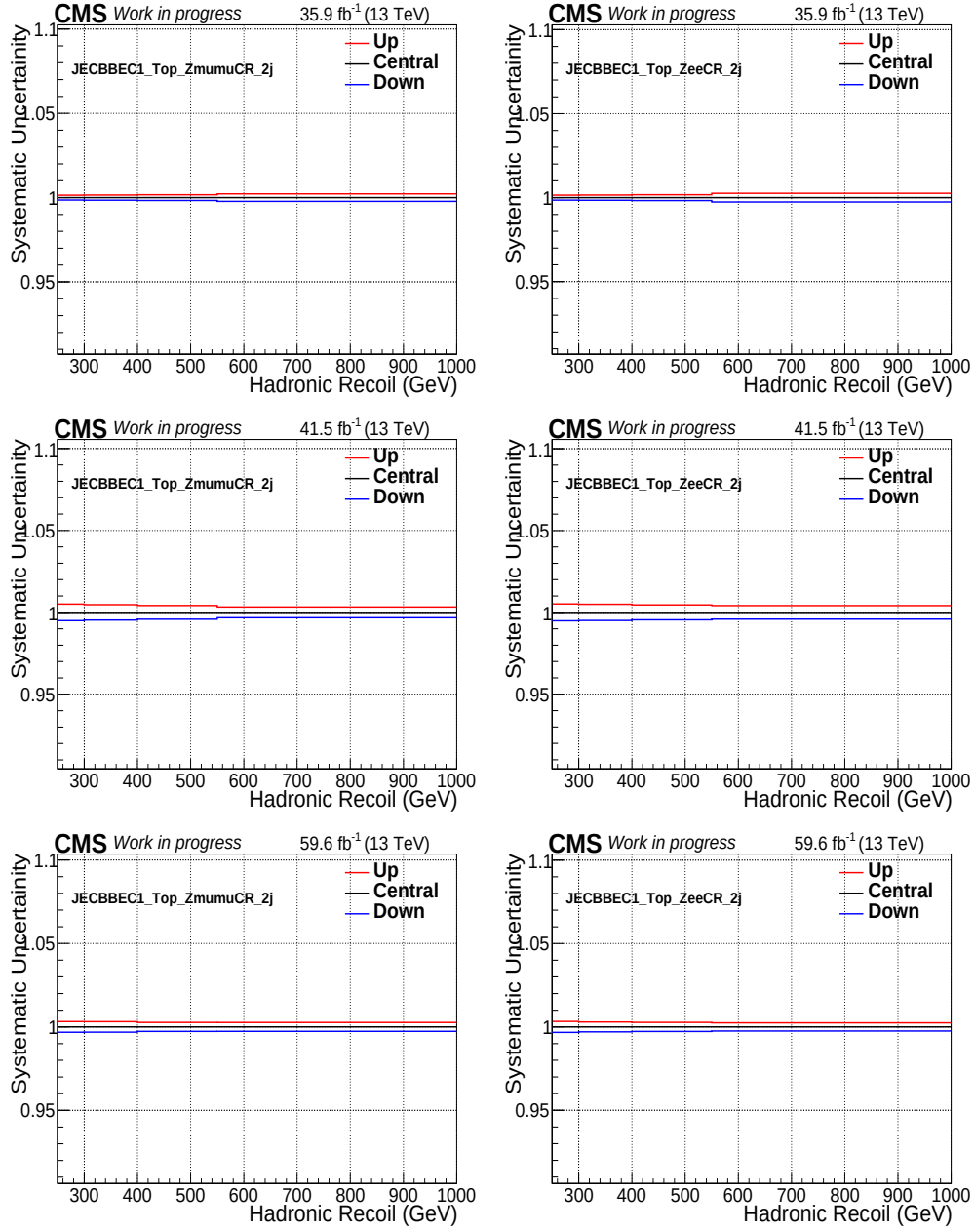


Figure 119: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

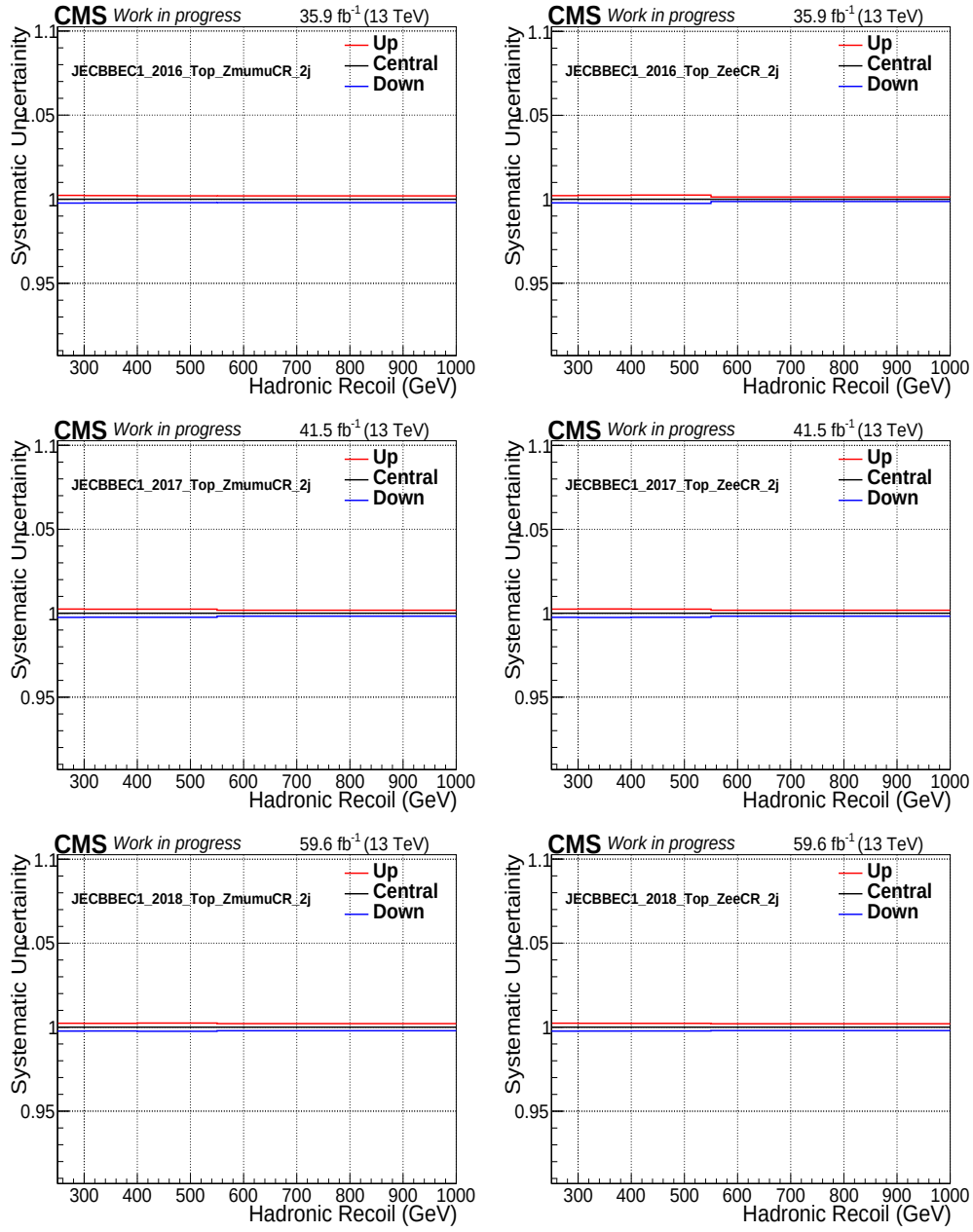


Figure 120: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

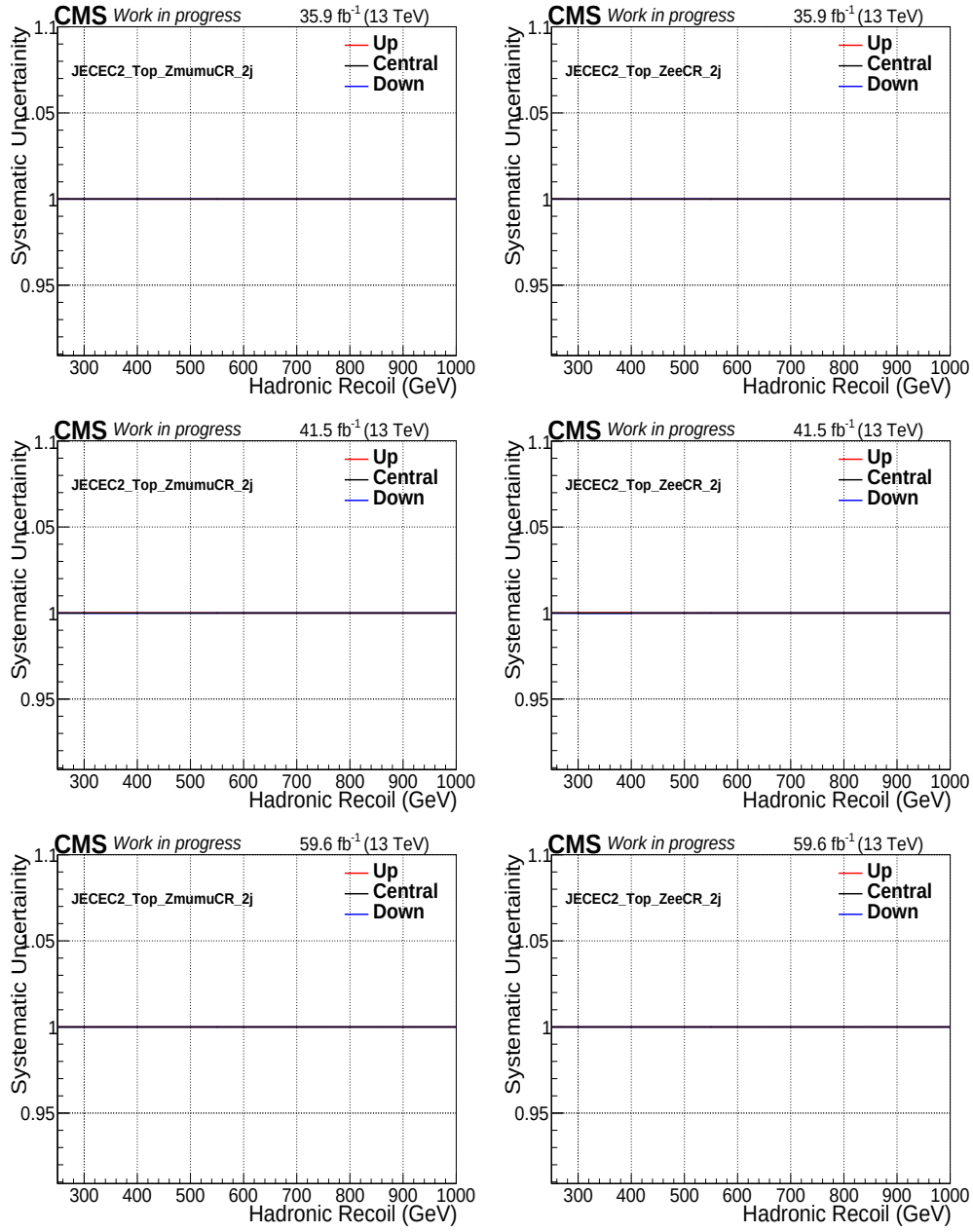


Figure 121: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

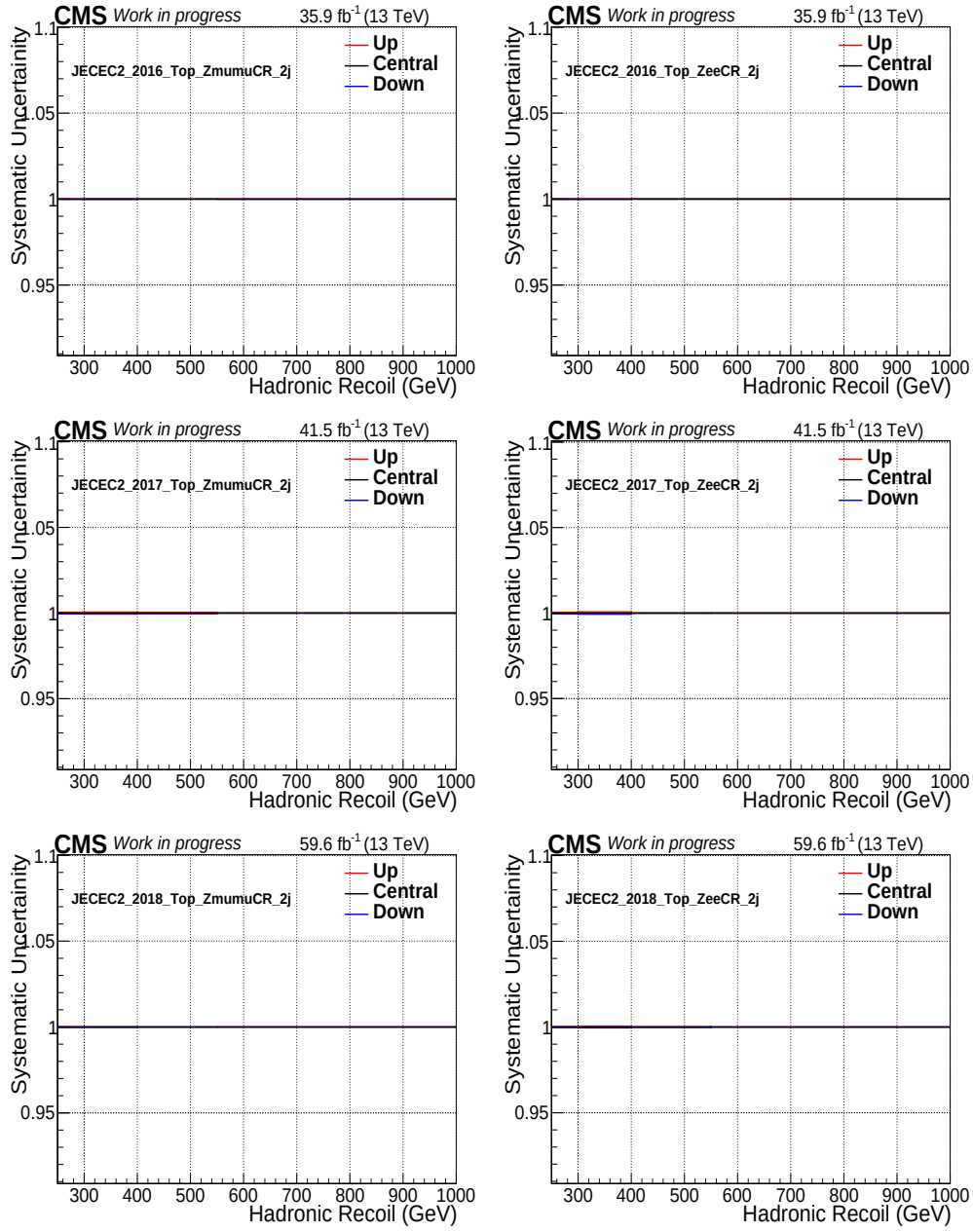


Figure 122: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

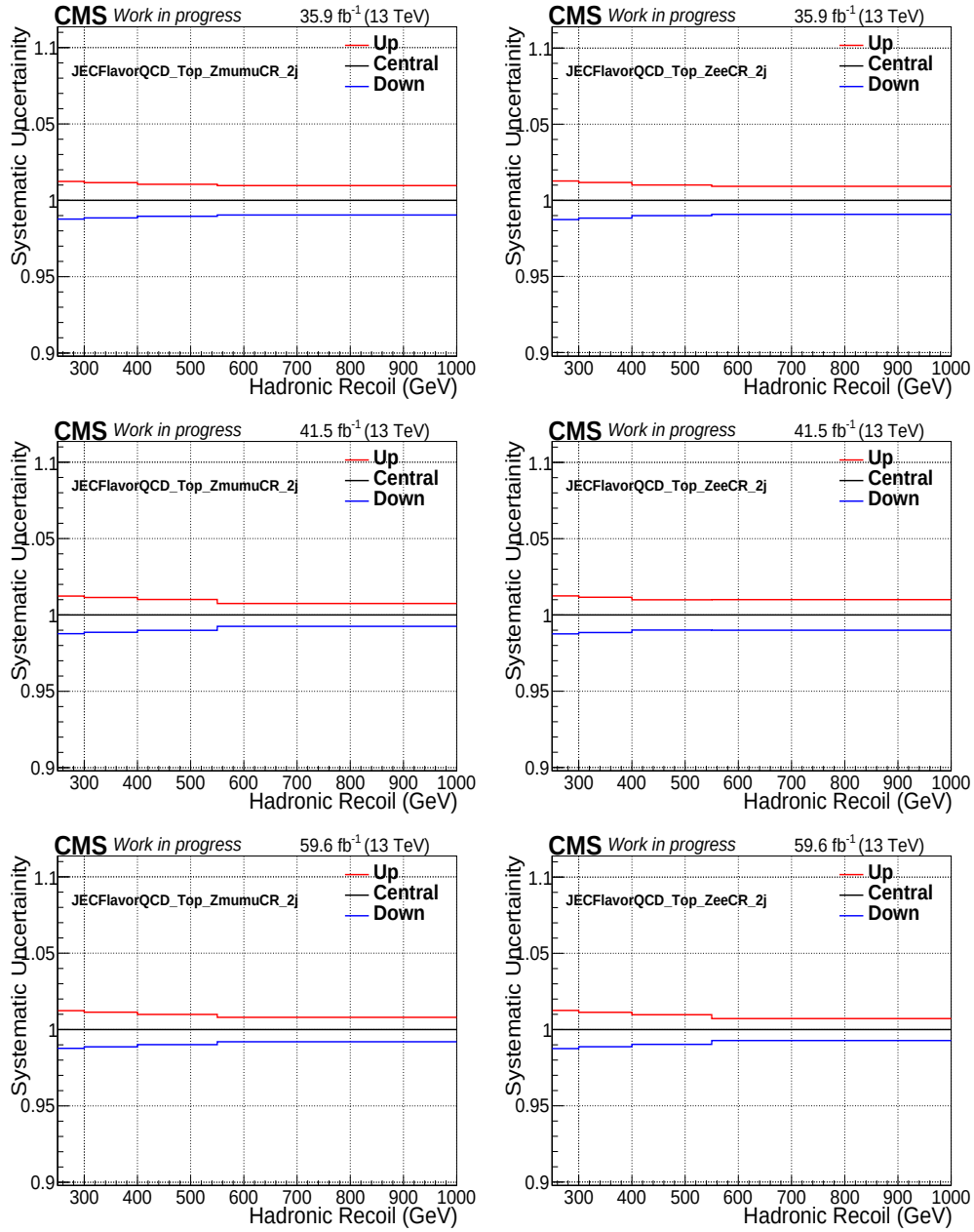


Figure 123: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

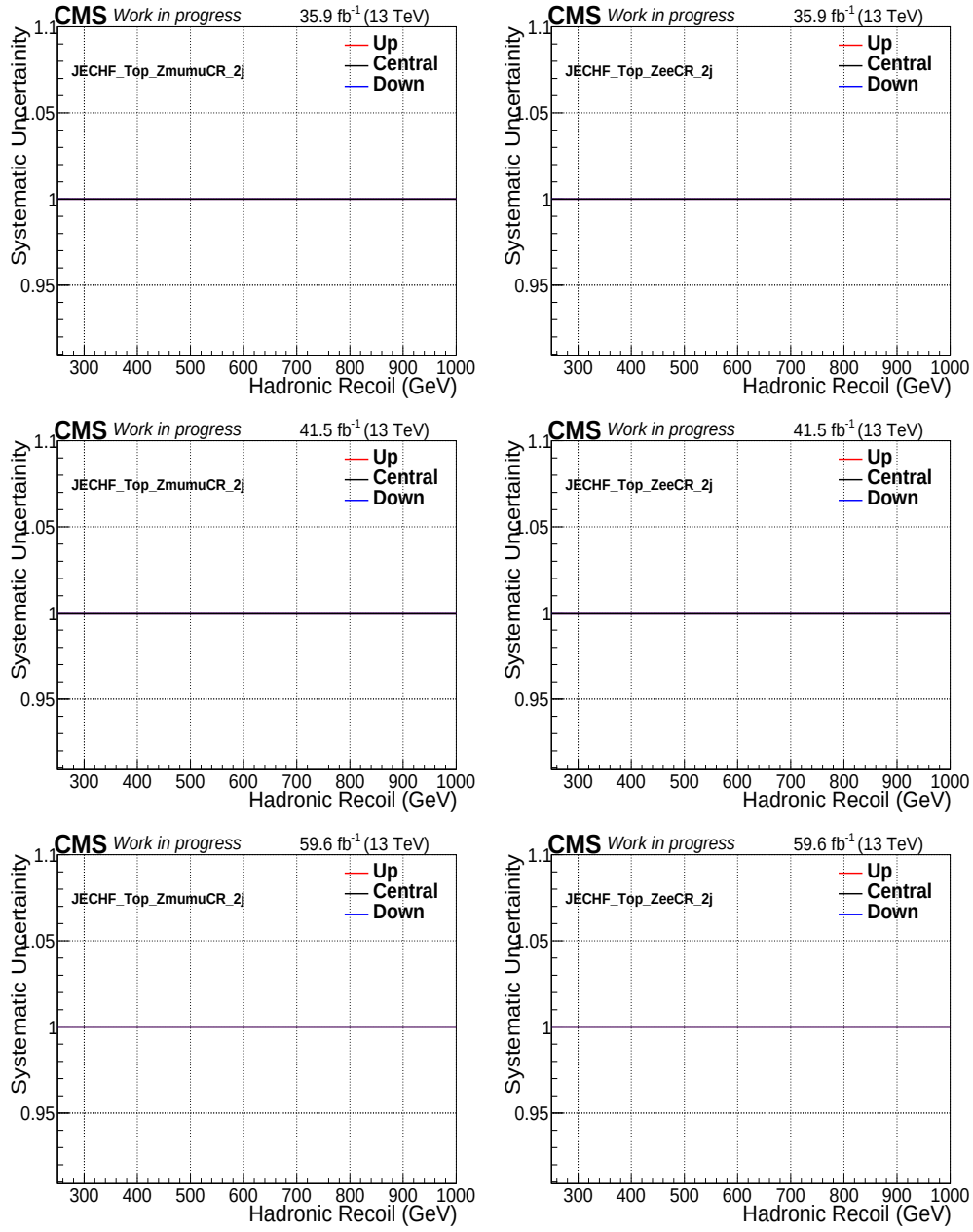


Figure 124: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

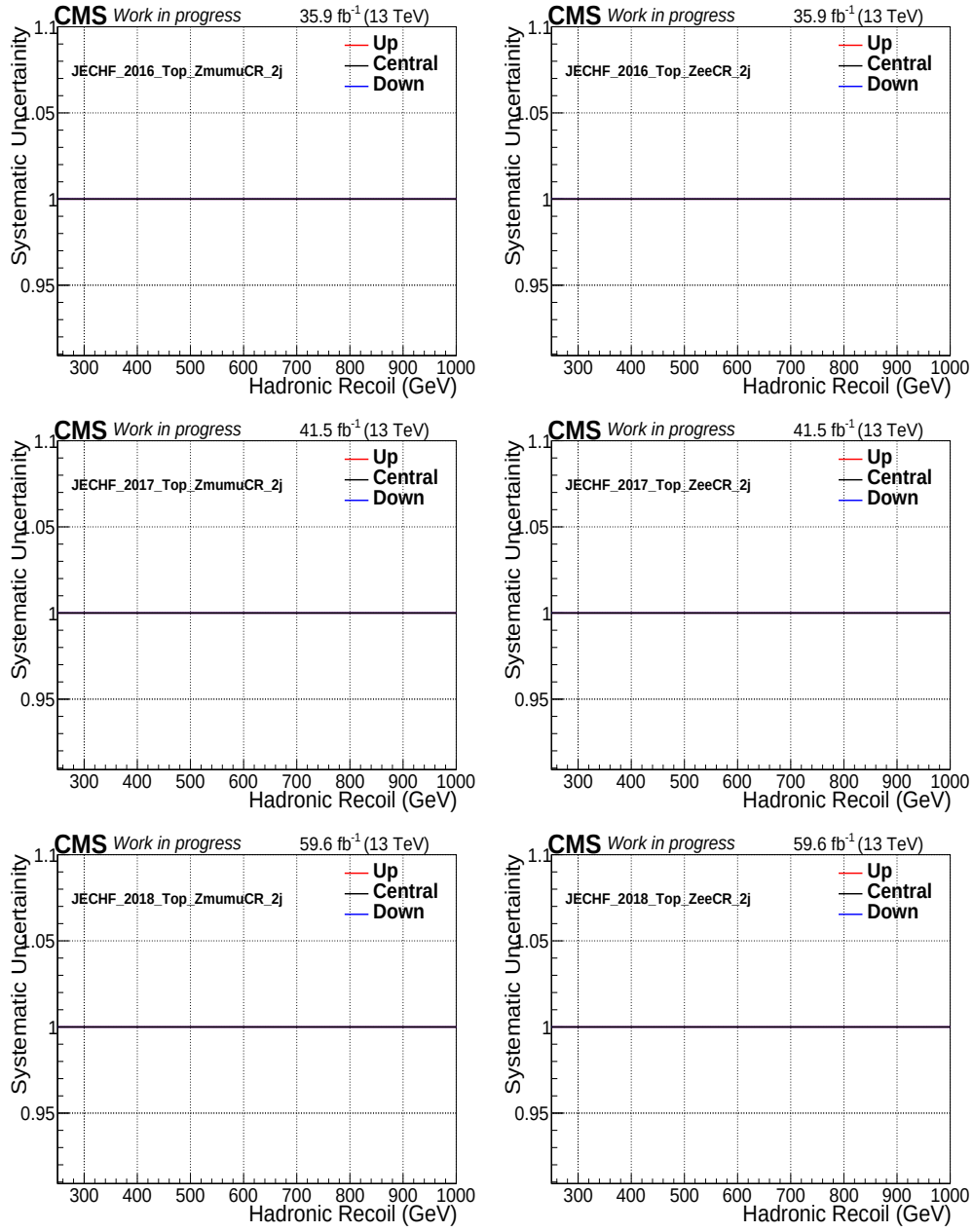


Figure 125: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

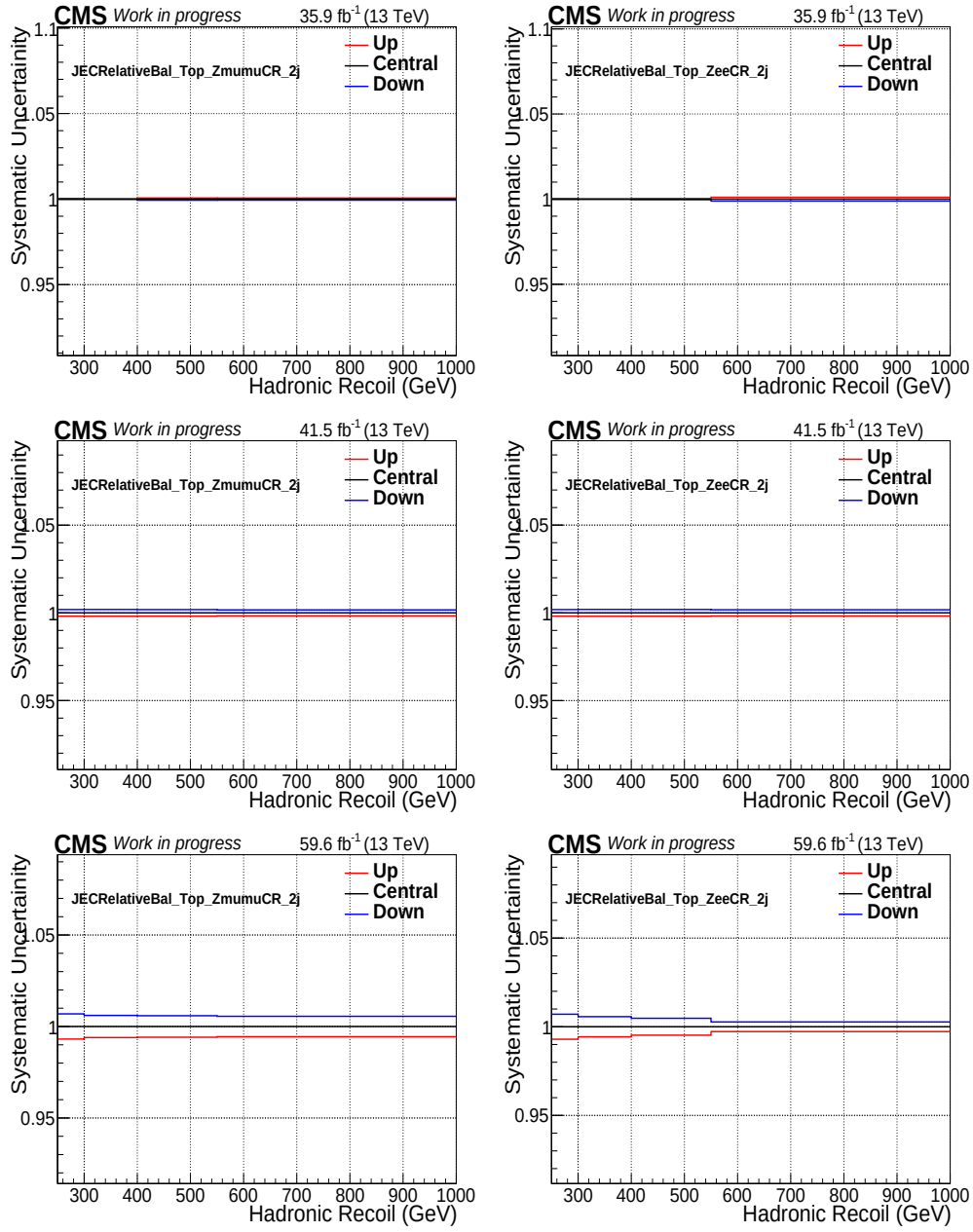


Figure 126: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

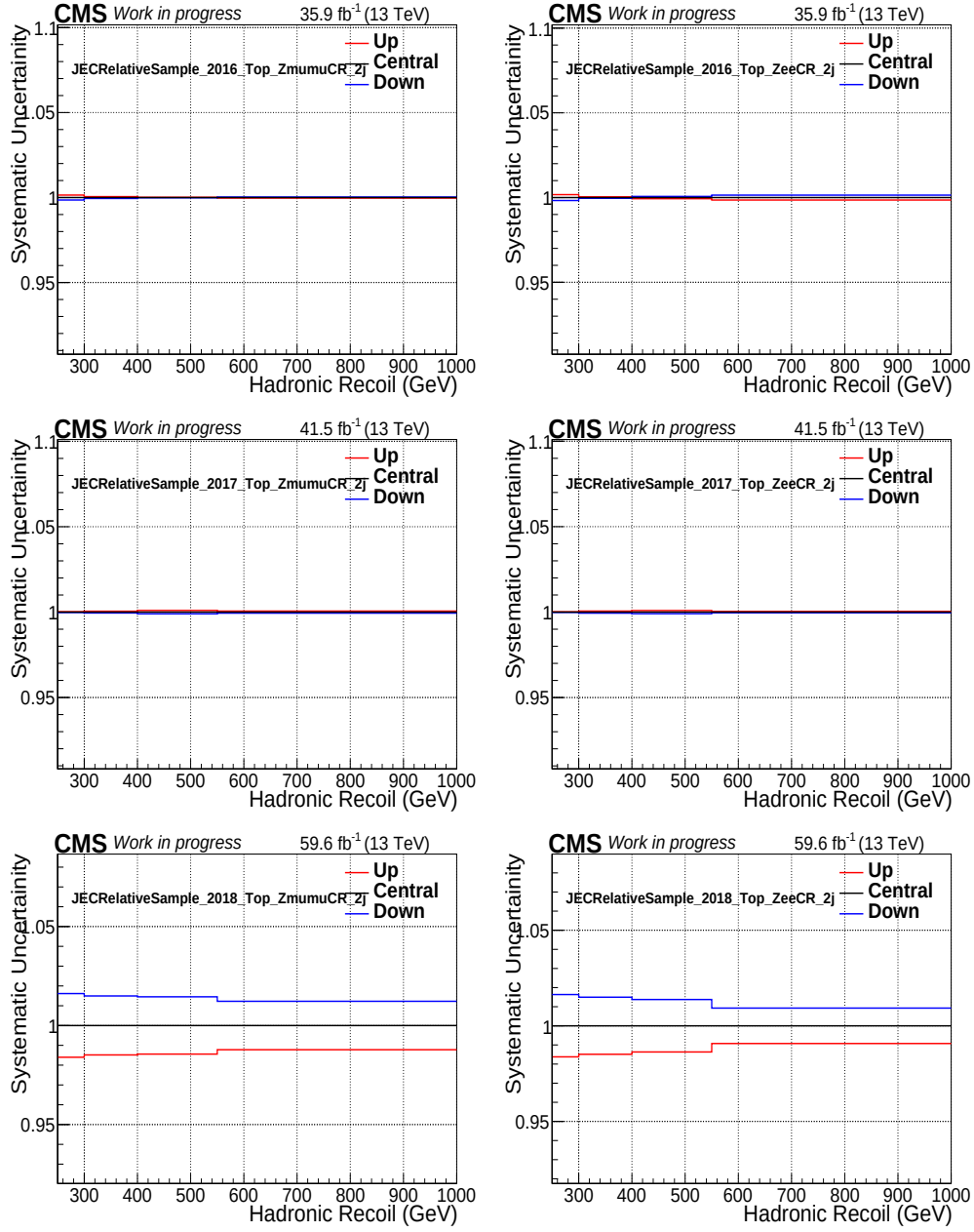


Figure 127: JEC uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

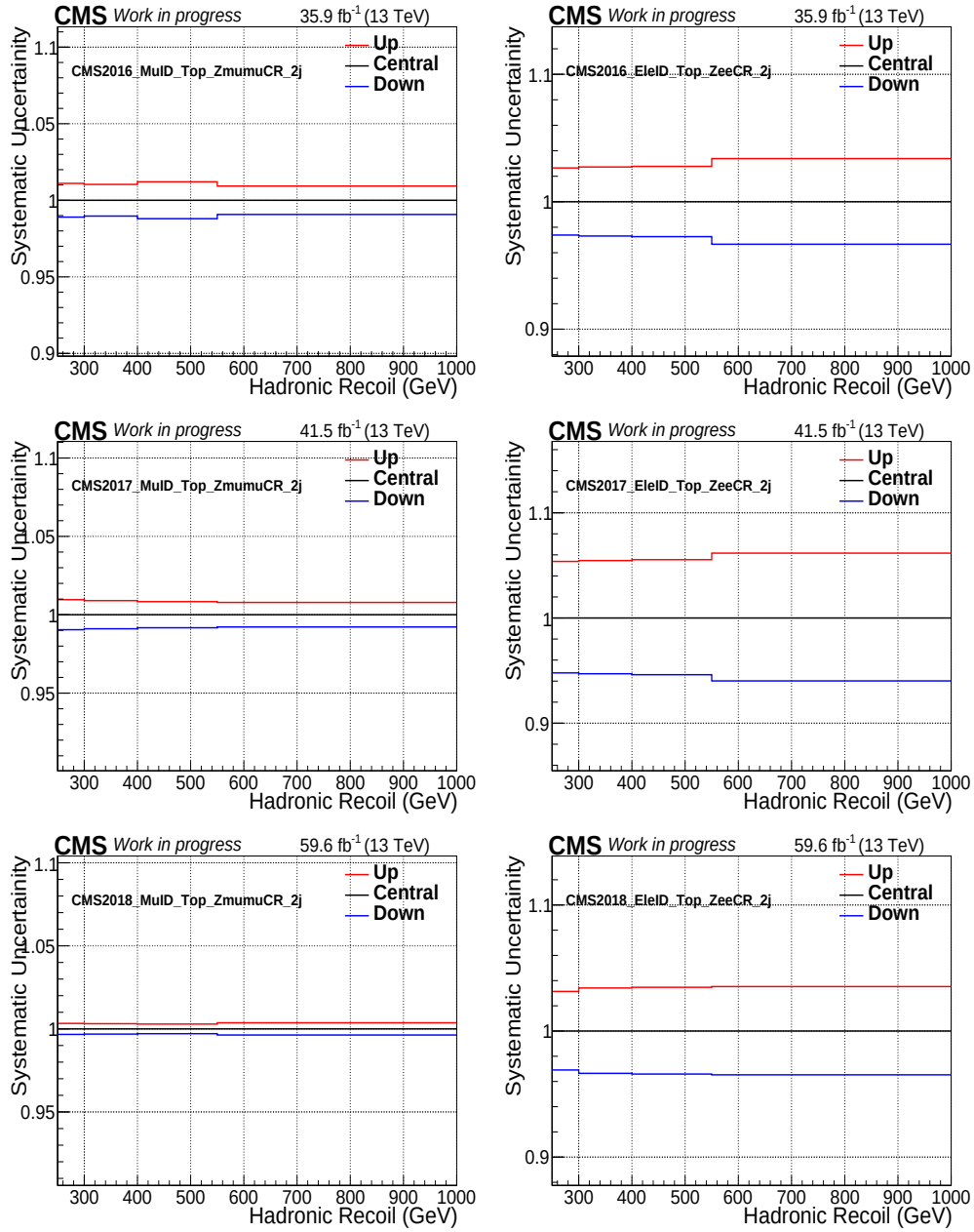


Figure 128: muon ID and electron ID uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

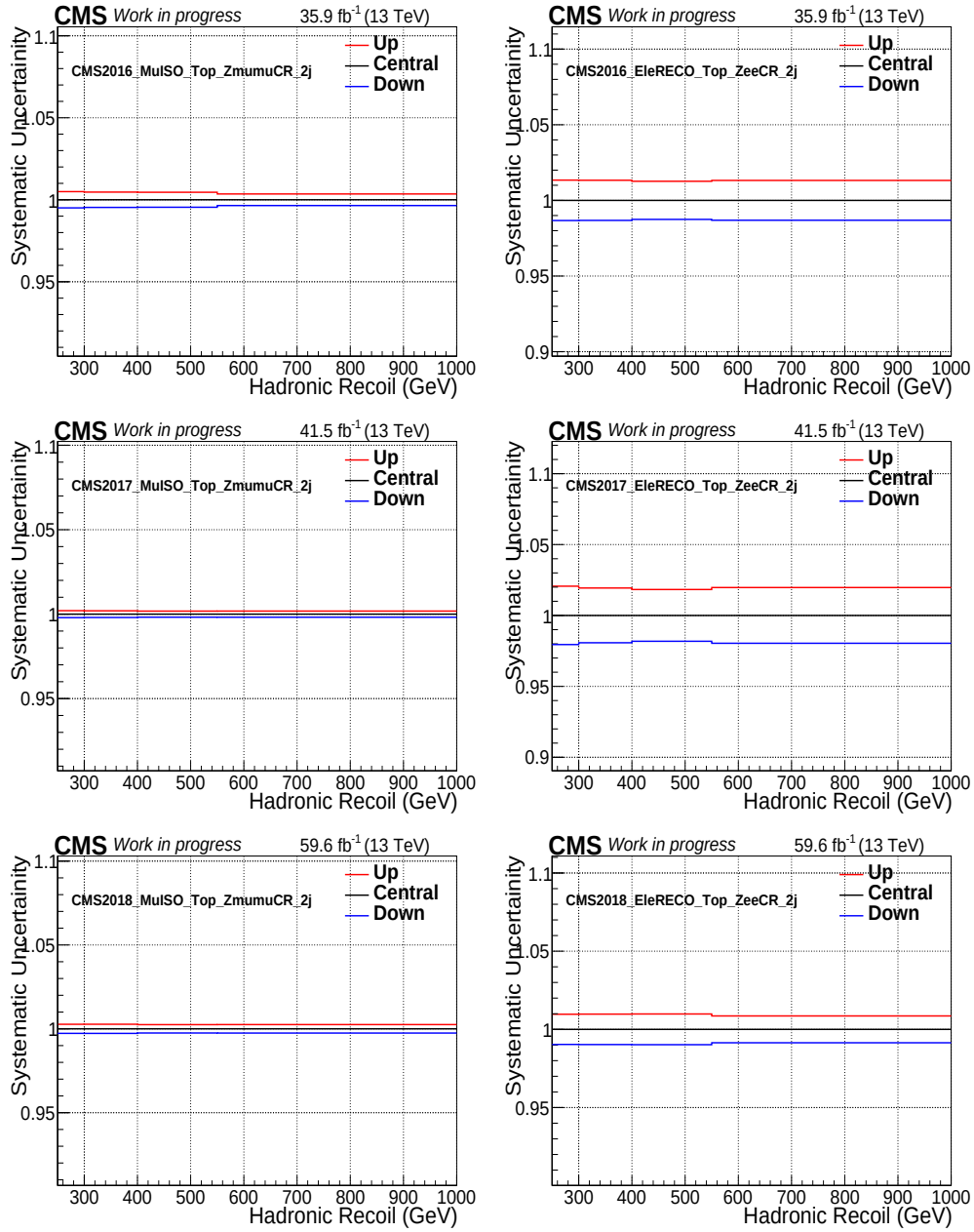


Figure 129: muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 2j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018

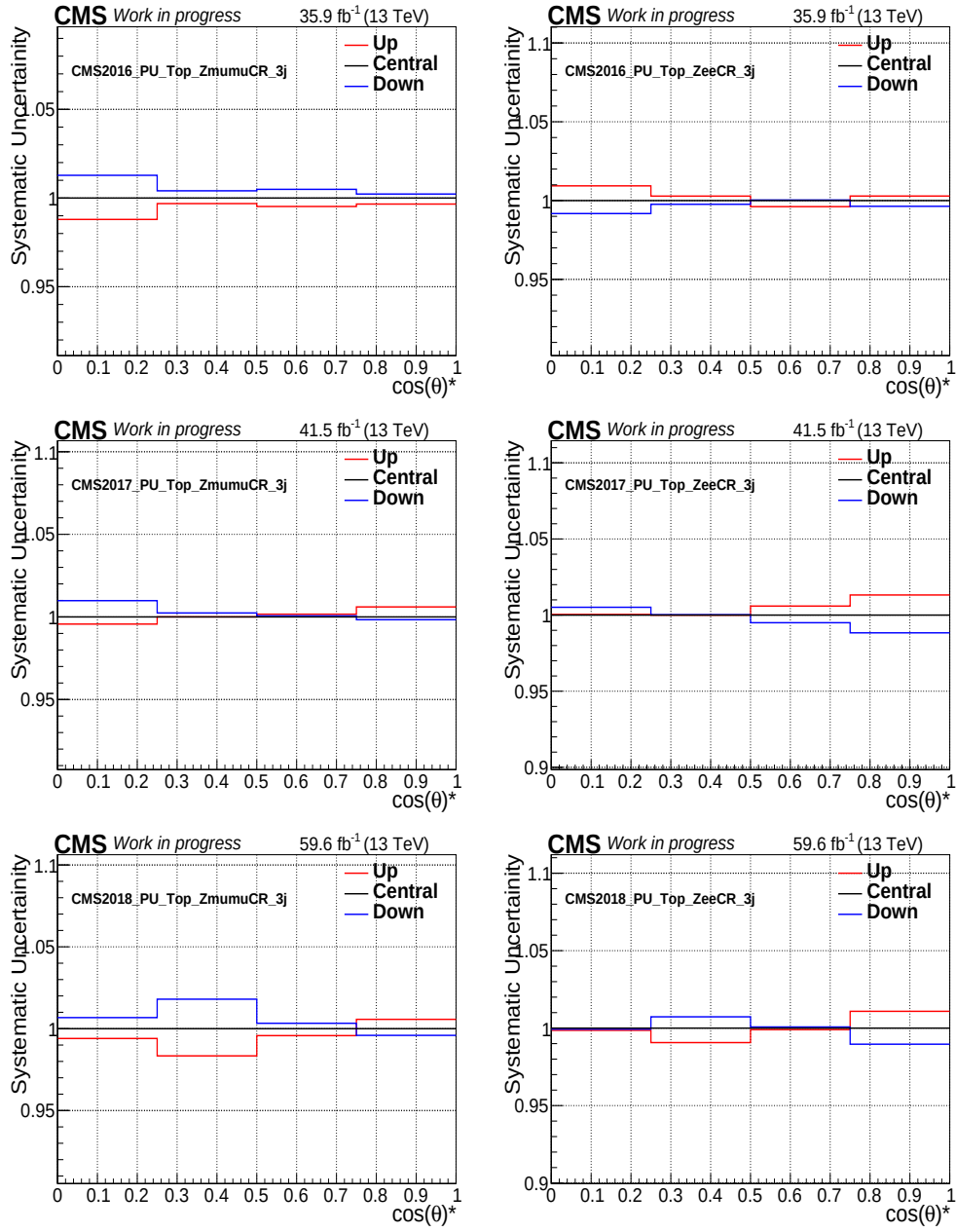


Figure 130: PileUP uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

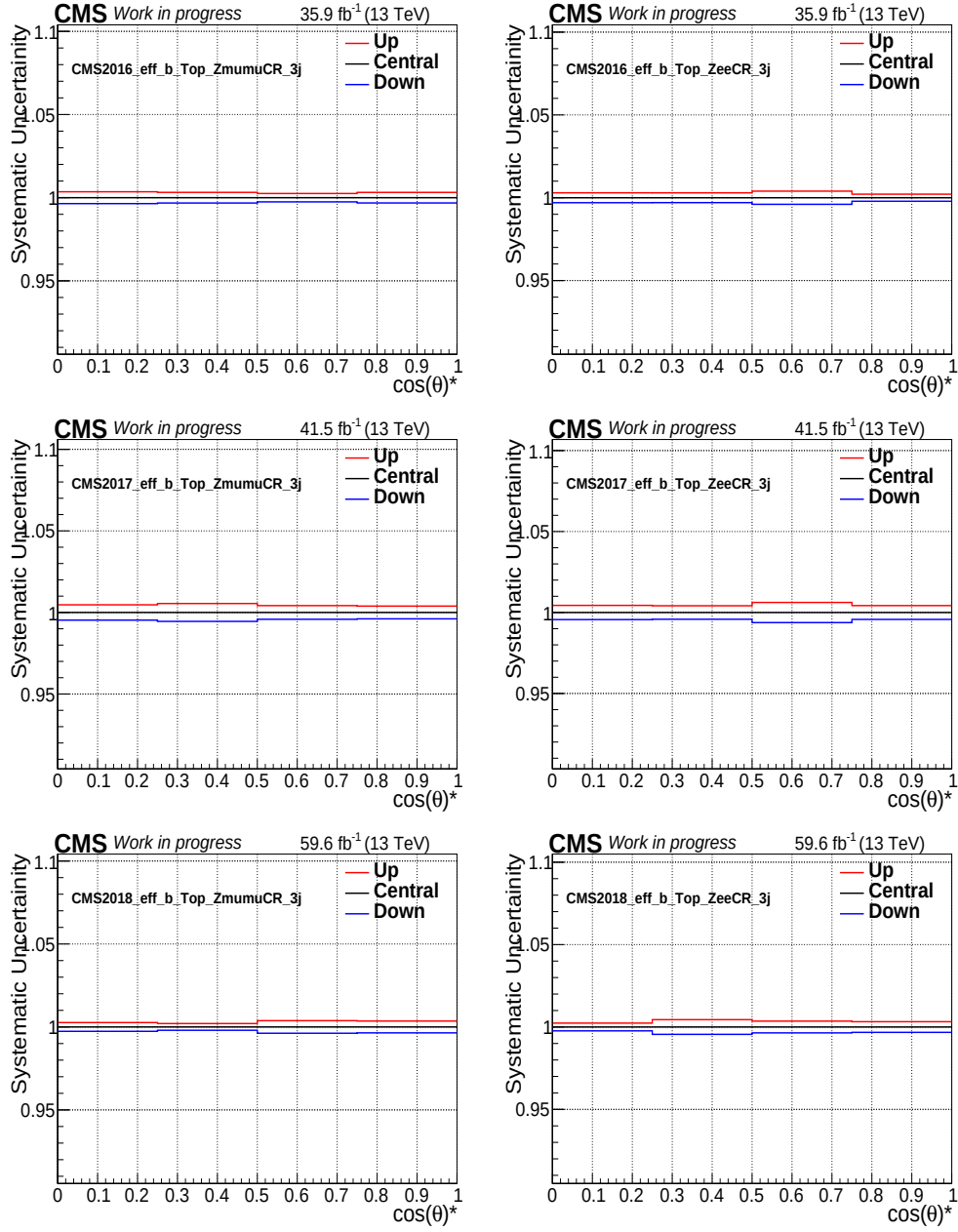


Figure 131: b-tag uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

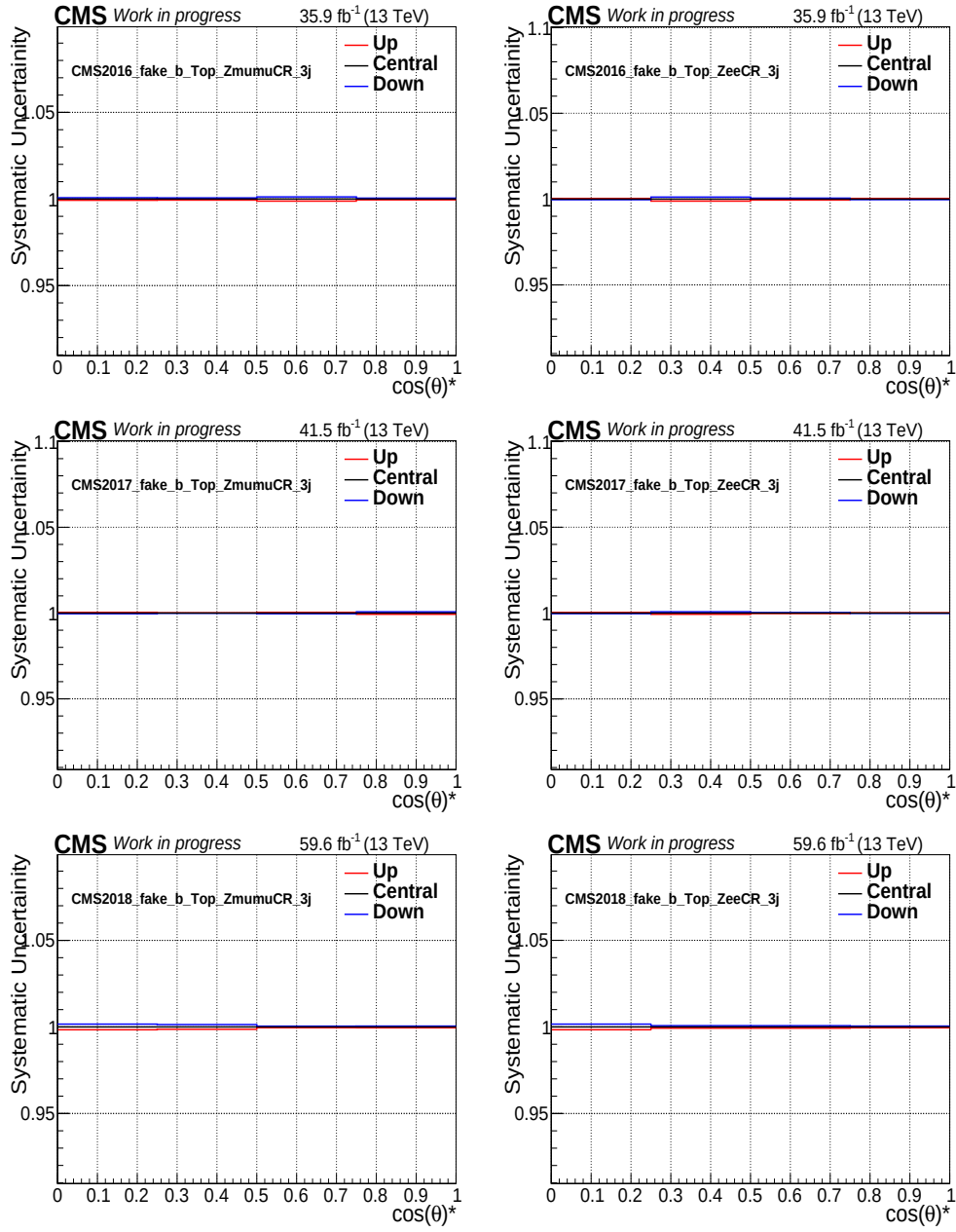


Figure 132: Fake b-tag uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

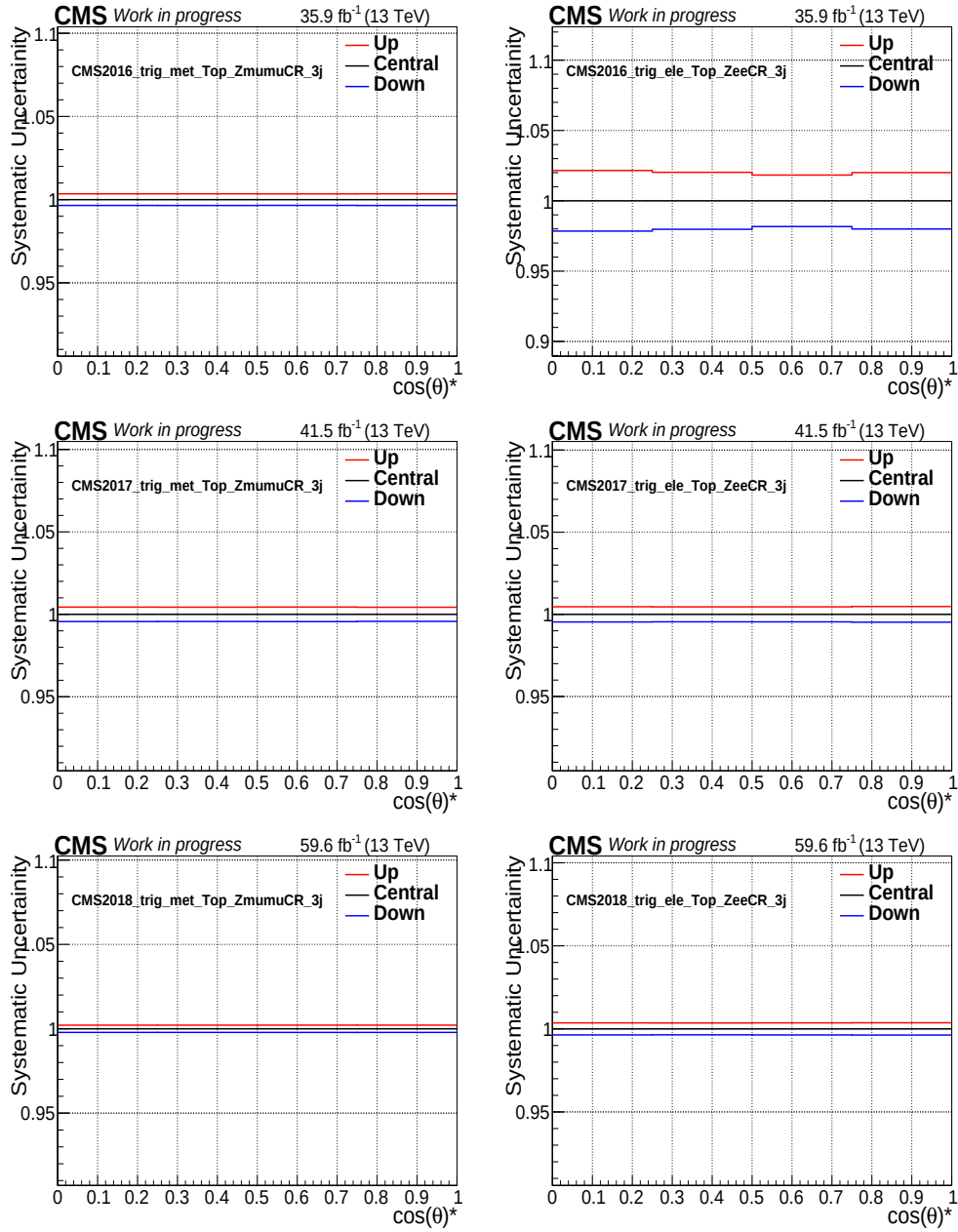


Figure 133: Trigger uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

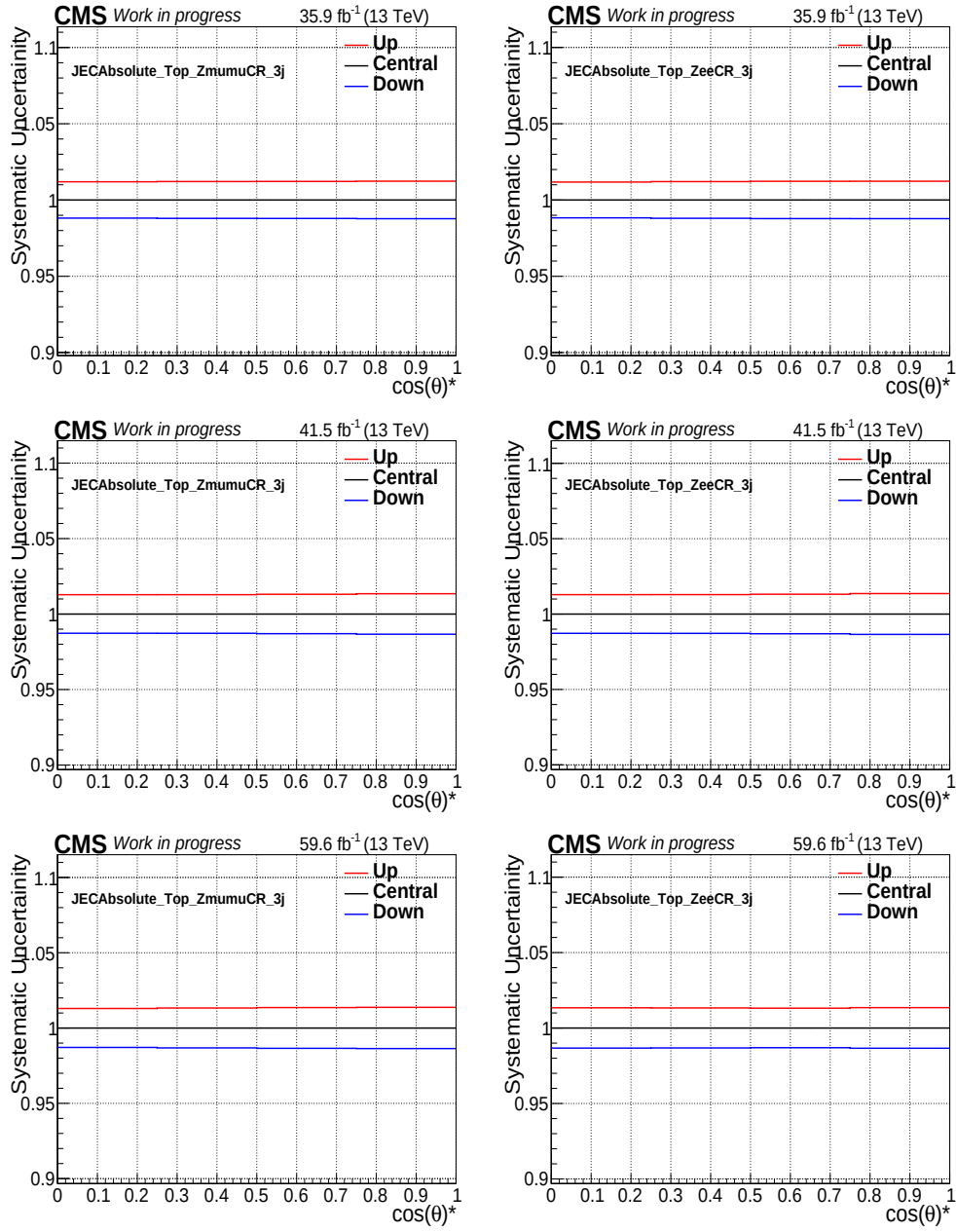


Figure 134: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

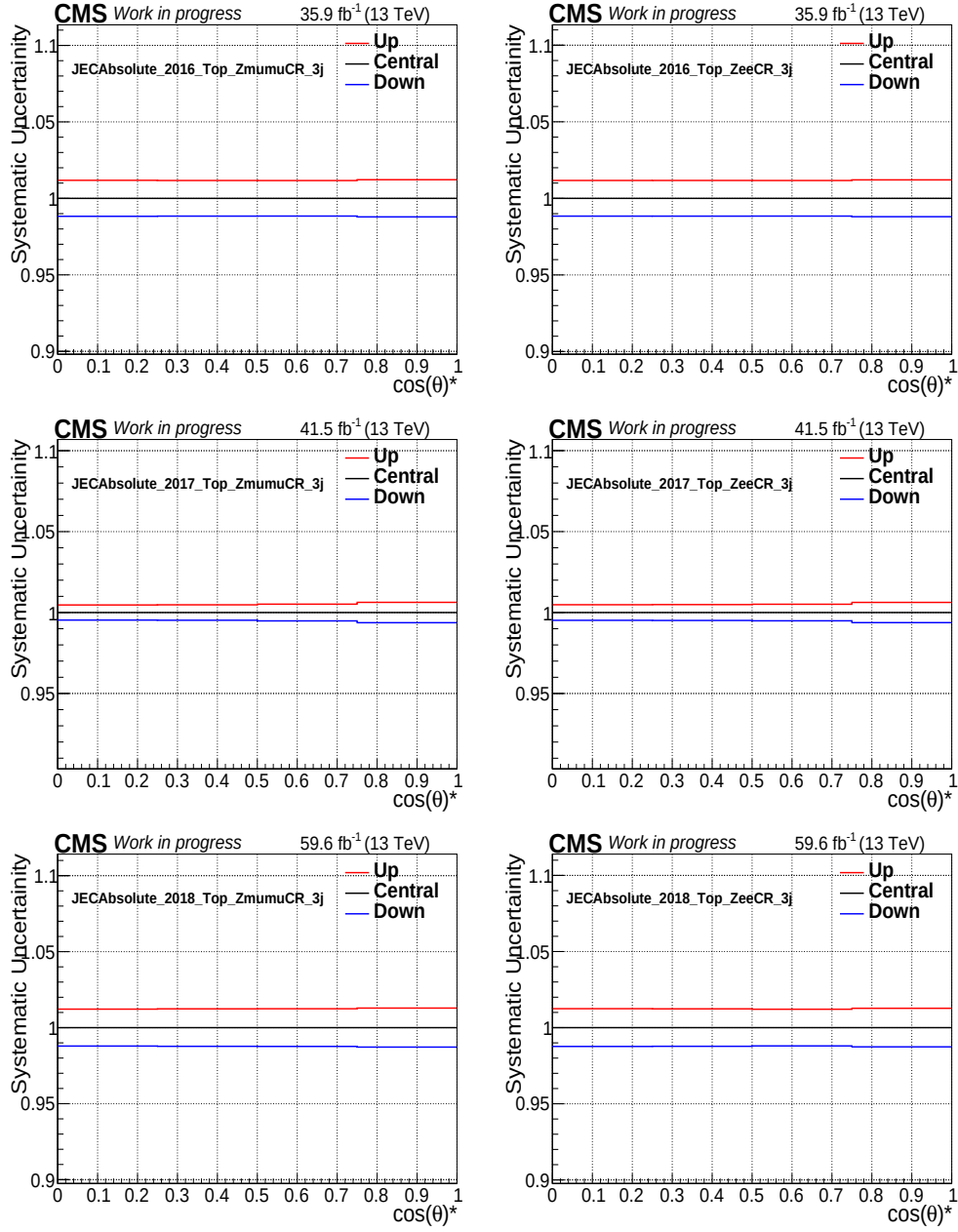


Figure 135: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

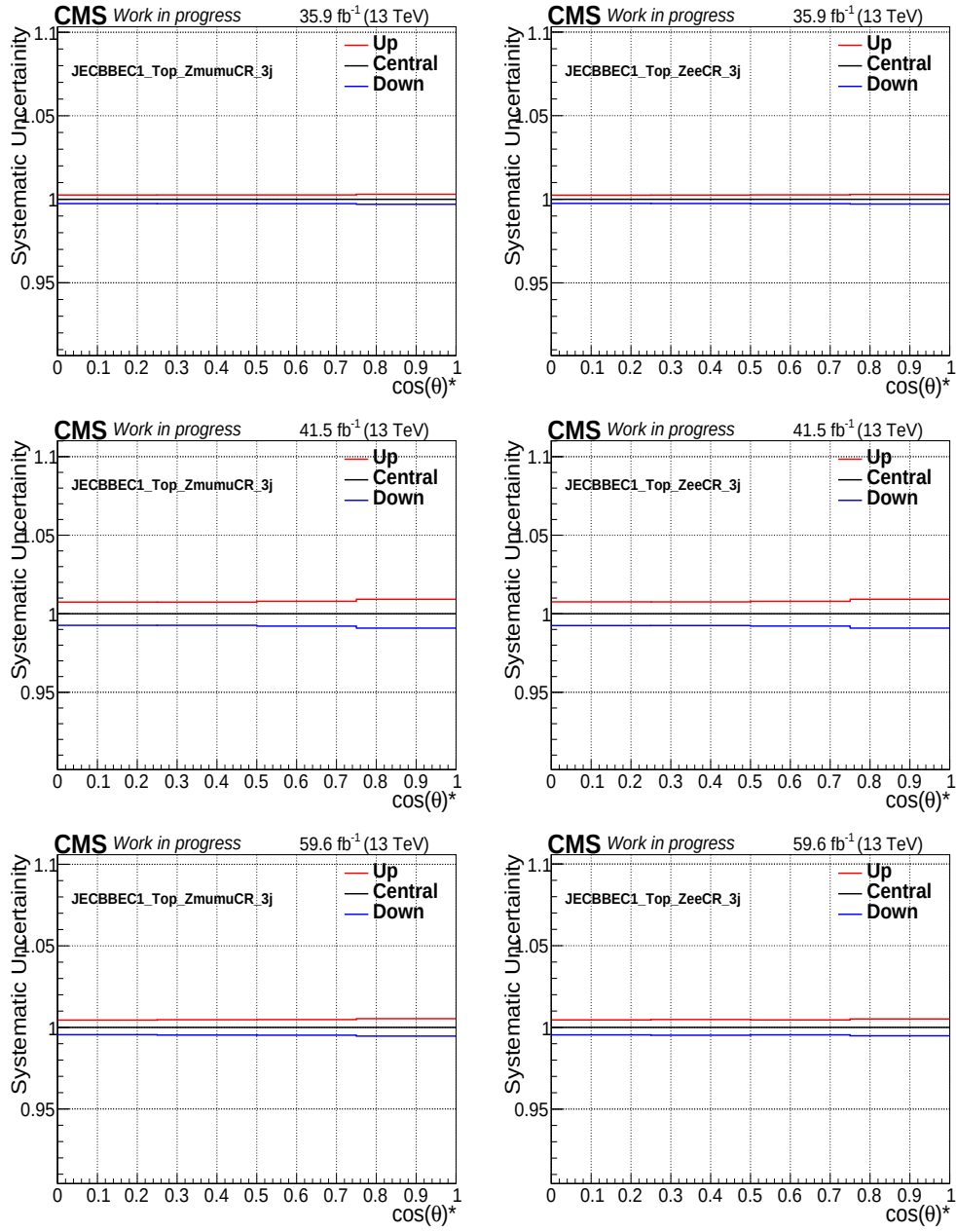


Figure 136: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

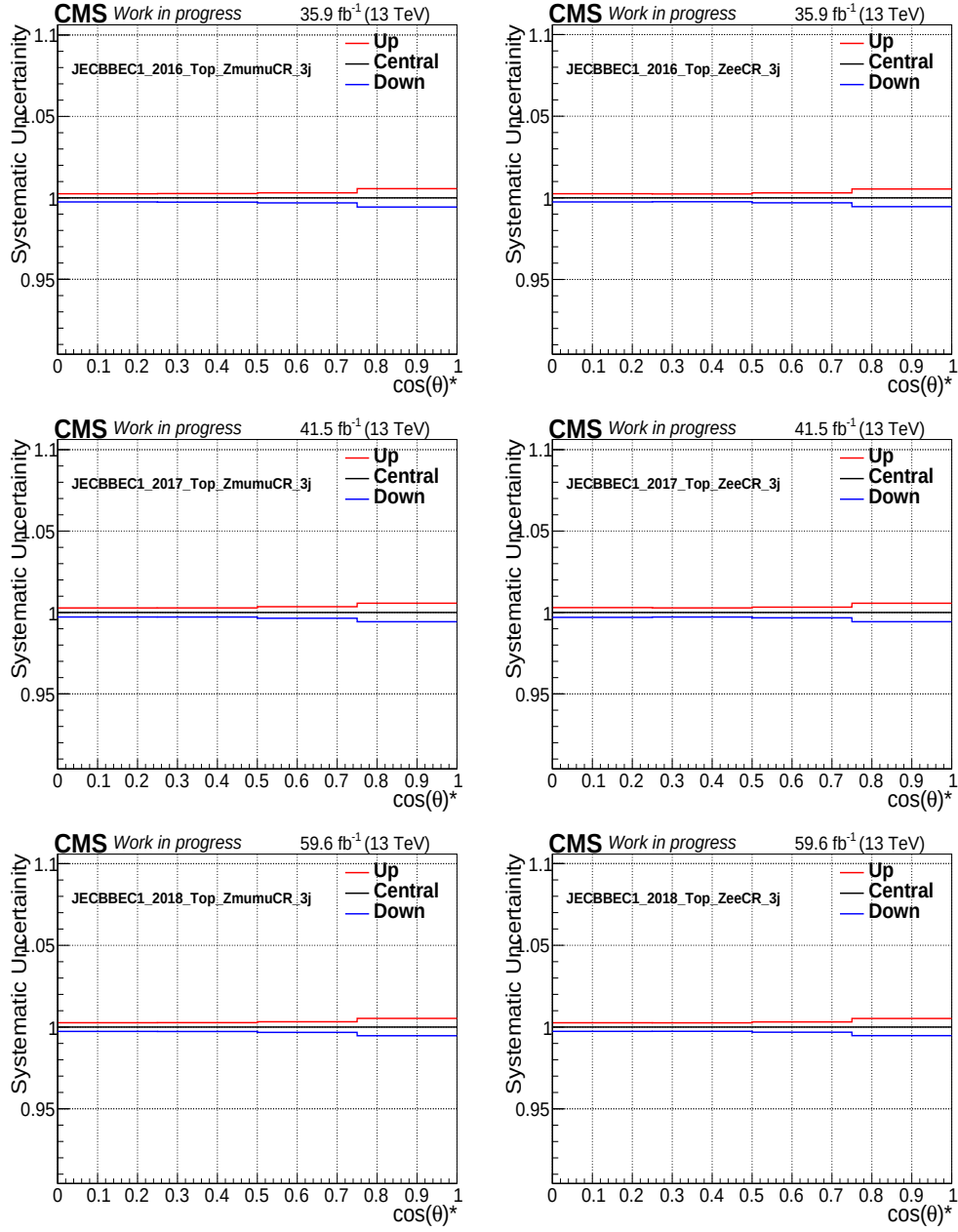


Figure 137: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

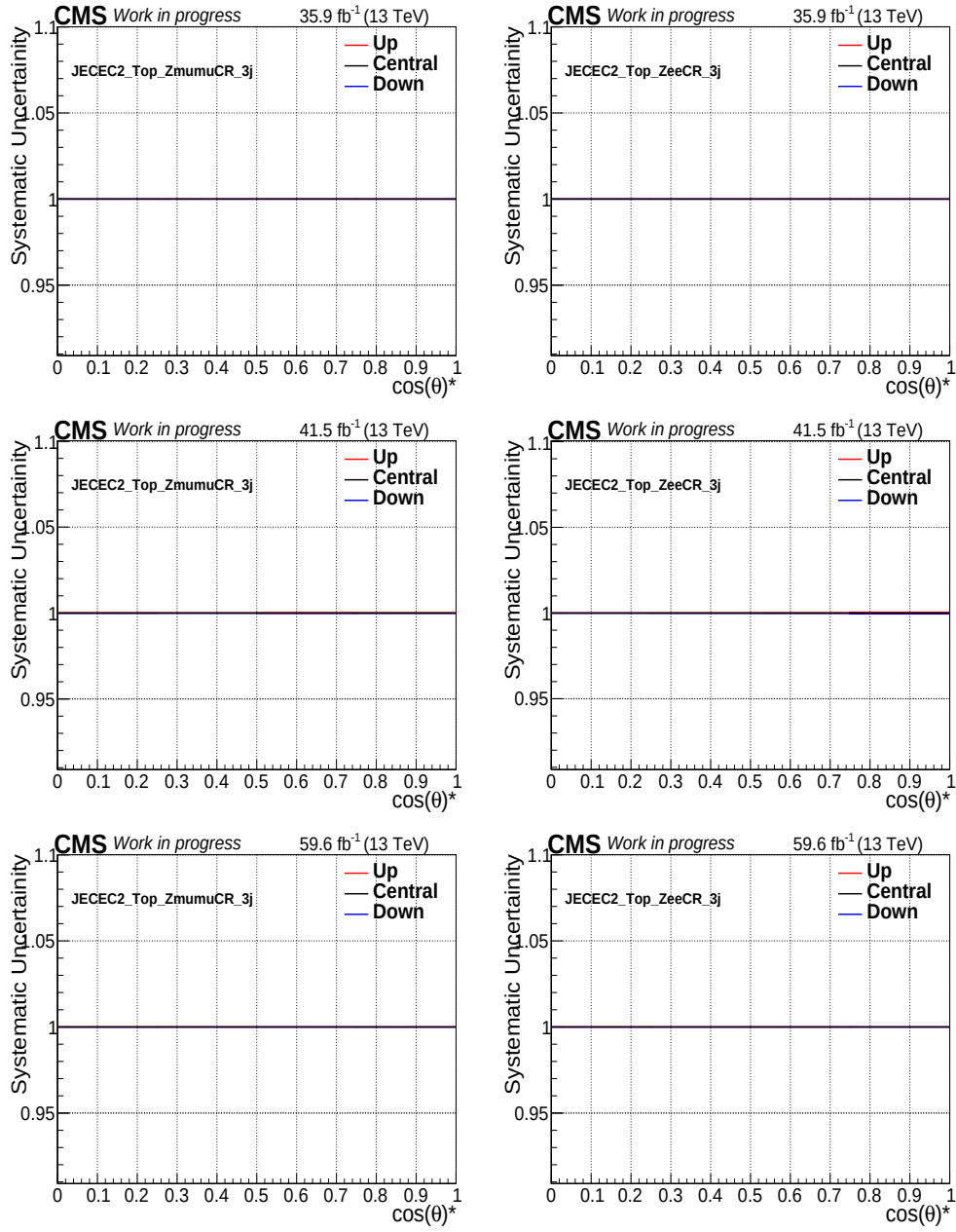


Figure 138: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

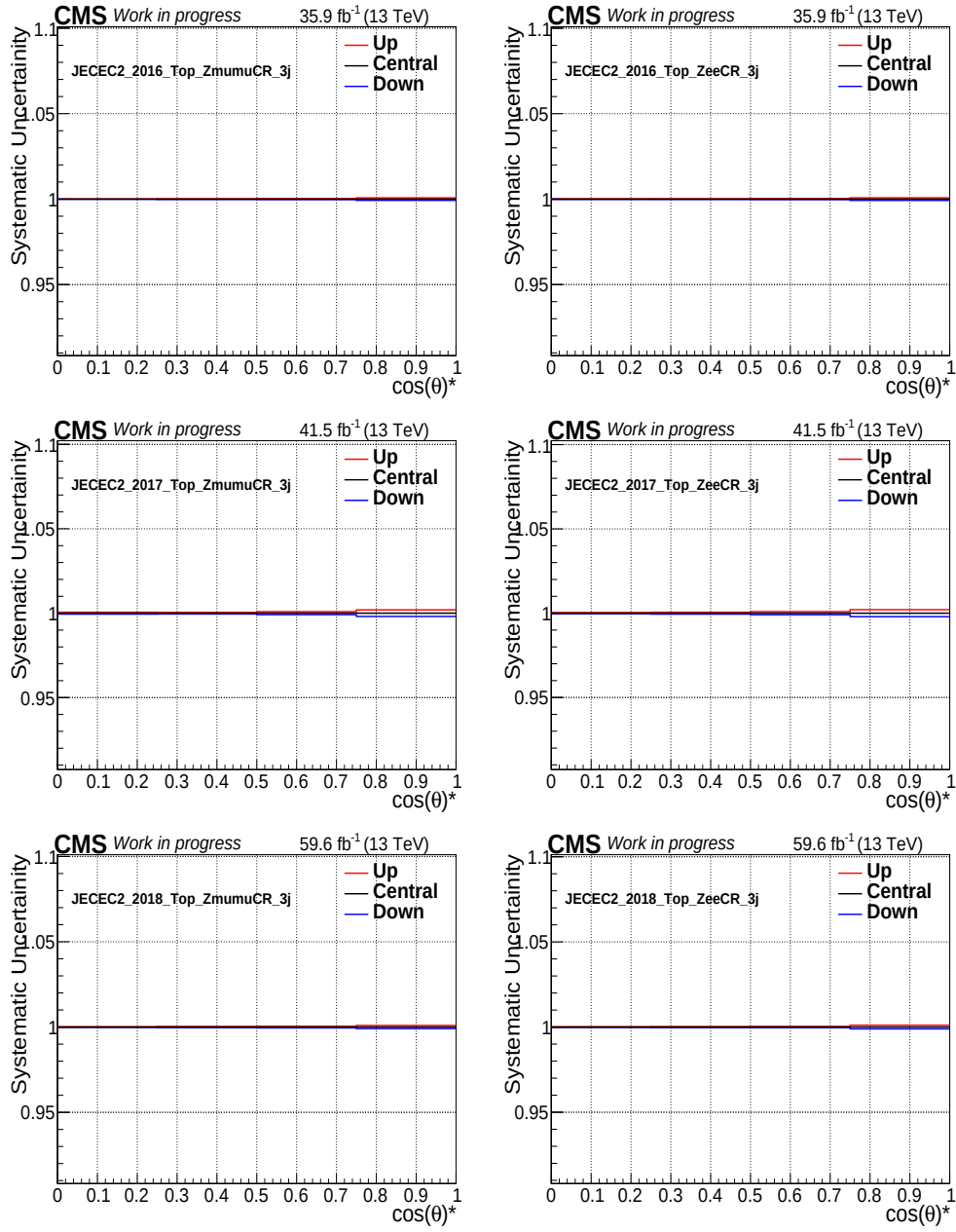


Figure 139: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

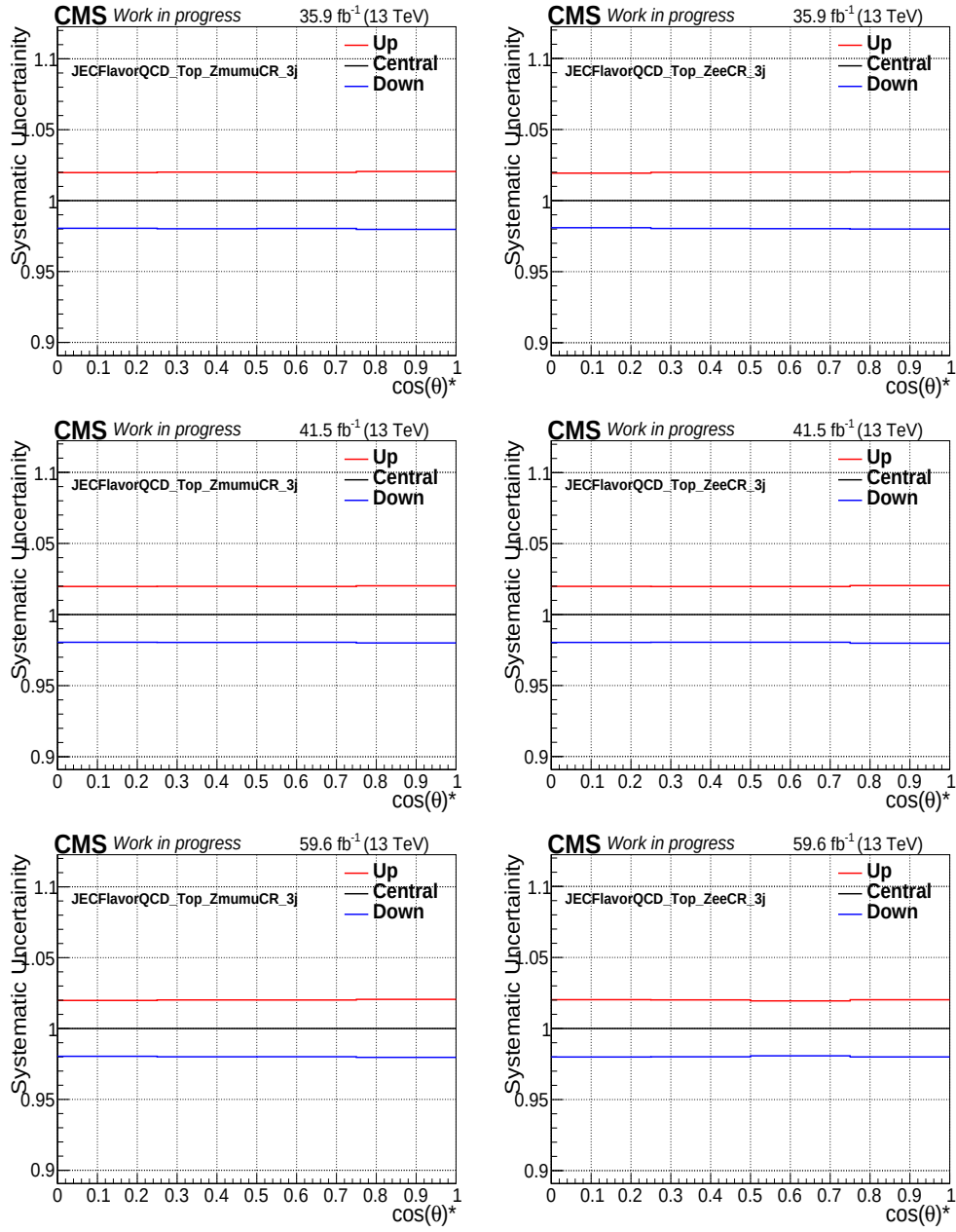


Figure 140: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

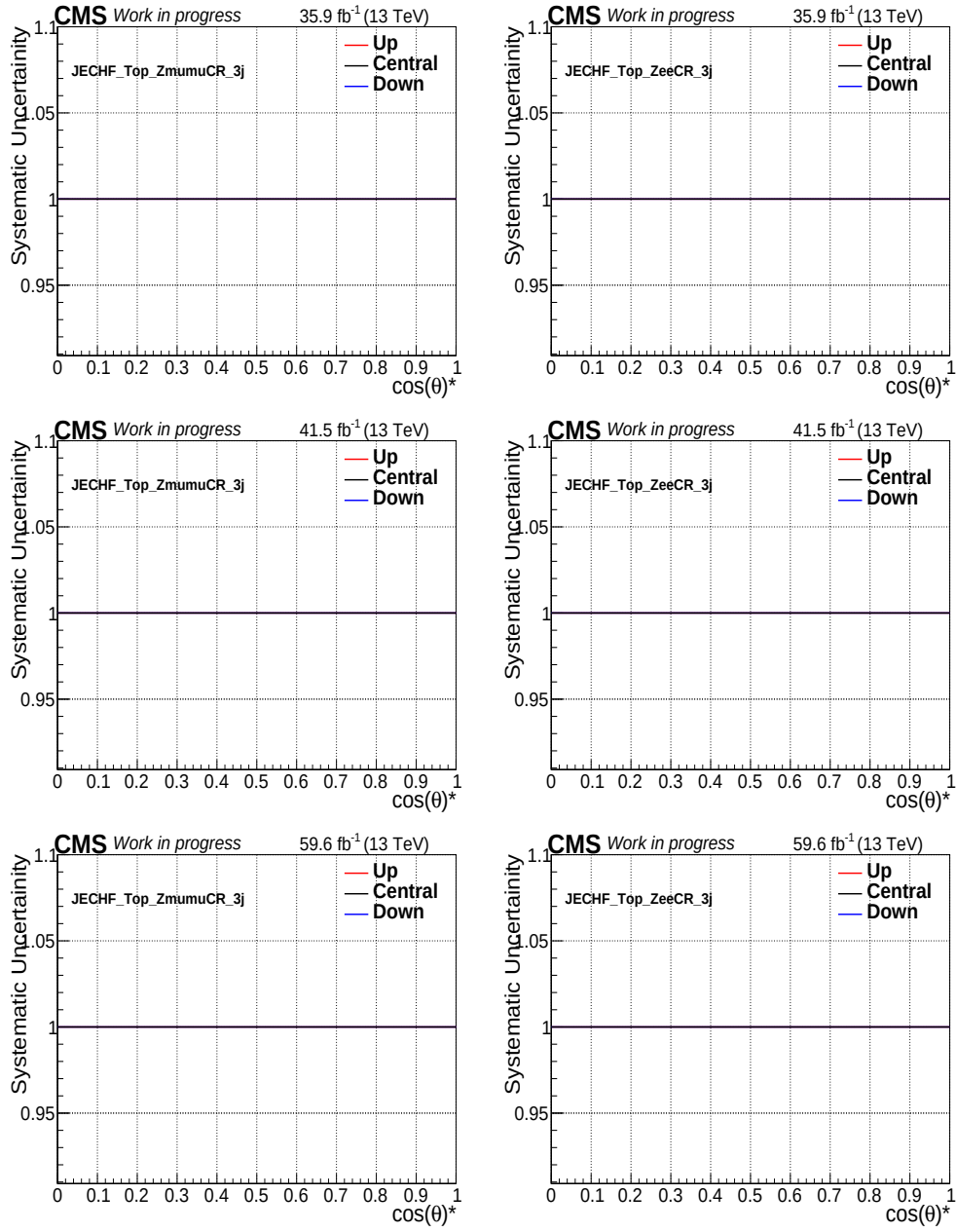


Figure 141: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

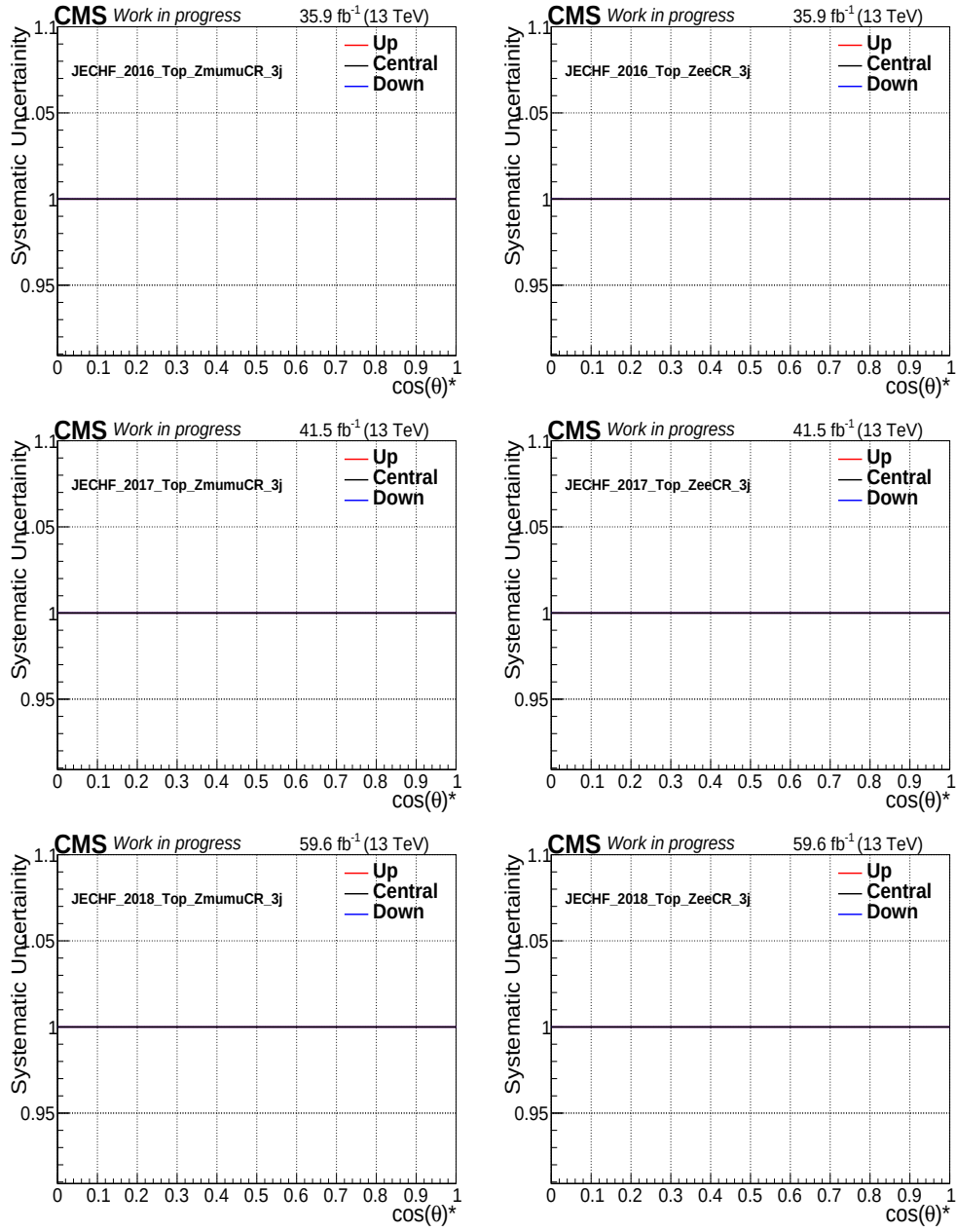


Figure 142: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

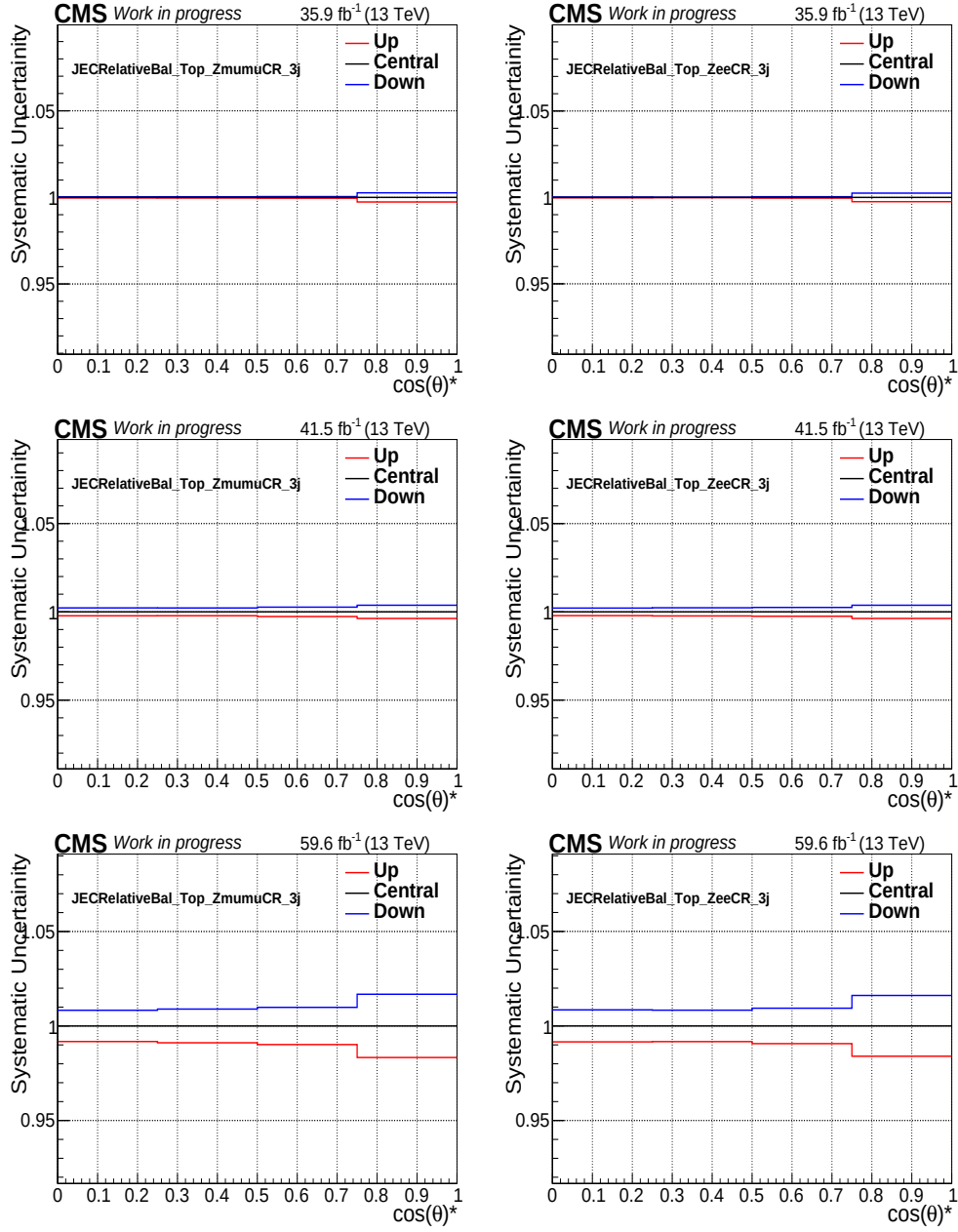


Figure 143: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

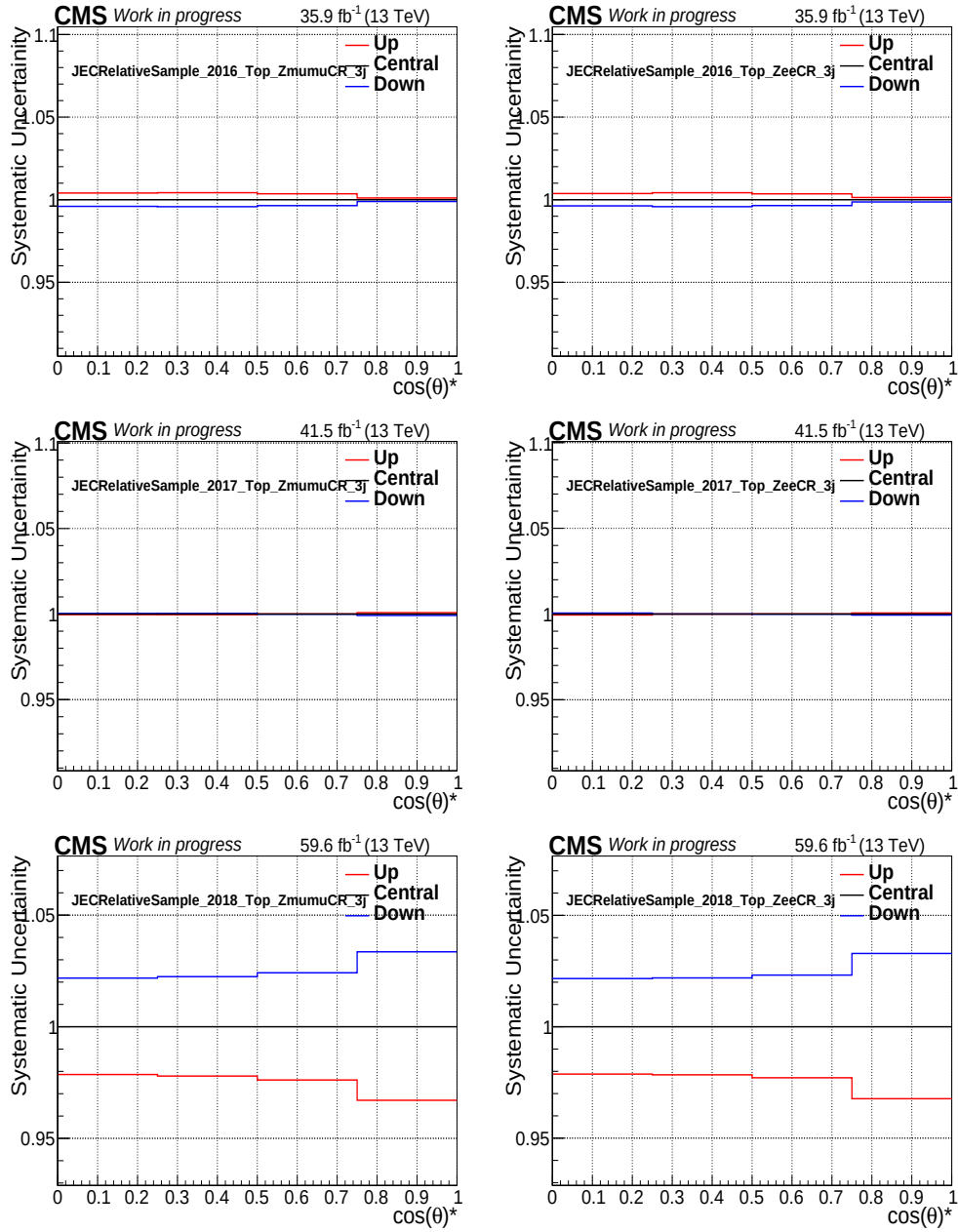


Figure 144: JEC uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ for 2016, 2017 and 2018

D. SYSTEMATIC UNCERTAINTY PLOTS

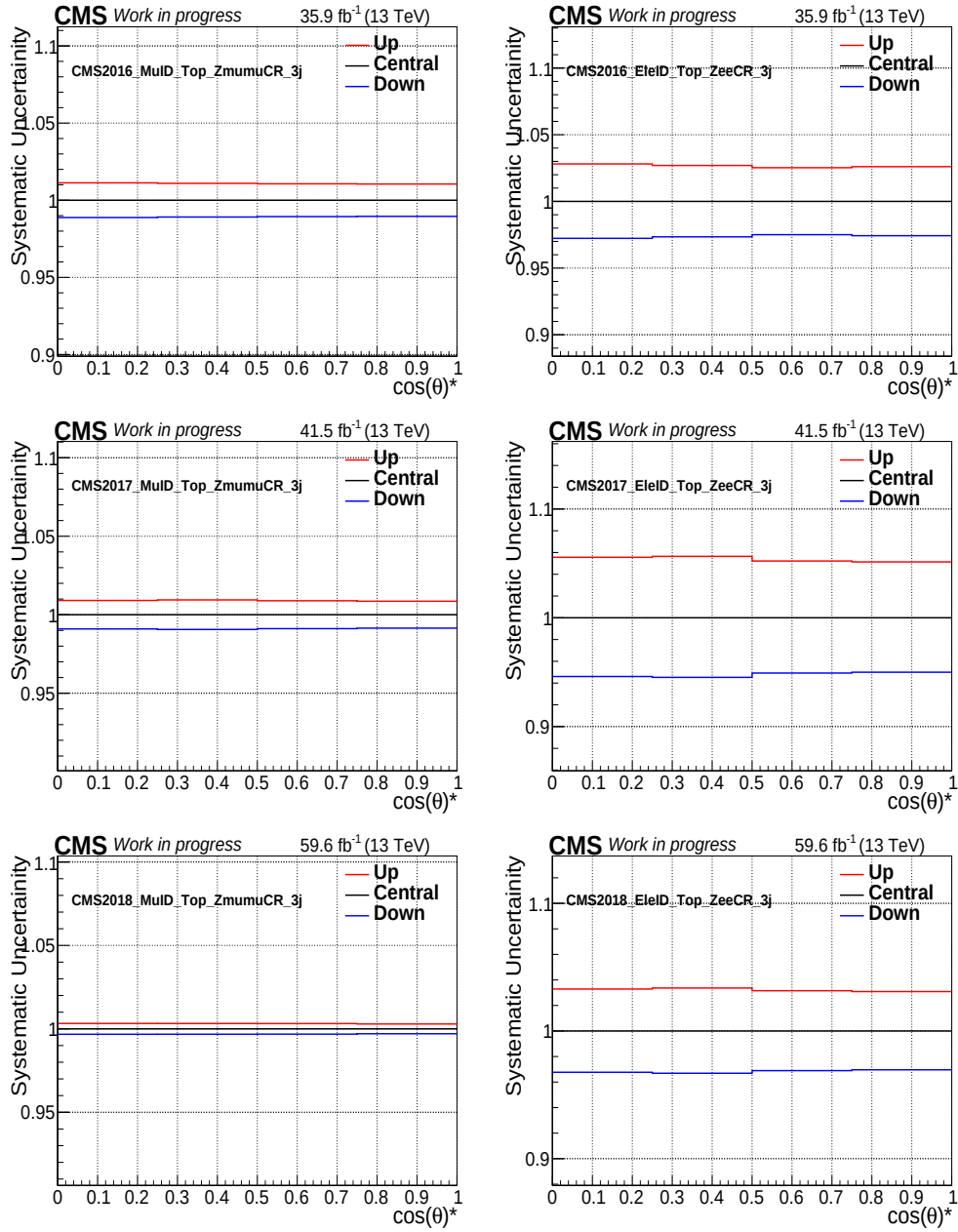


Figure 145: muon ID and electron ID uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018

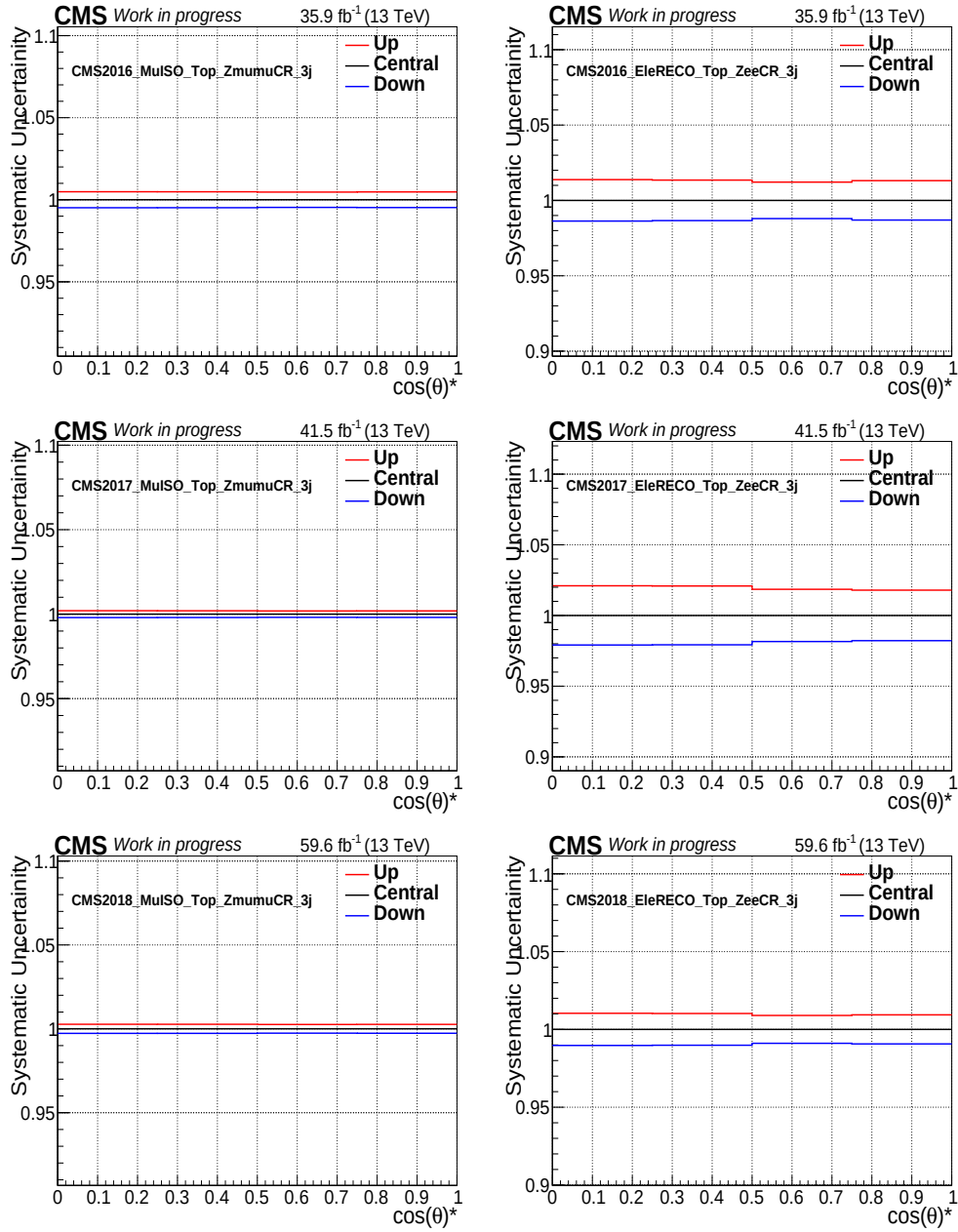


Figure 146: muon ISO and electron RECO uncertainty for $t\bar{t}$ background in 3j Category of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ respectively for 2016, 2017 and 2018

E Contributions at CMS

I have worked in the implementation of DeeCSV b-tagging algorithm in High Level Trigger for CMS, the details of all the steps involved is discussed in section [3.2.6.3.1](#). I have contributed to various service tasks in physics analysis groups (PAGs) and physics object groups (POGs) within the CMS collaboration.

- Service Tasks in CMS
 - MC generator contact for Higgs subgroup: The generation of Monte Carlo Simulations is generally done centrally within the CMS collaboration, I worked on the generation of these samples for analysis that come under Higgs PAG.
 - Jet Energy Correction(JEC): Worked in JetMET POG to implement the JECs as a monitoring tool at the HLT, using Z+Jet events to calibrate the Jet and calculate the jet response defined as the ratio of jet and Z boson momentum (Jet_{p_T}/Z_{p_T})
 - L1 Trigger Expert Analyst(TEA): I have also worked as L1 TEA which has following roles:
 - * CMS software release validation
 - Every CMS software release is validated first with existing release to check for inconsistencies. As L1 TEA we validated the CMS software release for L1 part of the Trigger
 - * To select the good runs to be used for physics analyses

-
- `btv@HLT` subgroup: `btv@HLT` is a subgroup of BTV POG of CMS Collaboration which deals with b-tagging and vertexing at HLT level where I have been a convenor for the duration of 1 year

- Conferences and Publications

- Tiwari, Praveen Chandra and On behalf Of the Cms collaboration. "Proceedings of The 39th International Conference on High Energy Physics — PoS(ICHEP2018)." Heavy flavour identification at CMS. SISSA Medialab, 2 Aug. 2019, p. 898, doi:10.22323/1.340.0898.
- Search for dark matter produced in association with bottom quarks at 13 TeV with CMS detector at DAE-HEP 2020: 24th DAE-BRNS High Energy Physics Symposium 2020, 14-18 Dec 2020, Bhubaneswar (India)
- Presented 'Dark Matter searches in CMS' on behalf of CMS collaboration, Pheno2021: Phenomenology 2021 Symposium, 24-26 May 2021, University of Pittsburgh.
- CMS author since 2017 ([Inspire Link](#))

BIBLIOGRAPHY

- [1] Taking a closer look at LHC - LHC parameters, Jun 2021. [Online;]. URL: http://lh-closer.es/taking_a_closer_look_at_lhc/1.lhc_parameters.
- [2] What Is the Universe Really Made Of? - OpenStax CNX, Nov 2020. [Online;]. URL: <https://cnx.org/contents/aiQnHGyf@10/What-Is-the-Universe-Really-Made-Of>.
- [3] Standard Model of Particle Physics [The Physics Travel Guide], Jun 2021. [Online;]. URL: https://physicstravelguide.com/models/standard_model.
- [4] Particle Data Group, P. A. Zyla, R. M. Barnett, J. Beringer, et al. Review of Particle Physics. *Prog Theor Exp Phys*, 2020(8), Aug 2020. doi:10.1093/ptep/ptaa104.
- [5] Stefano Giagu. WIMP Dark Matter Searches With the ATLAS Detector at the LHC. *Front Phys*, 7, May 2019. doi:10.3389/fphy.2019.00075.
- [6] The HL-LHC project | High Luminosity LHC Project, May 2021. [Online;]. URL: <https://hilumilhc.web.cern.ch/content/hl-lhc-project>.

BIBLIOGRAPHY

- [7] Eva Halkiadakis. INTRODUCTION TO THE LHC EXPERIMENTS. In *Physics of the Large and the Small*, pages 489–518. WORLD SCIENTIFIC, Singapore, Mar 2011. doi: 10.1142/9789814327183_0009.
- [8] CMS Detector Design, Jun 2021. [Online;]. URL: <https://cms.cern/book/export/html/1190>.
- [9] David Barney. CMS Detector Slice. Jan 2016. URL: <https://cds.cern.ch/record/2120661/?ln=en>.
- [10] The CMS Collaboration, S. Chatrchyan, G. Hmayakyan, V. Khachatryan, A. M. Sirunyan, et al. The CMS experiment at the CERN LHC. *J Instrum*, 3(08):S08004–S08004, Aug 2008. doi:10.1088/1748-0221/3/08/s08004.
- [11] Prasanna Kumar Siddireddy. The CMS ECAL Trigger and DAQ system: electronics auto-recovery and monitoring. Technical report, Jun 2018. 7 pages, 6 figures, 1 table, conference record. URL: <https://cds.cern.ch/record/2630618>, arXiv: 1806.09136.
- [12] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, et al. The CMS trigger system. 12(01):P01020–P01020, Jan 2017. doi:10.1088/1748-0221/12/01/p01020.
- [13] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambroggi, E. Asilar, et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. 13(05):P05011–P05011, May 2018. doi:10.1088/1748-0221/13/05/p05011.
- [14] Xl. cathode rays - by j.j. thomson. 87(5):303–326. arXiv:<https://doi.org/10.1080/09500830701306165>, doi:10.1080/09500830701306165.

- [15] J. J. Thomson. XXIV. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 7(39):237–265, Mar 1904. doi:[10.1080/14786440409463107](https://doi.org/10.1080/14786440409463107).
- [16] Professor E. Rutherford F.R.S. Lxxix. the scattering of α and β particles by matter and the structure of the atom. 21(125):669–688. arXiv:<https://doi.org/10.1080/14786440508637080>, doi:[10.1080/14786440508637080](https://doi.org/10.1080/14786440508637080).
- [17] E. RUTHERFORD. The structure of the atom. 92(2302):423–423. doi:[10.1038/092423a0](https://doi.org/10.1038/092423a0).
- [18] J. Chadwick. Possible existence of a neutron. 129(3252):312–312. doi:[10.1038/129312a0](https://doi.org/10.1038/129312a0).
- [19] Jerome I. Friedman and Henry W. Kendall. Deep inelastic electron scattering. 22:203–254. doi:[10.1146/annurev.ns.22.120172.001223](https://doi.org/10.1146/annurev.ns.22.120172.001223).
- [20] Lyndon Evans and Philip Bryant. LHC machine. 3(08):S08001–S08001. URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001>, doi:[10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001).
- [21] M. Tanabashi et al. Review of particle physics. 98:030001. URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>, doi:[10.1103/PhysRevD.98.030001](https://doi.org/10.1103/PhysRevD.98.030001).
- [22] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lh. 716(1):1–29. URL: <https://arxiv.org/abs/1207.3447>.

BIBLIOGRAPHY

- [//www.sciencedirect.com/science/article/pii/S037026931200857X](http://www.sciencedirect.com/science/article/pii/S037026931200857X),
doi:<https://doi.org/10.1016/j.physletb.2012.08.020>.
- [23] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. 716(1):30–61. URL: <https://www.sciencedirect.com/science/article/pii/S0370269312008581>,
doi:<https://doi.org/10.1016/j.physletb.2012.08.021>.
- [24] Edvige Corbelli and Paolo Salucci. The extended rotation curve and the dark matter halo of M33. 311(2):441–447. arXiv:<https://academic.oup.com/mnras/article-pdf/311/2/441/2881340/311-2-441.pdf>, doi:10.1046/j.1365-8711.2000.03075.x.
- [25] Sean Carroll. Dark matter is real. 2(10):653–654. doi:10.1038/nphys428.
- [26] M. J. Jee, H. C. Ford, G. D. Illingworth, R. L. White, T. J. Broadhurst, D. A. Coe, G. R. Meurer, A. van der Wel, N. Benitez, J. P. Blakeslee, and et al. Discovery of a ringlike dark matter structure in the core of the galaxy cluster cl 0024+17. 661(2):728–749. URL: <http://dx.doi.org/10.1086/517498>, doi:10.1086/517498.
- [27] Antonio Pich. The Standard Model of Electroweak Interactions. In *2010 European School of High Energy Physics*, 1 2012. arXiv:1201.0537.
- [28] Richard H. Parker, Chenghui Yu, Weicheng Zhong, Brian Estey, and Holger Müller. Measurement of the fine-structure constant as a test of the standard model. *Science*, 360(6385):191–195, Apr 2018. URL: <http://dx.doi.org/10.1126/science.aap7706>, doi:10.1126/science.aap7706.

- [29] J. M. Butterworth. The standard model: how far can it go and how can we tell? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2075):20150260, 2016. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2015.0260>, arXiv:<https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2015.0260>, doi:10.1098/rsta.2015.0260.
- [30] Michael Dine and Alexander Kusenko. The Origin of the matter - antimatter asymmetry. *Rev. Mod. Phys.*, 76:1, 2003. arXiv:hep-ph/0303065, doi:10.1103/RevModPhys.76.1.
- [31] Super-Kamiokande Collaboration. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81(8):1562–1567, Aug 1998. doi:10.1103/PhysRevLett.81.1562.
- [32] Dark energy, dark matter. URL: <https://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy>.
- [33] Marco Taoso, Gianfranco Bertone, and Antonio Masiero. Dark matter candidates: a ten-point test. *Journal of Cosmology and Astroparticle Physics*, 2008(03):022, mar 2008. doi:10.1088/1475-7516/2008/03/022.
- [34] Brian Fields and Subir Sarkar. Big-Bang nucleosynthesis (2006 Particle Data Group mini-review). 1 2006. arXiv:astro-ph/0601514.
- [35] G Jungman, M P Kamionkowski, and K Griest. Supersymmetric dark matter. *Phys. Rep.*, 267:195–373, Jun 1995. URL: <https://cds.cern.ch/record/283685>.
- [36] Katherine Freese. Status of Dark Matter in the Universe. 1(06):325–355. arXiv:1701.01840, doi:10.1142/S0218271817300129.

BIBLIOGRAPHY

- [37] Jonathan L. Feng. Dark Matter Candidates from Particle Physics and Methods of Detection. 48:495–545. [arXiv:1003.0904](#), [doi:10.1146/annurev-astro-082708-101659](#).
- [38] Virginia Trimble. Existence and Nature of Dark Matter in the Universe. 25:425–472. [doi:10.1146/annurev.aa.25.090187.002233](#).
- [39] Mark Srednicki, Richard Watkins, and Keith A. Olive. Calculations of Relic Densities in the Early Universe. 310:693. [doi:10.1016/0550-3213\(88\)90099-5](#).
- [40] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. 267:195–373. [arXiv:hep-ph/9506380](#), [doi:10.1016/0370-1573\(95\)00058-5](#).
- [41] Antonio Boveia and Caterina Doglioni. Dark Matter Searches at Colliders. 68:429–459. [arXiv:1810.12238](#), [doi:10.1146/annurev-nucl-101917-021008](#).
- [42] Albert M Sirunyan et al. Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV. 03:141. [arXiv:1901.01553](#), [doi:10.1007/JHEP03\(2019\)141](#).
- [43] Albert M. Sirunyan et al. Search for dark matter particles produced in association with a top quark pair at $\sqrt{s} = 13$ TeV. 122(1):011803. [arXiv:1807.06522](#), [doi:10.1103/PhysRevLett.122.011803](#).
- [44] Search for Dark Matter production associated with bottom quarks with 13.3 fb⁻¹ of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC.
- [45] Daniel Abercrombie et al. Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum. [arXiv:1507.00966](#).

- [46] Tomohiro Abe et al. LHC Dark Matter Working Group: Next-generation spin-0 dark matter models. 27:100351. [arXiv:1810.09420](#), [doi:10.1016/j.dark.2019.100351](#).
- [47] Martin Bauer, Ulrich Haisch, and Felix Kahlhoefer. Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators. 05:138. [arXiv:1701.07427](#), [doi:10.1007/JHEP05\(2017\)138](#).
- [48] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. URL: <http://cds.cern.ch/record/782076>, [doi:10.5170/CERN-2004-003-V-1](#).
- [49] A. Tricomi. Upgrade of the CMS tracker. 9(03):C03041–C03041, Mar 2014. [doi:10.1088/1748-0221/9/03/c03041](#).
- [50] *Technical proposal*. LHC technical proposal. CERN, Geneva, 1994. Cover title : CMS, the Compact Muon Solenoid : technical proposal. URL: <https://cds.cern.ch/record/290969>.
- [51] Sudeshna Banerjee. Performance of Hadron Calorimeter with and without HO. Technical report, CERN, Geneva, Nov 1999. URL: <https://cds.cern.ch/record/687178>.
- [52] CMS Collaboration. Performance of the CMS drift tube chambers with cosmic rays. 5(03):T03015–T03015, Mar 2010. [doi:10.1088/1748-0221/5/03/t03015](#).
- [53] CMS Collaboration. Performance of the CMS cathode strip chambers with cosmic rays. 5(03):T03018–T03018, Mar 2010. [doi:10.1088/1748-0221/5/03/t03018](#).

BIBLIOGRAPHY

- [54] M. Abbrescia, A. Colaleo, G. Iaselli, F. Loddo, M. Maggi, et al. The RPC system for the CMS experiment at the LHC. 508(1):137–141, Aug 2003. doi:[10.1016/S0168-9002\(03\)01336-6](https://doi.org/10.1016/S0168-9002(03)01336-6).
- [55] The CMS collaboration. Identification of b-quark jets with the cms experiment. *Journal of Instrumentation*, 8(04):P04013–P04013, Apr 2013. URL: <http://dx.doi.org/10.1088/1748-0221/8/04/P04013>, doi:[10.1088/1748-0221/8/04/p04013](https://doi.org/10.1088/1748-0221/8/04/p04013).
- [56] Daniel Guest, Julian Collado, Pierre Baldi, Shih-Chieh Hsu, Gregor Urban, et al. Jet flavor classification in high-energy physics with deep neural networks. 94(11):112002, Dec 2016. doi:[10.1103/PhysRevD.94.112002](https://doi.org/10.1103/PhysRevD.94.112002).
- [57] A.M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. 12(10):P10003. arXiv:[1706.04965](https://arxiv.org/abs/1706.04965), doi:[10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003).
- [58] CMS EGAMMA POG. Electron tag-and-probe. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/ElectronTagAndProbe>.
- [59] CMS EGAMMA POG. <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaPOG>.
- [60] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical report, CERN, Geneva, Apr 2009. URL: <https://cds.cern.ch/record/1194487>.
- [61] Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector. Technical report, 2010. URL: <http://cds.cern.ch/record/1247373>.

- [62] CMS Collaboration. Performance of τ -lepton reconstruction and identification in cms. 7(01):P01001. URL: <http://stacks.iop.org/1748-0221/7/i=01/a=P01001>.
- [63] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. 2008(04):063. URL: <http://stacks.iop.org/1126-6708/2008/i=04/a=063>.
- [64] The CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. 6(11):P11002. URL: <http://stacks.iop.org/1748-0221/6/i=11/a=P11002>.
- [65] The CMS collaboration. Identification of b-quark jets with the cms experiment. 8(04):P04013. URL: <http://stacks.iop.org/1748-0221/8/i=04/a=P04013>.
- [66] A.M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. 13(05):P05011. [arXiv:1712.07158](https://arxiv.org/abs/1712.07158), [doi:10.1088/1748-0221/13/05/P05011](https://doi.org/10.1088/1748-0221/13/05/P05011).
- [67] MET performance in 8 TeV data. URL: <https://cds.cern.ch/record/1543527>.
- [68] CMS MUON POG. Swguidemuonidrun2. URL: https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Loose_Muon.
- [69] CMS MUON POG. Swguidemuonidrun2. URL: https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight_Muon.
- [70] CMS MUON POG. Muon id and efficiency 2016. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonWorkInProgressAndPagResults>.
- [71] CMS MUON POG. Muon id and efficiency 2017. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs2017>.

BIBLIOGRAPHY

- [72] CMS MUON POG. Muon id and efficiency 2018. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs2018>.
- [73] CMS EGAMMA POG. Cut based electron id for run 2. <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>.
- [74] CMS EGAMMA POG. Egammaid recipes for run2. <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2>.
- [75] JetID13TeVRun2016 < CMS < TWiki, Jun 2021. [Online; accessed 30. Jun. 2021]. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2016>.
- [76] JetID13TeVRun2017 < CMS < TWiki, Jun 2021. [Online; accessed 30. Jun. 2021]. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2017>.
- [77] JetID13TeVRun2018 < CMS < TWiki, Jun 2021. [Online; accessed 30. Jun. 2021]. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2018>.
- [78] CMS TAU POG. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendationForRun2>.
- [79] CMS MUON POG. Btagrecommndation94x. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation94X>.
- [80] CMS MUON POG. Btagsfmethods. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods>.
- [81] CMS Jet/MET POG. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2>.

- [82] CMS DPG-PH JetMET. Met report. URL: <https://indico.cern.ch/event/759372/contributions/3149378/attachments/1721436/2802416/metreport.pdf>.
- [83] CMS JetMET POG. Ee noise mitigation. URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETUncertaintyPrescription#Instructions_for_2017_data_with.
- [84] CMS ECAL DPG. Level 1 ecal prefiring. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1ECALPrefiringWeightRecipe>.
- [85] Alan J. Barr. Measuring slepton spin at the LHC. *J. High Energy Phys.*, 2006(02):042, Feb 2006. doi:10.1088/1126-6708/2006/02/042.
- [86] Ulrich Haisch, Priscilla Pani, and Giacomo Polesello. Determining the CP nature of spin-0 mediators in associated production of dark matter and pairs. *J. High Energy Phys.*, 2017(2):131–24, Feb 2017. doi:10.1007/JHEP02(2017)131.
- [87] CMS Luminosity Measurements for the 2016 Data Taking Period. Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017. URL: <https://cds.cern.ch/record/2257069>.
- [88] CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-LUM-17-004, CERN, Geneva, 2018. URL: <https://cds.cern.ch/record/2621960>.
- [89] CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-LUM-18-002, CERN, Geneva, 2019. URL: <https://cds.cern.ch/record/2676164>.

BIBLIOGRAPHY

- [90] JECUncertaintySources < CMS < TWiki, Jun 2021. [Online; accessed 30. Jun. 2021]. URL: https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources#Run2_JEC_uncertainty_correlation.
- [91] PileupJSONFileforData < CMS < TWiki, Jun 2021. [Online; accessed 30. Jun. 2021]. URL: https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData#Pileup_JSON_Files_For_Run_II.
- [92] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, and et al. Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}=8$ tev. *Physical Review Letters*, 112(23), Jun 2014. URL: <http://dx.doi.org/10.1103/PhysRevLett.112.231802>, doi:10.1103/physrevlett.112.231802.
- [93] *Physics Letters B*, 772:884, Sep 2017. URL: <http://dx.doi.org/10.1016/j.physletb.2017.09.030>, doi:10.1016/j.physletb.2017.09.030.
- [94] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al. Measurement of the WZ production cross section in pp collisions at $\sqrt{s}=13$ tev. *Physics Letters B*, 766:268–290, Mar 2017. URL: <http://dx.doi.org/10.1016/j.physletb.2017.01.011>, doi:10.1016/j.physletb.2017.01.011.
- [95] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71(2):1554–19, Feb 2011. doi:10.1140/epjc/s10052-011-1554-0.