



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ  
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

---

**Μελετη γωνιακών κατανομών ζεύγους top quark με τον  
ανιχνευτή CMS στο LHC και αναζήτηση νέας φυσικής**

---

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΤΟΥ  
**Γεώργιου Μπάκα**

Πτυχιούχου Φυσικών Εφαρμογών του Εθνικού Μετσόβιου Πολυτεχνείου

Επιβλέπων:  
Γεώργιος Τσιπολίτης  
Καθηγητής ΕΜΠ

Συνεπιβλέπων:  
Κωνσταντίνος Κουσουρής  
Αναπληρωτής Καθηγητής ΕΜΠ

Αθήνα  
15 Ιουνίου 2023





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ  
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

---

**Μελετη γωνιακών κατανομών ζεύγους top quark με τον ανιχνευτή  
CMS στο LHC και αναζήτηση νέας φυσικής**

---

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΤΟΥ

**Γεώργιου Μπάκα**

Πτυχιούχου Φυσικών Εφαρμογών του Εθνικού Μετσόβιου Πολυτεχνείου

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ:

- 1.....Γεώργιος Τσιπολίτης, Καθηγητής ΕΜΠ  
(Επιβλέπων)
- 2.....Κωνσταντίνος Κουσουρής, Αν. Καθ. ΕΜΠ  
(Επιβλέπων)
- 3.....Κωνσταντίνος Φουντάς, Καθηγητής Πανεπιστήμιο  
Ιωαννίνων (Επιβλέπων)

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

- 1.....Γεώργιος Τσιπολίτης, Καθ. ΕΜΠ
- 2.....Κωνσταντίνος Κουσουρής, Αν. Καθ. ΕΜΠ
- 3.....Κωνσταντίνος Φουντάς, Καθηγητής Πανεπιστήμιο  
Ιωαννίνων
- 4.....Καραποστόλη Γεωργία, Επικ. Καθ. ΕΜΠ
- 5.....Αλέξανδρος Κεχαγιάς, Καθ. ΕΜΠ
- 6.....Νικόλαος Μαυρόματος, Καθ. ΕΜΠ
- 7.....Ευαγγελία Δρακοπούλου, Ερευνήτρια Γ -  
Ε.Κ.Ε.Φ.Ε 'Δημόκριτος'

Αθήνα, Ιούνιος 2023





NATIONAL TECHNICAL UNIVERSITY OF ATHENS  
SCHOOL OF APPLIED MATHEMATICS  
AND PHYSICAL SCIENCE

---

**Study of top quark pair angular distributions with the  
CMS detector of the LHC at CERN and search for new  
physics**

---

A dissertation submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
by

**Georgios Bakas**

Supervisor:  
Georgios Tsipolitis  
Professor, NTUA

Co-Supervisor:  
Konstantinos Kousouris  
Associate Professor, NTUA

Athens  
15 of June 2023





NATIONAL TECHNICAL UNIVERSITY OF ATHENS  
SCHOOL OF APPLIED MATHEMATICS  
AND PHYSICAL SCIENCE

---

**Study of top quark pair angular distributions with the  
CMS detector of the LHC at CERN and search for new  
physics**

---

DOCTORAL OF PHILOSOPHY PHYSICS  
by  
**Georgios Bakas**

ADVISORY COMMITTEE:

- 1.....Georgios Tsipolitis, Professor NTUA (Supervisor)
- 2..... Konstantinos Kousouris, Associate Professor  
NTUA (Supervisor)
- 3.....Konstantinos Fountas, Professor, University of  
Ioannina

EXAMINATION COMMITTEE:

- 1.....Georgios Tsipolitis, Prof. NTUA
- 2.....Konstantinos Kousouris, Associate Prof. NTUA
- 3.....Konstantinos Fountas, Prof. University of  
Ioannina
- 4.....Georgia Karapostoli, Assistant Prof. NTUA
- 5.....Alexandros Kechagias, Prof. NTUA
- 6.....Nikolaos Mavromatos, Prof. NTUA
- 7.....Evangelia Drakopoulou, Researcher N.C.S.R.  
'Demokritos'

Athens, June 2023

Στον πατέρα μου



# Ευχαριστίες

Η παρούσα διδακτορική διατριβή θα ήταν αδυνατό να υλοποιηθεί χωρίς την παρουσία, υποστήριξη και ανεκτικότητα ορισμένων ανθρώπων, καθ'όλη την διάρκειά της. Η προετοιμασία και εκπόνηση μιας διδακτορικής διατριβής είναι ένα αποτέλεσμα μακροσκελές που βασίζεται στον ενθουσιασμό και την αγάπη που έχει ένας άνθρωπος όταν καλείται να παράξει έρευνα. Σε όλη αυτή τη διαδικασία, συντέλεσαν πολλοί άνθρωποι οι οποίοι στάθηκαν δίπλα μου τόσο σε τεχνική όσο και ψυχολογική υποστήριξη.

Πρώτα από όλους θα ήθελα να ευχαριστήσω τους επιβλέποντές μου και πιο συγκεκριμένα τον Δρ. Κωνσταντίνο Κουσουρή Αναπληρωτή καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου του τμήματος Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών και τον Δρ. Γεώργιο Τσιπολίτη Καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου του τμήματος Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών για την υποστήριξη που μου παρείχαν καθ'όλη την διάρκεια της εκπόνησης της διδακτορικής μου διατριβής καθώς και για την καθοδήγηση, τις συμβουλές και την αμέριστη υλικοτεχνική υποστήριξη. Πέραν της καθοδήγησης στο τεχνικό κομμάτι, θα ήθελα να τους ευχαριστήσω ξεχωριστά για την στάση που είχαν απέναντί μου και την κατανόηση που έδειξαν τις δύσκολες στιγμές που παρουσιάστηκαν κατά την εκπόνηση αυτής της διατριβής.

Αισθάνομαι την ανάγκη να απευθύνω ένα μεγάλο ευχαριστώ στον συνάδελφο και φίλο μου Ιωάννη Παπακρισόπουλο για την τεχνική υποστήριξη, την συνεργασία μας και την καθοδήγησή του μέσω των οποίων κατάφερα να φτάσω στο σημείο να εκπονήσω την συγκεκριμένη διατριβή. Θα ήθελα επίσης να ευχαριστήσω όλη την υπόλοιπη ομάδα του Εθνικού Μετσόβιου Πολυτεχνείου: Άννα, Λίζα, Θεωρή και Ρένα οι οποίοι με βοήθησαν όποτε τους χρειάστηκα σε συζητήσεις φυσικής αλλά και τεχνικά θέματα.

Ένα πολύ μεγάλο ευχαριστώ σε όλο το CMS Collaboration, το πείραμα που ανήκει στο σύμπλεγμα επιταχυντικών διατάξεων του LHC του CERN, που δίχως αυτό αυτή η διδακτορική διατριβή θα ήταν αδύνατο να ολοκληρωθεί. Το CERN προσφέρει την δυνατότητα σε ερευνητές από όλο τον κόσμο να μπορέσουν να έρθουν σε επαφή ώστε να μπορέσουν να εξελίσουν την επιστήμη της φυσικής μέσα από την δική τους εξέλιξη. Θα ήθελα να ευχαριστήσω λοιπόν όλους τους συντονιστές της ομάδας που ασχολείται με την μελέτη του  $t\bar{t}$  ζεύγους για την καθοδήγηση και τα σχόλιά τους. Επίσης θα ήθελα να ευχαριστήσω τους Frank Glege και Cristina Vazquez Velez για την βοήθειά τους στο service work που πραγματοποίησα για 2 χρόνια στην ομάδα του Central DCS του CMS.

Τέλος θα ήθελα να ευχαριστήσω τους δικούς μου ανθρώπους που με στήριξαν και με βοήθησαν σε όλες τις στιγμές αυτού του διδακτορικού. Την σύντροφό μου Κωνσταντίνα που στάθηκε δίπλα μου σε όλη την προσπάθεια. Την μητέρα μου, τον πατέρα μου και την αδερφή μου, χωρίς την στήριξη των οποίων δεν θα μπορούσα να εκπονήσω αυτή την διδακτορική διατριβή.



# Περίληψη

Η παρούσα διατριβή περιγράφει τη δουλειά που πραγματοποιήθηκε κατά τη διάρκεια της διδακτορικής μου φοίτησης στο Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ). Η εργασία αποτελείται από δύο αναλύσεις σχετιζόμενες με την παραγωγή ζεύγους τοπ-αντιτοπ κουάρκ ( $t\bar{t}$ ) καθώς και ένα μέρος της συμβολής μου στο Κεντρικό Σύστημα Ελέγχου Ανιχνευτών CMS.

Το πρώτο μέρος της διατριβής περιλαμβάνει λεπτομέρειες σχετικά με το περίγραμμα του καθιερωμένου προτύπου καθώς και τεχνικά χαρακτηριστικά που σχετίζονται με τον ανιχνευτή CMS . Το Κεφάλαιο 1 περιγράφει τα βασικά στοιχεία του Καθιερωμένου Μοντέλου συμπεριλαμβανομένων των μηχανισμών παραγωγής ζεύγους τοπ-αντιτοπ κουάρκ καθώς και της παραγωγής μποζονίων  $Z$  που περιγράφονται από μοντέλα νέας φυσικής. Στο Κεφάλαιο 2, περιγράφονται συνοπτικά ο επιταχυντής LHC και ο ανιχνευτής CMS . Για το σύμπλεγμα του επιταχυντή καθώς και το πείραμα CMS περιγράφονται τα βασικά χαρακτηριστικά που επιτρέπουν στους ερευνητές να αναγνωρίσουν σωματίδια σε αυτή την κλίμακα. Τα βασικά υποσυστήματα ανιχνευτών CMS στα οποία βασίζεται η αναγνώριση των θεμελιωδών σωματιδίων προσδιορίζονται και περιγράφονται περαιτέρω τα βασικά στοιχεία κάθε υπο-ανιχνευτή. Το Κεφάλαιο 3 περιγράφει εν συντομία τους θεμελιώδεις αλγόριθμους που απαιτούνται για την ανακατασκευή αντικειμένων που αναγνωρίζονται στον ανιχνευτή CMS προκειμένου να αναγνωριστούν τα σωματίδια που βρίσκονται μέσα στον ανιχνευτή CMS , αναπτύσσονται και συντονίζονται διάφοροι αλγόριθμοι.

Το δεύτερο μέρος αυτής της διατριβής περιγράφει το Σύστημα Ελέγχου Ανιχνευτών CMS . Ένας ανιχνευτής τέτοιου μεγέθους απαιτεί ένα ισχυρό και αυτοματοποιημένο σύστημα για τον έλεγχο επίβλεψης και παρακολούθησης των διαφόρων στοιχείων και υποσυστημάτων του. Το Κεφάλαιο 4 εισάγει τον τρόπο με τον οποίο αναπτύσσεται και λειτουργεί ένα σύστημα ελέγχου ανιχνευτή για ένα μεγάλο πείραμα όπως το πείραμα CMS . Το Κεφάλαιο 4 δείχνει το έργο του συγγραφέα για την ανάπτυξη και τη συντήρηση διαφόρων λειτουργικών μονάδων που χρησιμοποιήθηκαν από την κεντρική ομάδα του CMS DCS.

Το τρίτο μέρος αυτής της διατριβής περιγράφει δύο  $t\bar{t}$  σχετικές αναλύσεις. Και οι δύο αναλύσεις χρησιμοποιούν δεδομένα σύγκρουσης πρωτονίων στο κέντρο μάζας ενέργειας  $\sqrt{s} = 13 \text{ TeV}$  , που συλλέχθηκαν το 2016, το 2017 και 2018 και παρουσιάζεται σε τελικό συνδυασμένο αποτέλεσμα με τον ανιχνευτή (CMS ). Τα αποτελέσματα αντιστοιχούν σε συνολική φωτεινότητα  $137,1 \text{ fb}^{-1}$  . Στο Κεφάλαιο 5 παρουσιάζεται η μέτρηση της διαφορικής ενεργού διατομής παραγωγής ζευγών τοπ-αντιτοπ κουάρκ υψηλής εγκάρσιας ορμής. Το σύνολο δεδομένων συλλέχθηκε κατά τη διάρκεια της πλήρους εκτέλεσης με τον ανιχνευτή CMS στο CERN από συγκρούσεις πρωτονίου-πρωτονίου σε ενέργεια κέντρου μάζας  $13 \text{ TeV}$  . Η μέτρηση χρησιμοποιεί γεγονότα όπου και τα δύο υποψήφια τοπ κουάρκ αποσυντίθενται αδρονικά και ανακατασκευάζονται ως μεγάλοι πίδακες ενέργειας. Η ενεργός διατομή εξάγεται διαφορικά ως συνάρτηση διαφόρων κινηματικών μεταβλητών του συστήματος ζεύγους τοπ κουάρκ ή τοπ κουάρκ. Τα αποτελέσματα αναλύονται σε δύο διαφορετικά επίπεδα ανιχνευτή και συγκρίνονται με διάφορα θεωρητικά μοντέλα. Το κεφάλαιο 6 βασίζεται στην ανάλυση που περιγράφεται στο Κεφάλαιο 5 και παρουσιάζει έναν συνδυασμό αναζητήσεων για βαρείς συντονισμούς που αποσυντίθενται σε ζεύγη τοπ-αντιτοπ κουάρκ, χρησιμοποιώντας το πλήρες σύνολο δεδομένων σύγκρουσης πρωτονίου-πρωτονίου που συλλέχθηκαν από τον ανιχνευτή CMS . Ο συνδυασμός περιλαμβάνει συνεισφορές από τα πλήρως αδρονικά διασπώμενα τοπ-αντιτοπ κουάρκ που έχουν υψηλή εγκάρσια ορμή ( $p_T$  ). Τα αποτελέσματα που παρουσιάζονται αυξάνουν περαιτέρω τα όρια αποκλεισμού μάζας για τα μοντέλα νέας φυσικής.

# Abstract

This dissertation describes the work during my PhD pursue that was carried out at the National And Technical University of Athens (NTUA). The work consists of two correlated  $t\bar{t}$  analyses as well as a part of contribution to the Central CMS Detector Control System.

The first part of the dissertation involves details regarding the thoretical Standard Model outline as well as technical characteristics related to the CMS detector. Chapter 1 outlines the basic elements of the Standard Model (SM) including the  $t\bar{t}$  production mechanisms as well as the production of  $Z'$  bosons found within models with a new  $U(1)$  gauge symmetry. The  $Z'$  is the gauge boson of the (broken)  $U(1)$  symmetry. In Chapter 2. the LHC and the CMS detector are briefly described. The accelerator complex as well as the CMS experiment is outlined in order to identify the key characteristics that allow researchers to identify particles at that scale. The basic CMS detector subsystems on which the identification of fundamental particles rely upon are identified and the basic components of each subdetector is further described. Chapter 3 briefly describes the fundamental algorithms required for the object reconstruction of objects that are identified within the CMS detector. In order to identify the particles found within the CMS detector, various algorithms are developed and tuned.

The second part of this dissertation discusses the CMS detector Control System. A detector of such size requires a powerful, adaptive and automated system for the control supervision and monitoring of its various components and subsystems. Chapter 4 introduces how a Detector control system for a large experiment such as the CMS experiment is developed and operated. Chapter 4 indicates the work of the author for the development and maintenance of various modules that were used by the CMS central DCS team.

The third part of this dissertation describes two  $t\bar{t}$  related analysis. Both analyses use proton-proton collision data at the center of mass energy  $\sqrt{s} = 13$  TeV, collected in 2016, 2017 and 2018 and presented in a final combined result with the Compact Muon Solenoid (CMS) detector at CERN LHC and correspond to a total integrated luminosity of  $137.1 \text{ fb}^{-1}$ . Chapter 5 introduces a measurement of the production cross section of high transverse momentum ( $p_T$ ) top quark pairs is reported. The dataset was collected during the full Run II with the CMS detector at the CERN LHC from proton-proton collisions at a center-of-mass energy of 13 TeV. The measurement uses events where either both top quark candidates decay hadronically and are reconstructed as large-R jets with  $p_T > 450 \text{ GeV}$ . The cross section is extracted differentially as a function of kinematic variables of the top quark or top quark pair system. The results are unfolded to the particle and parton levels, and are compared to various theoretical models. Chapter 6 is based on the analysis described in Chapter 4 and presents a combination of searches for heavy resonances decaying to top quark-antiquark pairs, using the full set of proton-proton collision data collected at a center-of-mass energy of  $\sqrt{s} = 13$  TeV in the Full Run II, corresponding to an integrated luminosity of  $137 \text{ fb}^{-1}$ . The combination includes contributions from the fully hadronic final state top and antitop quarks that have a high transverse momentum ( $p_T$ ). The results presented here further increase the mass exclusion limits for several models of new physics.

# Contents

<b>List of Figures</b>	<b>xvii</b>
<b>List of Tables</b>	<b>xxv</b>
<b>I Introduction</b>	<b>1</b>
<b>1 Theoretical Overview</b>	<b>3</b>
1.1 Standard Model of Particle Physics . . . . .	3
1.1.1 Quantum Electro Dynamics (QED) . . . . .	4
1.1.2 Quantum Chromo Dynamics (QCD) . . . . .	5
1.1.3 Weak Interaction . . . . .	5
1.1.4 Electroweak Unification . . . . .	6
1.1.5 Higgs Mechanism and Spontaneous Symmetry Breaking . . . . .	6
1.2 Top Quark . . . . .	7
1.2.1 Production in Hadron Colliders and Decay . . . . .	7
1.3 Physics of the Proton-Proton Collisions . . . . .	8
1.3.1 Pile-Up . . . . .	9
1.4 Z' physics . . . . .	9
<b>2 The CMS Experiment at the LHC</b>	<b>11</b>
2.1 THE LHC Complex . . . . .	11
2.2 The CMS Detector . . . . .	13
2.2.1 Coordinate System . . . . .	14
2.2.2 Solenoid . . . . .	15
2.2.3 Tracker . . . . .	15
2.2.4 Calorimeters . . . . .	16
2.2.5 Muon System . . . . .	18
2.2.6 Trigger . . . . .	19
2.2.7 Upgrades . . . . .	20
<b>3 Object Reconstruction</b>	<b>23</b>
3.1 Particle Flow Algorithm . . . . .	23
3.2 Particle Identification and Reconstruction . . . . .	23
3.2.1 Muon Identification . . . . .	23
3.3 Jet Algorithms . . . . .	24
3.3.1 Anti- $k_T$ . . . . .	24
3.3.2 Soft Drop . . . . .	26
3.3.3 PUPPI . . . . .	26
3.4 Heavy Flavor Jets . . . . .	26
3.5 Missing Transverse Momentum . . . . .	27

<b>II CMS Detector Control System</b>	<b>29</b>
<b>4 Control Systems</b>	<b>31</b>
4.1 SCADA systems . . . . .	31
4.2 WinCC_OA Toolkit . . . . .	31
4.2.1 Main Features . . . . .	31
4.3 Joint Control Project Framework (JCOP) . . . . .	33
4.3.1 Finite State Machine (FSM) . . . . .	33
4.3.2 Archiving Database . . . . .	34
4.3.3 Configuration Database . . . . .	34
4.3.4 Access Control . . . . .	35
4.4 The CMS Detector Control System . . . . .	35
4.5 The CMS Central DCS team . . . . .	35
4.6 CMS DCS Configuration Database . . . . .	36
4.7 CMS DCS Installation Tool . . . . .	42
4.8 MoxaE1240 Hardware Component . . . . .	43
4.9 Other projects . . . . .	46
<b>III <math>t\bar{t}</math> Analysis</b>	<b>49</b>
<b>5 Boosted <math>t\bar{t}</math> Differential Measurements</b>	<b>51</b>
5.1 Introduction . . . . .	51
5.2 Samples . . . . .	52
5.2.1 Data . . . . .	52
5.2.2 Simulation . . . . .	52
5.3 Trigger . . . . .	58
5.4 Reconstruction and Selection . . . . .	62
5.4.1 Object Reconstruction . . . . .	62
5.4.2 Selection . . . . .	62
5.4.3 Multivariate Discriminant . . . . .	65
5.4.4 Parton level . . . . .	68
5.4.5 Particle level . . . . .	68
5.5 Signal Extraction . . . . .	69
5.5.1 Inclusive cross section . . . . .	69
5.5.2 Differential cross sections . . . . .	70
5.6 Data vs Monte Carlo . . . . .	98
5.7 Fiducial Measurement . . . . .	109
5.8 Combination of different years . . . . .	129
5.9 Unfolded Measurement . . . . .	130
5.9.1 Parton Level . . . . .	130
5.9.2 Particle Level . . . . .	133
<b>6 <math>t\bar{t}</math> resonances in Search for new Physics</b>	<b>161</b>
6.1 Introduction to the analysis . . . . .	161
6.2 Simulated events . . . . .	162
6.3 Analysis Strategy . . . . .	162
6.4 Sensitivity Analysis . . . . .	174
6.5 Signal Discrimination . . . . .	183
6.6 Results . . . . .	192
6.7 Conclusions . . . . .	201

<b>IV</b>	<b>Ελληνική Περίληψη</b>	<b>203</b>
<b>7</b>	<b>Θεωρητική Εισαγωγή</b>	<b>205</b>
7.1	Το Καθιερωμένο Πρότυπο . . . . .	205
7.1.1	Quantum Electro Dynamics (QED) . . . . .	207
7.1.2	Quantum Chromo Dynamics (QCD) . . . . .	207
7.1.3	Ασθενής Δύναμη - Weak Interaction . . . . .	208
7.1.4	Electroweak Unification . . . . .	209
7.1.5	Μηχανισμός Higgs και αυθόρμητο σπάσιμο συμμετρίας . . . . .	209
7.2	Το Τοπ Κουαρκ . . . . .	209
7.2.1	Παραγωγή σε επιταχυντές αδρονίων και διάσπαση . . . . .	210
7.3	Φυσική των συγκρούσεων πρωτονίου-πρωτονίου . . . . .	211
7.3.1	Pile-Up . . . . .	211
7.4	Z φυσική . . . . .	212
<b>8</b>	<b>Το πείραμα CMS στο LHC</b>	<b>213</b>
8.1	Το σύμπλεγμα επιταχυντών LHC . . . . .	213
8.2	Ο ανιχνευτής CMS . . . . .	215
8.2.1	Σύστημα Συντεταγμένων . . . . .	216
8.2.2	Υπεραγωγίμος Σωληνοειδής Μαγνήτης . . . . .	217
8.2.3	Tracker . . . . .	217
8.2.4	Θερμιδόμετρα . . . . .	218
8.2.5	Σύστημα Ανίχνευσης Μιονίων . . . . .	220
8.2.6	Trigger . . . . .	221
<b>9</b>	<b>Ανακατασκευή Γεγονότων και Προσομοιώσεις</b>	<b>223</b>
9.1	Particle Flow Algorithm . . . . .	223
9.2	Αλγόριθμοι Jet . . . . .	223
9.2.1	Αλγόριθμος anti- $k_T$ . . . . .	224
9.2.2	Soft Drop . . . . .	225
9.2.3	PUPPI . . . . .	225
<b>10</b>	<b>Διαφορική Ενεργός Διατομή Παραγωγής Ζεύγους <math>t\bar{t}</math></b>	<b>227</b>
10.1	Ανάλυση Ζεύγους τοπ-αντιτοπ κουαρκ . . . . .	227
10.1.1	Στρατηγική Ανάλυσης . . . . .	227
10.1.2	$t\bar{t}$ Εξαγωγή Σήματος . . . . .	228
10.1.3	Συστηματικές αβεβαιότητες . . . . .	228
10.1.4	Συνδυασμός διαφορετικών ετών . . . . .	229
10.1.5	Unfolding . . . . .	229
10.1.6	Τελικά αποτελέσματα . . . . .	230
10.2	Z . . . . .	230
10.2.1	Στρατηγική Ανάλυσης . . . . .	231
10.2.2	Ανάλυση Ευαισθησίας . . . . .	231
10.2.3	Αποτελέσματα . . . . .	232
10.3	Συμπεράσματα . . . . .	232
	<b>Bibliography</b>	<b>233</b>
<b>A</b>	<b>Top Tagger Efficiency</b>	<b>241</b>

## CONTENTS



# List of Figures

1.1	Display of the particle content of the Standard Model divided into groups of bosons and fermions in their three generations. For every particle, mass, electromagnetic charge and spin is shown. . . . .	4
1.2	Top quark decay . . . . .	7
1.3	$t\bar{t}$ pair production . . . . .	8
2.1	The LHC accelerator complex. . . . .	11
2.2	This computer-generated image shows some of the parts vital for the operation of these components. The magnets must be cooled to 1.9 K (-270.3C) so that the superconducting coils can produce the required 8 T magnetic field strength. . . . .	12
2.3	Cumulative delivered and recorded luminosity versus time for 2010-2012 and 2015-2018 (pp data only). Delivered and recorded luminosity cumulative over all years during stable beams for pp collisions at nominal center-of-mass energy. . . . .	13
2.4	Full view of the CMS detector. Basic properties as weight and dimensions are summarised in a table in the left upper corner. Components of the detector are titled and shown in different colours. . . . .	14
2.5	Sketch of the CMS tracking detector. All subsystems namely pixel and strip detector with inner barrel (TIB), outer barrel (TOB) and outer endcaps (TEC) are shown. Additionally the range in $\eta$ , radius ( $r$ ) and $z$ -direction can be read off. . . . .	16
2.6	CMS ECAL geometry schema showing different components. . . . .	17
2.7	View of the CMS hadronic calorimeter in the $y$ - $z$ plane. . . . .	18
2.8	Cross-sectional view of a quadrant of the CMS detector in the $(r, z)$ plane, showing the layout of the muon detector. . . . .	19
2.9	The architecture of the CMS Level-1 trigger system in Run 2. . . . .	20
3.1	Slice through the CMS detector looking in the direction of the beam pipe. Tracks of different particles on their way through the layers of the detector are displayed. From left to right, tracking system, calorimeters, superconducting solenoid and muon system are shown. . . . .	24
3.2	A sample parton-level event (generated with Herwig [8]), together with many random soft “ghosts”, clustered with four different jets algorithms, illustrating the “active” catchment areas of the resulting hard jets. For $k_t$ and Cam/Aachen the detailed shapes are in part determined by the specific set of ghosts used, and change when the ghosts are modified. . . . .	25
3.3	Performance of the $b$ jet identification efficiency algorithms demonstrating the probability for non- $b$ jets to be misidentified as $b$ jet as a function of the efficiency to correctly identify $b$ jets [63]. . . . .	27
4.1	Schematic representation of a data point. . . . .	32
4.2	Schematic representation of a data point type. . . . .	33
4.3	A simple control system modelled using FSMs. . . . .	34
4.4	Main screen of the CMSfwInstallUtils component . . . . .	36
4.5	Inconsistency check between the project and the database (inconsistencies with red). . . . .	37

## LIST OF FIGURES

4.6	Main screen of the CMSfwInstallUtils Configuration Database component . . . . .	38
4.7	Provides the option of demonstrating only the inconsistent datapoints. . . . .	39
4.8	ConfDB Checks component utilities. . . . .	39
4.9	ConfDB Checks component utilities: When Inconsistency found or missing feature, the user is able to right click and check exactly which datapoint element is inconsistent or has missing configuration. . . . .	40
4.10	ConfDB Checks component allows the users report inconsistencies or missing datapoint elements that were not found. In this case the Central DCS team reproduces the case and incorporates configurations that are not found. . . . .	40
4.11	ConfDB Checks component utilities shows the user the entire DB vs Local project configuration when an inconsistency is found for the Address Configuration of a datapoint. . . . .	41
4.12	Central CMSfwMoxa component panel for the SNMP connection with the target agent, integrated into the Device Editor & Navigator. . . . .	44
4.13	CMSfwMoxa component panel showing temperature values connected to the MOXAE1240 hardware and communicating with WinCC_OA via SNMP protocol. . . . .	44
4.14	CMSfwMoxa Panels for the configuration of the peripheral address for the Channels of a specific agent. . . . .	45
4.15	CMSfwMoxa Panels for the configuration of the peripheral address for the Channels of a specific agent. . . . .	46
5.1	Pileup profile for the signal and control trigger paths. . . . .	58
5.2	Trigger efficiency for the signal path of the analysis vs second leading $p_T$ for 2016. . . . .	60
5.3	Trigger efficiency for the signal path of the analysis vs ht for 2017. . . . .	60
5.4	Trigger efficiency for the signal path of the analysis vs ht for 2018. . . . .	61
5.5	Graphical representation of the analysis region in the BDT-output vs jet softDrop mass. . . . .	64
5.6	BDT response for all years starting from the top: 2016 preVFP, 2016 postVFP, 2017 and 2018 in bottom). . . . .	66
5.7	ROC curve for all years starting from the top: 2016 preVFP, 2016 postVFP, 2017 and 2018 in bottom). . . . .	67
5.8	Contamination of the Control region . . . . .	69
5.9	Leading jet $p_T$ distributions in the QCD control region for all years. . . . .	73
5.10	Second leading jet $p_T$ distributions in the QCD control region for all years. . . . .	74
5.11	Leading jet absolute rapidity $y$ distributions in the QCD control region for all years. . . . .	75
5.12	Second leading jet absolute rapidity $y$ distributions in the QCD control region for all years. . . . .	76
5.13	Dijet mass $m^{t\bar{t}}$ distributions in the QCD control region for all years. . . . .	77
5.14	Dijet $p_T^{t\bar{t}}$ distributions in the QCD control region for all years. . . . .	78
5.15	Dijet rapidity $y^{t\bar{t}}$ distributions in the QCD control region for all years. . . . .	79
5.16	Dijet angular $\chi$ distributions in the QCD control region for all years. . . . .	80
5.17	Leading jet $ \cos(\theta^*) $ distributions in the QCD control region for all years. . . . .	81
5.18	Mass Fit Templates 2016 preVFP . . . . .	82
5.19	Mass Fit Templates 2016 postVFP . . . . .	83
5.20	Mass Fit Templates 2017 . . . . .	84
5.21	Mass Fit Templates 2018 . . . . .	85
5.22	Closure test for top quark leading SoftDrop Mass . . . . .	86
5.23	Closure test for top quark leading jet $p_T$ . . . . .	87
5.24	Closure test for top quark second leading jet $p_T$ . . . . .	88
5.25	Closure test for top quark leading absolute jet rapidity . . . . .	89
5.26	Closure test for top quark leading absolute jet rapidity . . . . .	90
5.27	Closure test for dijet $t\bar{t}$ $p_T$ . . . . .	91
5.28	Closure test for for dijet $t\bar{t}$ mass . . . . .	92

LIST OF FIGURES

5.29 Closure test for for dijet $t\bar{t}$ rapidity . . . . .	93
5.30 Result of the template fit on data for 2016 preVFP. The red line shows the $t\bar{t}$ contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds. . . . .	94
5.31 Result of the template fit on data for 2016. The red line shows the $t\bar{t}$ contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds. . . . .	95
5.32 Result of the template fit on data for 2017. The red line shows the $t\bar{t}$ contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds. . . . .	96
5.33 Result of the template fit on data for 2018. The red line shows the $t\bar{t}$ contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds. . . . .	97
5.34 SoftDrop mass of the leading AK8 jet after the baseline selection with both AK8 jets containing a b-tagged subjet, with the $BDT > 0.0$ cut (2016 preVFP top (left), 2016 postVFP (right)), $> 0.1$ for 2017 bottom (left) and $> 0.2$ for 2018 bottom (right). . . . .	99
5.35 Distributions of the absolute leading and second jet $p_T$ in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018. . . . .	100
5.36 Distributions of the absolute second leading jet $p_T$ in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018. . . . .	101
5.37 Distributions of the absolute leading jet rapidity $y$ in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018. . . . .	102
5.38 Distributions of the absolute second leading jet rapidity $y$ in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018. . . . .	103
5.39 Distributions of the $t\bar{t}$ system mass in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018). . . . .	104
5.40 Distributions of the $t\bar{t}$ system $p_T$ in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018). . . . .	105
5.41 Distribution of the $t\bar{t}$ system rapidity $y$ in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018). . . . .	106
5.42 Angular distribution of the $t\bar{t}$ $\chi$ in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018). . . . .	107
5.43 Angular distribution of the leading absolute $ \cos(\theta) $ in the signal region for all years (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018). . . . .	108
5.44 Fiducial differential cross section, absolute (left) and normalized (right), as a function of the leading top $p_T$ . The bottom panel shows the ratio (theory - data)/data. . . . .	109
5.45 Fiducial differential cross section, absolute (left) and normalized (right), as a function of the second top $p_T$ . The bottom panel shows the ratio (theory - data)/data. . . . .	110
5.46 Fiducial differential cross section, absolute (left) and normalized (right), as a function of the leading top $ y $ . The bottom panel shows the ratio (theory - data)/data. . . . .	111
5.47 Fiducial differential cross section, absolute (left) and normalized (right), as a function of second top $ y $ . The bottom panel shows the ratio (theory - data)/data. . . . .	112
5.48 Fiducial differential cross section, absolute (left) and normalized (right), as a function of $m^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . .	113
5.49 Fiducial differential cross section, absolute (left) and normalized (right), as a function of $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . .	114
5.50 Fiducial differential cross section, absolute (left) and normalized (right), as a function of top $y^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . .	115
5.51 Fiducial differential cross section, absolute (left) and normalized (right), as a function of top $\chi$ . The bottom panel shows the ratio (theory - data)/data. . . . .	116
5.52 Fiducial differential cross section, absolute (left) and normalized (right), as a function of top the leading $ \cos(\theta^*) $ . The bottom panel shows the ratio (theory - data)/data. . . . .	117
5.53 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the leading top $p_T$ . . . . .	120

LIST OF FIGURES

5.54 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the second leading top  $p_T$ . . . . . 121

5.55 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the leading top absolute rapidity  $y$ . . . . . 122

5.56 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the second leading top absolute rapidity  $y$ . . . . . 123

5.57 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $m_{t\bar{t}}$  dijet mass. 124

5.58 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $p_T^{t\bar{t}}$  dijet  $p_T$ . . . 125

5.59 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $y_{t\bar{t}}$  dijet rapidity. 126

5.60 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $chi_{t\bar{t}}$ . . . . . 127

5.61 Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $cos(\theta^*)$ . . . . 128

5.62 Acceptance and efficiency for the parton-level selection as a function of the various observables. . . . . 136

5.63 Acceptance and efficiency for the particle-level selection as a function of the various observables. . . . . 137

5.64 Simulated migration matrices at parton level. . . . . 138

5.65 Simulated migration matrices at particle level. . . . . 139

5.66 Purity and Stability for the parton-level selection as a function of the various observables. . . . . 140

5.67 Purity and Stability for the particle-level selection as a function of the various observables. . . . . 141

5.68 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of leading jet  $p_T$ . The bottom panel shows the ratio (theory - data)/data. . . . . 142

5.69 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of second leading jet  $p_T$ . The bottom panel shows the ratio (theory - data)/data. . . . . 143

5.70 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of absolute leading jet rapidity. The bottom panel shows the ratio (theory - data)/data. . . . . 144

5.71 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of absolute second leading jet rapidity. The bottom panel shows the ratio (theory - data)/data. . . . . 145

5.72 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the dijet  $t\bar{t}$  mass. The bottom panel shows the ratio (theory - data)/data. . . . . 146

5.73 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of  $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . . 147

5.74 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the dijet rapidity  $y^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . . 148

5.75 Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the dijet angular  $\chi^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . . 149

LIST OF FIGURES

5.76	Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the leading jet $ \cos\theta^* $ . The bottom panel shows the ratio (theory - data)/data. . . . .	150
5.77	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of leading jet $p_T$ . The bottom panel shows the ratio (theory - data)/data. . . . .	151
5.78	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of second leading jet $p_T$ . The bottom panel shows the ratio (theory - data)/data. . . . .	152
5.79	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of absolute leading jet rapidity. The bottom panel shows the ratio (theory - data)/data. . . . .	153
5.80	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of absolute second leading jet rapidity. The bottom panel shows the ratio (theory - data)/data. . . . .	154
5.81	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the dijet $t\bar{t}$ mass. The bottom panel shows the ratio (theory - data)/data. . . . .	155
5.82	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . .	156
5.83	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the dijet rapidity $y^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . .	157
5.84	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the dijet angular $\chi^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data. . . . .	158
5.85	Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the leading jet $\cos\theta^*$ . The bottom panel shows the ratio (theory - data)/data. . . . .	159
6.1	Graphical representation of the analysis region in the BDT-output vs jet softDrop mass. . . . .	164
6.2	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 1.2TeV and 1 % width. . . . .	166
6.3	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 1.4TeV and 1 % width. . . . .	167
6.4	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 1.6TeV and 1 % width. . . . .	168
6.5	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 1.8TeV and 1 % width. . . . .	169
6.6	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 2TeV and 1 % width. . . . .	170
6.7	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 2.5TeV and 1 % width. . . . .	171
6.8	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 3TeV and 1 % width. . . . .	172
6.9	Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the $Z'$ with mass 4TeV and 1 % width. . . . .	173
6.10	Significance for all years (2017 left and 2018 right) for the $Z'$ with mass 4.5TeV and 1 % width. . . . .	173
6.11	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.2TeV and 1 % width. . . . .	174

## LIST OF FIGURES

6.12	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.4TeV and 1 % width. . . . .	175
6.13	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.6TeV and 1 % width. . . . .	176
6.14	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.8TeV and 1 % width. . . . .	177
6.15	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 2TeV and 1 % width. . . . .	178
6.16	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 2.5TeV and 1 % width. . . . .	179
6.17	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 3TeV and 1 % width. . . . .	180
6.18	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 3.5TeV and 1 % width. . . . .	181
6.19	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 4TeV and 1 % width. . . . .	182
6.20	Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 4.5TeV and 1 % width. . . . .	182
6.21	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.2TeV and 1 % width. . . . .	183
6.22	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.4TeV and 1 % width. . . . .	184
6.23	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.6TeV and 1 % width. . . . .	185
6.24	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 1.8TeV and 1 % width. . . . .	186
6.25	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 2TeV and 1 % width. . . . .	187
6.26	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 2.5TeV and 1 % width. . . . .	188
6.27	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 3TeV and 1 % width. . . . .	189
6.28	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 3.5TeV and 1 % width. . . . .	190
6.29	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 4TeV and 1 % width. . . . .	191
6.30	Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the $Z'$ with mass 4.5TeV and 1 % width. . . . .	191
6.31	Datacard used for the combination of the 3 analyses. . . . .	192
6.32	Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a $Z'$ boson with $\Gamma/M$ of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for 2016 preVFP. . . . .	193
6.33	Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a $Z'$ boson with $\Gamma/M$ of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for 2017. . . . .	194

LIST OF FIGURES

6.34	Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a $Z'$ boson with $\Gamma/M$ of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for 2018. . . . .	195
6.35	Yearly Asymptotic limits results with respect to 2018 results. . . . .	196
6.36	Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a $Z'$ boson with $\Gamma/M$ of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for the combined 2016 preVFP, 2017 and 2018 analyses. . . . .	197
6.37	Brazilian Plots for combined years with respect to 2018 results. . . . .	198
7.1	Η διάσπαση τοπ κουαρκ . . . . .	210
7.2	$t\bar{t}$ παρ προδυστιον . . . . .	211
8.1	Το σύμπλεγμα επιταχυντών LHC . . . . .	213
8.2	Σωρευτική παραδοθείσα και καταγεγραμμένη φωτεινότητα σε σχέση με το χρόνο για τα έτη 2010-2012 και 2015-2018 (μόνο δεδομένα ππ). Παραδόθηκε και καταγράφηκε αθροιστική φωτεινότητα για όλα τα χρόνια κατά τη διάρκεια σταθερών δεσμών για συγκρούσεις pp σε ονομαστική ενέργεια κέντρου μάζας. . . . .	215
8.3	Πλήρης προβολή του ανιχνευτή CMS . Βασικές ιδιότητες όπως βάρος και διαστάσεις είναι συνοψίζονται σε έναν πίνακα στο αριστερό επάνω μέρος . . . . .	216
8.4	Σκίτσο του tracker . Εμφανίζονται όλα τα υποσυστήματα, συγκεκριμένα ο ανιχνευτής εικονοστοιχείων και ταινιών με εσωτερική κάθνη (TIB), εξωτερική κάθνη (TOB) και εξωτερικά τερματικά πόματα (TE"). Επιπλέον, το εύρος σε $\eta$ , ακτίνα ( $r$ ) και κατεύθυνση $\zeta$ μπορεί να διαβαστεί. . . . .	218
8.5	Σχήμα γεωμετρίας CMS ECAL που δείχνει διαφορετικά στοιχεία. . . . .	219
8.6	Προβολή του αδρονικού θερμοδομέτρου CMS στο επίπεδο $\psi$ - $\zeta$ . . . . .	220
8.7	Προβολή διατομής ενός τεταρτημορίου του ανιχνευτή CMS στο ( $r$ , $z$ ) επίπεδο, που δείχνει τη διάταξη του ανιχνευτή μιονίων. . . . .	221
8.8	Η αρχιτεκτονική του συστήματος ενεργοποίησης "ΜΣ ΛεεL-1 . . . . .	222
9.1	Ανακατασκευή Γεγονότων στον ανιχνευτή CMS . . . . .	224
A.1	Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 preVFP using a random jet. . . . .	245
A.2	Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 preVFP using both jets. . . . .	246
A.3	Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 postVFP using a random jet. . . . .	247
A.4	Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 postVFP using both jets. . . . .	248
A.5	Distributions of the probe jet and the jets used for the measurement in the related regions for 2017 using a random jet. . . . .	249
A.6	Distributions of the probe jet and the jets used for the measurement in the related regions for 2017 using both jets. . . . .	250
A.7	Distributions of the probe jet and the jets used for the measurement in the related regions for 2018 using a random jet. . . . .	251
A.8	Distributions of the probe jet and the jets used for the measurement in the related regions for 2018 using both jets. . . . .	252
A.9	Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per $p_T$ region and inclusive with statistical and systematic uncertainties . . . . .	253

LIST OF FIGURES

A.10 Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per JMAR proposed  $p_T$  region and inclusive with statistical and systematic uncertainties . . . . . 253

A.11 Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per  $p_T$  region and inclusive with statistical and systematic uncertainties . . . . . 254

A.12 Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per JMAR proposed  $p_T$  region and inclusive with statistical and systematic uncertainties . . . . . 254

A.13 Leading jet  $p_T$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 255

A.14 Second leading jet  $p_T$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 256

A.15 Leading jet rapidity distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 257

A.16 Second leading jet rapidity distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 258

A.17 Dijet mass (mJJ) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 259

A.18 Dijet rapidity (yJJ) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 260

A.19 Dijet  $p_T$  (ptJJ) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 261

A.20 Dijet chi ( $\chi$ ) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 262

A.21  $\cos(\theta^*)$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row). . . . . 263



# List of Tables

5.1	Data samples 2016. . . . .	52
5.2	Data samples 2017. . . . .	52
5.3	Data samples 2018. . . . .	53
5.4	Monte Carlo samples 2016 preVFP. . . . .	54
5.5	Monte Carlo samples 2016 postVFP. . . . .	55
5.6	Monte Carlo samples 2017 . . . . .	56
5.7	Monte Carlo samples 2018 . . . . .	57
5.8	Summary of triggers used in the analysis. . . . .	59
5.9	Summary of triggers used in the 2017 and 2018 analyses. . . . .	59
5.10	Baseline Selection Criteria . . . . .	63
5.11	Selection requirements per analysis region for 2016 pre and post VFP . . . . .	63
5.12	Selection requirements per analysis region for 2017 (Control and Signal triggers are the same) . . . . .	63
5.13	Selection requirements per analysis region for 2017 (Control and Signal triggers are the same) . . . . .	63
5.14	Definition of parton-level phase space. . . . .	68
5.15	Definition of particle-level phase space. . . . .	68
5.16	Results of the $t\bar{t}$ signal strength for each year. . . . .	70
5.17	Results of the fit in $SR_A$ for 2016 preVFP. . . . .	70
5.18	Results of the fit in $SR_A$ for 2016 postVFP. . . . .	71
5.19	Results of the fit in $SR_A$ for 2017. . . . .	71
5.20	Results of the fit in $SR_A$ for 2018. . . . .	72
5.21	Expected and observed event yields in the signal region for all analyzed years. . . . .	98
5.22	Correlations for systematic uncertainties between the various years . . . . .	129
5.23	$\chi^2$ Values of the Absolute differential cross sections at Parton level. . . . .	131
5.24	$\chi^2$ Values of Normalized differential cross sections at Parton level. . . . .	132
5.25	$\chi^2$ Values of the Absolute differential cross sections at Particle level. . . . .	134
5.26	$\chi^2$ Values of Normalized differential cross sections at Particle level. . . . .	135
6.1	Monte Carlo simulated 2016 $Z'$ samples. . . . .	162
6.2	Monte Carlo Simulated 2017 $Z'$ samples. . . . .	163
6.3	Monte Carlo Simulated 2018 $Z'$ samples. . . . .	163
6.4	Baseline Selection Criteria . . . . .	164
6.5	Selection requirements per analysis region for 2016 pre and post VFP. . . . .	164
6.6	Selection requirements per analysis region for 2017. . . . .	165
6.7	Selection requirements per analysis region for 2018. . . . .	165
6.8	Sliding $m_{JJ}$ cut for all $Z'$ masses. Same cut is applied for all years. . . . .	166
6.9	Expected and observed cross section limits at 95% CL, for the 1% width $Z'$ resonance hypothesis for the 2016 preVFP analysis. . . . .	199
6.10	Expected and observed cross section limits at 95% CL, for the 1% width $Z'$ resonance hypothesis for the 2017 analysis. . . . .	199
6.11	Expected and observed cross section limits at 95% CL, for the 1% width $Z'$ resonance hypothesis for the 2018 analysis. . . . .	200

LIST OF TABLES

6.12	Expected and observed cross section limits at 95% CL, for the 1% width $Z'$ resonance hypothesis for the combined (2016preVFP, 2017, 2018) analysis. . . . .	200
10.1	Μεταβλητό $m_{JJ}$ cut για όλες τις μάζες των διάφορων $Z$ . . . . .	231
A.1	Top Tagger efficiency Values for 2016 preVFP. . . . .	242
A.2	Top Tagger efficiency Values for 2016 preVFP using JMAR proposed $p_T$ regions. . . . .	242
A.3	Top Tagger efficiency Values for 2016 postVFP. . . . .	242
A.4	Top Tagger efficiency Values for 2016 postVFP using JMAR proposed $p_T$ regions. . . . .	243
A.5	Top Tagger efficiency Values for 2017. . . . .	243
A.6	Top Tagger efficiency Values for 2017 using JMAR proposed $p_T$ regions. . . . .	243
A.7	Top Tagger efficiency Values for 2018. . . . .	243
A.8	Top Tagger efficiency Values for 2018 using JMAR proposed $p_T$ regions. . . . .	243
A.9	Top Tagger SF Values for 2016 preVFP. . . . .	243
A.10	Top Tagger SF Values for 2016 preVFP using JMAR proposed $p_T$ regions. . . . .	243
A.11	Top Tagger SF Values for 2016 postVFP. . . . .	244
A.12	Top Tagger SF Values for 2016 postVFP using JMAR proposed $p_T$ regions. . . . .	244
A.13	Top Tagger SF Values for 2017. . . . .	244
A.14	Top Tagger SF Values for 2017 using JMAR proposed $p_T$ regions. . . . .	244
A.15	Top Tagger SF Values for 2018. . . . .	244
A.16	Top Tagger SF Values for 2018 using JMAR proposed $p_T$ regions. . . . .	244
A.17	Top Tagger SF Values for all years. . . . .	244

## **Part I**

# **Introduction**



# Chapter 1

## Theoretical Overview

The goal of particle physics is to explain the universe we observe with basic building blocks and fundamental interactions between them. Physicists have developed a theory answering many questions about the development and observable state of the universe. Up to now, the Standard Model of elementary particles, is the most precisely verified theory in physics. In the following sections, a basic introduction of this theory will be presented. Also, a special focus on the elementary particle that plays the most important part in this analysis, the top quark.

### 1.1 Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a quantum field theory describing elementary particles and their interactions, providing affiliated equations of motion. It was proposed by Glashow, Salam and Weinberg in the sixties [1], [2]. Its symmetry group is described by  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  representing the current model of strong interactions via the  $SU(3)_C$  and the unification of the electromagnetic and weak interactions into the electroweak via the  $SU(2)_L \otimes U(1)_Y$ .

All particles contained in this theory have been discovered and various experiments have confirmed the predictions of the SM at very high precision. Elementary particles of the SM and their properties are presented in Figure 1.1. They are ordered in two fundamental groups, depending on their spin. Spin- $\frac{1}{2}$  particles are called fermions while particles with integer spin are named bosons.

Fermions are further divided into quarks and leptons. While quarks are affected by the strong force, leptons aren't. Both of these groups consists of three generations. Each generation contains an isospin doublet with two particles. A quark generation consists of one up and down-type quarks distinguished by their electromagnetic charge of  $+\frac{2}{3}e$  and  $-\frac{1}{3}e$ , respectively. For leptons, a generation is built from a particle with charge  $-e$  and a neutral neutrino. Quarks and leptons from the first generation are the building blocks of atoms and thus form matter that we observe on earth. All other particles have a larger mass and therefore a finite lifetime. Additionally, every fermion has an anti-particle which has the same mass but opposite charge.

Bosons are carriers of the three fundamental forces included in the SM. The massless photons and gluons are the carriers of the electromagnetic and strong force, respectively.  $W^\pm$  and  $Z^0$  bosons are massive and transfer the weak interaction. While all called bosons are spin-1 particles, the Higgs boson is scalar. It is a consequence of the Higgs field, that gives elementary particles their mass. Every interaction of two particles is described by a boson emitted from one and absorbed by other particle. Thus, a boson changes energy and momentum of an absorbing emitting particle. For each force a charge is introduced. Only particles carrying the charge connected to a force can interact with the corresponding boson. Hence, the amount of charge carried by a particle is proportional to the probability of emitting or absorbing a boson.

With the discovery of the Higgs boson in 2012 [3], [4], all particles of the SM have been experimentally confirmed. Although the theory is believed to be incomplete because of known phenomena (like gravity and dark matter) are not included, the Standard Model is very successful describing outcomes of experiments with very high precision. A more detailed look following [1] on the three fundamental interactions is given in the sections below.

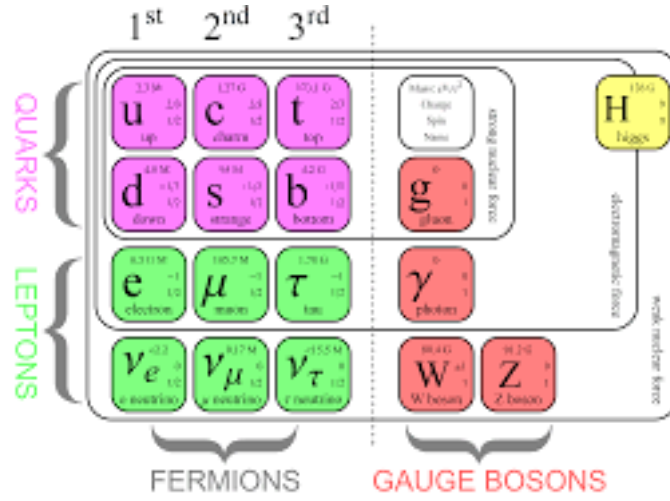


Figure 1.1: Display of the particle content of the Standard Model divided into groups of bosons and fermions in their three generations. For every particle, mass, electromagnetic charge and spin is shown.

### 1.1.1 Quantum Electro Dynamics (QED)

In the history of the Standard Model, QED was the first mechanism to describe known interactions in a quantum field theory [5]. The charge of QED is called the electromagnetic charge and it is usually stated in units of the elementary electron charge  $e$ . The mechanism to construct a quantum field theory is to take the Lagrange density for fermions, described by Dirac spinors and require local gauge invariance. The QED requires the Lagrange density to be invariant under transformations of a  $U(1)$  symmetry group. Equation 1.1 shows the Lagrange density of the Dirac equation.

$$\mathcal{L} = \underbrace{\bar{\psi}i\gamma^\mu\partial_\mu\psi}_{\text{kinetic term}} - \underbrace{m\bar{\psi}\psi}_{\text{mass term}} \quad (1.1)$$

Here  $\psi$  denotes the Dirac spinors and  $m$  the mass corresponding to the Dirac spinor. A vector of gamma matrices  $\gamma^\mu$  is used to write the Lagrangian in a four dimensional space-time representation. The index  $\mu$  indicates one of the four components of a four-vector. The  $U(1)$  transformation is:

$$\psi \rightarrow \psi' = e^{iq\alpha}\psi \quad (1.2)$$

where the parameter  $q$  will later be identified with a coupling strength and  $\alpha = \alpha(x)$  denotes the phase, dependent on space-time coordinates. While the mass term  $m\bar{\psi}\psi$  is invariant under a transformation  $\psi \rightarrow \psi'$ , the derivative  $\partial_\mu\psi$  is not. In order to construct an invariant Lagrange density, a new vector field  $A_\mu(x)$  is introduced, and defines a covariant derivative  $D_\mu$ :

$$D_\mu = \partial_\mu + iqA_\mu(x) \quad (1.3)$$

With the newly introduced covariant derivative, the Lagrange density becomes:

$$\mathcal{L} = \underbrace{\bar{\psi}i\gamma^\mu\partial_\mu\psi}_{\text{kinetic term}} - \underbrace{m\bar{\psi}\psi}_{\text{mass term}} - \underbrace{q\bar{\psi}\gamma^\mu\psi A_\mu(x)}_{\text{interaction term}} \quad (1.4)$$

By requiring the Lagrange density to be symmetric under  $U(1)$  gauge transformations, a new vector field  $A_\mu$  is predicted. This vector field is then associated with the photon. Because no mass term for the photon field occurs, it is expected to be massless. Additionally, the coupling between fermions  $\psi$  and photon  $A_\mu$  with a coupling strength  $q$  is included.

### 1.1.2 Quantum Chromo Dynamics (QCD)

The theory behind QCD is developed in analogy to the QED. The charge of the QCD is called color and exists in three states for particles: red, green and blue [5]. Additionally, every color has its anti-color carried by antiparticles. Thus, one has to consider rotations in the three dimensional color space, which is represented by an  $SU(3)$  group. Generators  $T_\alpha$ , with a running from 1 up to 8, are used to express these rotations and have to be considered. Similar to QED, a covariant derivative is defined to construct a symmetric Lagrange density. The covariant derivative for QCD is:

$$D_\mu = \partial_\mu + ig_s T^\alpha G_\mu^\alpha \quad (1.5)$$

where  $g_s$  denotes the coupling of quarks to the strong force. Secondly, eight new vector fields  $G_\mu^\alpha$  are introduced, which are associated with gluons, the bosons of the strong force. Because the generators of  $SU(3)$  do not commute, not only interaction between quarks and gluons but also gluon-gluon interaction is allowed. The next equation shows how the self interaction of gluons is combined with their kinetic energy term. The complete Lagrange density of QCD becomes:

$$\mathcal{L} = \underbrace{\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi}_{\text{mass and kinetic term quarks}} - \underbrace{g_s \bar{\psi}(\gamma^\mu T^\alpha G_\mu^\alpha)\psi}_{\text{quark-gluon interaction}} - \underbrace{\frac{1}{4} G_{\mu\nu}^\alpha G^{\alpha\mu\nu}}_{\text{kinetic term gluons and gluon-gluon interaction}} \quad (1.6)$$

Gluon-gluon interactions lead to a special effect in QCD; the energy density between two color charges increases with their distance. Thus, the energy stored in the field between the two charges will increase as well. At some point the energy is large enough to produce new color charged particles. For this reason, quarks can't exist freely but only in color neutral bound states, called hadrons.

### 1.1.3 Weak Interaction

The third force described in the SM is the weak force [5]. The charge of the weak force is called isospin and is carried by the  $W^\pm$  and the  $Z^0$  bosons. Because all weak bosons are massive, the weak interaction has a very limited range. Therefore, weak interaction only takes place on short distances. The corresponding symmetry group is a  $SU(2)$  group. While  $Z^0$  bosons also couple to right handed particles,  $W^\pm$  don't. This asymmetry in coupling strength is caused by a violation in parity. Through the  $W^\pm$  bosons, fermions can change their flavor. The left-handed quarks and leptons are arranged in isospin doublets, while the right-handed particles are represented by singlet. Because no right-handed neutrinos exist, the particle content is described as:

$$\left( \begin{matrix} u \\ d' \end{matrix} \right)_L, \left( \begin{matrix} c \\ s' \end{matrix} \right)_L, \left( \begin{matrix} t \\ b' \end{matrix} \right)_L, \left( \begin{matrix} \nu_e \\ e \end{matrix} \right)_L, \left( \begin{matrix} \nu_\mu \\ \mu \end{matrix} \right)_L, \left( \begin{matrix} \nu_\tau \\ \tau \end{matrix} \right)_L, u_R, d_R, c_R, s_R, b_R, t_R, e_R, \mu_R, \tau_R \quad (1.7)$$

The theory only allows flavor changing processes inside these doublets, although decays from one generation to another have been observed experimentally. The answer to this contradiction is that flavor states and mass states are not referring to the same state. The weak interaction changes the flavor but the observed mass state, is a superposition of different flavor states. By definition, flavor and mass states of up-type quarks are identical. For down-type quarks, the Cabibbo-Kobayash-Maskawa (CKM) matrix [6] is introduced to account for the mixing of the flavor states.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.8)$$

The probability that a quark of type  $i$  will transform into a quark of type  $j$  when it emits a  $W^\pm$  boson is calculated via  $|V_{ij}^2|$ . Because of the mixing states, not only a violation of parity, but also a CP-violation (charge and parity) is observed.

#### 1.1.4 Electroweak Unification

To get a better understanding of the forces in the SM, it would be a large success to lead the different forces back to one fundamental mechanism. A first step towards a unification of all three interactions can be realised by combined QED with the weak force. Electromagnetic charge  $Q$  and weak isospin  $T_3$  are combined to achieve a charge for the electroweak unification, called hypercharge  $Y$ :

$$Y = 2(Q - I_3) \quad (1.9)$$

Once again, the Lagrange density is required to be invariant under transformations of the underlying  $SU(2) \times U(1)$  group. This results into four new massless bosons:  $W_1, W_2, W_3$  and  $B$ .

#### 1.1.5 Higgs Mechanism and Spontaneous Symmetry Breaking

The bosons  $W_1, W_2, W_3$  and  $B$  found above in 1.1.4 are all massless. In contradiction, the masses of  $W$  and  $Z$  bosons are measured to values of 80 GeV and 91 GeV, respectively [7]. Therefore, a theory was developed which solves this problem via spontaneous symmetry breaking [8] [1] [2]. A field is introduced where the Lagrange density for local variations is not symmetric while the Lagrangian of the field itself is. This leads to mass terms for electro-weak bosons and a new boson, the Higgs boson, in the Standard Model Lagrange density. After spontaneous symmetry breaking,  $W_1$  and  $W_2$  are combined to get the  $W^\pm$  bosons of the weak interactions via:

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \quad (1.10)$$

The other two fields  $W_3$  and  $B$  mix together and result into the  $Z^0$  boson from the weak and the photon  $\gamma$  of the electromagnetic interaction:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad (1.11)$$

Here the weak mixing angle  $\theta_W$  is introduced, which also relates to the masses  $M_W$  and  $M_Z$  of the  $W^\pm$  and the  $Z^0$  bosons via:

$$M_Z = \frac{M_W}{\cos \theta_W} \quad (1.12)$$

Finally, also fermion masses can be included in this mechanism. Therefore, a Yukawa coupling to the Higgs field proportional to their masses is introduced. After symmetry breaking, the SM of particle physics can be written in a Lagrangian representation. For this, the Lagrangian density of QCD has to be combined with the full Lagrangian density of the electro-weak interaction after symmetry breaking. The latter, contains all kinetic terms of fermions, mass terms of electro-weak gauge bosons and the Higgs boson, couplings of fermions to electro-weak gauge bosons, self-interaction terms of called electro-weak bosons and Higgs boson, interactions between electro-weak bosons and the Higgs boson as well as the Yukawa couplings of fermions to the Higgs field.



## 1.2 Top Quark

The top quark is an up-type quark belonging to the third generation of the SM and carrying electromagnetic charge of  $+\frac{2}{3}e$  [9]. With its mass of about 173 GeV, the top quark is the heaviest particle in the Standard Model. A third generation of quarks was predicted in 1973 [10] to account for observed CP-violation that could not be explained with the two known quark generations. After finding the much lighter bottom quark only a few years after its prediction, the top quark was discovered in 1995 by the CDF [11] and D collaborations at Tevatron. Because of its high mass, it offers a large phase space for decays and has a short lifetime of approximately  $5 \times 10^{-25} s$  [9]. Because of its short lifetime, the top quark doesn't form bound hadronic states. Thus, measurements of the bare quark are possible, providing direct access to parameters of the Standard Model. Additionally, the top quark is of particular interest for searches of physics beyond the Standard Model since it is often part of the final state and/or a dominant background [12], [13].

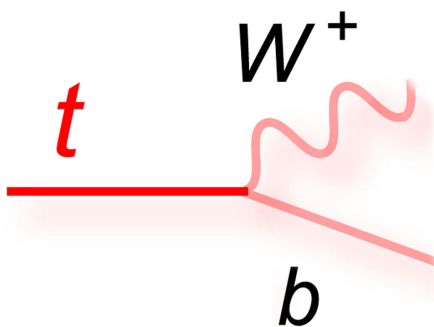


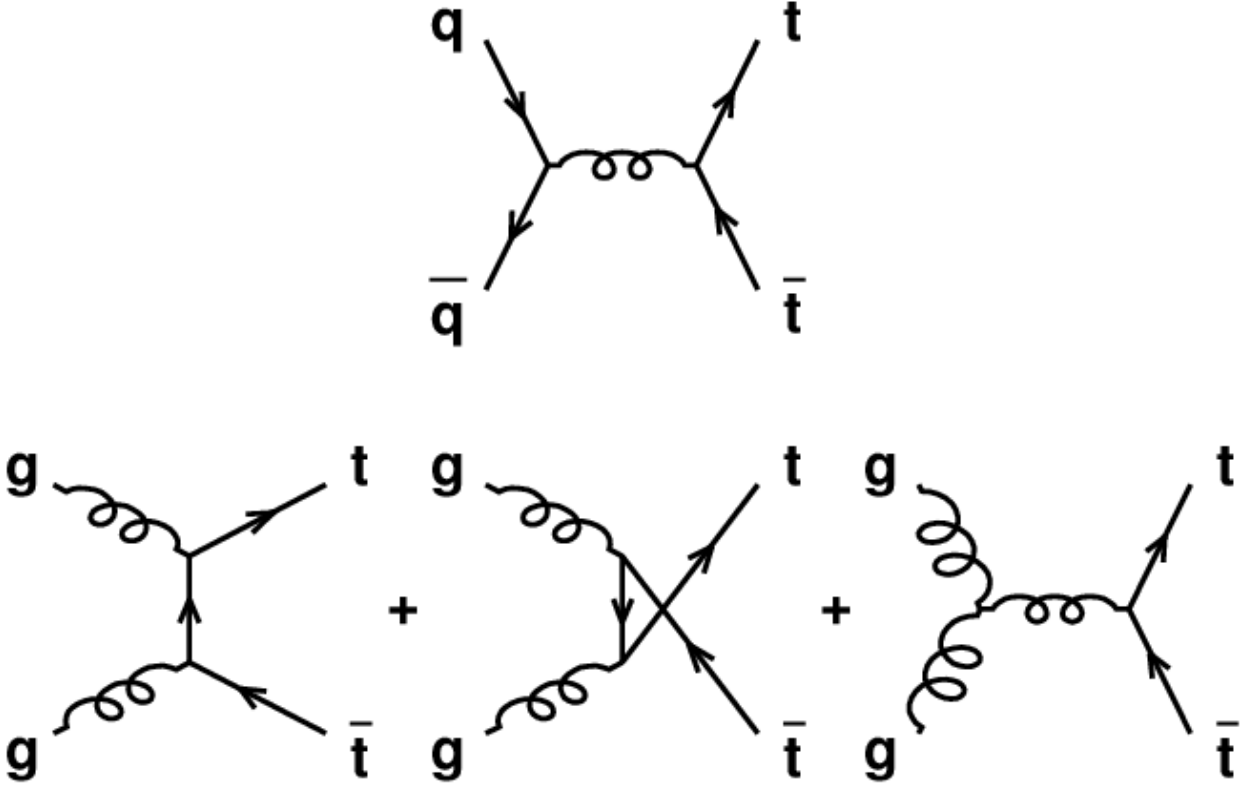
Figure 1.2: Top quark decay

### 1.2.1 Production in Hadron Colliders and Decay

In hadron colliders, the production of a  $t\bar{t}$  pair occurs via  $q\bar{q}$  annihilation or gluon fusion 1.3. At the centre-of-mass energy of 13 TeV, as reached within the LHC at CERN, gluon fusion is by far the dominant process. Top quarks can also be produced in single production, but this has a much smaller cross section. Hence, this analysis focuses on pair produced top quarks and treats single top production as a background process. The top quark decays via the weak interaction with a probability of almost 100% into a bottom quark and a W boson. This property arises from the entry in the CKM matrix where  $|V_{tb}^2| \approx 0.998$ . While the bottom quark is observed as a jet in the detector, the  $W^\pm$  boson further decays into a  $q\bar{q}$  pair, or into a lepton (anti-lepton) and an anti-neutrino (neutrino). Looking at the  $t\bar{t}$  production, this results into three possible channels:

- both W bosons decay into quarks (fully hadronic)
- one boson decays hadronically and the other leptonically (lepton + jets)
- both W bosons decays leptonically (dilepton)

The fully hadronic and lepton+jets channels are dominant and occur at 45.7% and 43.8% of the time respectively. 10.5% of all the  $t\bar{t}$  events results into two leptons in the final state.


 Figure 1.3:  $t\bar{t}$  pair production

### 1.3 Physics of the Proton-Proton Collisions

In collision experiments, the centre-of-mass energy  $\sqrt{s}$  defines the energy that is available to create particles and give them kinetic energy [14], [15]. At the LHC,  $s$  is defined by the momenta of the incoming protons, such as:

$$s = (p_1 + p_2)^2 = p_1^2 + p_2^2 + 2p_1p_2 = (E_1^2 - \vec{p}_1^2) + (E_2^2 - \vec{p}_2^2) + 2(E_1E_2 - \vec{p}_1\vec{p}_2) \quad (1.13)$$

where  $p_1$  and  $p_2$  are the four-momenta of two colliding protons with energies  $E$  and three-momentum vector  $\vec{p}$ . Because the two beams have the same energy but opposite directors, the assumption that  $E_1 = E_2 = E_{proton}$  and  $\vec{p}_1 = -\vec{p}_2$  holds. Given that, the centre-of-mass energy is:

$$\sqrt{s} = 2E_{proton} \quad (1.14)$$

Because protons are composite particles, the actual scattering involves quarks or gluons, called partons, carrying only a fraction of the nominal beam momentum,  $x p_{proton}$ . The centre-of-mass energy of the hard scattering process  $\sqrt{\hat{s}}$  is therefore only a fraction of the stated 13 TeV. Assuming the masses of partons to be small in comparison with the momentum, the centre-of-mass energy of the hard scattering process is:

$$\sqrt{\hat{s}} = \sqrt{x_1 x_2} \sqrt{s} \quad (1.15)$$

Here,  $x_1$  and  $x_2$  indicate the fraction of the proton momentum the two colliding partons carry. Because this fraction is not known, the center-of-mass energy of the hard process is not known either. Therefore, for data analysis one has to use variables that do not depend on the initial momentum in the direction of flight of the partons. Another peculiarity in hadron colliders is the

use of parton density functions (PDFs). PDFs return the probability to find a particular parton inside a given momentum interval in a proton. They are measured in electron-proton scattering experiments [16] and they are crucial to predict production cross sections in proton-proton collisions.

### 1.3.1 Pile-Up

The high collision rate at the LHC is required to collect a large amount of data and be able to observe even very rare processes. But the high rate also has a disadvantage. Since multiple proton-proton interactions take place per bunch crossing, not only one interesting but several other scattering events, mostly soft QCD processes, are seen simultaneously in the detector. This effect is called pile-up [17]. If not corrected for, energy measurements do always include particles not originating from the hard scattering one is interested in. To reduce pile-up effects, it is crucial to be able to identify additional energy as well as distinguish between hard scattering and vertices from additional interactions.

## 1.4 Z' physics

Since its discovery, the  $Z$  boson has been one of the most interesting topics in the high energy physics department [18].  $Z$  bosons are the particles that mediate the weak and nuclear force, and decay into any of the known quarks and leptons except from the top quark. Their versatility is what lead large experiments such as CERN CMS and ATLAS to investigate into the spectrum of different phenomena by making  $Z$  bosons.

One factor that excites the interest of physicists is any indication whether multiple generation that are observed in quarks and leptons are also applied to force-mediating bosons. Thus far, each of the force-carrying particles of the SM ( $W$  and  $Z$  bosons, photons and gluons) seem to be unique. There is no evidence yet to exist that even bares a hint at generation in the force-mediating bosons. On the other hand, there are several theories that predict the existence of heavier bosons of the  $W$  or  $Z$  type bosons. Such particles are the  $W'$  ( $W$  prime) and  $Z'$  ( $Z$  prime) [19], [20].

The CMS experiment has searched into the decay of a heavy  $Z'$  boson into a  $t\bar{t}$  pair [12] [21]. This particular decay mode is of much interest because producing top quarks via ordinary Standard Model physics is rather rare, making it easier to isolate and identify collisions in which a  $Z'$  might have been made. A new heavy vector boson would likely be one of the first clearly visible signals for new physics to be detected by an experiment, when a new accelerator switches on and/or higher centre of mass energies achieved.

The Sequential Standard Model (SSM) enriches the SM by depicting a new gauge boson, the  $Z'$  with the same couplings as the  $Z^0$  but a much higher polemass - of the order of TeV. The existing SM structure ( $SU(3)_C \times SU(2)_L \times U(1)_Y$ ) is added of an additional  $U(1)'$  symmetry. In nearly all models considered, this new boson interferes with its SM counterpart leading to an additional degree of deviation from the SM expectation in the mass spectrum preceding the peak. One interesting effect would be in the angular distributions of events in the Collins-Soper frame. The Collins-Soper frame is the rest frame of the  $t\bar{t}$  pair, wherein the two top quarks have equal and opposite momenta; thus, each makes the same angle  $\theta^*$  with the beam direction.



## Chapter 2

# The CMS Experiment at the LHC

The CMS detector (Compact Muon Solenoid) [22] is designed to study proton-proton as well as heavy ion collisions at the Large Hadron Collider (LHC) [23]. The LHC accelerator complex, together with the experiments was built to study particle collisions at energies significantly larger than the ones that have been achieved up to this date.

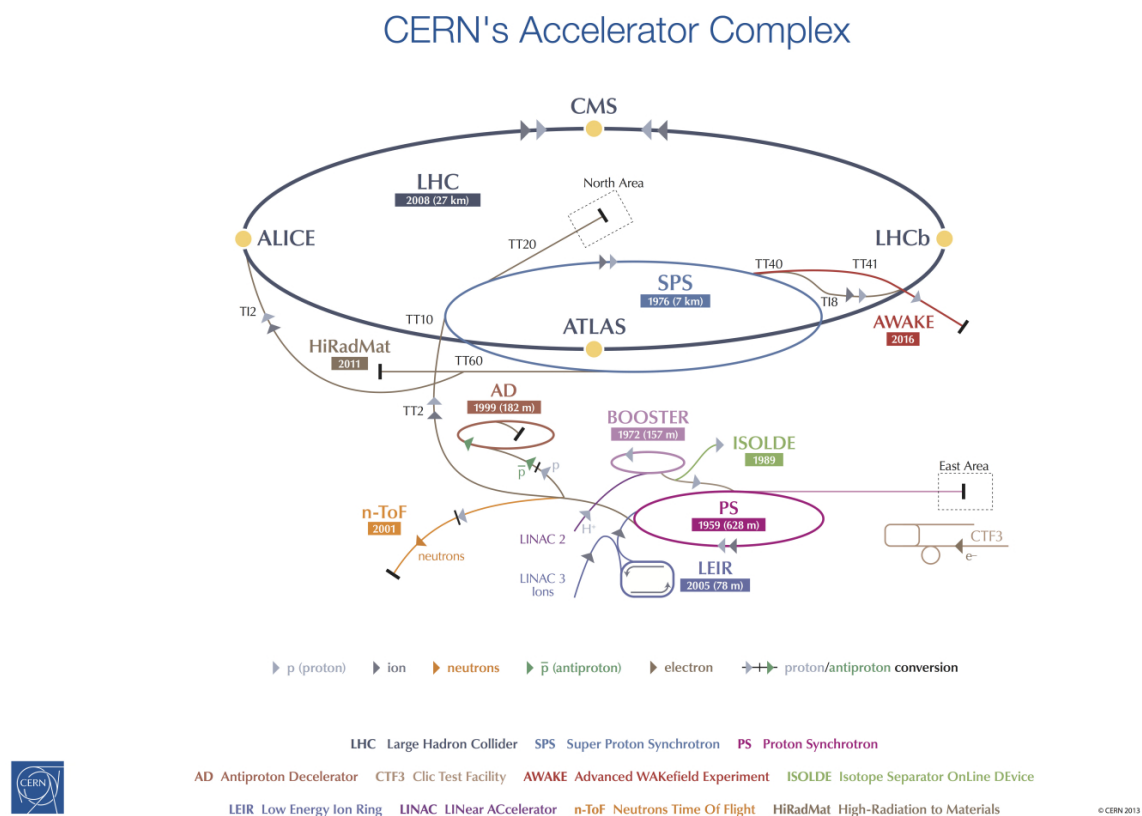


Figure 2.1: The LHC accelerator complex.

## 2.1 THE LHC Complex

The LHC [3] is the newest hadron particle accelerator. It is located within a tunnel which is 100 meters underground (depending on the inclination of the surface varying from 50 up to 175 meters) and 26.7 kilometers in circumference. It is hosted in the same tunnel that was used by the Large Electron Positron (LEP) collider from 1989 – 2000 at the Swiss-French border in Geneva - Switzerland at a laboratory called the “Conseil European pour la Recherche Nucleaire” (European Council for Nuclear Research), CERN. It consists of eight arcs and eight straight sections and its plane is inclined with 1.4 % slope.

The acceleration procedure is split into several stages. In the first stage, protons must be derived from the proton source, which is a simple bottle of hydrogen gas. An electric field is imposed to strip hydrogen atoms of their electrons and yield protons. The produced protons are accelerated in a 50 MeV linear accelerator (LINAC) [24]. Following down the acceleration chain, the 1.4 GeV Proton Synchrotron (PS) [25] booster followed by the 25 GeV PS deliver beam bunches of  $1.15 \times 10^{11}$  protons with 50 ns time difference with each other into the 6.9 km Super Proton Synchrotron (SPS) [26]. The energy reach of the beam bunches up to this point of the acceleration chain is up to 450GeV. Finally the beam bunches are injected into the LHC, which handles the focusing, final acceleration and eventually the collisions at the interaction points of the individual experiments with a bunch crossing rate of 40 MHz on designed energy (14 TeV).

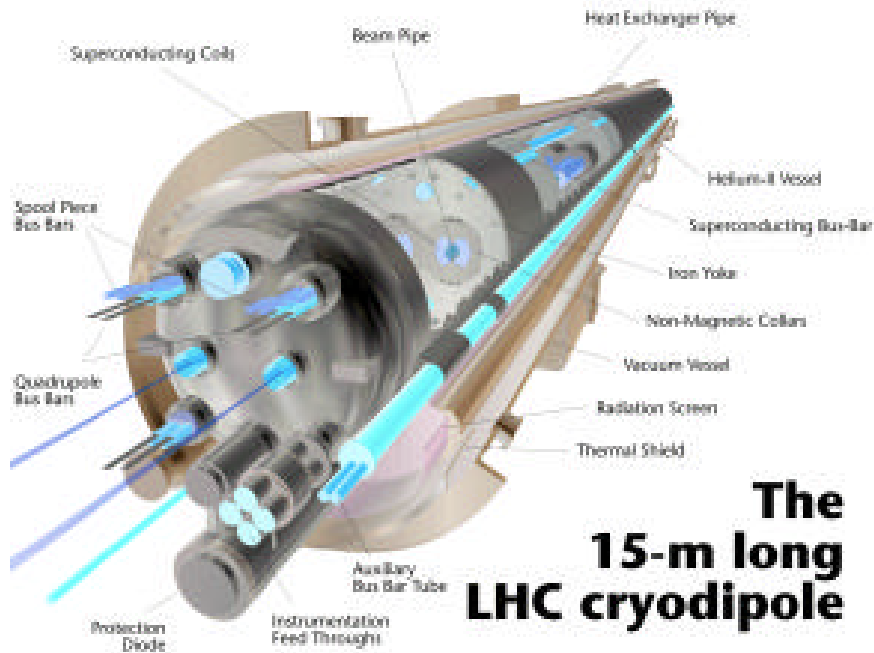


Figure 2.2: This computer-generated image shows some of the parts vital for the operation of these components. The magnets must be cooled to 1.9 K (-270.3C) so that the superconducting coils can produce the required 8 T magnetic field strength.

The beam has a bunch structure with a time interval of 25ns. This structure has been adopted in order to get the number of interactions passing the detector within certain limits, maximize the luminosity and keep track of the beam parameters.

The beam is driven through the accelerator using 9600 magnets of different varieties and size. These include 1232 dipole magnets with 15m length which bend the beams as shown in Fig.2.2 and 392 quadrupole magnets, 5-8 meters long each, which are used for focusing the beam as well as beam corrections. The acceleration of the beam is taken care with the use of 400MHz Radio-Frequency (RF) cavities.

The magnets have Niobium-Titanium (Nb-Ti) coils providing fields up to 8.33 T with a current of nearly 12 kA. The magnet and the beam lines are housed in the same cryostat and the system is cooled down to 1.9 K using super-fluid Helium (He). Just prior to the collision points, another type of magnet is used to "squeeze" the particles closer together to increase the probability of having a collision. The LHC is also capable of accelerating lead (Pb) ions to deliver both lead-lead and proton-lead collisions. It has four interaction points (IP), shown in Figure 2.1, each one accommodating a particle physics experiment. The ATLAS [27] and CMS are general purpose experiments designed for high-luminosity operation and are equipped to study a wide range of phenomena including precise measurement of Standard Model processes and searches for evidence of New Physics. Additionally there are two, more specialised, experiments: LHCb [28] with a focus on  $b$ -physics and ALICE [29], mainly aiming at research of quark-gluon plasma in heavy ion collisions.

A very important parameter of a particle collider is its luminosity. It provides a measure of how many collisions take place per area and second, thus, how well the beam is focused at the interaction points. It is calculated as follows:

$$L = \frac{nN_1N_2f}{4\pi\sigma_x\sigma_y} \quad (2.1)$$

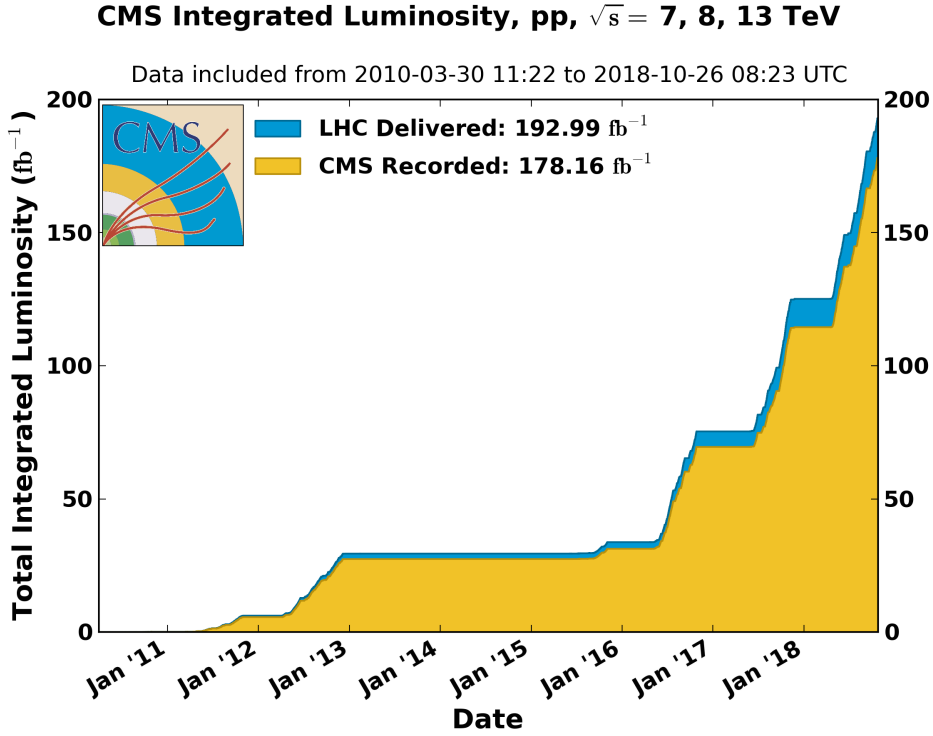


Figure 2.3: Cumulative delivered and recorded luminosity versus time for 2010-2012 and 2015-2018 (pp data only). Delivered and recorded luminosity cumulative over all years during stable beams for pp collisions at nominal center-of-mass energy.

where  $n$  denotes the number of bunches in the accelerator,  $N_1$  and  $N_2$  are the number of protons in the two colliding bunches and  $f$  is the collision frequency. The denominator denotes the cross sectional area where  $\sigma_x$  and  $\sigma_y$  describes the spread of the proton beam in the  $x$  and  $y$  direction respectively. The design luminosity of the LHC is  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  has been already exceeded since the 2016 run. By multiplying the production cross section of a particular process  $\sigma$  with luminosity  $L$  integrated over the data taking time, one gets an estimate of how many events  $N$  of a process took place in one data taking period,

$$L_{int} = \int L dt \quad (2.2)$$

$$N = L_{int}\sigma. \quad (2.3)$$

The LHC has performed very well and produced more collisions than what was anticipated during its desing. The total integrated luminosity for every year is shown in Fig. 2.3

## 2.2 The CMS Detector

The Compact Muon Solenoid (CMS) experiment is a multi purpose detector at the LHC. It is designed to measure momentum and energy of particles produced in proton-proton interactions. With a total weight of 14000 t, the mass of the CMS detector is dominated by a steel return yoke installed to lead the magnetic field originating from a solenoid. The detector has a cylindrical shape

with a length of 28.7 m and a diameter of 15.0 m. The CMS detector is built in onion-like layers of subdetectors, namely tracker, calorimeters, solenoid and muon system, with different purposes described in the following sections. The fact that different particles leave different signatures in the detector is utilised to identify and distinguish particles. The reconstruction of identified particles is further described in detail in chapter 3.

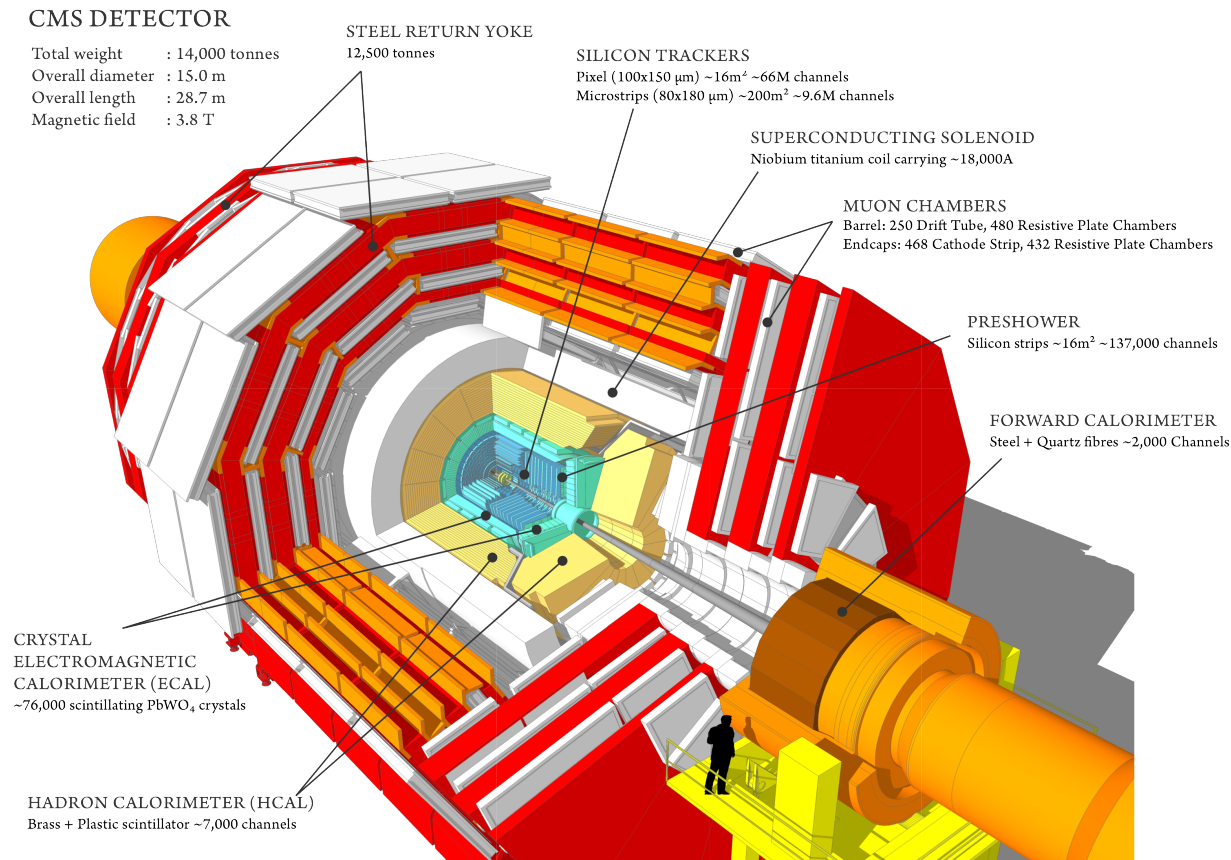


Figure 2.4: Full view of the CMS detector. Basic properties as weight and dimensions are summarised in a table in the left upper corner. Components of the detector are titled and shown in different colours.

Besides the detector systems, a very important component is the trigger, providing fast decisions if an event is discarded or interesting enough to be stored. A trigger system is essential because data cannot be stored with the same rate as collisions occur at LHC. The design of CMS was chosen to cover a wide range of physics approaches. Nevertheless, a focus was set on the discovery of the Higgs boson, which was announced in 2012 by ATLAS [4] and CMS [3]. Components, design and the general construction of the CMS detector are shown in Fig. 2.4.

### 2.2.1 Coordinate System

The coordinate system used in the CMS experiment is based on cartesian and right-handed coordinates. The origin is set in the centre of the CMS detector. Fix points are set to define the direction of the axes. The  $x$ -axis points in the direction of the centre of the LHC ring, the  $y$ -axis points up and the  $z$ -axis is defined parallel to the beam axis. Because LHC is a hadron collider and the  $z$ -momentum of the centre-of-mass system of a collision is unknown, a different set of coordinates is used in data analyses. A first approach are spherical coordinates with the angles  $\phi$  and  $\theta$ . The angle  $\phi$  is defined as the angle in the  $x - y$  plane measured from the  $x$ -axis. The angle  $\theta$  describes the angle from a given point in the  $xyz$  space to the beam axis ( $z$ -axis).  $\Theta$  is a very useful variable to measure distances in experiments like CMS because of likely boosts in



the  $z$ -direction. Therefore, a new variable is introduced, pseudorapidity  $\eta$ , whose differences are Lorentz invariant under longitudinal Lorentz boosts,

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \quad (2.4)$$

The distance  $\Delta R$  between two objects  $(i, j)$  is calculated using the differences  $\Delta\phi = \phi_i - \phi_j$  and  $\Delta\eta = \eta_i - \eta_j$ , thus

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad (2.5)$$

Also, when calculating momenta, it is once again unpractical to use a  $z$  component. Therefore, transverse momentum  $p_T$  is defined which is constructed using the  $x$  and  $y$ -components of the total momentum of an object  $(p_x, p_y)$

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (2.6)$$

A full set of coordinates to describe the momentum and mass of a detected particle are energy  $E$ , transverse momentum  $p_T$ , pseudorapidity  $\eta$  and angle  $\phi$ .

### 2.2.2 Solenoid

The CMS detector is equipped with a solenoid magnet that produces a magnetic field in order to bend paths of electromagnetic charged particles. With the resulting curvature radius and a known magnetic field strength, one is able to calculate the momenta and charge of those particles. Therefore, CMS uses a superconducting solenoid magnet to provide a magnetic field of 3.8T inside the tracking system. The solenoid has a length of 12.9m and a diameter of 5.9m. Tracker and calorimeters are placed within the radius of the solenoid while the muon system is installed outside the solenoid. The iron yokes are used to guide the magnetic flux outside the solenoid. The purpose of the magnetic field on the outside, is to bend tracks of muons within the muon calorimeters. The solenoid form of the magnet is chosen because it provides a constant magnetic flux inside the tracking system.

### 2.2.3 Tracker

The inner-tracking system of the CMS detector consists of the most inner part of the CMS detector, surrounding the interaction point. It is designed to provide precise and efficient measurements of the trajectories of charged particles (electrons, muons, hadrons) and precisely reconstruct primary and secondary interaction vertices. In general, its purpose is to measure the momentum and charge of particles. The tracker's intuition comes from the Lorentz's force which bends the trajectory of charged particles in a magnetic field. The bending radius is proportional to the momentum in a constant magnetic field.

The structure of the tracker system is shown in Fig 2.5 where different modules of the inner tracker are shown:

- pixel
- inner barrel (TIB)
- outer barrel (TOB)
- inner disks (TID)
- endcaps (TEC)

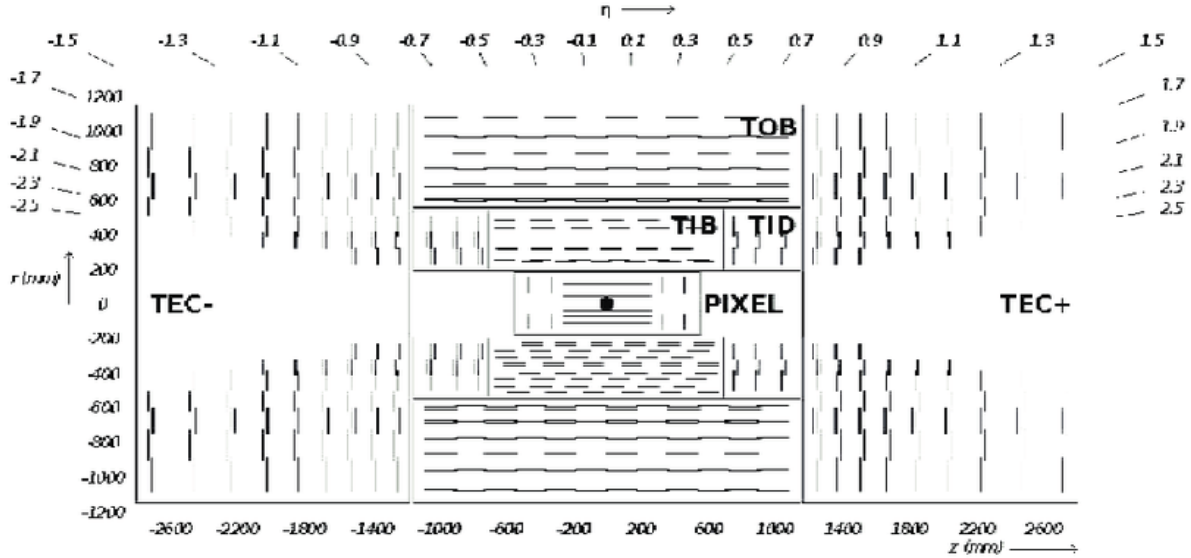


Figure 2.5: Sketch of the CMS tracking detector. All subsystems namely pixel and strip detector with inner barrel (TIB), outer barrel (TOB) and outer endcaps (TEC) are shown. Additionally the range in  $\eta$ , radius ( $r$ ) and  $z$ -direction can be read off.

The CMS tracker [30, 31] is composed of a pixel detector [32, 33] with three barrel layers at radii 4.4 cm, 7.3 cm and 10.2 cm and a silicon strip tracker [34, 35] with 10 barrel detection layers extending outwards to a radius of 1.1 m. Each system is completed by endcaps which consist of two disks in the pixel detector and 3 plus 9 disks in the strip tracker on each side of the barrel, extending the acceptance of the tracker up to a pseudorapidity of  $|\eta| < 2.5$ . Each of the 66 million silicon pixels extends to  $100\mu\text{m} \times 150\mu\text{m}$ .

## 2.2.4 Calorimeters

In order to define a particle type it is crucial that detectors are able to also measure the energy of a particle. Therefore, outside the tracker, a calorimetry system with two types of calorimeters has been installed. The basic principle of the calorimeters is to absorb all of the energy a particle carries with it while moving inside the detector and convert it into a measurable quantity. The energy that is deposited in the calorimeter is transformed into a light signal using scintillating materials. The intensity of the light signal is proportional to the energy deposited in the calorimeter cell. The layers of the detector must be dense enough in order to stop an incoming particle entirely and prevent any leakage. As a measure for calorimeter thickness, radiation length  $X_0$  for electrons, is defined as the path in which an electron loses  $\frac{1}{e}$  of its initial energy. In an analogous way, a nuclear interaction length  $\lambda_n$  for hadrons is introduced, taking into account the length scale of strong interactions. Both  $X_0$  and  $\lambda_n$  values depend on the used material and density of the detector layers. The final step is the collection of the light signal. This is collected and measured with photon multipliers and can be converted back into a measured energy value. Measurements from calorimeters typically get more precise while the energy is increasing. Thus, they provide a good alternative method in cases where the measurement of the tracking system has gotten worse. The resolution of a calorimeter is given by:

$$\frac{\sigma_E}{E[\text{GeV}]} = \frac{\alpha}{\sqrt{E[\text{GeV}]}} \oplus b \oplus \frac{c}{E[\text{GeV}]} \quad (2.7)$$

In this formula,  $\frac{\sigma_E}{E}$  denotes the relative energy resolution depending on the energy  $E$ , the stochastic term  $\alpha$ , the constant term  $b$  which includes inhomogeneities or shower leakage and the noise term  $c$ , containing electronics and pile-up noise. At high energies, the  $b$ -term is dominant but can be reduced by calibration.

### Electromagnetic Calorimeter - ECAL

The electromagnetic calorimeter (ECAL) [36, 37] of the CMS detector is purposed to measure the energy of electrons and photons in a range up to  $|\eta| = 3$ . It is a hermetic homogenous calorimeter which is radiation tolerant and total absorptive made up of tungstate crystals ( $\text{PbWO}_4$ ). It can also measure the EM components of jets and hadrons that deposit their energy in the ECAL crystals.

The operating principle is simple. A photon starts an electromagnetic shower in a electron-positron pair production while an electron starts as Bremsstrahlung radiation. Both develop a cascade of electrons, positrons and photons through the aforementioned processes. In order to measure the energies of photons and electrons with high resolution, ECAL is composed by materials that are ideal for scintillation. Lead tungstate is a inorganic crystal scintillator. Each crystal is 230 mm long, corresponding to  $25.8 X_0$  and has a front area of  $22\text{mm} \times 22\text{mm}$ . With this, photons which are not seen in the tracker can be reconstructed with a high precision and even resolve nearby showers. ( $\text{PbWO}_4$ ) has a Molière radius which is defined as the radius of a cylinder containing 90% of the shower's energy depositions, that is only 2.2 cm. This results to an excellent position resolution and separation between showers. This property is of particular importance since it allows to distinguish photons from isolated neutral pions ( $\pi^0 \rightarrow \gamma\gamma$  decays).

ECAL also gives the ability to provide information that are strongly related to the triggering system of CMS and is aiding particle identification.

In terms of positional structure, ECAL is divided into the barrel covering a pseudorapidity range of  $|\eta| < 1.479$  while the two endcaps cover the area at  $1.479 < |\eta| < 3.0$ . The endcap parts also include a preshower detector that covers the region  $1.65 < |\eta| < 2.6$ . Figure 2.6 depicts the ECAL geometry with its different components. The ECAL, forms a layer between the tracker and the HCAL detector modules. The cylindrical “barrel” consists of 61,200 crystals formed into 36 “supermodules”, each weighing around three tonnes and containing 1700 crystals. The flat ECAL endcaps seal off the barrel at either end and are made up of almost 15,000 further crystals.

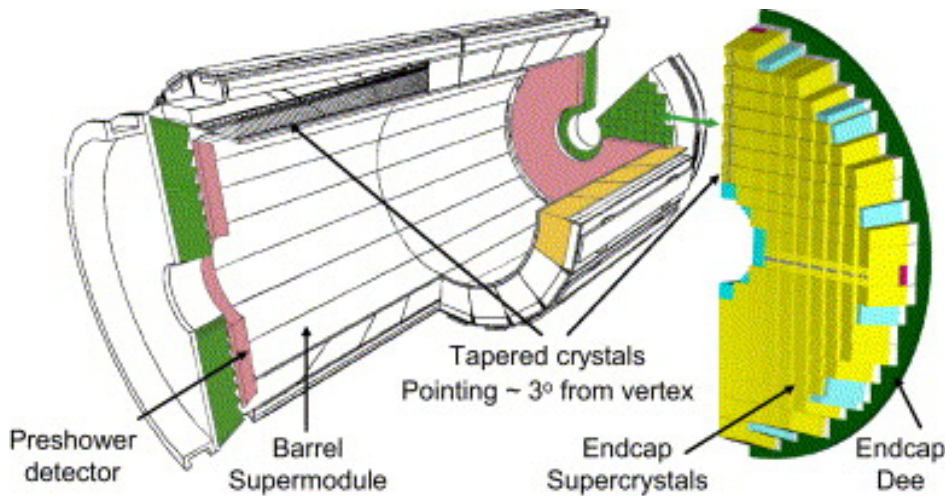


Figure 2.6: CMS ECAL geometry schema showing different components.

The preshower is installed in front of each ECAL endcap. It consists of two layers of lead radiation followed by silicon strip sensors. Its aim is to identify neutral pions within the region of  $1.653 < |\eta| < 2.6$ . Also, it helps the identification of electrons against minimum ionizing particles and improves the position determination of electrons and photons with high granularity.

The design choice of the ECAL is driven by the fact that a Higgs boson decays into two photons where the Higgs boson mass can be reconstructed from the four-momenta of the photon pair. A performance study [38] showed an energy resolution of less than 2% for 45GeV of electrons from Z boson decays in the barrel region. The energy resolution was found to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + 0.3\% \quad (2.8)$$

### Hadronic Calorimeter - HCAL

Since hadrons have a larger absorption length they mostly pass the electromagnetic calorimeter and are absorbed and measured in the hadronic calorimeter (HCAL) [39]. Two alternating layers of different materials are used. One is to absorb the majority of energy an incoming particle carries and consists of brass plates. The other one is made out of a plastic scintillator, that acts as active material. HCAL can provide supplementary measurements to the ECAL and the muon systems. The primary hadrons that have sufficiently long lifetime to traverse the CMS calorimeter are pions, kaons, protons and neutrons.

The energy resolution of HCAL [40] is given by:

$$\frac{\sigma}{E} = \frac{110\%}{\sqrt{E}} \oplus 9\% \quad (2.9)$$

The leading contribution to the HCAL energy resolution comes by effects from not fully containing the hadronic shower, while a stochastic noise term  $S$  of 110% and a constant term of 9%. Figure 2.7 shows a view of the HCAL in the  $y-z$  plane. The HCAL is composed of a barrel ( $HB$ ) and endcap ( $HE$ ) component, which are both contained inside the solenoidal magnet, and the outer ( $HO$ ) and forward ( $HF$ ), which are located outside the solenoid. The ( $HB$ ) covers the range  $|\eta| < 1.3$  while the  $HE$  covers  $1.3 < |\eta| < 3.0$ . Due to the limited space available for the HCAL within the solenoid, the  $HO$  is included outside the solenoid in order to increase the total interaction length.

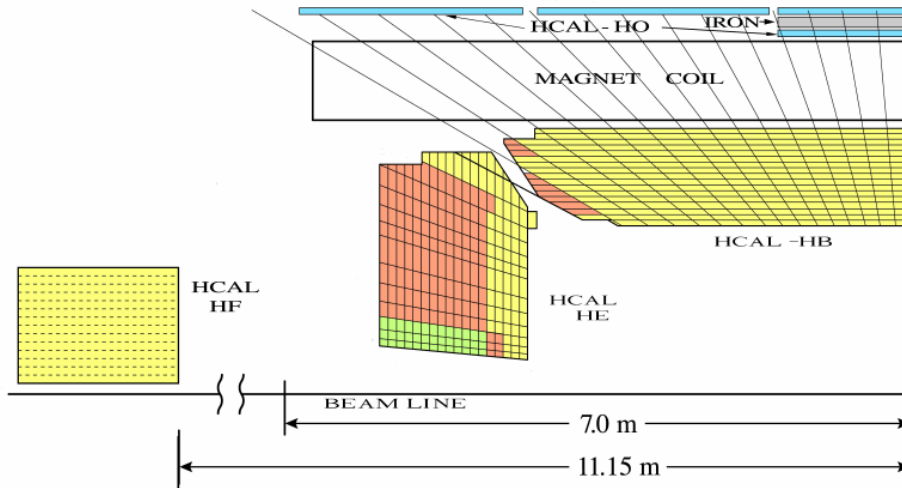


Figure 2.7: View of the CMS hadronic calorimeter in the  $y-z$  plane.

The barrel part of the HCAL begins in a distance of  $r = 1.77m$  to the interaction point and reaches up to the solenoid at  $r = 2.95m$  covering a range of  $|\eta| < 1.3$ . The thickness corresponds to about ten times the hadronic interaction length. Endcaps are installed to cover a range of  $1.3 < |\eta| < 3$ . Additionally, forward calorimeters at a distance of 11.1 m in  $z$ -direction to the interaction point are used to cover the range of  $3 < |\eta| < 5$ . A further outer calorimeter is installed outside the solenoid to account for and measure leakage.

### 2.2.5 Muon System

Because of the possibility to reconstruct muons with very high precision, they played an important role in the design of CMS since they also appear in an important channel of the Higgs boson decaying to four leptons. Thus, this is the reason why a whole system is installed to identify and measure muons with such precision. In general, muon detection is a powerful tool in terms of recognizing signatures of interesting processes over a high background rate that is expected in the high luminosities within the LHC.

The muon system [41] is the outermost component of the CMS detector, located outside the solenoid magnet, consisting of a barrel ( $|\eta| < 1.2$ ) and two endcaps ( $0.9 < |\eta| < 2.4$ ) which are visualized in figure 2.8.

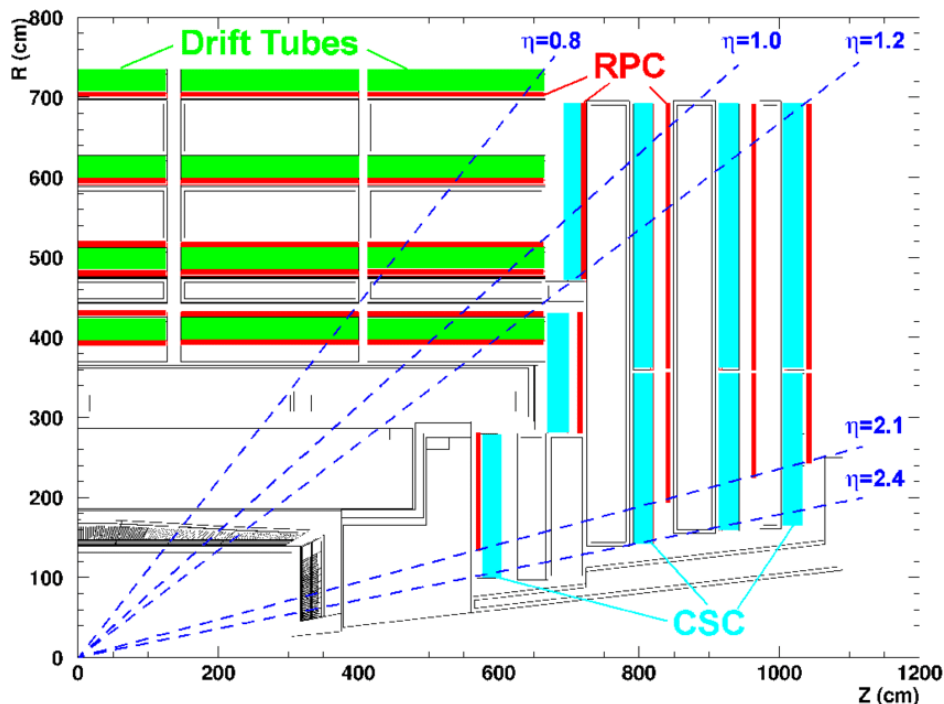


Figure 2.8: Cross-sectional view of a quadrant of the CMS detector in the  $(r, z)$  plane, showing the layout of the muon detector.

As a muon tracker various versions of gaseous detectors are installed. In the barrel part, drift tubes and resistive plate chambers are embedded in the iron return yoke of the magnet where a magnetic field of about 2 T is present. In addition, two endcaps with resistive plate chambers (RPCs) [42] in combination with cathode strip chambers (CSCs) [43] are installed. The design choice is due to the expected flux of muons in the different regions of the detector. While muons occur rarely in the barrel part, they are predicted to appear in high rates in the endcaps. The two gaseous detectors can also provide independently trigger information based on the  $p_T$  of muons with high efficiency and background rejection. The RPCs are installed both in the barrel and in the endcaps covering a region of  $|\eta| < 1.6$ . Also they provide a rapid and highly-segmented trigger with a sharp  $p_T$  threshold. They also help to resolve ambiguities while attempting to reconstruct tracks from multiple hits in a chamber.

## 2.2.6 Trigger

At the interaction points, proton bunches are brought to collision every 25ns. This means that almost  $8 \times 10^8$  interactions per second take place per second. Thus, it is impossible to store that amount of data with this rate while the total required storage capacity is either not feasible. For an accurate measurement of the event, all the information of the subdetector systems is necessary. It would need a tremendous amount of computing cores and computing power in general to analyse this data. In addition, only a small fraction of events are interesting for physics analyses. Therefore, the CMS detector is equipped with a trigger system [44] that is required in order to reduce the rate of the data acquisition by a factor of the order  $10^6$  to a few 100 Hz. This is achieved with two trigger stages:

1. Hardware based Level-1 Trigger (L1) [45, 46]
2. Software based High-Level Trigger (HLT) [47].

## Level-1 Trigger

The first level trigger is composed by custom hardware and uses important fast data coming from the muon and calorimeter subdetectors. It is responsible for identifying and selecting events that may contain muons, electrons, photons or jets at a rate up to 100 kHz within  $4\mu s$ . The first level trigger is further divided while having local, regional and global components. The local triggers (Trigger Primitive Generator) identifies energy depositions in calorimeter trigger towers and tracks segments or hit patterns in muon chambers. The regional triggers define ranking and sorting of trigger objects such as an electron or a muon candidate by combining their information and using pattern recognition. The Global Calorimeter (GCT) and Global Muon Triggers (GMT) determine the highest-rank calorimeter and muon objects across the entire experiment and transfer them to the global trigger (GT) which decides the rejection or the acceptance of an event for further evaluation by the High Level Trigger (HLT).

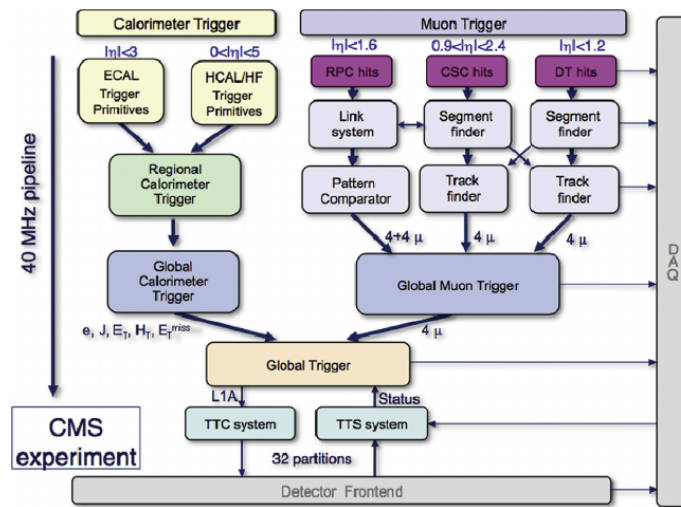


Figure 2.9: The architecture of the CMS Level-1 trigger system in Run 2.

## High Level Trigger

The second level also known as high-level trigger (HLT). It uses a partition of the full event reconstruction framework in order to process data in a very fast way in a farm of commercial processors. It is able to reduce the event rate up to 1kHz. The HLT has access to the complete read-out data and can therefore perform complex calculations similar to those made by the off-line analysis software. Like the level-1 trigger, the HLT is further divided into processing "steps", namely the L-2, L-2.5, L-3 steps. The L-2 step uses input originating from the calorimeter and muon detectors whereas the L-2.5 uses information originating from tracker such as pixel hits. The L-3 step refers to a selection that is able to include the fully reconstructed tracks in the tracker. The HLT is able to process and make decisions upon interesting data with a more efficient way in relation to the Level-1 trigger. It can include information from b-tagging discriminators and reconstructed jets. The next step of the processed data that pass both L-1 and HLT requirements is to store this data to physical memories (Hard drives) to perform offline processing which can take several seconds per event and performs a much more computational intensive event reconstruction which is then used for data analysis.

### 2.2.7 Upgrades

The first CMS *Phase 1* upgrade started at the end of Run1 and it is planned to continue until the start of LHC Run3 in 2021. It included a complete upgrade of the Level-1 trigger system, an upgrade of the photon sensors of the HCAL and HF scintillation system, as well as, an installation of a new pixel detector that replaced the old one in early 2017. The new pixel detector [48] contains four barrel layers and three discs in the endcaps, with double pixels. It is expected to be more

robust against high pile-up and to have a faster readout system with an increased number of channels.

The CMS *Phase II* upgrades [49] will replace or improve detectors and their associated hardware to provide the appropriate performance that is required to confront the challenging conditions of high luminosity when the HL-LHC runs begin. The installation of the upgraded systems has already begun and is planned to be completed during the LS3, presently scheduled for 2024 up to 2026 [50]. The proposed operating scenario is to level the instantaneous luminosity at  $5 \times 10^{34} \text{cm}^2 \text{s}^{-1}$  from potential peak value  $2 \times 10^{34} \text{cm}^2 \text{s}^{-1}$  at the beginning of the fills, and to deliver  $250 \text{fb}^{-1}$  per year for a further 10 year of operations. Under these conditions the event pile up will rise substantially and thus, the *Phase II upgrade* is crucial for the smooth operation of the detector.





## Chapter 3

# Object Reconstruction

At general purpose particle detectors, such as CMS, the output signal must be converted into objects with a physical meaning in order to perform studies and analyse recorded data. Thus, the CMS collaboration has developed a number of algorithms to identify, reconstruct and measure particles. The particles are returned in the form of lists of electrons, muons, photons and jets. The candidates for each particle type has to maintain certain criteria in order to be categorized as such. In this chapter, both the criteria and algorithms used to obtain objects for physics analyses are described.

### 3.1 Particle Flow Algorithm

The CMS detector is a detector which consists of many detector layers nested around the beam axis. Each detector layer provides information for a specific type of particle type. CMS uses a special algorithm, called Particle Flow (PF) [51] to combine information coming from every detector system, following the complete path of a particle through a detector. This algorithm is able to make use of the footprints of different particles passing through different layers of the detector. Muons are detected both in the inner tracker and the muon chambers, electrons are detected both in the inner tracker and ECAL detector, while charged hadrons leave tracks and shower in the HCAL. Photons and neutral hadrons only deposit energy in the ECAL and HCAL respectively. A schematic view of all detector layers is shown in Figure 3.1. The PF algorithm is able to capture and utilise all the aforementioned information in order to assign all showers from charged particles to their tracks, leaving untracked only the energy deposits originating from photons and neutral hadrons.

First, hits in the inner tracker and muon system are combined. If a track through the whole detector is found, the object is called a muon candidate and all belonging hits are removed from the algorithm input. Afterwards, the remaining tracks from electrons and charged hadrons are extrapolated into the calorimeter systems and assigned to showers. Here, photon radiation from electrons has to be treated with, since energy from these photons also need to be assigned to the initial electron. Finally, remaining showers in the ECAL and HCAL should originate from photons and neutral hadrons, respectively. To be able to use particle flow with high efficiency, the tracking systems need to have a good spatial resolution while the calorimeters also have to provide a high granularity. All these requirements are fulfilled in CMS, which makes the PF algorithm a powerful tool in object identification and reconstruction.

### 3.2 Particle Identification and Reconstruction

#### 3.2.1 Muon Identification

Muons have a very low probability to ionize CMS calorimeters. Thus, muons are identified by their hits in the inner tracker and the outer muon chambers. This results into a high identification rate while also maintaining a good reconstruction precision. In order to further improve the muon identification, additional criteria are applied to standalone, global and tracker muon candidates.

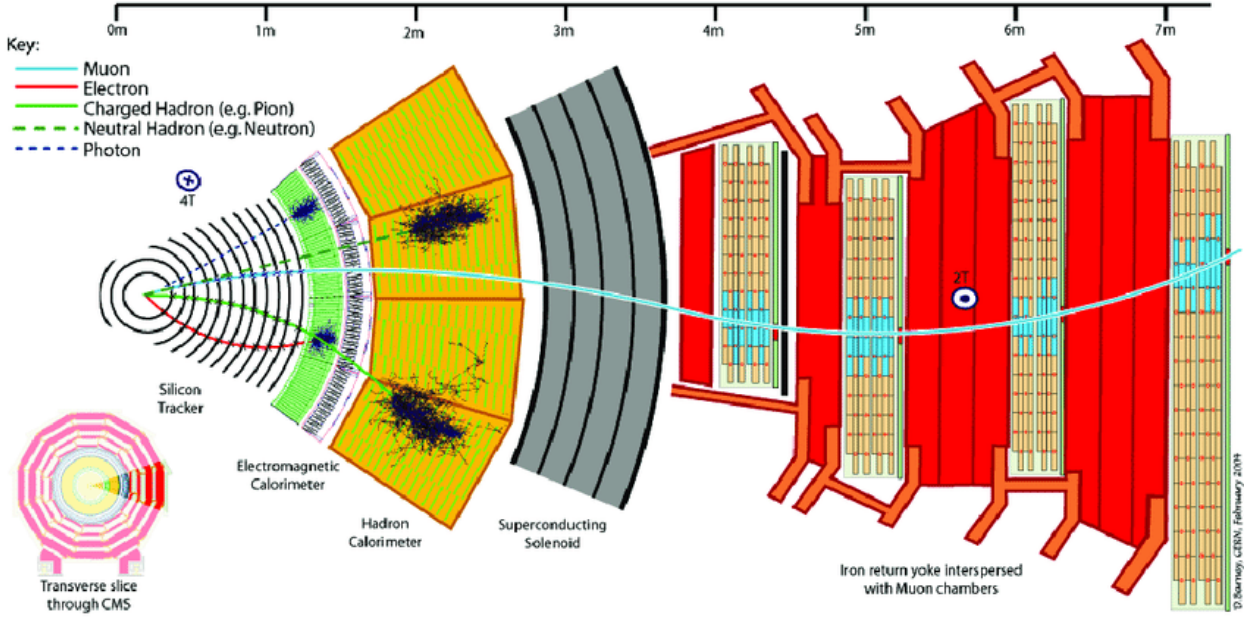


Figure 3.1: Slice through the CMS detector looking in the direction of the beam pipe. Tracks of different particles on their way through the layers of the detector are displayed. From left to right, tracking system, calorimeters, superconducting solenoid and muon system are shown.

Those are based on the various quality parameters referring to muon isolation and reconstruction and are used to suppress misidentified muons, mostly charged hadrons, while preserving a high efficiency for both isolated muons and muons inside jets.

### 3.3 Jet Algorithms

Jets are objects that are used in order to reconstruct high momentum partons that are produced in the proton-proton collisions. Because of confinement, quarks and gluons cannot exist isolated but hadronize to a cascade of colour neutral hadrons. Therefore, one needs to sum up all hadrons originating from the initial parton in order to reconstruct it. For these types of particles, jet clustering algorithms are used to combine particle flow candidates in a well defined way. Each defined jet must be infrared and collinear safe. The first criterion means that a jet should not change when including or excluding soft radiation, while the collinear safety criterion addresses the collinear splitting of particles in a jet which should also not change the jet.

#### 3.3.1 Anti- $k_T$

In the CMS experiment, the most common way to cluster jets from the detected particles is to use iterative jet algorithms. This means that PF candidate particles are clustered step by step until an abort criterion is reached using a group of jet algorithms called  $k_T$ -like [52] algorithms in the FastJet Package [53]. The anti- $k_T$  algorithm [54] works in the following manner. As an input, it uses a list of all reconstructed PF objects from the detector. It measures the distance between a PF candidate  $i$  and the surrounding particles  $j$  in the rapidity  $y$  and angle  $\phi$  plane. The algorithm then calculates two quantities,  $d_{ij}$  and  $d_{iB}$  for each pair of objects  $i$  and  $j$ :

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.1)$$

$$d_{iB} = p_{Ti}^2 \quad (3.2)$$

where:

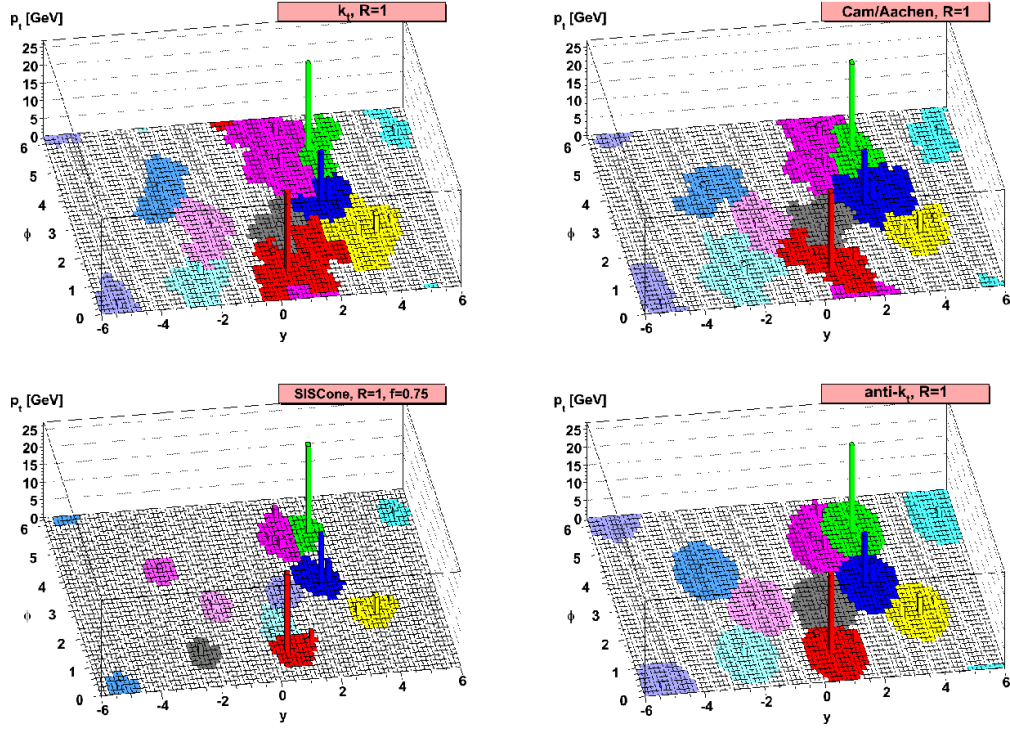


Figure 3.2: A sample parton-level event (generated with Herwig [8]), together with many random “ghosts”, clustered with four different jets algorithms, illustrating the “active” catchment areas of the resulting hard jets. For  $k_t$  and Cam/Aachen the detailed shapes are in part determined by the specific set of ghosts used, and change when the ghosts are modified.

- $\Delta R$  is defined as the distance of two objects  $ij$  in the rapidity-angle  $\phi$  plane and is:  $\Delta R_{ij} \equiv \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$
- $p_{T,i}$  is the transverse momentum of the  $i$  object
- $R$  is a constant parameter that defines the radius of the resulting jet
- $d_{iB}$  is defined as the beam distance

For each particle the algorithm calculated the beam distance  $d_{iB}$ . The clustering procedure continues by finding the minimum  $d_{min}$  of all particle distances  $d_{ij}$  and beam distance. If the smallest distance found for particle  $i$  with a particle  $j$  is  $d_{ij}$  the particles are combined to form a single entity. On the other hand, if the smallest distance satisfies  $d_{iB} < \min(d_{ij})$ , the particle  $i$  is considered a jet and is removed from the list of entities. The aforementioned step is repeated until no particles are left. The anti- $k_T$  algorithm outperforms similar algorithms such as  $k_T$ , Cambridge/Aachen and iterative-cone algorithms. This algorithm is the standard jet clustering algorithm used in the CMS experiment and results in particles with high  $p_T$  since the low- $p_T$  particles are considered subparticles of a jet.

The distance  $R$  is chosen based on the use cases of the resulting jet collection. The default jet collection in CMS consists of jets with small radius  $R$  of 0.4 and is referred as the “AK4” jet collection. These types of jets typically correspond to a single quark or a gluon originating from the hard-scattering process. Jet collections that use larger radii such as 0.8 (“AK8” jets) are used in order to reconstruct hadronically decaying particles with large Lorentz-boost, such as high- $p_T$  top quarks. High- $p_T$  particles decaying products can be collimated and thus may be reconstructed into a single- $R$  jet making it even more difficult to further distinguish particles lying within the large radius AK8 jet.

### 3.3.2 Soft Drop

When clustering particles into jets, algorithms cannot estimate whether a particle originates from the hard process or from pileup, underlying event or initial state radiation. Jet substructure techniques are further applied to remove soft radiation (grooming) in order to obtain a jet that in the best case scenario includes only energy from particles of the hard scattering process. The largest impact on jet mass measurements has wide angle radiation which is still reconstructed in the same jet. This results into extra particles like light quarks or gluons clustered within a jet that may be originating from a top quark decay and as a consequence adds more mass the larger the distance between top quark and light parton/gluon is. The CMS experiment uses the SoftDrop algorithm [55] in order to mitigate this effect. This method undoes the last clustering step, breaking the final jet into two constituents. After each de-clustering step, the softdrop criterion is checked:

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \left( \frac{\Delta R_{12}}{R_0} \right)^\beta \quad (3.3)$$

If the condition does not hold, the constituent with lower transverse momentum is removed, while on any other case the final jet remains unchanged. The  $p_{T1}, p_{T2}$  denote the transverse momentum of the two constituents while  $R_{12}$  denotes the distance between the constituents in the  $y - \phi$  plane. The parameter  $R_0$  is the radius of the final jet. A threshold value  $z_{cut}$  combined with an exponent  $\beta$  define the strength of the grooming. Performance studies within the CMS experiment have been done [56] resulting into values of  $z_{cut} = 0.1$  and  $\beta = 0$ .

### 3.3.3 PUPPI

An additional factor that may contaminate the reconstructed jets is additional energy attributed to jets that comes from the proton-proton interactions other than the hard-scatter event at the primary vertex. This additional energy from these events is called *pileup* and CMS has developed its own algorithms to counter-effect it. If a charged hadron is reconstructed in the tracker and is identified as originating from a pileup vertex, it is removed from the collection of particles used to form physics objects. This procedure is widely used in jet reconstruction and is referred to charged hadron subtraction (CHS) [57, 58].

An alternative way to mitigate pileup is the *pileup per particle identification* known as PUPPI [59]. In this scope, each reconstructed particle is applied a weight that refers to the probability that this particle originates from pileup interactions. The probability is based on local energy distribution around the particle, tracking information and event pileup properties. It uses this estimation to scale the four momentum of each PF particle candidate before it is clustered in to a jet.

## 3.4 Heavy Flavor Jets

In analyses involving top quarks, it is useful to identify bottom quarks since they are in almost every case one of the top quark decay products. The properties of the b-quark makes it possible to find and tag jets originating from a bottom quark. Since b-hadrons have a relatively large lifetime of about 1.5 ps, combined with time dilatation due to their velocity, they are likely to travel several millimetres in the detector before decaying. This leads to secondary vertices at the spatial point where the b-hadron decays. The CMS tracking system is able to resolve both primary and secondary vertices. Up to now, these characteristics were exploited by the *Combined Secondary Vertex (CSV)* [60, 61] algorithm and its updated version. The fraction of energy from charged hadrons in a jet and the invariant mass of all particles assigned to a secondary vertex are taken into account to distinguish them from non b-jets. The output of the CSVv2 algorithm is a number between 0 and 1 denoting how b-jet like a jet is (1 means very b-jet like). Different working points correspond to different minimum values that a jet must fulfil to be considered as a b-jet. In this analysis though not only the CSVv2 algorithm was used. CMS has developed yet another heavy flavor jet identification algorithm called the *DeepCSV* [62, 63]. The *DeepCSV* algorithm is a

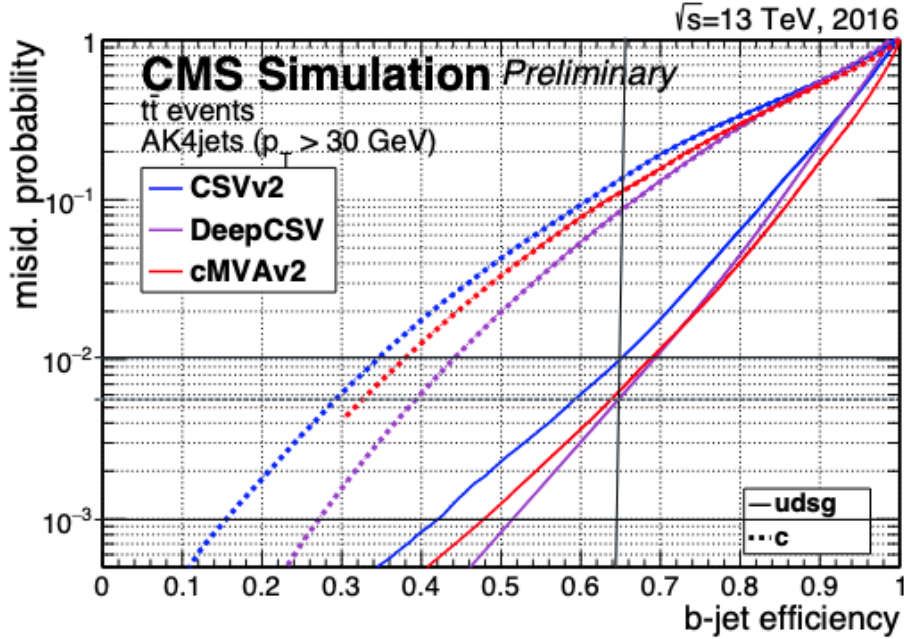


Figure 3.3: Performance of the b jet identification efficiency algorithms demonstrating the probability for non-b jets to be misidentified as b jet as a function of the efficiency to correctly identify b jets [63].

trained deep learning neural network. It classifies particles in each final layer (multi-classification problem) by outputting a single value for each class. Its input includes all the CSVv2 features while also adding few more "relative raw" information such as not only the 2D impact parameter significance but also its value. Also, *DeepCSV* algorithm uses more tracks in relation to the CSVv2 algorithm.

The output of the deep neural network consists of 5 exclusive categories:

- Exactly one **b hadron** in a jet
- Exactly one **c hadron** but no b-hadron in a jet
- Two or more b-hadrons in a jet
- Two or more c-hadrons, but no b-hadron in a jet
- Light jets (udsg)

### 3.5 Missing Transverse Momentum

Even though CMS is a full coverage detector that is able to interact with and reconstruct stable and long-lived particles produced in pp collisions, particles such as neutrinos or hypothetical weakly-interacting particles do not leave a signal footprint to the detector. Thus, their presence can only be inferred through the visible momentum imbalance in the transverse plane. One can define the missing transverse energy  $E_T$  as an estimate of the energy that is carried away by particles which leave the experiment undetected. The summation of all particles in the final state (both detectable and not) returns the transverse momentum ( $p_T$ ) of the system in the initial state in which  $p_T = 0$  at the LHC. Thus the transverse energy of all undetected particles is defined as the absolute value of the negative sum over the transverse momentum of the detected objects

$$E_T = \left| - \sum_{leptons, jets} p_T \right| \quad (3.4)$$



## **Part II**

# **CMS Detector Control System**





## Chapter 4

# Control Systems

This chapter focuses on the description and techniques that are used in designing and implementing a control system for the CMS detector. Also, a general introduction into the tools that are used to build such control systems is provided.

### 4.1 SCADA systems

The Detector Control Systems (DCS) belong in the subcategory of *Supervisory Control And Data Acquisition* (SCADA) [131] systems. SCADA are systems that are designed and implemented in order to monitor and control systems. The term data acquisition refers to the reading of values regarding the monitoring of the system which is performed in a very low rate, contrary to ex. physics experiments where the rate of data acquisition is very high (DAQ). For this reason, the SCADA systems are also called *SLOW Control SYstems* (SLOCSY). The described SCADA or SLOCSY systems are not only used in High Energy Physics, but also in a wide variety of factories, airports and many more.

A SCADA system consists of three layers:

- **Supervisory layer:** Visualization, control and sending of commands. Also, responsible for gathering data from all processes.
- **Front-end layer:** Between supervisory layer and hardware. Reads signals from hardware and sends them to the supervisory layer
- **Communication layer:** Intermediate layer between front-end and supervisory layer, while being responsible for their communication

SCADA systems are usually not complete control systems but are software toolkits used in order to develop the supervisory layer of a control system and also have the ability to connect to the front end layer. The importance of SCADA systems is automation, while allowing an organization to carefully study and anticipate the optimal response to measured conditions. Also, it allows to execute these responses in an automated manner.

### 4.2 WinCC\_OA Toolkit

CERN chose the product it would use for the development of control systems for its experiments. While many commercial and open-source products were available, the SCADA toolkit selected was the PVSS-II provided by the Austrian company ETM. In the long run, ETM was purchased by Siemens and the PVSS-II was renamed into WinCC\_OA [132].

#### 4.2.1 Main Features

WinCC\_OA is a software that is designed for being utilized in the field of automation engineering. Its main application is the operation and supervision of technical installations using full-graphics capabilities for the end-user and its operators. The toolkit is able to transfer input values and

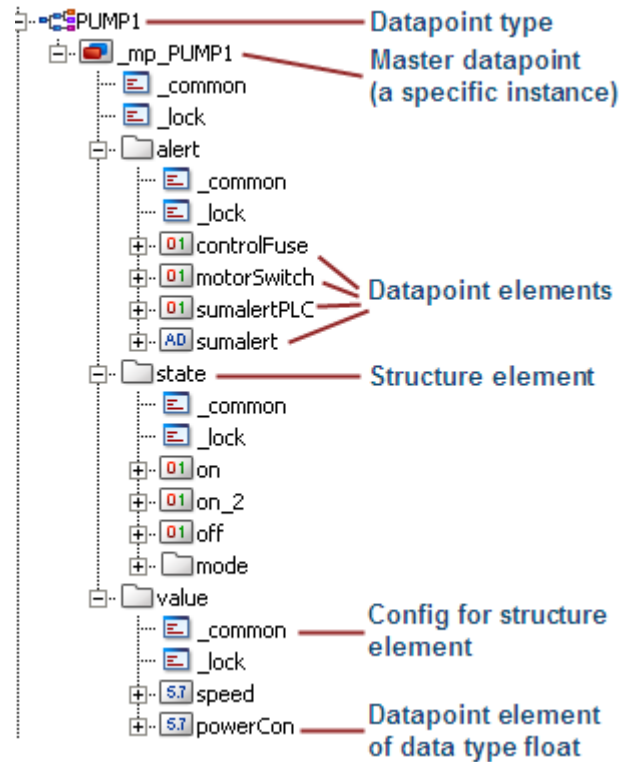


Figure 4.1: Schematic representation of a data point.

commands to the process and its control devices. The software also includes alerts that are sent to the user when critical states occur or predefined limits are exceeded, but also offers archiving features so that data are stored for later display or analysis.

One of the main features on WinCC that made CERN select it as the tool to develop its SCADA system was its ability to scale, while it enables to create distributed systems. A distributed system is not one system but a set of individual systems which are connected to each other in a hierarchical way forming a complete control system. In this way, the computational load is divided among many machines making the design and operation of a very large control system sustainable.

Another advantage of WinCC\_OA is that its equipped with a run-time database. WinCC\_OA by default comes with an internal built database, which schema consists of data points (DP), an abstract structure of data, as seen in Fig. 4.1. A data point structure is constructed from data point elements (DPE) which hold the values and states of a given data point. In addition to storing the value, a data point element includes attributes such as time stamp, quality information and originator. Each data point holds a certain type called data point Type (DPT) which gives the data point specific characteristics. The DPT defines the general structure of the data point as well as the name and type of its elements (Fig. 4.2). In this way, all data associated with a specific device are grouped together and are held in the same type of variables, thus enabling re-usability of code and components developed for these types of devices. In addition, there is a custom WinCC Application Programming Interface (API) which can be used to write WinCC managers and drivers.

WinCC\_OA can be easily scaled up and extended through the native programming language that is provided by the toolkit. The programming language that is provided by WinCC\_OA is called ConTRoL (CTRL). It is a C-like programming language used for creating both scripts as well as user interfaces (panels). The language is user friendly and easy to learn which allows new users to quickly get familiar with the toolkit and contribute with their own scripts and panels to large projects in a quick manner.

Finally, a great advantage that the WinCC\_OA offers is its modularity in terms of operating system. In contrary to most SCADA products, WinCC\_OA is a multi-operating system product offering the same functionality in Windows OS and Linux OS. This is important because, within the scientific community a great amount of software (ex. DAQ) use Linux as an operating system.

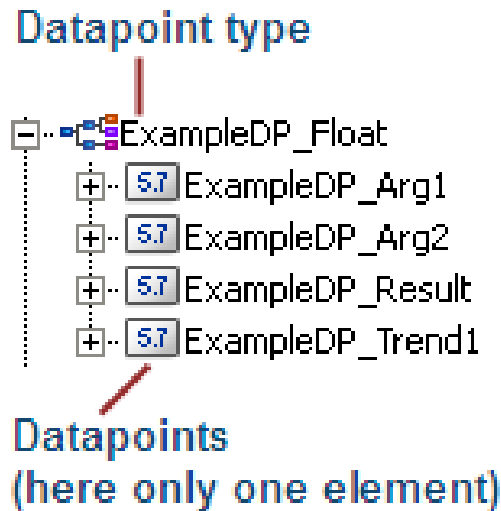


Figure 4.2: Schematic representation of a data point type.

### 4.3 Joint Control Project Framework (JCOP)

The DCS involves the control, configuration, readout and monitoring of hardware devices as well as monitoring of external systems such as the electrical system, cooling system etc. In order for all of the experiments at LHC to work in a coherent and safe way, the Joint Controls Project (JCOP) was introduced in 1998 aiming to provide common solutions for the control system for all LHC experiments.

The goal of the JCOP Framework [133] is to develop and provide components and tools that are used widely for the control and handling of the detectors. In order to perform such as task, JCOP introduces the common basis in order for each experiment to build and work upon.

#### 4.3.1 Finite State Machine (FSM)

The whole pipeline is built on the WinCC software that is provided by Siemens. The modelling of the control system has been implemented as a hierarchy of Finite State Machines (FSM) using the SMI++ toolkit [134, 135].

The complexity and size of the detector is so large that every device or software task needs to be very well organized. Through the SMI++ toolkit, this set of entities is described as a collection of objects behaving as Finite State Machines. To be more accurate, an FSM is a model of the detector objects, where each object can have an well defined, finite number of states where it can be found.

Each functional part is represented in the FSM by a "node", acquiring a "State" of operation but also a "Status" in order to monitor any anomalies. In the model of an FSM type system, there is a state/command interface between a parent and its children. Commands are passed from a parent to its children while states of the children are propagated to the parent. A visual representation of a typical Control System using FSMs can be found in Figure 4.3. Two types of objects are defined in the FSM hierarchy:

- **Control Units (CU):** Logical decision units that monitor the states of their children and report an overall state of their parent. They can model and control the sub-tree below them and typically model abstract or logical items. They can have children of all types but their parent must also be a CU. They derive their states from their children. In order for an FSM to exist, it must contain at least one Control Unit.
- **Device Units (DU)** Typically, they model a specific piece of equipment (hardware). They interface with the lower level components and represent hardware. They do not have logic

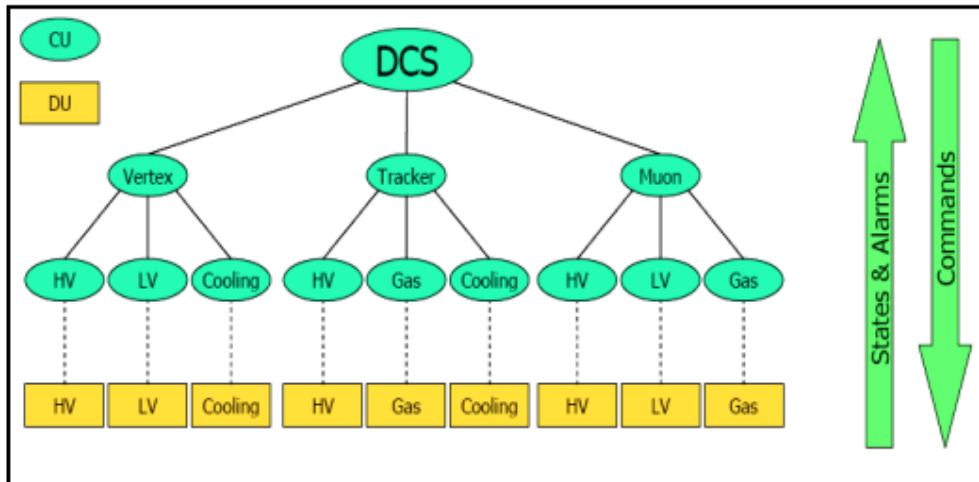


Figure 4.3: A simple control system modelled using FSMs.

behavior.

As mentioned above, all state information propagates from top to low (parent to children) while all status is propagated from low to top (children to parent).

- **State** defines the operational mode of the system
- **Status** gives the details of how well is the system working on a particular state (ex. the state of a HV chamber can be READY but the status can be in WARNING reflecting any anomalies in its values)

### 4.3.2 Archiving Database

One of the most important things regarding the correct and efficient functionality of the DCS is its ability to store and preserve data that relate to the state of the running system. It is crucial to store this type of data not only for monitoring and analyzing the detector's behaviour but also to ensure the quality of the data collected from the experiment. For example, a malfunctioning in a current detector may result into corrupt physics analysis data.

The WinCC\_OA control system is equipped with an internal built-in database that is used for archiving and storage. In order for the WinCC\_OA to be able to store such a large amount of data, with regard to a large physics experiment, Oracle databases are used for the archiving of the data. The archiving policy of the CMS experiment is to archive data *on change* but also using a *polling mechanism*. Regarding the *on change* method, this means that values are archived only if a change of a value is monitored, while *polling mechanism* archives values every certain amount of time regardless the change.

### 4.3.3 Configuration Database

The aim of the Configuration Database, provided by the JCOP framework it to provide means of storage and retrieval of different sets of configuration data that are related to control systems within an external database provided by Oracle. This implemented module includes both static and dynamic configuration data for the WinCC\_OA data points.

When controlling large scale projects such as the CMS detector, it is needed that values of power supplies, or sensor temperatures are changed throughout the lifetime of a project within a reasonable rapid time. Therefore, the Configuration Database addresses such issues by providing the storage of two types of data:

- **Static Data:** Image of the system (data points, addresses, etc)

- **Dynamic Data:** Values and alert limits of data points connected to hardware

Static data is used whenever a system failure occurs and thus restoring of the system is required. The configuration database allows the user to restore a failed system, since the system image is stored within the database. On the other hand, static data are used to configure parameters of the hardware. These parameters are called recipes. They are also used for changing the detector's state.

#### 4.3.4 Access Control

The framework also includes an access control component which is implemented in the UI level. It distinguishes the roles of the various users by allowing different actions to each one of them depending on their expertise, while it ensures the protection from non-malicious actions (such as misuse due to an operator's mistake) by enabling/disabling parts of the User Interface.

To achieve this, the framework and this module in particular defines *domains* that correspond to parts of a system. Each domain has a level of access (e.g. operator, root user, etc). A domain is an abstract term that represents a part of the detector control system, such as a subdetector. The roles are defined in such way so that each role is allowed only a set of current actions for this particular domain. As a result, each user that must perform an action on the current domain, is associated to a specific role.

### 4.4 The CMS Detector Control System

The following section describes the structure of the CMS Detector Control System as well as some tools and modules that were developed during my 2 year Central CMS DCS service as a control systems engineer.

The CMS is an experiment that requires various hardware and software systems to work at the same time in order for the detector to perform with such efficiency. Thus, all these systems are required to be monitored, in order to control and operate the CSM experiment. The interaction of the detector experts and the shifters with the detector hardware is performed with the help of a *Detector Control System* (DCS).

### 4.5 The CMS Central DCS team

The CMS Central DCS Team is responsible for the development and maintenance of the Detector Control System (DCS), the control and maintenance of the FSM (Finite State Machine model) and monitoring (sensors, crates, etc) of the CMS detector.

The CMS central DCS team makes use of frameworks such as the WinCC\_OA provided by Siemens. The WinCC\_OA is further optimized for CERN's detector needs by utilizing and building new tools via the JCOP (Joint Control Project) Framework. Also, the CMS central DCS team also provides a custom made framework for the CMS detector, the "CMS framework". The Central DCS team is responsible for the maintenance, upgrade and development of the CMS framework.

Each CMS subdetector develops and maintains its own system (DCS Subdetector Framework) which is integrated into the central DCS system. Every subdetector DCS must include the following:

- System with unique structure
- Monitor subdetector hardware
- Perform specific subdetector commands

The Central DCS team's role is to provide general rules and guidelines for the smooth integration of the subdetector tools. Also, the Central DCS team provides tools for the development

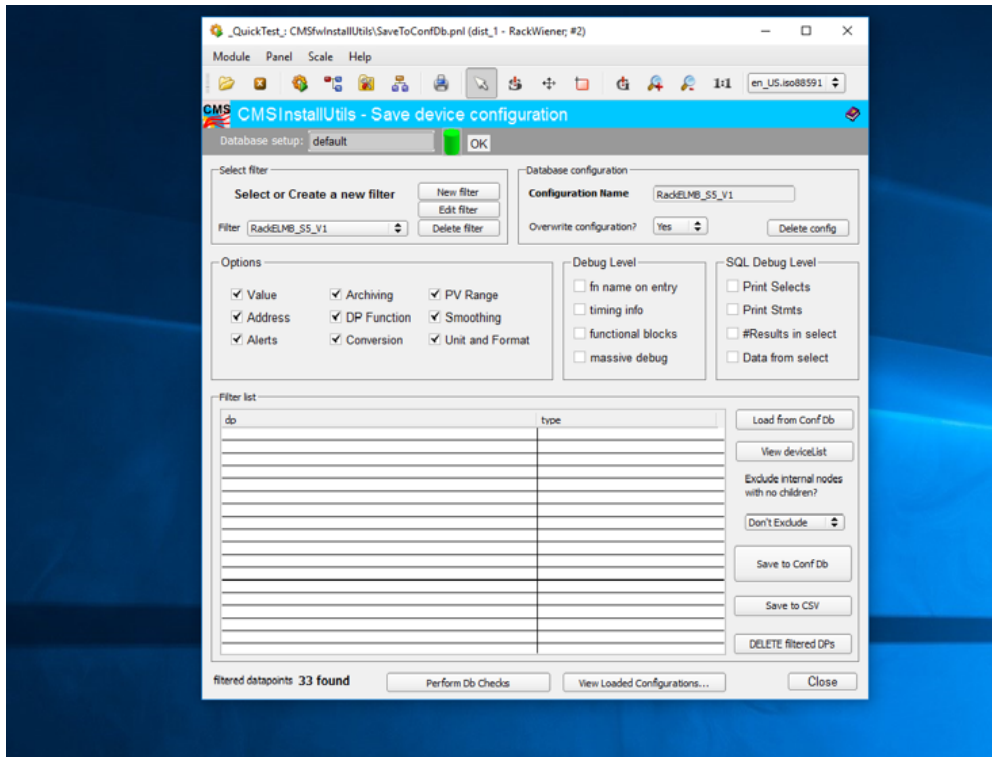


Figure 4.4: Main screen of the CMSfwInstallUtils component

to the subdetectors as well as monitoring for the CMS hardware (ELMB, crates @P5). Among the Central DCS team’s responsibilities, lies the general support regarding WinCC\_OA within the CMS collaboration, as well as other services such as database expertise or on-call service during detector running periods.

The National Technical University of Athens (NTUA) has been supporting and developing tools for the Central DCS team since 2016. Our involvement in the CMS DCS team includes the training on various framework components, tools, and system structures (both central and subdetector) with the aim to become DCS experts able to assist the on-call service of the central DCS team. Also, NTUA aids the development, support, improvement and testing of new or already available framework components or tools. The NTUA team participates in projects of the CMS DCS team including maintenance, development and improvement of the CMS framework.

## 4.6 CMS DCS Configuration Database

### Configuration database consistency checks

WinCC\_OA and the JCOP framework enable the connection between a developed project and a database. The CMS DCS team makes use of this functionality to store vital information regarding the structure of the monitoring project in a database. This allows the recreation of the whole CMS DCS or part of it, in case of any failure.

Since 2016, the NTUA team has created a component that cross checks between the information found in the system and the database. The tool comes handy when manipulating or interacting with a project that contains massive amounts of data with configurations. The tool is integrated in the already existing *CMSfwInstallUtils* component.

After the first check, the subcomponent allows the user to check for further inconsistencies. The newly developed consistency check checks the datapoint attributes (address, archiving, alerts, etc). The component also produces a table containing results, assisting the user to identify the problematic datapoints if any.

More specifically, the user can check the device List that is to be saved on database or it is already saved on the database. These panels give the ability to check for all datapoint elements

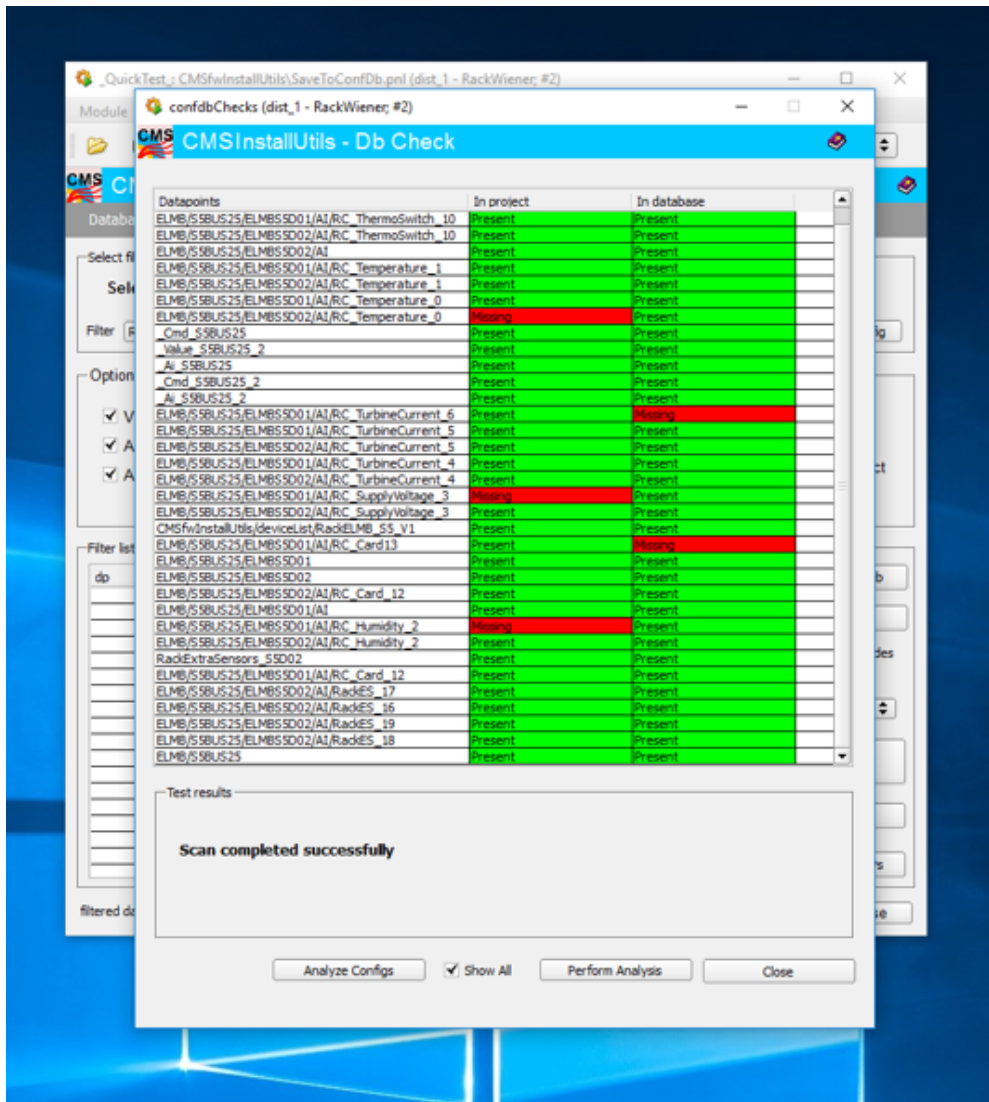


Figure 4.5: Inconsistency check between the project and the database (inconsistencies with red).

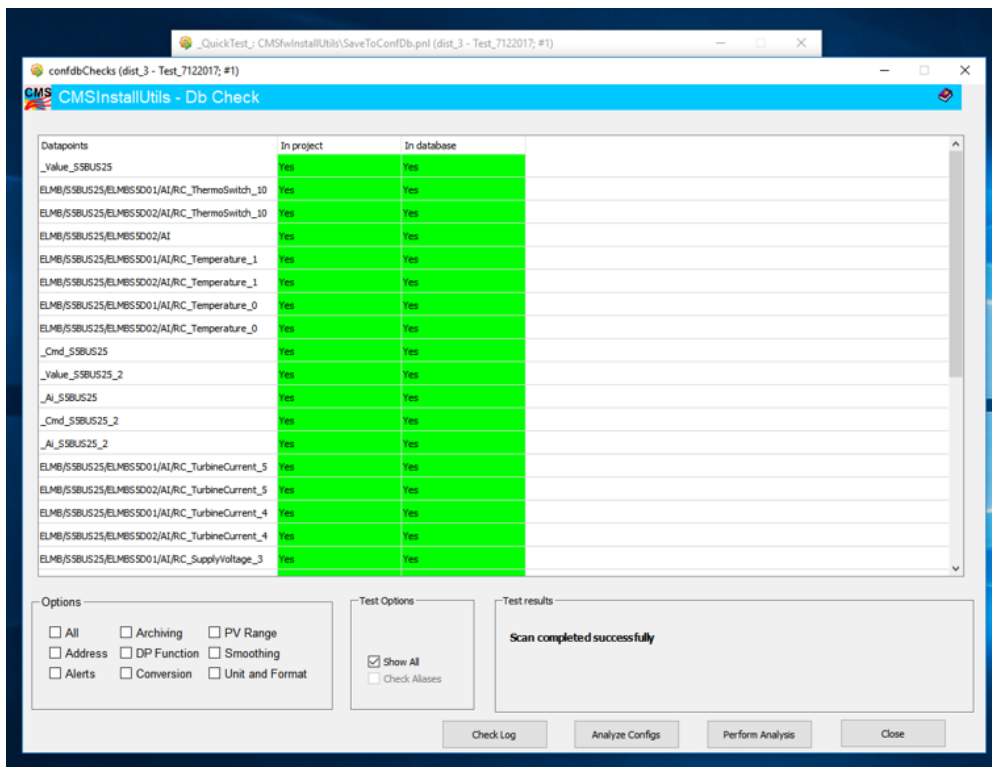


Figure 4.6: Main screen of the CMSfwInstallUtils Configuration Database component

whatever is missing either from project or from the database. The developed tool also checks each config (address, alerts, etc) individually for each datapoint element and does the same procedure. There are several possible outcomes:

- *Consistent*: This means that everything is ok
- *Inconsistent*: This means that config was found on the dpe in the db and in the project but it wasn't exactly the same
- *Missing*: Either missing from database or from project
- *Missing and inconsistent*: Found some inconsistent and missing datapoint elements

When *missing* or *inconsistency* is found the user can right click and an information panel opens with more info for the datapoint elements containing any issues as shown in Fig 4.9.



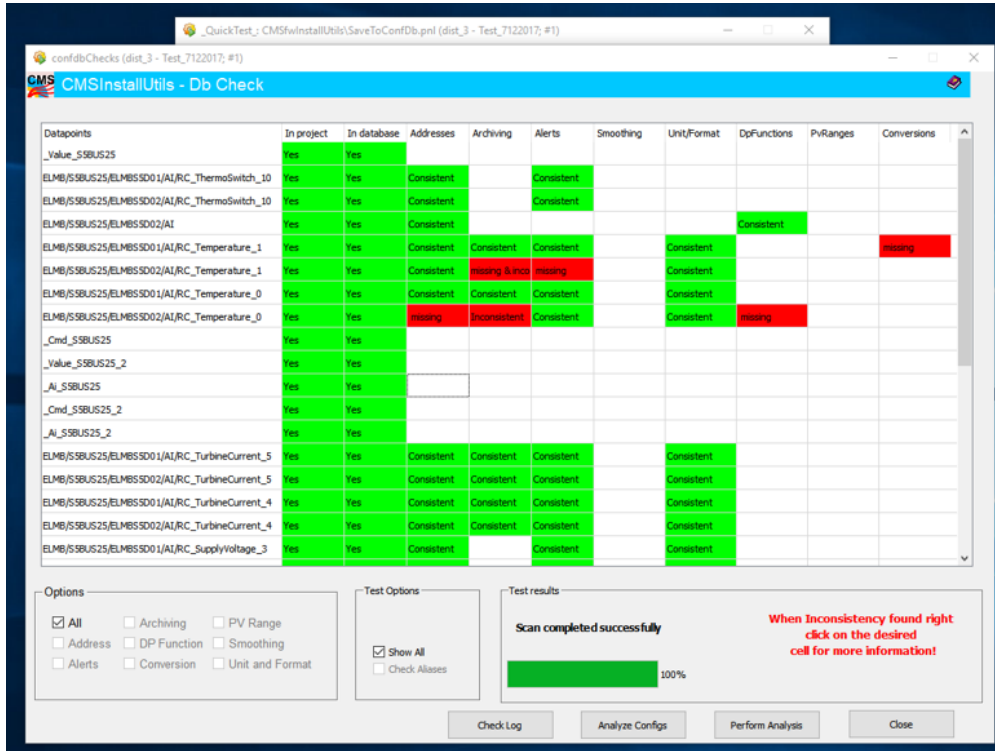


Figure 4.7: Provides the option of demonstrating only the inconsistent datapoints.

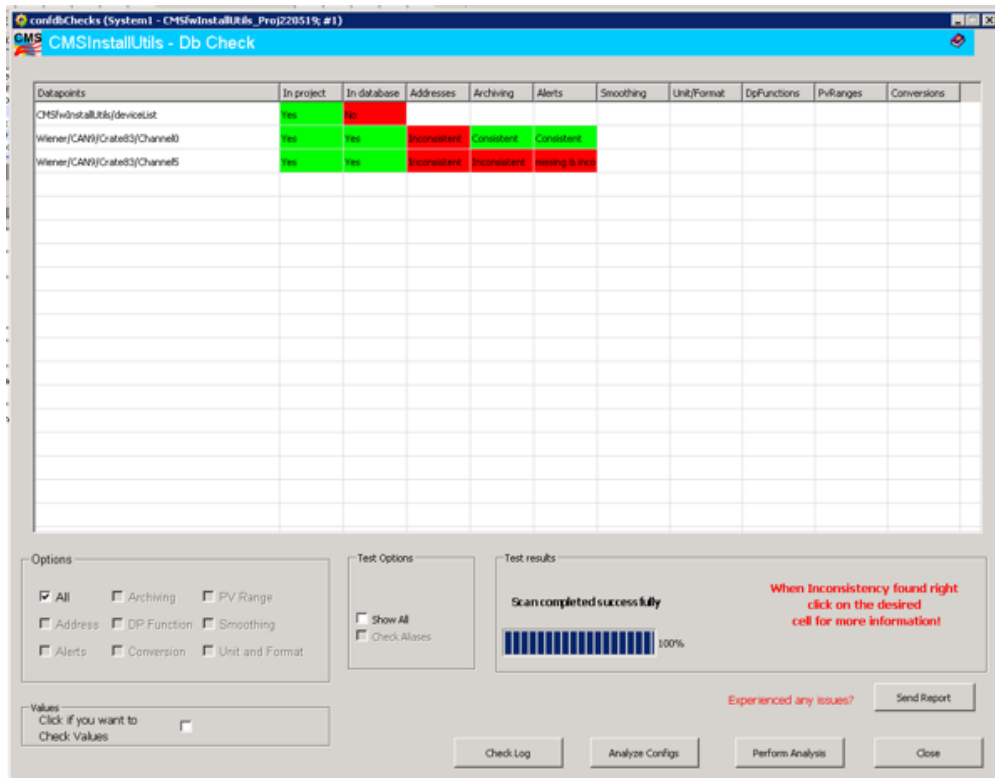


Figure 4.8: ConfDB Checks component utilities.

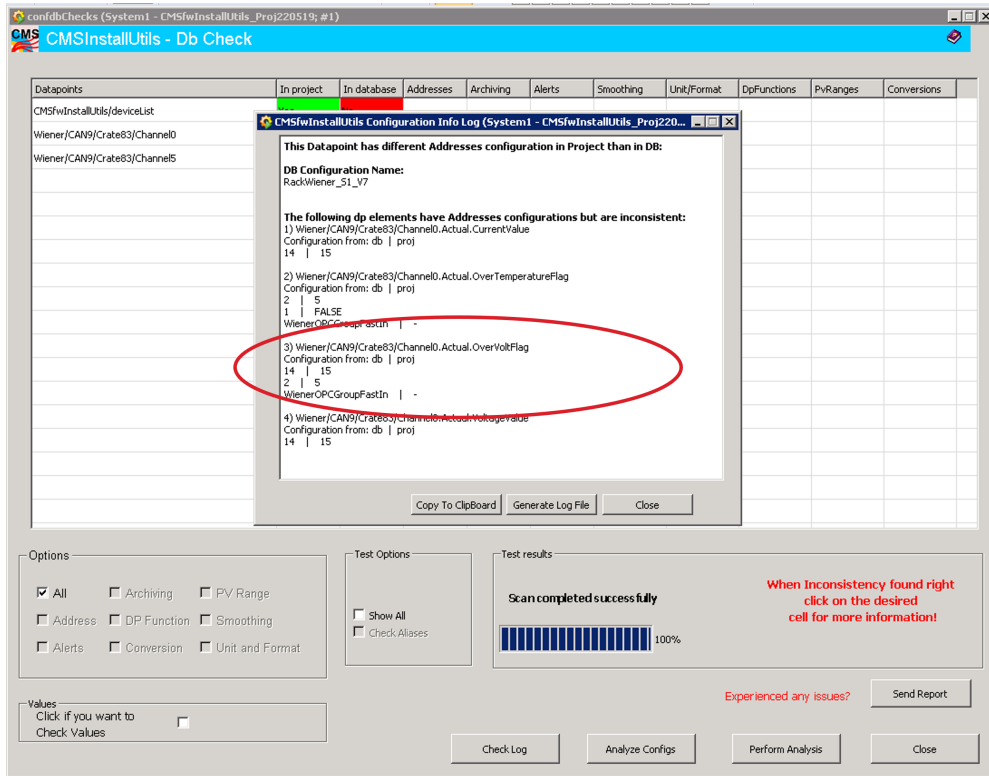


Figure 4.9: ConfDB Checks component utilities: When Inconsistency found or missing feature, the user is able to right click and check exactly which datapoint element is inconsistent or has missing configuration.

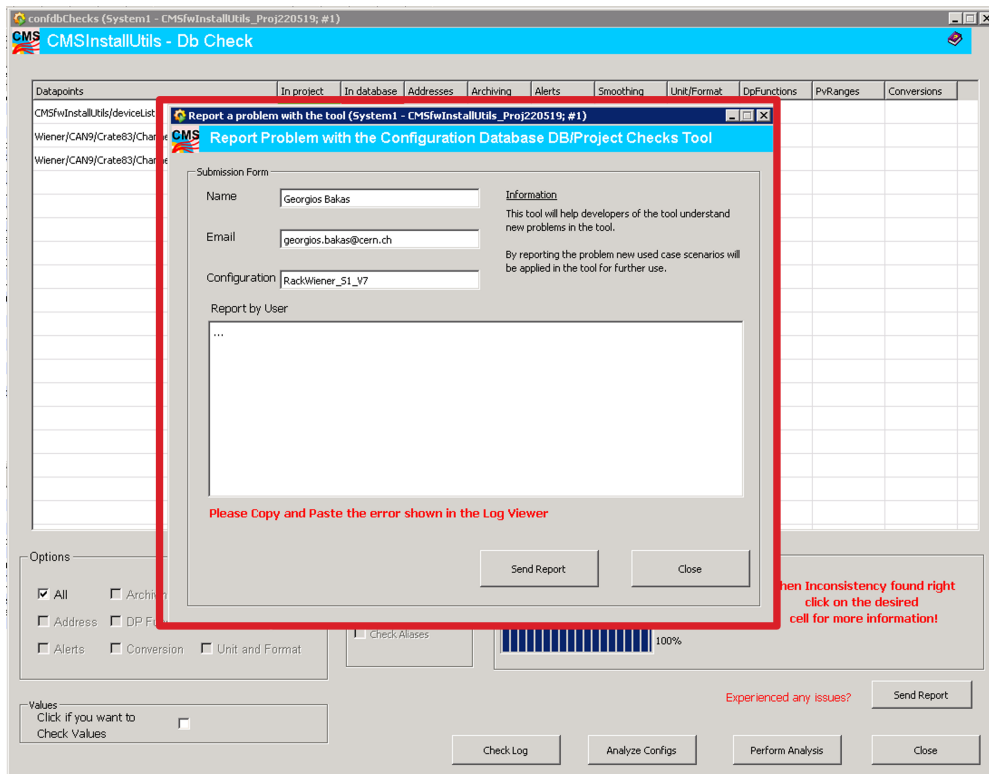


Figure 4.10: ConfDB Checks component allows the users report inconsistencies or missing datapoint elements that were not found. In this case the Central DCS team reproduces the case and incorporates configurations that are not found.

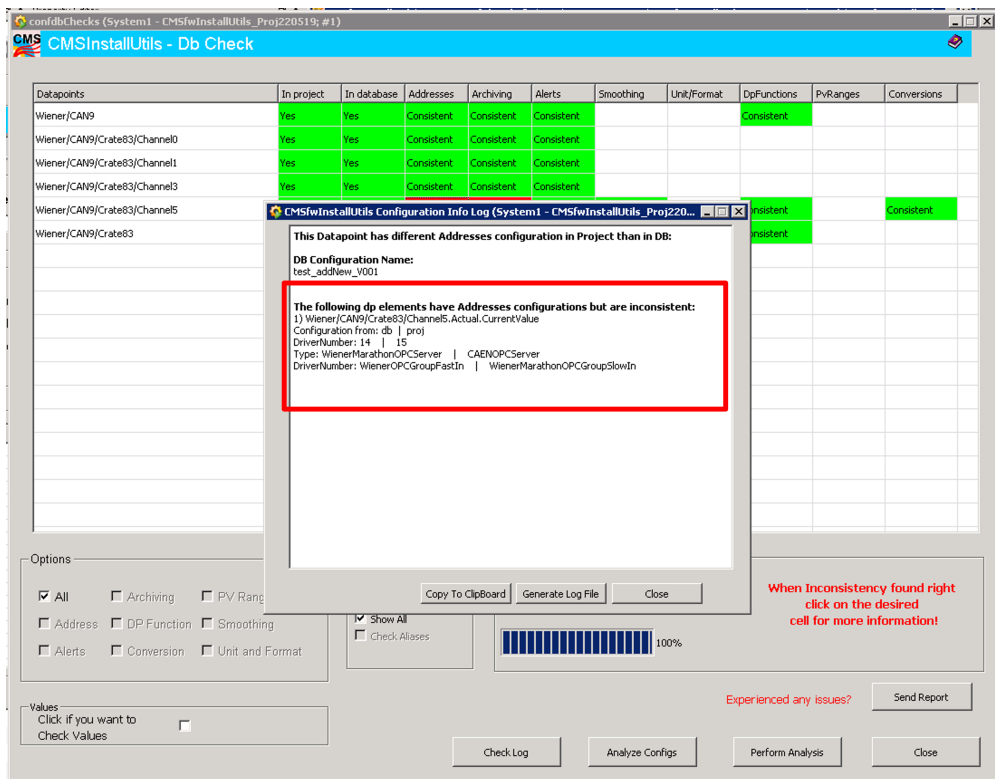


Figure 4.11: ConfDB Checks component utilities shows the user the entire DB vs Local project configuration when an inconsistency is found for the Address Configuration of a datapoint.

## 4.7 CMS DCS Installation Tool

The CMS DCS is a distributed system of 34 WinCC\_OA system that run on Dell Blades. Most of the systems are running on Windows Server operating systems, while the rest are running on Linux. WinCC\_OA systems are all redundant in order to protect the functioning of malfunctions and bugs in production.

For the production deployment, the subsystem developers and experts have no access to the CMS DCS computers. The solution to deploy custom made tools and modules has the following features:

- Package everything as a JCOP component
- Components are read from SVN (Code versioning, patching without re-installing the component while component is running)
- Installation tool & Configuration DB to recreate production projects locally and therefore test before deployment.

The JCOP team provides a CMS install library that standarizes and simplifies the **postInstall** scripts. The end-users are able to *Add/Stop/Restart* drivers and managers while at the same time target the Configuration Database (ConfigDB) configurations to the created system. Also everything is defined in a custom made XML file and only one line of code is needed for installation. The production Hardware and Logical views are also stored in the ConfigDB. Because there is no standard solution to store FSM trees, CMS has developed custom XML import/export files for the FSM restoration/reproduction.

All the projects are configured to run in the Central Management Mode. The System Configuration DB is the master, thus all changes in the configuration database are applied to the production system. The installation agent is always running in the production system checking for inconsistency every 5 minutes. Changes in the DB are pushed by subdetector experts that can update the version of their components using *cmsonline*. To target a new component, the intervention of the central DCS expert is required using the System Configuration Database WinCC\_OA interface. The CMS deployment policy ensures that all changes are reflected in the System Configuration Database, thus making any changes reproducible in case the system needs to be reconfigured or reinstalled. Also, this policy allows to upgrade to a new WinCC\_OA version or a subdetector component version in a relatively quick manner.

The JCOP team has been developing a framework to install and recreate a production project from scratch. The component is named *fwInstallation* and uses *.bat* files as well as queries to the Configuration DB. Specifically, the *CreateServiceProject.bat* follows specific steps:

- Connect to the System Configuration DB
- Retrieve all information (System Name, Config file, etc) depending on the hostname
- Create a default project with a predefined set of basic managers
- Start the installation tool that finds an empty project and installs all the components in a proper order to ensure that dependencies are configured correctly.
- Handle redundancy

The Cental CMS DCS team has been developing a component that allows the subdetector experts to recreate production-like systems in ther local systems. An end user is connected to the ??? and is able to navigate through a panel (dbEditorAndNavigator) which is incorporated in the *fwInstallationUtils* component. The NTUA CMS DCS team provided assistance and support for the developemnt and debugging of this new tool. Some solutions, and improvements are provided in the following:

- System properties selection not shown in some projects
- *ComputerProperties*: bug in registering a new host
- Problem with an sql query when a new project is created using this tool
- Fixes when someone tries to access system hierarchy properties
- Check user credentials so that the user cannot delete/create projects unless the owner
- Fix batch (.bat) files used for installation of service projects
- *reinstallComponentsIntoProjects*: User can specify which component to be reinstalled and if the files should be overwritten and the project to be restarted
- dbEditorNavigator
- Create project like

## 4.8 MoxaE1240 Hardware Component

The CMS DCS required the development of a tool in order for the CMS Tracker team to communicate with the MoxaE1240. MoxaE1240 is a device which allows the retrieval of I/O data, making it capable of handling a wide variety of applications. Applications such as factory automation, security and surveillance systems, and tunneled connections can make use of daisy-chained Ethernet for building multidrop I/O networks over standard Ethernet cables. The communication of the MoxaE1240 hardware with the end-user was performed via the SNMP communication protocol. Simple Network Management Protocol (SNMP) [136] is an Internet Standard protocol that allows different devices on a network to share information with one another. Devices that typically support SNMP include cable modems, routers, switches, servers, workstations, printers, and more.

The tracker team required a tool (both panel and backend) that has the following attributes:

- Uses an OID to connect each channel (e.g. Reference record for the SNMP protocol).
- Uses some kind of counter that is received by WinCC\_OA and is then converted to a Voltage and sometimes a temperature which can be plotted and archived
- User can configure and monitor the hardware from the Device Editor Navigator
- Possibility to save all the configurations to the database

The Central DCS NTUA team provided a tool that was incorporated in the CMS framework as an individual component. The following Figures show the output panel provided for the tracker end-users.

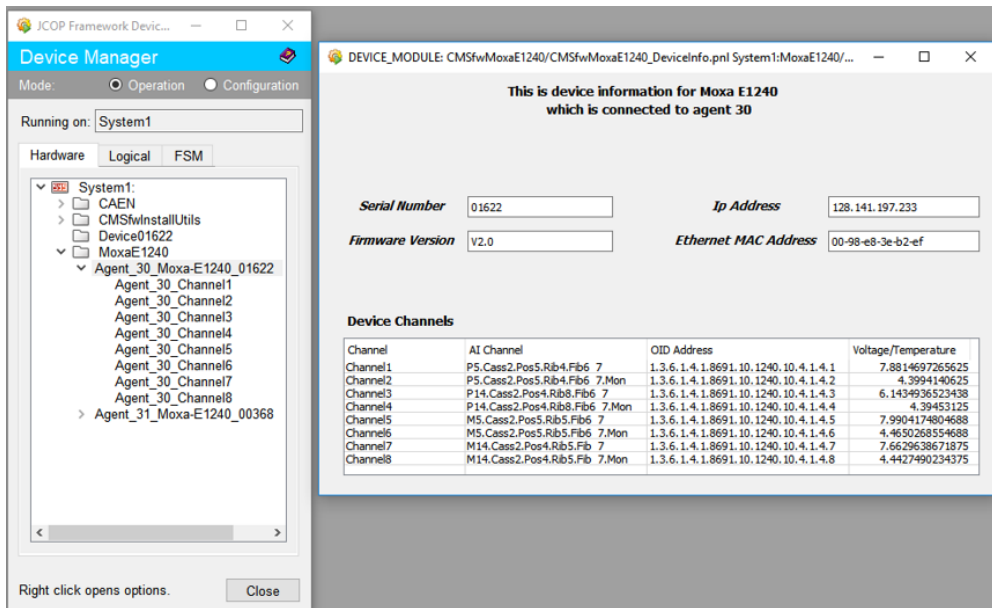


Figure 4.12: Central CMSfwMoxa component panel for the SNMP connection with the target agent, integrated into the Device Editor & Navigator.

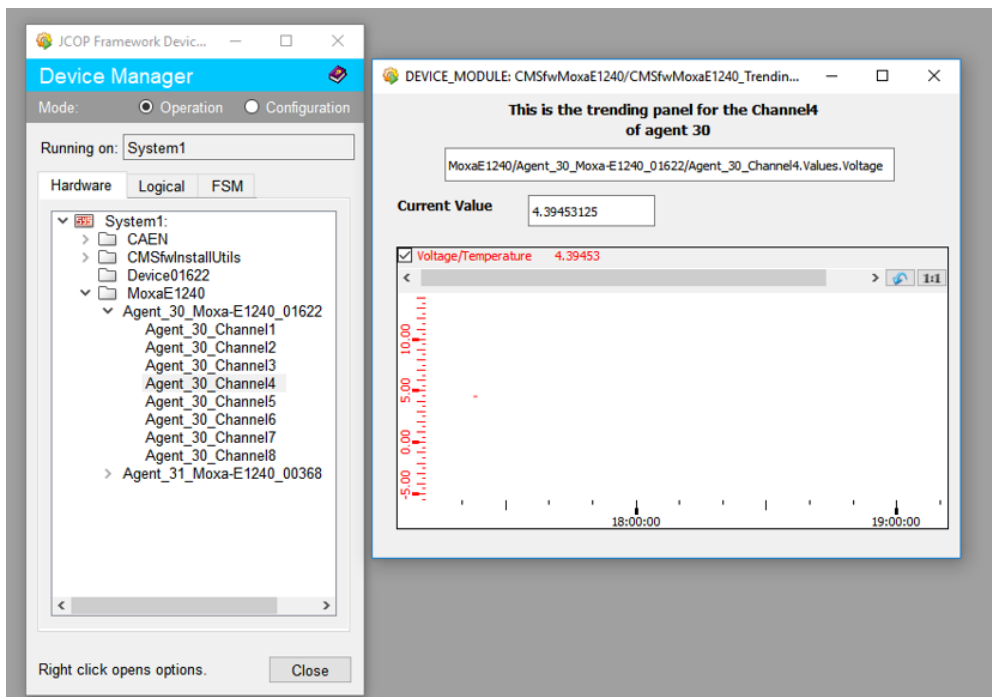


Figure 4.13: CMSfwMoxa component panel showing temperature values connected to the MOXAE1240 hardware and communicating with WinCC\_OA via SNMP protocol.

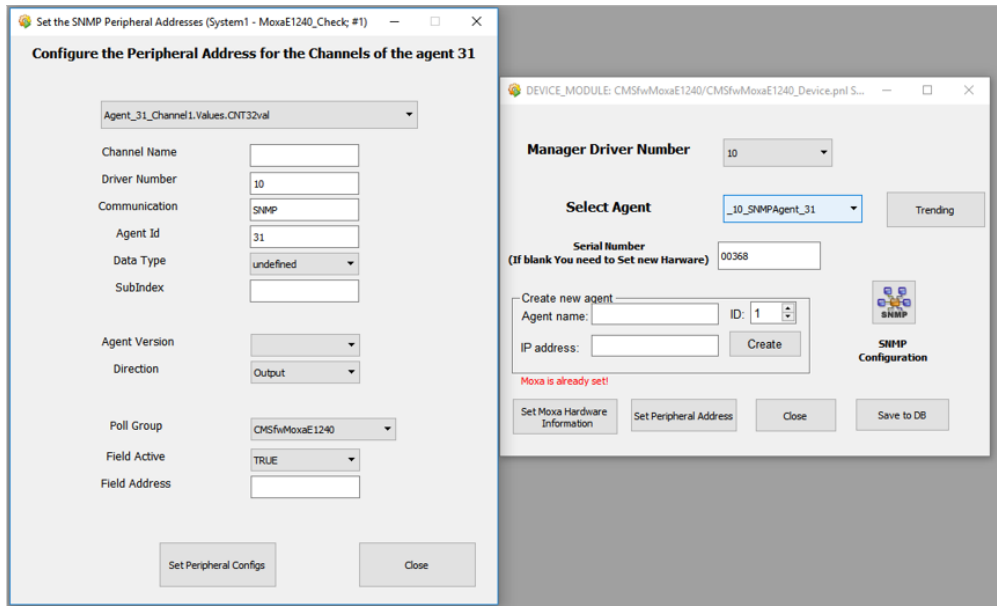


Figure 4.14: CMSfwMoxa Panels for the configuration of the peripheral address for the Channels of a specific agent.

## 4.9 Other projects

- **ELMB project:** The ELMBs for the RICH detectors are used to collect environmental information and interface sensors to WinCC\_OA . The ELMB is a 16 bit differential ADC with 64 inputs that are multiplexed. The range full input range can be adjusted from 25 mV to 5 V and set as unipolar or bipolar. The NTUA Central DCS team developed a component for the Tracker DCS team that offered the possibility to change the ELMB limits.
- **PLT HV Scans:** Develop a panel for the PLT detector. The users were able to select a bunch of HV channels that belonged to the PLT detector and make changes to the High Voltages while performing checks for their hardware.
- **CMSfwAlertSystem:** The NTUA DCS team updated this component in order to tackle a specific problem which occurred when an end-user installed the component locally and connected to the Configuration Database, thus reproducing and sending notifications to subscribers of a specific alert. Our team developed a methodology to shut notifications in non production systems while also fixing CMSfwAlertSystem main panel bugs and providing new implementations.

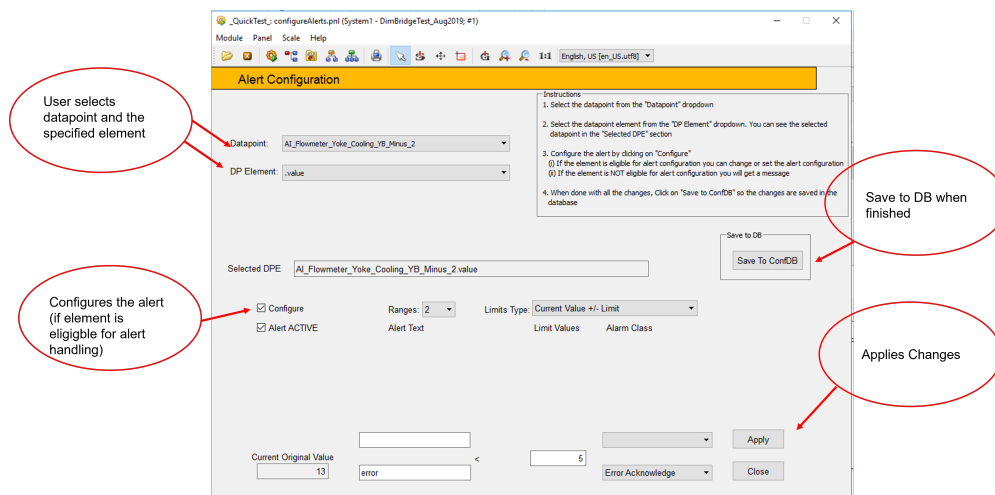


Figure 4.15: CMSfwMoxa Panels for the configuration of the peripheral address for the Channels of a specific agent.

- **CMSfwFsmXml component:** Reported bug when importing an already set device type in the project via XML file
- **DCS FSM main panel at Shifter's screen:** The Central FSM at P5, where the CMS experiment is located, is being operated by a shifter. In order to gain ownership of the central FSM, the shifter's credentials are read from a Card Reader. Sometimes shifters do not place their card correctly, resulting into sending out several emails that the FSM has no owner. The goal was to build a panel instructing the shifters how to correctly place their card while also implementing a filter for the CardReader changes to avoid any inconsistencies with the FSM ownership.
- **CMSfwRack:** Our team developed and implemented an expert panel that gives the users the ability to select several datapoint elements that have alarm configuration for a specific set of datapoint type Racks and activate or deactivate them. There was an existing panel which lets the expert user activate or deactivate all datapoint element alarms of a specific dp Type. The new panel also incorporated the implementation of import/export csv files for doing so.
- **Robocopy for linux machine:** Every 24 hours, all systems related to DCS machines make a copy of all of their local directories into DFS (CERN Distributed File System). For the Windows



machines, the windows scheduler is utilized, and the robocopy copies some files from the production projects to a directory in DFS. This tools does not work in linux machines. Our team developed a new methodology for copying linux production system files to the DFS by making use of the *rsync* command that runs as a cron-job every 24 hours.



## **Part III**

# **$t\bar{t}$ Analysis**



## Chapter 5

# Boosted $t\bar{t}$ Differential Measurements

A measurement of the production cross section of high transverse momentum ( $p_T$ ) top quark pairs is reported. The dataset was collected during the full Run II with the CMS detector at the CERN LHC from proton-proton collisions at a center-of-mass energy of 13 TeV. The measurement uses events where both top quark candidates decay hadronically and are reconstructed as large-R jets with  $p_T > 400$  GeV. The cross section is extracted differentially as a function of kinematic variables of the top quark or top quark pair system. The results are unfolded to the particle and parton levels, and are compared to various theoretical models.

### 5.1 Introduction

The top quark completes the third generation of quarks in the standard model (SM), and the precise knowledge of its properties is critical for the overall understanding of the theory. Measurements of the top-anti-top quark pair ( $t\bar{t}$ ) production cross section confront the predictions from quantum chromodynamics (QCD) and have the potential to constrain the QCD parameters, while being sensitive to physics beyond the SM. Also, the  $t\bar{t}$  production process is a dominant SM background to searches for new physics phenomena and therefore its precise knowledge is essential for new discoveries.

The large  $t\bar{t}$  yield expected in proton-proton (pp) collisions at the CERN LHC allows to perform measurements of the  $t\bar{t}$  production rate in a large phase space, and, more importantly, differentially, as a function of the  $t\bar{t}$  kinematic properties. Such measurements have been performed by the ATLAS [64, 65, 66, 67, 68, 69] and CMS [70, 71, 72, 73, 74, 75, 76, 77, 78] Collaborations at 7, 8, and 13 TeV center-of-mass energies, under the hypothesis of the resolved final state, where the decay products of the  $t\bar{t}$  pair can be reconstructed individually. This hypothesis is valid for top quark transverse momenta,  $p_T$ , up to approximately 500 GeV. However, at higher  $p_T$  ( $p_T/m \approx 1$ ), the top quark decay products are highly collimated ("boosted") and they can no longer be reconstructed separately. In order to explore the highly boosted phase space, hadronic top quark decays are reconstructed as large-radius jets. Previous efforts in this domain by the ATLAS [79, 80] and CMS [81, 82, 83] Collaborations confirm that it is feasible to perform precise differential measurements of the  $t\bar{t}$  production and have also shown interesting deviations from the theory predictions.

In this note, a measurement of the differential boosted  $t\bar{t}$  production cross section in the hadronic final state is presented, using  $pp$  collisions at  $\sqrt{s} = 13$  TeV recorded with the CMS detector during the 2016, 2017 and 2018 LHC full Run II and amounting to a total integrated luminosity of  $35.9 \text{ fb}^{-1}$ ,  $41.5 \text{ fb}^{-1}$  and  $59.7 \text{ fb}^{-1}$  respectively. In the hadronic decay channel, each  $W$  boson arising from the top quark decays into a pair of light quarks. As a result, the final state consists of at least six partons (more are possible due to initial- and final-state radiation), two of which are b quarks. Due to the high boost considered in this measurement (leading jet  $p_T > 450$  GeV), the top quarks are reconstructed unambiguously as large-radius jets and the final state consists of at least two such jets.

## 5.2 Samples

In this section we describe the data and Monte Carlo samples that are used in the analysis.

### 5.2.1 Data

The collision events used for the measurement of the  $t\bar{t}$  cross section have been collected with the triggers described in Section 5.3 that are part of the JetHT primary dataset. Tables 5.1-5.3 list the samples that correspond to different data-taking eras and have been reconstructed in the 21Feb2020, 31Mar2018, 17Sep2018 reprocessing for 2016 2017 and 2018 respectively. The 2016 data are splitted into two periods.

The good Run and luminosity section list used is contained in the certification file for each year:

- **2016\_preVFP:** Cert\_271036-284044\_13TeV\_Legacy2016\_Collisions16\_JSON.txt. The total integrated luminosity of the analyzed data is  $19.5 \text{ fb}^{-1}$ .
- **2016\_postVFP:** Cert\_271036-284044\_13TeV\_Legacy2016\_Collisions16\_JSON.txt. The total integrated luminosity of the analyzed data is  $16.8 \text{ fb}^{-1}$ .
- **2017:** Cert\_294927-306462\_13TeV\_UL2017\_Collisions17\_GoldenJSON.txt. The total integrated luminosity of the analyzed data is  $41.5 \text{ fb}^{-1}$ .
- **2018:** Cert\_314472-325175\_13TeV\_Legacy2018\_Collisions18\_JSON.txt. The total integrated luminosity of the analyzed data is  $59.7 \text{ fb}^{-1}$ .

Except from the JetHT dataset, we have used the SingleMuon dataset for the measurement of the trigger efficiency. The eras and ranges are exactly the same as the ones in Tables 5.1, 5.2, 5.3.

Table 5.1: Data samples 2016.

Sample	Run range	Lumi ( $\text{pb}^{-1}$ )
JetHT/Run2016B-21Feb2020_ver2_UL2016_HIPM-v1/MINIAOD	273150-275376	5750
JetHT/Run2016C-21Feb2020_UL2016_HIPM-v1/MINIAOD	275656-276283	2573
JetHT/Run2016D-21Feb2020_UL2016_HIPM-v1/MINIAOD	276315-276811	4242
JetHT/Run2016E-21Feb2020_UL2016_HIPM-v1/MINIAOD	276947-277420	4025
JetHT/Run2016F-21Feb2020_UL2016_HIPM-v1/MINIAOD	277932-278807	3105
JetHT/Run2016F-21Feb2020_UL2016-v1/MINIAOD	278769-278808	7576
JetHT/Run2016G-21Feb2020_UL2016-v1/MINIAOD	278820-284035	8435
JetHT/Run2016H-21Feb2020_UL2016-v1/MINIAOD	281613-284044	216

Table 5.2: Data samples 2017.

Sample	Run range	Luminosity ( $\text{pb}^{-1}$ )
/JetHT/Run2017B-31Mar2018-v1/MINIAOD	297047-299329	4793
/JetHT/Run2017C-31Mar2018-v1/MINIAOD	299368-302029	9755
/JetHT/Run2017D-31Mar2018-v1/MINIAOD	302031-302663	4320
/JetHT/Run2017E-31Mar2018-v1/MINIAOD	303824-304797	9422
/JetHT/Run2017F-31Mar2018-v1/MINIAOD	305040-306460	13568

### 5.2.2 Simulation

Monte Carlo simulation is used to generate samples for the  $t\bar{t}$  signal and to model the kinematic distributions of some of the background processes. Samples of simulated  $t\bar{t}$  events have been

Table 5.3: Data samples 2018.

Sample	Run range	Luminosity ( $\text{pb}^{-1}$ )
/JetHT/Run2018A-17Sep2018-v1/MINIAOD	315257-316995	13530
/JetHT/Run2018B-17Sep2018-v1/MINIAOD	317080-319310	6788
/JetHT/Run2018C-17Sep2018-v1/MINIAOD	319337-320065	6612
/JetHT/Run2018D-PromptReco-v2/MINIAOD	320497-325175	31947

generated at next-to-leading order (NLO) in QCD using POWHEG v2 [84, 85, 86, 87, 88], assuming a top quark of  $m_t = 172.5$  GeV. Single top quark production in the  $t$  channel or in association with a  $W$  boson are simulated at NLO with POWHEG [89]. The production of  $W$  or  $Z$  bosons in association with jets ( $V$ +jets), as well as QCD multijet events, are simulated with MG5\_AMC@NLO [90] at leading order (LO), with the MLM matching algorithm [91].

All simulated events are processed with PYTHIA 8.240 [92, 93] for modeling of the parton showering, hadronization, and underlying event (UE). The NNPDF 3.1 NLO [94, 95] Parton Distribution Functions (PDF) are used throughout, and the CP5 tune [96] is used for all processes. The Simulation of the CMS detector response is based on GEANT4 [97]. Additional  $pp$  interactions in the same or neighbouring bunch crossings (pileup) are simulated with PYTHIA and overlaid with generated events according to the pileup distribution measured in data.

The various simulated processes are normalized to the best known theoretical cross sections, namely the  $t\bar{t}$ ,  $V$ +jets and single top [98, 99, 100].

The measured cross sections for the  $t\bar{t}$  process are compared to theoretical predictions provided by the following Monte Carlo models: POWHEG combined with PYTHIA for the parton showering, as aforementioned in the previous paragraph and NLO MG5 aMC@NLO[FxFx] again as mentioned above. The matching of the matrix-element jets to parton showers is performed using the FxFx [101] prescription. For both POWHEG combined with PYTHIA and the NLO MG5 aMC@NLO[FxFx] predictions, only the Parton Distribution Function uncertainties, the Parton Shower uncertainties (ISR, FSR) and the Renormalization and factorization scales as described in Chapter 8.

In the first part of the Tables 5.4- 5.7 we show the signal samples, while in the second part we show the background samples that include QCD multijet production,  $V$ +jets, and single-top production. The list shows the total number of events analyzed, including all samples of the same kind (nominal, extensions, backup). The reconstruction of each year Monte Carlo samples belongs to the associated eras:

- 2016 preVFP: RunIISummer20UL16MiniAODAPV-106X\_mcRun2\_asymptotic\_preVFP
- 2016 postVFP: RunIISummer16MiniAODv3-PUMoriond17\_94X\_mcRun2\_asymptotic
- 2017: RunIIFall17MiniAODv2-PU2017\_12Apr2018\_94X\_mc2017\_realistic
- 2018: RunIIAutumn18MiniAOD-102X\_upgrade2018\_realistic

In general, we have used the MINIAODSIM data tier. The reported cross sections have been used in the normalization of the various processes in the data vs MC comparison plots.

Table 5.4: Monte Carlo samples 2016 preVFP.

Sample	Events ( $\times 10^6$ )	$\sigma$ (pb)
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	98.1	377.96
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	138.6	365.34
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	41.6	88.29
TTToHadronic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	40.0	377.96
TTToSemiLeptonic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	56.8	365.34
TTTo2L2Nu_mtop166p5_TuneCP5_13TeV-powheg-pythia8	16.7	88.29
TTToHadronic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	39.6	377.96
TTToSemiLeptonic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	57.7	365.34
TTTo2L2Nu_mtop169p5_TuneCP5_13TeV-powheg-pythia8	16.4	88.29
TTToHadronic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	38.9	377.96
TTToSemiLeptonic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	57.9	365.34
TTTo2L2Nu_mtop171p5_TuneCP5_13TeV-powheg-pythia8	16.9	88.29
TTToHadronic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	39.5	377.96
TTToSemiLeptonic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	57.5	365.34
TTTo2L2Nu_mtop173p5_TuneCP5_13TeV-powheg-pythia8	16.9	88.29
TTToHadronic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	39.8	377.96
TTToSemiLeptonic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	55.5	365.34
TTTo2L2Nu_mtop175p5_TuneCP5_13TeV-powheg-pythia8	16.9	88.29
TTToHadronic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	38.3	377.96
TTToSemiLeptonic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	57.9	365.34
TTTo2L2Nu_mtop178p5_TuneCP5_13TeV-powheg-pythia8	16.9	88.29
TTToHadronic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	40.0	377.96
TTToSemiLeptonic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	56.9	365.34
TTTo2L2Nu_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	17.0	88.29
TTToHadronic_hdampUP_TuneCP5_13TeV-powheg-pythia8	40.0	377.96
TTToSemiLeptonic_hdampUP_TuneCP5_13TeV-powheg-pythia8	57.0	365.34
TTTo2L2Nu_hdampUP_TuneCP5_13TeV-powheg-pythia8	17.0	88.29
TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8	99.3	832
QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	52.6	315400
QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	58.5	32260
QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	45.5	6830
QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	14.1	1207
QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	10.3	119.1
QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	5.1	25.16
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	3.4	38.09
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	3.3	38.09
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	31.0	35.6
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	56.0	35.6
ST_t-channel_antitop_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	31.0	119.7
ST_t-channel_top_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	57.1	82.52
WJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	8	2549.0
WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	5.14	276.5
WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	7.62	59.25
WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8	7.71	28.75



Table 5.5: Monte Carlo samples 2016 postVFP.

Sample	Events ( $\times 10^6$ )	$\sigma$ (pb)
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	112.6	377.96
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	158.8	365.34
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	48.3	88.29
TTToHadronic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	43.6	377.96
TTToSemiLeptonic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	64.0	365.34
TTTo2L2Nu_mtop166p5_TuneCP5_13TeV-powheg-pythia8	18.7	88.29
TTToHadronic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	44.8	377.96
TTToSemiLeptonic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	55.1	365.34
TTTo2L2Nu_mtop169p5_TuneCP5_13TeV-powheg-pythia8	18.3	88.29
TTToHadronic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	45.0	377.96
TTToSemiLeptonic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	63.9	365.34
TTTo2L2Nu_mtop171p5_TuneCP5_13TeV-powheg-pythia8	18.9	88.29
TTToHadronic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	44.8	377.96
TTToSemiLeptonic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	64.0	365.34
TTTo2L2Nu_mtop173p5_TuneCP5_13TeV-powheg-pythia8	18.8	88.29
TTToHadronic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	43.5	377.96
TTToSemiLeptonic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	64.0	365.34
TTTo2L2Nu_mtop175p5_TuneCP5_13TeV-powheg-pythia8	19.0	88.29
TTToHadronic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	44.6	377.96
TTToSemiLeptonic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	64.0	365.34
TTTo2L2Nu_mtop178p5_TuneCP5_13TeV-powheg-pythia8	19.0	88.29
TTToHadronic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	45.0	377.96
TTToSemiLeptonic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	63.7	365.34
TTTo2L2Nu_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	18.3	88.29
TTToHadronic_hdampUP_TuneCP5_13TeV-powheg-pythia8	45.0	377.96
TTToSemiLeptonic_hdampUP_TuneCP5_13TeV-powheg-pythia8	63.0	365.34
TTTo2L2Nu_hdampUP_TuneCP5_13TeV-powheg-pythia8	19.0	88.29
TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8	110.1	832
QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	57.6	315400
QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	59.7	32260
QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	47.4	6830
QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	15.3	1207
QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	10.4	119.1
QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	5.3	25.16
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	10.3	38.09
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	3.8	38.09
ST_t-channel_antitop_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	4.0	35.6
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	22.8	35.6
ST_t-channel_top_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	6.0	119.7
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	46.5	82.52
WJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	7.07	2549.0
WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	4.49	276.5
WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	6.76	59.25
WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8	6.85	28.75

Table 5.6: Monte Carlo samples 2017

Sample	Events ( $\times 10^6$ )	$\sigma$ (pb)
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	249.6	377.96
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	355.9	365.34
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	107.0	88.29
TTToHadronic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	97.8	377.96
TTToSemiLeptonic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	135.7	365.34
TTTo2L2Nu_mtop166p5_TuneCP5_13TeV-powheg-pythia8	42.5	88.29
TTToHadronic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	98.3	377.96
TTToSemiLeptonic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	138.6	365.34
TTTo2L2Nu_mtop169p5_TuneCP5_13TeV-powheg-pythia8	41.0	88.29
TTToHadronic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	97.9	377.96
TTToSemiLeptonic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	141.0	365.34
TTTo2L2Nu_mtop171p5_TuneCP5_13TeV-powheg-pythia8	43.0	88.29
TTToHadronic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	99.9	377.96
TTToSemiLeptonic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	141.7	365.34
TTTo2L2Nu_mtop173p5_TuneCP5_13TeV-powheg-pythia8	42.7	88.29
TTToHadronic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	100.0	377.96
TTToSemiLeptonic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	141.6	365.34
TTTo2L2Nu_mtop175p5_TuneCP5_13TeV-powheg-pythia8	42.9	88.29
TTToHadronic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	99.6	377.96
TTToSemiLeptonic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	141.3	365.34
TTTo2L2Nu_mtop178p5_TuneCP5_13TeV-powheg-pythia8	42.7	88.29
TTToHadronic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	99.6	377.96
TTToSemiLeptonic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	141.6	365.34
TTTo2L2Nu_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	41.1	88.29
TTToHadronic_hdampUP_TuneCP5_13TeV-powheg-pythia8	99.7	377.96
TTToSemiLeptonic_hdampUP_TuneCP5_13TeV-powheg-pythia8	140.9	365.34
TTTo2L2Nu_hdampUP_TuneCP5_13TeV-powheg-pythia8	42.2	88.29
TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8	249.1	832
QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	57.2	315400
QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	57.9	32260
QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	45.8	6830
QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	15.3	1207
QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	10.6	119.1
QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	5.5	25.16
ST_t-channel_antitop_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	3.7	35.6
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	24.3	35.6
ST_t-channel_top_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	5.5	119.7
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	46.2	82.52
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	10.0	38.09
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	9.2	38.09
WJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	15.42	2549.0
WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	9.91	276.5
WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	14.4	59.25
WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8	14.75	28.75

Table 5.7: Monte Carlo samples 2018

Sample	Events ( $\times 10^6$ )	$\sigma$ (pb)
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	347.4	377.96
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	495.0	365.34
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	148.5	88.29
TTToSemiLeptonic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	191.1	365.34
TTToHadronic_mtop166p5_TuneCP5_13TeV-powheg-pythia8	137.8	377.96
TTTo2L2Nu_mtop166p5_TuneCP5_13TeV-powheg-pythia8	60.0	88.29
TTTo2L2Nu_mtop169p5_TuneCP5_13TeV-powheg-pythia8	59.8	88.29
TTToHadronic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	138.7	377.96
TTToSemiLeptonic_mtop169p5_TuneCP5_13TeV-powheg-pythia8	194.7	365.34
TTTo2L2Nu_mtop171p5_TuneCP5_13TeV-powheg-pythia8	59.8	88.29
TTToHadronic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	135.3	377.96
TTToSemiLeptonic_mtop171p5_TuneCP5_13TeV-powheg-pythia8	195.2	365.34
TTTo2L2Nu_mtop173p5_TuneCP5_13TeV-powheg-pythia8	59.9	88.29
TTToHadronic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	139.6	377.96
TTToSemiLeptonic_mtop173p5_TuneCP5_13TeV-powheg-pythia8	200.0	365.34
TTTo2L2Nu_mtop175p5_TuneCP5_13TeV-powheg-pythia8	60.0	88.29
TTToHadronic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	138.4	377.96
TTToSemiLeptonic_mtop175p5_TuneCP5_13TeV-powheg-pythia8	199.8	365.34
TTTo2L2Nu_mtop178p5_TuneCP5_13TeV-powheg-pythia8	59.9	88.29
TTToHadronic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	139.7	377.96
TTToSemiLeptonic_mtop178p5_TuneCP5_13TeV-powheg-pythia8	199.9	365.34
TTTo2L2Nu_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	60.0	88.29
TTToHadronic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	139.7	377.96
TTToSemiLeptonic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8	193.2	365.34
TTTo2L2Nu_hdampUP_TuneCP5_13TeV-powheg-pythia8	59.8	88.29
TTToHadronic_hdampUP_TuneCP5_13TeV-powheg-pythia8	138.3	377.96
TTToSemiLeptonic_hdampUP_TuneCP5_13TeV-powheg-pythia8	199.4	365.34
TTTo2L2Nu_TuneCP5down_13TeV-powheg-pythia8	60.0	88.29
TTToHadronic_TuneCP5down_13TeV-powheg-pythia8	139.8	377.96
TTToSemiLeptonic_TuneCP5down_13TeV-powheg-pythia8	190.3	365.34
TTTo2L2Nu_TuneCP5up_13TeV-powheg-pythia8	57.9	88.29
TTToHadronic_TuneCP5up_13TeV-powheg-pythia8	139.9	377.96
TTToSemiLeptonic_TuneCP5up_13TeV-powheg-pythia8	199.7	365.34
TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8	340.5	832
QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	55.1	315400
QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	58.5	32260
QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	47.7	6830
QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	15.7	1207
QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	10.6	119.1
QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	4.5	25.16
ST_t-channel_antitop_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	3.7	35.6
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	23.4	35.6
ST_t-channel_top_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	5.7	119.7
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	47.7	82.52
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	10.1	38.09
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	9.1	38.09
WJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	14.39	2549.0
WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	9.26	276.5
WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	13.54	59.25
WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8	13.6	28.75

### 5.3 Trigger

For 2016, the trigger paths employed for the collection of signal events use single-jet L1 seeds that require the presence of a jet with  $p_T > 180$  GeV. At HLT, jets are reconstructed from (online) particle flow candidates using the anti-kt algorithm with distance parameter  $R = 0.8$ . Their mass, after trimming of soft particles, must be greater than 30 GeV. In order to classify an event as interesting, it is required to have at least two such jets with  $p_T > 280(200)$  GeV for the leading (trailing) respectively. Also at least one of the two jets should be tagged as a b-jet, using the online CSV algorithm. The trigger path ran unprescaled for the duration of the 2016 run, collected an integrated luminosity of  $35.9 \text{ fb}^{-1}$ .

A second, prescaled, path, using the same L1 seed, was also employed, with identical kinematic requirements but no b-tagging cut, which ran in parallel and collected and integrated luminosity of  $1.67 \text{ fb}^{-1}$ . This path is used for the selection of a control QCD sample, as described later. All the triggers described above are summarized in Table 5.8. Finally, it should be noted that the pileup profile of the prescaled control trigger is shifted to lower number of interactions (Fig. 5.1), because the path tended to collect more data towards the end of the fills when the instantaneous luminosity was lower.

For the 2017 and 2018 measurements, a more complex L1 seed is used, see Table 5.9. The big or statement requires that the total hadronic transverse energy (HTT) is at least 120 GeV. Or the presence of 4 jets with HTT at least of 280 and the energy of the leading jet to be greater than 70 GeV, the second 55 GeV, the third 40 GeV and the fourth 34 GeV. All of them must have a pseudorapidity in the range  $[-2.4, 2.4]$ . In the HLT events are required to have  $H_T > 800$  GeV ( $p_T$  sum of all the jets) and that the mass of each jet is greater than 50 GeV after trimming of soft particles. In contrast to the HLT used for 2016, the triggers for 2017 and 2018 do not use b-tagging selection at HLT level. The triggers ran unprescaled for the duration of both 2017 and 2018, collecting an integrated luminosity of  $41.5 \text{ fb}^{-1}$  and  $59.7 \text{ fb}^{-1}$  respectively. The triggers are fully described in Table 5.9.

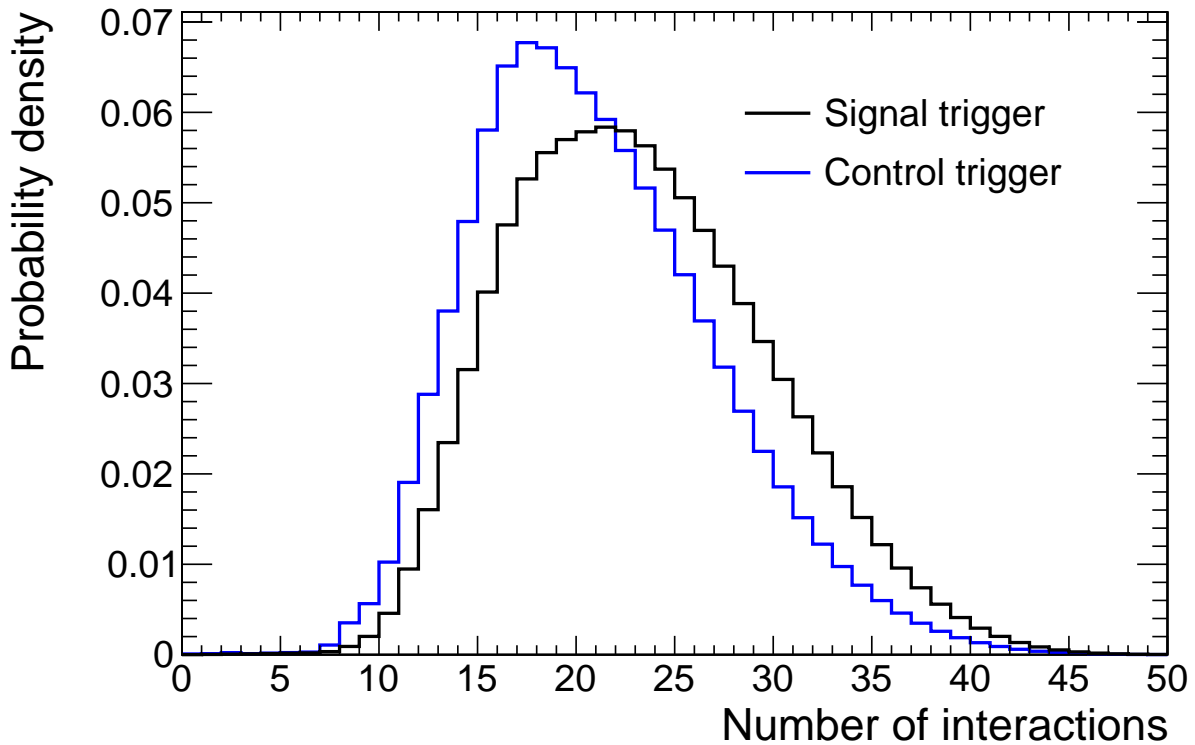


Figure 5.1: Pileup profile for the signal and control trigger paths.

The efficiency of the signal trigger path is measured with respect to an orthogonal path that

Table 5.8: Summary of triggers used in the analysis.

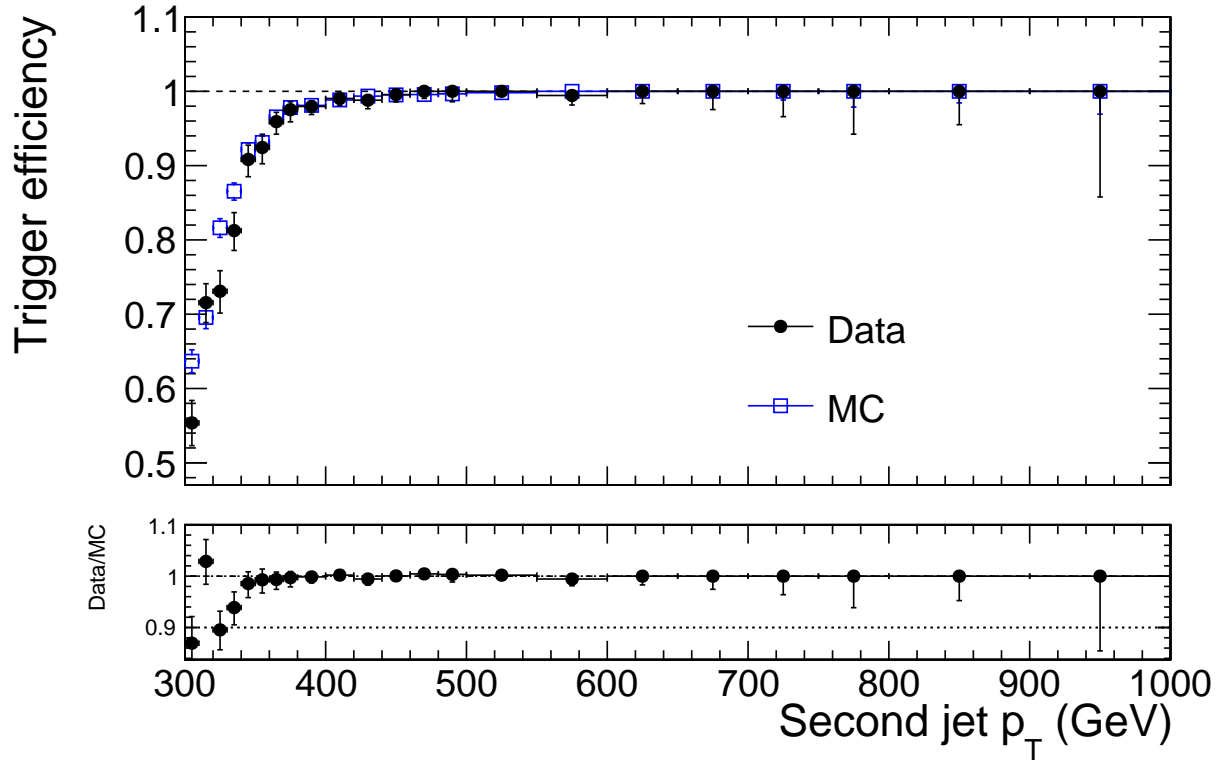
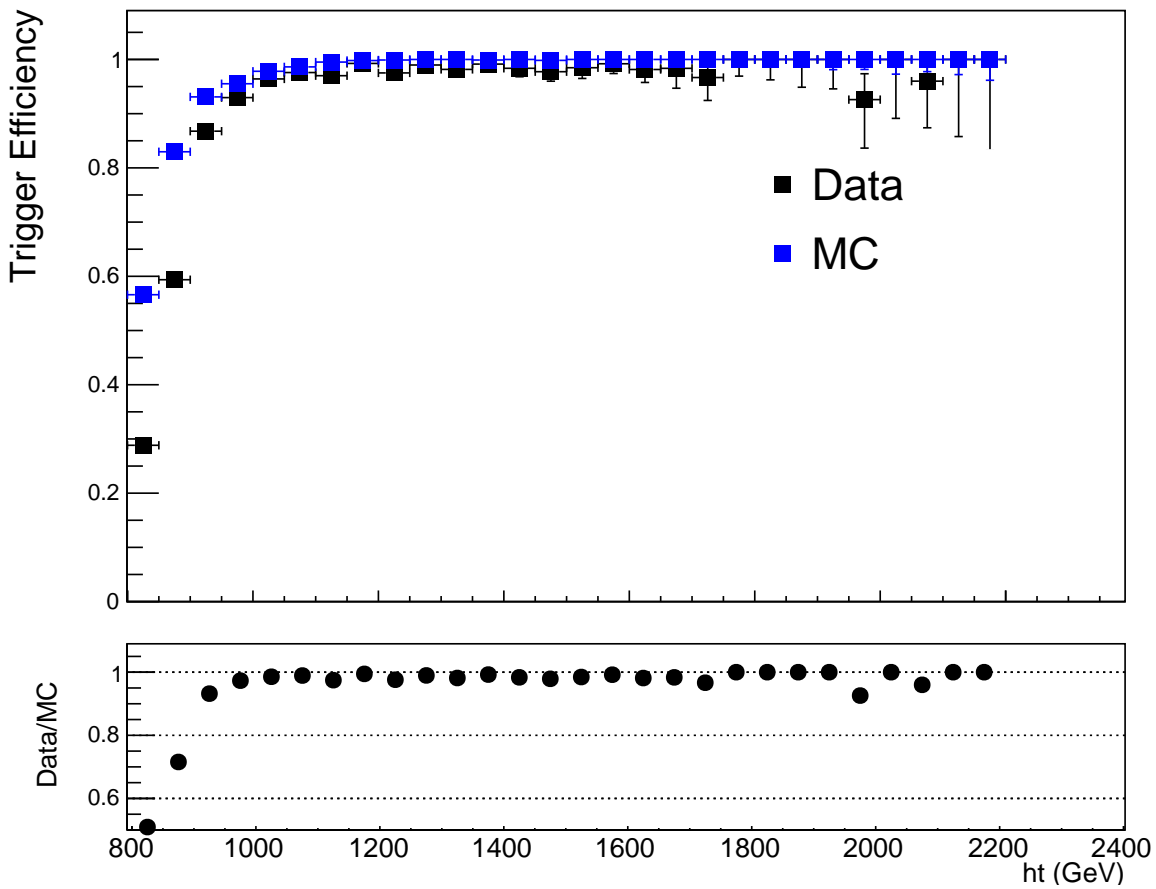
Trigger	Purpose
L1_SingleJet180 OR L1_SingleJet200	L1 seed
HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20	signal HLT path
HLT_AK8PFJet140	control HLT path

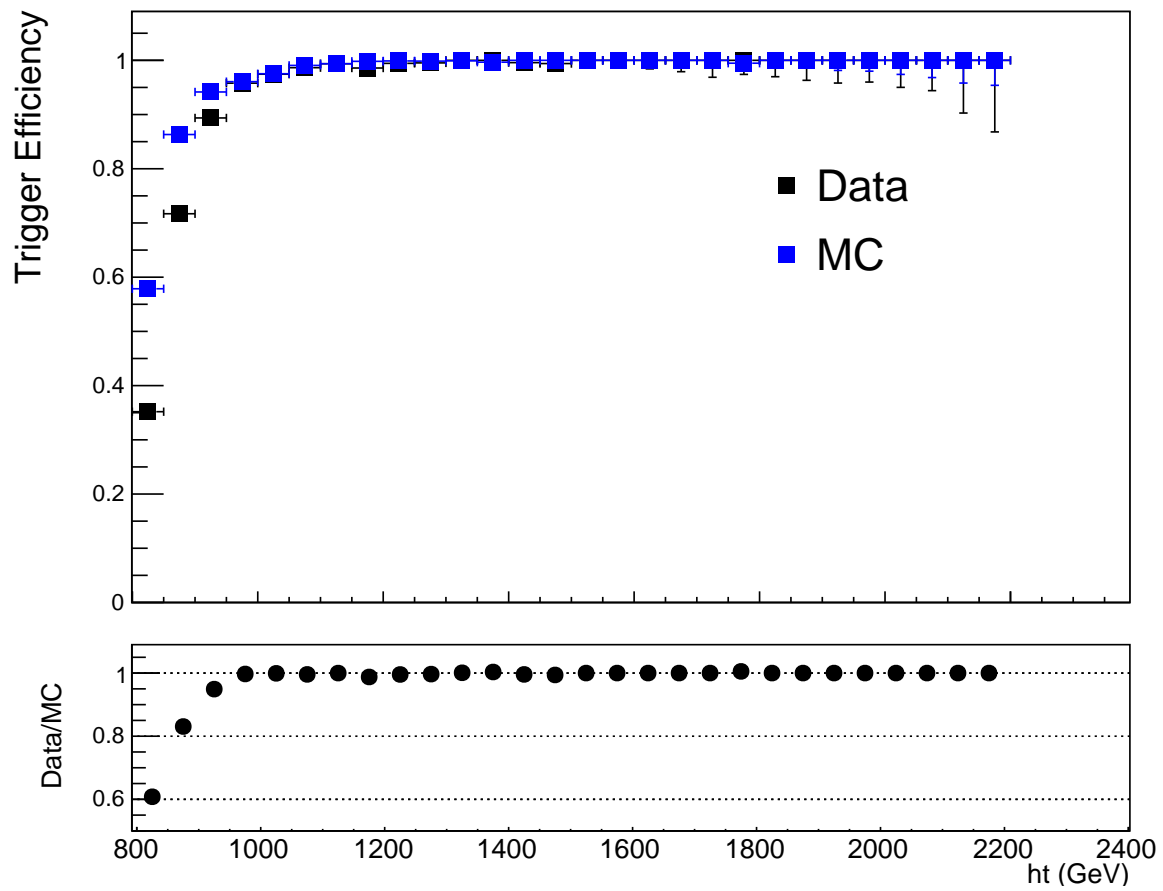
Table 5.9: Summary of triggers used in the 2017 and 2018 analyses.

Trigger	Purpose
L1_HTT120er OR L1_HTT160er OR L1_HTT200er OR L1_HTT220er OR L1_HTT240er OR L1_HTT255er OR L1_HTT270er OR L1_HTT280er OR L1_HTT280er_QuadJet_70_55_40_35_er2p5 OR L1_HTT300er OR L1_HTT300er_QuadJet_70_55_40_35_er2p5 OR L1_HTT320er OR L1_HTT340er OR L1_HTT380er OR L1_HTT400er OR L1_HTT450er OR L1_HTT500er	L1 seed
HLT_AK8PFHT800_TrimMass50	signal & control HLT path

requires the presence of an isolated muon with  $p_T > 27$  GeV (HLT\_IsoMu27). Figure 5.2 shows the trigger efficiency as a function of the second jet  $p_T$  in events with at least two reconstructed jets, with at least one of them contains a b-tagged subjet. For details about the offline reconstruction see Section 5.4. The efficiency measured in data is compared to the simulated efficiency, showing an excellent agreement. The offline selection requires that the second jet  $p_T$  is greater than 400 GeV, which is at the beginning of the efficiency plateau.

For 2017 and 2018 the efficiency is measured with respect to the same orthogonal trigger HLT\_IsoMu27. Figures 5.3 and 5.4 show the trigger efficiency as a function of  $H_T$ . The offline selection requires the  $H_T$  of the event to be greater than 850 GeV. The efficiency measured in data is represented in comparison with the one from simulation and seem to have a fairly good agreement.

Figure 5.2: Trigger efficiency for the signal path of the analysis vs second leading  $p_T$  for 2016.Figure 5.3: Trigger efficiency for the signal path of the analysis vs  $ht$  for 2017.

Figure 5.4: Trigger efficiency for the signal path of the analysis vs  $ht$  for 2018.

## 5.4 Reconstruction and Selection

In this section we present the objects used in the analysis, the multivariate method that discriminates  $t\bar{t}$  events from the QCD multijet background, and we describe the event selection.

### 5.4.1 Object Reconstruction

The leptons (muons, electrons) used in this analysis come from the default reconstructed collections in CMS samples ('slimmedMuons' and 'slimmedElectrons') and must have  $p_T > 20$  GeV. Muons are required to pass the medium ID working point and electrons must pass the tight working point, while both lepton types must have a relative mini-isolation less than 0.1.

Jets are reconstructed from particle-flow (PF) candidates. These can be further classified based on different pileup reduction techniques. In this analysis we use the PUPPI jets (Pileup Per Particle Identification). The momentum 4-vectors of the PF candidates are clustered using the anti-kt algorithm where the distance parameter  $R$  is set to  $R = 0.8$ . Hence we will refer to these candidates as AK8 PFPUPPI jets. The aforementioned jets are required to pass the tight jet ID. Also, an algorithm runs with purpose to identify subjets of distance parameter  $R = 0.4$  within the AK8 jets. Regarding the reconstructed AK8 jet mass, the soft drop technique is used for its calculation. For all the aforementioned algorithms, our group has used the default definitions in the CMS software and the default collections in the MINIAOD data and Monte Carlo samples.

The selection of the AK8 jets originating from the top (anti-top) decay relies on the identification of at least one b-jet within the respective large- $R$  jet. For this purpose, we use the DeepCSV b-tagging algorithm that is applied on the AK4 subjets of each AK8 jet. More specifically, we use the medium working point, which requires the value of the tagger to be greater than: 0.6001, 0.5847, 0.4506, 0.4168 for 2016\_preVFP, 2016\_postVFP, 2017 and 2018 respectively.

Since leptons are also reconstructed as jets, we perform a cross cleaning by removing each identified lepton from the jet collection with geometrical matching in the  $\eta - \phi$  space: if a jet has  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$  from any accepted lepton candidate it is removed from the jet collection.

### 5.4.2 Selection

The baseline selection, summarized in Table 5.10, is common for all regions used in the analysis. It requires at least two large- $R$  jets ( $R = 0.8$ ) in the event with leading jet  $p_T > 450$  GeV and second leading jet with  $p_T > 400$  GeV and softDrop mass values within the range (50, 300) GeV. Furthermore, a veto on the leptons is applied in order to minimize the probability to select leptonic top decays. On top of the baseline selection, we define four specific regions based on the output of our BDT (Boosted Decision Tree) that discriminates  $t\bar{t}$  signal from QCD background, the jets' softDrop masses and the number of the b-tagged subjets in each jet, that serve different analysis purposes. The signal region ( $SR$ ) is where we perform the differential measurements and it requires both jets to have a b-tagged subjet and a tighter selection of the jet mass softDrop value 120 – 220 GeV. The pure QCD control region ( $CR$ ) has the same selection criteria with the aforementioned Signal Region, with the only difference being that the b-tagging requirement is inverted (the jets must not contain a b-tagged subjet). This region is used to receive the shape of the QCD background from the data (data driven method) for each variable of interest. Finally, the signal region A ( $SR_A$ ) is defined. This is the region where the fit is performed and is used to determine the normalization of the QCD background. The  $SR_A$  is a superset of the signal region having exactly the same cuts but a looser mass cut [50 – 300 GeV]



Table 5.10: Baseline Selection Criteria

Observable	Requirement
$N_{\text{jets}}$	$> 1$
$N_{\text{leptons}}$	$= 0$
$p_T^{\text{jet1}}$	$> 450 \text{ GeV}$
$p_T^{\text{jet2}}$	$> 400 \text{ GeV}$
$ \text{et al.}^{\text{jet1,2}} $	$< 2.4$
$m_{SD}^{\text{jet1,2}}$	$(50, 300) \text{ GeV}$
$m_{t\bar{t}}$	$> 1000 \text{ GeV}$

Table 5.11: Selection requirements per analysis region for 2016 pre and post VFP

Region	Trigger	Offline	Purpose
$SR$	Signal	$\text{Baseline}+BDT > 0.2+2btag+m_{SD}^{\text{jet1,2}} \in (120, 220) \text{ GeV}$	Signal Region
$SR_A$	Signal	$\text{Baseline}+BDT > 0.2+2btags$	QCD fit region
$CR$	Control	$\text{Baseline}+BDT > 0.2+0btag+m_{SD}^{\text{jet1,2}} \in (120, 220) \text{ GeV}$	QCD control region
$CR_A$	Control	$\text{Baseline}+BDT > 0.2+0btags$	QCD extended CR

Table 5.12: Selection requirements per analysis region for 2017 (Control and Signal triggers are the same)

Region	Trigger	Offline	Purpose
$SR$	Signal	$\text{Baseline}+BDT > 0.0+2btag+m_{SD}^{\text{jet1,2}} \in (120, 220) \text{ GeV}$	Signal Region
$SR_A$	Signal	$\text{Baseline}+BDT > 0.0+2btags$	QCD fit region
$CR$	Signal	$\text{Baseline}+BDT > 0.0+0btag+m_{SD}^{\text{jet1,2}} \in (120, 220) \text{ GeV}$	QCD control region
$CR_A$	Control	$\text{Baseline}+BDT > 0.0+0btags$	QCD extended CR

Table 5.13: Selection requirements per analysis region for 2017 (Control and Signal triggers are the same)

Region	Trigger	Offline	Purpose
$SR$	Signal	$\text{Baseline}+BDT > 0.1+2btag+m_{SD}^{\text{jet1,2}} \in (120, 220) \text{ GeV}$	Signal Region
$SR_A$	Signal	$\text{Baseline}+BDT > 0.1+2btags$	QCD fit region
$CR$	Signal	$\text{Baseline}+BDT > 0.1+0btag+m_{SD}^{\text{jet1,2}} \in (120, 220) \text{ GeV}$	QCD control region
$CR_A$	Control	$\text{Baseline}+BDT > 0.1+0btags$	QCD extended CR

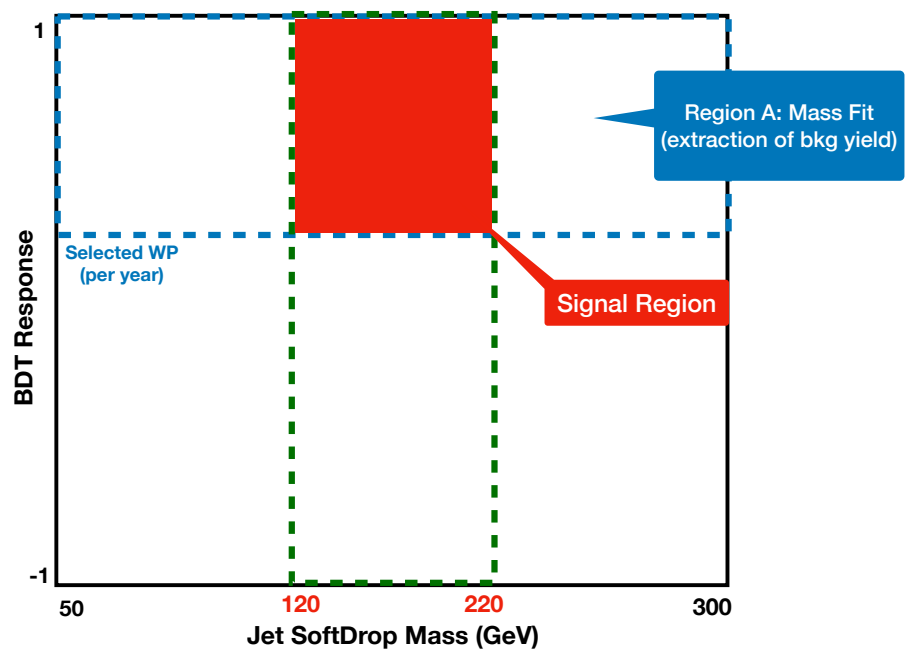


Figure 5.5: Graphical representation of the analysis region in the BDT-output vs jet softDrop mass.

### 5.4.3 Multivariate Discriminant

In order to discriminate between events that come from  $t\bar{t}$  decays and QCD multijet production, we rely on variables that reveal the substructure of the jet. In particular, we use the "N-subjettiness"  $\tau_N$  as well as the "Energy Correlation Functions"  $ECFB(N, \beta)$ . The aforementioned "N-subjettiness" is defined as:

$$\tau_N = \frac{1}{\sum_k p_{T,k} R} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\}, \quad (5.1)$$

where  $N$  denotes the reconstructed candidate subjets and  $k$  runs over the constituent particles in the jet. The variable  $\Delta R_{i,k} = \sqrt{(\Delta y_{i,k})^2 + (\Delta \phi_{i,k})^2}$  is the angular distance between the candidate subjet  $i$  and the particle  $k$ . The variable  $R$  is the characteristic jet radius ( $R = 0.8$ ).

As for the Energy Correlation Functions they are based on the energies and pair-wise angles of particles within a jet, with  $(N+1)$ -point correlators sensitive to  $N$ -prong substructure. For hadron colliders, these functions are defined better as a function of the transverse momentum. They are defined as:

$$ECF(N, \beta) = \sum_{i_1 < i_2 \dots < i_N \in J} \left( \prod_{a=1}^N p_{T,i_a} \right) \left( \prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta \quad (5.2)$$

where  $R_{ij}$  is the Euclidean distance between  $i$  and  $j$  in the rapidity-azimuth angle plane,  $R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ , with  $y_i = \frac{1}{2} \ln \frac{E_i + p_{z_i}}{E_i - p_{z_i}}$ . In this analysis we only use the ECFB1N2, ECFB1N3, ECFB2N2, ECFB2N3. The developed BDT discriminates the jets within each event based on the jet's  $\tau_{1,2,3}$ ,  $ECF(N, \beta)$ , the mass of the subjets within the jet (AK4 jets) and the fraction of the jet  $p_T$  to the sum of the  $p_T$ 's of all the jets in the event. As a result our Multivariate Discriminant Analysis requires 10 variables to decide whether a jet is a top (antitop) or not. In comparison with previous analysis [102] which discriminates a top-antitop pair from QCD multijets, the developed top tagger decides whether a jet is a top jet candidate or not.

As a baseline we consider a simple Fisher discriminant, which is a linear combination of the variables. We also deploy a Boosted Decision Tree composed of 500 trees and trained with the Gradient Boost method with a shrinkage parameter equal to 0.1. The training performed here is a categorical training. The jets are trained based on their  $p_T$ . There are 4 categories: (400,600) GeV, (600,800) GeV (800,1200) GeV and (1200,Inf) GeV. The distribution of the response of the developed BDT can be seen in 5.6. In Appendix the distribution of the variables that were used for the training of the Boosted Decision Trees are depicted. The training is performed using  $t\bar{t}$  MC samples for the signal and QCD samples for the background. The signals are split in half using the first part for the training and the second for the testing of the response of the BDT.

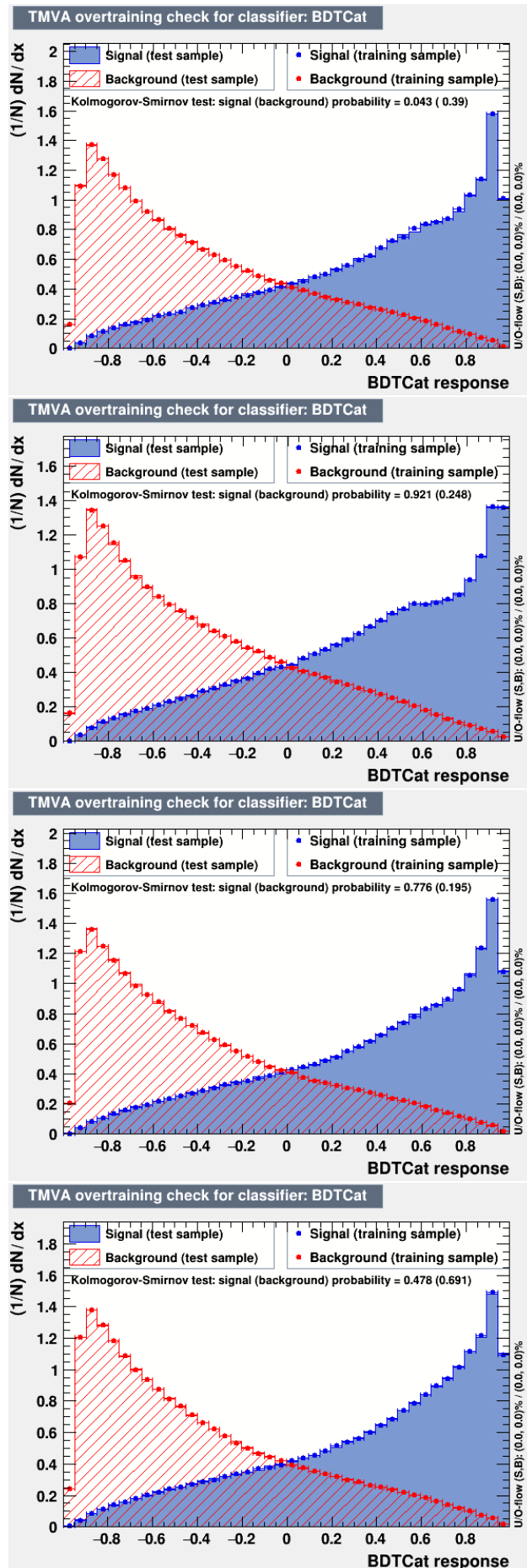


Figure 5.6: BDT response for all years starting from the top: 2016 preVFP, 2016 postVFP, 2017 and 2018 in bottom).

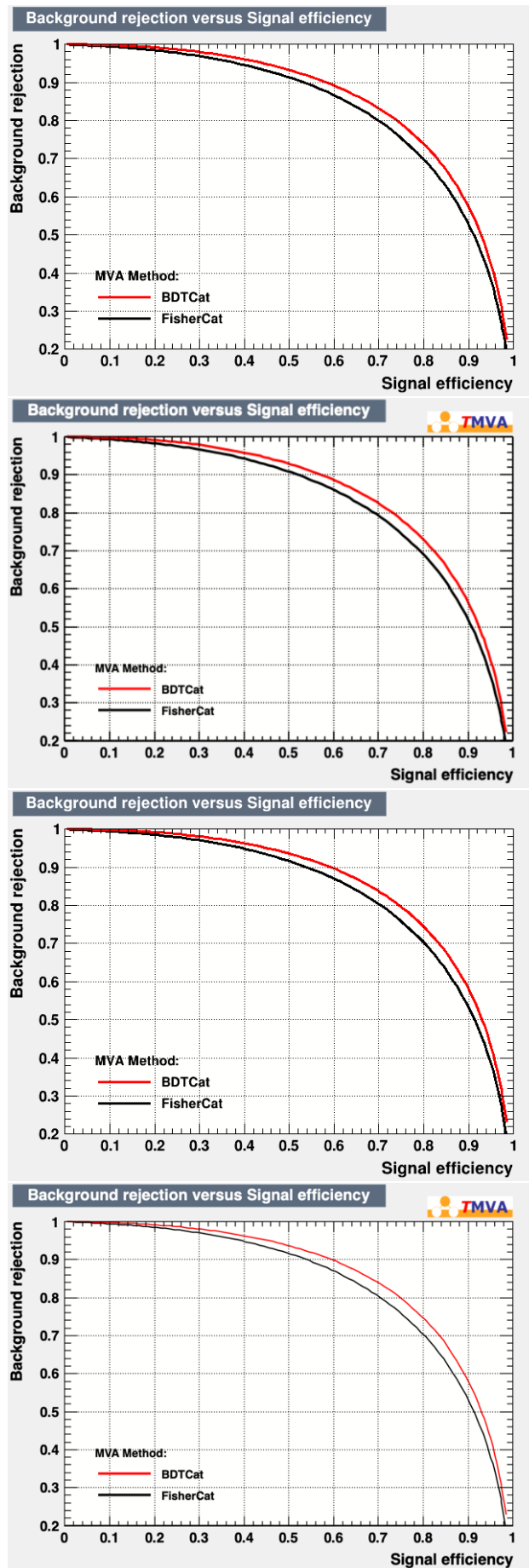


Figure 5.7: ROC curve for all years starting from the top: 2016 preVFP, 2016 postVFP, 2017 and 2018 in bottom).

### 5.4.4 Parton level

The partonic phase space to which reconstructed level data is unfolded to is constrained by the same kinematic requirements of the fiducial level region. Specifically, the top (antitop) partons must fulfil the following requirements:

Leading Parton with  $p_T > 450$  GeV, second leading Parton with  $p_T > 400$  GeV and both candidates with  $|\eta| < 2.4$ , while the invariant mass of the  $t\bar{t}$  system must be greater than 1000 GeV in order to avoid extreme events with high top (antitop)  $p_T$  and very low  $m_{t\bar{t}}$ . The aforementioned selection criteria are summarized in Table 5.14.

Table 5.14: Definition of parton-level phase space.

Observable	Requirement
Leading Parton $p_T^{t,\bar{t}}$	$> 450$ GeV
Second leading Parton $p_T^{t,\bar{t}}$	$> 400$ GeV
$ \eta^{t,\bar{t}} $	$< 2.4$
$m_{t\bar{t}}$	$> 1000$ GeV

### 5.4.5 Particle level

The particle level represents the state that consists of stable particles originating from the proton-proton collisions after the hadronization process and just before the interaction of these particles with the detector. We study these variables because the observables computed from the particles' momenta are thought to be better defined compared to the ones computed from parton information and are accompanied by smaller theoretical uncertainties. Also the associated phase space is closer to the fiducial phase space of the measurement at detector-level. In this analysis, particle jets are reconstructed from stable particles (excluding neutrinos) using the anti-kt algorithm with a distance parameter  $R = 0.8$ , same as the detector-level reconstruction. Here we should indicate that only jets originating from the primary interaction are selected. For that reason, jets that are geometrically matched within  $\Delta R < 0.4$  in  $\eta - \phi$  phase space from generated leptons are removed from the particle-jet collection. The two particle jets with the highest  $p_T$  are considered the particle-level top quark candidates for our analysis. The particle level kinematic requirements are chosen with respect to the parton phase space requirements. These requirements are summarized in Table 5.15.

Table 5.15: Definition of particle-level phase space.

Observable	Requirement
$N_{\text{jets}}$	$> 1$
Leading particle jet $p_T^{\text{jet1}}$	$> 450$ GeV
Second Leading particle jet $p_T^{\text{jet2}}$	$> 400$ GeV
$ \eta^{\text{jet1,2}} $	$< 2.4$
$m_{SD}^{\text{jet1,2}}$	(120, 220) GeV
$m_{jj}$	$> 1000$ GeV

## 5.5 Signal Extraction

The dominant background in this analysis is the QCD multijet production. There is a finite probability that ordinary jets, originating from single parton radiation, mimic the topological substructure of a top jet. The QCD multijet production is suppressed significantly using the developed BDT, which uses jet substructure information to distinguish signal from background. Also, the b-tagging requirements play a huge role in suppressing the QCD multijet production. In order to estimate the remaining contribution of the QCD multijet production we employ a data-driven technique based on the assumption that if we revert the b-tagging requirement (asking for 0 btagged subjets in the event) we a) get a pure QCD multijet sample and b) the jet kinematic properties are not affected. Our CR works as a phase space, where (theoretically) only QCD multijets exist. In reality, the CR consists of QCD multijet production events but also from  $t\bar{t}$  and subdominant background events (W+jets, etc). The subdominant background contribution can be as high as 10%, Figure 5.8. For that reason, subdominant background contributions are removed from the CR before the QCD multijet estimation.

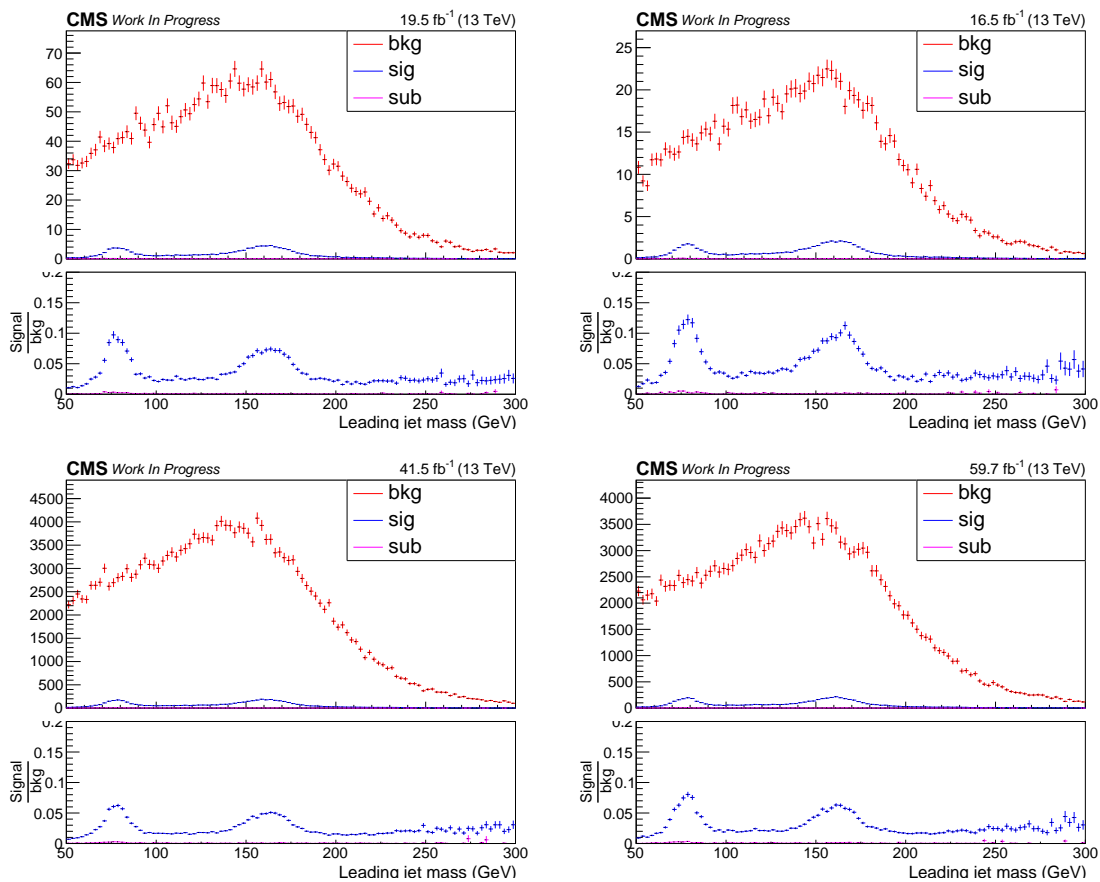


Figure 5.8: Contamination of the theoretically "pure" QCD Control Region. We present the contamination of the CR from  $t\bar{t}$  and subdominant backgrounds as a function of the SoftDropMass of the leading  $p_T$  jet. On the top (left)for 2016 preVFP, and top (right)2016 postVFP and in the bottom (left)for 2017 and (right)for 2018. On top we present the expected yields for each process in the Control Region. The ratio of the plots shows the percentage of contamination in the CR with respect to the present QCD in the CR for each year.

### 5.5.1 Inclusive cross section

Before attempting to extract the differential cross sections, it is first mandatory to measure the inclusive, fiducial cross section. This is done in an extension of our SR,  $SR_A$ , which contains a QCD dominated sideband in the top candidate mass (SoftDrop mass of the leading jet) and allows

a simultaneous fit for the  $t\bar{t}$  signal and the QCD background yields. The fit that we employ is described by the equation below:

$$D_{SR_A}(m^t) = N_{t\bar{t}}T(m^t; k_{\text{scale}}, k_{\text{res}}) + N_{\text{qcd}}(1 + k_{\text{slope}}m^t)Q_{CR_A}(m^t) + N_{\text{bkg}}B(m^t), \quad (5.3)$$

The shapes (templates)  $T(m^t)$ ,  $B(m^t)$  of the signal and subdominant backgrounds, are taken from simulation respectively. The shape  $Q(m^t)$  of the QCD is taken from the control sample in data ( $CR_A$ ). The templates of various components are shown in Figs. 5.18, 5.19, 5.20, 5.21. To account for any differences observed in the closure test of QCD for the  $m^t$  variable (Fig. 5.22), we introduce the linear modification factor  $(1 + k_{\text{slope}}m^t)$  inspired by the simulation, but with the slope parameter  $k_{\text{slope}}$  left free in the fit. The normalization factors  $N_{t\bar{t}}$ ,  $N_{\text{QCD}}$  and  $N_{\text{bkg}}$  are also left free. We finally introduce two more nuisance parameters in the  $t\bar{t}$  simulation. These two nuisances are  $k_{\text{scale}}$  and  $k_{\text{res}}$  which account for possible differences between data and simulation in the scale and resolution of the  $m^t$  parameter. The fit model is imported to the RooFit package [103] and the fit result is shown in Fig. 5.30, 5.31, 5.32, 5.33 for all years ('16,'17,'18). The fitted parameters are summarized in Tables. 5.17, 5.18, 5.19, 5.20. We observe for all the years, that the fitted  $t\bar{t}$  yield is significantly lower than the expectation. This implies that the fiducial cross section is  $\approx 35\%$  lower compared to the nominal MC (Powheg + Pythia8) predictions. We also define the  $t\bar{t}$  signal strength as a measure of agreement between the expected yield from simulation and the data.

$$r_{t\bar{t}} = \frac{N_{\text{data}}}{N_{\text{simulation}}} \quad (5.4)$$

The respective values of the  $t\bar{t}$  signal strength are given in Table 5.16:

Table 5.16: Results of the  $t\bar{t}$  signal strength for each year.

Year	Value	Error
2016 preVFP	0.691	0.028
2016 postVFP	0.640	0.029
2017	0.665	0.018
2018	0.675	0.016

Table 5.17: Results of the fit in  $SR_A$  for 2016 preVFP.

Parameter	Value	Error
$k_{\text{res}}$	1.032	0.045
$k_{\text{scale}}$	0.987	3.36e-03
$k_{\text{slope}}$	6.91e-02	0.14
$N_{\text{bkg}}$	108	183
$N_{\text{qcd}}$	1321	212
$N_{t\bar{t}}$	2543	104

## 5.5.2 Differential cross sections

In this analysis, the  $t\bar{t}$  cross section is also reported differentially in the fiducial detector phase space, in the signal region ( $SR$ ) as a function of several variables. The results are presented with respect to the following variables: the leading and second leading top jet candidate  $p_T$  and  $|y|$ , the



Table 5.18: Results of the fit in  $SR_A$  for 2016 postVFP.

Parameter	Value	Error
$k_{\text{res}}$	0.976	0.049
$k_{\text{scale}}$	0.979	3.61e-03
$k_{\text{slope}}$	0.185	0.136
$N_{\text{bkg}}$	10	848
$N_{\text{qcd}}$	951	639
$N_{t\bar{t}}$	1977	92

 Table 5.19: Results of the fit in  $SR_A$  for 2017.

Parameter	Value	Error
$k_{\text{res}}$	1.03	0.03
$k_{\text{scale}}$	1.025	2.74e-03
$k_{\text{slope}}$	0.2	0.07
$N_{\text{bkg}}$	526	159
$N_{\text{qcd}}$	2514	210
$N_{t\bar{t}}$	6008	160

$t\bar{t}$  system mass,  $p_T$  and rapidity ( $y$ ). Also part of this analysis deals with the  $\chi$  dijet as a variable of interest as well as the  $|\cos(\theta^*)|$ , the angle between the two leading jets in the Zero Momentum Frame (ZMF). The variable  $\chi$  is computed by measuring the difference of the rapidities of the two leading jets such as the corresponding rapidity in the ZMF is :

$$y^* = \frac{1}{2}(y_1 - y_2) \quad (5.5)$$

where  $y_{1,2}$  denote the rapidity of the leading and subleading selected jets respectively. Variable  $\chi$  is defined as:

$$\chi = e^{|2y^*|} = e^{|y_1 - y_2|}. \quad (5.6)$$

In order to derive the differential cross sections, the background contributions are subtracted from the data:

$$S(x) = D(x) - R_{\text{yield}} N_{\text{qcd}} Q(x) - B(x) \quad (5.7)$$

where  $x = p_T^{t1,2}, |y^{t1,2}|, m^{t\bar{t}}, p_T^{t\bar{t}}, y^{t\bar{t}}$ ,  $S(x)$  is the signal,  $D(x)$  is the measured distribution in data in SR,  $Q(x)$  is the QCD shape (taken from the data control region),  $B(x)$  is the subdominant backgrounds' contribution (both the shape and normalization are taken from simulation).  $R_{\text{yield}}$  is the yield ratio of the signal region (SR) and the fit region ( $SR_A$ ) and works as a transfer factor from the fit region ( $SR_A$ ) to the signal region, where the measurement is performed. The  $N_{\text{qcd}}$  is the fitted number of QCD events in  $SR_A$ .

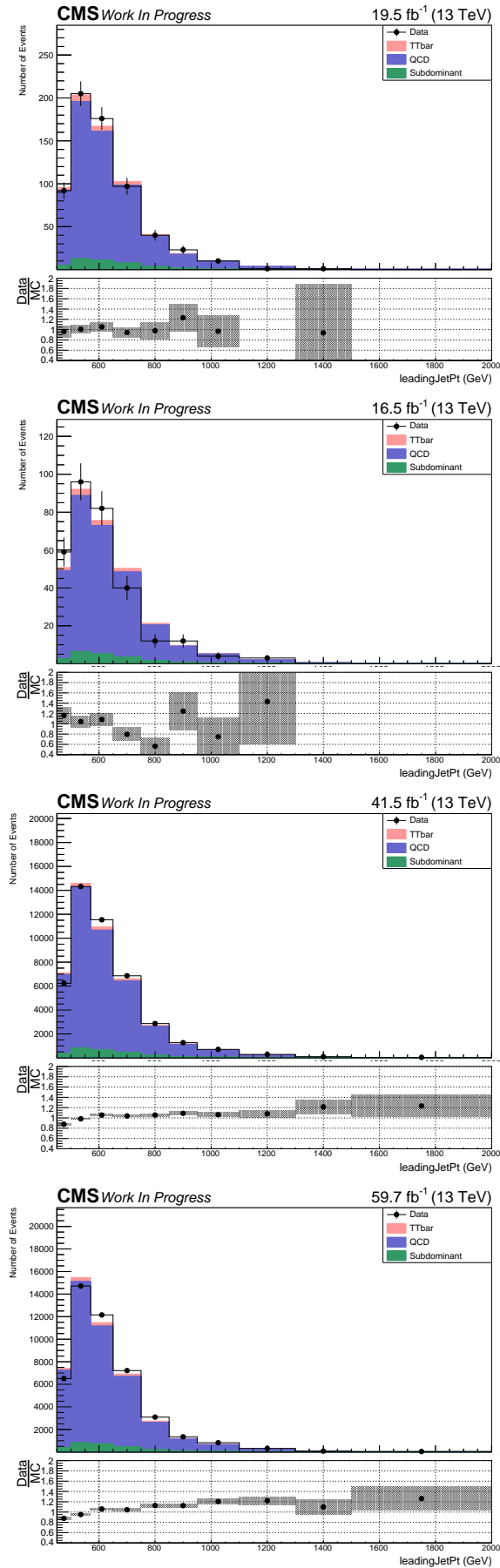
In order to apply and use the previously mentioned equation, two elements are needed. The first is the shape  $Q(x)$  of the QCD background as a function of an observable of interest  $x$ . This is taken from the QCD control sample by applying the signal region selection and reverting the b-tagging requirement (none of the selected jets is allowed to contain a b-tagged subjet). Figures 5.22, 5.23, 5.25, 5.27, 5.28, 5.29 show the closure test in MC as a function of jet and  $t\bar{t}$

Table 5.20: Results of the fit in  $SR_A$  for 2018.

Parameter	Value	Error
$k_{\text{res}}$	1.01	0.026
$k_{\text{scale}}$	0.992	0.002
$k_{\text{slope}}$	5e-02	3.24e-03
$N_{\text{bkg}}$	479.3	222
$N_{\text{qcd}}$	3900	251
$N_{t\bar{t}}$	7664	177

system variables. One can observe that within statistical precision of the simulation, the shapes are compatible, regardless the b-tagging requirement.

The other element needed for the differential cross section measurement is the QCD background normalization. This element is determined from the fit to the data in  $SR_A$ . However, since the fit region ( $SR_A$ ) is an extension of the signal region, a transfer factor  $R_{\text{yield}} = \frac{N_{SR}}{N_{SR_A}}$  is needed in order to transfer the variable of interest (QCD normalization) to the signal region.


 Figure 5.9: Leading jet  $p_T$  distributions in the QCD control region for all years.

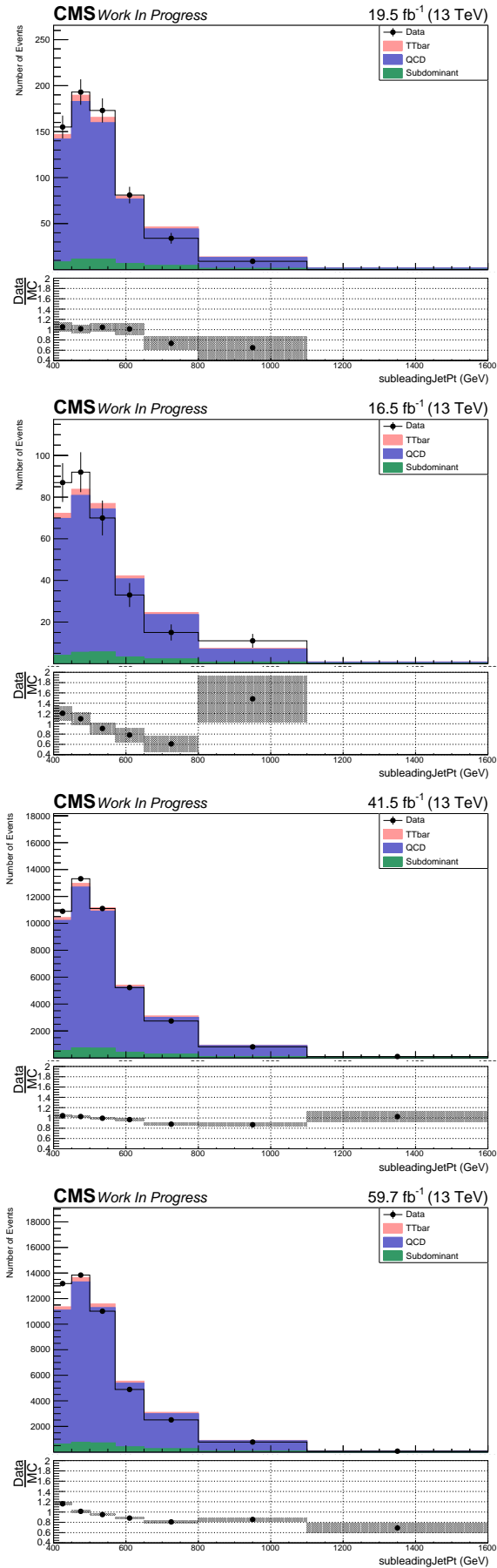
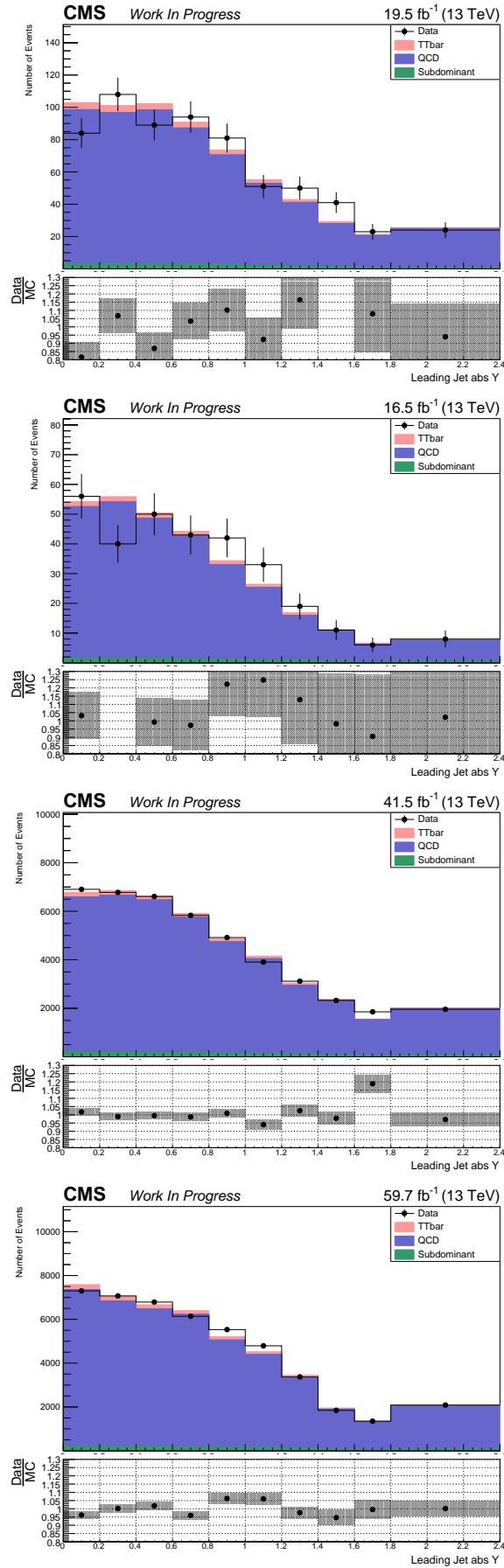


Figure 5.10: Second leading jet  $p_T$  distributions in the QCD control region for all years.


 Figure 5.11: Leading jet absolute rapidity  $y$  distributions in the QCD control region for all years.

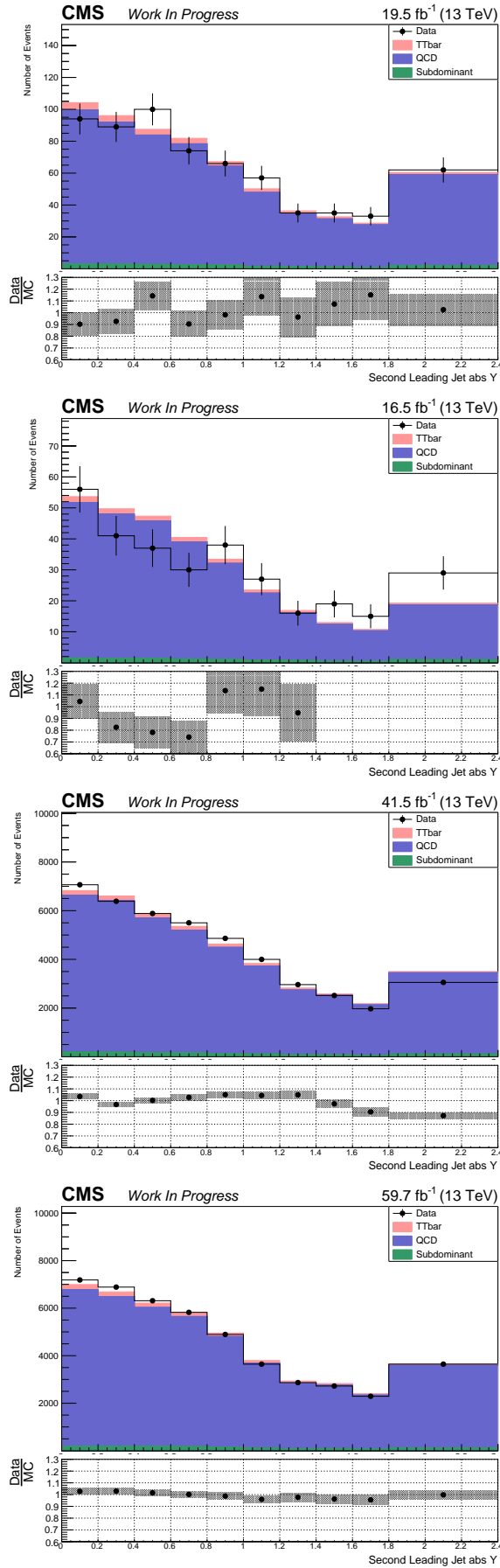
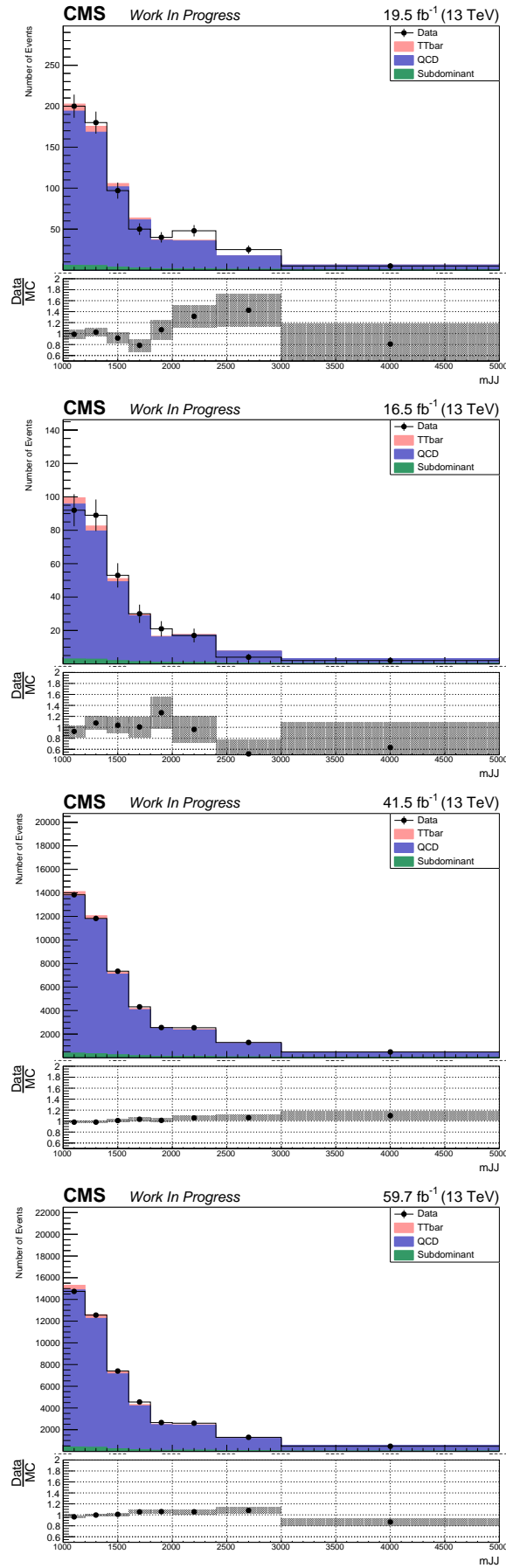


Figure 5.12: Second leading jet absolute rapidity  $y$  distributions in the QCD control region for all years.


 Figure 5.13: Dijet mass  $m_{t\bar{t}}$  distributions in the QCD control region for all years.

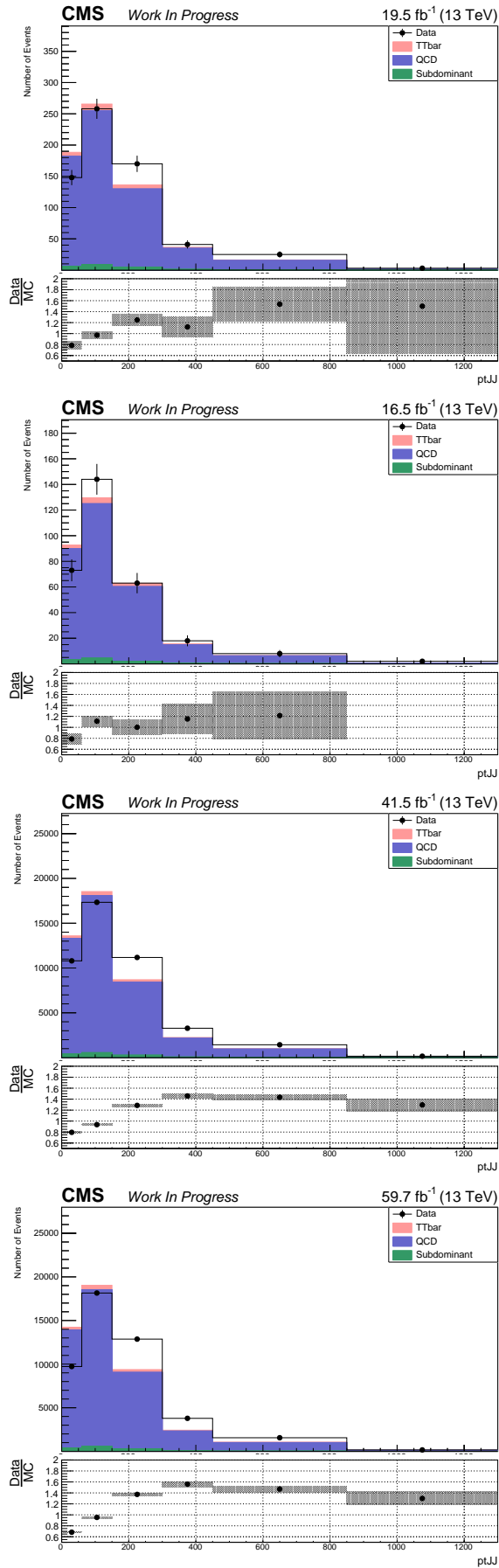
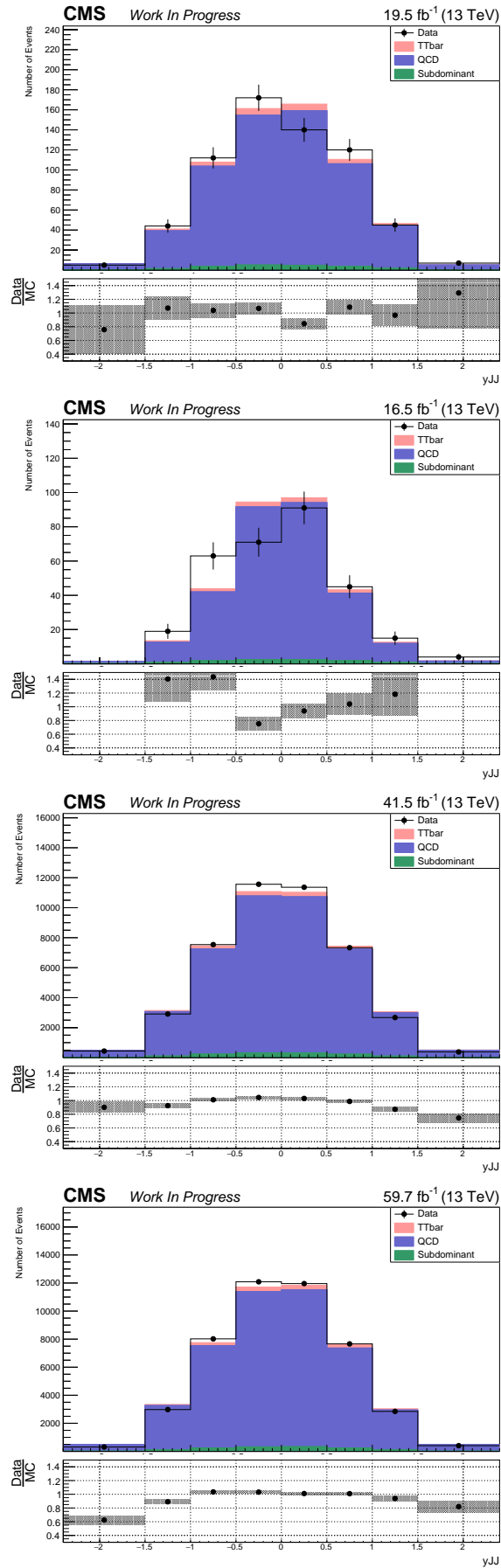
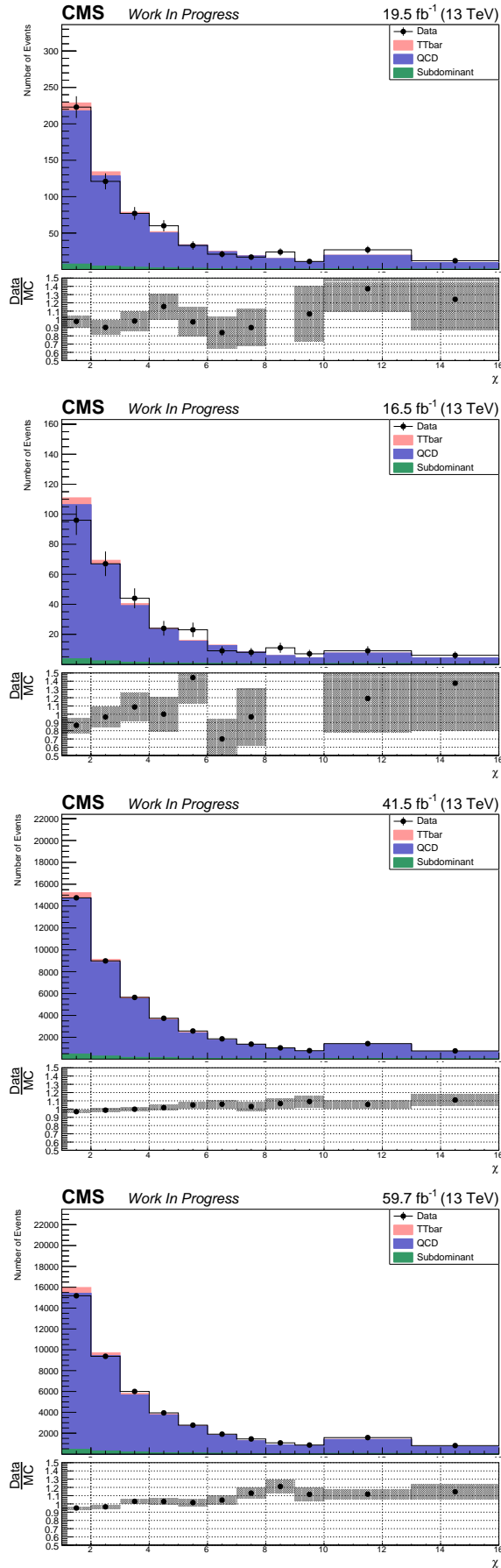
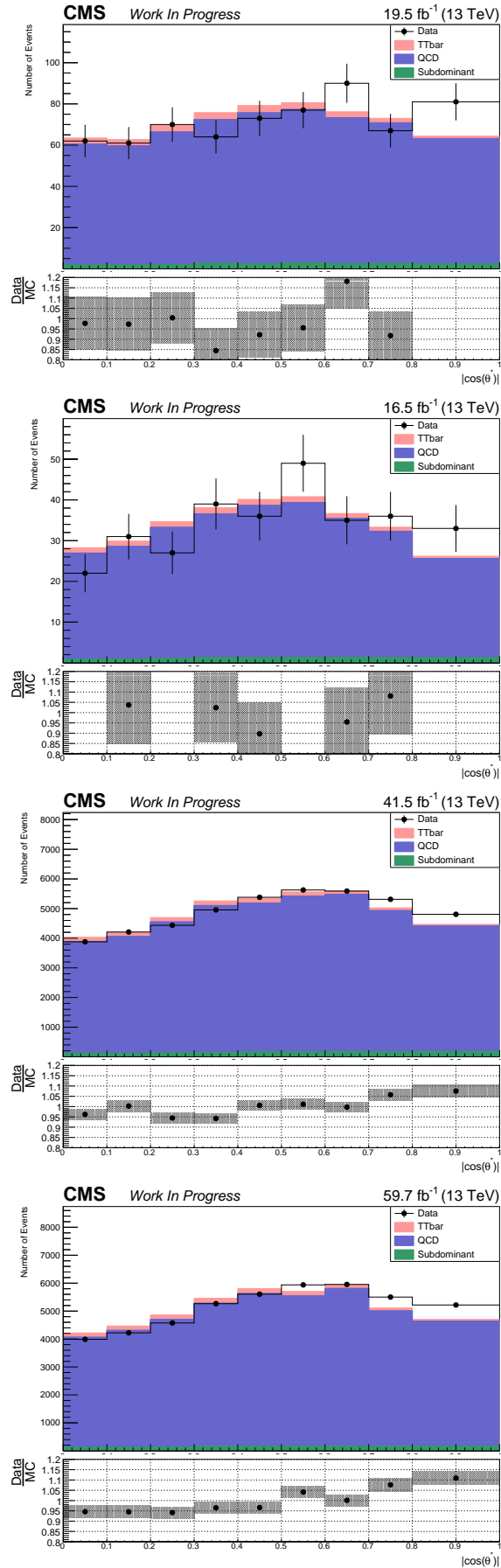


Figure 5.14: Dijet  $p_T^{t\bar{t}}$  distributions in the QCD control region for all years.




 Figure 5.15: Dijet rapidity  $y^{t\bar{t}}$  distributions in the QCD control region for all years.


 Figure 5.16: Dijet angular  $\chi$  distributions in the QCD control region for all years.


 Figure 5.17: Leading jet  $|\cos(\theta^*)|$  distributions in the QCD control region for all years.

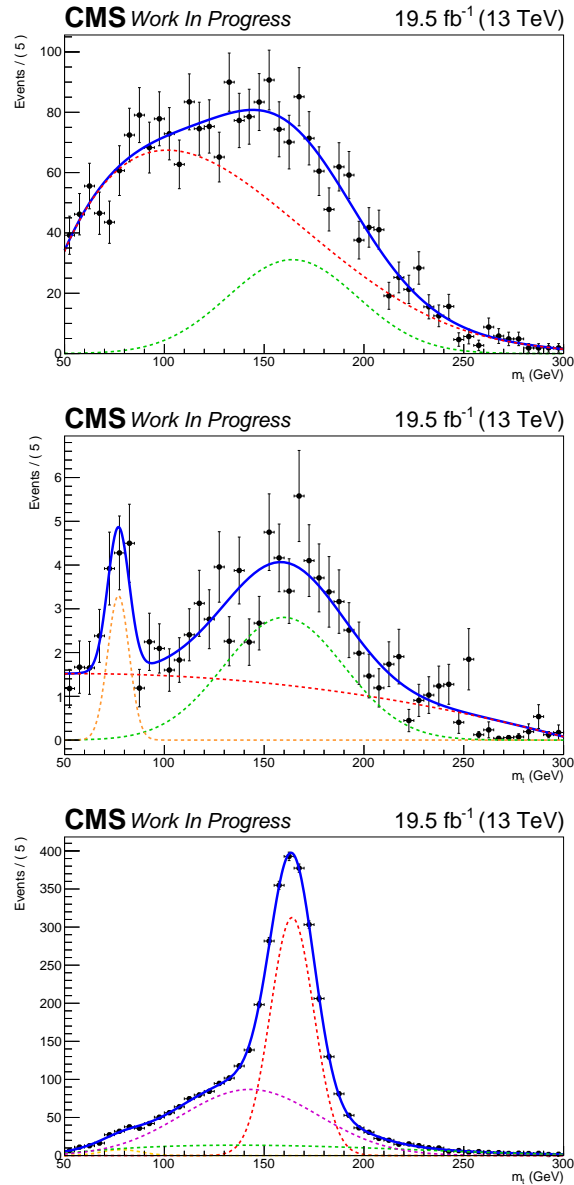


Figure 5.18: Templates of QCD (left), taken from data, of the subdominant backgrounds (center), taken from the simulation and templates of the  $t\bar{t}$  (right) taken from simulation for 2016 preVFP. The different lines show the individual components (frozen in the fit) used to describe the shapes. The QCD shape is composed of a smooth polynomial and a Gaussian, while the shape of the subdominant backgrounds contains a smooth polynomial and two Gaussians (one describes the W resonance from the single top and WJets processes and the other describes the broader peak from the kinematic selections). On the (right), template of the  $t\bar{t}$  signal taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shape. The shape consists of a smooth polynomial and two Gaussians (one describes the W resonance from unmerged top decays and the other describes the fully merged top resonance).

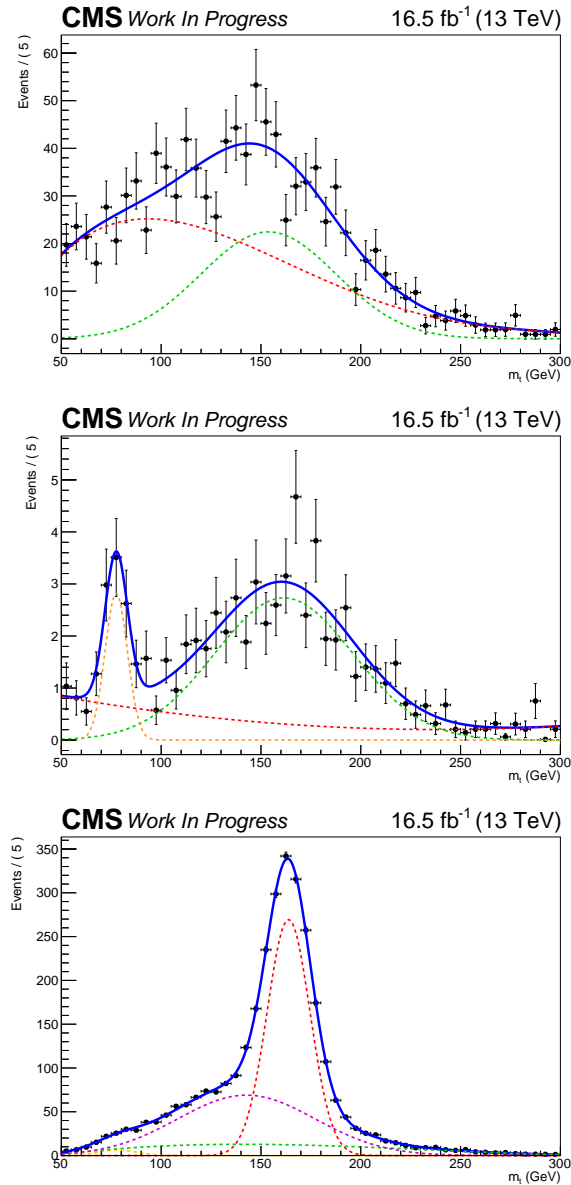


Figure 5.19: Templates of QCD ((left)), taken from data, of the subdominant backgrounds (center), taken from the simulation and templates of the  $t\bar{t}$  ((right)) taken from simulation for 2016 postVFP. The different lines show the individual components (frozen in the fit) used to describe the shapes. The QCD shape is composed of a smooth polynomial and a Gaussian, while the shape of the subdominant backgrounds contains a smooth polynomial and two Gaussians (one describes the W resonance from the single top and WJets processes and the other describes the broader peak from the kinematic selections). On the (right), template of the  $t\bar{t}$  signal taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shape. The shape consists of a smooth polynomial and two Gaussians (one describes the W resonance from unmerged top decays and the other describes the fully merged top resonance).

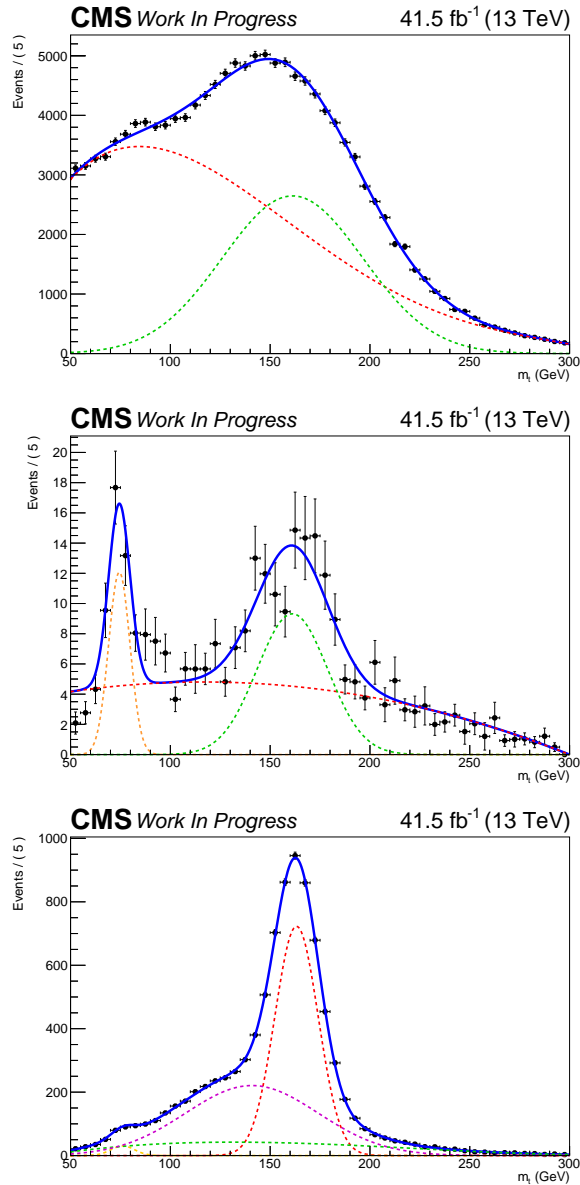


Figure 5.20: Templates of QCD (left), taken from data, of the subdominant backgrounds (center), taken from the simulation and templates of the  $t\bar{t}$  (right) taken from simulation for 2017. The different lines show the individual components (frozen in the fit) used to describe the shapes. The QCD shape is composed of a smooth polynomial and a Gaussian, while the shape of the subdominant backgrounds contains a smooth polynomial and two Gaussians (one describes the W resonance from the single top and WJets processes and the other describes the broader peak from the kinematic selections). On the (right), template of the  $t\bar{t}$  signal taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shape. The shape consists of a smooth polynomial and two Gaussians (one describes the W resonance from unmerged top decays and the other describes the fully merged top resonance).

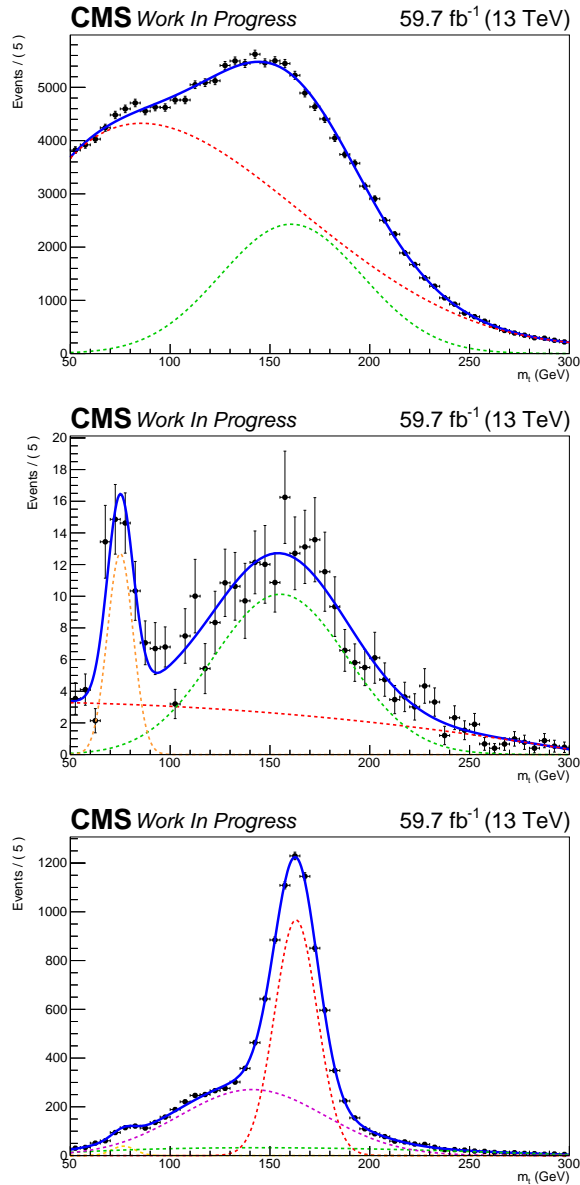


Figure 5.21: Templates of QCD ((left)), taken from data, of the subdominant backgrounds (center), taken from the simulation and templates of the  $t\bar{t}$  ((right)) taken from simulation for 2018. The different lines show the individual components (frozen in the fit) used to describe the shapes. The QCD shape is composed of a smooth polynomial and a Gaussian, while the shape of the subdominant backgrounds contains a smooth polynomial and two Gaussians (one describes the W resonance from the single top and WJets processes and the other describes the broader peak from the kinematic selections). On the (right), template of the  $t\bar{t}$  signal taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shape. The shape consists of a smooth polynomial and two Gaussians (one describes the W resonance from unmerged top decays and the other describes the fully merged top resonance).

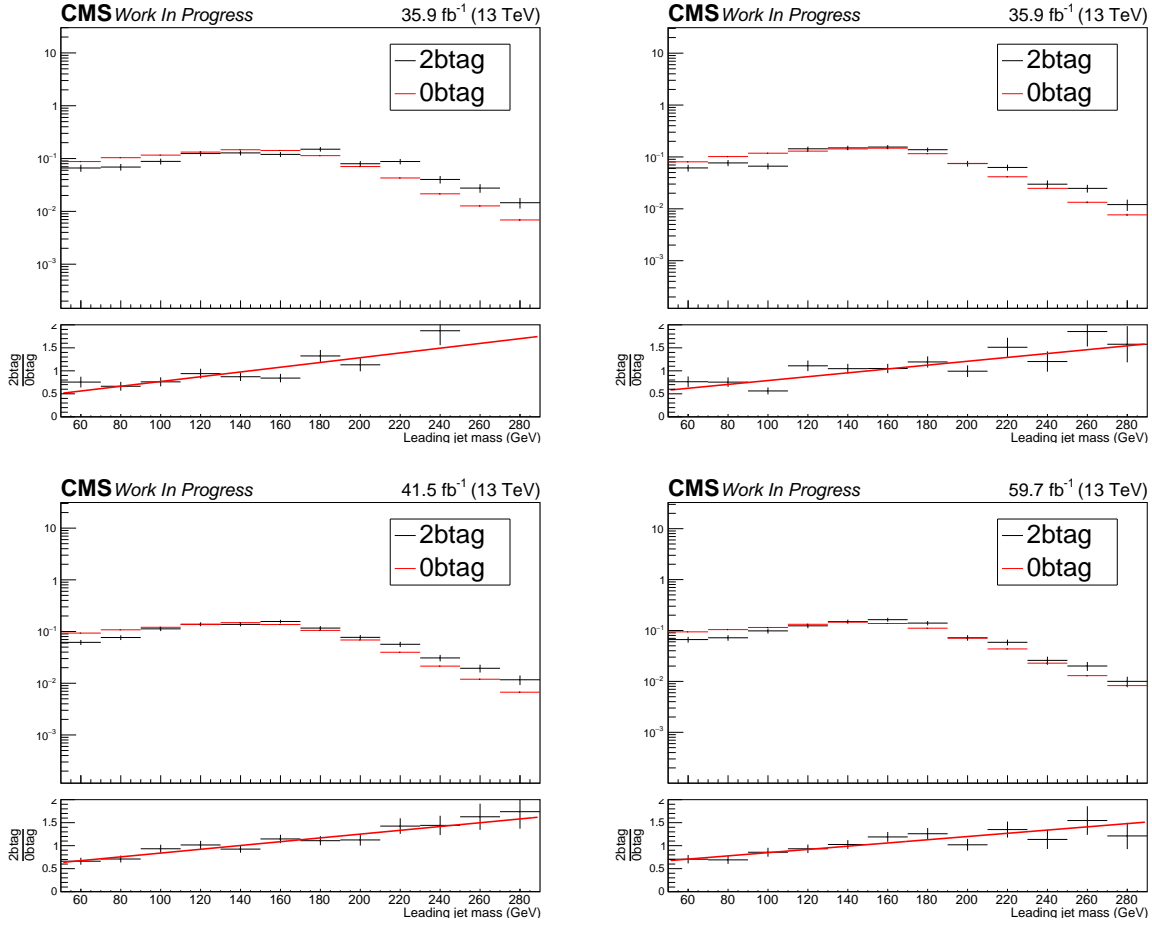


Figure 5.22: Closure test in the QCD simulation for the shape of the leading (left) and second leading (right)  $p_T$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.



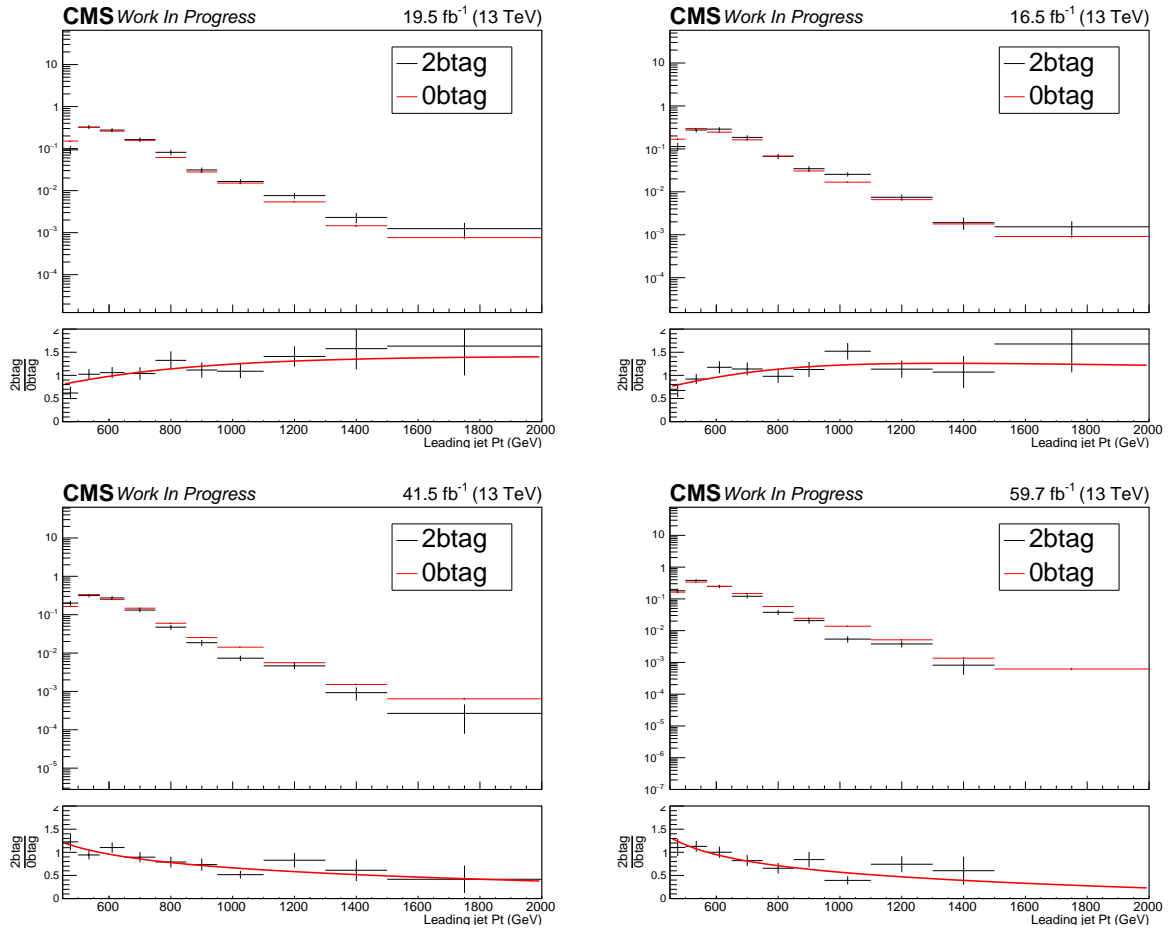


Figure 5.23: Closure test in the QCD simulation for the shape of leading  $p_T$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

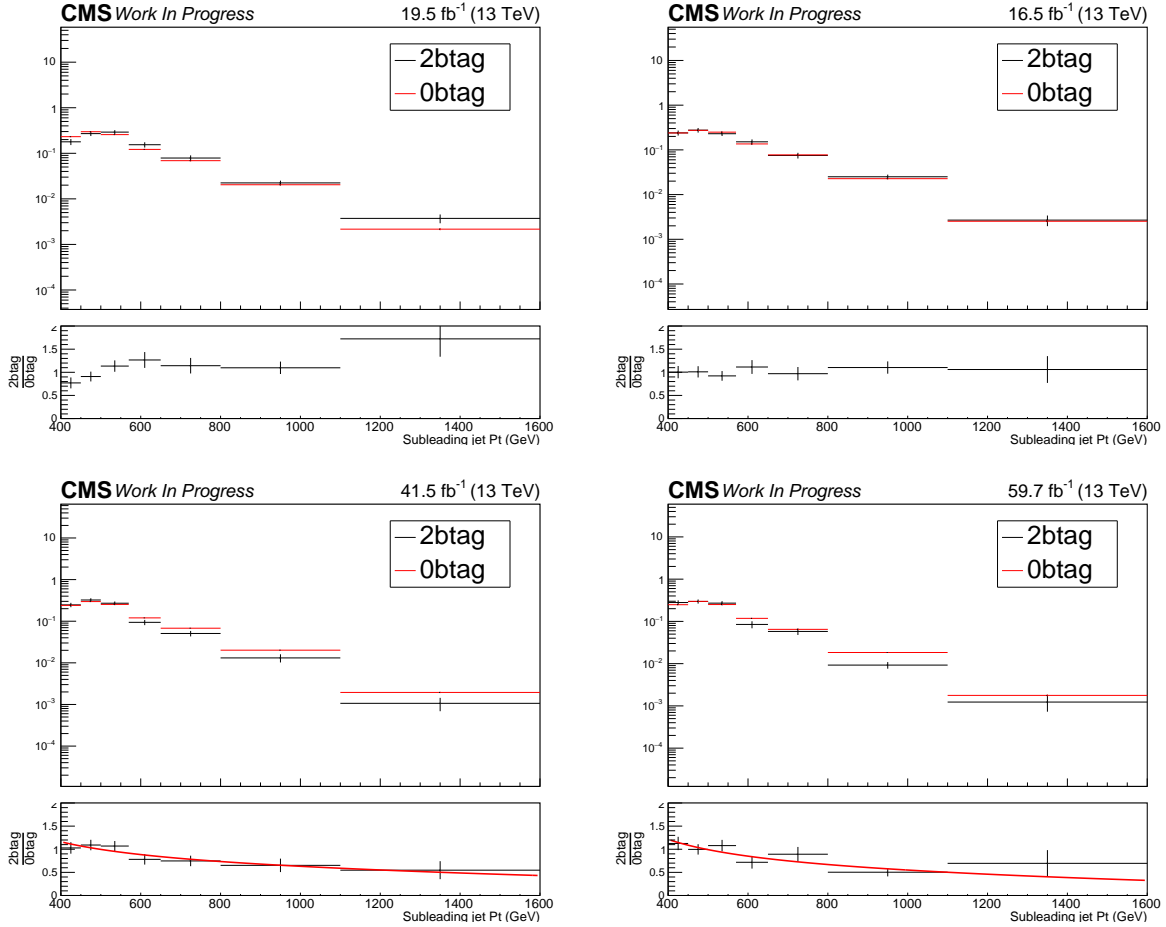


Figure 5.24: Closure test in the QCD simulation for the shape of second leading  $p_T$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

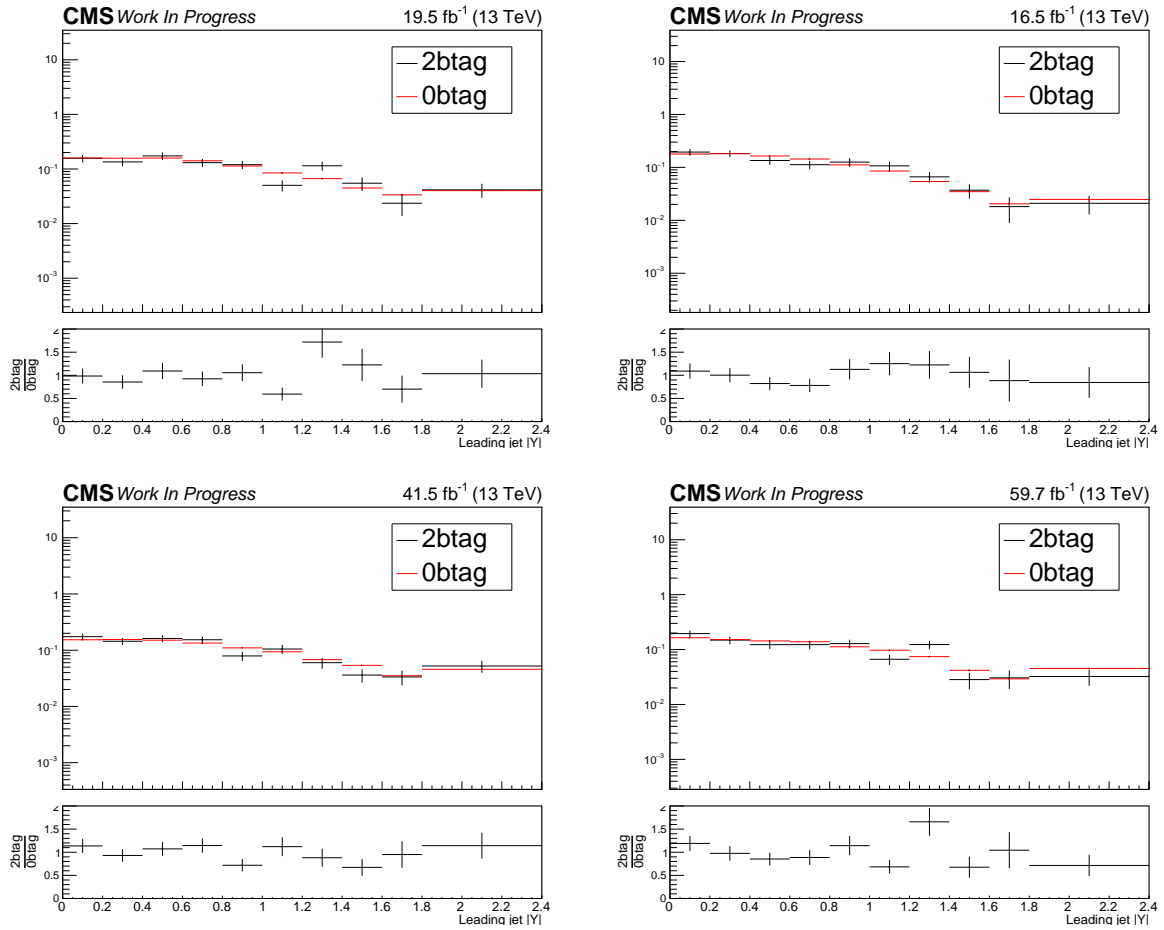


Figure 5.25: Closure test in the QCD simulation for the shape of the leading absolute rapidity  $|Y|$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

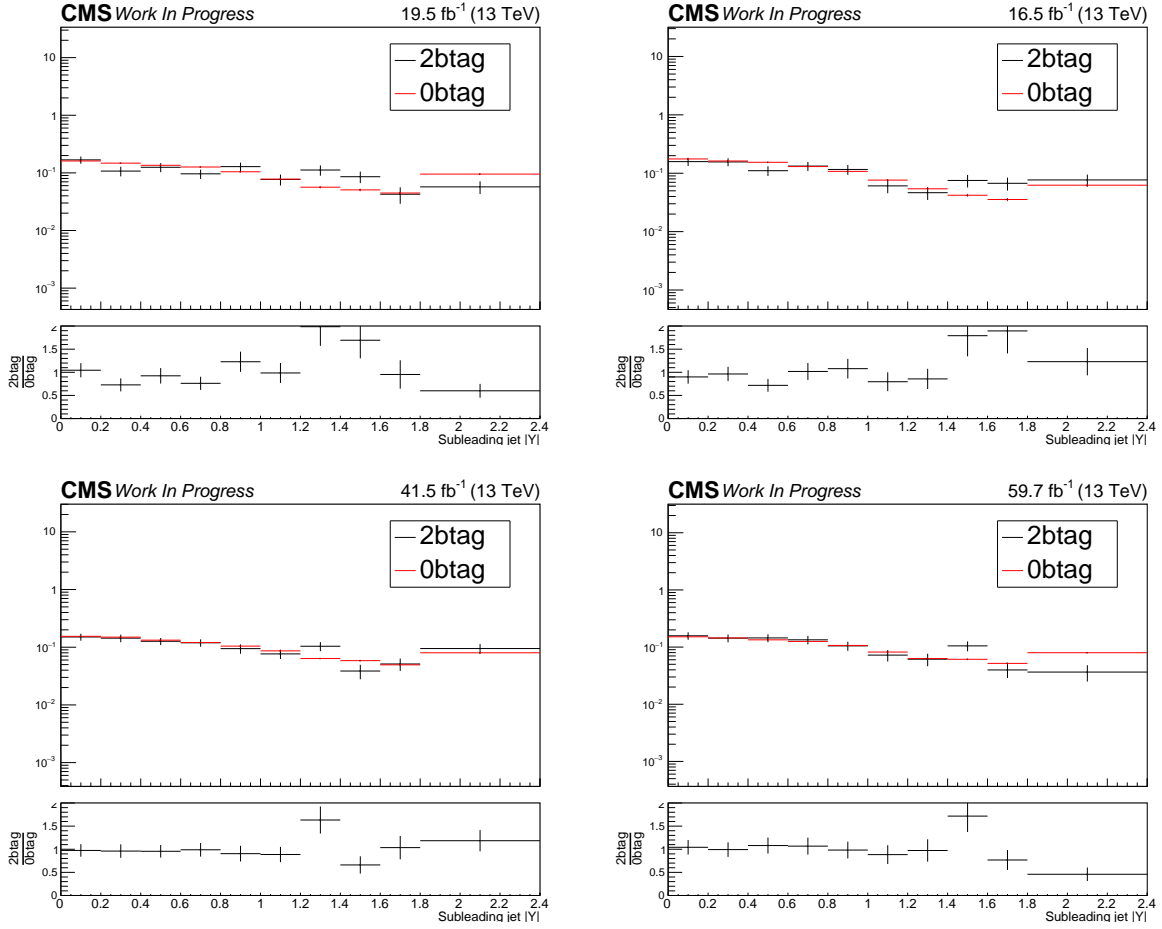


Figure 5.26: Closure test in the QCD simulation for the shape of the second leading absolute rapidity  $|y|$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

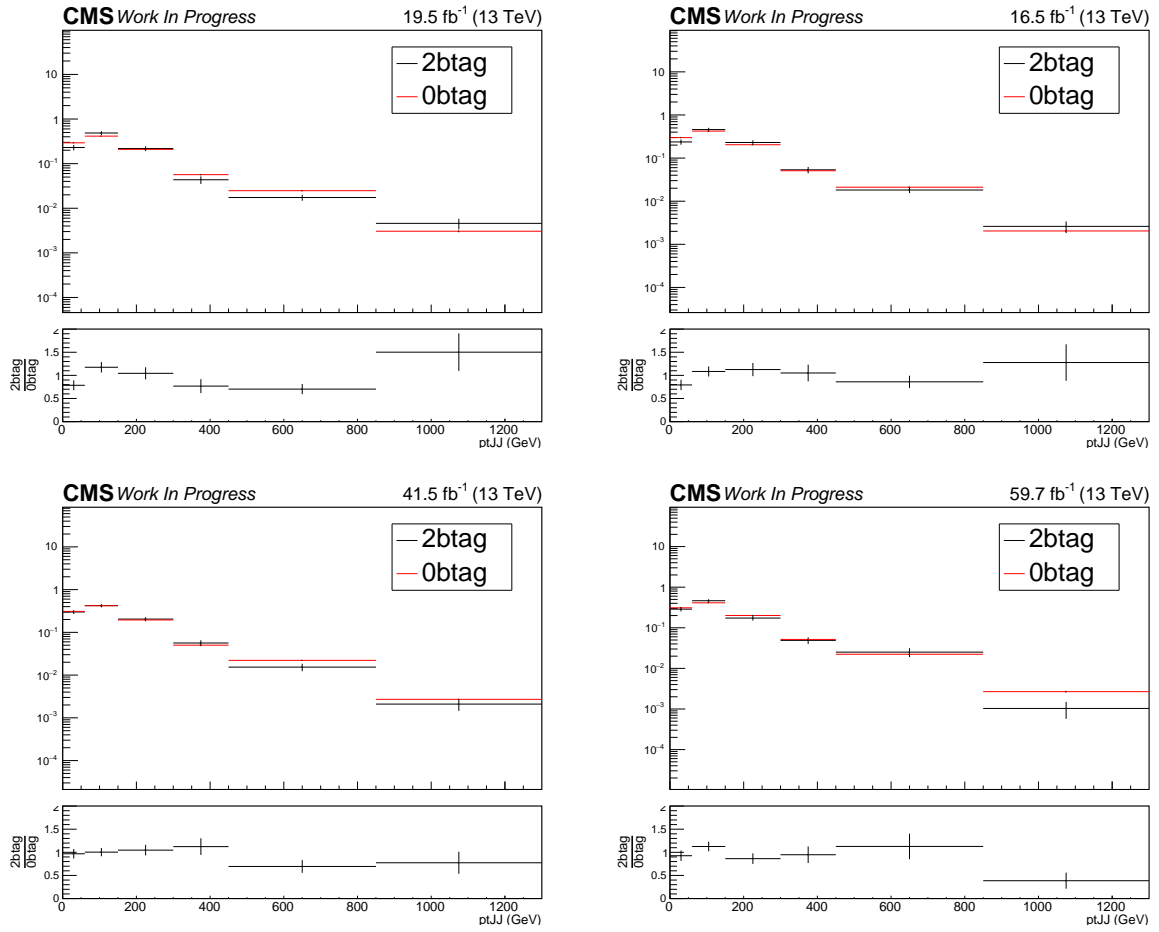


Figure 5.27: Closure test in the QCD simulation for the shape of the dijet  $pt_{JJ}$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

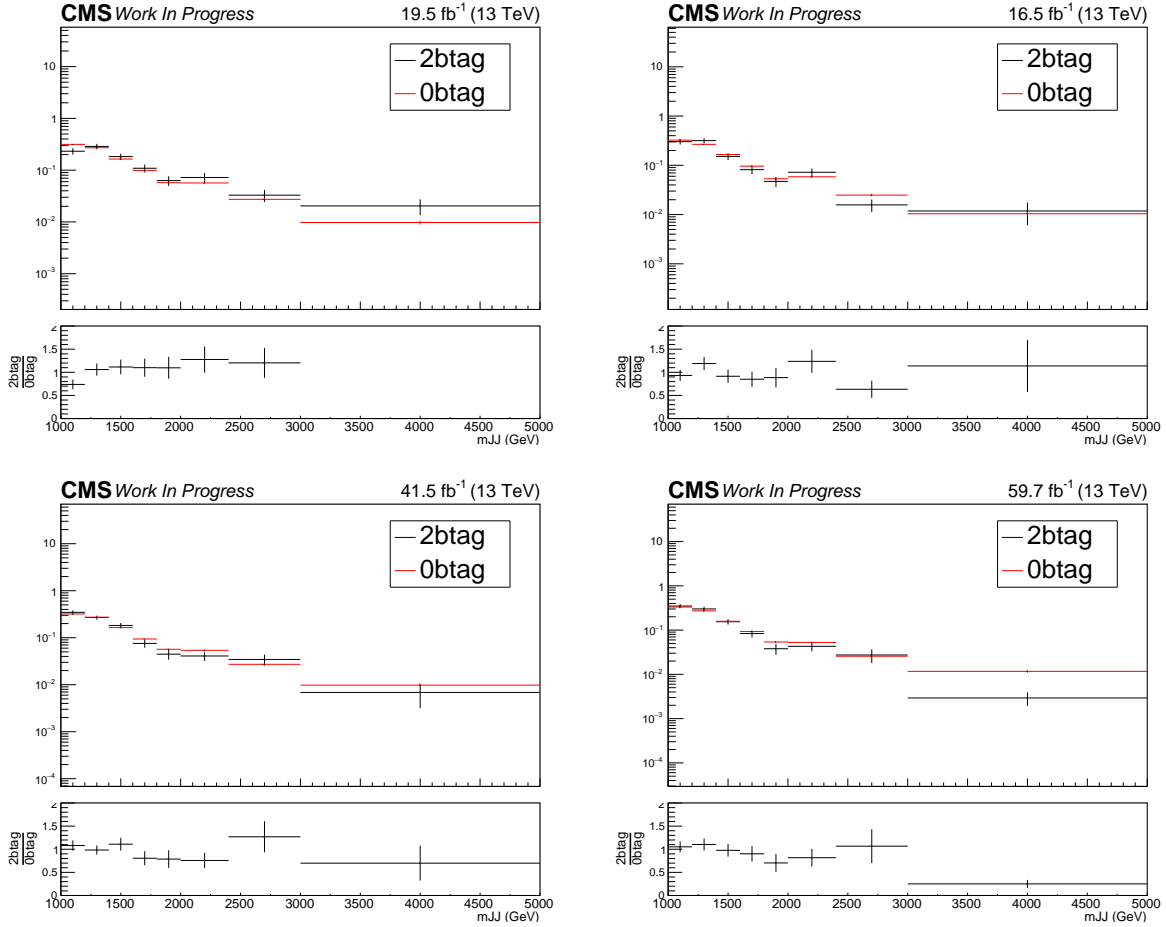


Figure 5.28: Closure test in the QCD simulation for the shape of the dijet  $t\bar{t}$  mass  $m_{JJ}$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

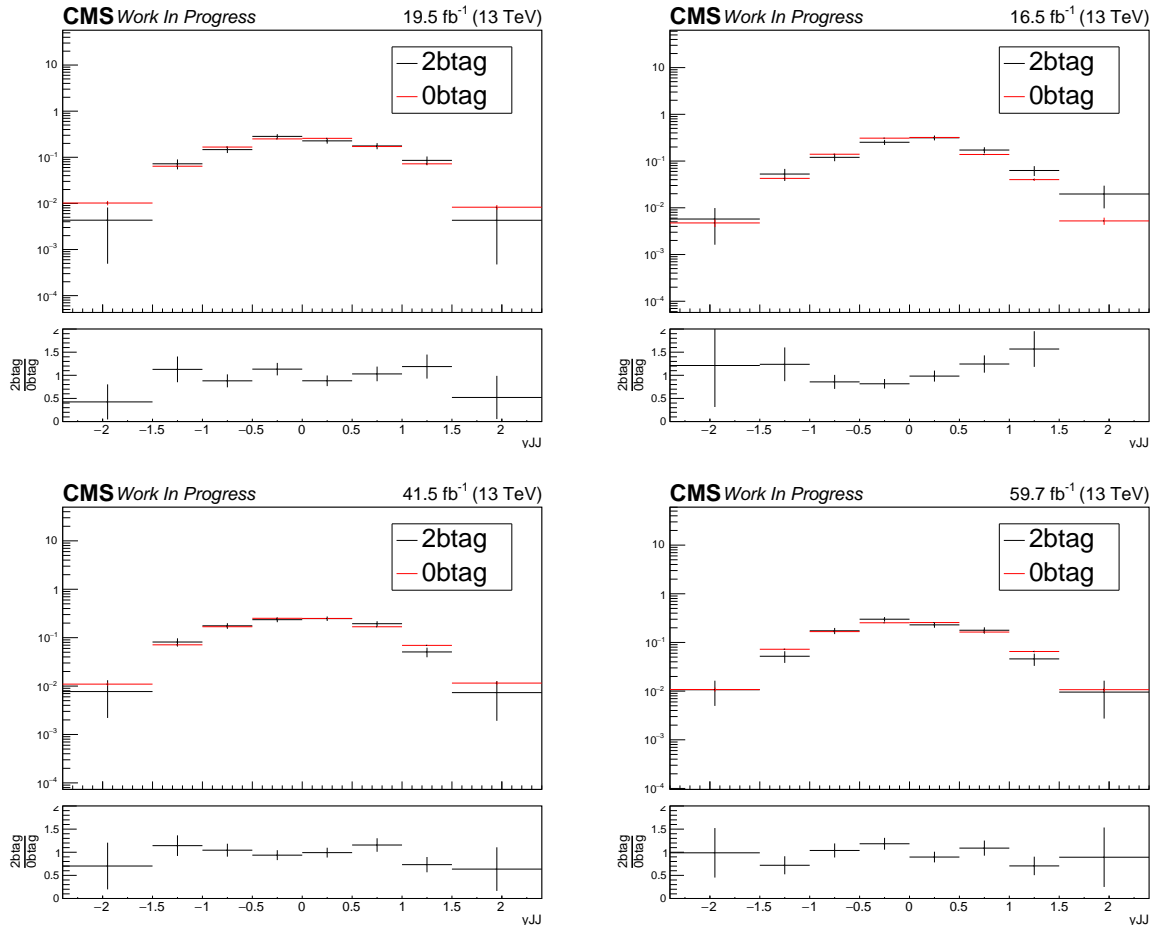


Figure 5.29: Closure test in the QCD simulation for the shape of the dijet  $t\bar{t}$  rapidity  $y_{JJ}$  variable in the two possible b-tagging requirements (none, or both jets contain a b-tagged subjet). All shapes have been corrected with a fit taken from the ratio of the QCD simulation in the different phase spaces.

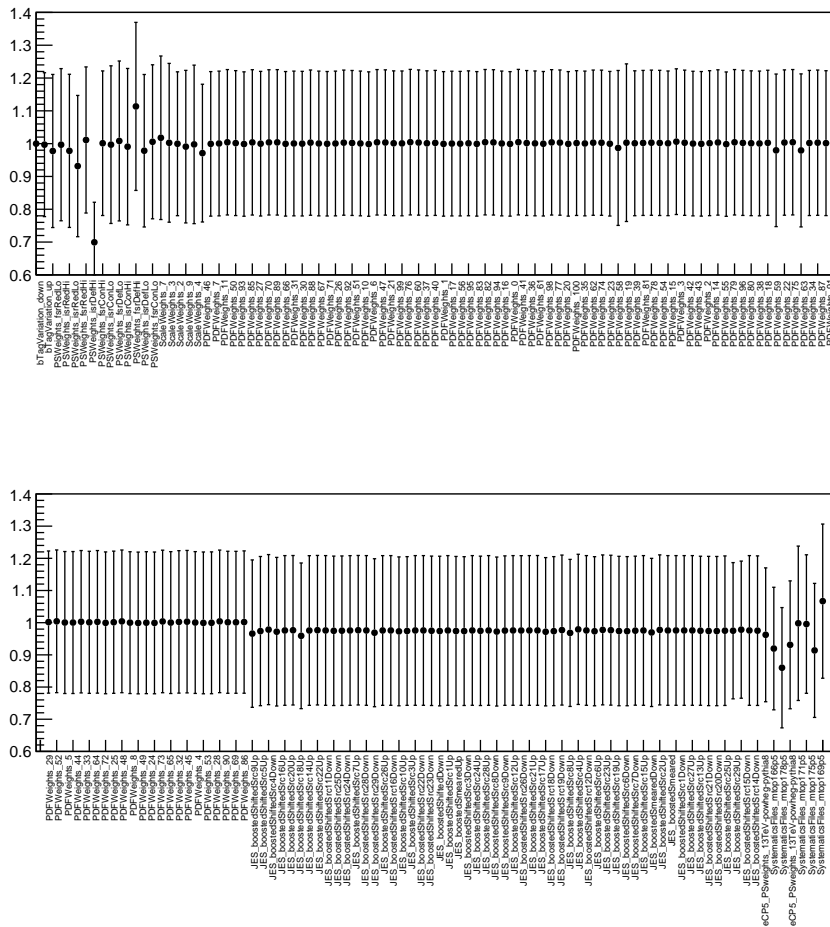
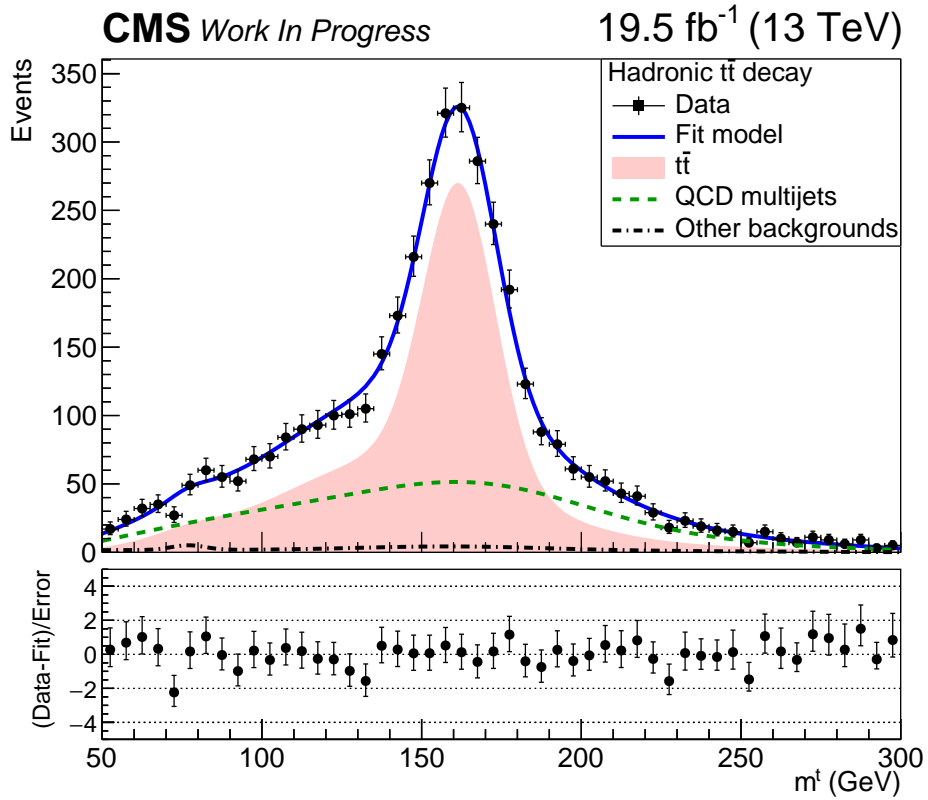


Figure 5.30: Result of the template fit on data for 2016 preVFP. The red line shows the  $t\bar{t}$  contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds.



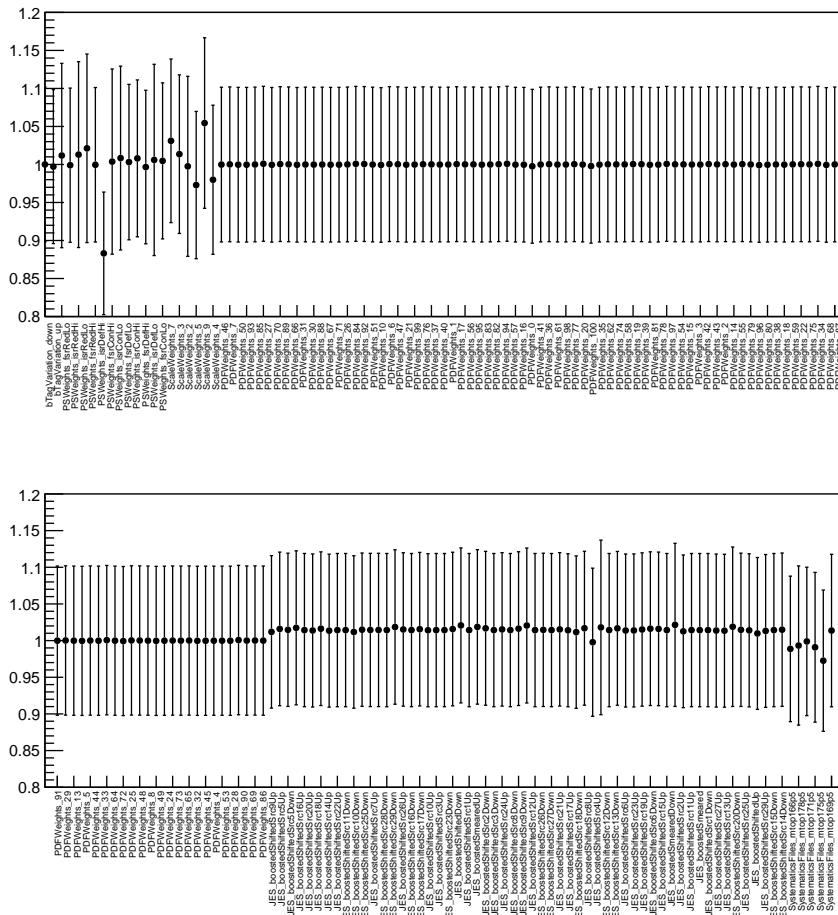
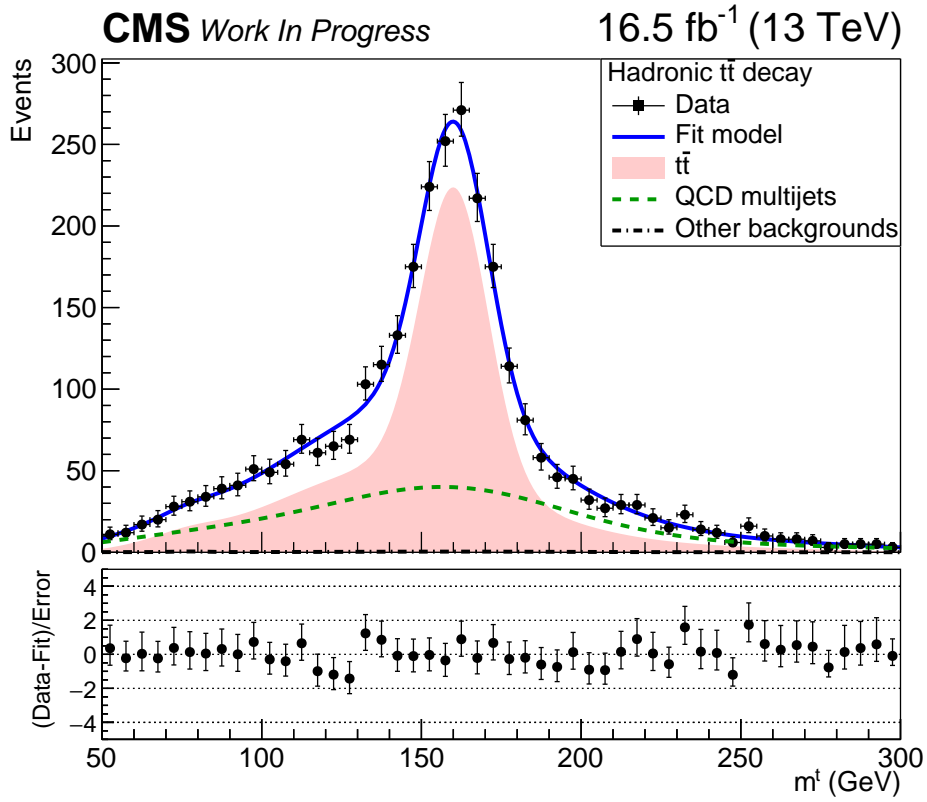


Figure 5.31: Result of the template fit on data for 2016. The red line shows the  $t\bar{t}$  contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds.

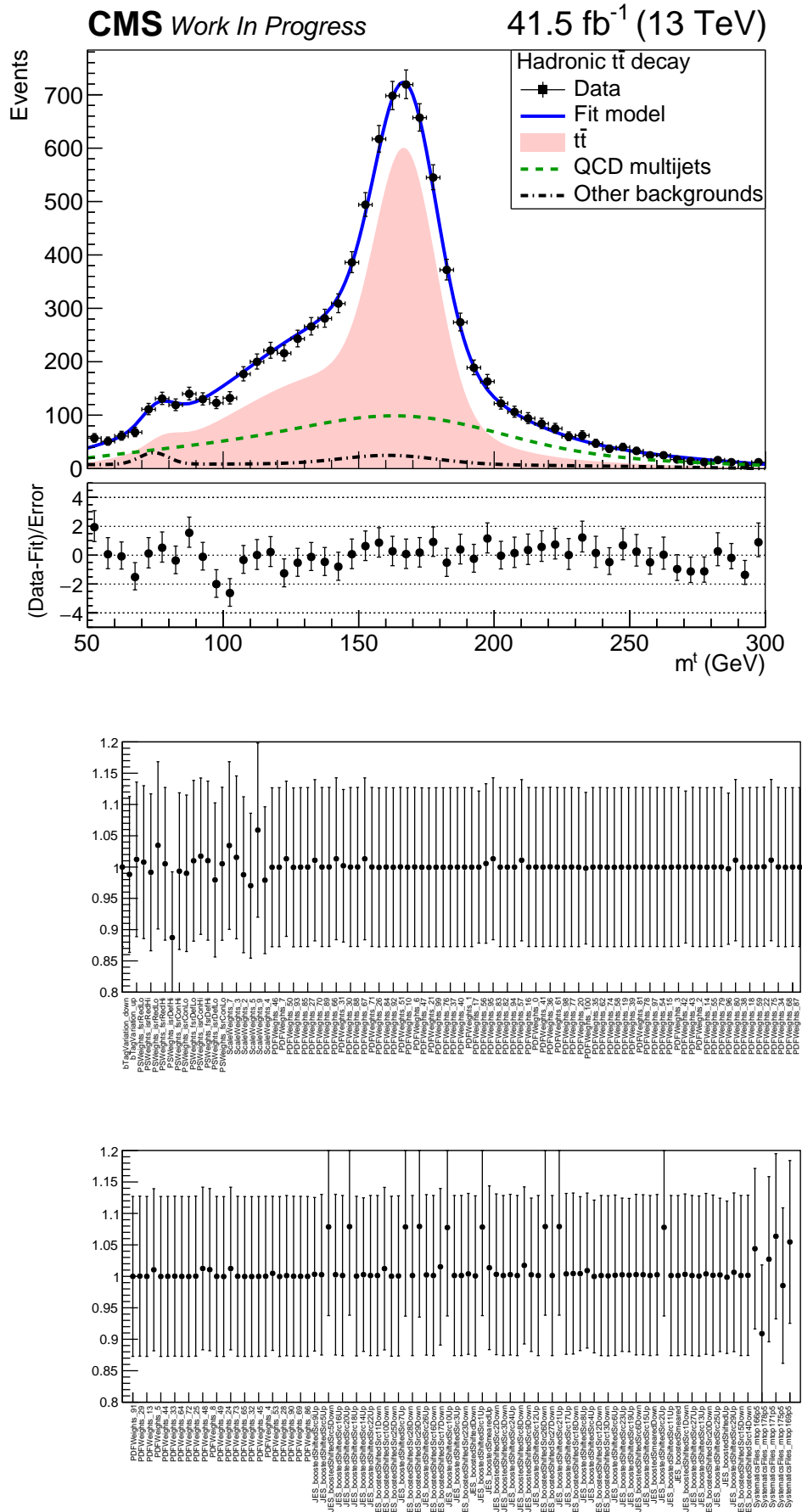


Figure 5.32: Result of the template fit on data for 2017. The red line shows the  $t\bar{t}$  contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds.

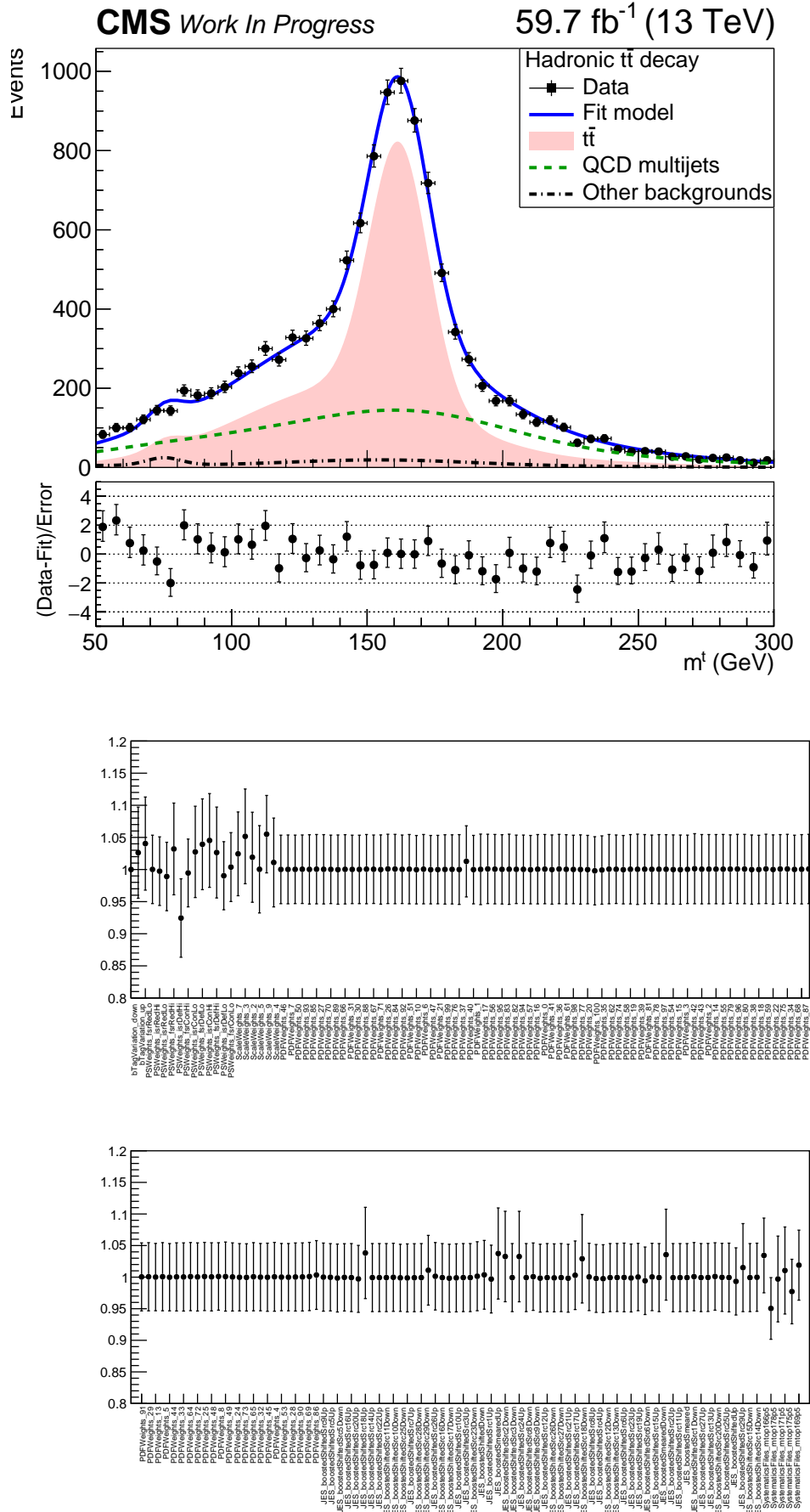


Figure 5.33: Result of the template fit on data for 2018. The red line shows the  $t\bar{t}$  contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds.

## 5.6 Data vs Monte Carlo

In this section we present a comparison between data and Monte Carlo for various observables. It should be noted that the  $t\bar{t}$  normalization is scaled by a factor 0.69 for 2016 preVFP, 0.64 for 2016 postVFP 0.64 for 2017 and 0.65 for 2018, consistent with the findings of Section 5.5. The QCD background yield is adjusted such that the total Monte Carlo events are equal to the events in data. Table 5.21 shows the event yields for the various processes after applying the baseline selection plus the requirement that both AK8 jets contain a b-tagged subjet. Table 5.21 shows the expected and observed event yields in the signal region for all analyzed years. Also, a Data vs MC comparison is provided in Figures 5.34-5.43.

Table 5.21: Expected and observed event yields in the signal region for all analyzed years.

Process	Yield 2016 preVFP	Yield 2016 postVFP	Yield 2017	2018
$t\bar{t}$	2252	1952	5358	6840
QCD	434	329	1182	1280
Subdominant (W+jets, Z+jets, Single Top)	10	8	38	47
Data	2187	1654	4818	6205

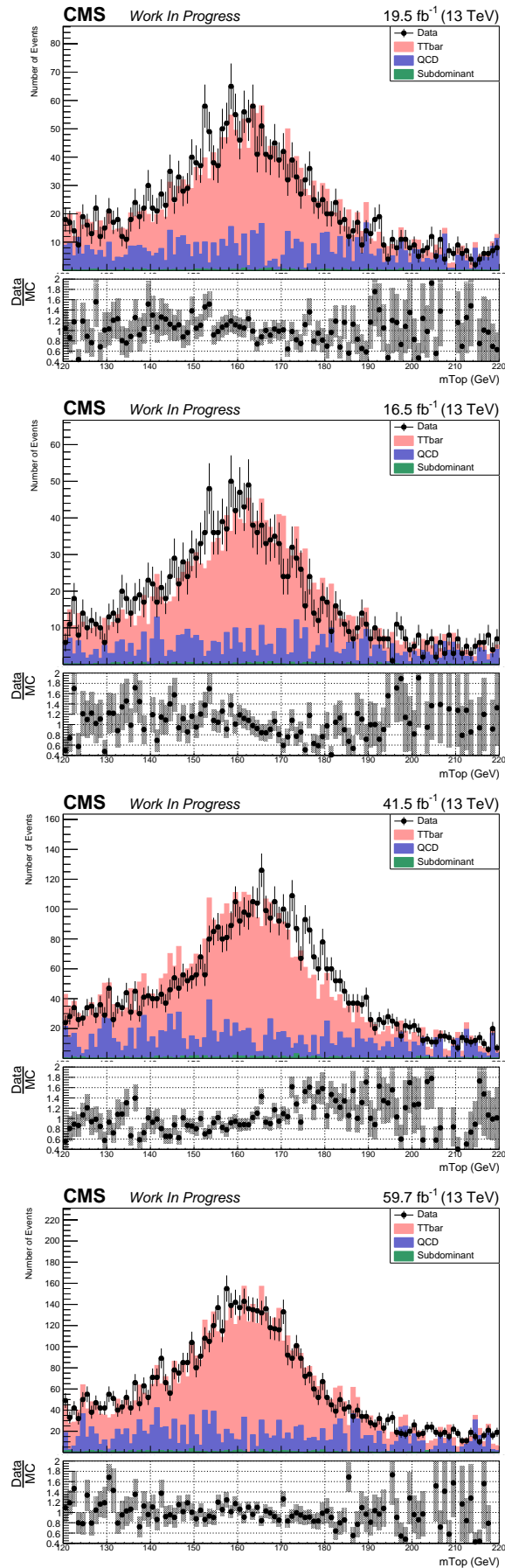


Figure 5.34: SoftDrop mass of the leading AK8 jet after the baseline selection with both AK8 jets containing a b-tagged subjet, with the  $BDT > 0.0$  cut (2016 preVFP top (left), 2016 postVFP (right)),  $> 0.1$  for 2017 bottom (left) and  $> 0.2$  for 2018 bottom (right).

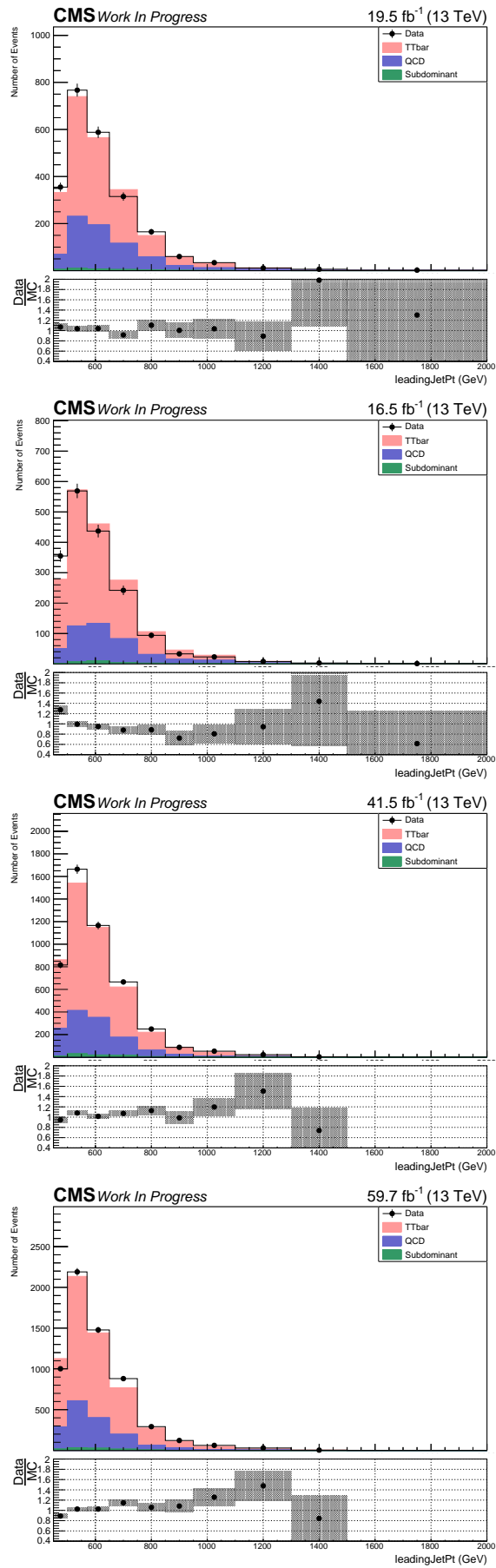


Figure 5.35: Distributions of the absolute leading and second jet  $p_T$  in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018.

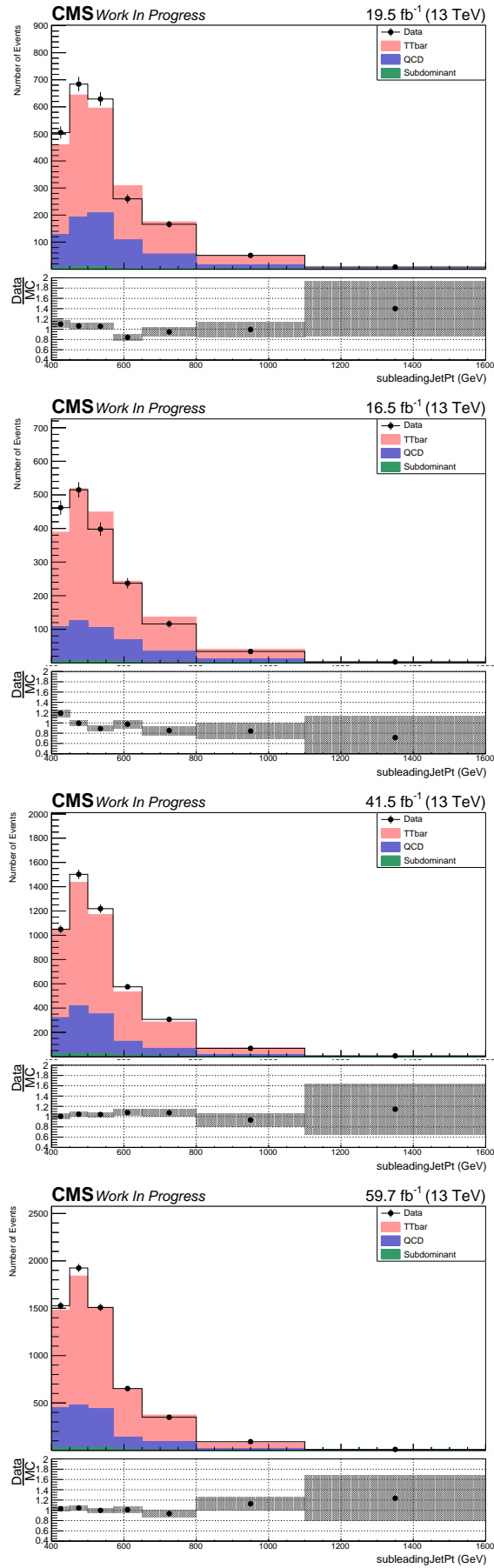


Figure 5.36: Distributions of the absolute second leading jet  $p_T$  in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018.

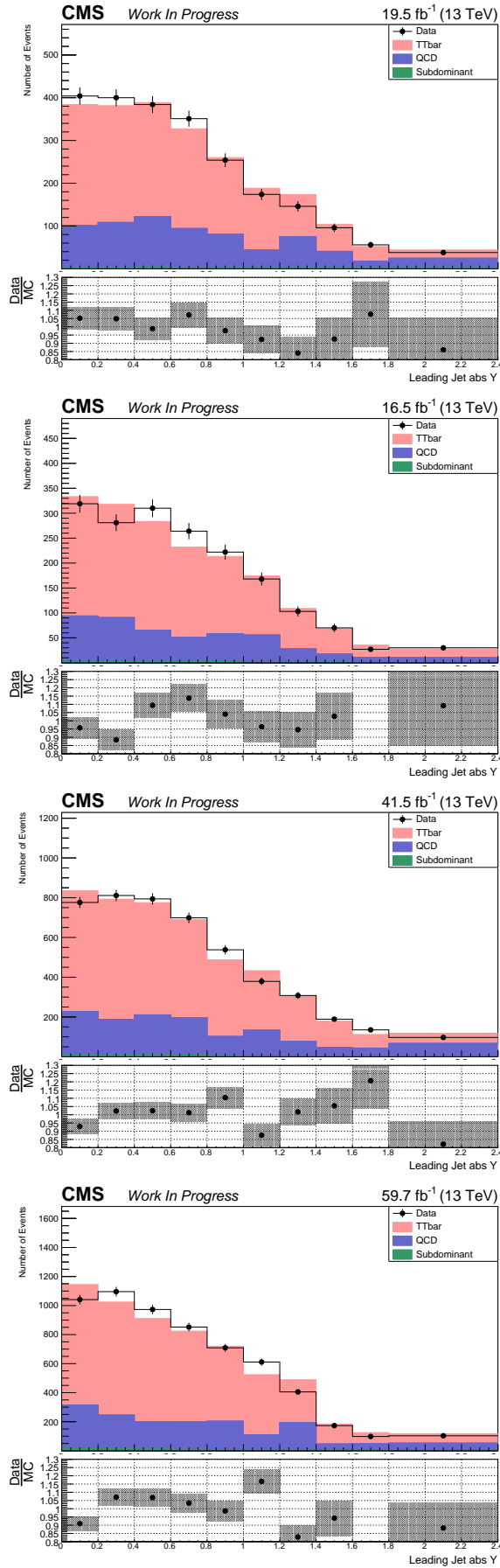


Figure 5.37: Distributions of the absolute leading jet rapidity  $y$  in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018.



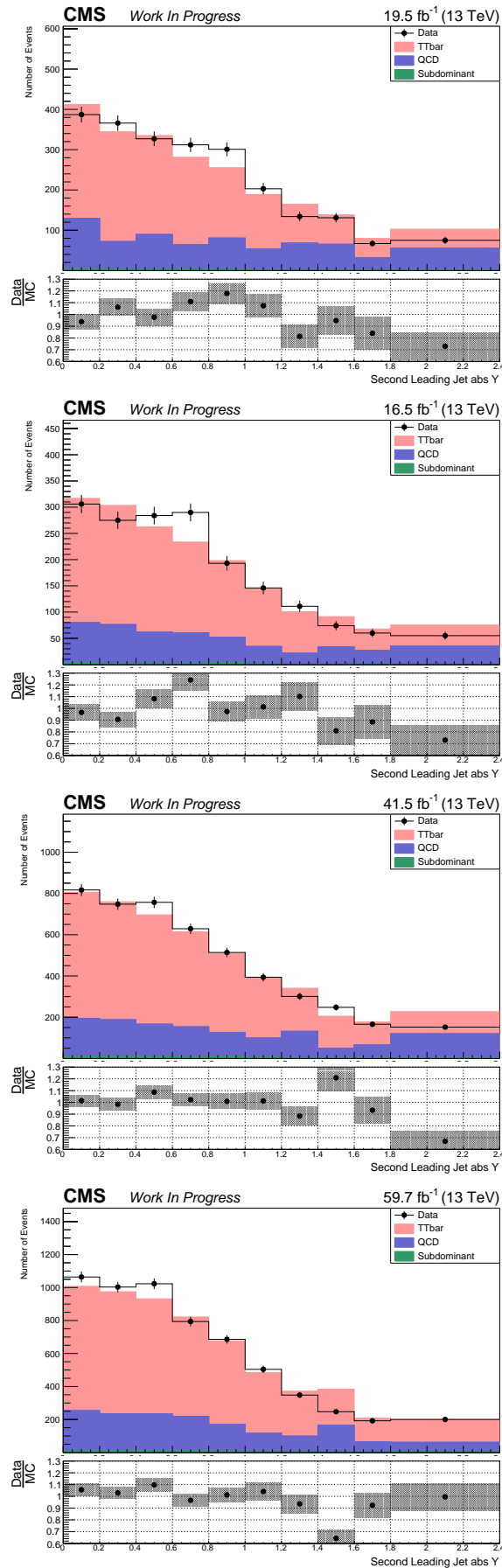


Figure 5.38: Distributions of the absolute second leading jet rapidity  $y$  in the signal region. Starting from top: 2016 preVFP, 2016 postVFP, 2017, and 2018.

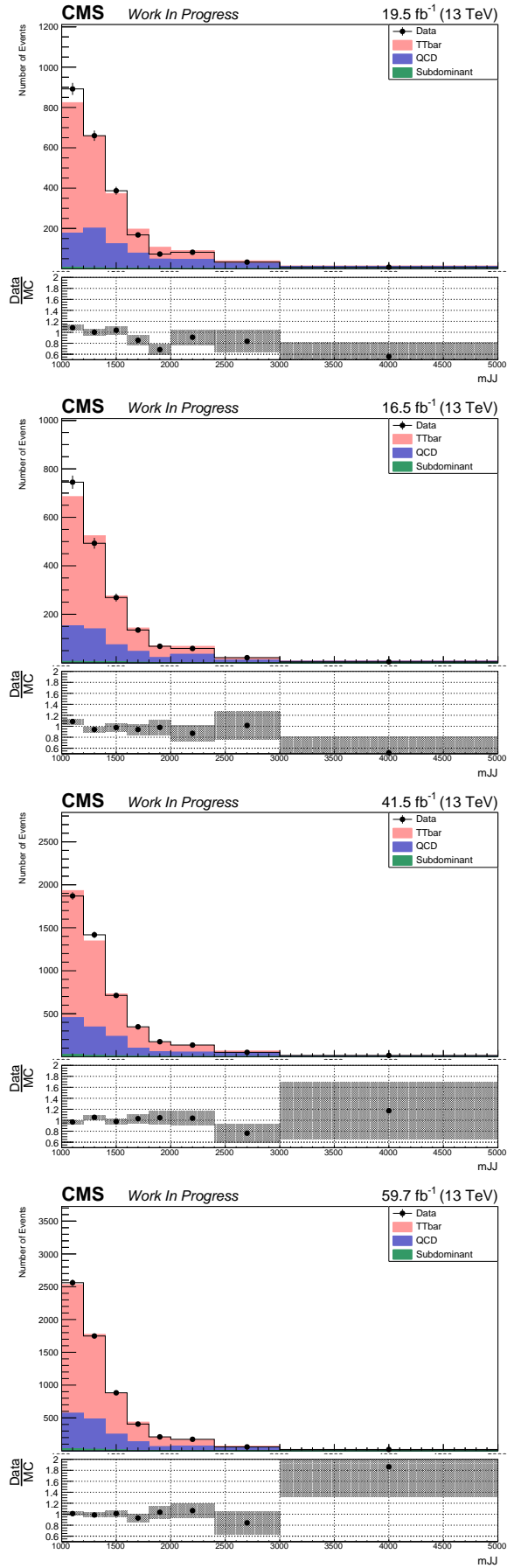


Figure 5.39: Distributions of the  $t\bar{t}$  system mass in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018).

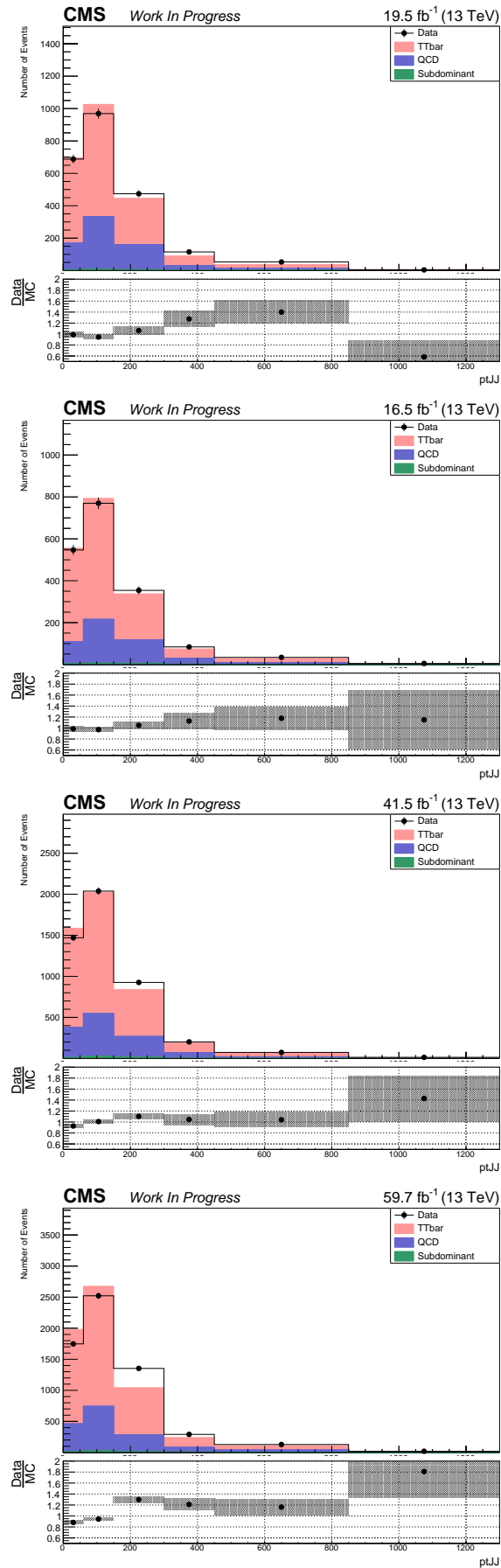


Figure 5.40: Distributions of the  $t\bar{t}$  system  $p_T$  in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018).

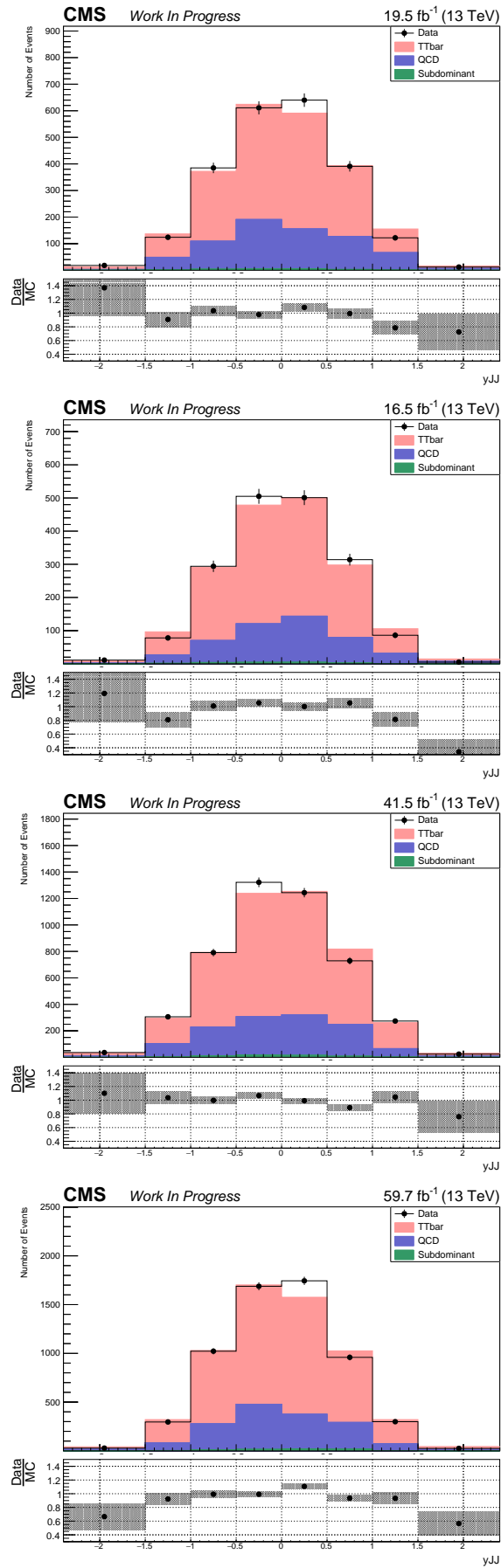


Figure 5.41: Distribution of the  $t\bar{t}$  system rapidity  $y$  in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018).

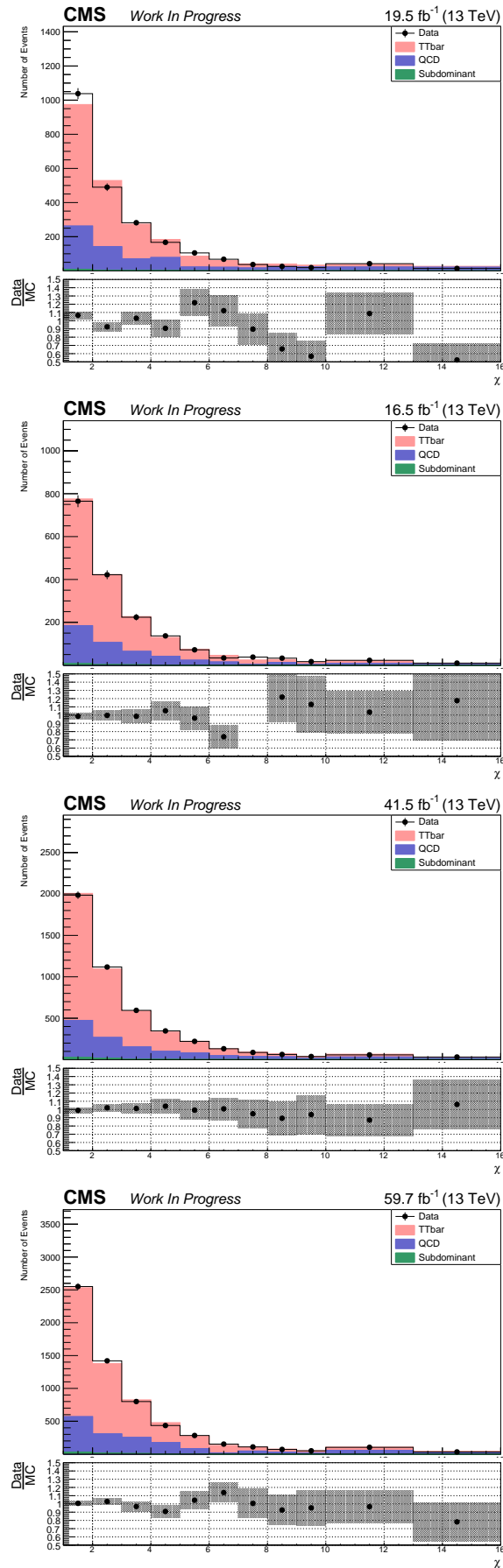


Figure 5.42: Angular distribution of the  $t\bar{t}$   $\chi$  in the signal region for all years. (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018).

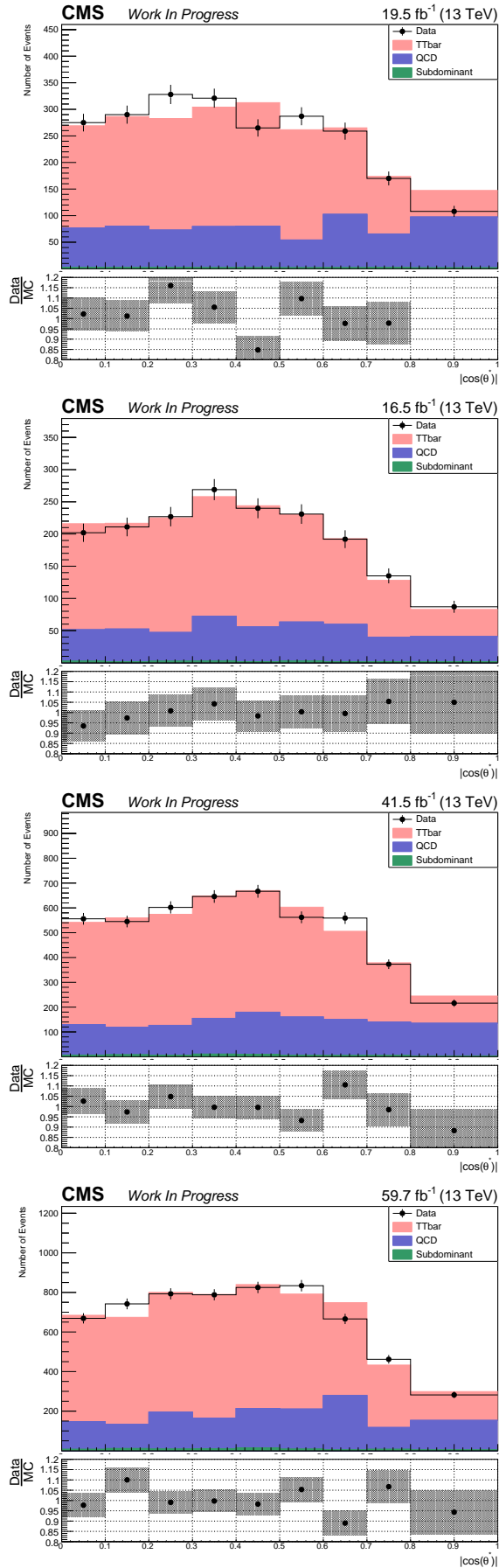


Figure 5.43: Angular distribution of the leading absolute  $|\cos(\theta)|$  in the signal region for all years (Top: (left)2016 preVFP, (right)2016 postVFP. Bottom: (left)2017, (right)2018).

## 5.7 Fiducial Measurement

The fiducial differential cross section is derived per bin  $i$  of the variable of interest  $x$  from the signal yield  $S_i$  as follows:

$$\frac{d\sigma_i^{\text{fid}}}{dx} = \frac{S_i}{\mathcal{L} \cdot \Delta x_i}, \quad (5.8)$$

where  $\mathcal{L}$  is the total integrated luminosity for the full Run II and  $\Delta x_i$  is the width of the  $i$ -th bin of the observable  $x$ . We also calculate and present the normalized differential cross section. This is calculated using the following equation:

$$\frac{1}{\sigma^{\text{fid}}} \frac{d\sigma_i^{\text{fid}}}{dx} = \frac{1}{\sum_k S_k} \cdot \frac{S_i}{\mathcal{L} \cdot \Delta x_i}, \quad (5.9)$$

The normalized differential cross section is used to confront the modelling of the differential cross section regardless of the overall normalization. In order to estimate the uncertainty of the measurement, the whole procedure is repeated for every source of uncertainty.

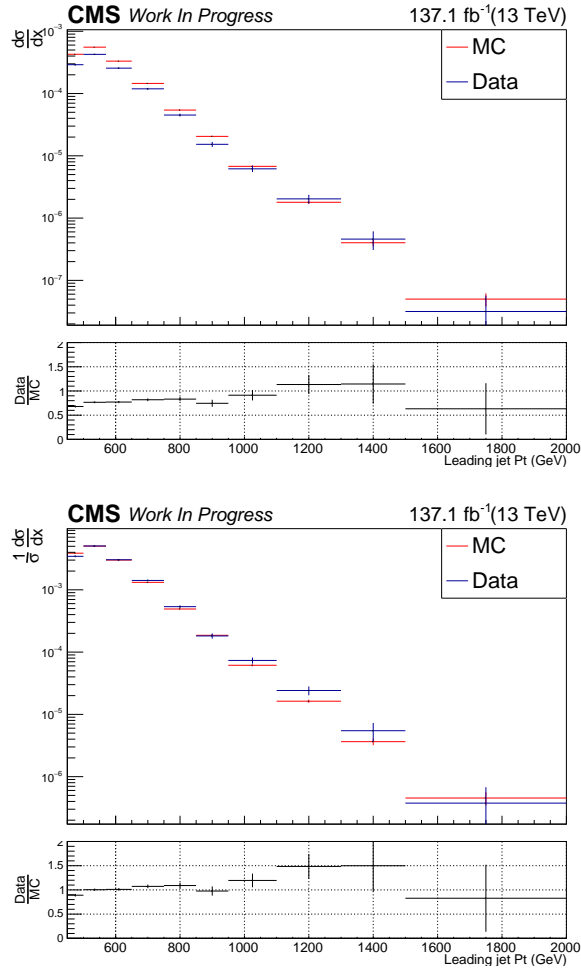


Figure 5.44: Fiducial differential cross section, absolute (left) and normalized (right), as a function of the leading top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

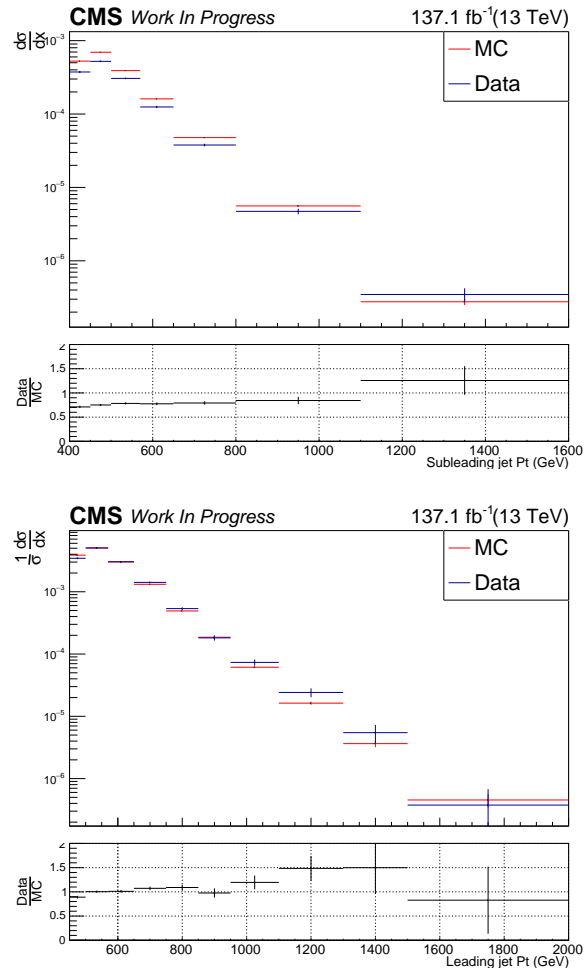


Figure 5.45: Fiducial differential cross section, absolute (left) and normalized (right), as a function of the second top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.



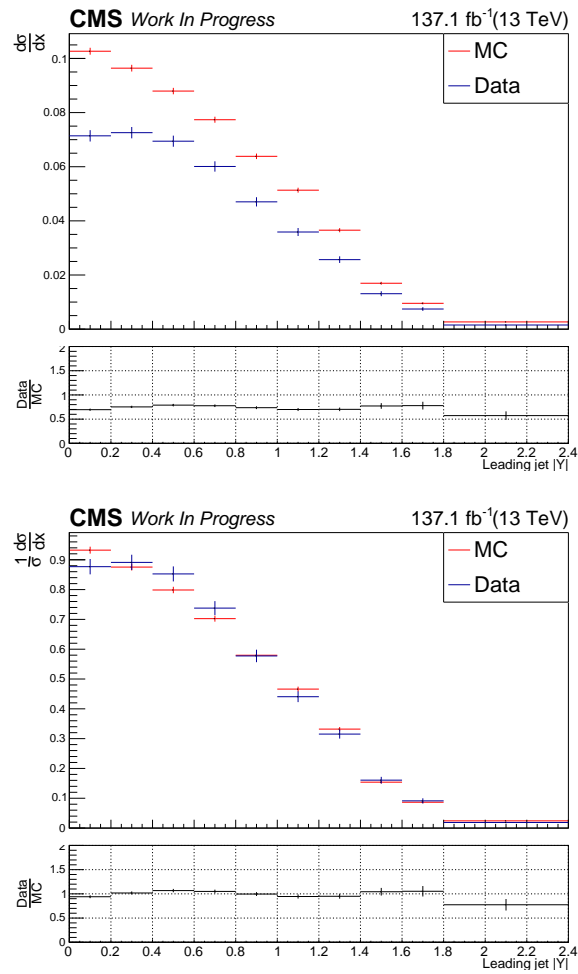


Figure 5.46: Fiducial differential cross section, absolute (left) and normalized (right), as a function of the leading top  $p_T$ . The bottom panel shows the ratio  $(\text{theory} - \text{data})/\text{data}$ .

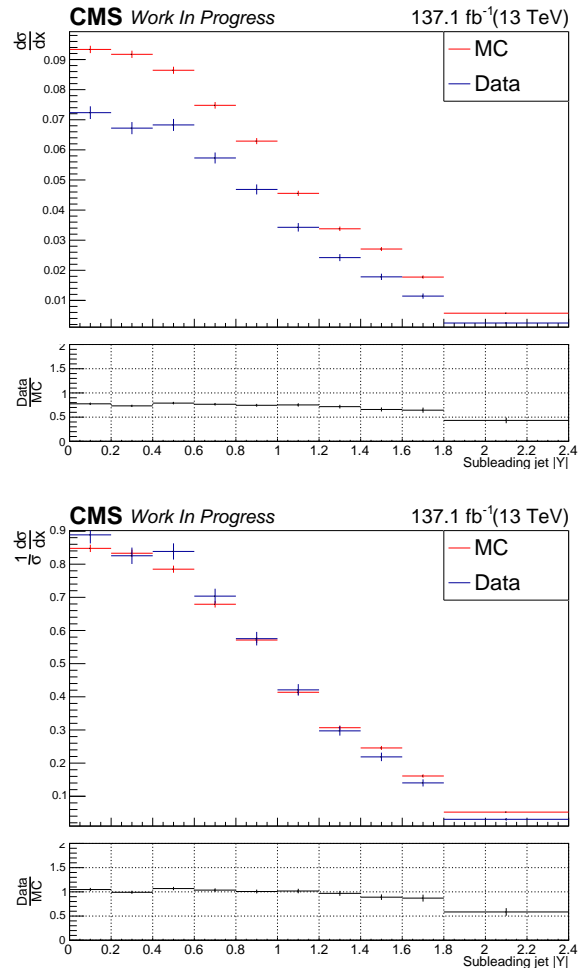


Figure 5.47: Fiducial differential cross section, absolute (left) and normalized (right), as a function of second top  $|y|$ . The bottom panel shows the ratio (theory - data)/data.

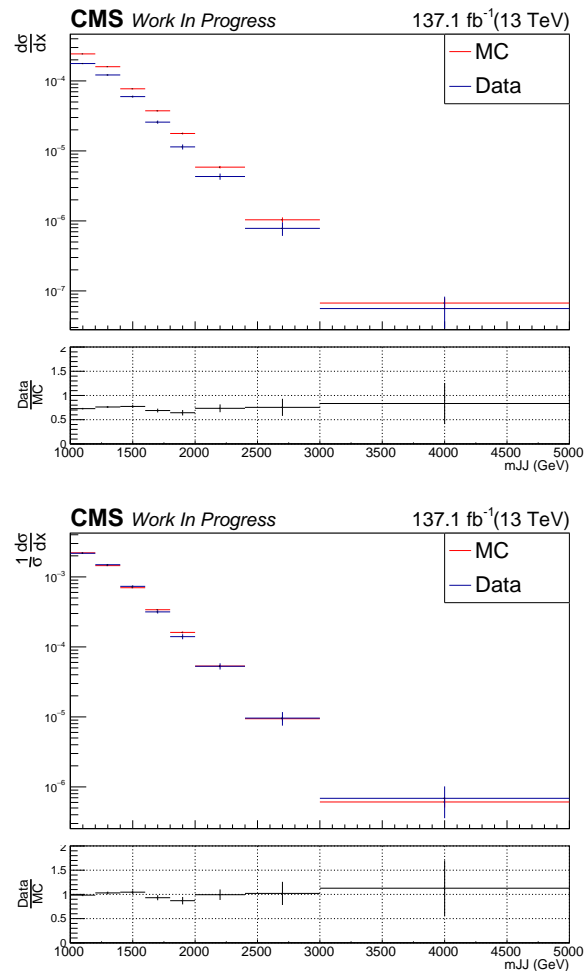


Figure 5.48: Fiducial differential cross section, absolute (left) and normalized (right), as a function of  $m_{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

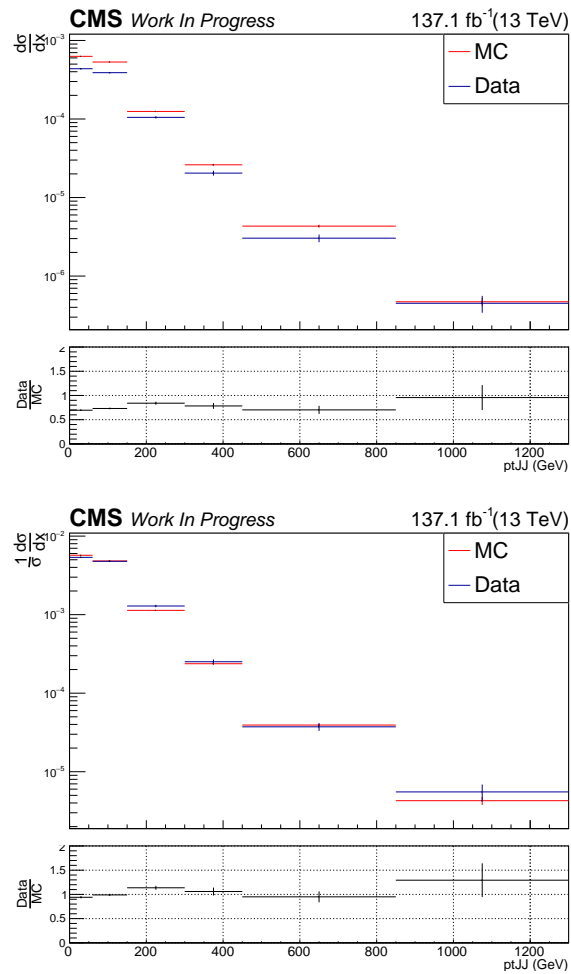


Figure 5.49: Fiducial differential cross section, absolute (left) and normalized (right), as a function of  $p_T^{tt}$ . The bottom panel shows the ratio (theory - data)/data.

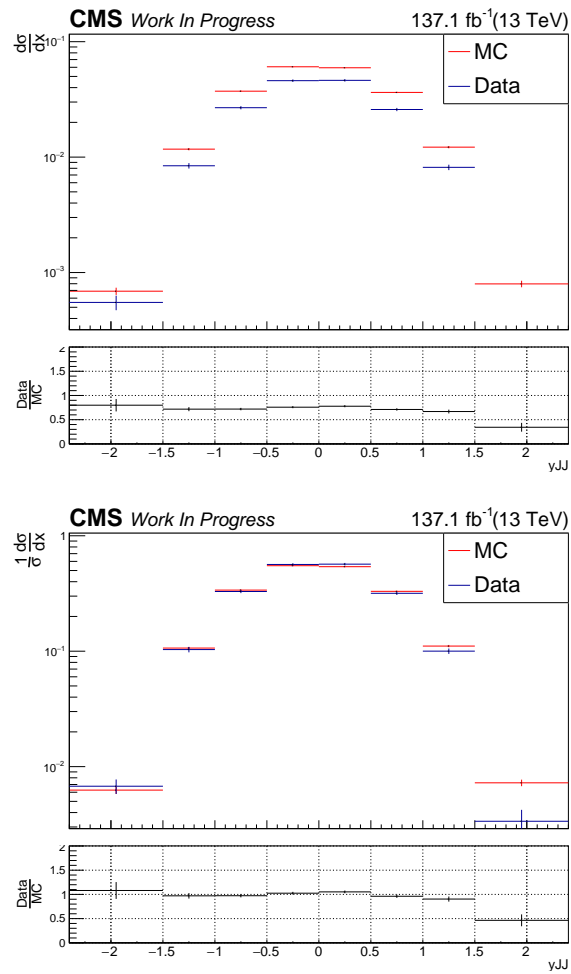


Figure 5.50: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top  $y^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

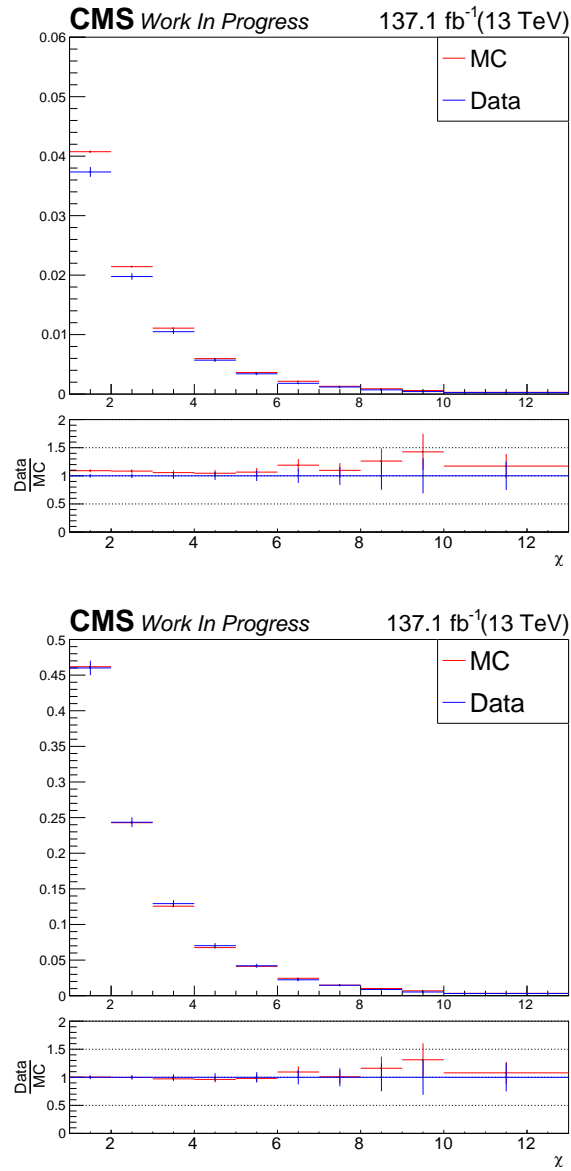


Figure 5.51: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top  $\chi$ . The bottom panel shows the ratio (theory - data)/data.

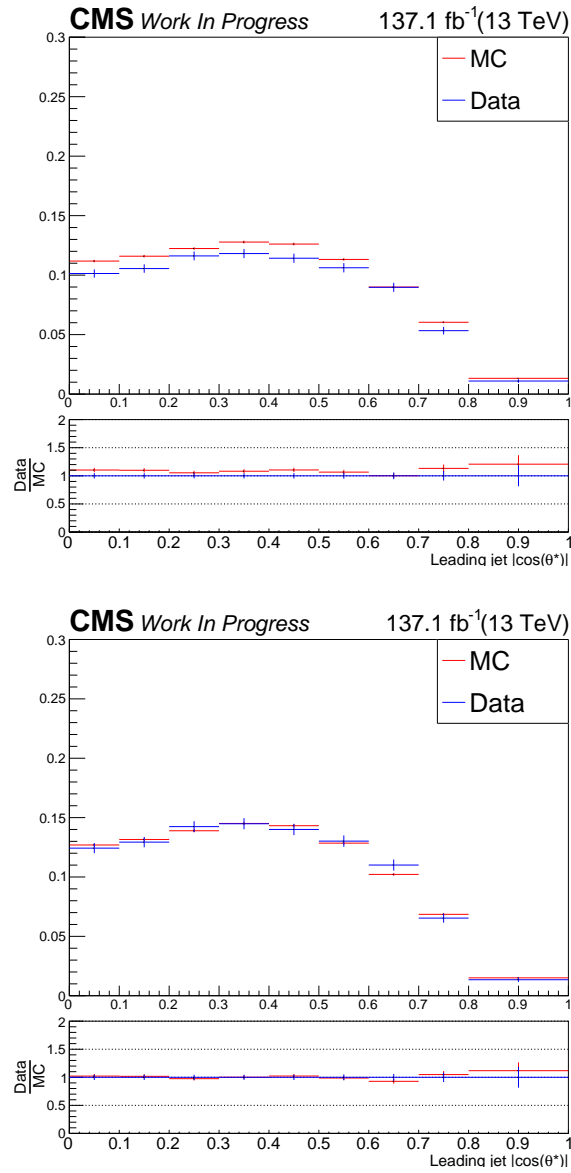


Figure 5.52: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top the leading  $|\cos(\theta^*)|$ . The bottom panel shows the ratio (theory - data)/data.

The systematic uncertainties considered in this analysis are divided in two categories: experimental and theoretical. The former includes all the uncertainties related to the differences in the object performance between data and simulation. The latter are related to the simulation itself and affect primarily the unfolded results through the acceptance, efficiency, and migration matrix. The list below describes briefly these uncertainties and the way they have been handled here. It should be noted that for each systematic variation the differential cross sections (fiducial and unfolded) are re-measured and the difference with respect to the nominal result is taken as the effect of this variation to the measurement. It is worth mentioning that the same procedure used for applying the systematic uncertainties to the calculated cross section values are also applied to the theoretical values. This is true only for the Theoretical uncertainties described below.

### 1. Experimental Uncertainties

- *QCD background prediction*: we use the fitted QCD yield uncertainty. The shape uncertainties due to the closure test in the simulation and the different pileup profiles in the control and signal regions are very small, at the 1% level.
- *Top Tagging Efficiency*: The uncertainty related to the identification of jets originating from top quarks candidates using the in-house developed top tagger. The mismatch in the performance of the tagger in data and simulation samples is corrected with the use of a scale factor. The uncertainty in the calculation of the scale factor is used to calculate the effect of the use of the scale factor in the cross section measurement. The scale factor as well as the uncertainty are calculated with the method described in detail in Appendix A. Using this method, an up and a down variation are calculated. Since the uncertainty is  $p_T$  dependent it leads to an effect of 5 – 25% based on the jets present in the event. This is a leading experimental uncertainty.
- *Jet energy scale (JES)*: this is the uncertainty on the energy scale of each reconstructed jet and it is a leading experimental uncertainty. Following the recommendations of the JME group, we have considered 30 independent JES sources [104] as follows: for each variation a new jet collection is created and the event interpretation is repeated. This results not only in variations of the  $p_T$  scale itself, but may also lead to different top candidates. The JES uncertainty, per jet, is of the order 1 – 2%,  $p_T$  and  $\eta$  dependent. The effect on the measured cross section is typically of the order 10% but it can be much larger at very high jet  $p_T$ .
- *Jet energy resolution*: The impact on the measurement due to the jet energy resolution (JER) is determined by smearing the jets according to the JER uncertainty. The effect on the cross section is relatively small, at the level of 2%.
- *Subjet b tagging efficiency (hadronic)*: The uncertainty in the identification of b-subjets within the large- $R$  jets (estimated in [105]) is a leading experimental in the hadronic channel. The effect on the cross sections is of the order of 7% relatively flat in all the observables. Unlike the uncertainty associated with the JES, the b-subjet tagging uncertainty therefore largely cancels in the normalized cross sections.
- *Pileup*: The uncertainty related to the modeling of additional pileup interactions is a sub-dominant uncertainty. The impact on the measurement is estimated by varying the total inelastic cross section used to weight the simulated events (69.2 mb) by  $\pm 4.6\%$  [106]. The effect on the cross sections is negligible (below 1%).
- *Trigger*: this accounts for the difference between the simulated and observed trigger efficiency. Based on Figure 5.2, 5.3, 5.4 the uncertainty is well below 1% in the phase space of this analysis.
- *Luminosity*: The uncertainty in the measurement per year is 1.2%, 2.3% and 2.5% for 2016, 2017 and 2018 respectively. The uncertainty in the measurement of the total integrated luminosity is 1.6% [107, 108, 109, 110].



## 2. Theoretical Uncertainties

The theoretical uncertainties are divided into two sub-categories: the ones related to the matrix element of the hard process and the ones related to the modelling of the parton shower and the underlying event. Practically, the first category (consisting of the first three sources below) is evaluated by variations of LHE event weights stored in the nominal MC simulation, while the second category is evaluated with dedicated, alternative MC samples.

- *Parton distribution functions*: The uncertainty due to parton distribution functions (PDFs) is estimated by applying event weights corresponding to the 100 replicas of the NNPDF set [94]. For each observable we compute its standard deviation from the 100 variations.
- *Renormalization and factorization scales*: The uncertainty in the choice of renormalization and factorization scale is assessed by separate and simultaneous variations of the renormalization and factorization scales. The unphysical anticorrelated variations are discarded, yielding a total of 6 combinations of the renormalization and factorization scales. The event weight is determined for each variation and the envelope is taken as the scale uncertainty, according to the prescription in [<https://indico.cern.ch/event/459797/contributions/1961581/attachments/1181Feb15-2016.pdf>]. These variations are applied using the event weights provided in the LHEEventInfo [<https://twiki.cern.ch/twiki/bin/viewauth/CMS/LHEReaderCMSW>].
- *Strong coupling constant ( $\alpha_S$ )*: The uncertainty associated with the  $\alpha_S$  is estimated by applying event weights corresponding to higher and lower values of  $\alpha_S$  for the matrix element using the variations of the NNPDF set [95].
- *Initial and Final state radiation (ISR, FSR)*: these uncertainties are used to evaluate the impact of the string coupling constant  $\alpha_S$  in the parton shower simulation. This is done by varying the  $\alpha_S^{ISR}, \alpha_S^{FSR}$  parameters by a factor of 2 or 0.5. The variations are applied using LHE weights provided in the samples.
- *Matrix element - parton shower matching*: In the POWHEG matrix element to parton shower (ME-PS) matching scheme, the resummation damping factor  $h_{\text{damp}}$  is used to regulate high- $p_T$  radiation. Uncertainties in  $h_{\text{damp}}$  are parameterized by considering alternative simulated samples with  $h_{\text{damp}}$  varied by  $h_{\text{damp}} = m_t$  and  $h_{\text{damp}} = 1.379^{+0.926}_{-0.5052} m_t$ .
- *Underlying event tune*: This uncertainty is estimated from alternative Monte Carlo samples with the tune CP5 parameters varied by  $\pm 1\sigma$ .

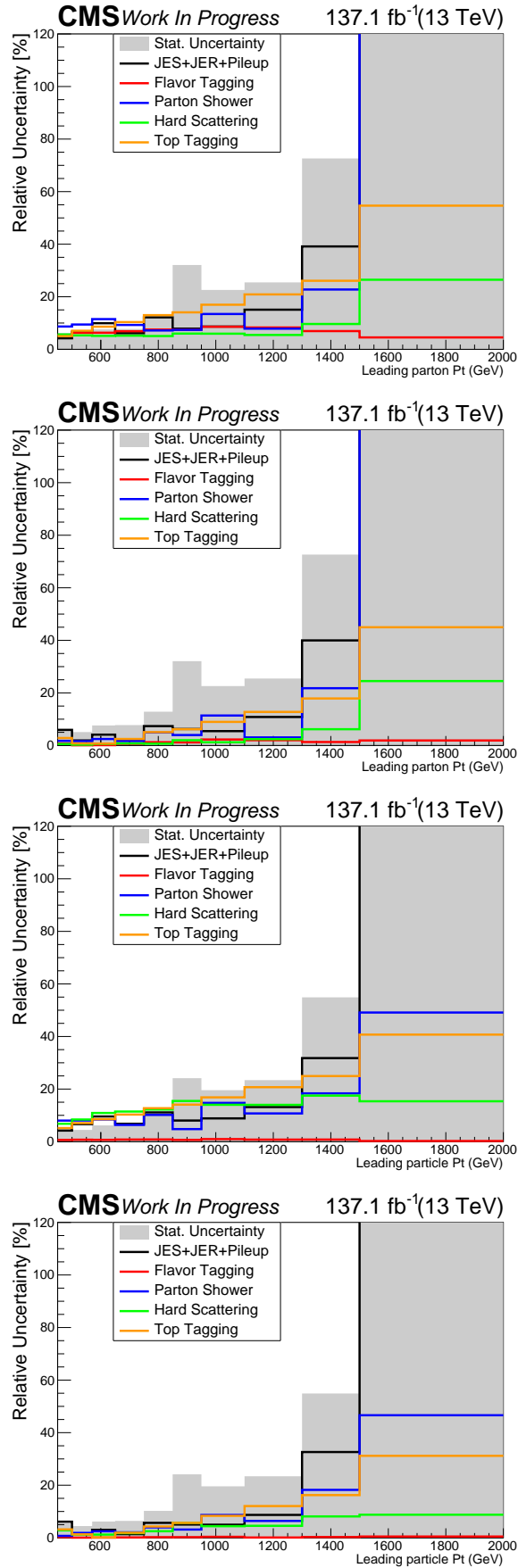


Figure 5.53: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the leading top  $p_T$ .

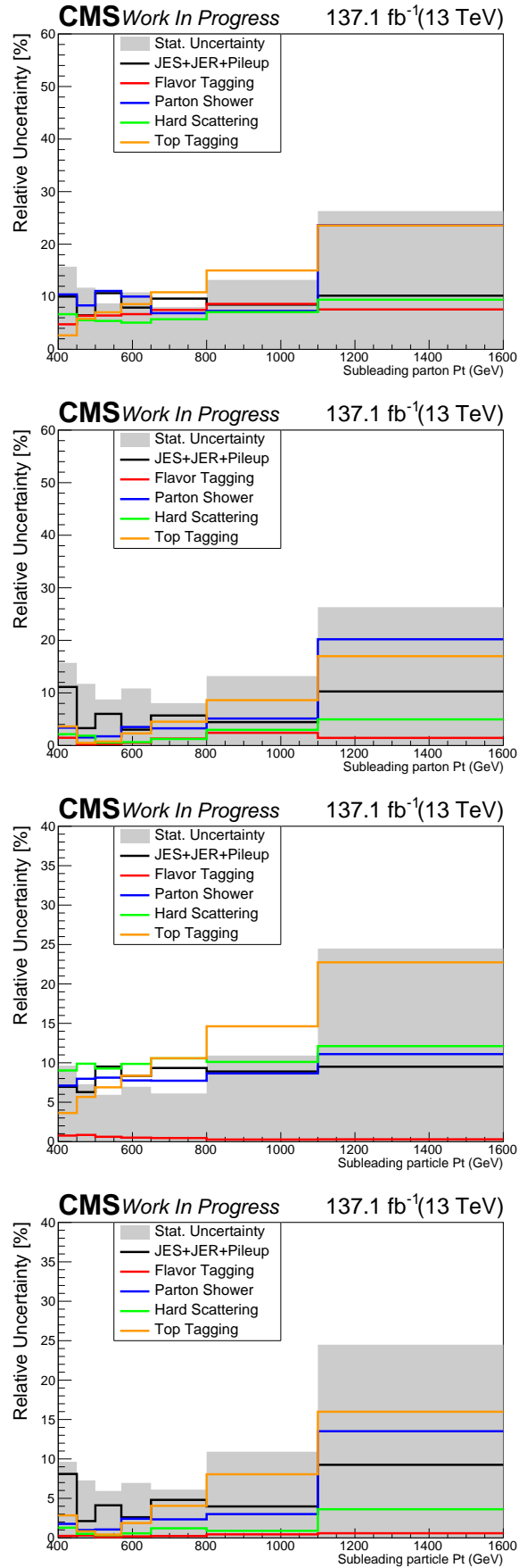


Figure 5.54: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the second leading top  $p_T$ .

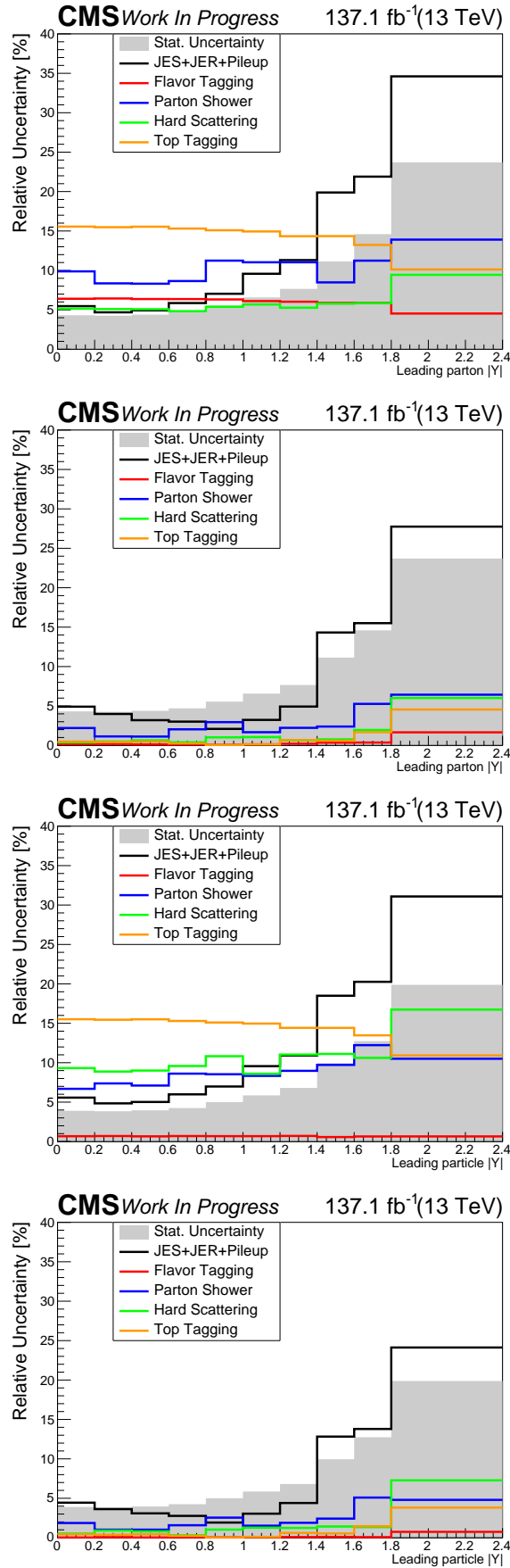


Figure 5.55: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the leading top absolute rapidity  $y$ .

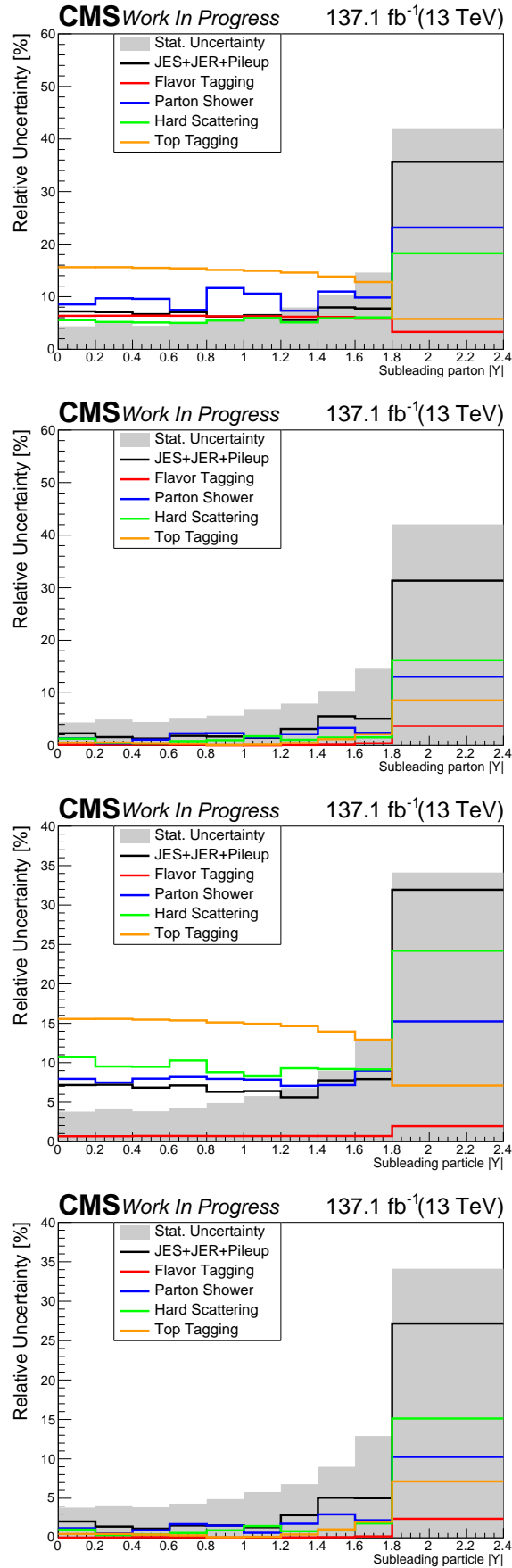


Figure 5.56: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the second leading top absolute rapidity  $y$ .

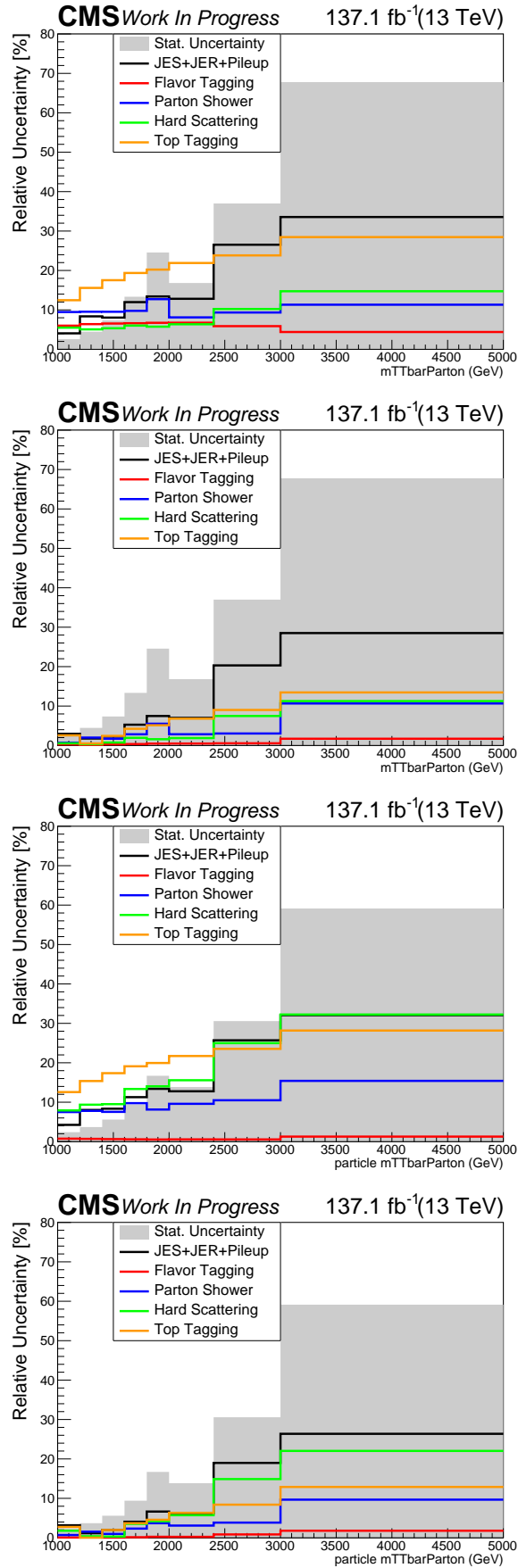


Figure 5.57: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $m_{t\bar{t}}$  dijet mass.

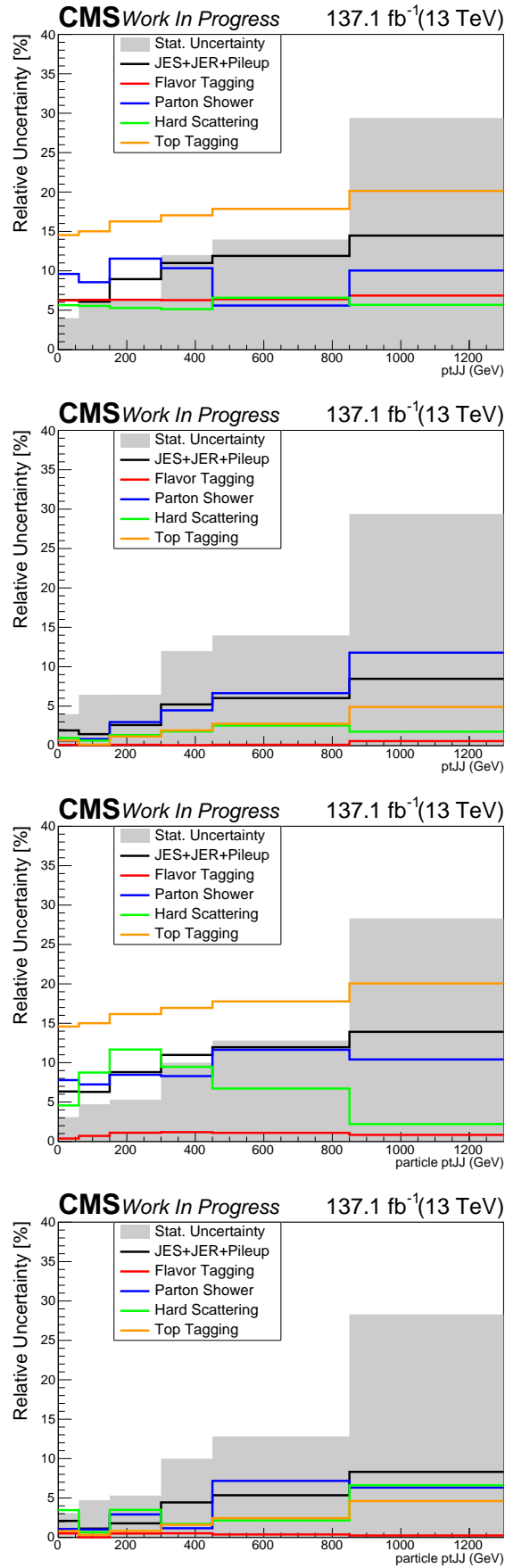


Figure 5.58: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $p_T^{t\bar{t}}$  dijet  $p_T$ .

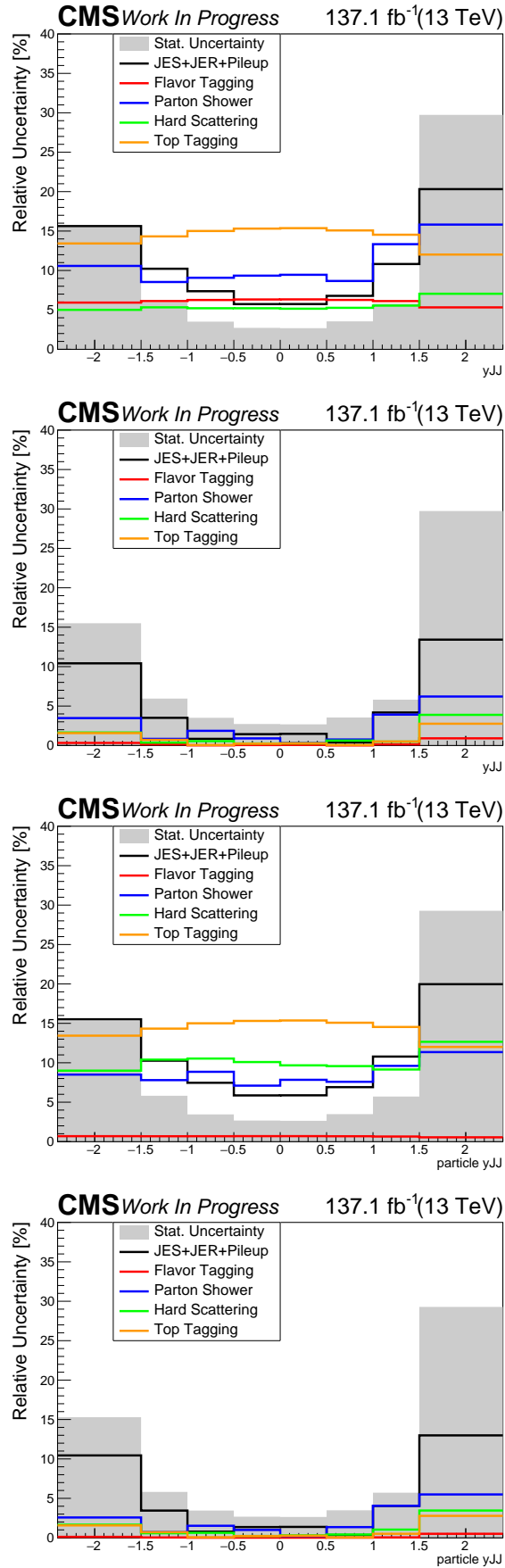


Figure 5.59: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $y_{t\bar{t}}$  dijet rapidity.



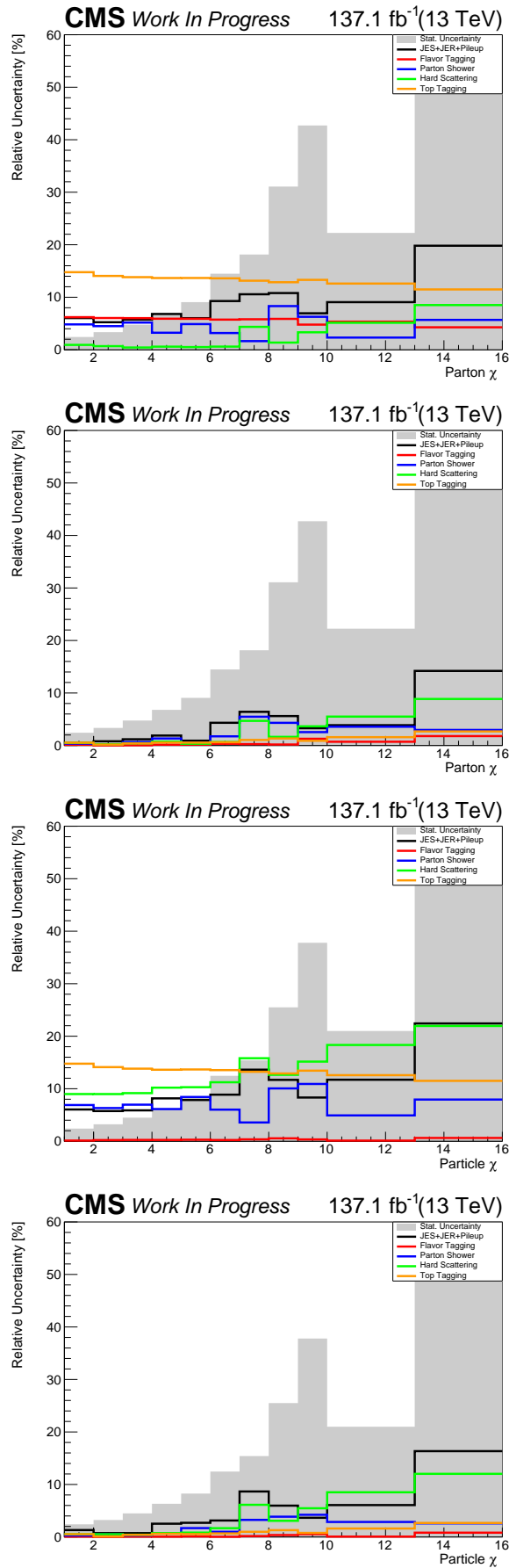


Figure 5.60: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $\chi_{t\bar{t}}$ .

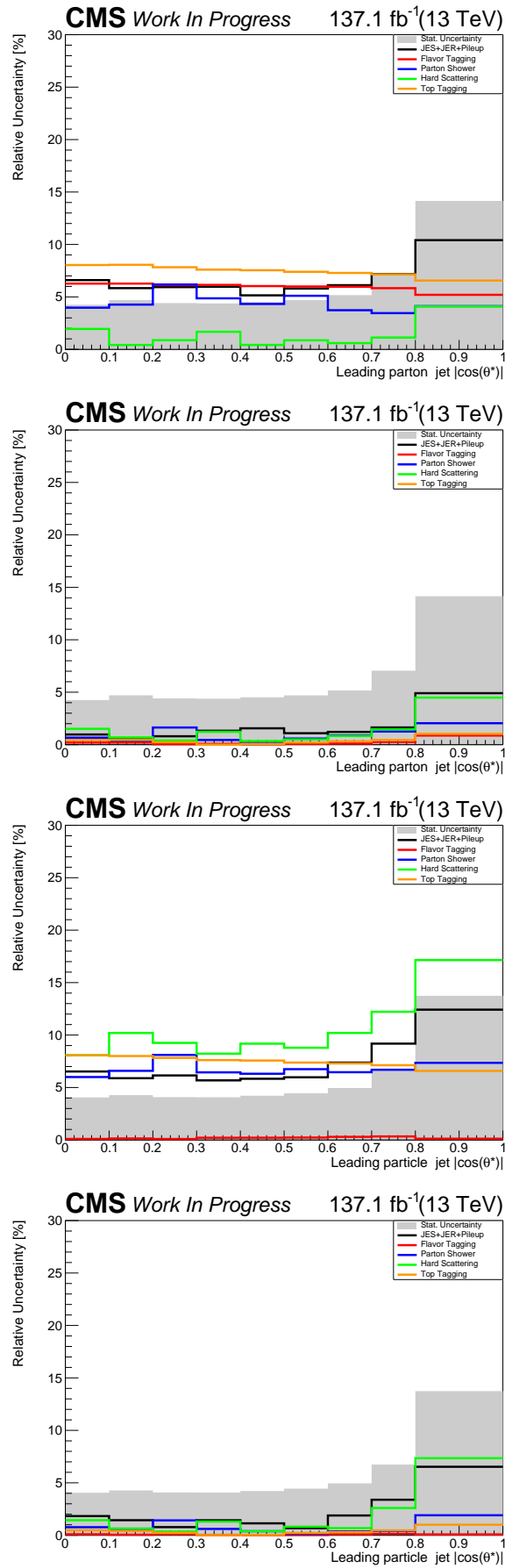


Figure 5.61: Decomposition of uncertainties: top part absolute, and normalized for parton and bottom part absolute and normalized for particle as a function of the  $\cos(\theta^*)$ .

## 5.8 Combination of different years

The goal of this analysis is to combine the measurements of each individual year to a single ("combined") measurement. The combination is done in the following way. For each year the signal extraction, yield, is computed separately. At the same time for each source of uncertainty the yield is also computed. At this point we have for each year one central (nominal) value for the yield and one value for the yield of each uncertainty source. We then combine the yearly values for each source. This way we end up with one nominal value which is the sum of the four individual central values and one value for each uncertainty which is also the sum of the four individual values for each uncertainty source. During the addition of the uncertainty sources the correlation between the yearly measurements are taken into account as shown in Table 5.22.

Table 5.22: Correlations for systematic uncertainties between the various years

Source	2016_pre-2016_post	2016_pre-2017	2016_pre-2018	2016_post-2017	2016_post-2018	2017-2018
pile - up	100%	100%	100%	100%	100%	100%
b-Tagging Unc.	See [111]					
Parton Shower Unc.	100%	100%	100%	100%	100%	100%
CP5 Tune Unc.	100%	100%	100%	100%	100%	100%
hDamp Samples Unc.	100%	100%	100%	100%	100%	100%
JES	See [112]					
Scale Unc.	100%	100%	100%	100%	100%	100%
PDF Unc.	100%	100%	100%	100%	100%	100%
Luminosity	30%	30%	30%	30%	30%	30%

To compute the differential cross section (Eq. 10.1) a combined response matrix needs to be computed and also a combined fraction  $f_1$  and  $f_2$ . For the response matrix the four individual response matrices are scaled to their respective luminosities and then added. During the unfolding procedure the matrix is normalized to unity so the luminosity of each year acts as a weight, where the years with more luminosity contribute more to the final response matrix. This is done for each source of uncertainty as well as the nominal value. For the fractions  $f_1$  and  $f_2$  we follow the same approach as for the signal yield. The numerators denominators for each year are added. From the common numerator and denominator the combined fraction is calculated. The procedure is repeated for each systematic uncertainty.

## 5.9 Unfolded Measurement

The fiducial measurement at detector level is unfolded and reported to the parton and particle phase spaces. The unfolded differential cross section is defined as:

$$\frac{d\sigma_i^{\text{unf}}}{dx} = \frac{1}{\mathcal{L} \cdot \Delta x_i} \cdot \frac{1}{f_{2,i}} \cdot \sum_j \left( R_{ij}^{-1} \cdot f_{1,j} \cdot S_j \right), \quad (5.10)$$

where  $\mathcal{L}$  is the total integrated luminosity of all years and  $\Delta x_i$  is the width of the  $i$ -th bin of the respective observable  $x$ . The quantity  $f_{1,j}$ , is the fraction of reconstructed events in the  $j$ -th bin that have an equivalent event at the unfolded level (parton or particle), whereas the quantity  $f_{2,i}$ , is defined as the fraction of events at the unfolded level that have an equivalent reconstructed event. Figures 5.62 and 5.63 show the previously defined fractions. The quantity  $R_{ij}^{-1}$  is the inverse of the migration matrix between the  $i$ -th and  $j$ -th bins of the unfolded (parton or particle) and reconstructed level respectively. The migration matrix is non-diagonal, due to the finite resolution of the detector as depicted in Figures 5.64 and 5.65. Thus, an unfolding procedure is necessary. The binning of the various observables has been chosen in such way so that the purity (fraction of reconstructed events that the true value of the observable lies in the same bin) and stability (fraction of true events that the reconstructed observable lies in the same bin) are well above 50%. The choice of this binning results in highly diagonal migration matrices, as shown in Figs. 5.64 and 5.65. In order to avoid any biases introduced by the various unfolding techniques with some type of regularization, the simple migration matrix inversion method has been used, as written in Eq. 10.1. The price we pay is that we have a moderate increase of the statistical uncertainty.

The results also include comparison with the theoretical models. Two theoretical models are used, POWHEG with PYTHIA 8 as well as aMC@NLO[FxFx] with PYTHIA 8. The comparison is presented in the form of  $\chi^2$  values calculated using the covariance matrix. Using the  $\chi^2$  values and the number of degrees of freedom (NDOF) p-values are also calculated. It is also worth mentioning here that systematic uncertainties are applied in the calculation of the theoretical values for the absolute cross section. The uncertainties are also taken into account in the calculation of the  $\chi^2$  values as they are incorporated in the covariance matrix. The  $\chi^2$  values are calculated using

$$\chi^2 = V_{N_b}^T \cdot C_{N_b}^{-1} \cdot V_{N_b} \quad (5.11)$$

where  $V_{N_b}$  is the vector of differences between the measured values and the corresponding theoretical predictions and  $V_{N_b}^T$  its transpose.  $N_b$  is the length of the vector and corresponds to the number of bins for each measurement.  $C_{N_b}^{-1}$  is the inverse of the covariance matrix. The covariance matrix is calculated using the following formula

$$C = C_{stat} + C_{syst} \quad (5.12)$$

where  $C_{stat}$  is the covariance matrix for the statistical uncertainties and  $C_{syst}$  is the covariance matrix for the systematic. The covariance matrix for the systematic uncertainties is calculated as

$$C_{ij}^{syst} = \sum_{k,l} \frac{1}{N_k} C_{j,k,l} C_{i,k,l}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N, \quad (5.13)$$

where  $C_{i,k,l}$  is the systematic uncertainty in the  $i$ -th bin of the  $l$ -th variation of the  $k$ -th source and  $N_k$  is the number of variations for the source  $k$ . The sum runs over all the sources of systematic uncertainties and their variations.

### 5.9.1 Parton Level

The results of the unfolded measurement at parton level are shown in Figs. 5.68- 5.76. The comparison with the theory predictions follows the trends observed in the fiducial measurement. That is, there is an 20 – 40% lower inclusive cross section, while the shapes of the differential distributions are reasonably reproduced by all models.

Table 5.23:  $\chi^2$  Values of the Absolute differential cross sections at Parton level.

Variable	NDOF	$\chi^2$ Theory	$\chi^2$ Theory AMC@NLO	$\chi^2$ Theory Powheg+Herwig	P-value Theory	P-value AMC@NLO	P-value Powheg+Herwig
Leading Parton $p_T$	6	8.2	16	5.3	0.22	0.016	0.51
Second Leading Parton $p_T$	7	5.7	8.6	3.8	0.57	0.28	0.8
Parton $y_{t\bar{t}}$	8	8.9	6.8	9	0.35	0.56	0.35
Parton $p_T^{t\bar{t}}$	6	3.2	7.5	3.1	0.78	0.27	0.8
Parton $m_{t\bar{t}}$	8	4.5	6.9	3.5	0.8	0.55	0.9
Leading Parton $ Y $	10	4.7	7	3.7	0.91	0.73	0.96
Second Leading Parton $ Y $	10	12	13	11	0.26	0.21	0.35
Parton $\chi$	11	1.4	9.7	4.2	1.	0.56	0.96
Parton $ \cos(\theta^*) $	9	1.2	4.2	2.3	1.	0.9	0.99

Table 5.24:  $\chi^2$  Values of Normalized differential cross sections at Parton level.

Variable	NDOF	$\chi^2$ Theory	$\chi^2$ Theory AMC@NLO	$\chi^2$ Theory Powheg+Herwig	P-value Theory	P-value AMC@NLO	P-value Powheg+Herwig
Leading Parton $p_T$	5	2.8	2.2	2.7	0.73	0.82	0.74
Second Leading Parton $p_T$	6	1	1.1	0.96	0.98	0.98	0.99
Parton $\theta_{t\bar{t}}$	7	12	12	12	0.1	0.09	0.094
Parton $p_T^{t\bar{t}}$	5	4.8	1.3	5.1	0.44	0.93	0.4
Parton $m_{t\bar{t}}$	7	1.1	1.1	0.86	0.99	0.99	1
Leading Parton $ Y $	9	1.9	1.7	2	0.99	1	0.99
Second Leading Parton $ Y $	9	12	12	12	0.23	0.22	0.24
Parton $\chi$	10	2.3	1.8	2.	0.99	1.	1.
Parton $ \cos(\theta^*) $	8	1.3	0.98	1.1	1	1.	1.

### 5.9.2 Particle Level

The results of the unfolded measurement at particle level are shown in Figs. 5.77- 5.85. The comparison with the theory predictions follows the trends observed in the fiducial measurement. That is, there is an 20 – 40% lower inclusive cross section, while the shapes of the differential distributions are reasonably reproduced by all models.

Table 5.25:  $\chi^2$  Values of the Absolute differential cross sections at Particle level.

Variable	NDOF	$\chi^2$ Theory	$\chi^2$ Theory AMC@NLO	$\chi^2$ Theory Powheg+Herwig	P-value Theory	P-value AMC@NLO	P-value Powheg+Herwig
Leading Particle $p_T$	6	4.7	8.4	6.1	0.59	0.21	0.41
Second Leading Particle $p_T$	7	4.7	4.9	5.5	0.7	0.67	0.59
Particle $y_{t\bar{t}}$	8	11	6.2	19	0.18	0.62	0.017
Particle $p_T^{t\bar{t}}$	6	3	6.8	4.3	0.81	0.34	0.63
Particle $m_{t\bar{t}}$	8	4.5	5.4	4.8	0.81	0.71	0.78
Leading Particle $ Y $	10	5.3	6	6.8	0.87	0.81	0.74
Second Leading Particle $ Y $	10	14	10	17	0.19	0.41	0.079
Particle $\chi$	11	4.31	5.1	5.93	0.96	0.93	0.88
Particle $ \cos(\theta^*) $	9	3.66	2.55	4.28	0.93	0.98	0.89



Table 5.26:  $\chi^2$  Values of Normalized differential cross sections at Particle level.

Variable	NDOF	$\chi^2$ Theory	$\chi^2$ Theory AMC@NLO	$\chi^2$ Theory Powhcg+Herwig	P-value Theory	P-value AMC@NLO	P-value Powhcg+Herwig
Leading Particle $p_T$	5	1.8	1.9	1.8	0.88	0.86	0.88
Second Leading Particle $p_T$	6	0.73	1.2	0.68	0.99	0.98	0.99
Particle $y_{t\bar{t}}$	77	13	11	12	0.071	0.12	0.098
Particle $p_T^{t\bar{t}}$	5	3.4	4.2	4.1	0.64	0.52	0.54
Particle $m_{t\bar{t}}$	7	1.8	2.5	1.9	0.97	0.92	0.97
Leading Particle $ Y $	9	1.9	2.6	2.7	0.99	0.98	0.98
Second Leading Particle $ Y $	9	13	11	12	0.16	0.26	0.23
Particle $\chi$	10	1.65	1.29	2.6	0.99	0.99	0.97
Particle $ \cos(\theta^*) $	8	1.6	1.3	2.6	0.99	0.99	0.98

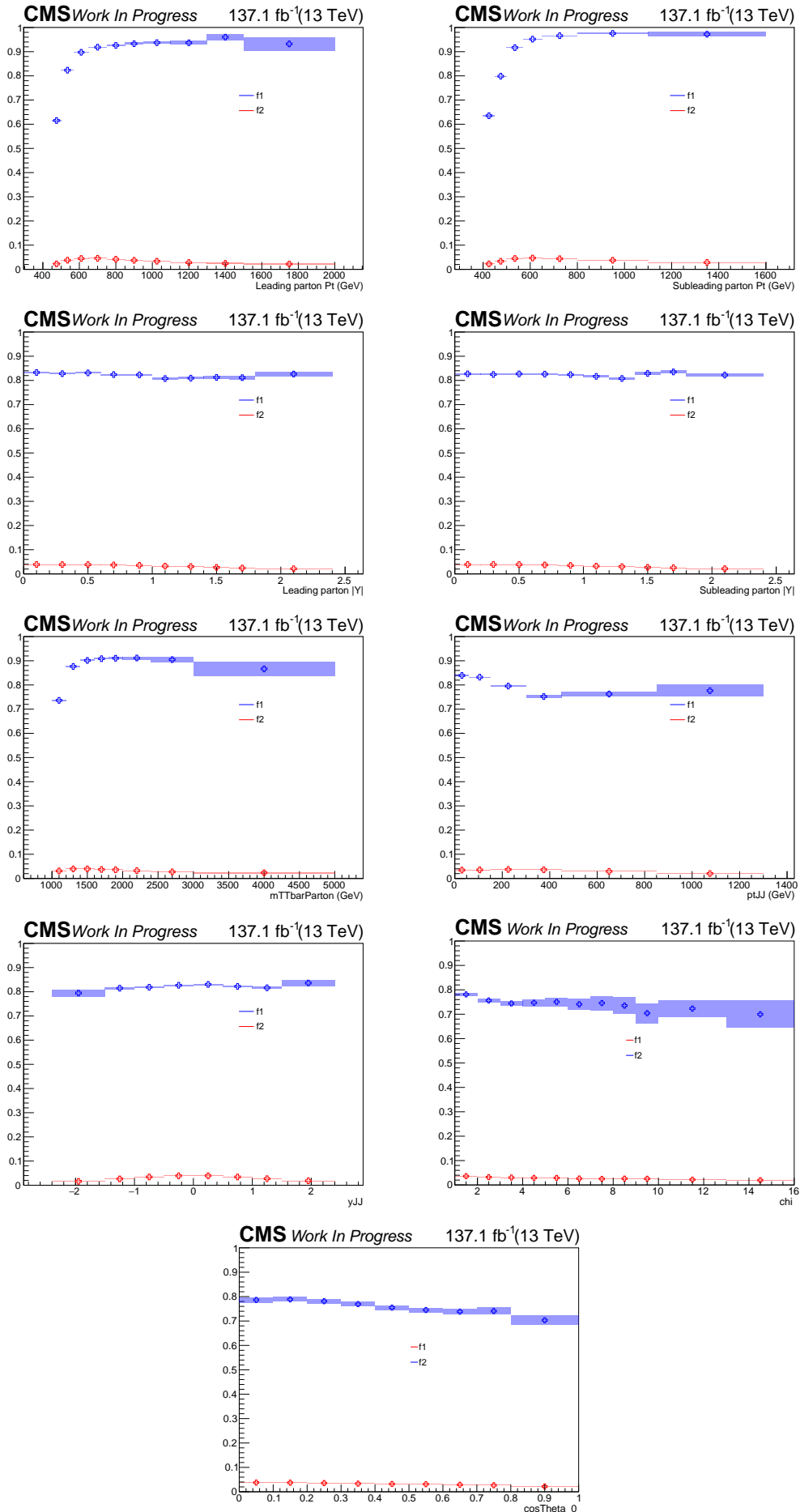


Figure 5.62: Acceptance and efficiency for the parton-level selection as a function of the various observables.

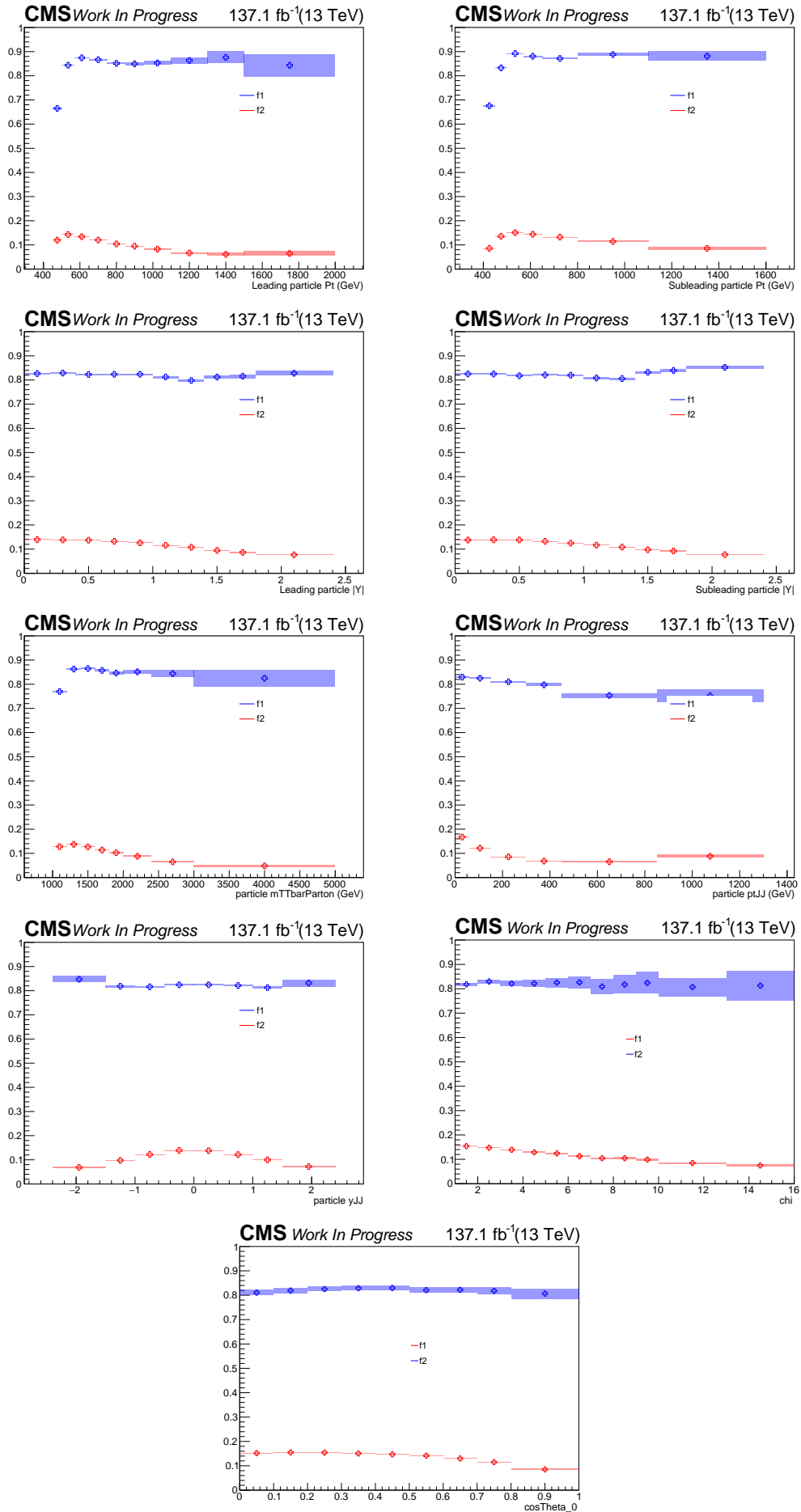


Figure 5.63: Acceptance and efficiency for the particle-level selection as a function of the various observables.

# CHAPTER 5. BOOSTED $t\bar{t}$ DIFFERENTIAL MEASUREMENTS

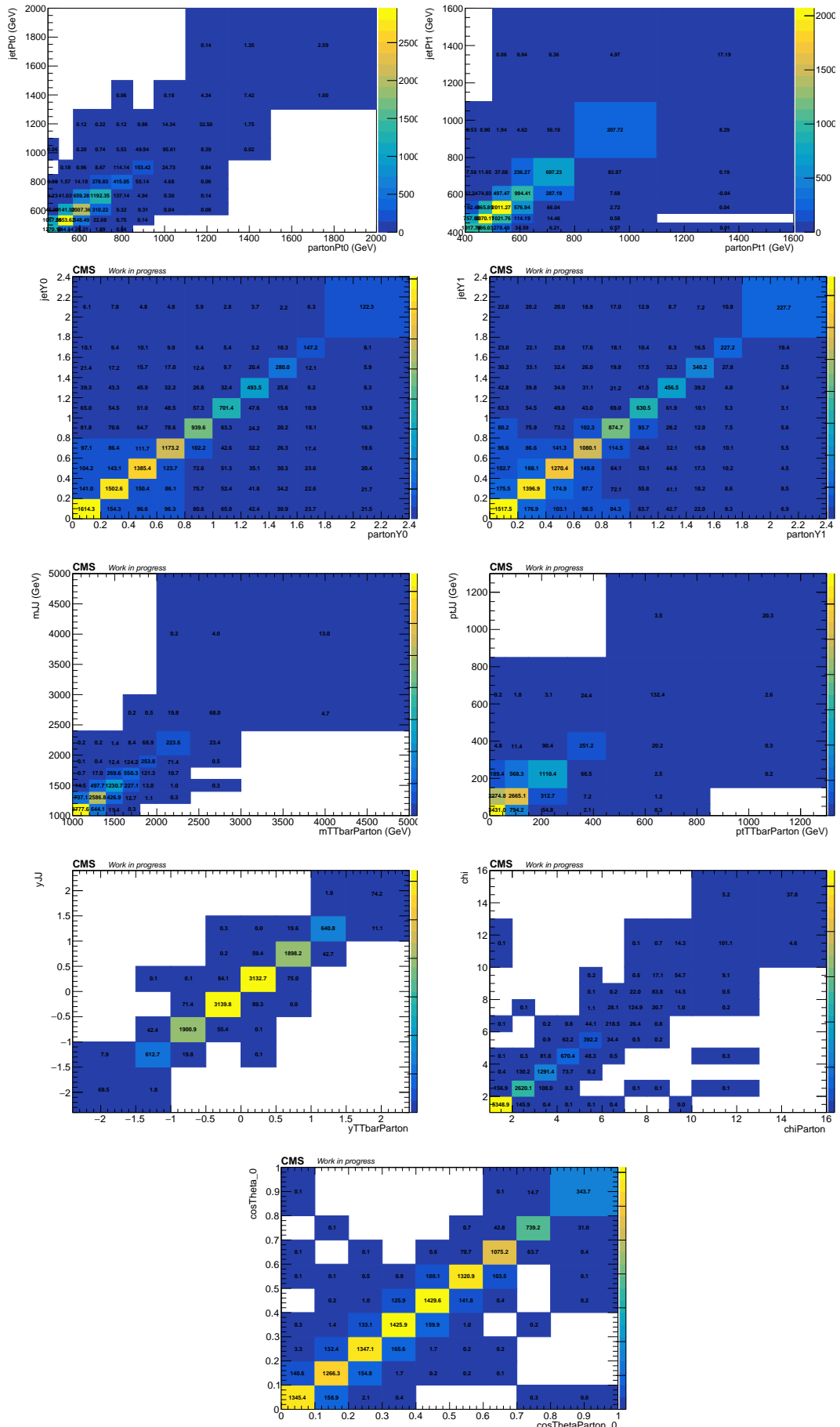


Figure 5.64: Simulated migration matrices at parton level.

# CHAPTER 5. BOOSTED $t\bar{t}$ DIFFERENTIAL MEASUREMENTS

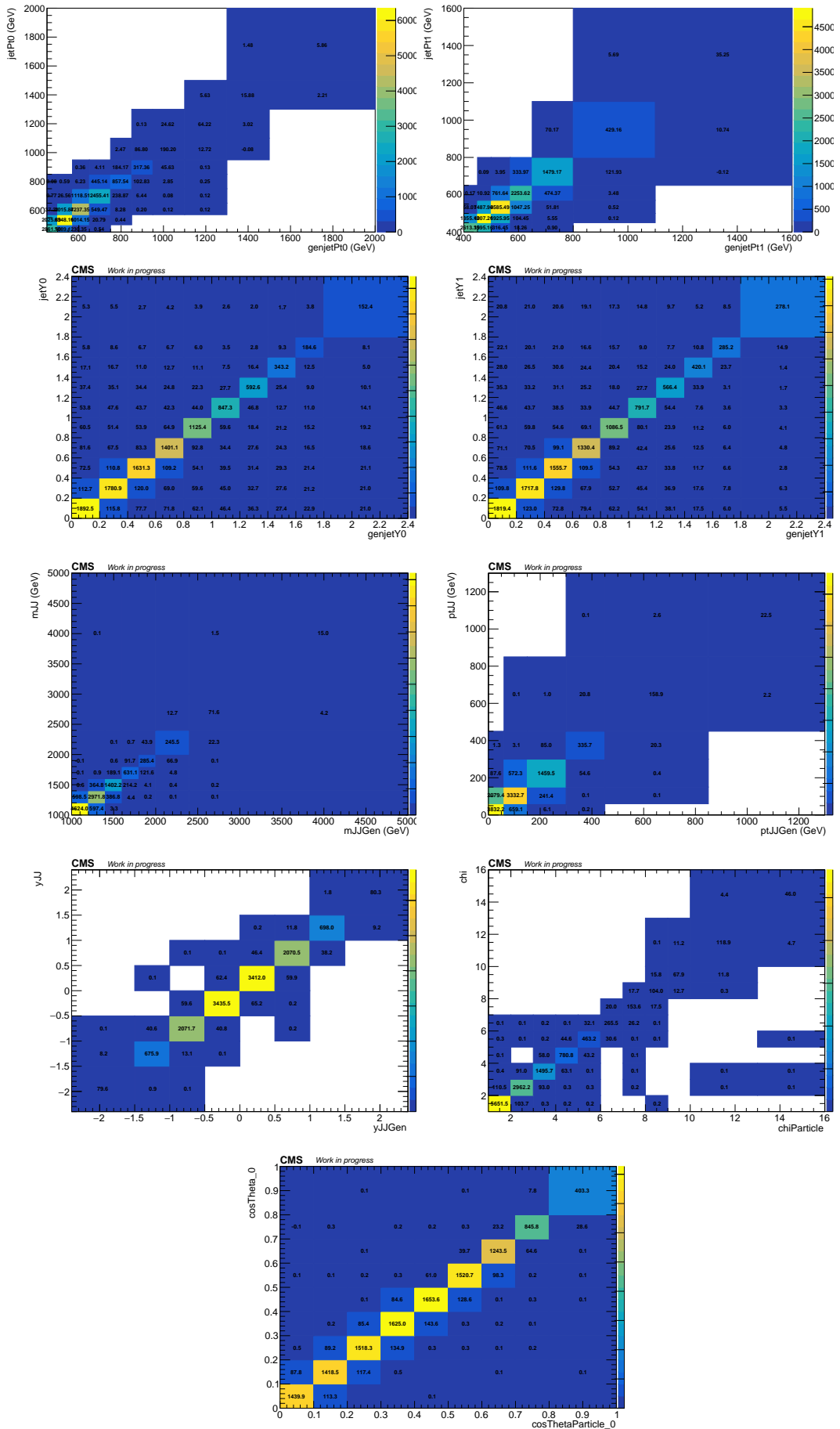


Figure 5.65: Simulated migration matrices at particle level.

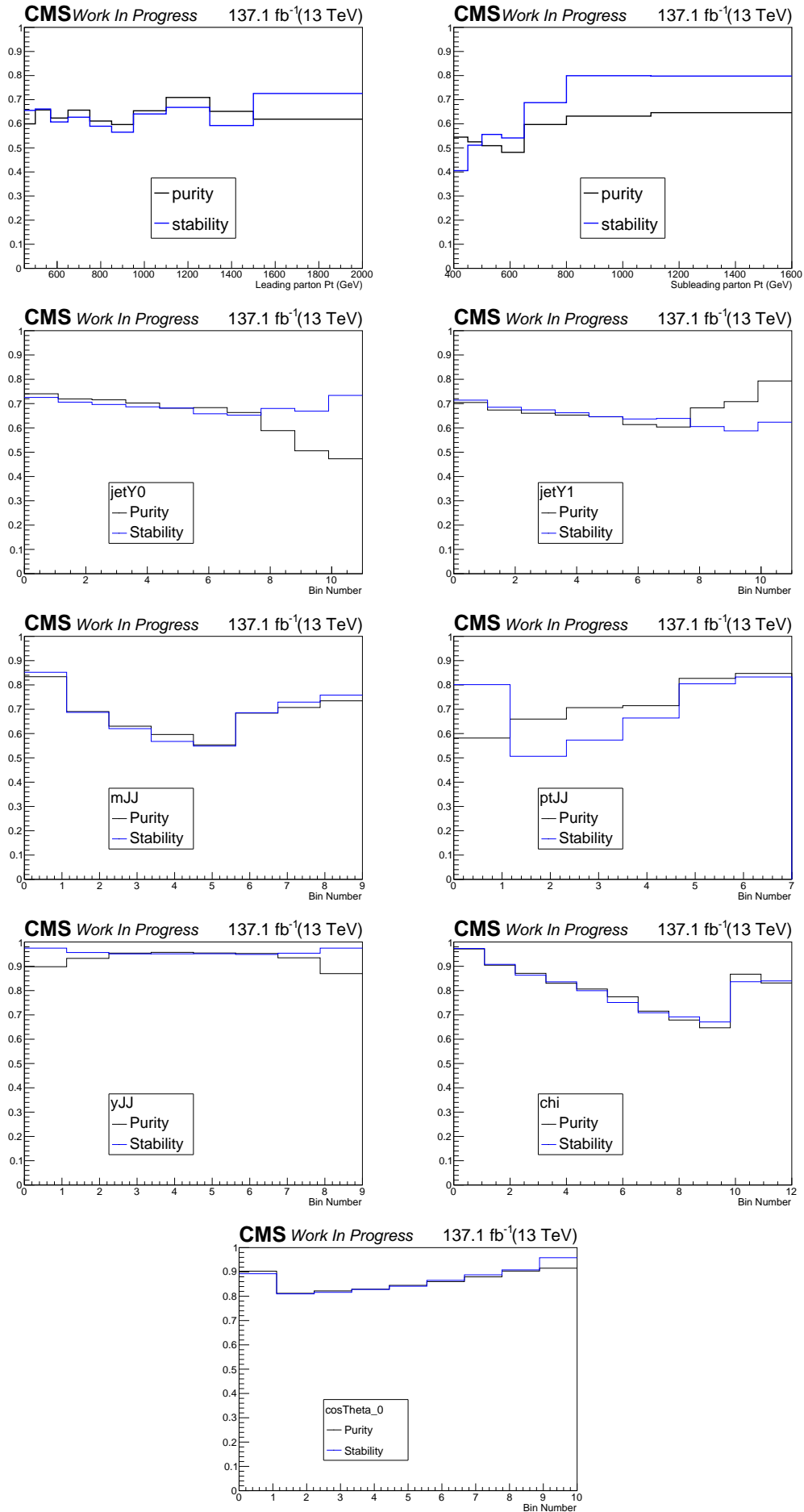


Figure 5.66: Purity and Stability for the parton-level selection as a function of the various observables.

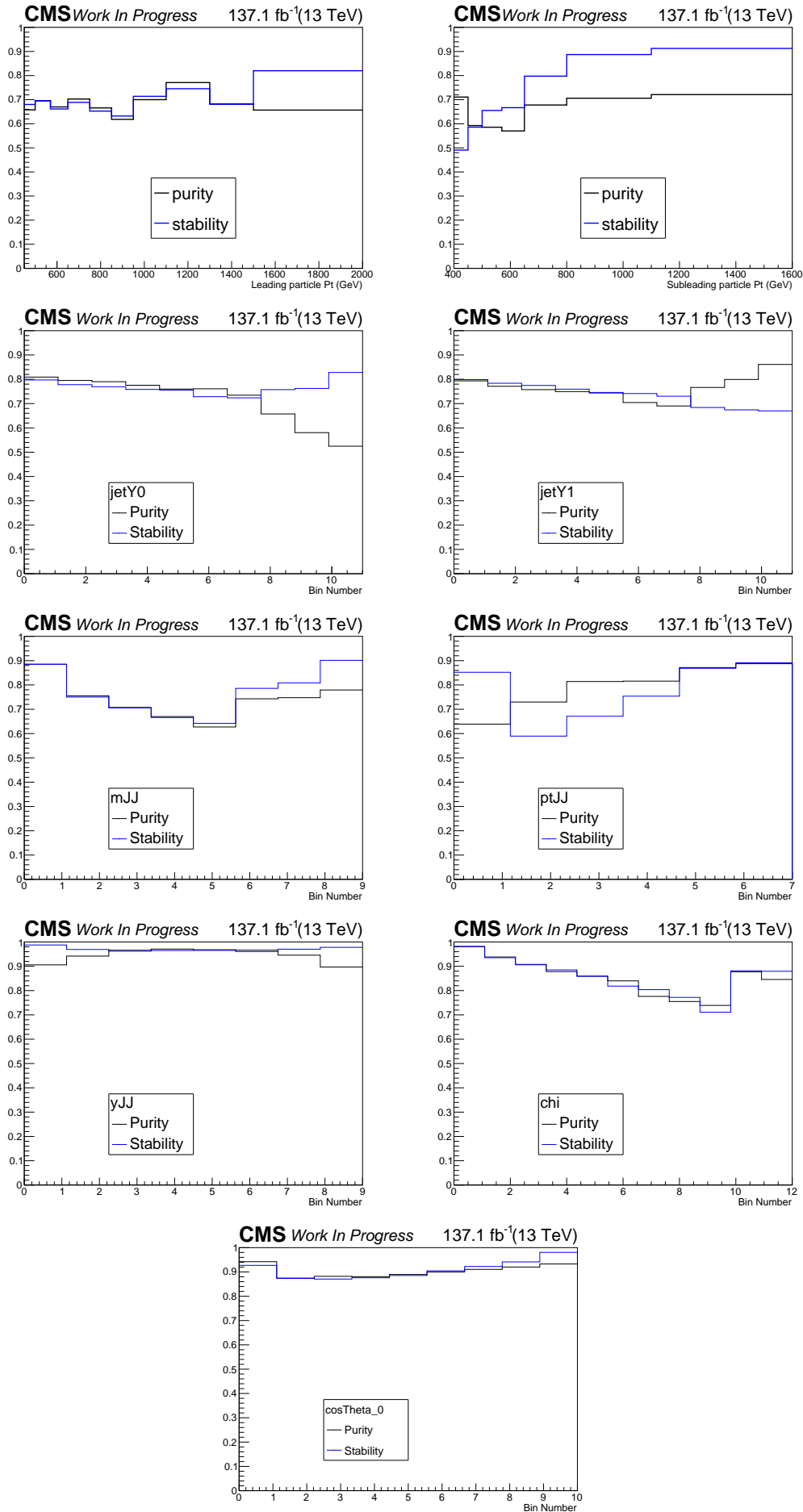


Figure 5.67: Purity and Stability for the particle-level selection as a function of the various observables.

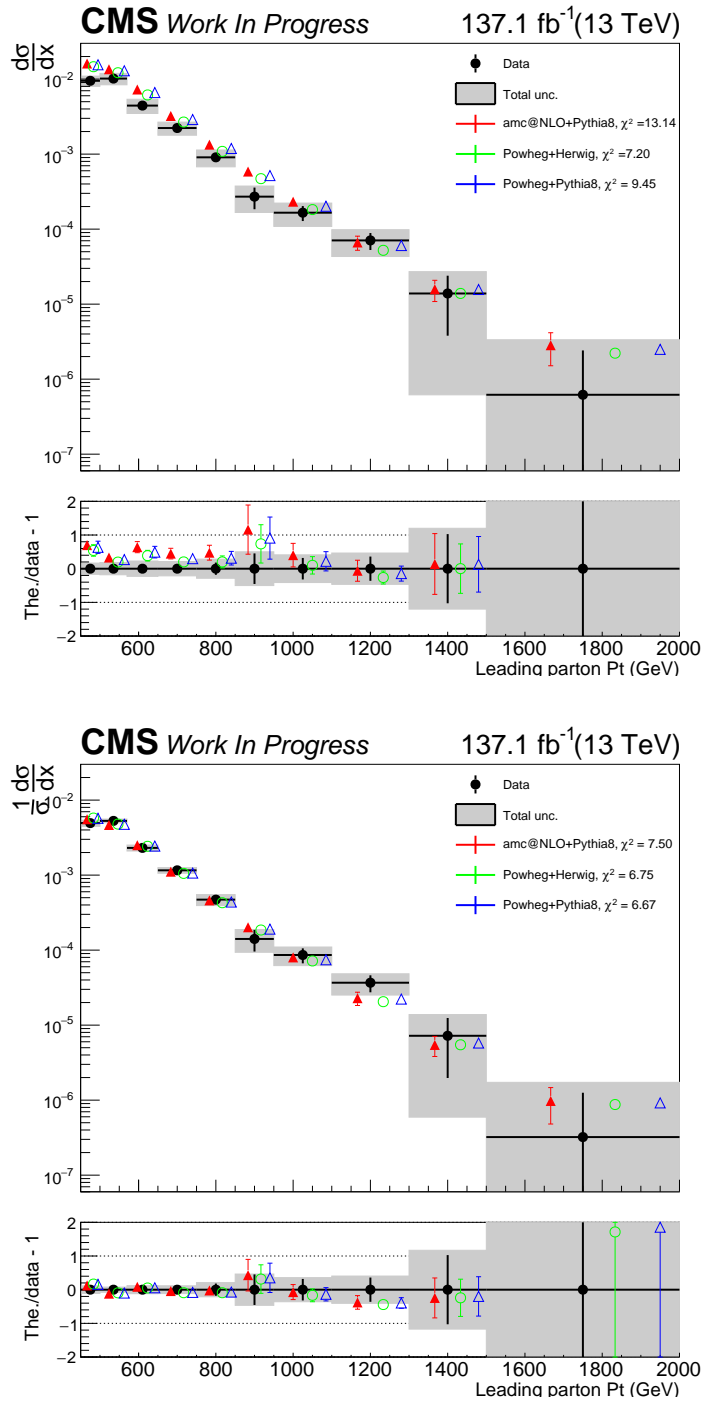


Figure 5.68: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of leading jet  $p_T$ . The bottom panel shows the ratio (theory - data)/data.



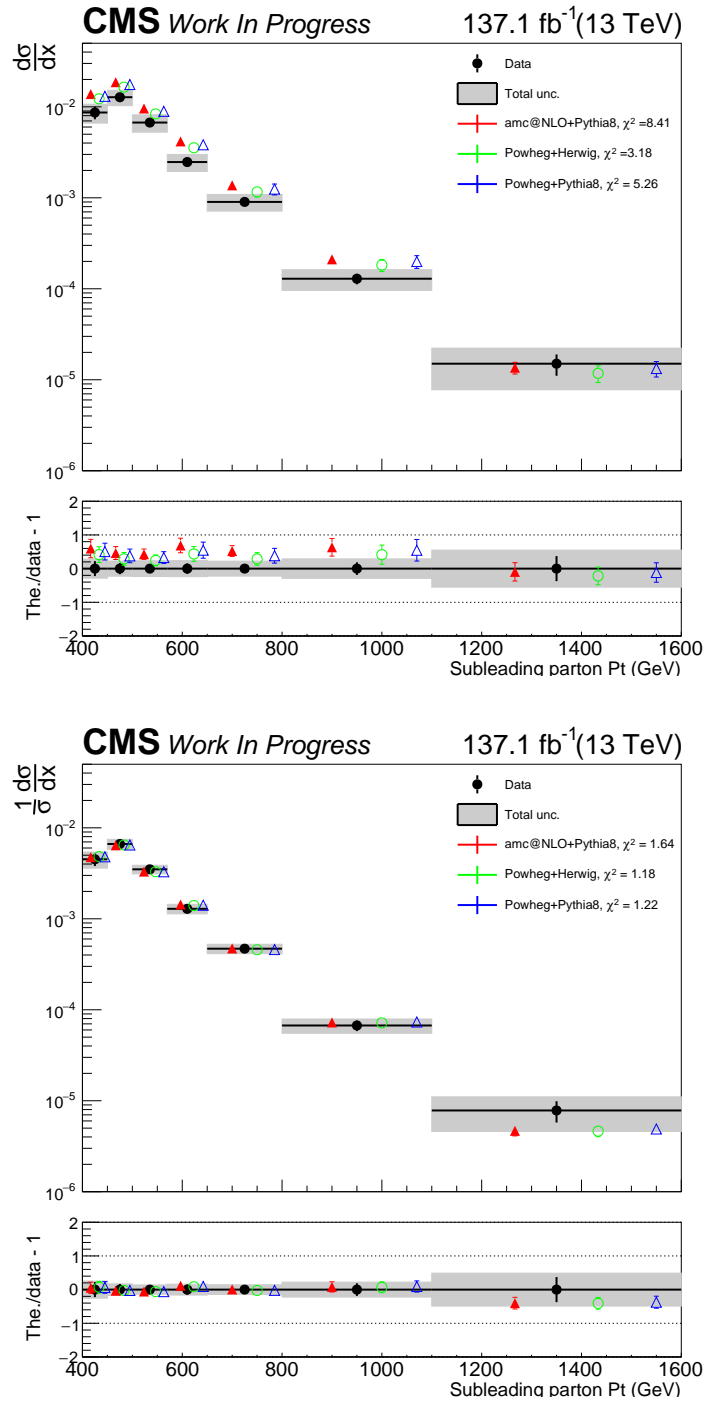


Figure 5.69: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of second leading jet  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

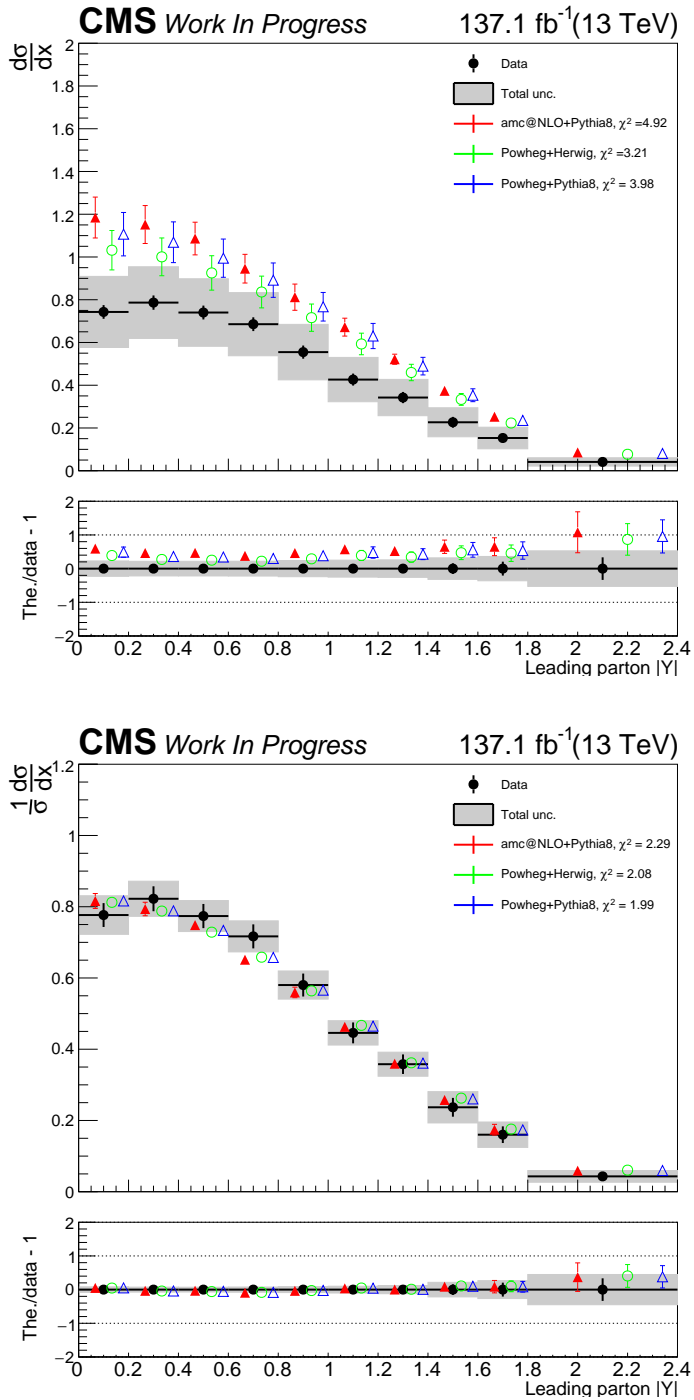


Figure 5.70: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of absolute leading jet rapidity. The bottom panel shows the ratio (theory - data)/data.

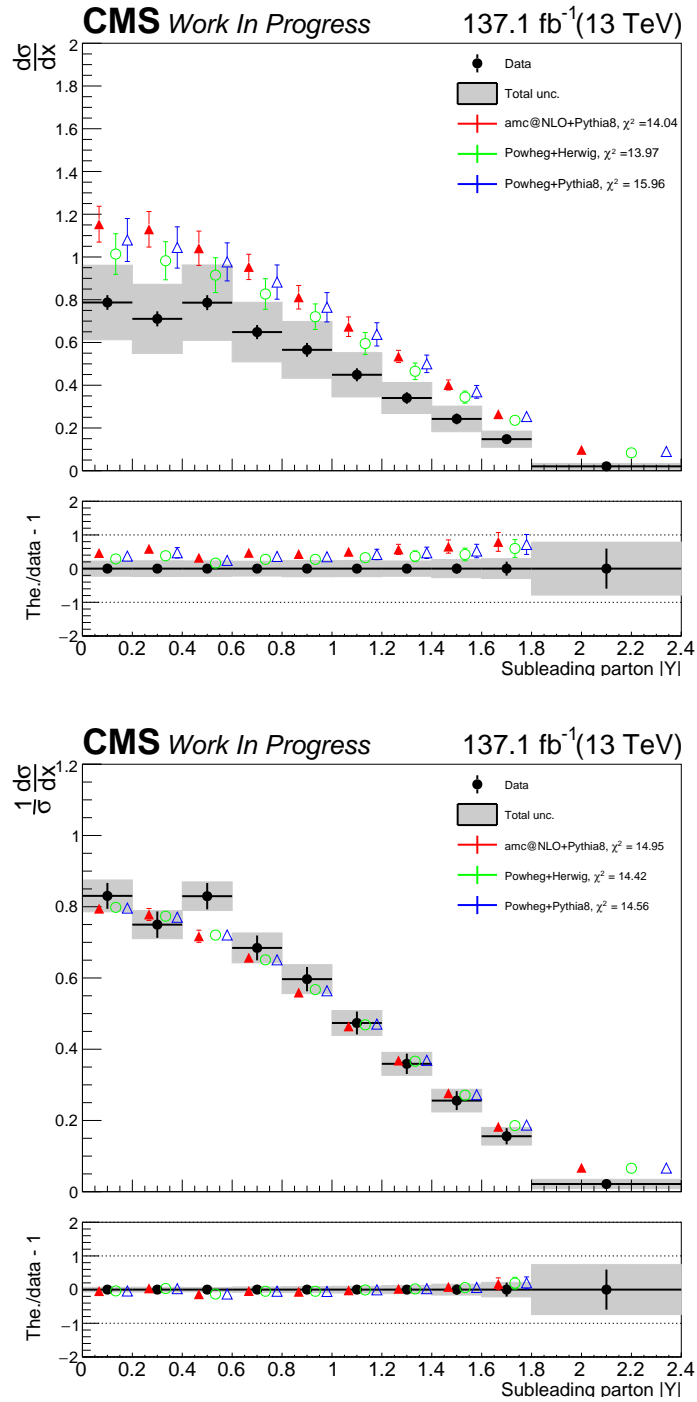


Figure 5.71: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of absolute second leading jet rapidity. The bottom panel shows the ratio (theory - data)/data.

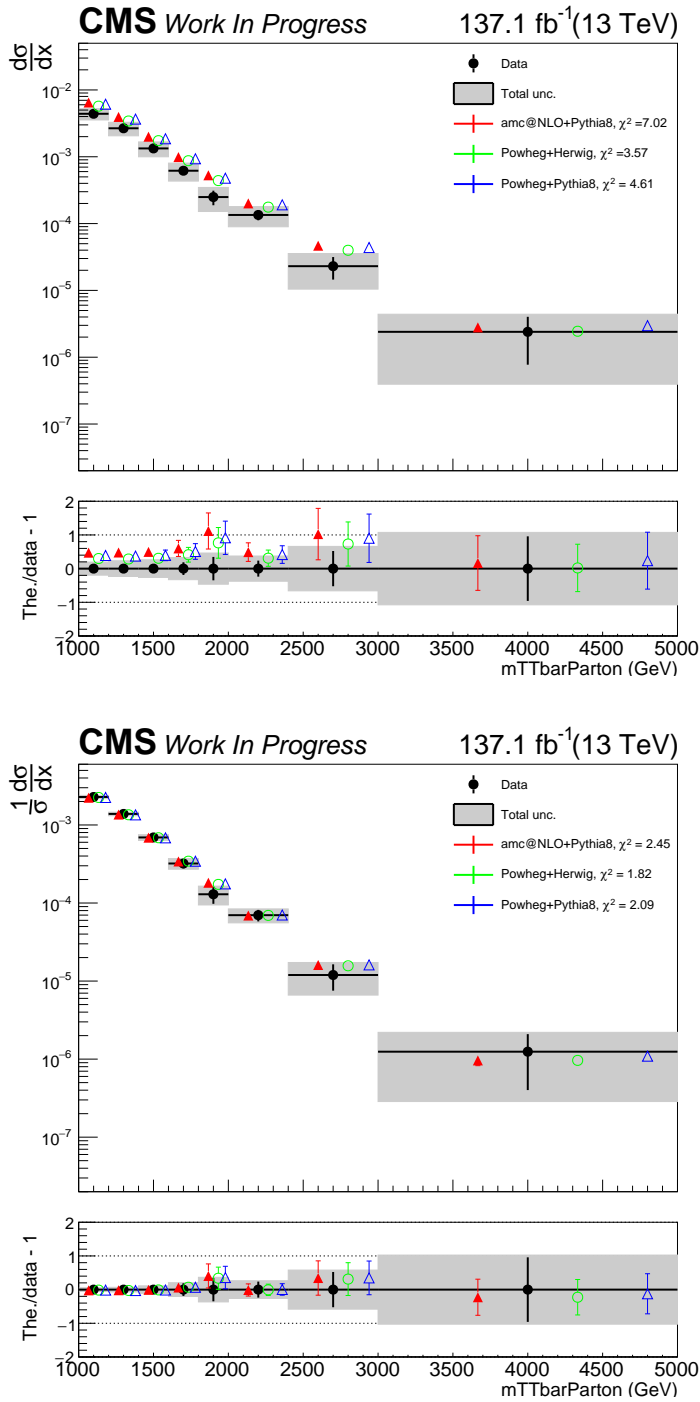


Figure 5.72: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the dijet  $t\bar{t}$  mass. The bottom panel shows the ratio (theory - data)/data.

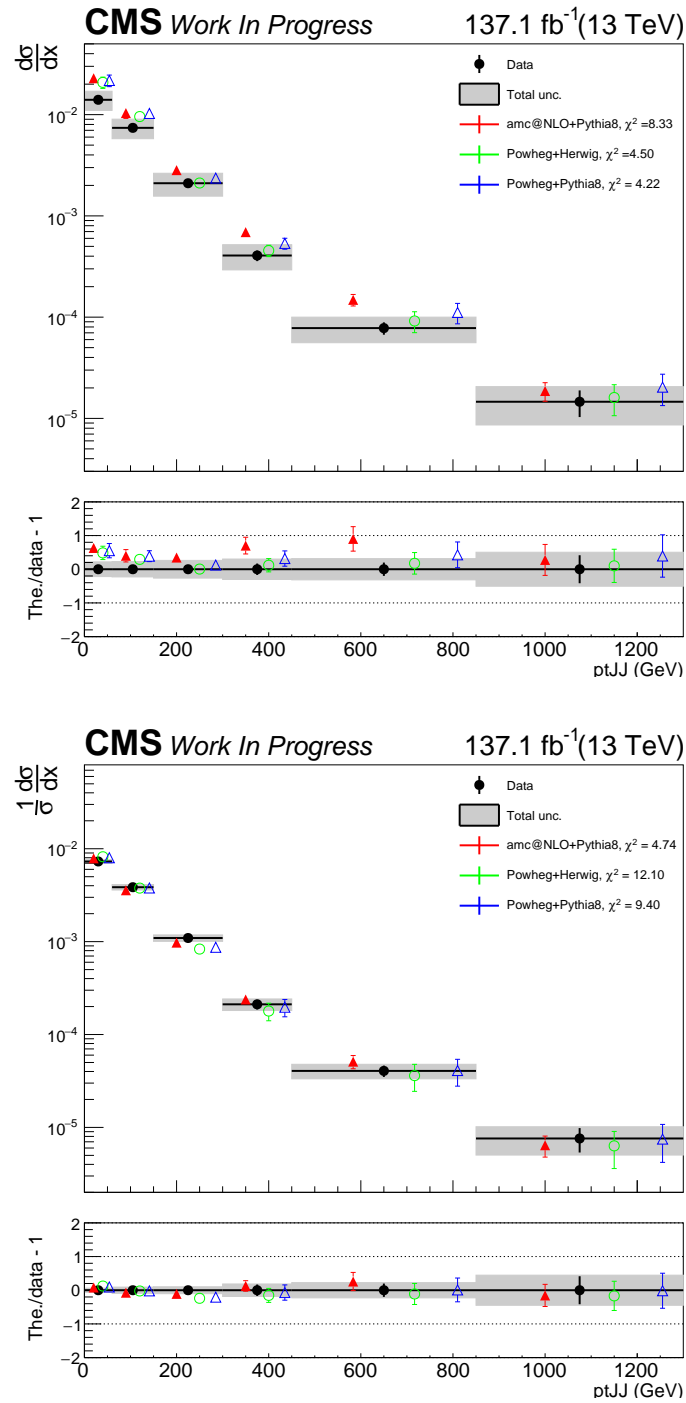


Figure 5.73: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of  $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

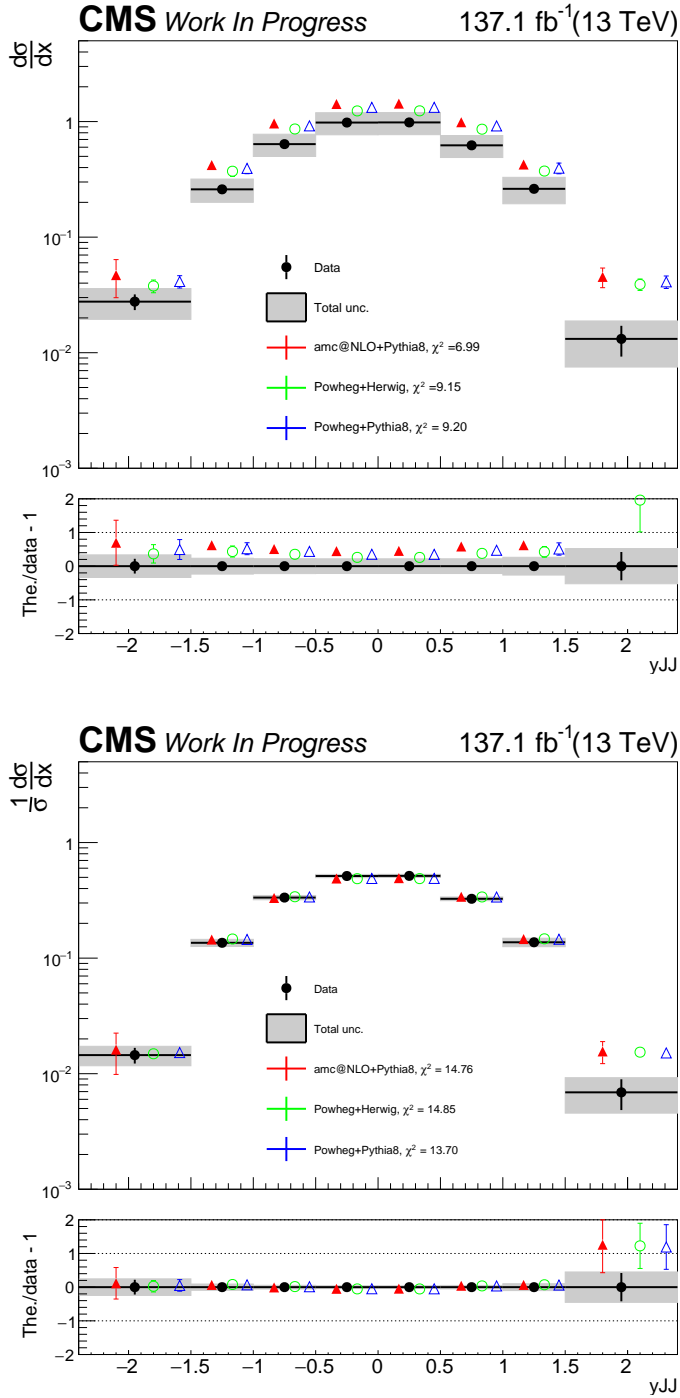


Figure 5.74: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the dijet rapidity  $y_{JJ}^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

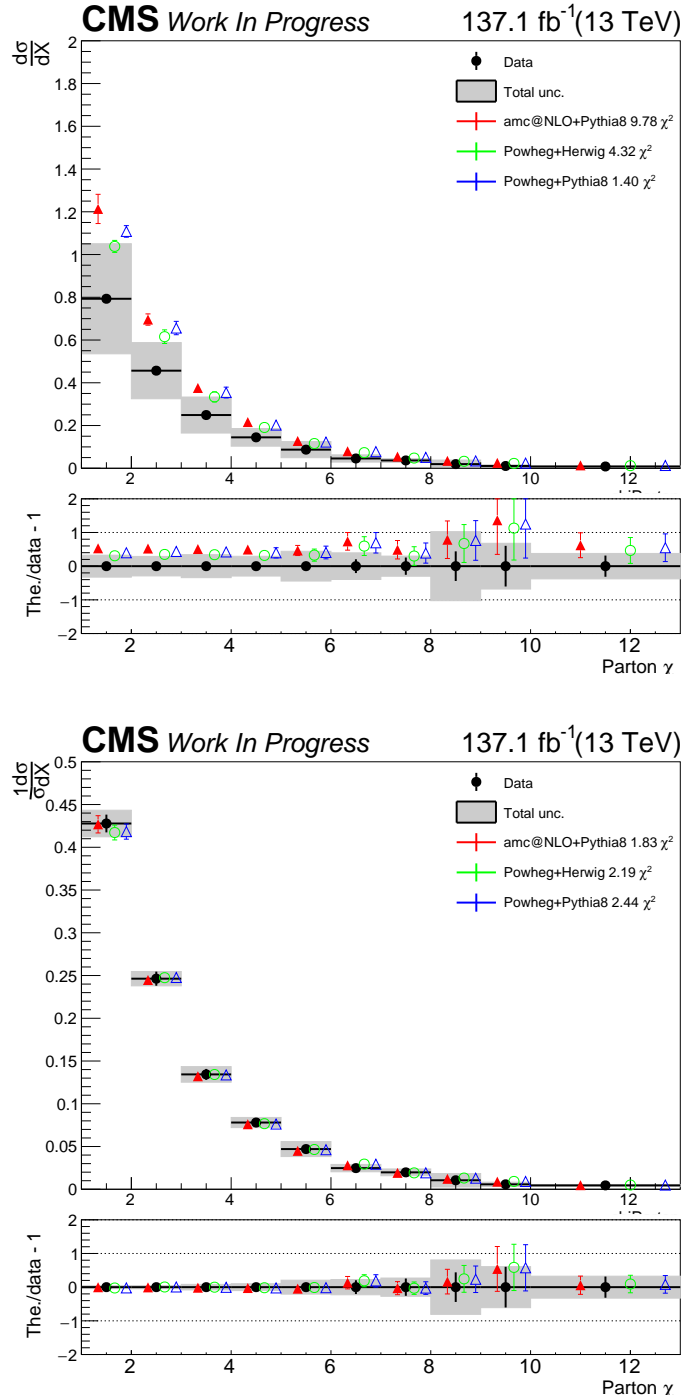


Figure 5.75: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the dijet angular  $\chi^{\text{tt}}$ . The bottom panel shows the ratio (theory - data)/data.

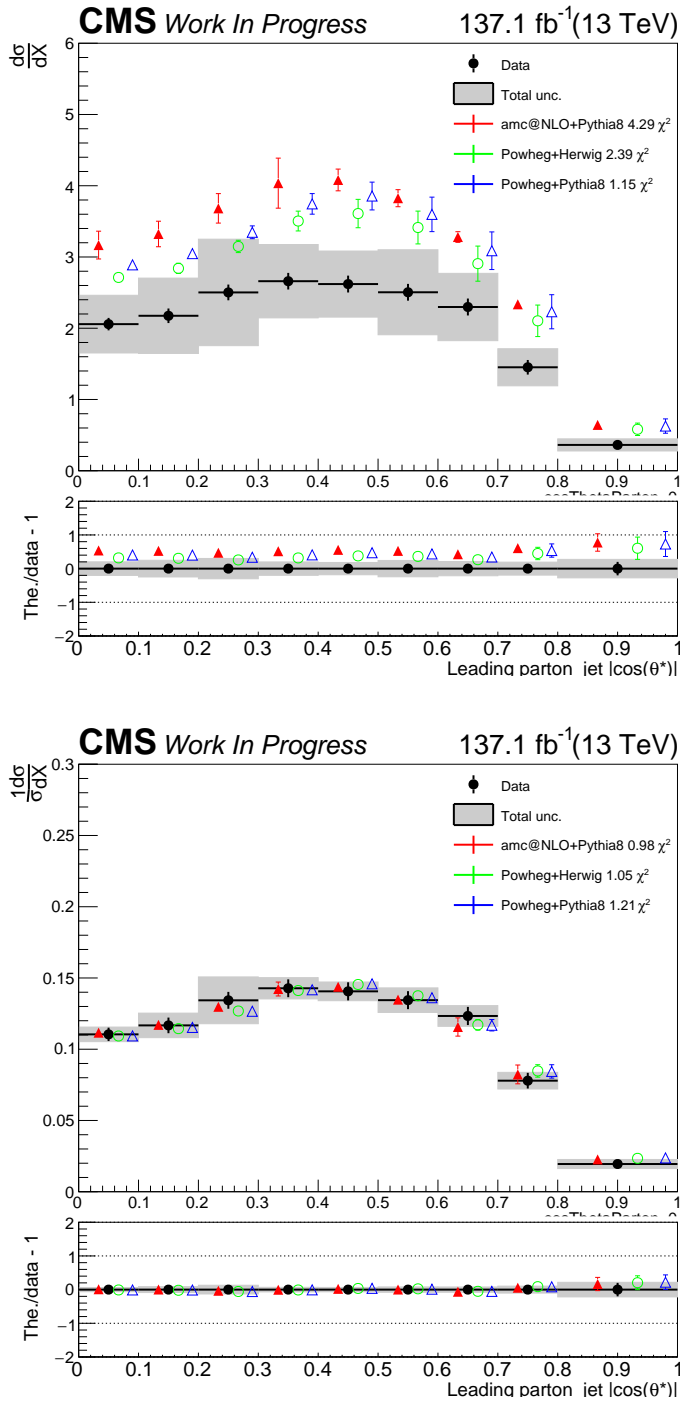


Figure 5.76: Differential cross section unfolded to parton level, absolute (top part) and normalized (bottom part), as a function of the leading jet  $|\cos \theta^*|$ . The bottom panel shows the ratio (theory - data)/data.



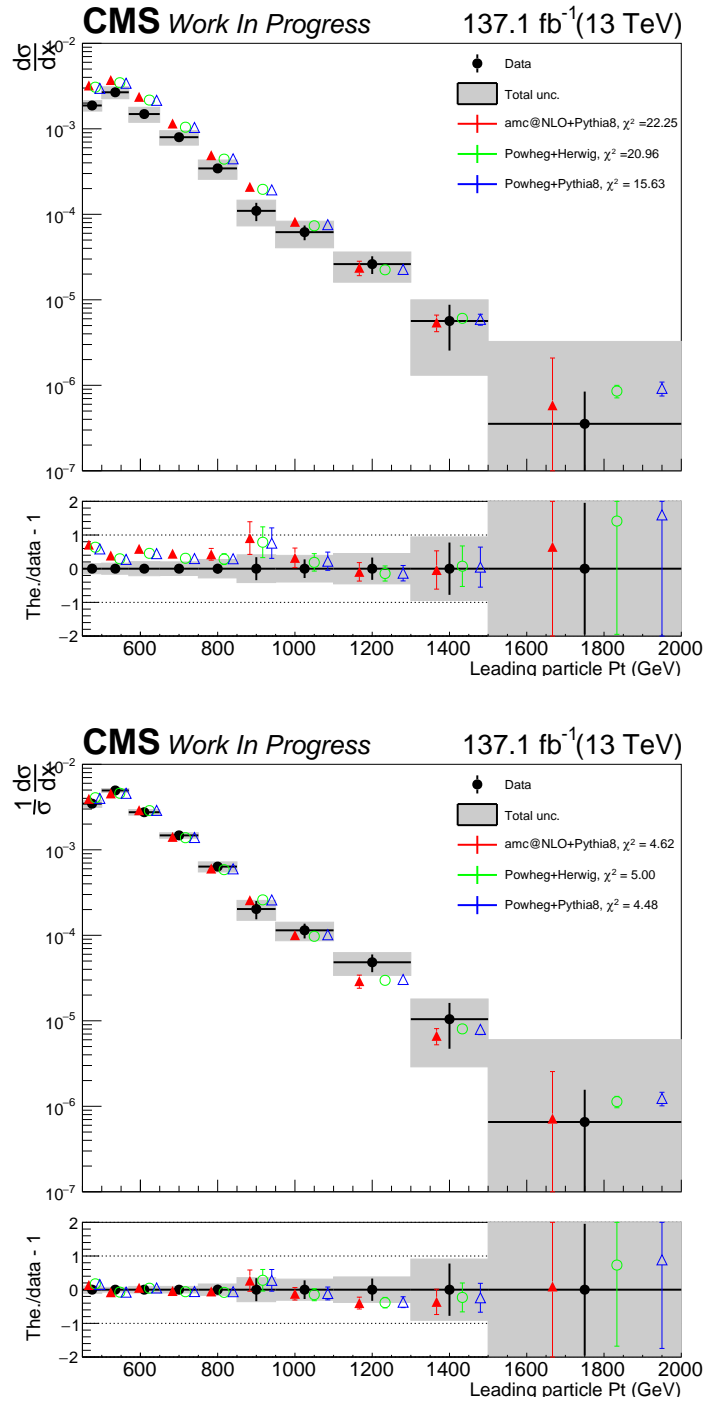


Figure 5.77: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of leading jet  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

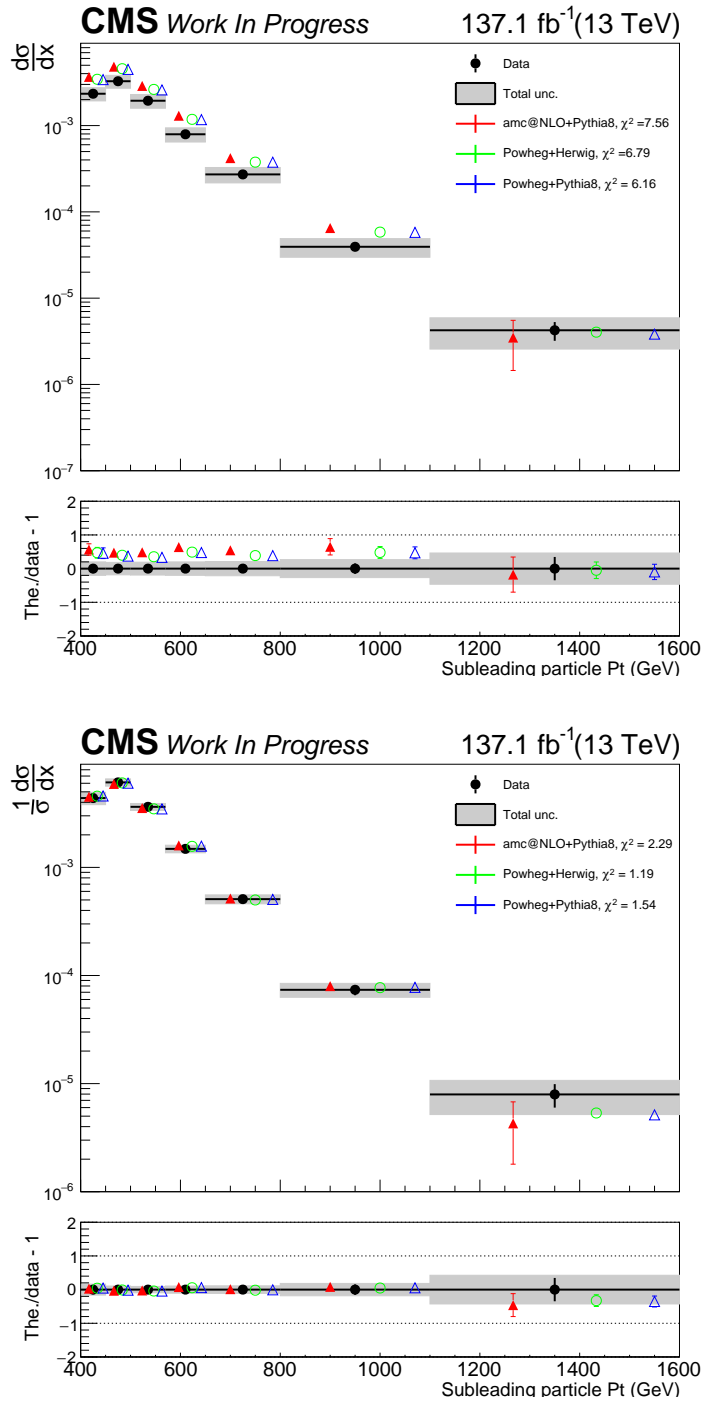


Figure 5.78: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of second leading jet  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

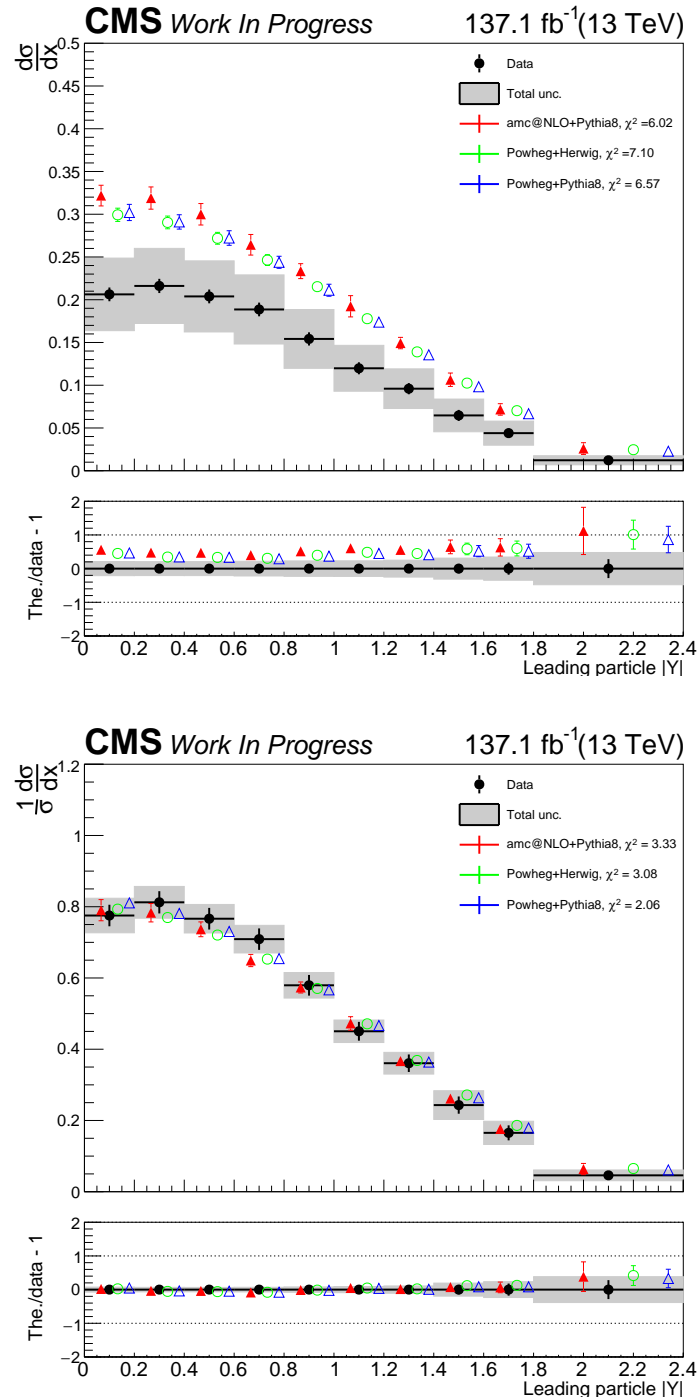


Figure 5.79: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of absolute leading jet rapidity. The bottom panel shows the ratio (theory - data)/data.

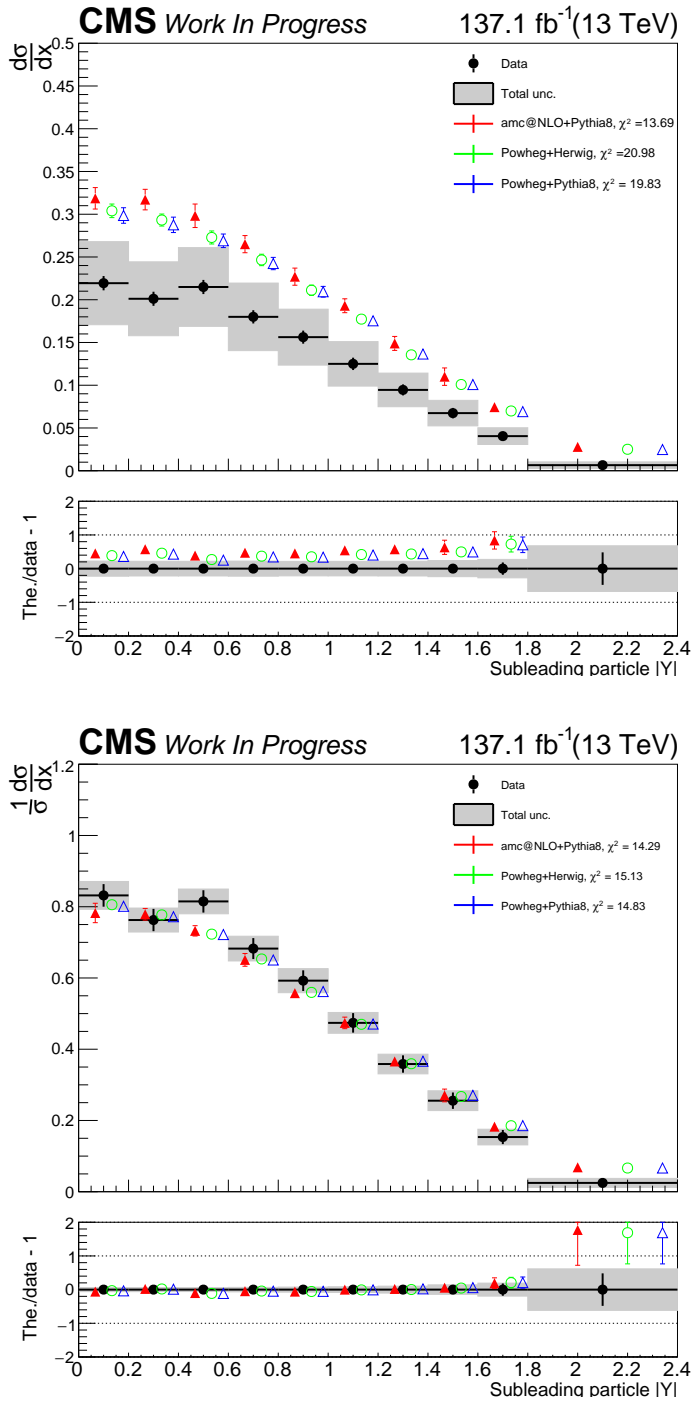


Figure 5.80: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of absolute second leading jet rapidity. The bottom panel shows the ratio (theory - data)/data.

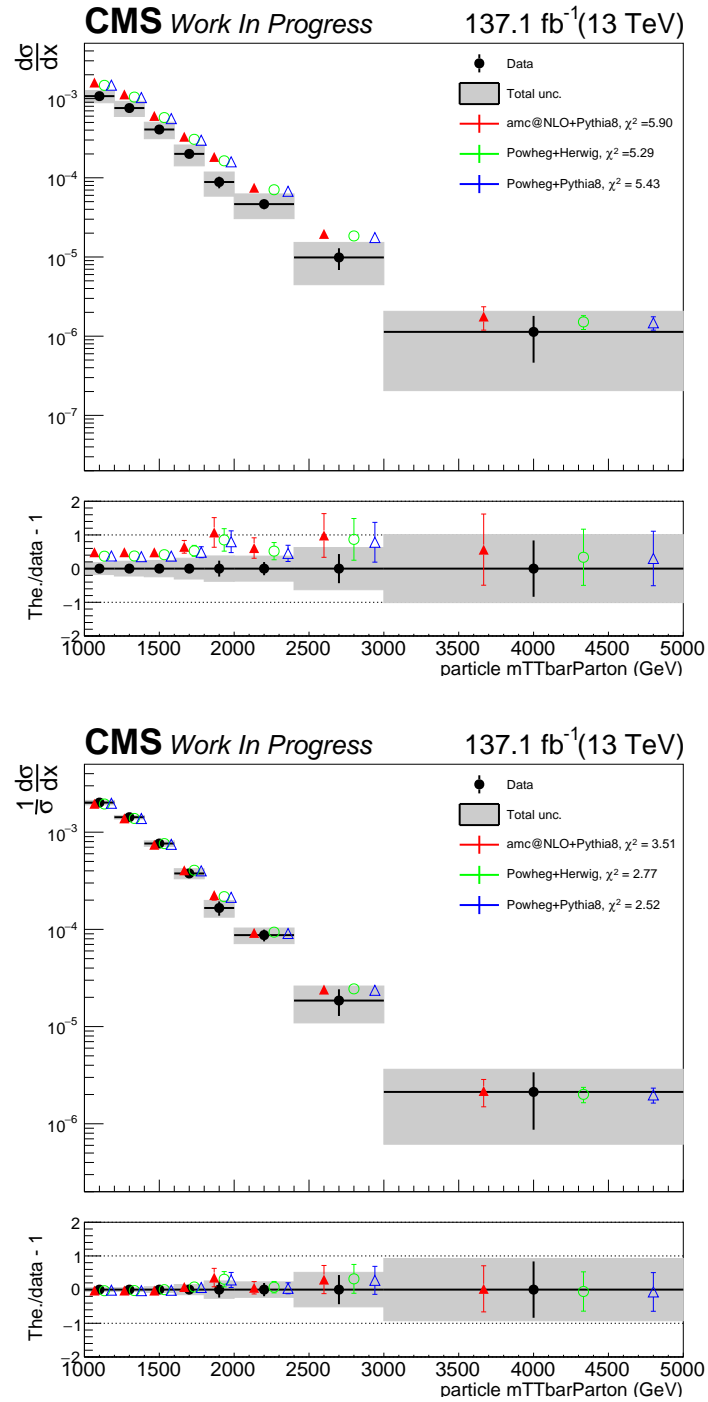


Figure 5.81: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the dijet  $t\bar{t}$  mass. The bottom panel shows the ratio (theory - data)/data.

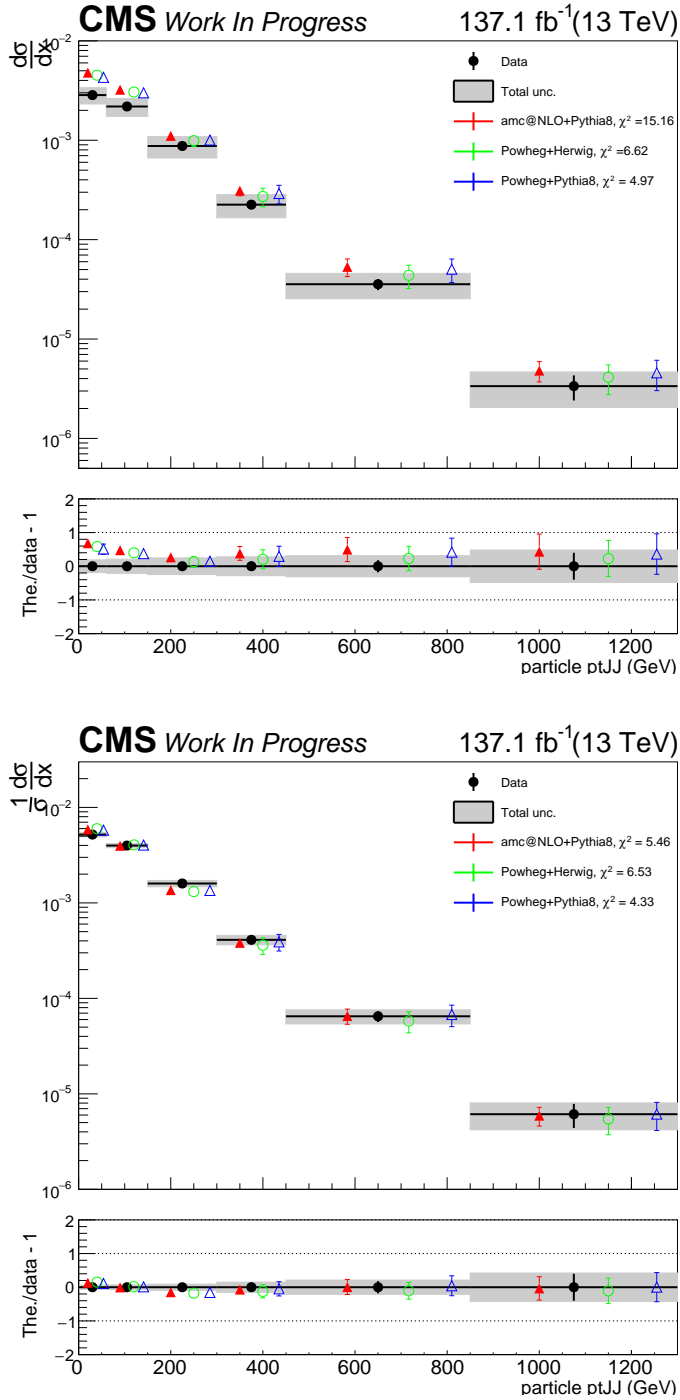


Figure 5.82: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of  $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

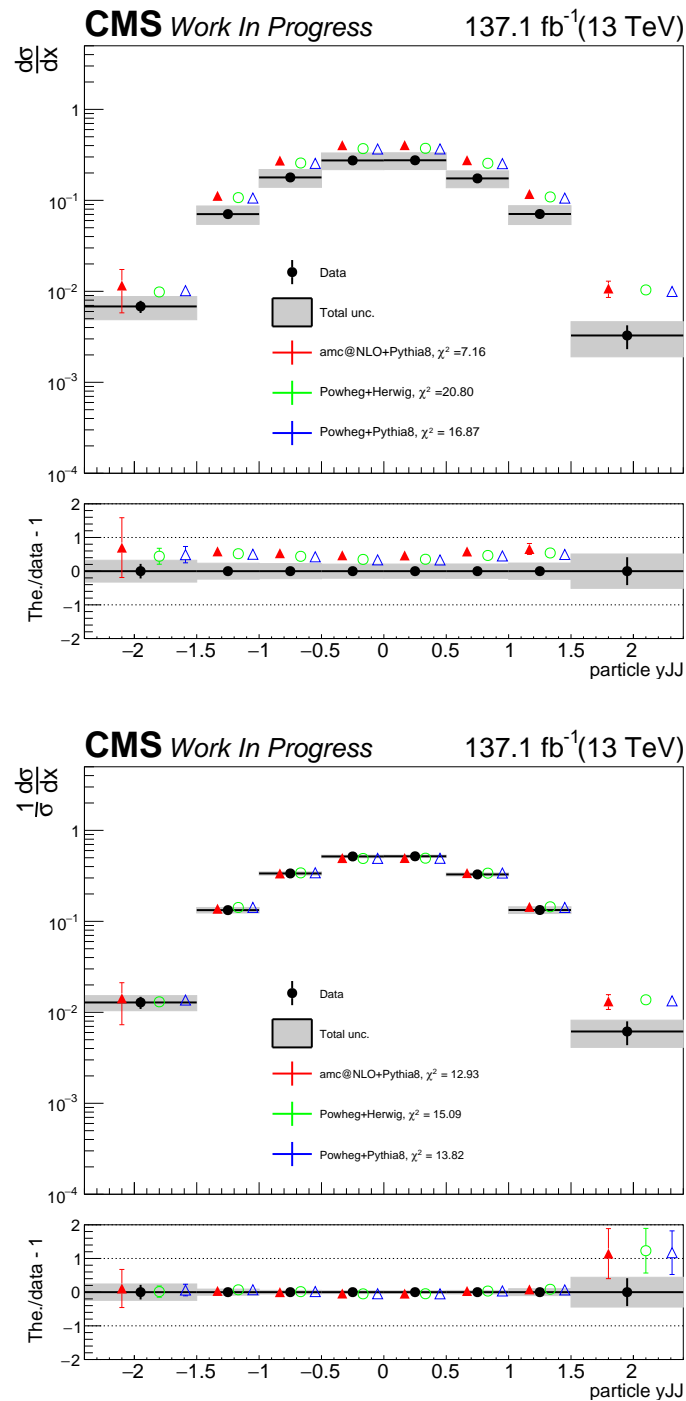


Figure 5.83: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the dijet rapidity  $y_{JJ}^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

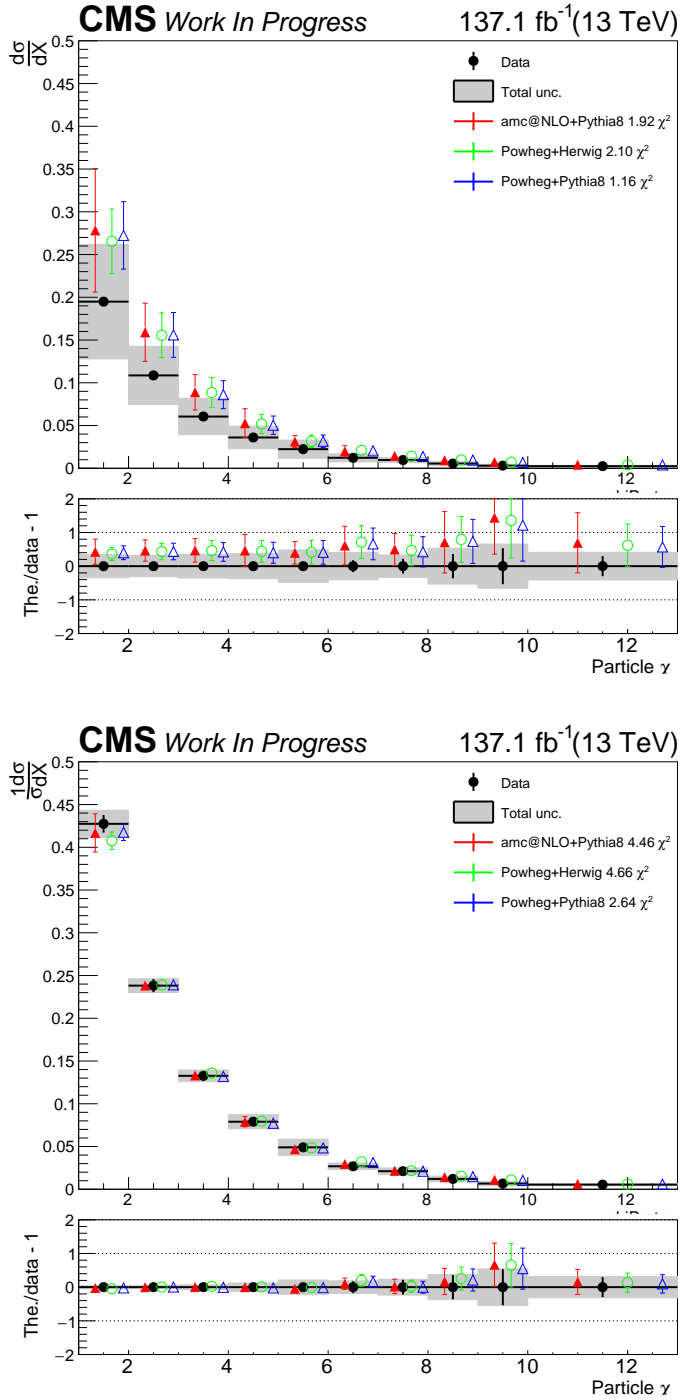


Figure 5.84: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the dijet angular  $\chi^{\text{t}\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.



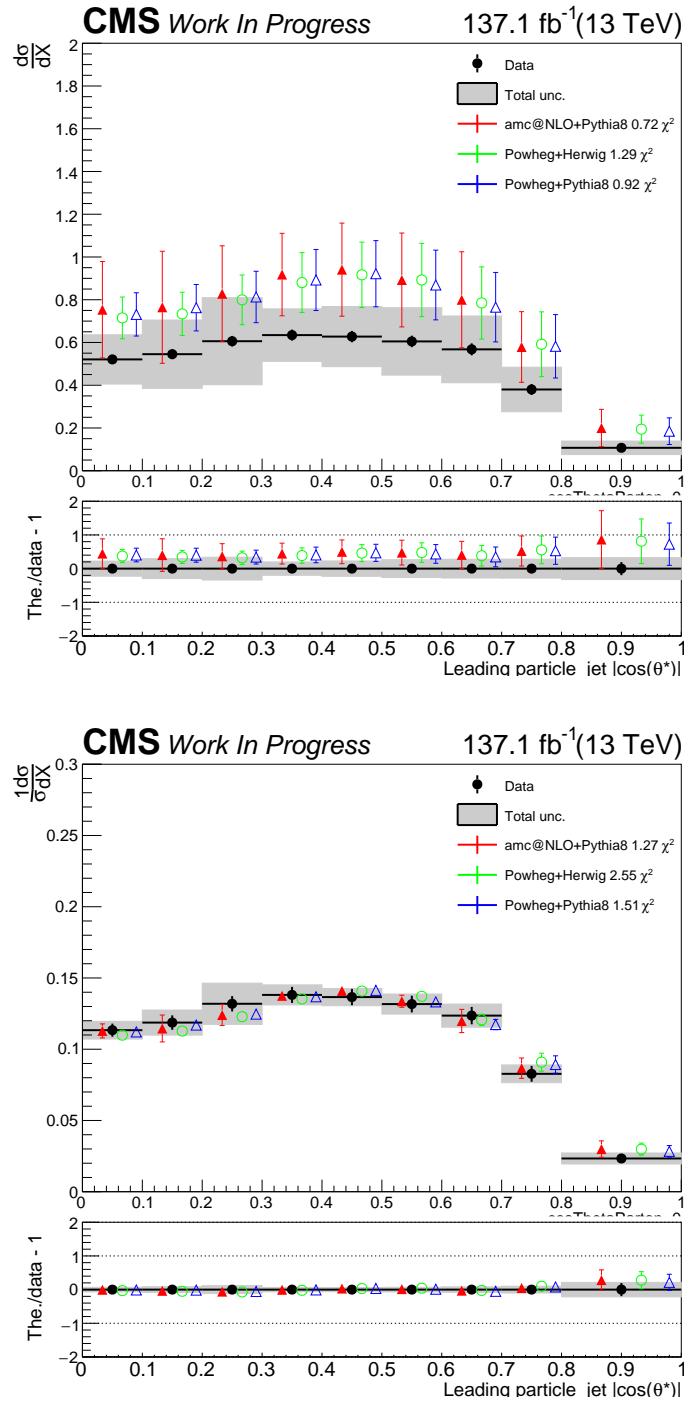


Figure 5.85: Differential cross section unfolded to particle level, absolute (top part) and normalized (bottom part), as a function of the leading jet  $\cos\theta^*$ . The bottom panel shows the ratio (theory - data)/data.



## Chapter 6

# Search for top-antitop resonances in the all-hadronic final state using the full Run II proton-proton collision data at $\sqrt{s} = 13$ TeV

In this chapter, we present a combination of searches for heavy resonances decaying to top quark-antiquark pairs, using the full set of proton-proton collision data collected at a center of-mass energy of  $\sqrt{s} = 13$  TeV in the Full Run II, corresponding to an integrated luminosity of  $137 \text{ fb}^{-1}$ . The combination includes contributions from the fully hadronic final state top and antitop quarks that have a high transverse momentum ( $p_T$ ). The results presented here further increase the mass exclusion limits for several models of new physics.

### 6.1 Introduction to the analysis

The top quark completes the third generation of quarks in the standard model (SM), and the precise knowledge of its properties is critical for the overall understanding of the theory. Measurements of the top-anti-top quark pair ( $t\bar{t}$ ) production cross section confront the predictions from quantum chromodynamics (QCD) and have the potential to constrain the QCD parameters, while being sensitive to physics beyond the SM. Also, the  $t\bar{t}$  production process is a dominant SM background to searches for new physics phenomena and therefore its precise knowledge is essential for new discoveries.

Numerous extensions of the standard model (SM) predict the existence of new interactions with enhanced couplings to third-generation quarks, especially the top quark. The associated massive new particle contained in these theories could be observed as a  $t\bar{t}$  resonance in experiments at LHC. Examples of such resonances are massive color-singlet  $Z$ -like bosons ( $Z'$ ) in extended gauge theories [113, 114], colorons [115] and axigluons [116] in models with extended strong interaction sectors, heavier Higgs siblings in models with extended Higgs sectors and Kaluza-Klein (KK) excitations of gluons [117], electroweak gauge bosons [118], and gravitons in various extensions of the Randall-Sundrum (RS) model [119]. These models predict the existence of TeV-scale resonances with production cross sections of a few picobarns at  $\sqrt{s} = 13$  TeV. In all of these examples, resonant  $t\bar{t}$  production could be observable in the exponent of the absolute rapidity difference between the two leading jets, named  $\chi$ .

The CMS experiment has produced several public results of searches for heavy resonances decaying into  $t\bar{t}$  pairs using datasets collected in 2016 preVFP, 2017 and 2018 for the full Run II analysis which amount to a total integrated luminosity of  $19.5 \text{ fb}^{-1}$ ,  $41.5 \text{ fb}^{-1}$  and  $59.7 \text{ fb}^{-1}$  respectively. These searches probe the same physics models, including generic  $Z'$  resonances having widths of 1% of the  $Z'$  mass [?, ?]. In each of these models, the heavy particle ( $Z'$ ) decays into a  $t\bar{t}$  pair. Due to the high mass regime probed (0.5-4.5 TeV), the top quarks are expected to be highly boosted, requiring special reconstruction techniques. The search to be presented in this result includes the search probing the all-hadronic analysis, in which both top products decay into hadrons using top-tagging in conjunction with subjet b-tagging.

This analysis note will not describe in detail the aforementioned channel, but instead will

focus on the strategy and preliminary results mentioned above. Further details on the individual analysis can be found in the documentation of the all-hadronic channel chapter.

## 6.2 Simulated events

The simulation of  $Z'$  resonances is performed with the leading-order MADGRAPH v5.2.2.2 [126] Monte Carlo (MC) program using SM values for the left- and right-handed  $Z'$  couplings to top quarks. The simulation is performed for a range of  $Z'$  masses between 0.5 and 4.0 TeV, and for one relative width hypothesis of 1%. Higher-order QCD multijet processes for up to three extra partons are simulated at tree level. The  $Z'$  boson is required to decay into a  $t\bar{t}$  pair in all generated events. The parton showering and hadronization is modeled with PYTHIA 8.205 [120, 121], and the MLM algorithm [122] is used to match the parton shower to the matrix element calculation with a merging scale of 35 GeV.

Background events from  $t\bar{t}$  production via QCD interactions and electroweak production of single top quarks in the  $tW$  channel are simulated with the next-to-leading order (NLO) generator POWHEG (v2) [123, 124, 125]. The s- and t-channel processes of single top quark production are simulated with MADGRAPH5 aMC@NLO v5.2.2.2. All events are interfaced with PYTHIA for the description of fragmentation and hadronization.

The associated production of W or Z boson and jets is simulated using MADGRAPH. The MLM matching scheme is applied to match the showers generated with PYTHIA. Up to four additional partons in the matrix element calculations are included. The  $t\bar{t}$ , W/Z+jets, and singletop-quark samples are normalized to the theoretical predictions described in Refs. [127, 128]. Diboson processes ( $VV = WW, WZ, \text{ and } ZZ$ ) are simulated with PYTHIA for both the matrix element and parton showering calculations. The event rates are normalized to the NLO crosssections from Ref. [129].

For this analysis, only 2016 preVFP, 2017 AND 2018 files were available ranging the  $Z'$  masses between 0.5 and 4.0 TeV, for one relative mass width hypothesis of 1%.

Table 6.1: Monte Carlo simulated 2016  $Z'$  samples.

Sample	Events	$\sigma$ (pb)
ZprimeToTT_M1200_W12_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	497321	1.73
ZprimeToTT_M1400_W14_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	560171	9.095e-01
ZprimeToTT_M1600_W16_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	548012	5.002e-01
ZprimeToTT_M1800_W18_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	533048	2.833e-01
ZprimeToTT_M2000_W20_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	527279	1.662e-01
ZprimeToTT_M2500_W25_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	508484	4.749e-02
ZprimeToTT_M3000_W30_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM -pythia8	490008	1.494e-02
ZprimeToTT_M3500_W35_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	473382	5.105e-03
ZprimeToTT_M4000_W40_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM -pythia8	460416	1.900e-03

## 6.3 Analysis Strategy

The baseline selection, summarized in Table 6.4 is common for all regions used in the analysis. It requires at least two large-R jets ( $R = 0.8$ ) in the event with  $p_T > 400$  GeV and softDrop mass values within the range (50, 300) GeV. Furthermore, a veto on the leptons is applied in order to minimize the probability to select leptonic top decays. On top of the baseline selection, we define four specific regions (Table 6.4) based on the output of our BDT (Boosted Decision Tree) that discriminates  $t\bar{t}$  signal from QCD background, the jets' softDrop masses and the number of the b-tagged subjects in each jet, that serve different analysis purposes. The signal region ( $SR$ ) is where we perform the differential measurements and it required both jets to have a b-tagged subset, a tighter selection

Table 6.2: Monte Carlo Simulated 2017  $Z'$  samples.

Sample	Events	$\sigma$ (pb)
ZprimeToTT_M1200_W12_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	195613	1.73
ZprimeToTT_M1400_W14_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	208708	9.095e-01
ZprimeToTT_M1600_W16_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	219802	5.002e-01
ZprimeToTT_M1800_W18_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	197407	2.833e-01
ZprimeToTT_M2000_W20_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	211425	1.662e-01
ZprimeToTT_M2500_W25_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	203518	4.749e-02
ZprimeToTT_M3000_W30_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM -pythia8	192535	1.494e-02
ZprimeToTT_M3500_W35_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	189565	5.105e-03
ZprimeToTT_M4000_W40_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM -pythia8	184083	1.900e-03
ZprimeToTT_M4500_W45_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	174039	7.613e-04

Table 6.3: Monte Carlo Simulated 2018  $Z'$  samples.

Sample	Events	$\sigma$ (pb)
ZprimeToTT_M1200_W12_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	198217	1.73
ZprimeToTT_M1400_W14_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	217763	9.095e-01
ZprimeToTT_M1600_W16_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	207930	5.002e-01
ZprimeToTT_M1800_W18_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	203473	2.833e-01
ZprimeToTT_M2000_W20_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	196868	1.662e-01
ZprimeToTT_M2500_W25_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	201970	4.749e-02
ZprimeToTT_M3000_W30_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM -pythia8	194745	1.494e-02
ZprimeToTT_M3500_W35_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	187591	5.105e-03
ZprimeToTT_M4000_W40_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM -pythia8	184021	1.900e-03
ZprimeToTT_M4500_W45_TuneCP2_PWeights_13TeV-madgraph-pythiaMLM-pythia8	179341	7.613e-04

of the jet mass softDrop value 120 – 220 GeV. The pure QCD control region ( $CR$ ) has the same selection criteria with the aforementioned Signal Region, with the only difference being that the b-tagging requirement is reverted (the jets must not contain a b-tagged subjet). This region is used to get the QCD shape of the QCD background from the data (data driven method) for each variable of interest. Then, we employ the signal region A ( $SR_A$ ), which is used to determine the normalization of the QCD background, and finally the signal region B ( $SR_B$ ), which is used to constrain some of the signal modelling uncertainties.

The dominant background in the analysis is the QCD multijet production. There is a finite probability that ordinary jets, originating from single parton radiation, mimic the topological substructure of a top jet. The QCD multijet production is suppressed significantly using the aforementioned developed Boosted Decision Tree, which uses jet substructure information to distinguish signal from background. Also, the b-tagging requirements play a huge role in suppressing the QCD multijet production. In order to estimate the remaining contribution of the QCD multijet production we employ a data-driven technique based on the assumption that if we revert the b-tagging requirement (asking for 0 btagged subjets in the event) we a) get a pure QCD multijet sample and b) the jet kinematic properties are not affected. Our Control Region works as a phase space, where (theoretically) only QCD multijets exists. In contrast, our CR consists not only from QCD multijet production contribution but also from  $t\bar{t}$  and Subdominant backgrounds contribution ( $W$ +jets, etc) to a factor of 10%. For that reason, we remove these contributions from the CR before we continue with QCD multijet estimation data-driven method. In this way we are able to extract the remaining signal in our SR.

The variable of interest for this search is the  $\chi$  dijet variable. We choose this variable because its distributions associated with the final states produced via QCD interactions are relatively flat compared to the distributions of the BSM models or new particles, which typically peak at low

Table 6.4: Baseline Selection Criteria

Observable	Requirement
$N_{\text{jets}}$	$> 1$
$N_{\text{leptons}}$	$0$
$p_T^{\text{jet}1}$	$> 450 \text{ GeV}$
$p_T^{\text{jet}2}$	$> 400 \text{ GeV}$
$ \text{et al.}^{\text{jet}1,2} $	$< 2.4$
$m_{SD}^{\text{jet}1,2}$	$(50, 300) \text{ GeV}$
$m_{t\bar{t}}$	$> 1000 \text{ GeV}$

Table 6.5: Selection requirements per analysis region for 2016 pre and post VFP.

Region	Trigger	Offline	Purpose
$SR$	Signal	Baseline+ $BDT > 0.2 + 2btag + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	Signal Region
$SR_A$	Signal	Baseline+ $BDT > 0.2 + 2btags$	QCD fit region
$CR$	Control	Baseline+ $BDT > 0.2 + 0btag + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	QCD control region

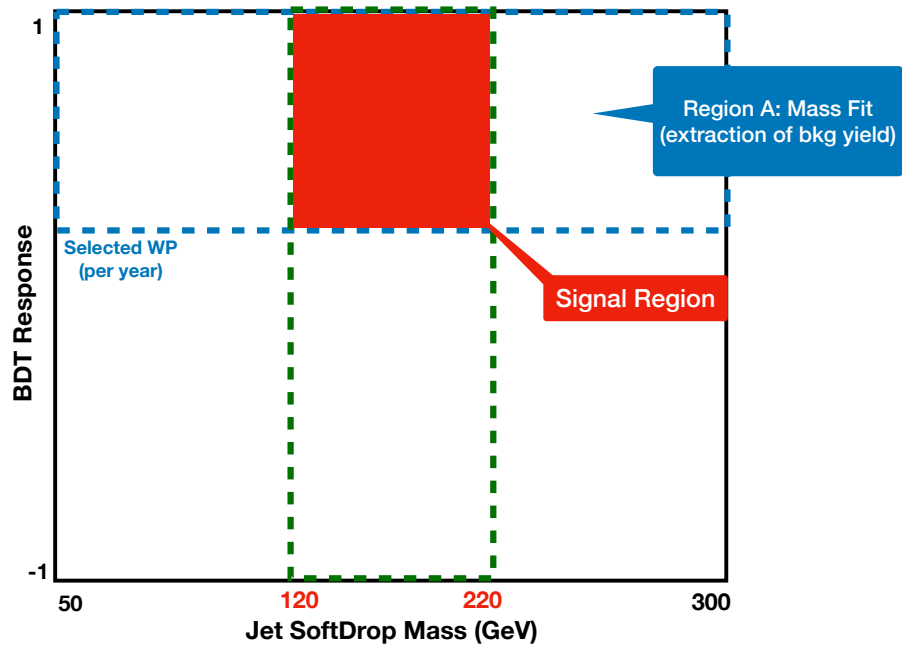


Figure 6.1: Graphical representation of the analysis region in the BDT-output vs jet softDrop mass.

values of  $\chi$ . The choice of this variable, stems from the fact that the dijet mass of the  $t\bar{t}$  pair suffers from Energy Scale uncertainties while the angular dijet variable  $\chi$  is not affected.

The variable is computed by measuring the difference of the rapidities of the two leading jets

Table 6.6: Selection requirements per analysis region for 2017.

Region	Trigger	Offline	Purpose
$SR$	Signal	$\text{Baseline} + BDT > 0.0 + 2btag + m_{SD}^{jet1,2} \in (120, 220) \text{ GeV}$	Signal Region
$SR_A$	Signal	$\text{Baseline} + BDT > 0.0 + 2btags$	QCD fit region
$CR$	Control	$\text{Baseline} + BDT > 0.0 + 0btag + m_{SD}^{jet1,2} \in (120, 220) \text{ GeV}$	QCD control region

Table 6.7: Selection requirements per analysis region for 2018.

Region	Trigger	Offline	Purpose
$SR$	Signal	$\text{Baseline} + BDT > 0.1 + 2btag + m_{SD}^{jet1,2} \in (120, 220) \text{ GeV}$	Signal Region
$SR_A$	Signal	$\text{Baseline} + BDT > 0.1 + 2btags$	QCD fit region
$CR$	Control	$\text{Baseline} + BDT > 0.1 + 0btag + m_{SD}^{jet1,2} \in (120, 220) \text{ GeV}$	QCD control region

such as the corresponding rapidity in the Zero Momentum Frame is :

$$y^* = \frac{1}{2}(y_1 - y_2) \quad (6.1)$$

where  $y_{1,2}$  denote the rapidity of the leading and subleading selected jets respectively. Variable  $\chi$  is defined as:

$$\chi = e^{|2y_*|} = e^{|y_1 - y_2|}. \quad (6.2)$$

In order to further enhance the existence of the  $Z'$  signal within the proposed signal region, different sets of dijet mass cuts were applied for every  $Z'$  mass respectively. We have observed that in lower masses, the angular distribution of  $\chi$  is more competitive than the dijet mass  $m_{JJ}$  distribution. To quantify the optimal dijet mass cut, we calculate the significance per  $Z'$  mass. We define significance as the following ratio:

$$\text{Significance} = \frac{\text{Signal}}{\sqrt{\text{Bkg} + \text{Signal}}}. \quad (6.3)$$

The following figures ( 6.6-6.10) show the Significance for each dijet mass ( $m_{JJ}$ ) cut. From this plot, the cut that maximizes the Significance for each year is derived, for each  $Z'$  (mass, width). Based on the  $m_{JJ}$  cuts that result into maximum significance, the following dijet mass cut are imposed as shown in Table 6.8:

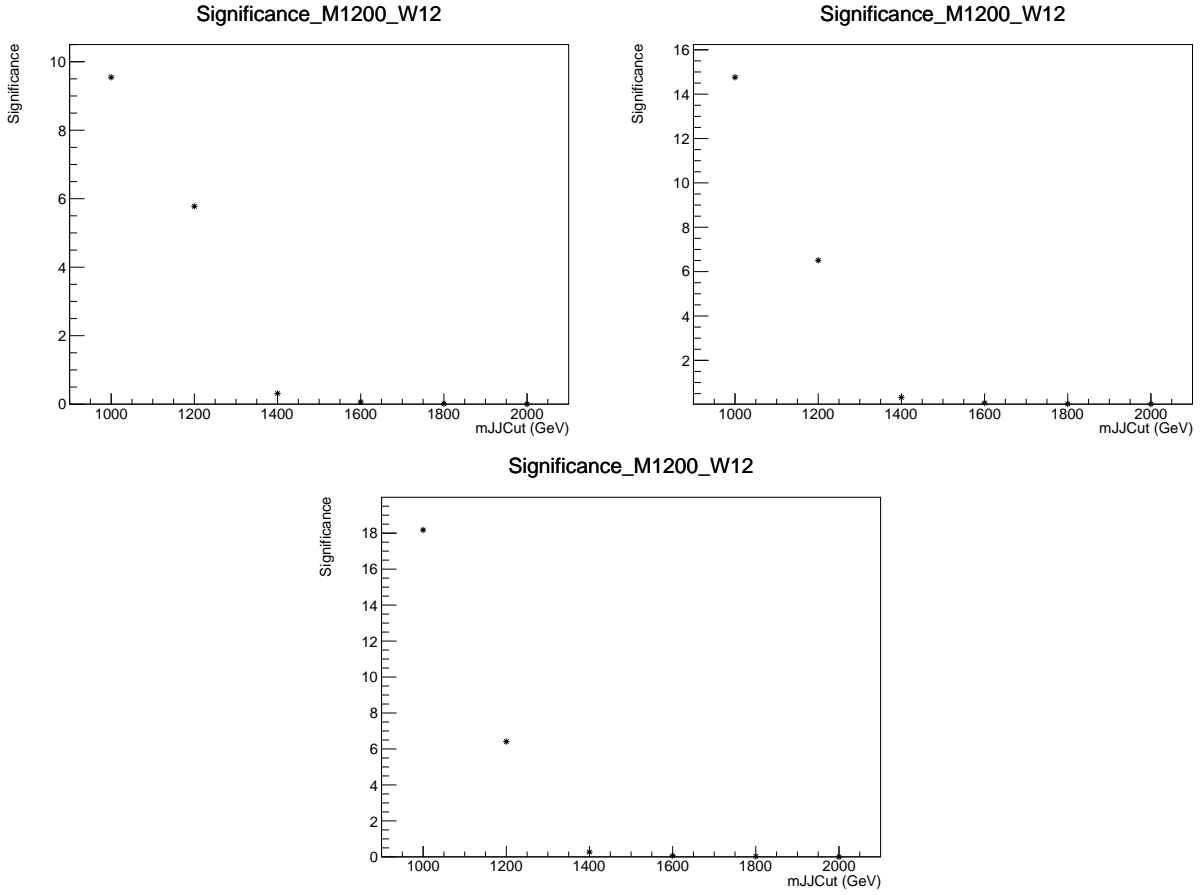


Figure 6.2: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 1.2TeV and 1 % width.

Mass (GeV)	Width (1%) (GeV)	$m_{JJ}$ Cut (GeV)
1200	12	1000
1400	14	1200
1600	16	1600
1800	18	1600
2000	20	1600
2500	25	2000
3000	30	2000
3500	35	2000
4000	40	2000
4500	45	2000

Table 6.8: Sliding  $m_{JJ}$  cut for all  $Z'$  masses. Same cut is applied for all years.



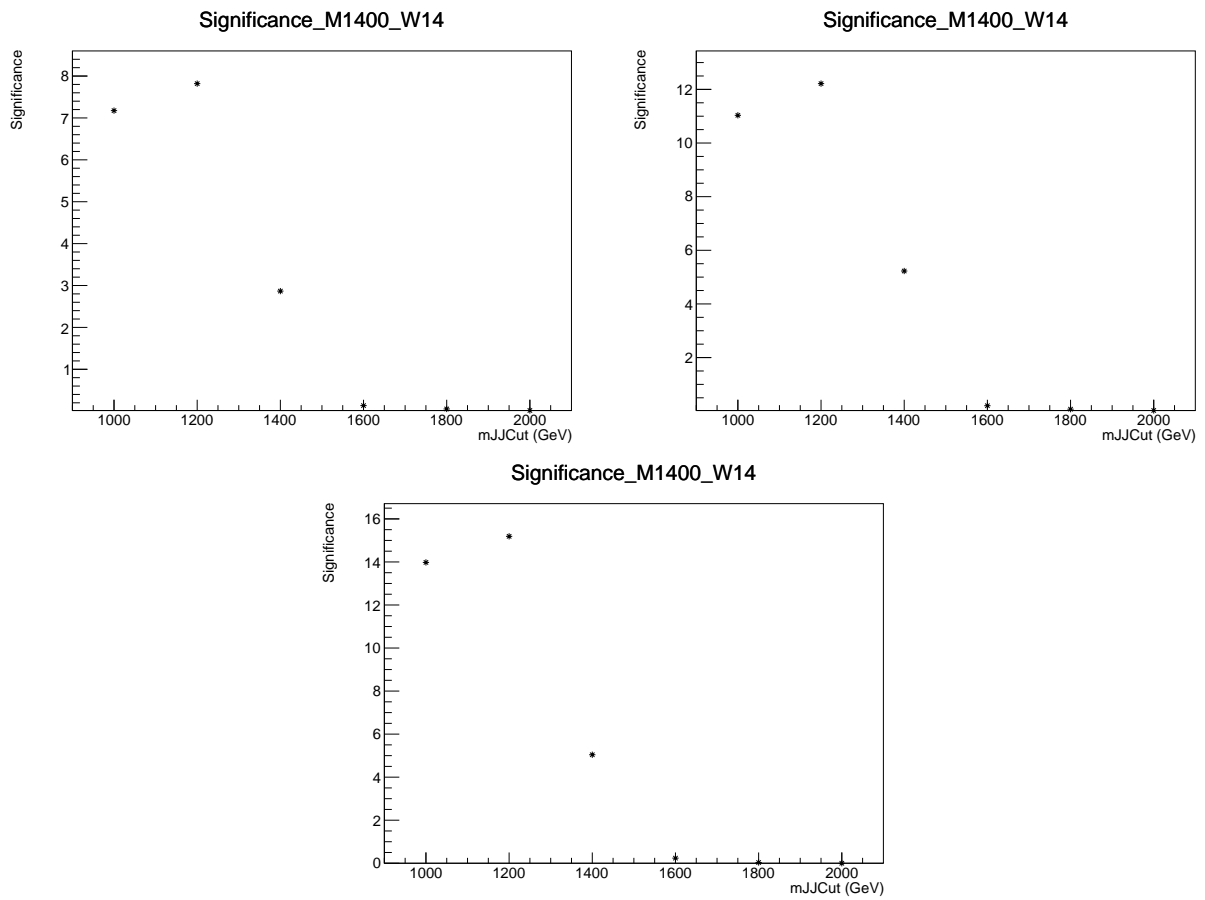


Figure 6.3: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 1.4TeV and 1 % width.

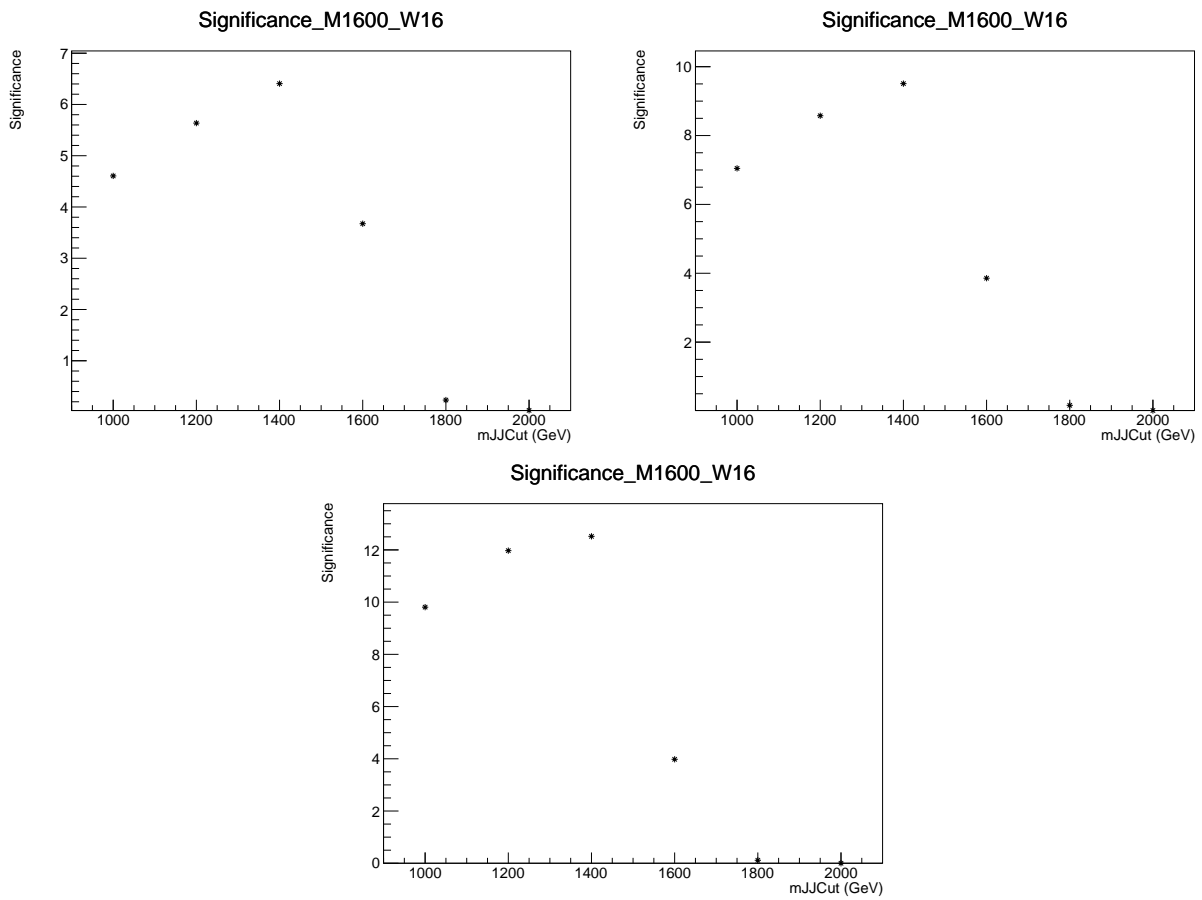


Figure 6.4: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 1.6TeV and 1 % width.

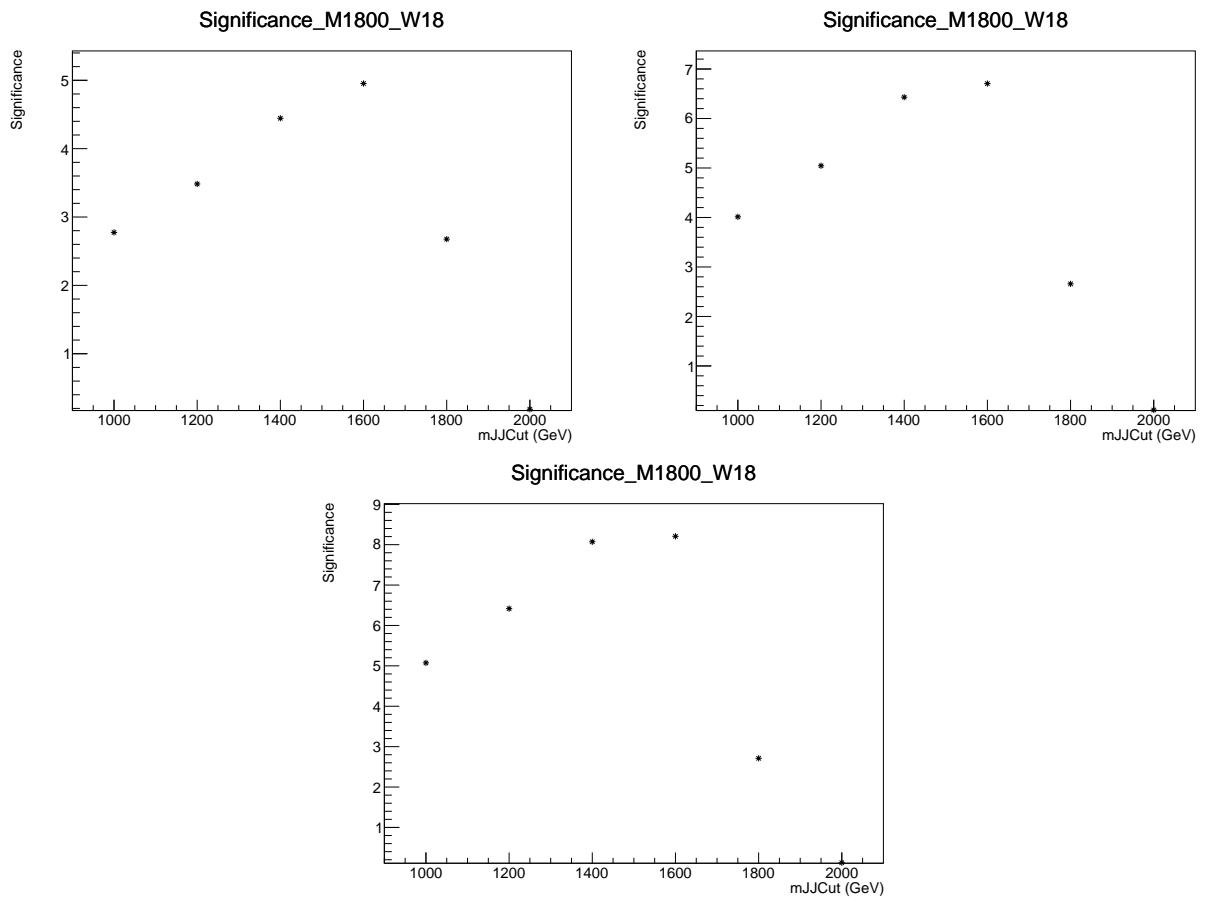


Figure 6.5: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 1.8TeV and 1 % width.

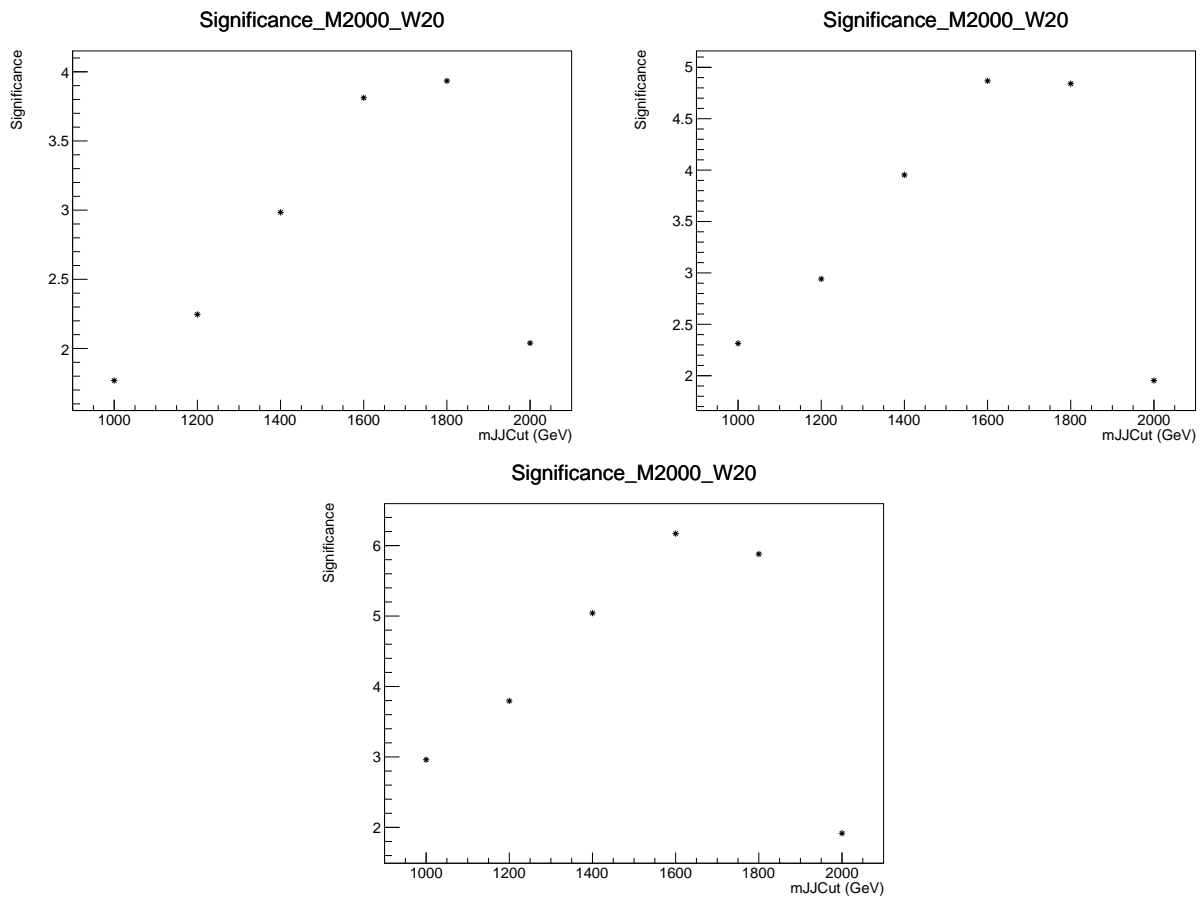


Figure 6.6: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 2TeV and 1 % width.

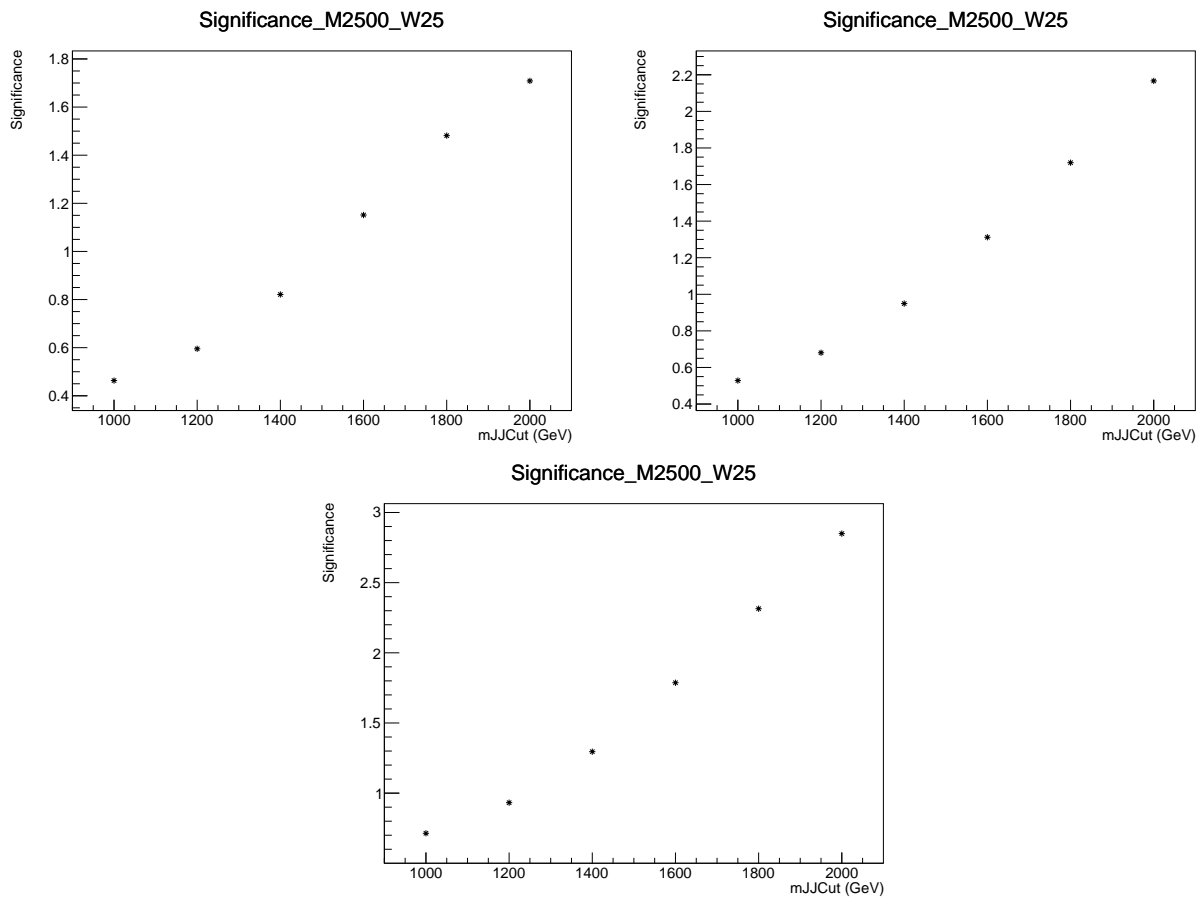


Figure 6.7: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 2.5TeV and 1 % width.

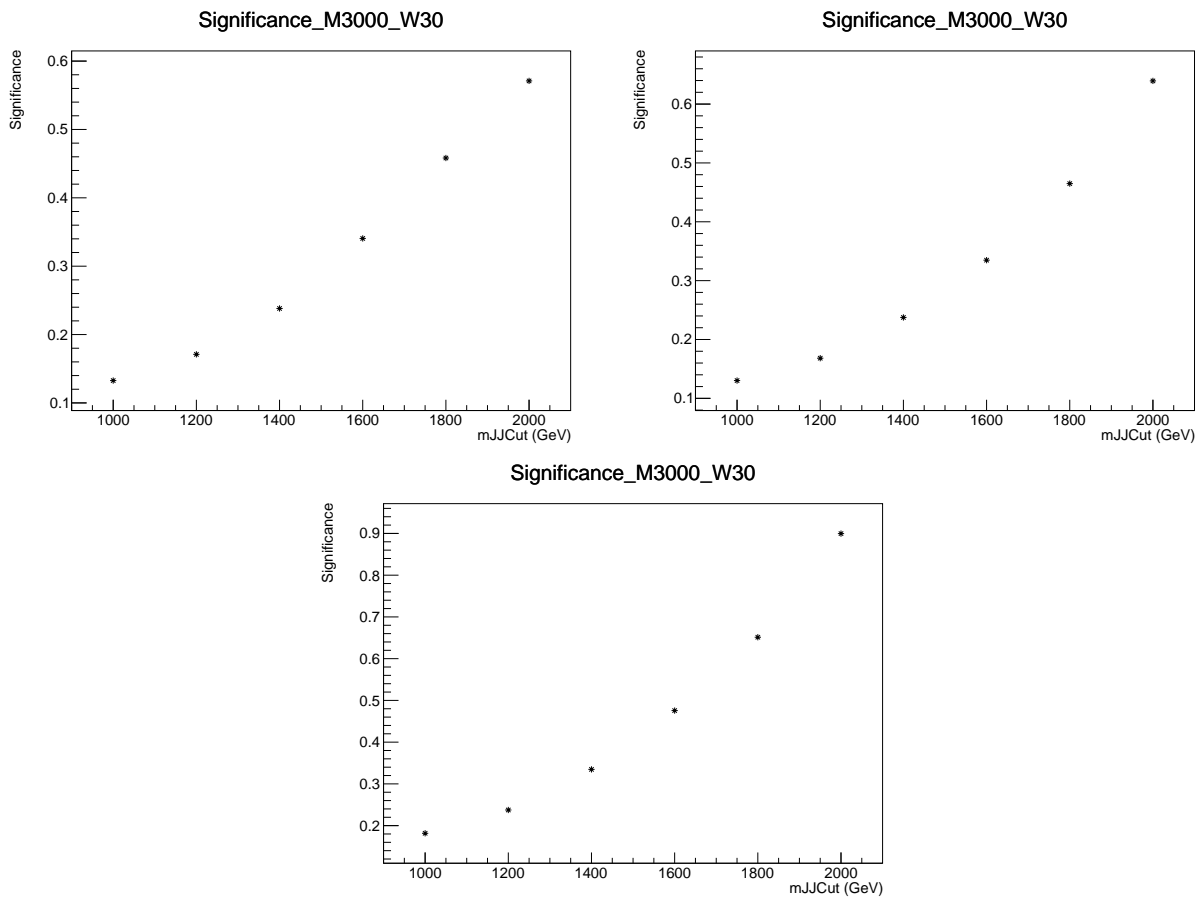


Figure 6.8: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 3TeV and 1 % width.

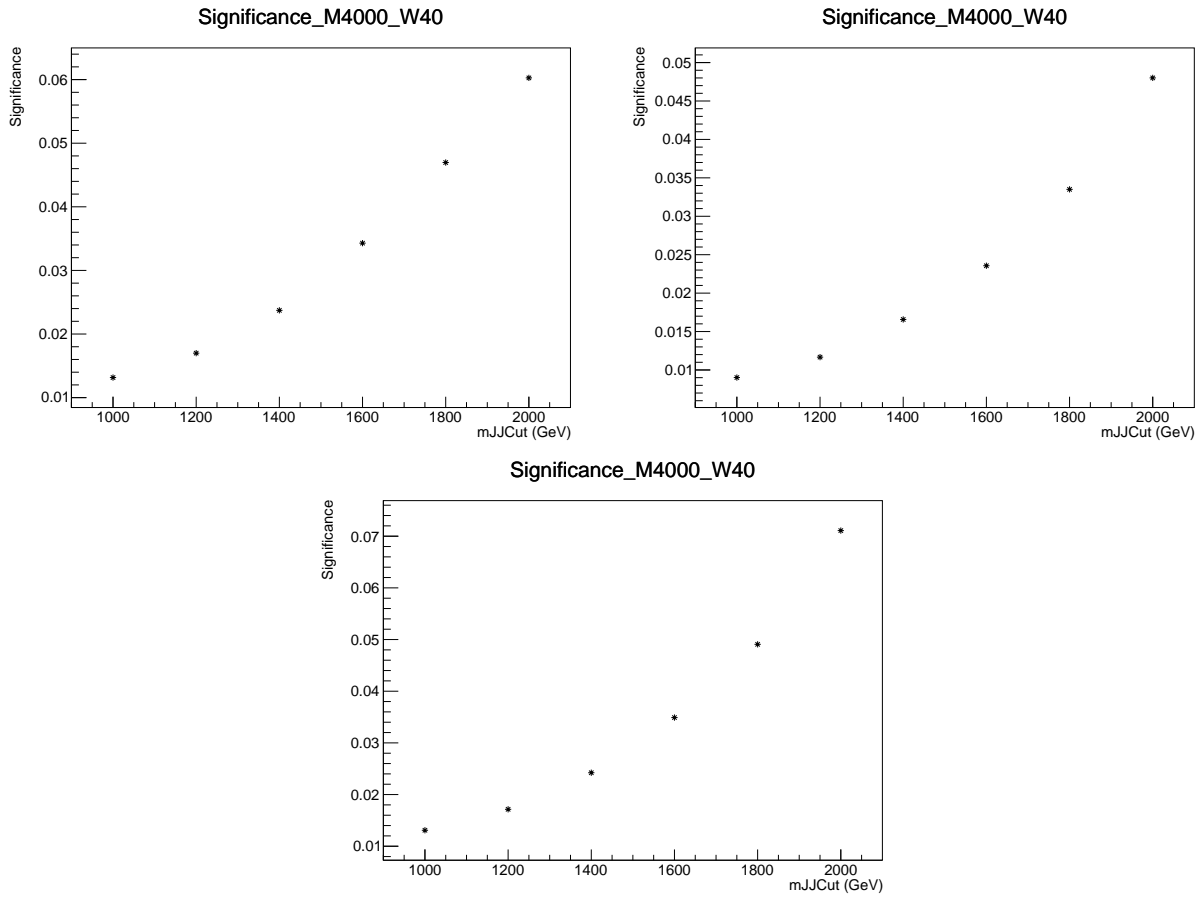


Figure 6.9: Significance for all years (top left 2016 preVFP, 2017 bottom left and 2018 bottom right) for the  $Z'$  with mass 4TeV and 1 % width.

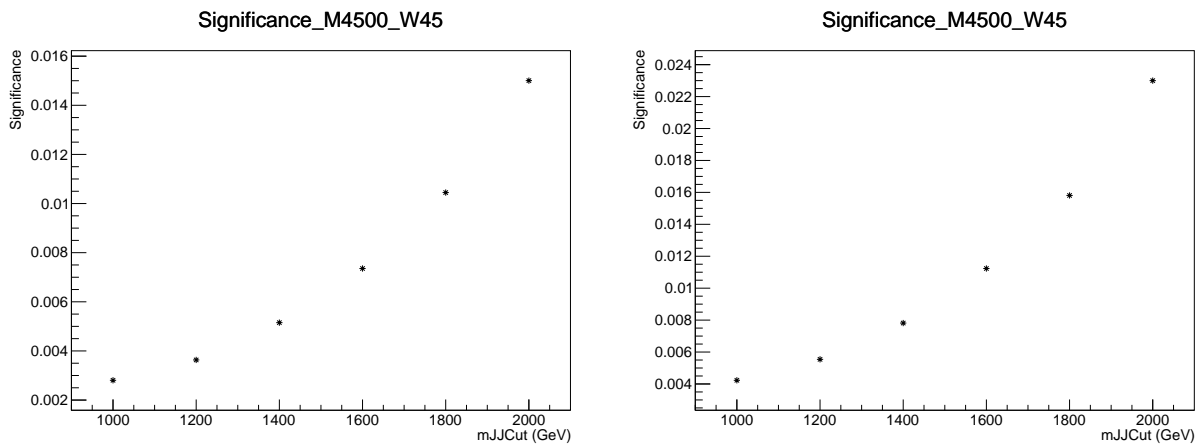


Figure 6.10: Significance for all years (2017 left and 2018 right) for the  $Z'$  with mass 4.5TeV and 1 % width.

## 6.4 Sensitivity Analysis

In this section, a sensitivity analysis is performed in order to detect shape differences between the  $t\bar{t}$  and the  $Z'$   $\chi$  distributions. Both files are derived from MC generated events.

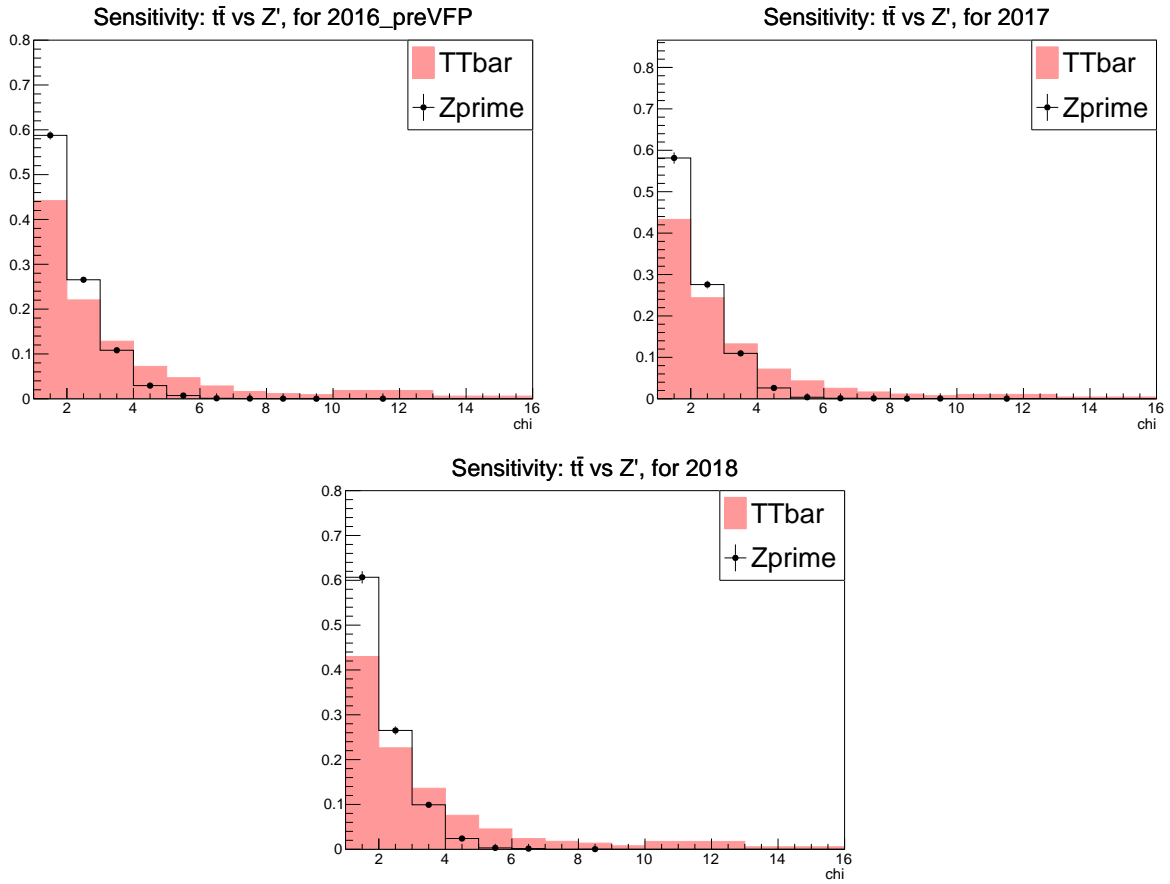


Figure 6.11: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.2TeV and 1 % width.



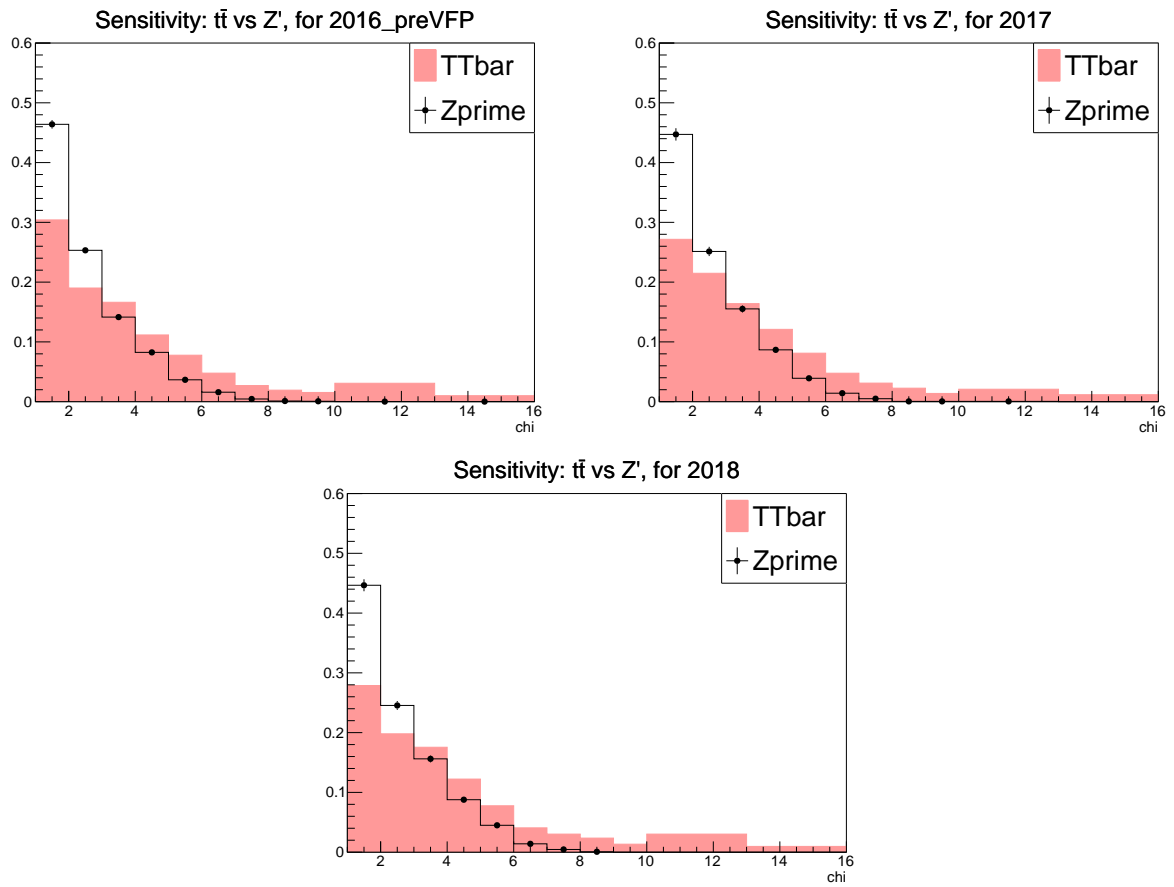


Figure 6.12: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.4TeV and 1 % width.

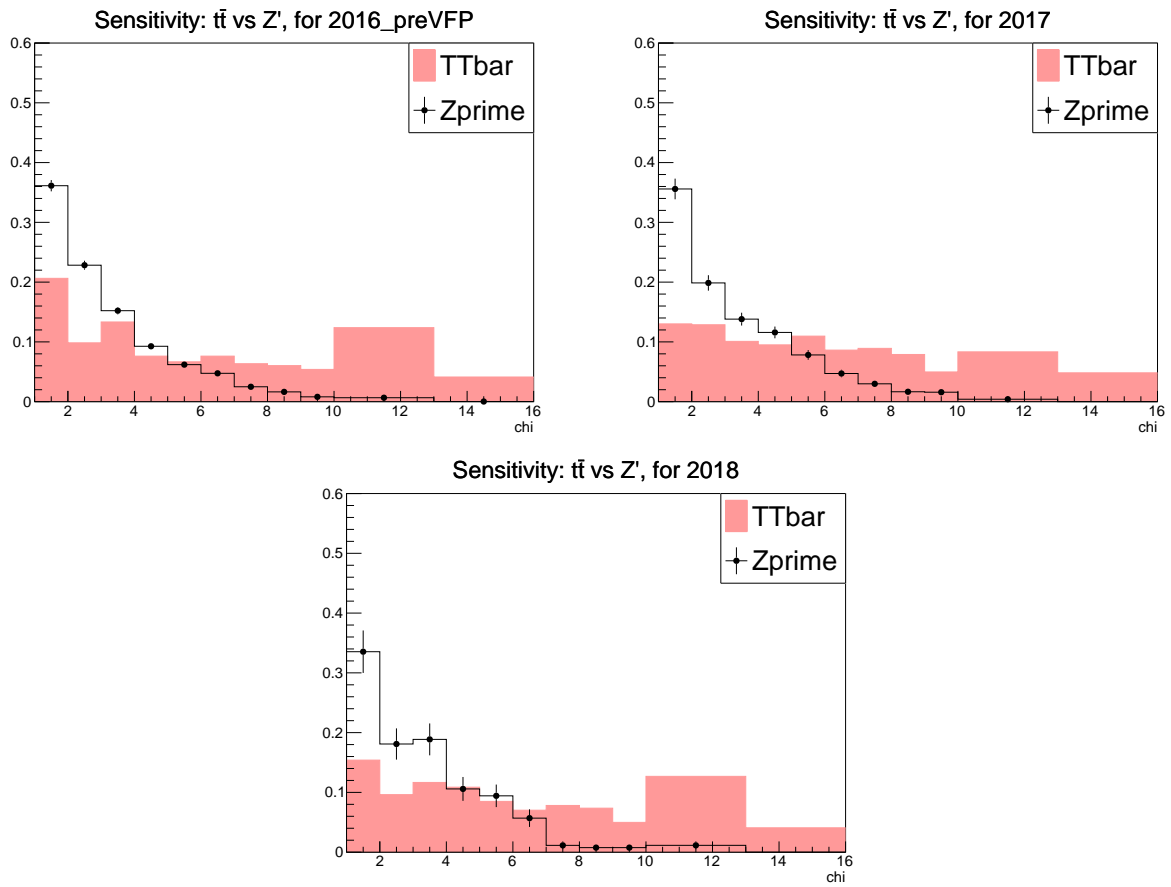


Figure 6.13: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.6TeV and 1 % width.

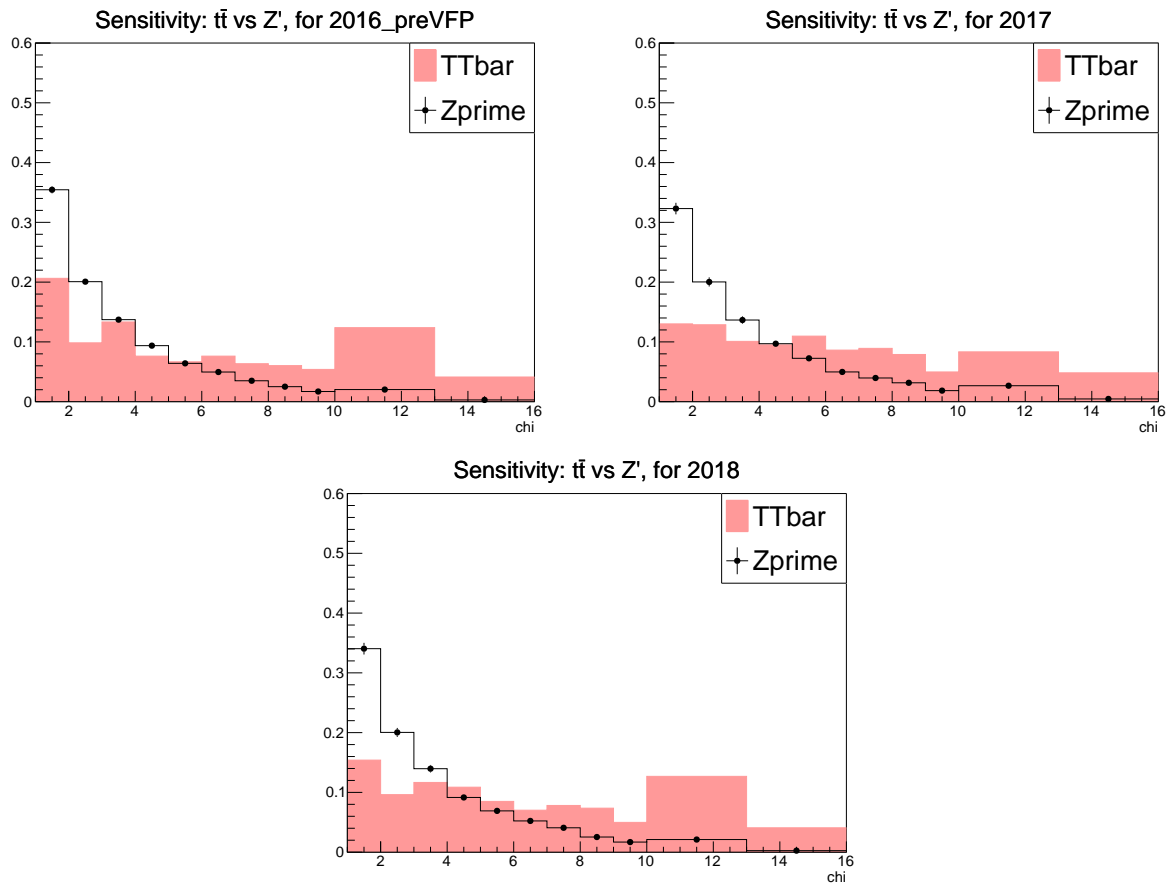


Figure 6.14: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.8TeV and 1 % width.

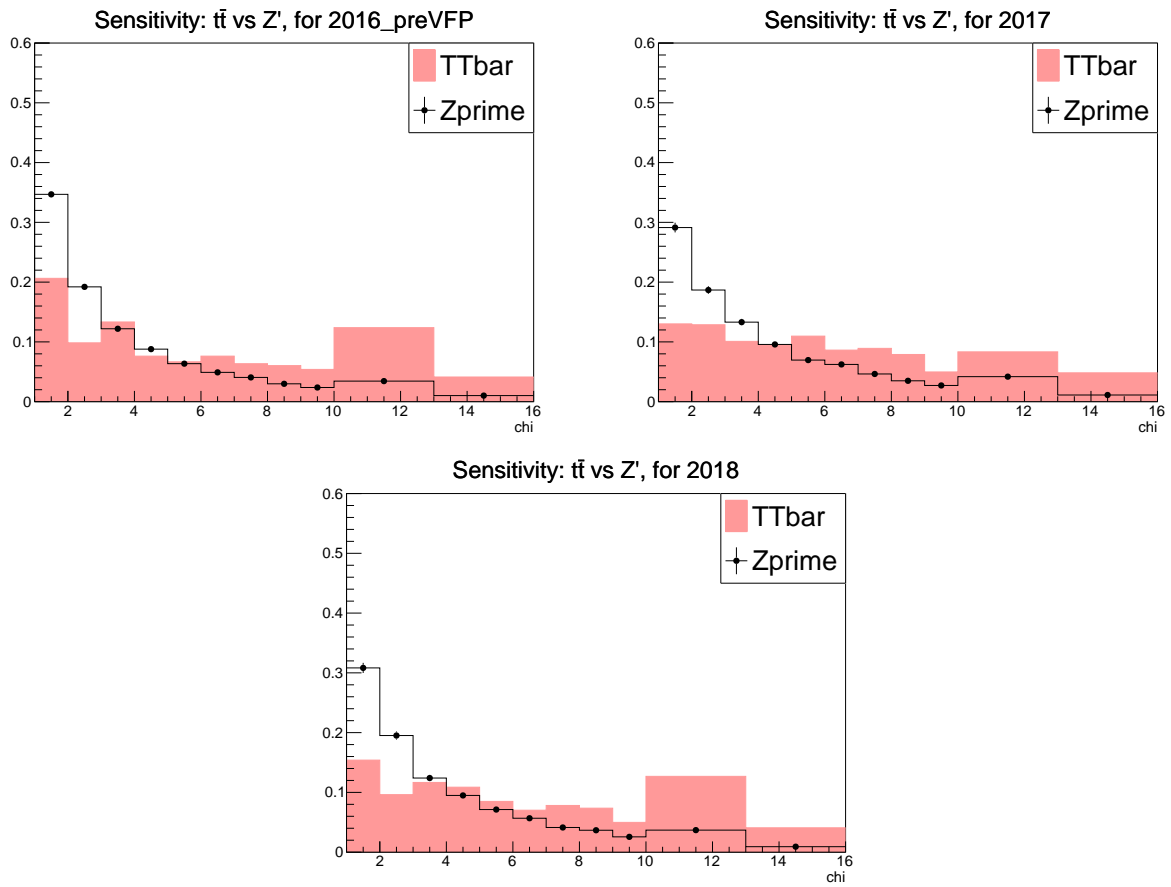


Figure 6.15: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 2TeV and 1 % width.

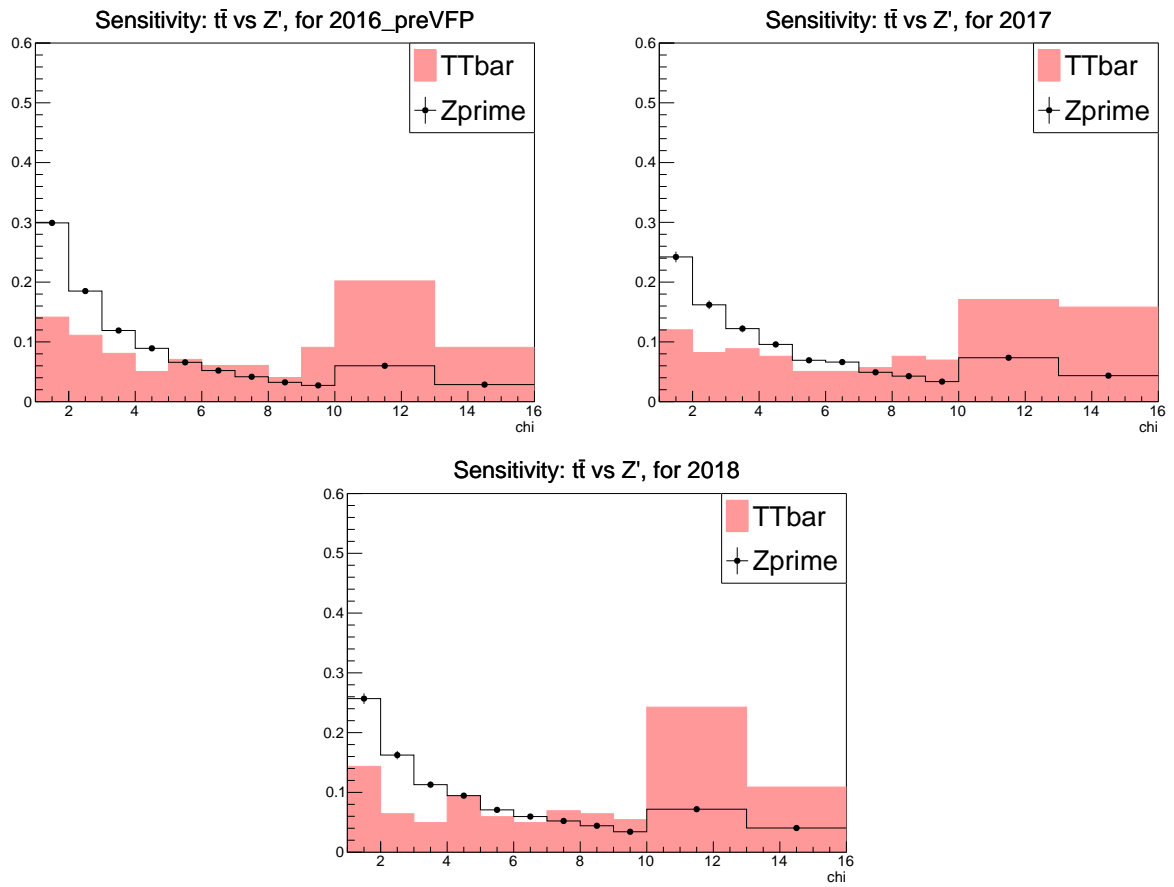


Figure 6.16: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 2.5TeV and 1 % width.

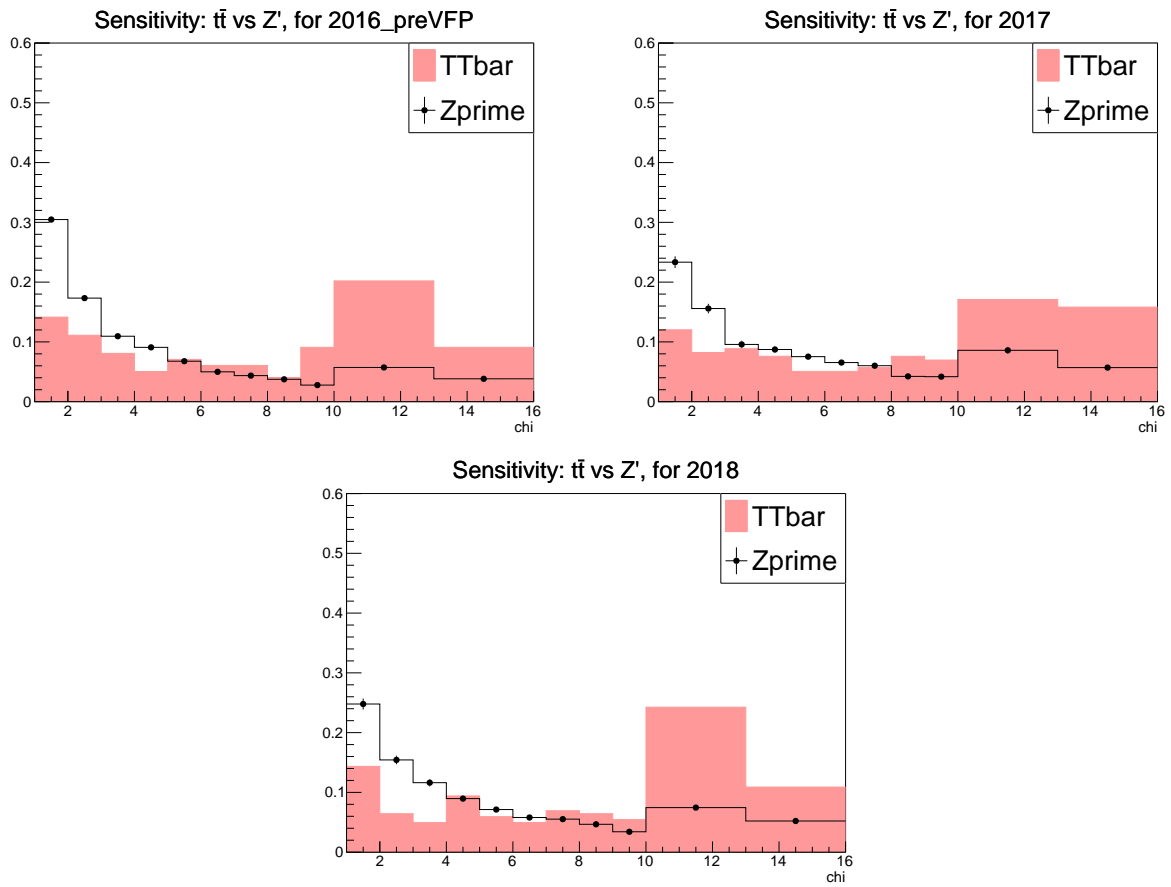


Figure 6.17: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 3TeV and 1 % width.

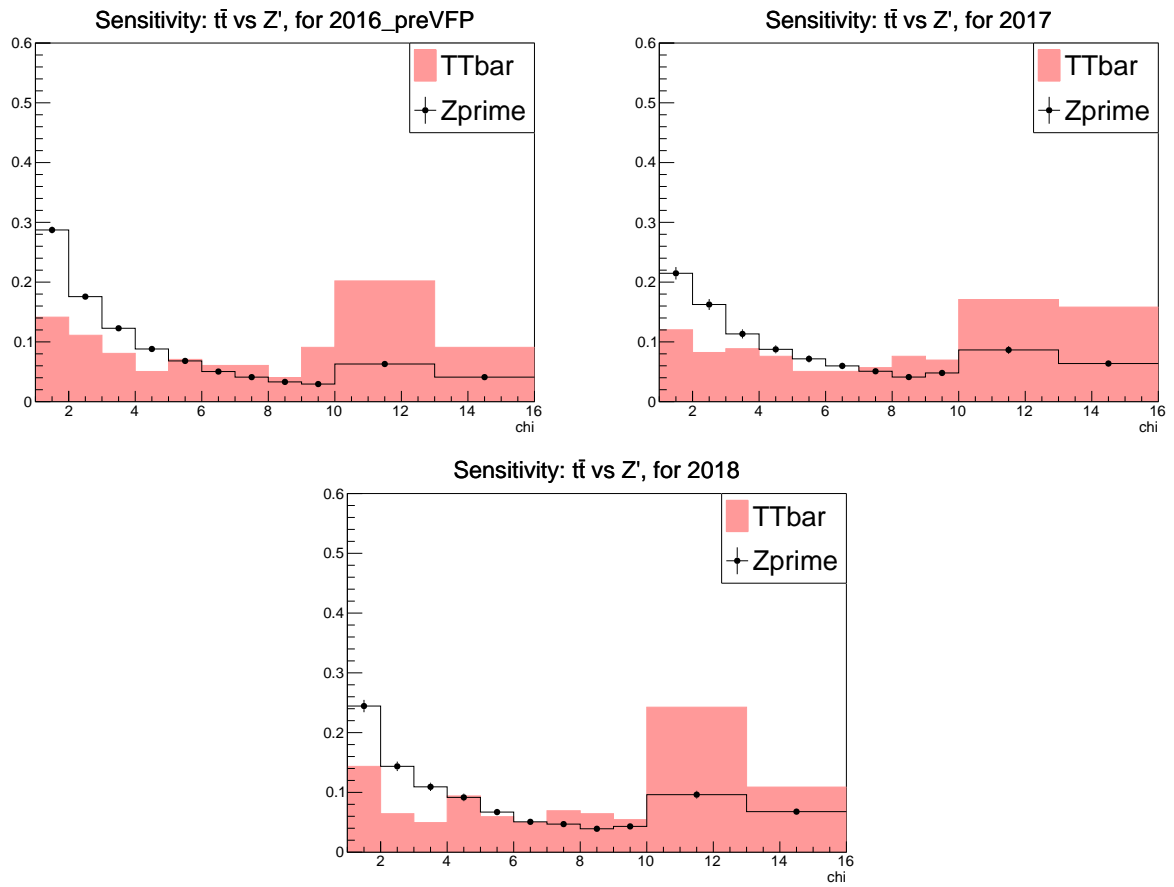


Figure 6.18: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 3.5TeV and 1 % width.

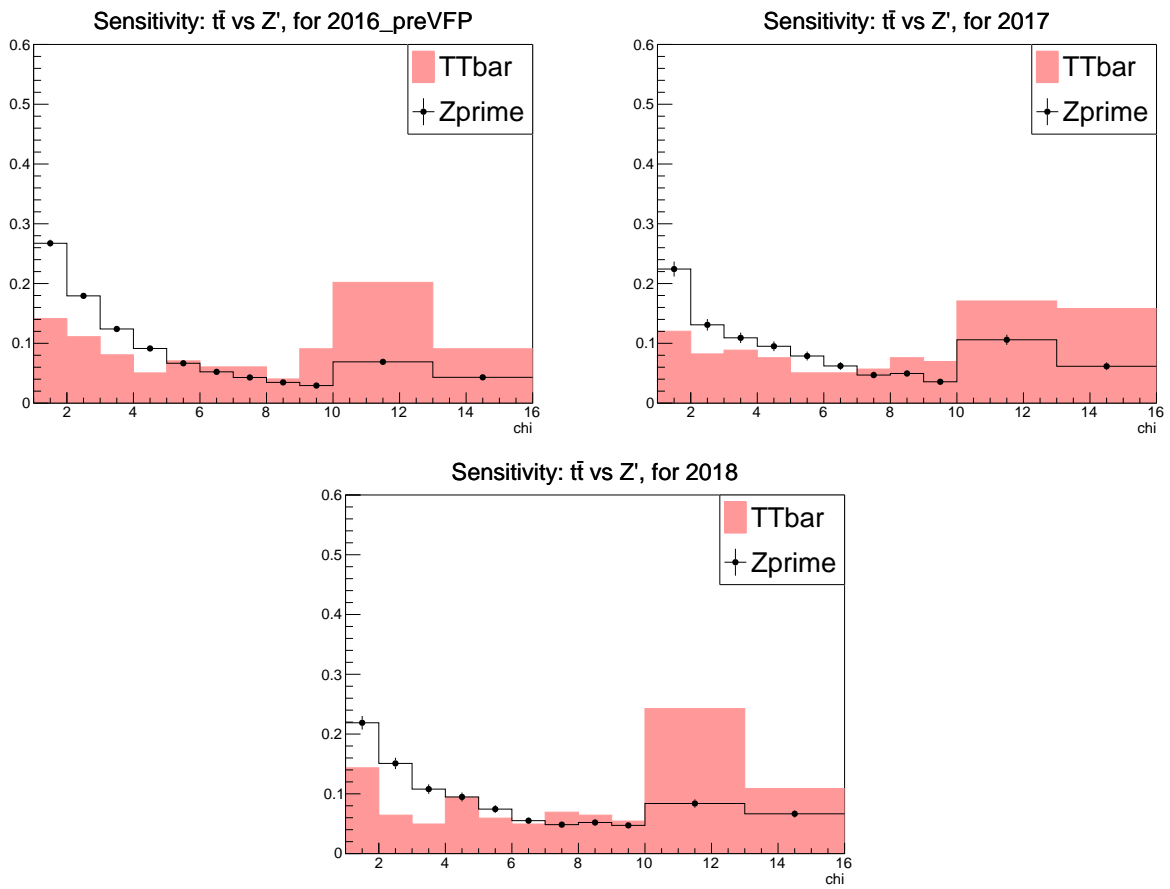


Figure 6.19: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 4TeV and 1 % width.

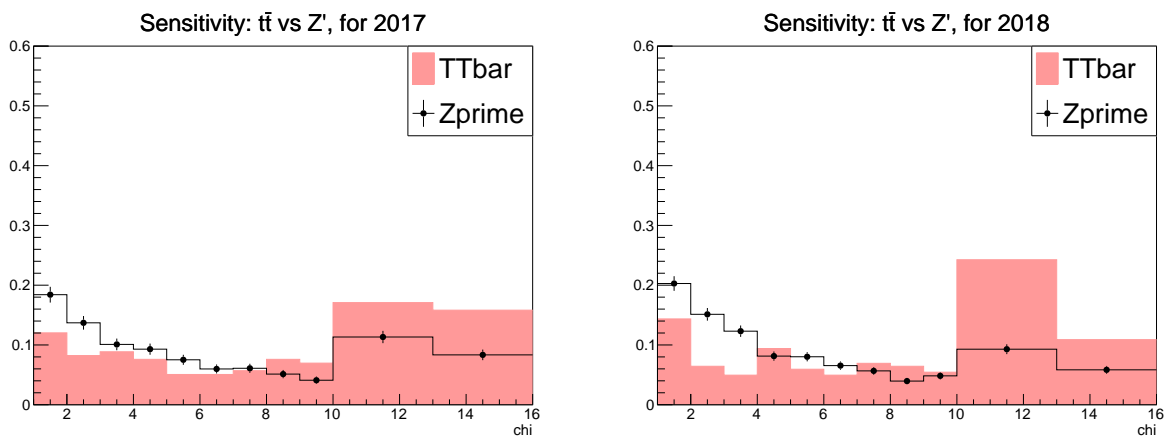


Figure 6.20: Sensitivity for all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 4.5TeV and 1 % width.



## 6.5 Signal Discrimination

Figures 6.21-6.30 show the distributions for the  $\chi$  variable in the all-hadronic channel. In these plots, the  $t\bar{t}$  signal is scaled to the signal strength obtained in the previous section 6.3 while the QCD is scaled to the data. The independent contributions seem to agree with the data, thus not needing any scaling. The plots are also shown in the previous chapters based on the all hadronic  $t\bar{t}$  analysis. The  $\chi$  distributions show the  $Z'$  contribution interpolated in the stack distribution.

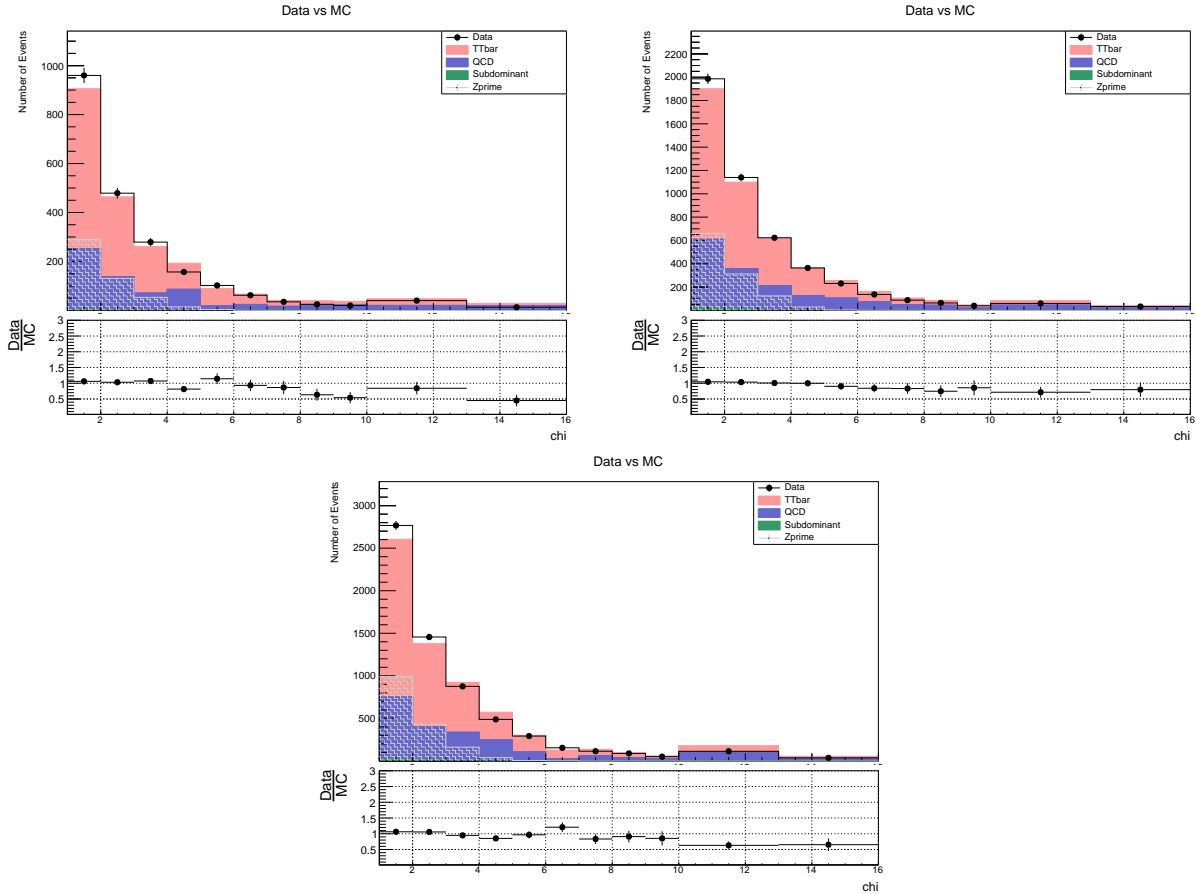


Figure 6.21: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.2TeV and 1 % width.

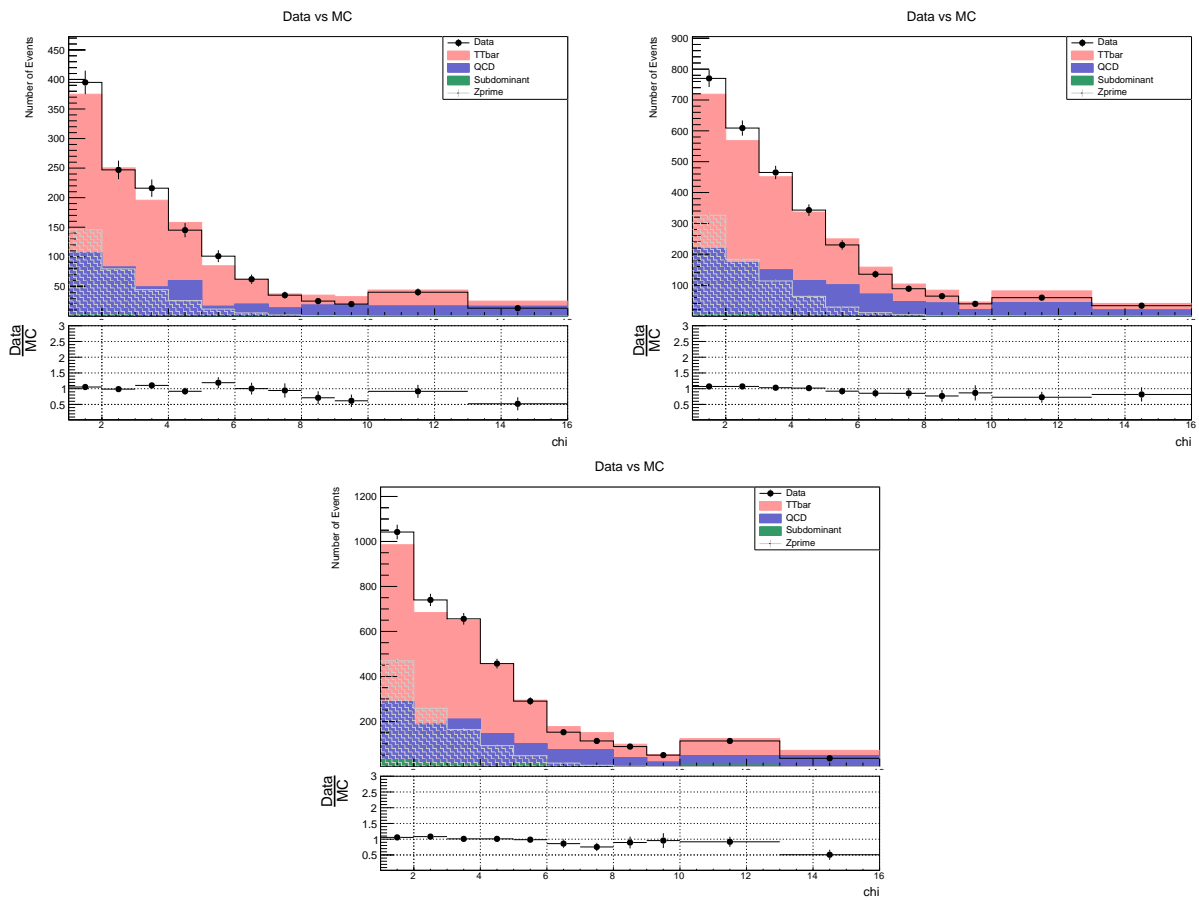


Figure 6.22: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.4 TeV and 1 % width.

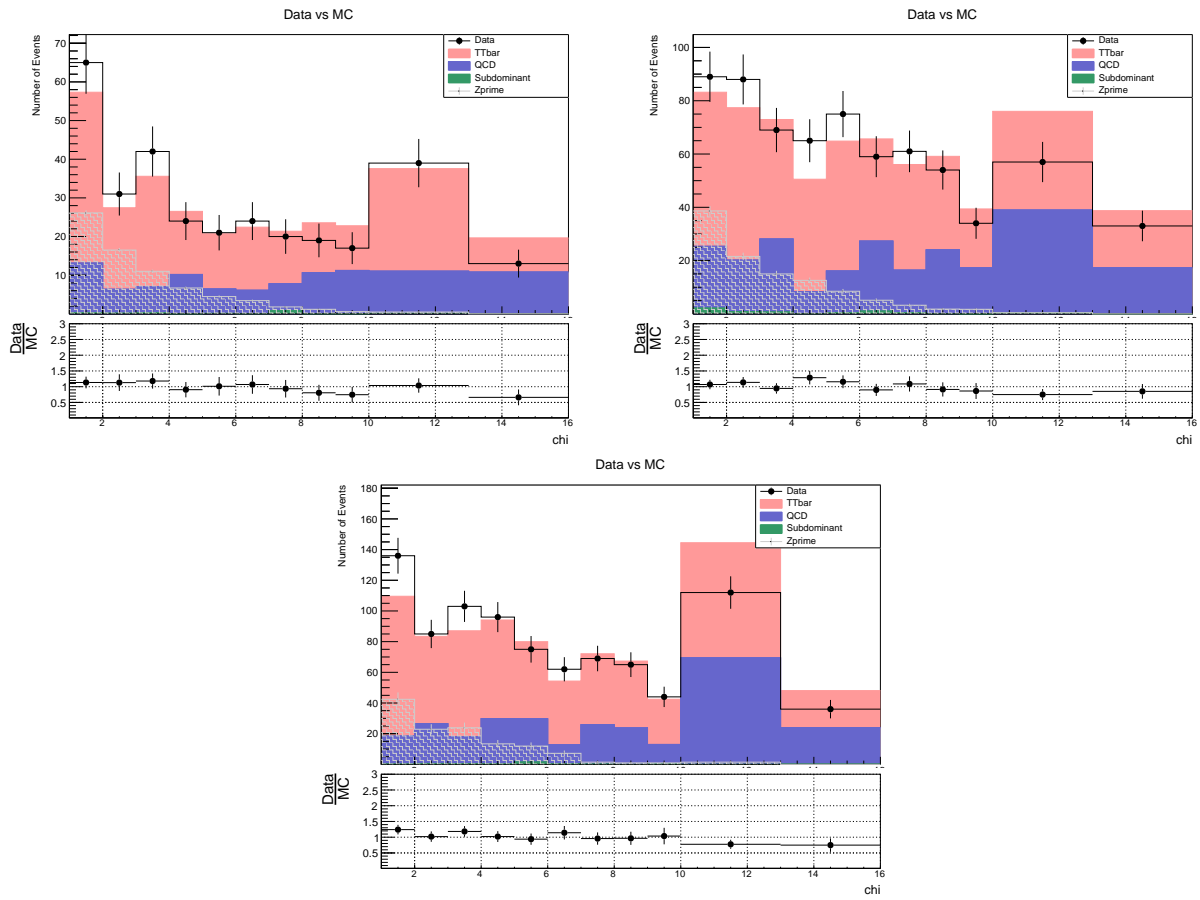


Figure 6.23: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.6TeV and 1 % width.

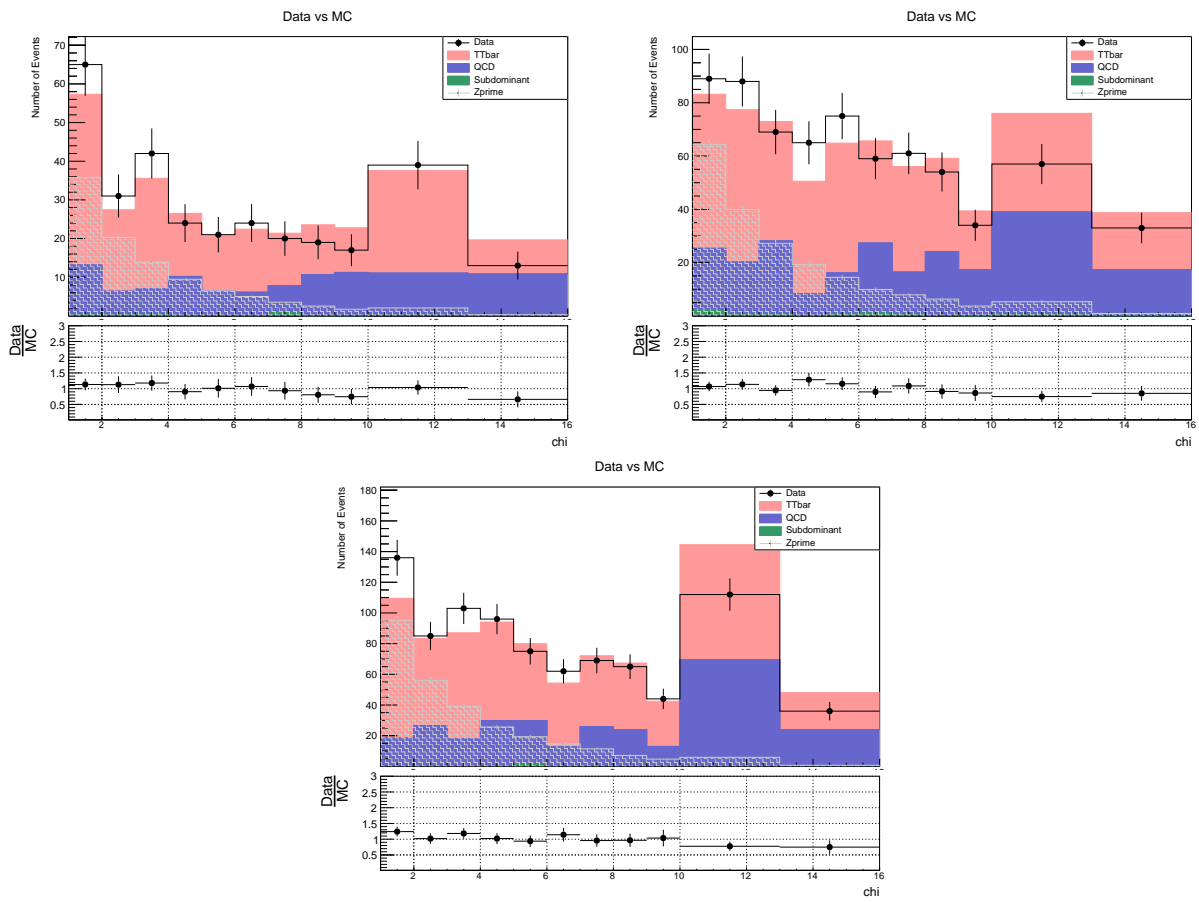


Figure 6.24: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 1.8 TeV and 1 % width.

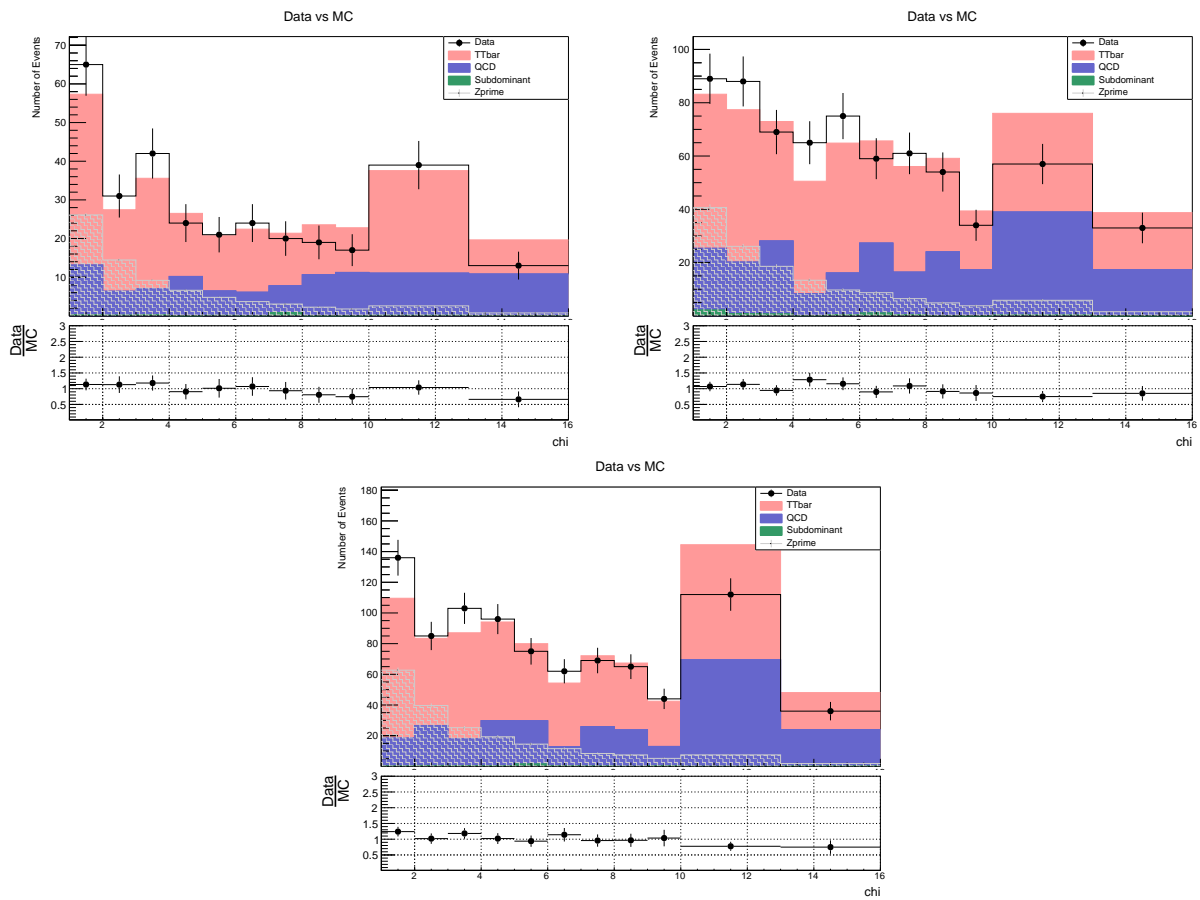


Figure 6.25: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass  $2\text{TeV}$  and  $1\%$  width.

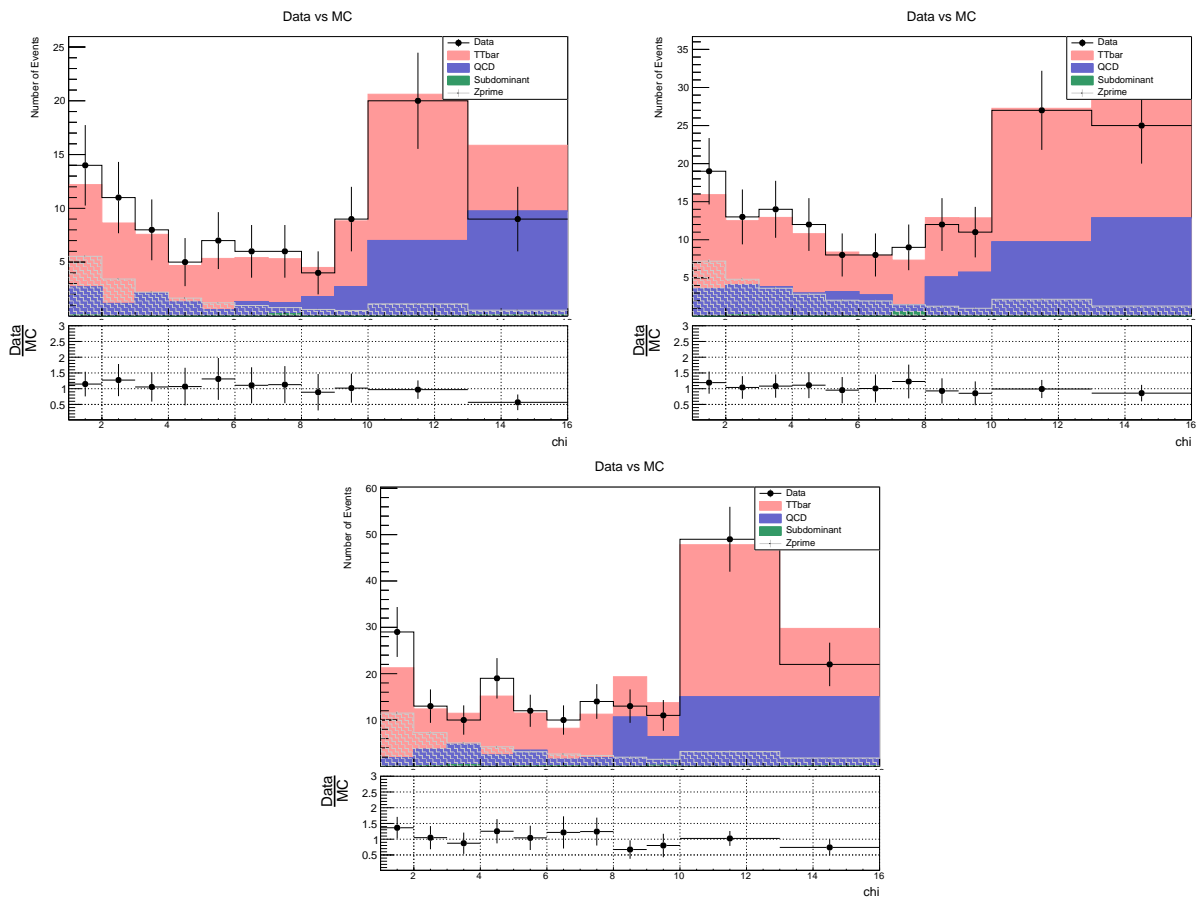


Figure 6.26: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 2.5 TeV and 1 % width.

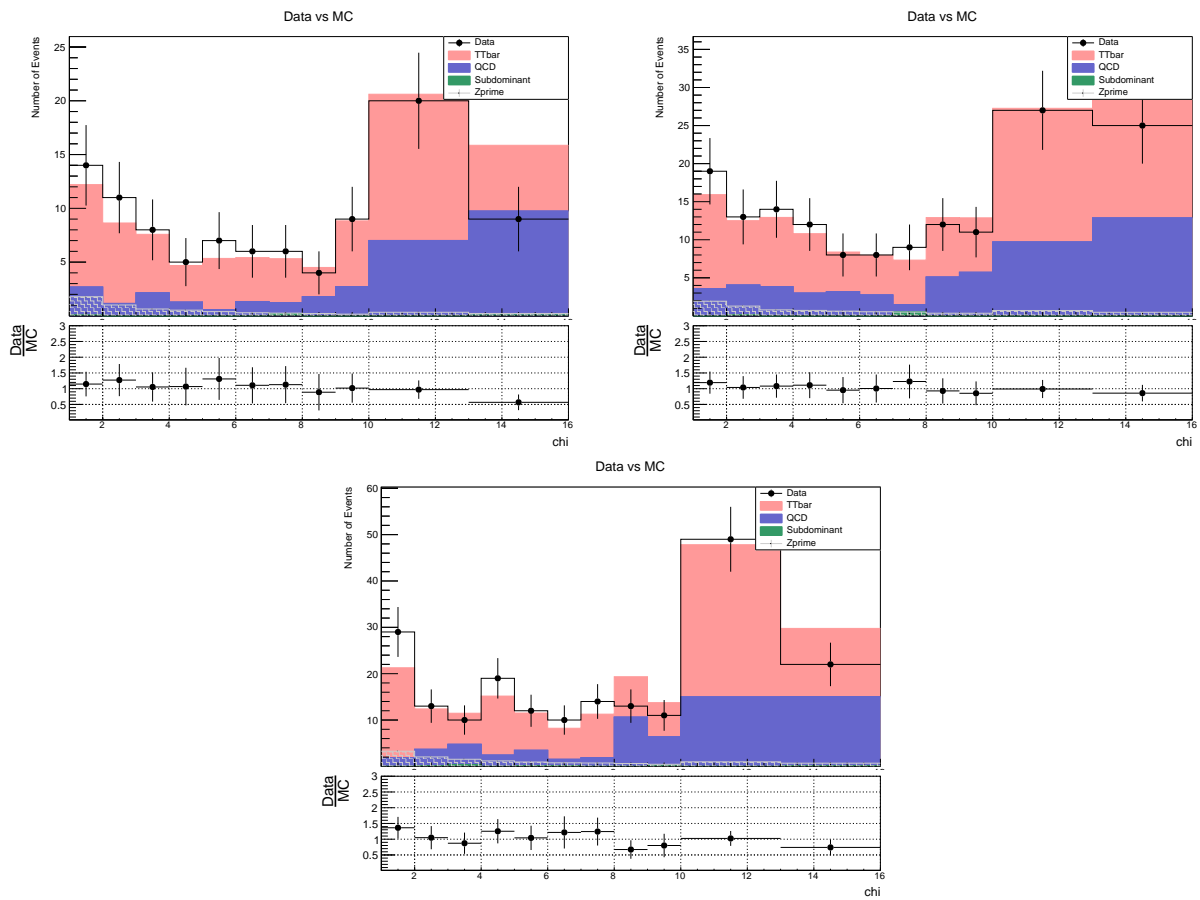


Figure 6.27: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass  $3\text{TeV}$  and  $1\%$  width.

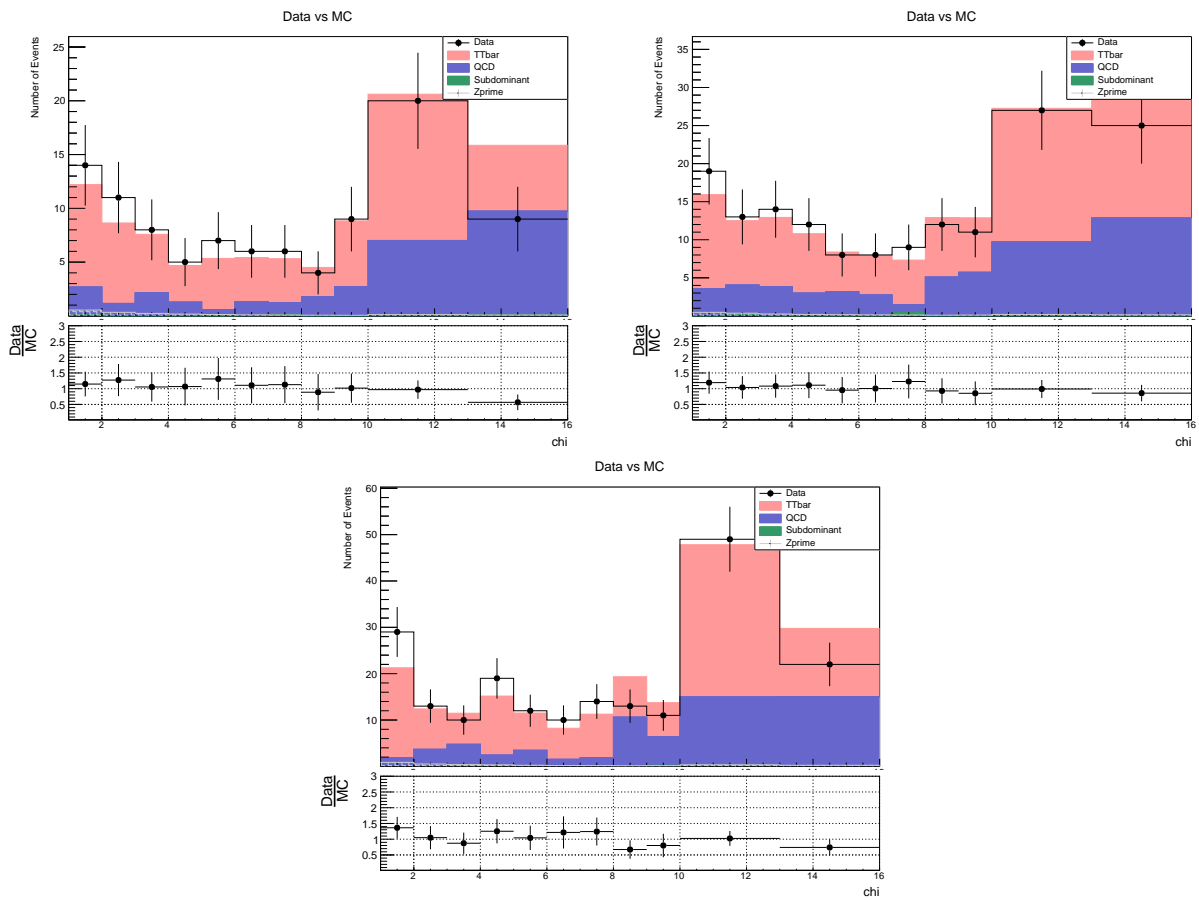


Figure 6.28: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 3.5 TeV and 1 % width.



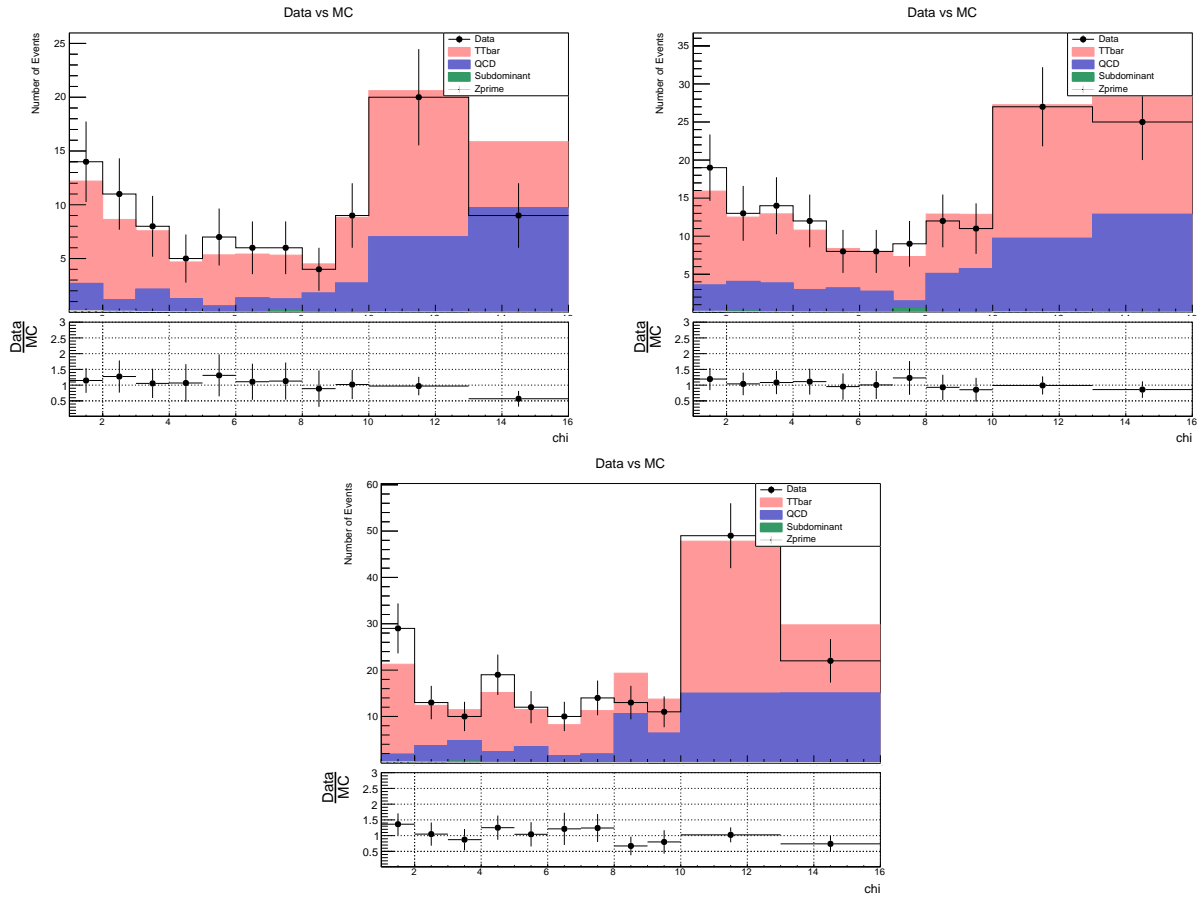


Figure 6.29: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 4 TeV and 1 % width.

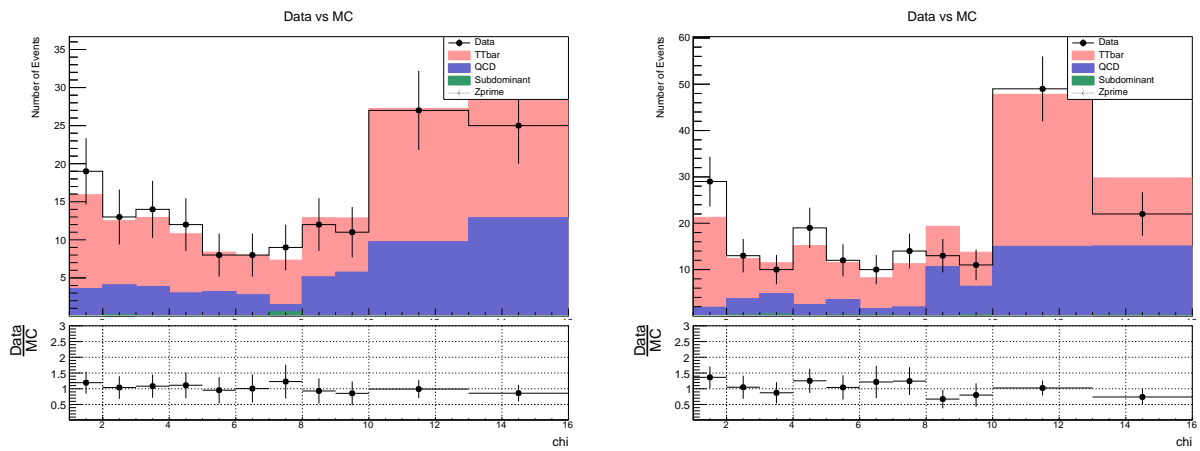


Figure 6.30: Data vs MC stack plots all years (top left 2016 preVFP, top right 2017 and bottom 2018) for the  $Z'$  with mass 4.5 TeV and 1 % width.

## 6.6 Results

In this section, the combined results are provided. The y-axis represents confidence levels for the value of the ratio of the measured cross section over the standard model predicted value. The plots in the y axis are representing confidence levels for the value of the ratio of measured cross section over the standard model predictions.

The Asymptotic Limits method allows to compute quickly an estimate of the observed and expected limits, which is fairly accurate when the event yields are not too small and the systematic uncertainties don't play a major role in the result. The limit calculation relies on an asymptotic approximation of the distributions of the LHC test-statistic, which is based on a profile likelihood ratio, under signal and background hypotheses to compute two p-values  $p_\mu, p_b$  and thus  $CL_s = \frac{p_\mu}{(1-p_b)}$ , where  $CL_s$  is the asymptotic approximation of computing limits with frequentist toys.

For this analysis, the *Combine* Statistical Tool [130] provided by the CMS Statistics Committee is used. This framework provides the user with the ability to run an analysis on Asymptotic limits. The program prints out the limit on the signal strength  $\mathbf{r}$  defined as (number of signal events / number of expected signal events) e.g. Observed Limit:  $r < 1.6297$  @ 95% CL, the median expected limit Expected 50.0%:  $r < 2.3111$  and edges of the 68% and 95% ranges for the expected limits. By default, the limits are calculated using the CLs prescription, as noted in the output, which takes the ratio of p-values under the signal plus background and background only hypothesis.

The analysis is proceeded as a first step per year. For each  $Z'$  mass variation a *datacard* is composed. The datacard includes all the information that the Combine tool requires for the analysis such as a Data file, Signal files, background files as well as variations. In this analysis, the Signal file is the  $Z'$  MC generated events, while  $t\bar{t}$ , QCD and subdominant (Single top, etc) are defined as background. As for the variations, only the  $t\bar{t}$  variations are taken into account while they compose the largest background in this phase space.

A typical data card is shown in Figure 6.31:

```

1 Combination of ../2016_preVFP_test/datacard_chi_SR_mZ_1200_12_cut_1000.txt ../2017_test/datacard_chi_SR_mZ_1200_12_cut_1000.txt ../2018_test/datacard_chi_SR_mZ_1200_12_cut_1000.txt
2 lmax 3 number of bins
3 jmax 3 number of processes minus 1
4 kmax 11 number of nuisance parameters
5
6 shapes * ch1 ../2016_preVFP_test/ProcessesFile_1000.root_h_chi_$PROCESS_h_chi_$PROCESS_$SYSTEMATIC
7 shapes Zprime ch1 ../2016_preVFP_test/ZprimeFile_1200_12_massCut1000.root_h_chi_$PROCESS
8 shapes data_obs ch1 ../2016_preVFP_test/DataFile_1000.root_h_Data
9 shapes * ch2 ../2017_test/ProcessesFile_1000.root_h_chi_$PROCESS_h_chi_$PROCESS_$SYSTEMATIC
10 shapes Zprime ch2 ../2017_test/ZprimeFile_1200_12_massCut1000.root_h_chi_$PROCESS
11 shapes data_obs ch2 ../2017_test/DataFile_1000.root_h_Data
12 shapes * ch3 ../2018_test/ProcessesFile_1000.root_h_chi_$PROCESS_h_chi_$PROCESS_$SYSTEMATIC
13 shapes Zprime ch3 ../2018_test/ZprimeFile_1200_12_massCut1000.root_h_chi_$PROCESS
14 shapes data_obs ch3 ../2018_test/DataFile_1000.root_h_Data
15
16 bin ch1 ch2 ch3
17 observation -1 -1 -1
18
19 bin ch1 ch1 ch1 ch1 ch2 ch2 ch2 ch2 ch3 ch3 ch3 ch3
20 process Zprime ttbar Subdominant qcd Zprime ttbar Subdominant qcd Zprime ttbar Subdominant qcd
21 process 0 1 2 3 0 1 2 3 0 1 2 3
22 rate -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
23
24 btag shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
25 fsr shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
26 isr shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
27 lumi_13TeV lnN - 1.025 1.025 1.025 - 1.025 1.025 - 1.025 1.025
28 pdf shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
29 scale shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
30 shifted shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
31 smeared shapeN - 1.0 - - - 1.0 - - - 1.0 - - -
32 yield_Subdominant lnN - - 1.5 - - - 1.5 - - - 1.5 - - -
33 yield_qcd lnN - - - 1.5 - - - 1.5 - - - 1.5 - - -
34 yield_ttbar lnN - 1.5 - - - 1.5 - - - 1.5 - - -
35 ch1 autoMCStats 10 1 1
36 ch2 autoMCStats 10 1 1
37 ch3 autoMCStats 10 1 1

```

Figure 6.31: Datacard used for the combination of the 3 analyses.

For  $Z'$  mass variation the analysis is performed based on the selected mapped mass seen in 6.8. All of the results are combined into a graph. Each graph includes the observed and expected upper limits including systematic variations.

In order to analyze the outcomes of multiple searches together, it is necessary to combine their datacards into a single one. This is done by the *combineCards.py* script. The combined Brazilian plots are shown in 6.36

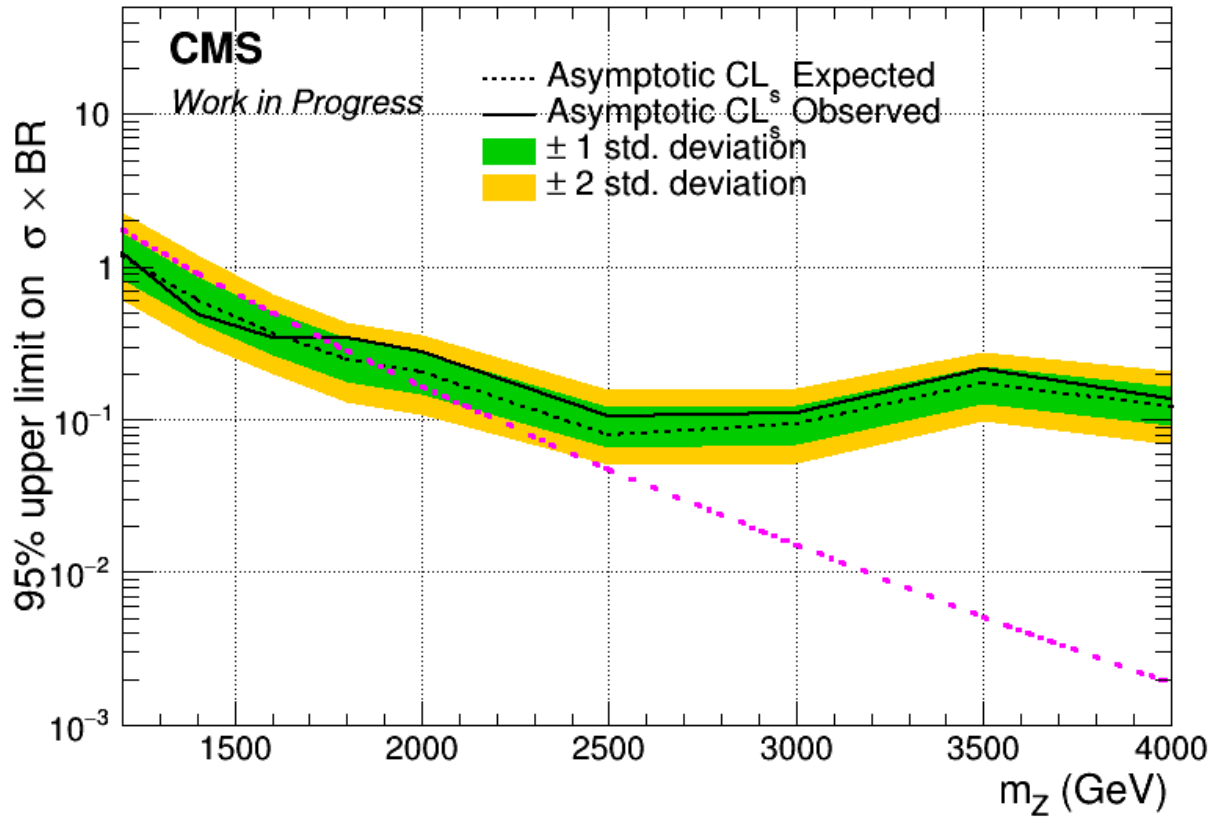


Figure 6.32: Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a  $Z'$  boson with  $\Gamma/M$  of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for 2016 preVFP.

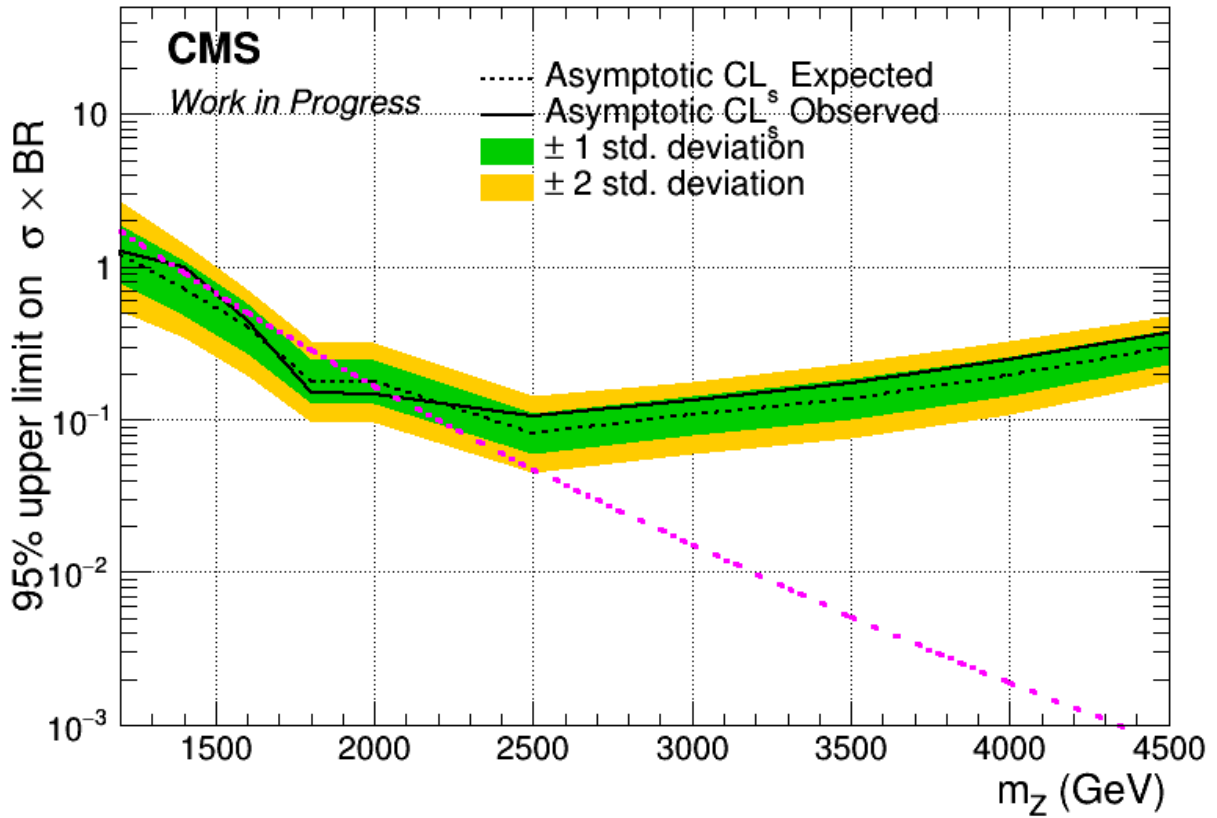


Figure 6.33: Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a  $Z'$  boson with  $\Gamma/M$  of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for 2017.

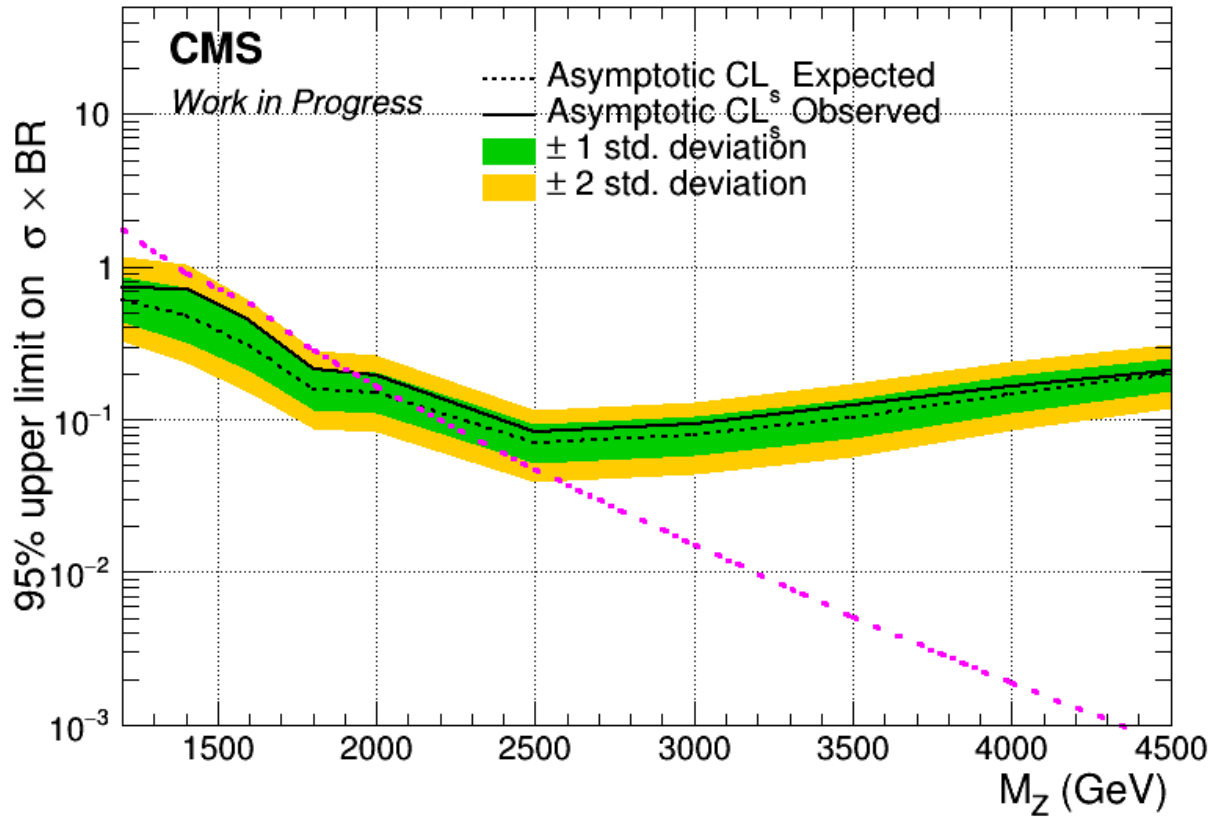


Figure 6.34: Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a  $Z'$  boson with  $\Gamma/M$  of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for 2018.

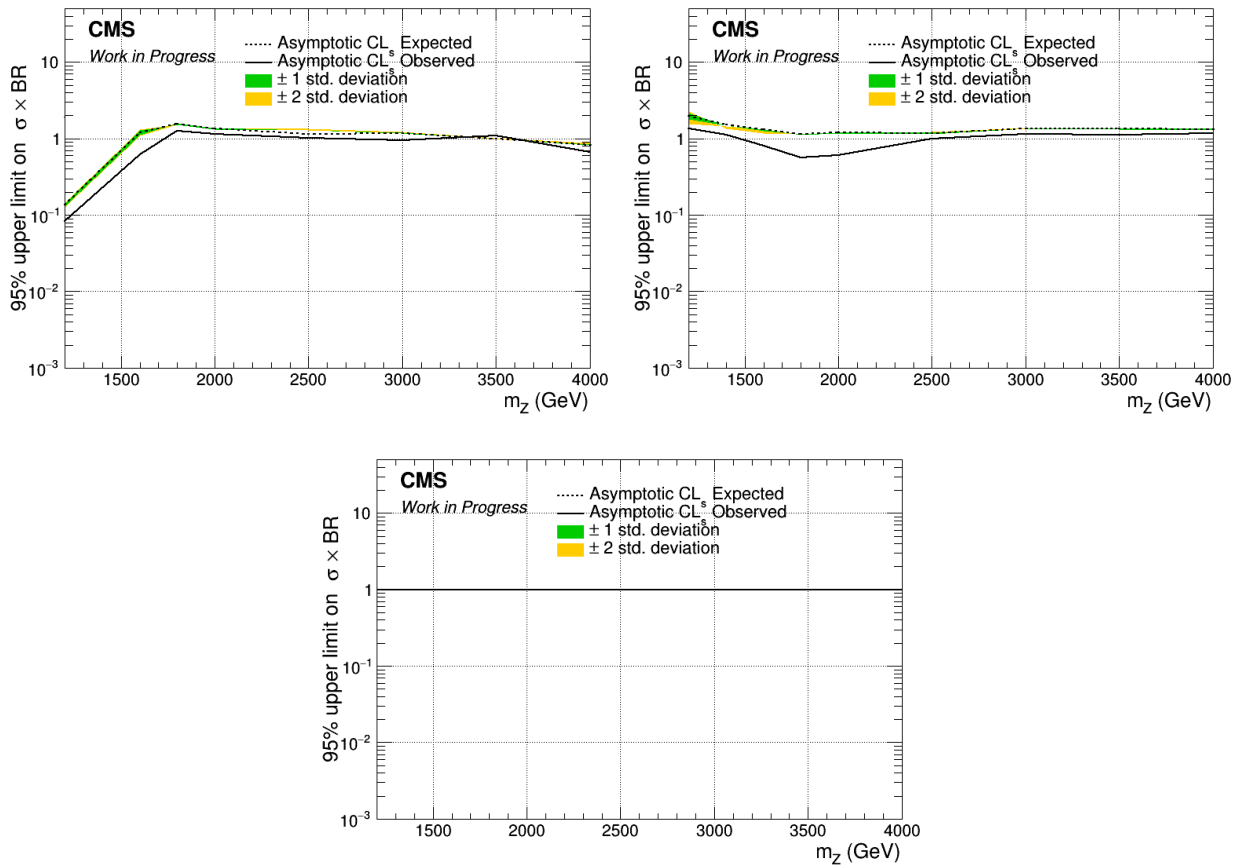


Figure 6.35: Yearly Asymptotic limits results with respect to 2018 results.

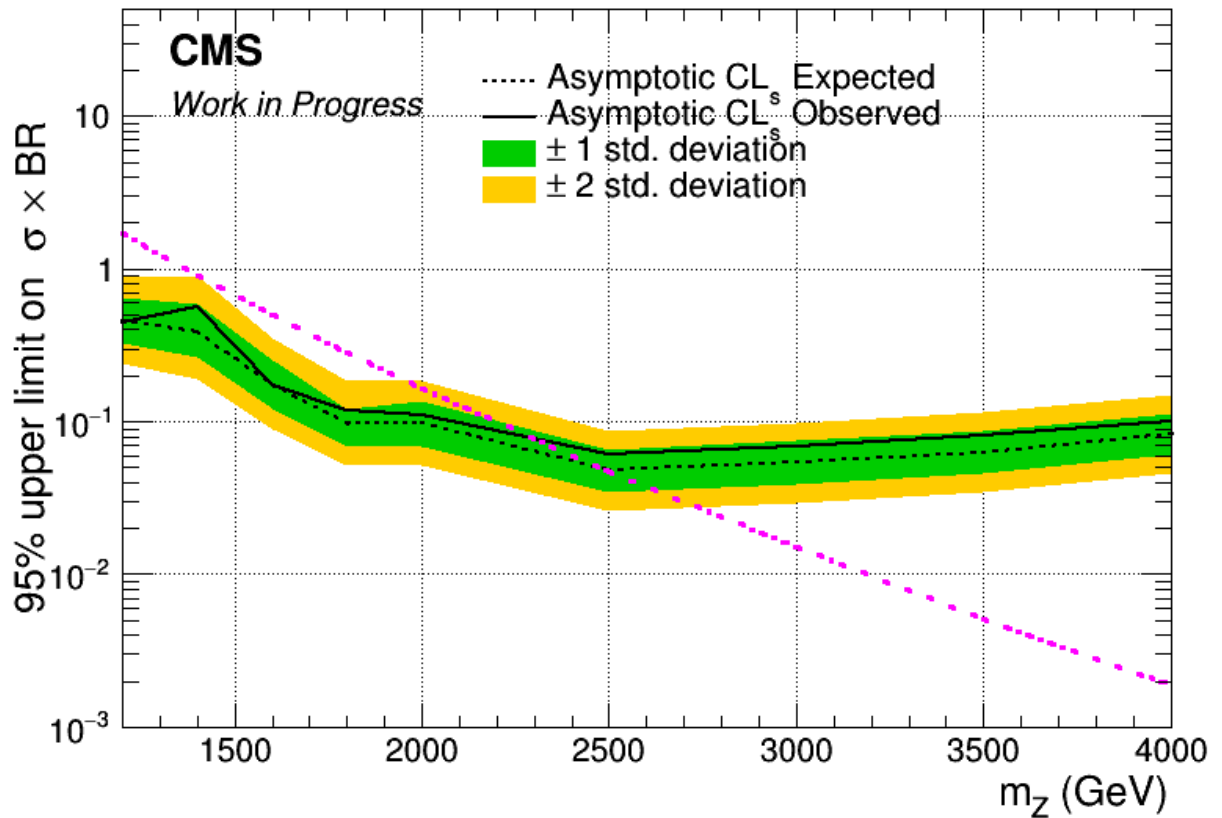


Figure 6.36: Observed and expected upper limits at 95% CL on the product of the production cross section and branching fractions for the full combination of the analysis results, shown as function of the resonance mass. Limits are set using one extensions to the SM a  $Z'$  boson with  $\Gamma/M$  of 1%. The corresponding theoretical prediction as a function of the resonance mass is shown as a dot-dashed curve for the combined 2016 preVPF, 2017 and 2018 analyses.

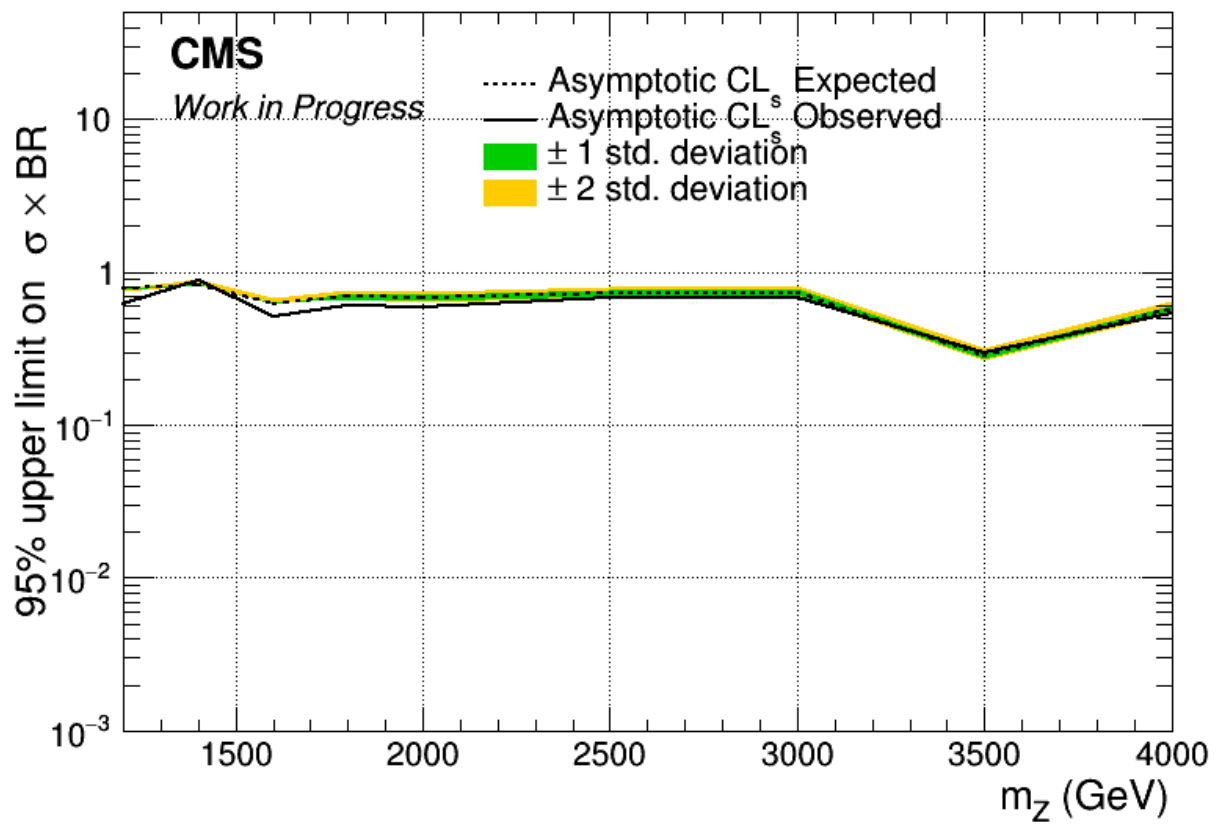


Figure 6.37: Brazilian Plots for combined years with respect to 2018 results.



Table 6.9: Expected and observed cross section limits at 95% CL, for the 1% width  $Z'$  resonance hypothesis for the 2016 preVFP analysis.

Mass [TeV]	Observed [pb]	Expected [pb]				
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
1.2	1.23	0.609	0.821	1.16	1.66	2.29
1.4	0.485	0.321	0.429	0.609	0.869	1.2
1.6	0.346	0.197	0.264	0.368	0.511	0.666
1.8	0.34	0.130	0.175	0.245	0.339	0.433
2.0	0.278	0.108	0.145	0.203	0.28	0.359
2.5	0.106	0.05	0.0659	0.0795	0.12	0.156
3.0	0.109	0.052	0.068	0.0923	0.125	0.161
3.5	0.212	0.0974	0.129	0.171	0.225	0.279
4.0	0.137	0.0692	0.0907	0.122	0.164	0.21

Table 6.10: Expected and observed cross section limits at 95% CL, for the 1% width  $Z'$  resonance hypothesis for the 2017 analysis.

Mass [TeV]	Observed [pb]	Expected [pb]				
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
1.2	1.26	0.518	0.791	1.19	1.87	2.66
1.4	0.993	0.339	0.475	0.7178	1.08	1.41
1.6	0.446	0.197	0.272	0.403	0.574	0.714
1.8	0.151	0.0957	0.128	0.178	0.248	0.329
2.0	0.147	0.0962	0.128	0.177	0.244	0.322
2.5	0.105	0.0444	0.0591	0.0813	0.111	0.142
3.0	0.132	0.0593	0.0787	0.107	0.141	0.179
3.5	0.172	0.0747	0.1	0.137	0.184	0.234
4.0	0.244	0.109	0.144	0.195	0.259	0.328
4.5	0.366	0.175	0.228	0.301	0.388	0.477

Table 6.11: Expected and observed cross section limits at 95% CL, for the 1% width  $Z'$  resonance hypothesis for the 2018 analysis.

Mass [TeV]	Observed [pb]	Expected [pb]				
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
1.2	0.923	0.329	0.439	0.61.	0.858	1.17
1.4	0.881	0.234	0.322	0.474	0.726	1.04
1.6	0.553	0.15	0.207	0.305	0.453	0.603
1.8	0.267	0.0862	0.115	0.156	0.214	0.285
2.0	0.241	0.0825	0.109	0.149	0.203	0.266
2.5	0.105	0.0392	0.0514	0.07	0.0928	0.117
3.0	0.115	0.0443	0.0582	0.0787	0.104	0.13
3.5	0.155	0.0573	0.0753	0.102	0.136	0.172
4.0	0.207	0.0847	0.111	0.148	0.193	0.234
4.5	0.259	0.118	0.153	0.2	0.254.	0.308

Table 6.12: Expected and observed cross section limits at 95% CL, for the 1% width  $Z'$  resonance hypothesis for the combined (2016preVFP, 2017, 2018) analysis.

Mass [TeV]	Observed [pb]	Expected [pb]				
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$
1.2	0.57	0.249	0.337	0.475	0.675	0.927
1.4	0.784	0.193	0.266	0.393	0.609	0.942
1.6	0.283	0.0955	0.13	0.188	0.28	0.41
1.8	0.161	0.0569	0.0771	0.11	0.157	0.216
2.0	0.141	0.0527	0.0710	0.1	0.144	0.198
2.5	0.072	0.027	0.0362	0.0505	0.0706	0.0936
3.0.	0.0778	0.0309	0.0415	0.0581	0.0808	0.107
3.5	0.0461	0.0153	0.0204	0.0287	0.0407	0.0553
4.0	0.111	0.0456	0.0608	0.0846	0.117	0.155

## 6.7 Conclusions

A model-independent search for the production of heavy spin-1 resonances decaying into  $t\bar{t}$  final states has been conducted. The data correspond to an integrated luminosity of 120.3 fb<sup>-1</sup> collected with the CMS detector in proton-proton collisions at  $\sqrt{s} = 13$  TeV at the LHC.

The analysis is designed to have high sensitivity at resonance masses above 1 TeV, where final state decay products become collimated because of the large Lorentz boosts of the top quarks. The analysis method provides an in-situ measurement of the data-to-simulation scale factor for the normalization of the main backgrounds. No evidence for massive resonances that decay to  $t\bar{t}$  is found. Limits at 95% CL are set on the production cross section of new spin-1 particles decaying to  $t\bar{t}$  with relative decay widths that are either narrow or wide compared with the detector resolution.

In addition, limits are set on the production of particles in benchmark models beyond the standard model.  $Z'$  bosons with relative widths  $\Gamma/M$  of 1% are excluded for mass ranges of 1.0-4.0 TeV, respectively. This search presents limits on  $Z'$  bosons as a function of the relative width of the resonance in the range from 1%.



## **Part IV**

# **Ελληνική Περίληψη**



## Κεφάλαιο 7

# Θεωρητική Εισαγωγή

Ο στόχος της σωματιδιακής φυσικής είναι να εξηγήσει το σύμπαν που παρατηρούμε με βασικά δομικά στοιχεία και θεμελιώδεις αλληλεπιδράσεις μεταξύ τους. Οι φυσικοί έχουν αναπτύξει μια θεωρία που απαντά σε πολλά ερωτήματα σχετικά με την ανάπτυξη και την παρατηρήσιμη κατάσταση του σύμπαντος. Μέχρι τώρα, το Καθιερωμένο Μοντέλο των στοιχειωδών σωματιδίων, είναι η πιο ακριβής επαληθευμένη θεωρία στη φυσική. Στις επόμενες ενότητες, θα παρουσιαστεί μια βασική εισαγωγή αυτής της θεωρίας. Επίσης, ιδιαίτερη εστίαση στο στοιχειώδες σωματίδιο που παίζει τον πιο σημαντικό ρόλο σε αυτή την ανάλυση, το τοπ κουάρκ.

### 7.1 Το Καθιερωμένο Πρότυπο

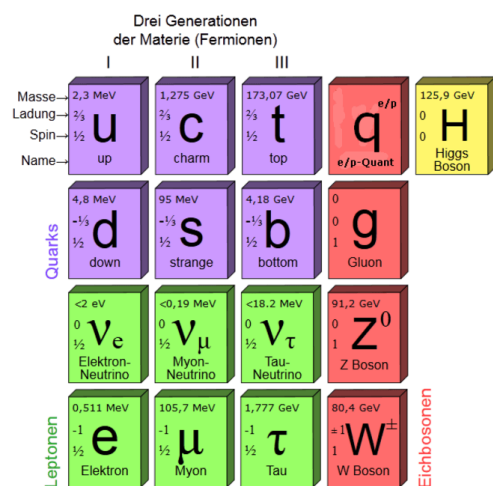
Το Καθιερωμένο Μοντέλο (SM) της σωματιδιακής φυσικής είναι μια κβαντική θεωρία πεδίου που περιγράφει τα στοιχειώδη σωματίδια και τις αλληλεπιδράσεις τους, παρέχοντας συναφείς εξισώσεις κίνησης. Προτάθηκε από τους Glashow, Salam Weinberg στη δεκαετία του εξήντα [1], [2]. Η ομάδα συμμετρίας του περιγράφεται από  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  που αντιπροσωπεύει το τρέχον μοντέλο ισχυρών αλληλεπιδράσεων μέσω του  $SU(3)_C$  και την ενοποίηση των ηλεκτρομαγνητικών και ασθενών αλληλεπιδράσεων μέσω του  $SU(2)_L \times U(1)_Y$ .

Όλα τα σωματίδια που περιέχονται σε αυτή τη θεωρία έχουν ανακαλυφθεί και διάφορα πειράματα έχουν επιβεβαιώσει τις προβλέψεις του SM με πολύ υψηλή ακρίβεια. Τα στοιχειώδη σωματίδια του SM κατατάσσονται σε δύο θεμελιώδεις ομάδες, ανάλογα με το spin τους. Τα σωματίδια Spin- $\frac{1}{2}$  ονομάζονται φερμιόνια ενώ τα σωματίδια με ακέραιο spin ονομάζονται μποζόνια.

Τα φερμιόνια χωρίζονται περαιτέρω σε κουάρκ και λεπτόνια. Ενώ τα κουάρκ επηρεάζονται από την ισχυρή δύναμη, τα λεπτόνια δεν επηρεάζονται. Και οι δύο αυτές ομάδες αποτελούνται από τρεις γενιές. Κάθε γενιά περιέχει ένα διπλό ισοσπιν με δύο σωματίδια. Μια γενιά κουάρκ αποτελείται από ένα κουάρκ άνω και κάτω τύπου που διακρίνεται από το ηλεκτρομαγνητικό τους φορτίο  $+\frac{2}{3}e$  και  $-\frac{1}{3}e$ , αντίστοιχα. Για τα λεπτόνια, μια γενιά δημιουργείται από ένα σωματίδιο με φορτίο  $-e$  και ένα ουδέτερο νεutrino. Τα κουάρκ και τα λεπτόνια από την πρώτη γενιά είναι τα δομικά στοιχεία των ατόμων και έτσι σχηματίζουν ύλη που παρατηρούμε στη γη. Όλα τα άλλα σωματίδια έχουν μεγαλύτερη μάζα και επομένως πεπερασμένη διάρκεια ζωής. Επιπλέον, κάθε φερμιόνιο έχει ένα αντισωματίδιο που έχει την ίδια μάζα αλλά αντίθετο φορτίο.

Τα μποζόνια είναι φορείς των τριών θεμελιωδών δυνάμεων που περιλαμβάνονται στο SM. Τα φωτόνια χωρίς μάζα και τα γκλουόνια είναι οι φορείς της ηλεκτρομαγνητικής και της ισχυρής δύναμης, αντίστοιχα. Τα μποζόνια  $W^\pm$  και  $Z^0$  μεταφέρουν την ασθενή αλληλεπίδραση. Ενώ όλα τα ονομαζόμενα μποζόνια είναι σωματίδια spin-1, το μποζόνιο Higgs είναι βαθμωτό. Είναι συνέπεια του πεδίου Higgs, που δίνει στα στοιχειώδη σωματίδια τη μάζα τους. Κάθε αλληλεπίδραση δύο σωματιδίων περιγράφεται από ένα μποζόνιο που εκπέμπεται από το ένα και απορροφάται από ένα άλλο σωματίδιο. Έτσι, ένα μποζόνιο αλλάζει ενέργεια και ορμή ενός απορροφητικού εκπεμπόμενου σωματιδίου. Για κάθε δύναμη εισάγεται ένα φορτίο. Μόνο τα σωματίδια που φέρουν το φορτίο που συνδέονται με μια δύναμη μπορούν να αλληλεπιδράσουν με το αντίστοιχο μποζόνιο. Ως εκ τούτου, η ποσότητα φορτίου που μεταφέρεται από ένα σωματίδιο είναι ανάλογη με την πιθανότητα να εκπέμψει ή να απορροφήσει ένα μποζόνιο.

Το δεύτερο μέρος περιλαμβάνει τις θεμελιώδεις αλληλεπιδράσεις (Πίνακας 1.3) στα οποία αντι-στοιχούν τα μποζόνια με spin-1. Οι δυνάμεις αυτές είναι διαφορετικής εμβέλειας και δύναμης. Οι



φορείς τους είναι: το γλουόνιο για την ισχυρή αλληλεπίδραση, το φωτόνιο για την ηλεκτρομαγνητική, τα δυο W και Z για την ασθενή και το υποθετικό γκραβιτόνιο για τη βαρύτητα.

Το τρίτο μέρος περιλαμβάνει τον μηχανισμό Higgs που εισήχθει από τον Englert-Brout-Higgs (1964) και επιτρέπει στα σωματίδια να αποκτούν μάζα ανάλογα με την επίδρασή τους με το πεδίο Higgs . Οι μάζες καθορίζονται ανάλογα με το πόσο ισχυρή είναι η σύζευξη (coupling) με το πεδίο Higgs. Το SM Higgs έχει μάζα 125 GeV και είναι αποτέλεσμα του αυθόρμητου σπασίματος της ηλεκτρασθενούς συμμετρίας.

Με την ανακάλυψη του μποζονίου Higgs το 2012 [5], [6], όλα τα σωματίδια του SM έχουν επιβεβαιωθεί πειραματικά. Αν και η θεωρία πιστεύεται ότι είναι ελλιπής επειδή δεν περιλαμβάνονται γνωστά φαινόμενα (όπως η βαρύτητα και η σκοτεινή ύλη), το Καθιερωμένο Πρότυπο είναι πολύ επιτυχημένο μοντέλο που περιγράφει τα αποτελέσματα των πειραμάτων με πολύ υψηλή ακρίβεια. Μια πιο λεπτομερής ματιά μετά το [1] για τις τρεις θεμελιώδεις αλληλεπιδράσεις δίνεται στις παρακάτω ενότητες.

Πίνακας 1.1: Κουάρκς

Πρώτη γεννιά	up (u)	down (d)
Δεύτερη γεννιά	charm (c)	strange (s)
Τρίτη γεννιά	top (t)	bottom (b)

Πίνακας 1.2: Λεπτόνια

Πρώτη γεννιά	electron (e)	electron neutrino ( $\nu_e$ )
Δεύτερη γεννιά	muon ( $\mu$ )	muon neutrino ( $\nu_\mu$ )
Τρίτη γεννιά	tau ( $\tau$ )	tau neutrino ( $\nu_\tau$ )

Πίνακας 1.3: Θεμελιώδεις αλληλεπιδράσεις



Interaction	Strength	Theory	Mediator boon	Charge
Strong	1	Chromodynamics	8 colored gluons	0
Electromagnetic	$10^{-2}$	Electrodynamics	$\gamma$	0
Weak	$10^{-7}$	Flavordynamics	$W^+, W^-, Z$	+1,-1,0
Gravitational	$10^{-39}$	General Relativity	Graviton	0

### 7.1.1 Quantum Electro Dynamics (QED)

Στην ιστορία του Καθιερωμένου Μοντέλου, το QED ήταν ο πρώτος μηχανισμός που περιέγραψε γνωστές αλληλεπιδράσεις σε μια κβαντική θεωρία πεδίου. Το φορτίο του QED ονομάζεται ηλεκτρομαγνητικό φορτίο και δηλώνεται συνήθως σε μονάδες του στοιχειώδους φορτίου ηλεκτρονίων  $e$ . Ο μηχανισμός για την κατασκευή μιας κβαντικής θεωρίας πεδίου είναι να ληφθεί η πυκνότητα Lagrange για τα φερμιόνια, που περιγράφεται από τους σπίνορες Dirac και απαιτεί τοπική αναλλοίωτη μέτρηση. Το QED απαιτεί η πυκνότητα Lagrange να είναι αμετάβλητη στους μετασχηματισμούς μιας ομάδας συμμετρίας  $U(1)$ . Η εξίσωση 1.1 δείχνει την πυκνότητα Lagrange της εξίσωσης Dirac .

$$\mathcal{L} = \underbrace{\bar{\psi}i\gamma^\mu\partial_\mu\psi}_{\text{κινητικός όρος}} - \underbrace{m\bar{\psi}\psi}_{\text{όρος μάζας}} - \underbrace{q\bar{\psi}\gamma^\mu\psi A_\mu(x)}_{\text{όρος αλληλεπίδρασης}} \quad (7.1)$$

Εδώ το  $\psi$  υποδηλώνει τους σπίνορες Dirac και το  $m$  τη μάζα που αντιστοιχεί στον σπινόρα Dirac . Ένα διάνυσμα πινάκων γάμμα  $\gamma^\mu$  χρησιμοποιείται για την εγγραφή της Λαγρανζιανής σε μια τετραδιάστατη αναπαράσταση χωροχρόνου. Ο δείκτης  $\mu$  υποδεικνύει ένα από τα τέσσερα συστατικά ενός τεσσάρων διανύσματος. Ο μετασχηματισμός  $Y(1)$  είναι:

$$\psi \rightarrow \psi' = e^{iq\alpha}\psi \quad (7.2)$$

όπου η παράμετρος  $\chi$  αργότερα θα προσδιοριστεί με μια δύναμη σύζευξης και το  $\alpha = \alpha(x)$  υποδηλώνει τη φάση, ανάλογα με τις χωροχρονικές συντεταγμένες. Ενώ ο όρος μάζας  $m\bar{\psi}\psi$  είναι αμετάβλητος κάτω από έναν μετασχηματισμό  $\psi \rightarrow \psi'$ , η παράγωγος  $\partial_\mu\psi$  δεν είναι. Προκειμένου να κατασκευαστεί μια αμετάβλητη πυκνότητα Λαγρανγε, εισάγεται ένα νέο διανυσματικό πεδίο  $A_\mu(x)$  και ορίζει μια παράγωγο συμμεταβλητής  $D_\mu$ :

$$D_\mu = \partial_\mu + iqA_\mu(x) \quad (7.3)$$

Με την πρόσφατα εισαχθείσα παράγωγο συμμεταβλητής, η πυκνότητα Lagrange γίνεται:

$$\mathcal{L} = \underbrace{\bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi}_{\text{mass and kinetic term quarks}} - \underbrace{g_s\bar{\psi}(\gamma^\mu T^\alpha G_\mu^\alpha)\psi}_{\text{quark-gluon interaction}} - \underbrace{\frac{1}{4}G_{\mu\nu}^\alpha G^{\alpha\mu\nu}}_{\text{kinetic term gluons and gluon-gluon interaction}} \quad (7.4)$$

Απαιτώντας η πυκνότητα Lagrange να είναι συμμετρική κάτω από μετασχηματισμούς  $U(1)$ , προβλέπεται ένα νέο διανυσματικό πεδίο  $A_\mu$ . Αυτό το διανυσματικό πεδίο συνδέεται στη συνέχεια με το φωτόνιο. Επειδή δεν υπάρχει όρος μάζας για το πεδίο φωτονίων, αναμένεται να είναι χωρίς μάζα. Επιπλέον, περιλαμβάνεται η σύζευξη μεταξύ φερμιονίων  $\psi$  και φωτονίου  $A_\mu$  με δύναμη σύζευξης  $\chi$ .

### 7.1.2 Quantum Chromo Dynamics (QCD)

Η θεωρία πίσω από την QCD αναπτύσσεται σε αναλογία με την QED . Το φορτίο στην θεωρία της QCD ονομάζεται χρώμα και αποτελείται από τρεις καταστάσεις για τα σωματίδια: κόκκινο, πράσινο και μπλε. Επιπλέον, κάθε χρώμα έχει το αντιφορτίο του που μεταφέρεται από αντισωματίδια. Έτσι, πρέπει κανείς να εξετάσει τις περιστροφές στον τρισδιάστατο "χρωματικό" χώρο, ο οποίος αντιπροσωπεύεται από μια ομάδα  $SU(3)$  . Οι γεννήτορες  $T_\alpha$ , με λειτουργία από 1 έως 8, χρησιμοποιούνται για να εκφράσουν αυτές τις περιστροφές και πρέπει να ληφθούν υπόψη. Παρόμοια με την θεωρία QED ,

μια συμμεταβλητή παράγωγος ορίζεται για την κατασκευή μιας συμμετρικής πυκνότητας Lagrange . Η συμμεταβλητή παράγωγος για το QCD είναι :

$$D_\mu = \partial_\mu + ig_s T^\alpha G_\mu^\alpha \quad (7.5)$$

όπου το  $g_s$  δηλώνει τη σύζευξη των κουάρκ με την ισχυρή δύναμη. Δεύτερον, εισάγονται οκτώ νέα διανυσματικά πεδία  $G_\mu^\alpha$ , τα οποία σχετίζονται με τα γκλουόνια, τα μποζόνια της ισχυρής δύναμης. Επίσης επιτρέπεται όχι μόνο η αλληλεπίδραση μεταξύ κουάρκ και γκλουονίων αλλά και η αλληλεπίδραση γκλουονίου-γκλουονίου. Η επόμενη εξίσωση δείχνει πώς συνδυάζεται η αλληλεπίδραση των γλουονίων με τον όρο της κινητικής τους ενέργειας. Η πλήρης πυκνότητα Lagrange του QCD γίνεται :

$$\mathcal{L} = \underbrace{\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi}_{\text{mass and kinetic term quarks}} - \underbrace{g_s \bar{\psi}(\gamma^\mu T^\alpha G_\mu^\alpha)\psi}_{\text{quark-gluon interaction}} - \underbrace{\frac{1}{4} G_{\mu\nu}^\alpha G^{\alpha\mu\nu}}_{\text{kinetic term gluons and gluon-gluon interaction}} \quad (7.6)$$

Οι αλληλεπιδράσεις γλουονίου-γκλουονίου οδηγούν σε ένα ειδικό αποτέλεσμα στην θεωρία QCD . η ενεργειακή πυκνότητα μεταξύ δύο φορτίων (χρωμάτων) αυξάνεται με την απόστασή τους. Έτσι, η ενέργεια που αποθηκεύεται στο πεδίο μεταξύ των δύο φορτίων θα αυξηθεί επίσης. Σε κάποιο σημείο η ενέργεια είναι αρκετά μεγάλη ώστε να παράγει φορτισμένα (χρώμα) σωματίδια. Για το λόγο αυτό, τα κουάρκ δεν μπορούν να υπάρχουν ελεύθερα αλλά μόνο σε χρωματικές ουδέτερες δεσμευμένες καταστάσεις, που ονομάζονται αδρόνια.

### 7.1.3 Ασθενής Δύναμη - Weak Interaction

Η τρίτη δύναμη που περιγράφεται στο SM είναι η ασθενής δύναμη. Το φορτίο της ασθενούς δύναμης ονομάζεται ισοσπιν και μεταφέρεται από τα μποζόνια  $W^\pm$  και  $Z^0$ . Επειδή όλα τα ασθενή μποζόνια είναι μαζικά, η ασθενής αλληλεπίδραση έχει πολύ περιορισμένο εύρος. Επομένως, η ασθενής αλληλεπίδραση λαμβάνει χώρα μόνο σε μικρές αποστάσεις. Η αντίστοιχη ομάδα συμμετρίας είναι μια ομάδα  $SU(2)$ . Ενώ τα μποζόνια  $Z^0$  συζευγνύουν επίσης σωματίδια με τα δεξιά, τα  $W^\pm$  όχι. Αυτή η ασυμμετρία στη δύναμη σύζευξης προκαλείται από παραβίαση της ισοτιμίας. Μέσω των  $W^\pm$  μποζονίων, τα φερμιόνια μπορούν να αλλάξουν τη γεύση τους. Τα αριστερόστροφα κουάρκ και τα λεπτόνια είναι διατεταγμένα σε διπλά ισοσπινικά, ενώ τα δεξιόστροφα σωματίδια αντιπροσωπεύονται με μονήρη. Επειδή δεν υπάρχουν δεξιόχειρα νετρίνα, το περιεχόμενο των σωματιδίων περιγράφεται ως :

$$\left( \begin{matrix} u \\ d' \end{matrix} \right)_L, \left( \begin{matrix} c \\ s' \end{matrix} \right)_L, \left( \begin{matrix} t \\ b' \end{matrix} \right)_L, \left( \begin{matrix} \nu_e \\ e \end{matrix} \right)_L, \left( \begin{matrix} \nu_\mu \\ \mu \end{matrix} \right)_L, \left( \begin{matrix} \nu_\tau \\ \tau \end{matrix} \right)_L, u_R, d_R, c_R, s_R, b_R, t_R, e_R, \mu_R, \tau_R \quad (7.7)$$

Η θεωρία επιτρέπει μόνο διαδικασίες αλλαγής γεύσης μέσα σε αυτά τα διπλά, αν και έχουν παρατηρηθεί πειραματικά διασπάσεις από τη μια γενιά στην άλλη. Η απάντηση σε αυτή την αντίφαση είναι ότι οι γευστικές καταστάσεις και οι καταστάσεις μάζας δεν αναφέρονται στην ίδια κατάσταση. Η ασθενής αλληλεπίδραση αλλάζει τη γεύση, αλλά η παρατηρούμενη κατάσταση μάζας είναι μια υπέρθεση διαφορετικών γευστικών καταστάσεων. Εξ ορισμού, οι καταστάσεις γεύσης και μάζας των κουάρκ υπ-τυπε είναι πανομοιότυπες. Για τα κουάρκ τύπου down , εισάγεται η μήτρα Cabibbo-Kobayash-Maskawa (CKM) [6] για να ληφθεί υπόψη η ανάμειξη των καταστάσεων.

$$\left( \begin{matrix} d' \\ s' \\ b' \end{matrix} \right) = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \left( \begin{matrix} d \\ s \\ b \end{matrix} \right) \quad (7.8)$$

Η πιθανότητα ένα κουάρκ τύπου  $i$  να μετατραπεί σε κουάρκ τύπου  $j$  όταν εκπέμπει ένα μποζόνιο  $W^\pm$  υπολογίζεται μέσω  $|V_{ij}^2|$ . Λόγω των καταστάσεων ανάμειξης, παρατηρείται όχι μόνο παραβίαση ισοτιμίας, αλλά και παραβίαση CP (charge and parity) .

### 7.1.4 Electroweak Unification

Για να κατανοήσουμε καλύτερα τις δυνάμεις στο SM , θα ήταν μεγάλη επιτυχία να οδηγήσουμε τις διαφορετικές δυνάμεις πίσω σε έναν θεμελιώδη μηχανισμό. Ένα πρώτο βήμα προς την ενοποίηση και των τριών αλληλεπιδράσεων μπορεί να πραγματοποιηθεί με συνδυασμό QED με την ασθενή δύναμη. Το ηλεκτρικό φορτίο και το ασθενές ισοσπίν συνδυάζονται στο υπερφορτίο  $Y$ :

$$Y = 2(Q - I_3) \quad (7.9)$$

Για άλλη μια φορά, η πυκνότητα Lagrange απαιτείται να είναι αμετάβλητη στους μετασχηματισμούς της υποκείμενης ομάδας  $SU(2) \times U(1)$ . Αυτό έχει ως αποτέλεσμα τέσσερα νέα μποζόνια χωρίς μάζα:  $W_1, W_2, W_3$  και  $B$ .

### 7.1.5 Μηχανισμός Higgs και αυθόρμητο σπάσιμο συμμετρίας

Τα μποζόνια  $W_1, W_2, W_3$  και  $B$  που βρέθηκαν παραπάνω στο 1.1.4 είναι όλα χωρίς μάζα. Σε αντίθεση, οι μάζες των μποζονίων  $W$  και  $Z$  μετρώνται σε τιμές 80 και 91 GeV αντίστοιχα. Ως εκ τούτου, αναπτύχθηκε μια θεωρία που λύνει αυτό το πρόβλημα μέσω της αυθόρμητης διακοπής της συμμετρίας. Εισάγεται ένα πεδίο όπου η πυκνότητα Lagrange για τοπικές παραλλαγές δεν είναι συμμετρική ενώ η Lagrange του ίδιου του πεδίου είναι. Αυτό οδηγεί σε όρους μάζας για ηλεκτροασθενή μποζόνια και ένα νέο μποζόνιο, το μποζόνιο Higgs . Μετά από αυθόρμητο σπάσιμο της συμμετρίας, τα  $W_1$  και  $W_2$  συνδυάζονται για να λάβουν τα  $W^\pm$  μποζόνια των ασθενών αλληλεπιδράσεων μέσω:

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \quad (7.10)$$

Τα άλλα δύο πεδία  $W_3$  και  $B$  αναμειγνύονται και καταλήγουν στο ασθενές  $Z^0$  μποζόνιο και στο φωτόνιο  $\gamma$  της ηλεκτρομαγνητικής αλληλεπίδρασης:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad (7.11)$$

Εδώ εισάγεται η ασθενής γωνία ανάμειξης  $\theta_W$ , η οποία σχετίζεται επίσης με τις μάζες  $M_W$  και  $M_Z$  των  $W^\pm$  και των  $Z^0$  μποζονίων μέσω:

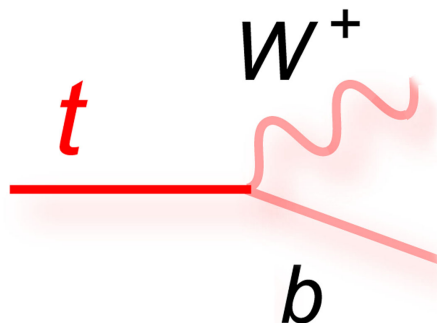
$$M_Z = \frac{M_W}{\cos \theta_W} \quad (7.12)$$

Τέλος, σε αυτόν τον μηχανισμό μπορούν να συμπεριληφθούν και μάζες φερμιονίων. Επομένως, εισάγεται μια σύζευξη Yukawa με το πεδίο Higgs ανάλογη με τις μάζες τους. Μετά το σπάσιμο της συμμετρίας, το SM της σωματιδιακής φυσικής μπορεί να γραφτεί σε μια εξίσωση Lagrange . Για αυτό, η Lagrangian πυκνότητα της QCD πρέπει να συνδυαστεί με την πλήρη Lagrangian πυκνότητα της ηλεκτροασθενούς αλληλεπίδρασης μετά το σπάσιμο της συμμετρίας. Το τελευταίο περιέχει όλους τους κινητικούς όρους των φερμιονίων, τους όρους μάζας των ηλεκτροασθενών μποζονίων και το μποζόνιο Higgs , τις αλληλεπιδράσεις μεταξύ ηλεκτροασθενών μποζονίων και το μποζόνιο Higgs καθώς και τις συζεύξεις Yukawa φερμιονίων στο πεδίο Higgs .

## 7.2 Το Τοπ Κουάρκ

Το τοπ κουάρκ είναι ένα κουάρκ up -τύπου που ανήκει στην τρίτη γενιά του SM και φέρει ηλεκτρομαγνητικό φορτίο  $+\frac{2}{3}$  [9]. Με τη μάζα του να είναι περίπου 173 GeV , το τοπ κουάρκ είναι το βαρύτερο σωματίδιο στο Καθιερωμένο Μοντέλο. Μια τρίτη γενιά κουάρκ είχε προβλεφθεί το 1973 για να εξηγήσει την παρατηρούμενη παραβίαση CP που δεν μπορούσε να εξηγηθεί με τις δύο γνωστές γενιές κουάρκ. Μετά την εύρεση του πολύ ελαφρύτερου κουάρκ κάτω μόνο λίγα χρόνια μετά την πρόβλεψή του, το τοπ κουάρκ ανακαλύφθηκε το 1995 από τις συνεργασίες CDF και DO στο Tevatron . Λόγω

της υψηλής μάζας του, προσφέρει μεγάλο χώρο φάσης για αποσύνθεση και έχει μικρή διάρκεια ζωής περίπου  $5 \times 10^{-25} s$ . Λόγω της μικρής διάρκειας ζωής του, το τοπ κουάρκ δεν σχηματίζει δεσμευμένες αδρονικές καταστάσεις. Έτσι, είναι δυνατές μετρήσεις του κουάρκ, παρέχοντας άμεση πρόσβαση στις παραμέτρους του Καθιερωμένου Μοντέλου. Επιπλέον, το τοπ κουάρκ έχει ιδιαίτερο ενδιαφέρον για αναζητήσεις της φυσικής πέρα από το Καθιερωμένο Μοντέλο, καθώς είναι συχνά μέρος της τελικής κατάστασης ή ένα κυρίαρχο υπόβαθρο.



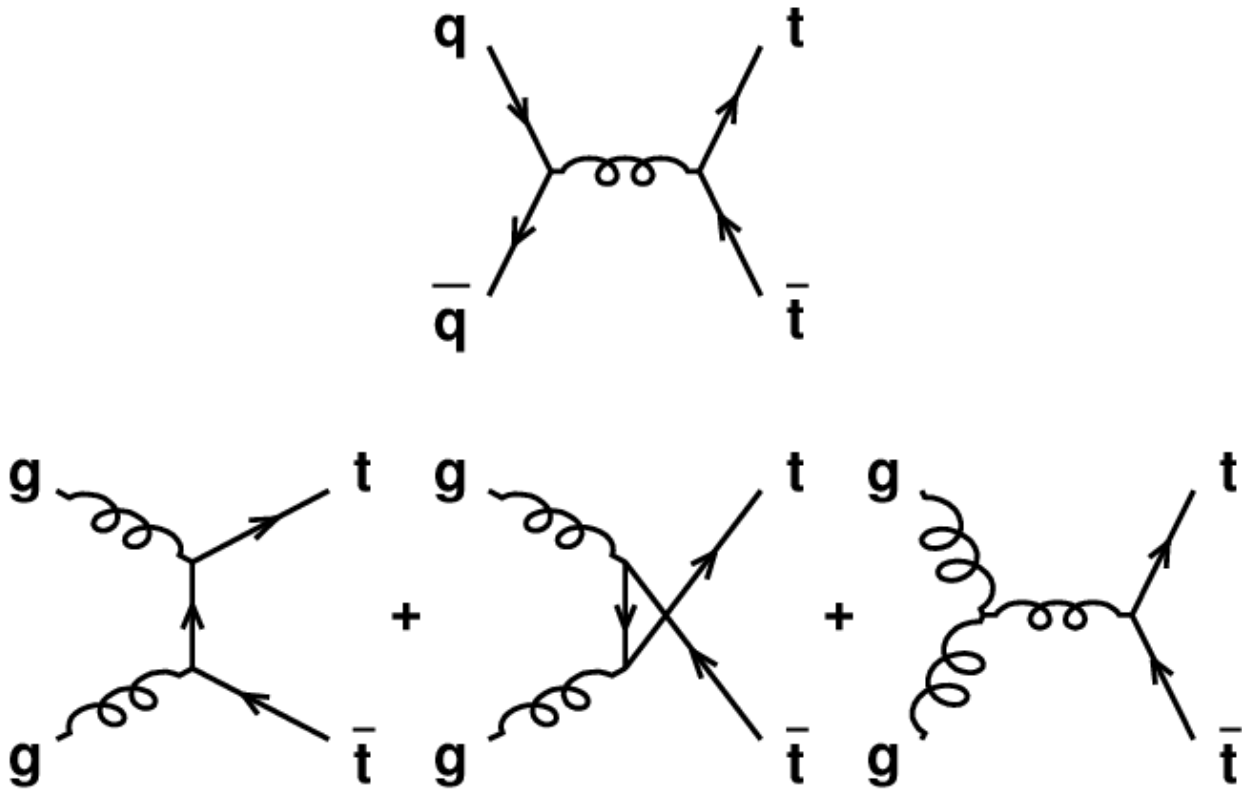
Σχήμα 7.1: Η διάσπαση τοπ κουαρκ

### 7.2.1 Παραγωγή σε επιταχυντές αδρονίων και διάσπαση

Στους επιταχυντές αδρονίων, η παραγωγή ενός ζεύγους  $t\bar{t}$  λαμβάνει χώρα μέσω αλληλεπίδρασης  $q\bar{q}$  ή σύντηξης γκλουονίων (βλ. 7.2). Στο κέντρο μάζας της ενέργειας των 13 TeV, όπως επιτυγχάνεται εντός του LHC στο CERN, η σύντηξη γκλουονίου είναι μακράν η κυρίαρχη διαδικασία. Τα τοπ κουάρκ μπορούν επίσης να παραχθούν σε απλή παραγωγή, αλλά αυτό έχει πολύ μικρότερη διατομή. Ως εκ τούτου, αυτή η ανάλυση εστιάζει στα τοπ κουάρκ που παράγονται από ζευγάρια και αντιμετωπίζει την παραγωγή μεμονωμένων τοπ ως διαδικασία υποβάθρου. Το τοπ κουάρκ διασπάται μέσω της ασθενούς αλληλεπίδρασης με πιθανότητα σχεδόν 100% σε ένα κουάρκ κάτω και ένα μποζόνιο W. Αυτή η ιδιότητα προκύπτει από την καταχώρηση στον πίνακα CKM όπου  $|V_{tb}^2|$  περίπου 0,998. Ενώ το κάτω κουάρκ παρατηρείται ως πίδακας στον ανιχνευτή, το μποζόνιο  $W^\pm$  διασπάται περαιτέρω σε ένα ζεύγος  $q\bar{q}$  ή σε ένα λεπτόνιο (αντι-λεπτόνιο) και ένα αντι-νεutrino (νεutrino). Εξετάζοντας την παραγωγή  $t\bar{t}$ , αυτό καταλήγει σε τρία πιθανά κανάλια:

- και τα δύο μποζόνια  $W$  διασπώνται σε κουάρκ (πλήρως αδρονικά)
- ένα μποζόνιο διασπάται αδρονικά και το άλλο λεπτονικά (λεπτόνιο + πίδακες)
- και τα δύο μποζόνια  $W$  διασπώνται λεπτονικά (δίλεπτον)

Τα πλήρως αδρονικά κανάλια και τα κανάλια lepton+jets είναι κυρίαρχα και εμφανίζονται στο 45,7% και 43,8% του χρόνου αντίστοιχα. Το 10,5% όλων των συμβάντων  $t\bar{t}$  καταλήγει σε δύο λεπτόνια στην τελική κατάσταση. Αυτή η ανάλυση θα επικεντρωθεί στο κανάλι πλήρως αδρονικής διάσπασης.



Σχήμα 7.2:  $t\bar{t}$  παρ προδυστιον

### 7.3 Φυσική των συγκρούσεων πρωτονίου-πρωτονίου

Σε πειράματα σύγκρουσης, η ενέργεια του κέντρου μάζας  $\sqrt{s}$  καθορίζει την ενέργεια που είναι διαθέσιμη για τη δημιουργία σωματιδίων και την απόδοση κινητικής ενέργειας. Στο LHC, το  $s$  ορίζεται από τη ροπή των εισερχόμενων πρωτονίων, όπως:

$$s = (p_1 + p_2)^2 = p_1^2 + p_2^2 + 2p_1p_2 = (E_1 - \vec{p}_1^2) + (E_2 - \vec{p}_2^2) + 2(E_1E_2 - \vec{p}_1\vec{p}_2) \quad (7.13)$$

όπου  $p_1$  και  $p_2$  είναι οι τετραορμές δύο πρωτονίων που συγκρούονται με ενέργειες  $E$  και διάνυσμα ορμών  $\vec{p}$ . Επειδή οι δύο δέσμες έχουν την ίδια ενέργεια αλλά και αντίθετες κατευθύνσεις, ισχύει η υπόθεση ότι  $E_1 = E_2 = E_{proton}$  και  $\vec{p}_1 = -\vec{p}_2$ . Δεδομένου ότι, η ενέργεια κέντρου μάζας έχει τιμή:

$$\sqrt{s} = 2E_{proton} \quad (7.14)$$

Εδώ, τα  $x_1$  και  $x_2$  υποδεικνύουν το κλάσμα της ορμής πρωτονίων που φέρουν τα δύο συγκρουσιακά παρτόνια. Επειδή αυτό το κλάσμα δεν είναι γνωστό, η ενέργεια του κέντρου μάζας δεν είναι γνωστή. Επομένως, για την ανάλυση δεδομένων πρέπει να χρησιμοποιηθούν μεταβλητές που δεν εξαρτώνται από την αρχική ορμή προς την κατεύθυνση των παρτονίων. Μια άλλη ιδιαιτερότητα στους επιταχυντές αδρονίων είναι η χρήση συναρτήσεων πυκνότητας παρτονίου (PDF). Τα PDF επιστρέφουν την πιθανότητα να βρεθεί ένα συγκεκριμένο παρτόνιο μέσα σε ένα δεδομένο διάστημα ορμής σε ένα πρωτόνιο. Μετρώνται σε πειράματα σκέδασης ηλεκτρονίων-πρωτονίων και είναι ζωτικής σημασίας για την πρόβλεψη των διατομών παραγωγής σε συγκρούσεις πρωτονίου-πρωτονίου.

#### 7.3.1 Pile-Up

Ο υψηλός ρυθμός σύγκρουσης στο LHC απαιτείται για τη συλλογή μεγάλου όγκου δεδομένων και τη δυνατότητα παρατήρησης ακόμη και πολύ σπάνιων διεργασιών. Όμως το υψηλό ποσοστό έχει και

ένα μειονέκτημα. Δεδομένου ότι λαμβάνουν χώρα πολλαπλές αλληλεπιδράσεις πρωτονίου-πρωτονίου ανά διασταύρωση δέσμης, όχι μόνο ένα ενδιαφέρον αλλά πολλά άλλα γεγονότα σκέδασης, κυρίως μαλακές διαδικασίες QCD, εμφανίζονται ταυτόχρονα στον ανιχνευτή. Αυτό το φαινόμενο ονομάζεται συσσώρευση. Εάν δεν διορθωθούν, οι μετρήσεις ενέργειας περιλαμβάνουν πάντα σωματίδια που δεν προέρχονται από τη σκληρή σκέδαση που ενδιαφέρει κάποιος. αλληλεπιδράσεις.

## 7.4 Ζ' φυσική

Από την ανακάλυψή του, το μποζόνιο  $Z$  είναι ένα από τα πιο ενδιαφέροντα θέματα στο τμήμα φυσικής υψηλής ενέργειας. Τα μποζόνια  $Z$  είναι τα σωματίδια που μεσολαβούν στην ασθενή και πυρηνική δύναμη και διασπώνται σε οποιοδήποτε από τα γνωστά κουάρκ και λεπτόνια, εκτός από το τοπ κουάρκ. Η ευελιξία τους είναι αυτή που οδηγεί μεγάλα πειράματα όπως το CERN CMS και το ATLAS να διερευνήσουν το φάσμα διαφορετικών φαινομένων φτιάχνοντας μποζόνια  $Z$ .

Ένας παράγοντας που διεγείρει το ενδιαφέρον των φυσικών είναι οποιαδήποτε ένδειξη εάν η πολλαπλή γενιά που παρατηρείται σε κουάρκ και λεπτόνια εφαρμόζεται επίσης σε μποζόνια που διαμεσολαβούν τη δύναμη. Μέχρι στιγμής, καθένα από τα σωματίδια που φέρουν δύναμη του SM (μποζόνια  $W$  και  $Z$ , φωτόνια και γκλουόνια) φαίνεται να είναι μοναδικό. Δεν υπάρχουν ακόμη στοιχεία που να δείχνουν ακόμη και τη δημιουργία στα μποζόνια που διαμεσολαβούν τη δύναμη. Από την άλλη πλευρά, υπάρχουν αρκετές θεωρίες που προβλέπουν την ύπαρξη βαρύτερων μποζονίων των μποζονίων τύπου  $W$  ή  $Z$ . Τέτοια σωματίδια είναι τα  $W'$  ( $W$  prime) και  $Z'$  ( $Z$  prime).

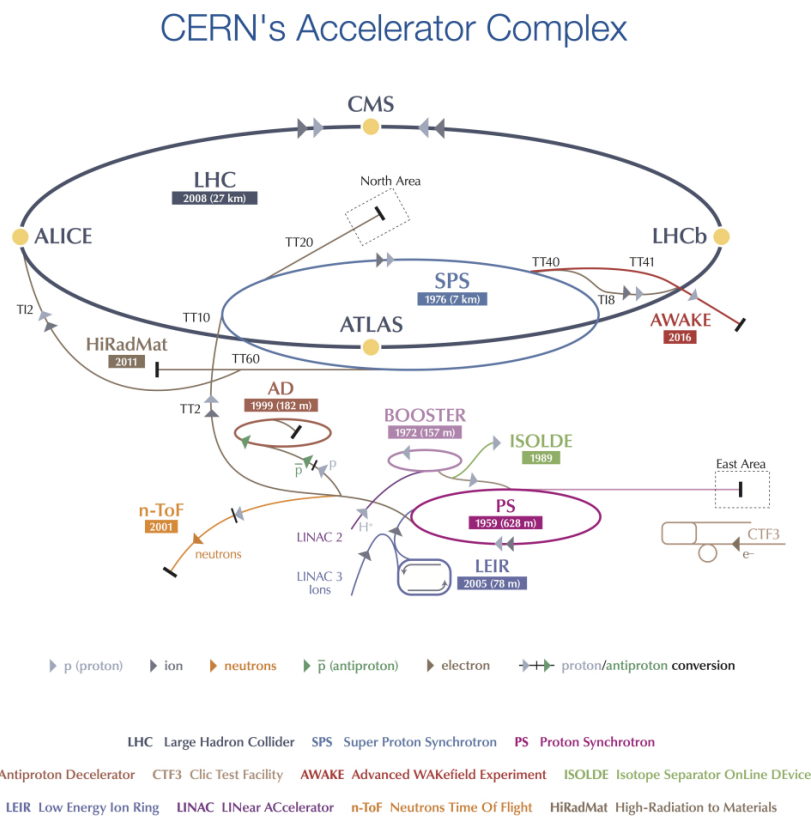
Το πείραμα CMS έψαξε τη διάσπαση ενός μποζονίου  $Z$  σε ένα ζεύγος  $t\bar{t}$ . Αυτός ο συγκεκριμένος τρόπος αποσύνθεσης παρουσιάζει μεγάλο ενδιαφέρον επειδή η παραγωγή τοπ κουάρκ μέσω της συνηθισμένης φυσικής του τυπικού μοντέλου είναι μάλλον σπάνια, καθιστώντας ευκολότερη την απομόνωση και τον εντοπισμό συγκρούσεων στις οποίες μπορεί να έχει γίνει ένα  $Z$ , ένα νέο βαρύ  $Z$  prime. Το διανυσματικό μποζόνιο θα είναι πιθανότατα ένα από τα πρώτα σαφώς ορατά σήματα για τη νέα φυσική που θα ανιχνευθεί από ένα πείραμα, όταν ένας νέος επιταχυντής ενεργοποιείται και/ή επιτυγχάνεται υψηλότερο κέντρο μάζας ενεργειών.

Οι Beyond Standard Model (BSM) εμπλουτίζουν το SM απεικονίζοντας ένα νέο μποζόνιο, το  $Z'$  με τις ίδιες ζεύξεις με το  $Z^0$  αλλά πολύ υψηλότερη μάζα - της τάξης των TeV. Η υπάρχουσα δομή SM ( $SU(3)_C \times SU(2)_L \times U(1)_Y$ ) προστίθεται με μια επιπλέον συμμετρία  $U(1)$ . Σχεδόν σε όλα τα εξεταζόμενα μοντέλα, αυτό το νέο μποζόνιο παρεμβαίνει στο αντίστοιχο του SM οδηγώντας σε έναν επιπλέον βαθμό απόκλισης από την προσδοκία SM στο φάσμα μάζας που προηγείται της κορυφής. Ένα ενδιαφέρον αποτέλεσμα θα ήταν οι γωνιακές κατανομές των γεγονότων στο πλαίσιο Collins-Soper, στο οποίο το ζεύγος  $t\bar{t}$ , έχουν ίση και αντίθετη ορμή. Έτσι, καθένα έχει την ίδια γωνία  $\theta^*$  με την κατεύθυνση της δέσμης.

## Κεφάλαιο 8

# Το πείραμα CMS στο LHC

Ο ανιχνευτής CMS (Compact Muon Solenoid) έχει σχεδιαστεί για να μελετά συγκρούσεις πρωτονίου-πρωτονίου καθώς και βαριές συγκρούσεις ιόντων στον Μεγάλο Επιταχυντή Αδρονίων (LHC) . Το σύμπλεγμα επιταχυντή LHC , μαζί με τα πειράματα κατασκευάστηκε για να μελετήσει τις συγκρούσεις σωματιδίων σε ενέργειες πολύ μεγαλύτερες από αυτές που έχουν επιτευχθεί μέχρι σήμερα.



Σχήμα 8.1: Το σύμπλεγμα επιταχυντών LHC .

### 8.1 Το σύμπλεγμα επιταχυντών LHC

Ο LHC είναι ο νεότερος επιταχυντής σωματιδίων αδρονίων. Βρίσκεται μέσα σε σήραγγα που είναι 100 μέτρα κάτω από τη γη (ανάλογα με την κλίση του επιφάνεια που κυμαίνεται από 50 έως 175 μέτρα) και έχει περίμετρο 26,7 χιλιόμετρα. Φιλοξενείται στην ίδια σήραγγα που χρησιμοποιήθηκε από ο επιταχυντής LEP (Large Electron Positron) επιταχυντής από το 1989 έως το 2000 στα σύνορα Ελβετίας-Γαλλίας στη Γενεύη - Ελβετία στο ένα εργαστήριο που ονομάζεται «Conseil Europeen pour la Recherche Nuclaire» (Ευρωπαϊκό « Συμβούλιο Πυρηνικής Έρευνας), CERN . Αποτελείται από οκτώ τόξα και οκτώ ευθεία τμήματα και το επίπεδό του είναι κεκλιμένο με 1,4% κλίση.

Η διαδικασία επιτάχυνσης χωρίζεται σε διάφορα στάδια. Στο πρώτο στάδιο, τα πρωτόνια ξεκινούν από την πηγή πρωτονίων, η οποία είναι μια απλή φιάλη αερίου υδρογόνου. Ένα ηλεκτρικό πεδίο βάλλεται για να αφαιρέσει τα άτομα υδρογόνου από τα ηλεκτρόνια τους και να δώσει πρωτόνια. Τα παραγόμενα πρωτόνια επιταχύνονται σε έναν γραμμικό επιταχυντή 50 MeV (LINAC) . Ακολουθώντας την αλυσίδα επιτάχυνσης, ο ενισχυτής 1,4 GeV Proton Synchrotron (PS) ακολουθούμενος από το 25 GeV PS παρέχει δέσμες  $1,15 \times 10^{11}$  πρωτονίων με διαφορά χρόνου 50 ns μεταξύ τους στο Super Proton Synchrotron 6,9 km (SPS) . Η ενέργεια των δεσμίδων μέχρι αυτό το σημείο της αλυσίδας επιτάχυνσης είναι έως και 450 GeV. Τέλος, οι δέσμες εισέρχονται στον LHC , ο οποίος είναι υπεύθυνος για την εστίαση, την τελική επιτάχυνση και τελικά τις συγκρούσεις στα σημεία αλληλεπίδρασης των επιμέρους πειραμάτων με ρυθμό διέλευσης δέσμης 40 MHz σε σχεδιασμένη ενέργεια 14 TeV.

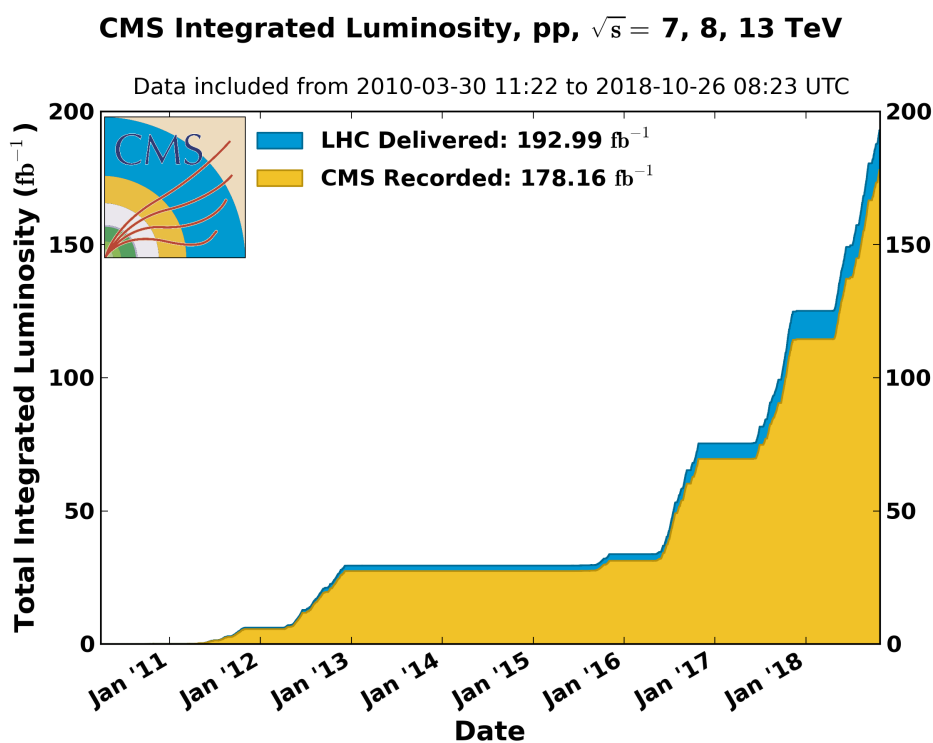
Η δέσμη κινείται μέσω του επιταχυντή χρησιμοποιώντας 9600 μαγνήτες διαφορετικών ποικιλιών και μεγεθών. Αυτοί περιλαμβάνουν 1232 διπολικούς μαγνήτες με μήκος 15 μέτρα που λυγίζουν τις δέσμες όπως φαίνεται στο Σχ.2.2 και 392 τετραπολικούς μαγνήτες, μήκους 5-8 μέτρων ο καθένας, που χρησιμοποιούνται για την εστίαση της δέσμης καθώς και για τις διορθώσεις της δέσμης. Η επιτάχυνση της δέσμης λαμβάνεται μέριμνα με τη χρήση κοιλοτήτων Ραδιοσυχνότητας (RF) 400 MHz.

Οι μαγνήτες έχουν πηνία νιοβίου-τιτανίου (Nb-Ti) που παράγουν μαγνητικό πεδίο έως 8,33 T με ρεύμα σχεδόν 12 kA. Ο μαγνήτης και οι γραμμές δέσμης βρίσκονται στον ίδιο κρουστάτη και το σύστημα ψύχεται στους 1,9 K χρησιμοποιώντας υπερρευστό Ήλιο (He) . Λίγο πριν από τα σημεία σύγκρουσης, ένας άλλος τύπος μαγνήτη χρησιμοποιείται για να «συμπιέσει» τα σωματίδια πιο κοντά μεταξύ τους για να αυξήσει την πιθανότητα σύγκρουσης. Ο LHC είναι επίσης ικανός να επιταχύνει τα ιόντα μολύβδου (Pb) για να παρέχει τόσο συγκρούσεις μολύβδου-μολύβδου όσο και πρωτονίου-μολύβδου. Έχει τέσσερα σημεία αλληλεπίδρασης, με το καθένα να φιλοξενεί ένα πείραμα φυσικής σωματιδίων. Το ATLAS και το CMS είναι πειράματα γενικής χρήσης σχεδιασμένα για να μελετούν ένα ευρύ φάσμα φαινομένων, συμπεριλαμβανομένης της ακριβούς μέτρησης των διαδικασιών του Καθιερωμένου Μοντέλου και των αναζητήσεων για στοιχεία Νέας Φυσικής. Επιπλέον, υπάρχουν δύο, πιο εξειδικευμένα, πειράματα: LHCb με εστίαση στη *b* -φυσική και ALICE , που στοχεύουν κυρίως στην έρευνα του πλάσματος κουάρκ-γλουονίων σε βαριές συγκρούσεις ιόντων.

Μια πολύ σημαντική παράμετρος ενός επιταχυντή σωματιδίων είναι η φωτεινότητά του. Παρέχει ένα μέτρο για το πόσες συγκρούσεις λαμβάνουν χώρα ανά περιοχή και το δεύτερο, επομένως, πόσο καλά εστιάζεται η δέσμη στα σημεία αλληλεπίδρασης. Υπολογίζεται ως εξής:

$$L = \frac{nN_1N_2f}{4\pi\sigma_x\sigma_y} \quad (8.1)$$





Σχήμα 8.2: Σωρευτική παραδοθείσα και καταγεγραμμένη φωτεινότητα σε σχέση με το χρόνο για τα έτη 2010-2012 και 2015-2018 (μόνο δεδομένα pp). Παραδόθηκε και καταγράφηκε αθροιστική φωτεινότητα για όλα τα χρόνια κατά τη διάρκεια σταθερών δεσμών για συγκρούσεις pp σε ονομαστική ενέργεια κέντρου μάζας.

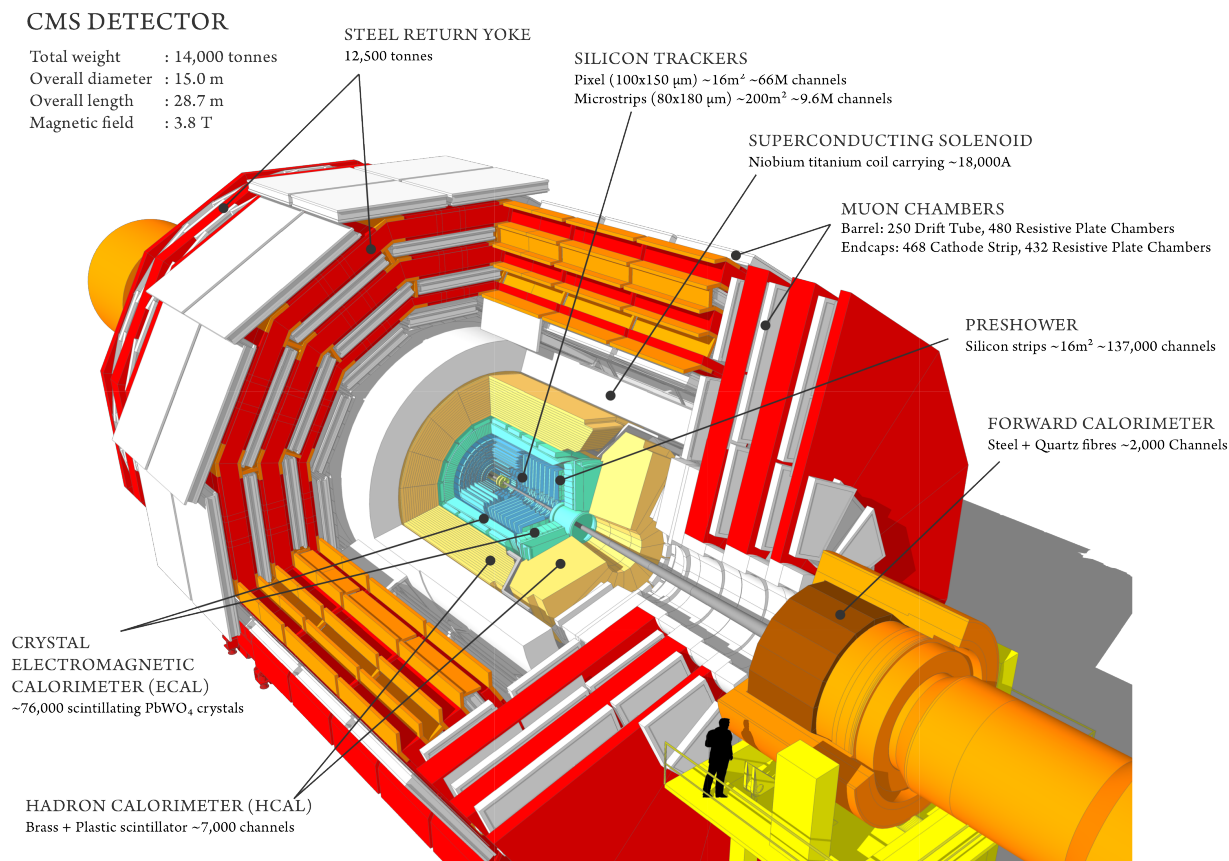
όπου  $n$  υποδηλώνει τον αριθμό των δεσμίδων στον επιταχυντή,  $N_1$  και  $N_2$  είναι ο αριθμός των πρωτονίων στις δύο δέσμες που συγκρούονται και  $f$  είναι η συχνότητα σύγκρουσης. Ο παρονομαστής υποδηλώνει την περιοχή διατομής όπου τα  $\sigma_x$  και  $\sigma_y$  περιγράφουν την εξάπλωση της δέσμης πρωτονίων προς την κατεύθυνση  $x$  και  $y$  αντίστοιχα. Η σχεδιαστική φωτεινότητα του LHC είναι  $10^3 4cm^{-2}s^{-1}$  έχει ήδη ξεπεραστεί από το 2016. Πολλαπλασιάζοντας τη διατομή παραγωγής μιας συγκεκριμένης διεργασίας  $\sigma$  με την φωτεινότητα  $L$  ενσωματωμένη στο χρόνο που απαιτείται για τα δεδομένα, λαμβάνεται μια εκτίμηση για το πόσα συμβάντα  $N$  μιας διαδικασίας πραγματοποιήθηκαν σε μια περίοδο λήψης δεδομένων,

$$L_{int} = \int L dt \quad (8.2)$$

$$N = L_{int} \sigma. \quad (8.3)$$

## 8.2 Ο ανιχνευτής CMS

Το πείραμα CMS είναι ένας ανιχνευτής πολλαπλών χρήσεων στο LHC. Έχει σχεδιαστεί για τη μέτρηση της ορμής και της ενέργειας των σωματιδίων που παράγονται στις αλληλεπιδράσεις πρωτονίου-πρωτονίου. Με συνολικό βάρος 14000 T, η μάζα του ανιχνευτή CMS κυριαρχείται από έναν χαλύβδινο ζυγό επιστροφής που είναι εγκατεστημένος για να οδηγεί το μαγνητικό πεδίο που προέρχεται από μια ηλεκτρομαγνητική βαλβίδα. Ο ανιχνευτής έχει κυλινδρικό σχήμα με μήκος 28,7 μέτρα και διάμετρο 15,0 μέτρα. Ο ανιχνευτής είναι κατασκευασμένος σε στρώματα υποανιχνευτών σε στρώσεις, δηλαδή τον tracker, θερμιδόμετρα, μαγνήτης και σύστημα μιονίων, με διαφορετικούς σκοπούς που περιγράφονται στις ακόλουθες ενότητες. Το γεγονός ότι διαφορετικά σωματίδια αφήνουν διαφορετικές τροχιές και υπογραφές στον ανιχνευτή χρησιμοποιείται για την αναγνώριση και διάκριση των σωματιδίων.



Σχήμα 8.3: Πλήρης προβολή του ανιχνευτή CMS . Βασικές ιδιότητες όπως βάρος και διαστάσεις είναι συνοψίζονται σε έναν πίνακα στο αριστερό επάνω μέρος

Εκτός από τα συστήματα ανιχνευτών, ένα πολύ σημαντικό στοιχείο είναι το σύστημα σκανδαλισμού (trigger), το οποίο παρέχει γρήγορες αποφάσεις εάν ένα συμβάν απορρίπτεται ή είναι αρκετά ενδιαφέρον για να αποθηκευτεί. Ένα τέτοιο σύστημα είναι απαραίτητο επειδή τα δεδομένα δεν μπορούν να αποθηκευτούν με τον ίδιο ρυθμό που συμβαίνουν οι συγκρούσεις στο LHC . Ο σχεδιασμός του CMS επιλέχθηκε για να καλύψει ένα ευρύ φάσμα προσεγγίσεων φυσικής. Ωστόσο, επικεντρώθηκε στην ανακάλυψη του μποζονίου Higgs , το οποίο ανακοινώθηκε το 2012 από το ATLAS και το CMS . Τα εξαρτήματα, ο σχεδιασμός και η γενική κατασκευή του ανιχνευτή CMS φαίνονται στο Σχ. 8.3 .

### 8.2.1 Σύστημα Συντεταγμένων

Το σύστημα συντεταγμένων που χρησιμοποιείται στο πείραμα CMS βασίζεται σε καρτεσιανές και δεξιόστροφες συντεταγμένες. Η αρχή του ορίζεται στο κέντρο του ανιχνευτή. Τα σημεία σταθεροποίησης ρυθμίζονται για να ορίσουν την κατεύθυνση των αξόνων. Ο άξονας  $x$  δείχνει προς την κατεύθυνση του κέντρου του δακτυλίου LHC , ο άξονας  $y$  δείχνει προς τα πάνω και ο άξονας  $z$  ορίζεται παράλληλα με τον άξονα της δέσμης. Επειδή ο LHC είναι ένας επιταχυντής αδρονίων και η ορμή  $z$ -ορμής του συστήματος κέντρου μάζας μιας σύγκρουσης είναι άγνωστη, ένα διαφορετικό σύνολο συντεταγμένων χρησιμοποιείται στις αναλύσεις δεδομένων. Μια πρώτη προσέγγιση είναι οι σφαιρικές συντεταγμένες με τις γωνίες  $\phi$  και  $\theta$ . Η γωνία  $\phi$  ορίζεται ως η γωνία στο επίπεδο  $x - y$  που μετράται από τον άξονα  $x$ . Η γωνία  $\theta$  περιγράφει τη γωνία από ένα δεδομένο σημείο το διάστημα  $xyz$  προς τον άξονα της δέσμης ( $z$ -άξονας). Η  $\theta$  είναι μια πολύ χρήσιμη μεταβλητή για τη μέτρηση αποστάσεων σε πειράματα όπως το CMS . Επομένως, εισάγεται μια νέα μεταβλητή, η ψευδοταχύτητα  $\eta$ , της οποίας οι διαφορές είναι αμετάβλητες σε μετασχηματισμούς Λορεντζ,

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \quad (8.4)$$

Η απόσταση  $\Delta R$  μεταξύ δύο αντικειμένων ( $i, j$ ) υπολογίζεται χρησιμοποιώντας τις διαφορές  $\Delta\phi =$

$\phi_i - \phi_j$  και  $\Delta\eta = \eta_i - \eta_j$ , επομένως

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad (8.5)$$

Επίσης, κατά τον υπολογισμό της ορμής, δεν είναι και πάλι πρακτικό να χρησιμοποιούνται στοιχεία του άξονα  $z$ . Επομένως, ορίζεται η εγκάρσια ορμή  $p_T$  η οποία κατασκευάζεται χρησιμοποιώντας τις συνιστώσες  $x$  και  $y$  της συνολικής ορμής ενός αντικειμένου  $(p_x, p_y)$   $\varsigma$

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (8.6)$$

Ένα πλήρες σύνολο συντεταγμένων για την περιγραφή της ορμής και της μάζας ενός ανιχνευόμενου σωματιδίου είναι η ενέργεια  $E$ , η εγκάρσια ορμή  $p_T$ , η ψευδοταχύτητα  $\eta$  και η γωνία  $\phi$ .

## 8.2.2 Υπεραγωγίμος Σωληνοειδής Μαγνήτης

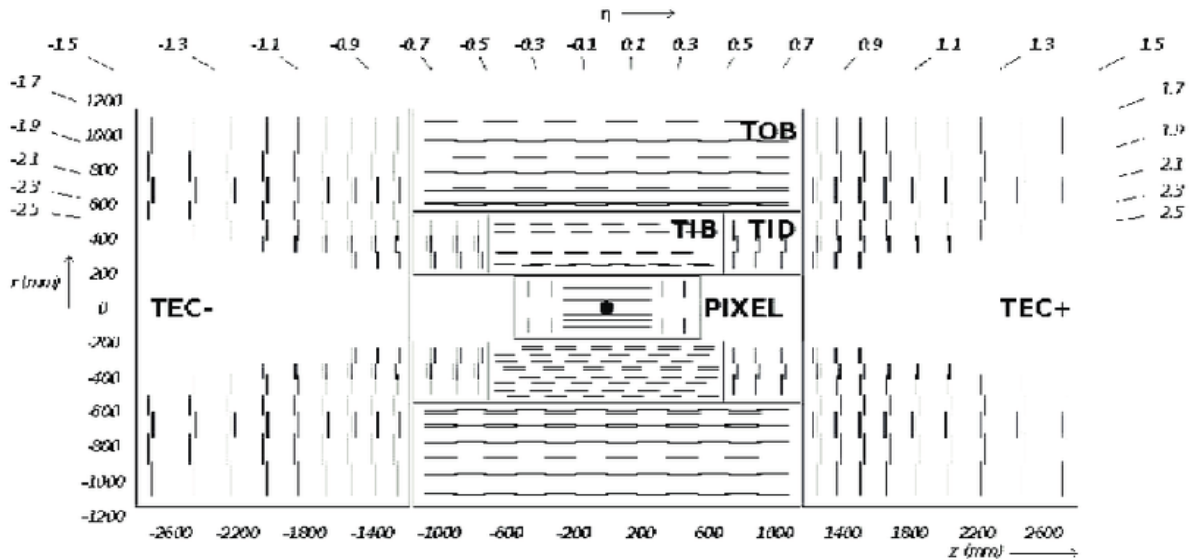
Ο ανιχνευτής CMS είναι εξοπλισμένος με έναν υπεραγωγίμο σωληνοειδή μαγνήτη που παράγει ένα μαγνητικό πεδίο ικανό να κάμπτει την τροχιά ηλεκτρομαγνητικών φορτισμένων σωματιδίων. Μέσω της ακτίνας καμπυλότητας και της έντασης του μαγνητικού πεδίου, μπορεί κανείς να υπολογίσει τη ροπή και το φορτίο των φορτισμένων σωματιδίων. Επομένως, το CMS χρησιμοποιεί έναν υπεραγωγίμο μαγνήτη σωληνοειδούς για να παρέχει ένα μαγνητικό πεδίο 3,8 T στα εσωτερικά στρώματα του ανιχνευτή. Το σωληνοειδές έχει μήκος 12,9 μέτρα και διάμετρο 5,9 μέτρα. Ο tracker και τα θερμιδόμετρα τοποθετούνται εντός του μαγνήτη, ενώ το σύστημα μιονίων είναι εγκατεστημένο έξω από αυτόν. Ο σκοπός του μαγνητικού πεδίου στο εξωτερικό, είναι να κάμψει τη διαδρομή των μιονίων εντός των θερμιδομέτρων μιονίων. Η ηλεκτρομαγνητική μορφή του μαγνήτη επιλέγεται επειδή παρέχει σταθερή μαγνητική ροή μέσα στο σύστημα παρακολούθησης.

## 8.2.3 Tracker

Το σύστημα εσωτερικής παρακολούθησης του ανιχνευτή CMS αποτελείται από το πιο εσωτερικό μέρος του ανιχνευτή CMS, που περιβάλλει το σημείο αλληλεπίδρασης. Έχει σχεδιαστεί για να παρέχει ακριβείς και αποτελεσματικές μετρήσεις των τροχιών των φορτισμένων σωματιδίων (ηλεκτρόνια, μίονια, αδρόνια) και να ανακατασκευάζει με ακρίβεια τις κορυφές πρωτογενούς και δευτερεύουσας αλληλεπίδρασης. Γενικά, ο σκοπός του είναι να μετρήσει την ορμή και το φορτίο των σωματιδίων. Η ιδέα για την κατασκευή του tracker προέρχεται από τη δύναμη του Λορεντζ που κάμπτει την τροχιά των φορτισμένων σωματιδίων σε ένα μαγνητικό πεδίο. Η ακτίνα κάμψης είναι ανάλογη με την ορμή σε ένα σταθερό μαγνητικό πεδίο.

Η δομή του συστήματος παρακολούθησης φαίνεται στο Σχήμα 8.4 όπου φαίνονται διαφορετικές μονάδες του εσωτερικού τραςκερ:

- pixel
- inner barrel (TIB)
- outer barrel (TOB)
- inner disks (TID)
- endcaps (TEC)



Σχήμα 8.4: Σκίτσο του tracker . Εμφανίζονται όλα τα υποσυστήματα, συγκεκριμένα ο ανιχνευτής εικονοστοιχείων και ταινιών με εσωτερική κάμψη (TIB), εξωτερική κάμψη (TOB) και εξωτερικά τερματικά πώματα (TEC<sup>-</sup>). Επιπλέον, το εύρος σε  $\eta$ , ακτίνα (r) και κατεύθυνση ζ μπορεί να διαβαστεί.

### 8.2.4 Θερμιδόμετρα

Προκειμένου να καθορισθεί ένας τύπος σωματιδίου, είναι σημαντικό οι ανιχνευτές να μπορούν επίσης να μετρήσουν την ενέργεια ενός σωματιδίου. Επομένως, έξω από τον ιχνηλάτη έχει εγκατασταθεί ένα σύστημα θερμιδομετρίας με δύο τύπους θερμιδομέτρων. Η βασική αρχή των θερμιδομέτρων είναι να απορροφούν όλη την ενέργεια που μεταφέρει ένα σωματίδιο ενώ κινείται μέσα στον ανιχνευτή και να τη μετατρέπουν σε μετρήσιμη ποσότητα. Η ενέργεια που εναποτίθεται στο θερμιδόμετρο μετασχηματίζεται σε φωτεινό σήμα χρησιμοποιώντας σπινθηριστικά υλικά. Η ένταση του φωτεινού σήματος είναι ανάλογη με την ενέργεια που εναποτίθεται στο θερμιδομετρικό στοιχείο. Τα στρώματα του ανιχνευτή πρέπει να είναι αρκετά πυκνά ώστε να σταματήσει εντελώς ένα εισερχόμενο σωματίδιο και να αποτραπεί οποιαδήποτε διαρροή. Ως μέτρο για το πάχος του θερμιδομέτρου, το μήκος ακτινοβολίας  $X_0$  για τα ηλεκτρόνια, ορίζεται ως η διαδρομή στην οποία ένα ηλεκτρόνιο χάνει  $\frac{1}{e}$  της αρχικής του ενέργειας. Με ανάλογο τρόπο, εισάγεται ένα μήκος πυρηνικής αλληλεπίδρασης  $\lambda_n$  για τα αδρόνια, λαμβάνοντας υπόψη την κλίμακα μήκους των αλληλεπιδράσεων στρυνίων. Και οι δύο τιμές  $X_0$  και  $\lambda_n$  εξαρτώνται από το χρησιμοποιούμενο υλικό και την πυκνότητα των στρωμάτων ανιχνευτή. Το τελευταίο βήμα είναι η συλλογή του φωτεινού σήματος. Αυτό συλλέγεται και μετράται με πολλαπλασιαστές φωτονίων και μπορεί να μετατραπεί ξανά σε μετρούμενη ενεργειακή τιμή. Οι μετρήσεις από θερμιδομετρικές συνήθως γίνονται πιο ακριβείς ενώ η ενέργεια αυξάνεται. Έτσι, παρέχουν μια καλή εναλλακτική μέθοδο σε περιπτώσεις όπου η μέτρηση του συστήματος παρακολούθησης έχει χειροτερέψει. Η ανάλυση ενός θερμιδομέτρου δίνεται από:

$$\frac{\sigma_E}{E[\text{GeV}]} = \frac{\alpha}{\sqrt{E[\text{GeV}]}} \oplus b \oplus \frac{c}{E[\text{GeV}]} \quad (8.7)$$

Σε αυτόν τον τύπο, το  $\frac{\sigma_E}{E}$  υποδηλώνει τη σχετική ενεργειακή ανάλυση ανάλογα με την ενέργεια E, τον στοχαστικό όρο  $\alpha$ , τον σταθερό όρο  $\beta$  που περιλαμβάνει ανομοιογένειες και τον όρο θορύβου  $c$ , που περιέχει θόρυβο από ηλεκτρονικά. Σε υψηλές ενέργειες, ο όρος  $\beta$  είναι κυρίαρχος αλλά μπορεί να μειωθεί με βαθμονόμηση.

### Ηλεκτρομαγνητικό θερμιδόμετρο - ECAL

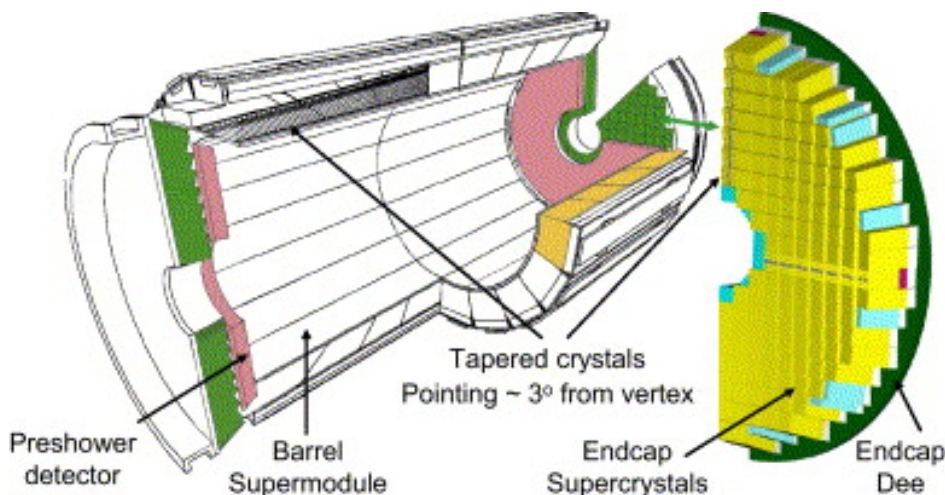
Το ηλεκτρομαγνητικό θερμιδόμετρο (ECAL) του ανιχνευτή CMS προορίζεται για τη μέτρηση της ενέργειας των ηλεκτρονίων και των φωτονίων σε ένα εύρος έως  $|\eta| = 3$ . Είναι ένα ομοιογενές θερμιδόμετρο

ανθεκτικό στην ακτινοβολία που αποτελείται από κρυστάλλους βολφραμικού ( $PbWO_4$ ). Μπορεί να μετρήσει την ενέργεια που εναποθέτουν τα φορτισμένα σωματίδια στους κρυστάλλους ECAL .

Η αρχή λειτουργίας είναι απλή. Ένα φωτόνιο εκκινεί την παραγωγή ζεύγους ηλεκτρονίων-ποζιτρονίων ενώ ένα ηλεκτρόνιο ξεκινά ως ακτινοβολία Bremsstrahlung . Και οι δύο αναπτύσσουν έναν “καταρράκτη” ηλεκτρονίων, ποζιτρονίων και φωτονίων μέσω των προαναφερθέντων διεργασιών. Προκειμένου να μετρηθούν οι ενέργειες φωτονίων και ηλεκτρονίων με υψηλή ανάλυση, το ECAL αποτελείται από υλικά που είναι ιδανικά για σπινθηρισμό. Ο βολφραμικός μόλυβδος είναι ένας ανόργανος σπινθηριστής κρυστάλλων. Κάθε κρύσταλλος έχει μήκος 230 μμ, που αντιστοιχεί σε  $25,8 X_0$  και έχει μπροστινή επιφάνεια  $22 \text{ mm} \times 22\text{mm}$ . Με αυτό, τα φωτόνια που δεν φαίνονται στον ανιχνευτή μπορούν να ανακατασκευαστούν με μεγάλη ακρίβεια. Το ( $PbWO_4$ ) έχει μια ακτίνα Molière η οποία ορίζεται ως η ακτίνα ενός κυλίνδρου που περιέχει 90% των εναποθέσεων ενέργειας, δηλαδή μόνο 2,2 cm . Αυτό οδηγεί σε εξαιρετική ανάλυση θέσης και διαχωρισμό μεταξύ των showers . Αυτή η ιδιότητα είναι ιδιαίτερης σημασίας καθώς επιτρέπει τη διάκριση φωτονίων από απομονωμένα ουδέτερα πόνια ( $\pi^0 \rightarrow \gamma\gamma$  decays).

Το ECAL δίνει επίσης τη δυνατότητα παροχής πληροφοριών που σχετίζονται στενά με το σύστημα ενεργοποίησης του CMS και βοηθούν στην αναγνώριση σωματιδίων.

Όσον αφορά τη δομή θέσης, το ECAL χωρίζεται στην κάννη που καλύπτει ένα εύρος ψευδοταχύτητας  $|\eta| < 1,479$  ενώ τα δύο άκρα καλύπτουν την περιοχή στο  $1,479 < |\eta| < 3,0$ . Τα εξαρτήματα του τελικού καλύμματος περιλαμβάνουν επίσης έναν ανιχνευτή προ ντους που καλύπτει την περιοχή  $1,65 < |\eta| < 2,6$ . Το Σχήμα 8.5 απεικονίζει τη γεωμετρία ECAL με τα διάφορα στοιχεία της. Το ECAL , σχηματίζει ένα στρώμα μεταξύ του ανιχνευτή και των μονάδων ανιχνευτή HCAL . Το κυλινδρικό «βαρέλι» αποτελείται από 61.200 κρυστάλλους που σχηματίζονται σε 36 «υπερμονάδες», το καθένα ζυγίζει περίπου τρεις τόνους και περιέχει 1700 κρυστάλλους. Τα επίπεδα ακραία καπάκια ECAL σφραγίζουν την κάννη σε κάθε άκρο και αποτελούνται από σχεδόν 15.000 επιπλέον κρυστάλλους.



Σχήμα 8.5: Σχήμα γεωμετρίας CMS ECAL που δείχνει διαφορετικά στοιχεία.

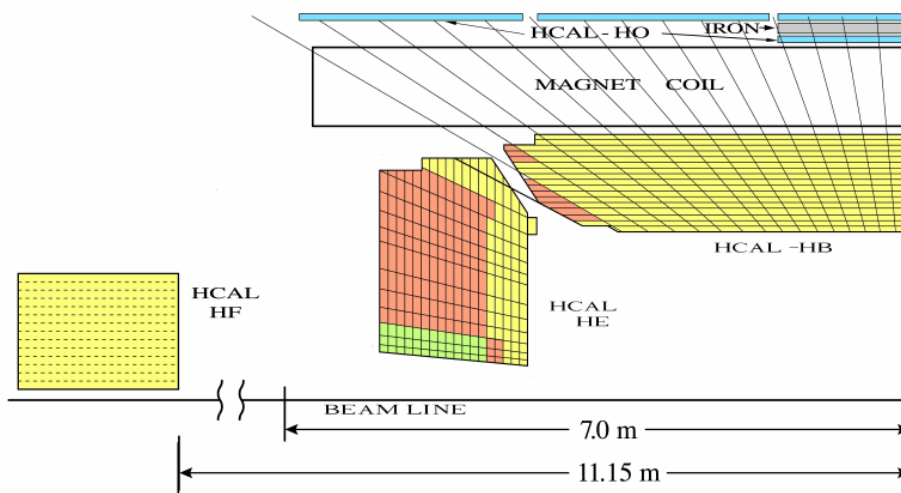
### Αδρονικό Θερμιδόμετρο - HCAL

Δεδομένου ότι τα αδρόνια έχουν μεγαλύτερο μήκος απορρόφησης, περνούν κυρίως το ηλεκτρομαγνητικό θερμιδόμετρο και απορροφώνται και μετρώνται στο αδρονικό θερμιδόμετρο (HCAL ). Χρησιμοποιούνται δύο εναλλασσόμενες στρώσεις διαφορετικών υλικών. Το ένα απορροφά το μεγαλύτερο μέρος της ενέργειας που μεταφέρει ένα εισερχόμενο σωματίδιο και αποτελείται από ορειχάλκινες πλάκες. Το άλλο είναι κατασκευασμένο από πλαστικό σπινθηριστή, που λειτουργεί ως ενεργό υλικό. Το HCAL μπορεί να παρέχει συμπληρωματικές μετρήσεις στα συστήματα ECAL και μιονίων. Τα κύρια αδρόνια που έχουν αρκετά μεγάλη διάρκεια ζωής για να διασχίσουν το θερμιδόμετρο CMS είναι τα πόνια, τα καόνια, τα πρωτόνια και τα νετρόνια.

Η ενεργειακή ακρίβεια του HCAL δίνεται από :

$$\frac{\sigma}{E} = \frac{110\%}{\sqrt{E}} \oplus 9\% \quad (8.8)$$

Η κορυφαία συνεισφορά στην ενεργειακή ανάλυση HCAL προέρχεται από τα αποτελέσματα από τη μη πλήρη συγκράτηση του αδρονικού "ντους", ενώ ο όρος στοχαστικού θορύβου S 110% και ένας σταθερός όρος 9%. Το σχήμα 8.6 δείχνει μια άποψη του HCAL στο επίπεδο  $y - z$ . Το HCAL αποτελείται από ένα εξάρτημα κάννης (HB) και τελικού πώματος (HE), τα οποία περιέχονται και τα δύο μέσα στον ηλεκτρομαγνήτη και το εξωτερικό (HO) και το εμπρός (HF), που βρίσκονται έξω από την ηλεκτρομαγνητική βαλβίδα. Το (HB) καλύπτει το εύρος  $|\eta| < 1,3$  ενώ το HE καλύπτει  $1,3 < |\eta| < 3,0$ . Λόγω του περιορισμένου διαθέσιμου χώρου για το HCAL εντός της ηλεκτρομαγνητικής βαλβίδας, το HO περιλαμβάνεται εκτός της ηλεκτρομαγνητικής βαλβίδας προκειμένου να αυξηθεί το συνολικό μήκος αλληλεπίδρασης.



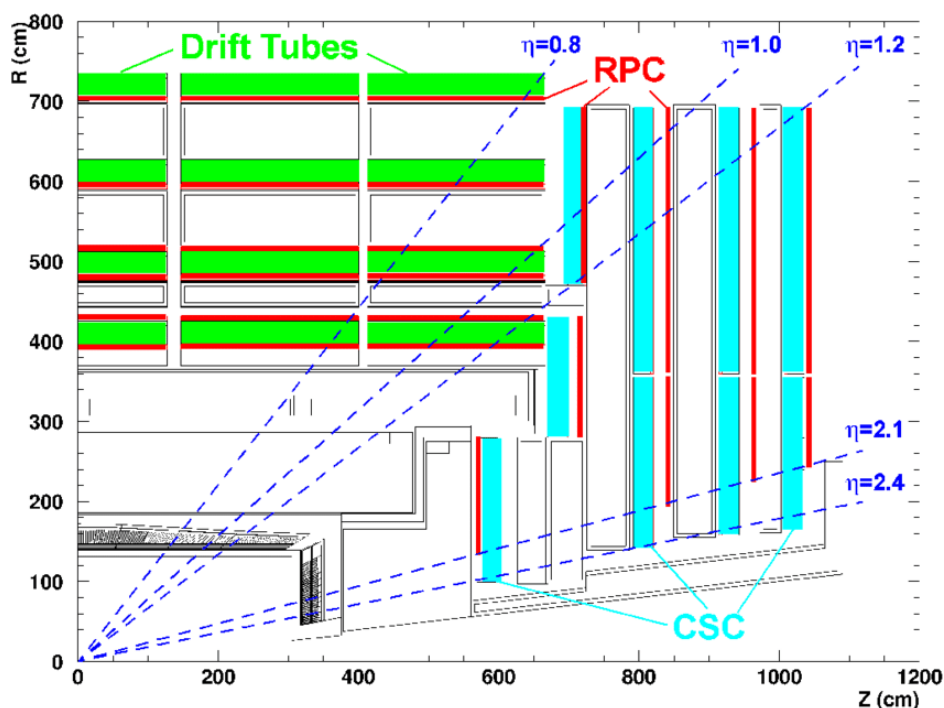
Σχήμα 8.6: Προβολή του αδρονικού θερμιδομέτρου CMS στο επίπεδο  $y-z$ .

Το τμήμα της κάννης του HCAL ξεκινά σε απόσταση  $r = 1,77m$  προς το σημείο αλληλεπίδρασης και φτάνει μέχρι την ηλεκτρομαγνητική βαλβίδα σε  $r = 2,95m$  καλύπτοντας ένα εύρος  $|\eta| < 1,3$ . Το πάχος αντιστοιχεί σε περίπου δέκα φορές το μήκος αδρονικής αλληλεπίδρασης. Τα ενδραπς εγκαθίστανται για να καλύπτουν ένα εύρος  $1,3 < |\eta| < 3$ . Επιπλέον, χρησιμοποιούνται θερμιδόμετρα προς τα εμπρός σε απόσταση 11,1 μ στην κατεύθυνση  $\zeta$  προς το σημείο αλληλεπίδρασης για να καλύψουν το εύρος των  $3 < |\eta| < 5$ . Ένας επιπλέον εξωτερικός θερμιδομέτρης εγκαθίσταται έξω από την ηλεκτρομαγνητική βαλβίδα για να υπολογίζει και να μετράει τη διαρροή.

### 8.2.5 Σύστημα Ανίχνευσης Μιονίων

Λόγω της δυνατότητας ανακατασκευής μιονίων με πολύ υψηλή ακρίβεια, έπαιξαν ένα σημαντικό ρόλο στο σχεδιασμό του CMS αφού εμφανίζονται επίσης σε ένα σημαντικό κανάλι του μποζονίου Higgs που διασπάται σε τέσσερα λεπτόνια. Έτσι, αυτός είναι ο λόγος που εγκαθίσταται ένα ολόκληρο σύστημα για τον εντοπισμό και τη μέτρηση μιονίων με τέτοια ακρίβεια. Γενικά, η ανίχνευση μιονίων είναι ένα ισχυρό εργαλείο από την άποψη της αναγνώρισης υπογραφών ενδιαφέρουσες διεργασίες σε υψηλό ποσοστό υποβάθρου που αναμένεται στις υψηλές ανοσοποιήσεις εντός του LHC .

Το σύστημα μιονίων είναι το πιο εξωτερικό στοιχείο του ανιχνευτή CMS , που βρίσκεται έξω από τον μαγνήτη σωληνοειδούς, που αποτελείται από μια κάννη ( $|\eta| < 1,2$ ) και δύο τερματικά πώματα ( $0,9 < |\eta| < 2,4$ ) που απεικονίζονται στο σχήμα 8.7.



Σχήμα 8.7: Προβολή διατομής ενός τεταρτημορίου του ανιχνευτή CMS στο  $(r, z)$  επίπεδο, που δείχνει τη διάταξη του ανιχνευτή μιονίων.

Ως ανιχνευτής μιονίων διάφορες εκδόσεις αερίων έχουν εγκατασταθεί ανιχνευτές. Στο τμήμα της κάννης, σωλήνες ολίσθησης και θάλαμοι πλάκας αντίστασης είναι ενσωματωμένοι στον σιδερένιο ζυγό επιστροφής του μαγνήτη όπου υπάρχει μαγνητικό πεδίο περίπου 2 T. Επιπλέον, τοποθετούνται δύο τερματικά πώματα με θαλάμους πλάκας αντίστασης σε συνδυασμό με θαλάμους ταινίας καθόδου. Η επιλογή σχεδιασμού οφείλεται στην αναμενόμενη ροή μιονίων στις διάφορες περιοχές του ανιχνευτή. Ενώ τα μίονια εμφανίζονται σπάνια στο τμήμα του κυλίνδρου, προβλέπεται ότι θα εμφανιστούν σε υψηλά ποσοστά στα τελικά πώματα. Οι δύο ανιχνευτές αερίων μπορούν επίσης να παρέχουν ανεξάρτητα πληροφορίες ενεργοποίησης με βάση το  $p_T$  των μιονίων με υψηλή απόδοση και απόρριψη φόντου. Τα ΡΠΣ είναι εγκατεστημένα τόσο στο βαρέλι όσο και στα τελικά καπάκια που καλύπτουν μια περιοχή  $|\eta| < 1,6$ . Επίσης, παρέχουν ένα γρήγορο και άκρως τμηματοποιημένο έναυσμα με ένα απότομο όριο  $p_T$ . Βοηθούν επίσης στην επίλυση ασάφειων κατά την προσπάθεια ανακατασκευής κομματιών από πολλές επιτυχίες σε μια αίθουσα.

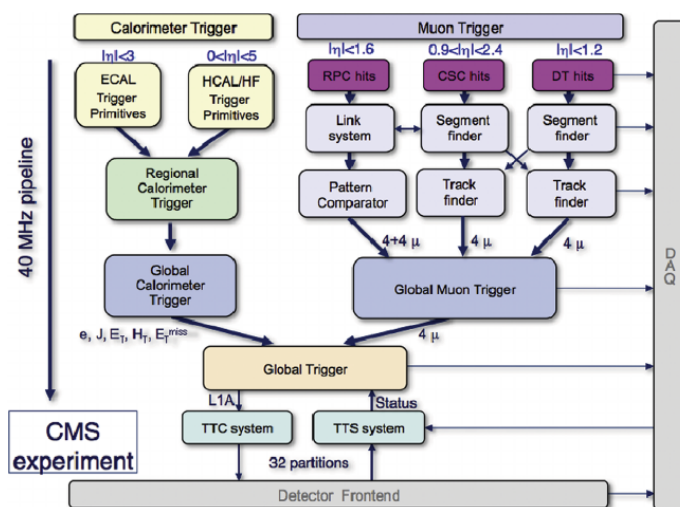
### 8.2.6 Trigger

Στα σημεία αλληλεπίδρασης, οι δέσμες πρωτονίων έρχονται σε σύγκρουση κάθε 25 ns . Αυτό σημαίνει ότι σχεδόν  $8 \times 10^8$  αλληλεπιδράσεις πραγματοποιούνται ανά δευτερόλεπτο. Έτσι, είναι αδύνατο να αποθηκευτεί αυτή η ποσότητα δεδομένων με αυτόν τον ρυθμό, ενώ η συνολική απαιτούμενη χωρητικότητα αποθήκευσης είναι μεγάλη. Για την ακριβή μέτρηση του συμβάντος είναι απαραίτητες όλες οι πληροφορίες των συστημάτων υποανιχνευτών. Θα χρειαζόταν τεράστια ποσότητα υπολογιστικών πυρήνων και υπολογιστικής ισχύος γενικότερα για να αναλυθούν αυτά τα δεδομένα. Επιπλέον, μόνο ένα μικρό κλάσμα γεγονότων είναι ενδιαφέρον για αναλύσεις φυσικής. Επομένως, ο ανιχνευτής CMS είναι εξοπλισμένος με ένα σύστημα ενεργοποίησης που απαιτείται προκειμένου να μειωθεί ο ρυθμός απόκτησης δεδομένων κατά έναν παράγοντα της τάξης των  $10^6$  σε 100HZ . Αυτό επιτυγχάνεται σε δύο στάδια ενεργοποίησης:

1. Ενεργοποίηση επιπέδου 1 βάσει υλικού (Level -1 trigger (L1) )
2. Ενεργοποίηση υψηλού επιπέδου βάσει λογισμικού (High Level Trigger (HLT) )

### Leve L-1 Trigger

Η σκανδάλη πρώτου επιπέδου αποτελείται από προσαρμοσμένο hardware και χρησιμοποιεί σημαντικά γρήγορα δεδομένα που προέρχονται από τους υποανιχνευτές μιονίου και θερμιδόμετρου. Είναι υπεύθυνος για τον εντοπισμό και την επιλογή γεγονότων που μπορεί να περιέχουν μίονια, ηλεκτρόνια, φωτόνια ή πίδακες σε ρυθμό έως 100 kHz εντός 4μs. Η σκανδάλη πρώτου επιπέδου διαιρείται περαιτέρω. Οι τοπικές σκανδάλες (Trigger Primitive Generator) προσδιορίζουν την εναπόθεση ενέργειας σε πύργους ενεργοποίησης θερμιδομετρητών και παρακολουθεί τμήματα ή μοτίβα χτυπημάτων σε θαλάμους μιονίων. Οι τοπικοί ενεργοποιητές καθορίζουν την κατάταξη και την ταξινόμηση αντικειμένων ενεργοποίησης, όπως ένα ηλεκτρόνιο ή ένα υποψήφιο μίονιο, συνδυάζοντας τις πληροφορίες τους και χρησιμοποιώντας αναγνώριση προτύπων. Το Global Calorimeter (GCT) και οι Global Muon Triggers (GMT) καθορίζουν τα αντικείμενα θερμιδομέτρου και μιονίων υψηλότερης κατάταξης σε ολόκληρο το πείραμα και τα μεταφέρουν στον καθολικό έναυσμα (GT) που αποφασίζει την απόρριψη ή την αποδοχή ενός γεγονότος για περαιτέρω αξιολόγηση από το Υψηλό Επίπεδο Trigger (HLT).



Σχήμα 8.8: Η αρχιτεκτονική του συστήματος ενεργοποίησης “ΜΣ ΛεεL-1 .

### High Level Trigger (HLT)

Το δεύτερο επίπεδο γνωστό και ως σκανδάλη υψηλού επιπέδου (HLT), χρησιμοποιεί ένα διαμέρισμα του πλήρους πλαισίου ανακατασκευής συμβάντων για την επεξεργασία δεδομένων με πολύ γρήγορο τρόπο σε μια φάρμα επεξεργασιών. Είναι σε θέση να επαναχρησιμοποιήσει τον ρυθμό συμβάντων έως και 1kHz. Το HLT έχει πρόσβαση στα πλήρη δεδομένα ανάγνωσης και ως εκ τούτου μπορεί να εκτελέσει σύνθετους υπολογισμούς παρόμοιους με αυτούς που γίνονται από το λογισμικό ανάλυσης εκτός σύνδεσης. Όπως η σκανδάλη επιπέδου-1, το HLT χωρίζεται περαιτέρω σε “βήματα” επεξεργασίας, δηλαδή στα βήματα L-2, L-2.5, L-3. Το βήμα L-2 χρησιμοποιεί είσοδο που προέρχεται από το θερμιδόμετρο και τους ανιχνευτές μιονίων, ενώ το L-2.5 χρησιμοποιεί πληροφορίες που προέρχονται από tracker, όπως χτυπήματα pixel. Το βήμα L-3 αναφέρεται σε μια επιλογή που μπορεί να συμπεριλάβει τα πλήρως ανακατασκευασμένα κομμάτια στον tracker. Το HLT είναι σε θέση να επεξεργάζεται και να λαμβάνει αποφάσεις σχετικά με ενδιαφέροντα δεδομένα με πιο αποτελεσματικό τρόπο σε σχέση με το έναυσμα Επιπέδου-1. Μπορεί να περιλαμβάνει πληροφορίες από b-tagging, αλλά και ανακατασκευασμένους πίδακες ενέργειας. Το επόμενο βήμα των επεξεργασμένων δεδομένων που πληρούν τις απαιτήσεις L-1 και HLT είναι η αποθήκευση αυτών των δεδομένων σε φυσικές μνήμες (σκληροί δίσκοι) για την εκτέλεση επεξεργασίας εκτός σύνδεσης που μπορεί να διαρκέσει αρκετά δευτερόλεπτα ανά συμβάν και εκτελεί μια πολύ πιο εντατική υπολογιστική ανακατασκευή συμβάντων που είναι στη συνέχεια χρησιμοποιείται για ανάλυση δεδομένων.



## Κεφάλαιο 9

# Ανακατασκευή Γεγονότων και Προσομοιώσεις

Σε ανιχνευτές σωματιδίων γενικής χρήσης, όπως το CMS, το σήμα εξόδου πρέπει να μετατραπεί σε αντικείμενα με φυσικό νόημα προκειμένου να πραγματοποιηθούν μελέτες και να αναλυθούν τα καταγεγραμμένα δεδομένα. Έτσι, το πείραμα CMS έχει αναπτύξει έναν αριθμό αλγορίθμων για τον εντοπισμό, την ανακατασκευή και τη μέτρηση σωματιδίων. Τα σωματίδια επιστρέφονται με τη μορφή "καταλόγων" ηλεκτρονίων, μιονίων, φωτονίων και πίδακες ενέργειας. Οι υποψήφιοι για κάθε τύπο σωματιδίου πρέπει να διατηρούν ορισμένα κριτήρια για να κατηγοριοποιηθούν ως τέτοια. Σε αυτό το κεφάλαιο, περιγράφονται τόσο τα κριτήρια όσο και οι αλγόριθμοι που χρησιμοποιούνται για τη λήψη αντικειμένων για αναλύσεις φυσικής.

### 9.1 Particle Flow Algorithm

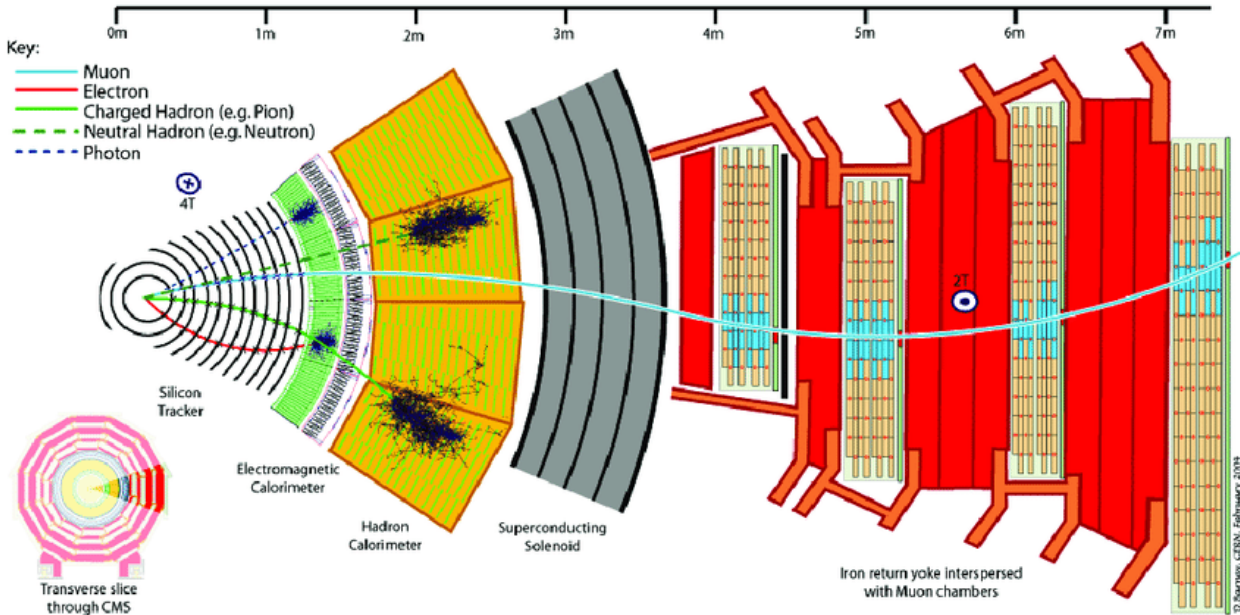
Ο ανιχνευτής CMS είναι ένας ανιχνευτής που αποτελείται από πολλά στρώματα ανιχνευτή γύρω από τον άξονα της δέσμης. Κάθε στρώμα ανιχνευτή παρέχει πληροφορίες για έναν συγκεκριμένο τύπο τύπου σωματιδίων. Το CMS χρησιμοποιεί έναν ειδικό αλγόριθμο, που ονομάζεται αλγόριθμος "Ροή σωματιδίων" Particle Flow algorithm (PF) για να συνδυάσει πληροφορίες που προέρχονται από κάθε σύστημα ανιχνευτή, ακολουθώντας την πλήρη διαδρομή ενός σωματιδίου μέσω ενός ανιχνευτή. Αυτός ο αλγόριθμος είναι σε θέση να κάνει χρήση των αποτυπωμάτων διαφορετικών σωματιδίων που περνούν από διαφορετικά στρώματα του ανιχνευτή. Τα μίονια ανιχνεύονται τόσο στον εσωτερικό tracker όσο και στους θαλάμους μιονίων, τα ηλεκτρόνια ανιχνεύονται τόσο στον εσωτερικό tracker όσο και στον ανιχνευτή ECAL, ενώ τα φορτισμένα αδρόνια αφήνουν ίχνη και βρέχονται στο HCAL. Τα φωτόνια και τα ουδέτερα αδρόνια εναποθέτουν την ενέργειά τους μόνο στο ECAL και στο HCAL αντίστοιχα. Μια σχηματική όψη όλων των επιπέδων ανιχνευτή φαίνεται στο Σχήμα 9.1. Ο αλγόριθμος PF είναι σε θέση να συλλάβει και να χρησιμοποιήσει όλες τις προαναφερθείσες πληροφορίες προκειμένου να αντιστοιχίσει όλες τις ντους από φορτισμένα σωματίδια στις διαδρομές τους, αφήνοντας απαρατήρητες μόνο τις ενεργειακές εναποθέσεις που προέρχονται από φωτόνια και ουδέτερα αδρόνια.

Τα χτυπήματα στον tracker και το σύστημα μιονίων συνδυάζονται. Εάν βρεθεί ένα ίχνος σε ολόκληρο τον ανιχνευτή, το αντικείμενο ονομάζεται υποψήφιο μίονιο και όλα τα χτυπήματα που ανήκουν αφαιρούνται από την είσοδο του αλγορίθμου. Στη συνέχεια, οι υπόλοιπες διαδρομές από ηλεκτρόνια και φορτισμένα αδρόνια προεκτείνονται στα συστήματα θερμιδομέτρησης και εκχωρούνται σε "ντους". Εδώ, η ακτινοβολία φωτονίων από τα ηλεκτρόνια πρέπει να αντιμετωπιστεί, καθώς η ενέργεια από αυτά τα φωτόνια πρέπει επίσης να εκχωρηθεί στο αρχικό ηλεκτρόνιο. Τέλος, οι καταγίγισμοι στο ECAL και το HCAL θα πρέπει να προέρχονται από φωτόνια και ουδέτερα αδρόνια, αντίστοιχα. Για να είναι δυνατή η χρήση της ροής σωματιδίων με υψηλή απόδοση, τα συστήματα παρακολούθησης πρέπει να έχουν καλή χωρική και ενεργειακή ανάλυση. Όλες αυτές οι απαιτήσεις πληρούνται στο CMS, γεγονός που καθιστά τον αλγόριθμο PF ένα ισχυρό εργαλείο για την αναγνώριση και την ανακατασκευή αντικειμένων.

### 9.2 Αλγόριθμοι Jet

Οι πίδακες είναι αντικείμενα που χρησιμοποιούνται για την ανακατασκευή παρτιονίων υψηλής ορμής που παράγονται στις συγκρούσεις πρωτονίου-πρωτονίου. Τα κουάρκ και τα γκλουόνια δεν μπορούν να

Σχήμα 9.1: Ανακατασκευή Γεγονότων στον ανιχνευτή CMS .



υπάρχουν μεμονωμένα, αλλά “αδρονίζονται” (hadronize) σε έναν καταρράκτη χρωματικά ουδέτερων αδρονίων. Επομένως, χρειάζεται να αθροιστούν όλα τα αδρόνια που προέρχονται από το αρχικό parton για να το ανακατασκευάσουμε. Για αυτούς τους τύπους σωματιδίων, χρησιμοποιούνται αλγόριθμοι ομαδοποίησης πίδακα για να συνδυάσουν υποψήφιες ροές σωματιδίων. Το πρώτο κριτήριο σημαίνει είναι ένας πίδακας να μην αλλάζει όταν συμπεριλαμβάνεται ή αποκλείεται η “μαλακή” ακτινοβολία (soft radiation), ενώ το κριτήριο συγγραμμικής ασφάλειας αντιμετωπίζει τη συγγραμμική διάσπαση σωματιδίων σε έναν πίδακα που επίσης δεν πρέπει να αλλάξει τον πίδακα.

### 9.2.1 Αλγόριθμος anti- $k_T$

Στο πείραμα CMS, ο πιο συνηθισμένος τρόπος ομαδοποίησης των ενεργειακών πιδάκων από τα ανιχνευμένα σωματίδια είναι η χρήση επαναληπτικών αλγορίθμων πίδακα. Αυτό σημαίνει ότι τα υποψήφια σωματίδια PF ομαδοποιούνται βήμα προς βήμα μέχρι να επιτευχθεί ένα κριτήριο ματαίωσης χρησιμοποιώντας μια ομάδα αλγορίθμων πίδακα που ονομάζονται αλγόριθμοι τύπου  $k_T$  στο πακέτο FastJet. Ο αλγόριθμος anti- $k_T$  λειτουργεί με τον ακόλουθο τρόπο. Ως είσοδο, χρησιμοποιεί μια λίστα με όλα τα ανακατασκευασμένα αντικείμενα PF από τον ανιχνευτή. Μετρά την απόσταση μεταξύ ενός υποψηφίου PF  $i$  και των γύρω σωματιδίων  $j$  στο επίπεδο ταχύτητας  $y$  και γωνίας  $\phi$ . Στη συνέχεια, ο αλγόριθμος υπολογίζει δύο ποσότητες,  $d_{ij}$  και  $d_{iB}$  για κάθε ζεύγος αντικειμένων  $i$  και  $j$ :

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \quad (9.1)$$

$$d_{iB} = p_{T,i}^2 \quad (9.2)$$

όπου:

- $\Delta R$  ορίζεται ως η απόσταση δύο αντικειμένων  $i, j$  στο επίπεδο γωνίας ταχύτητας  $\phi$  και είναι:  $\Delta R_{ij} \equiv (y_i - y_j)^2 + (\phi_i - \phi_j)^2$
- $p_{T,i}$  είναι η εγκάρσια ορμή του αντικειμένου  $i$
- $R$  είναι μια σταθερή παράμετρος που ορίζει την ακτίνα του πίδακα που προκύπτει
- $d_{iB}$  ορίζεται ως η απόσταση δέσμης

Για κάθε σωματίδιο, ο αλγόριθμος υπολογίζει την απόσταση δέσμης  $d_{iB}$ . Η διαδικασία ομαδοποίησης συνεχίζεται βρίσκοντας το ελάχιστο  $d_{min}$  όλων των αποστάσεων σωματιδίων  $d_{ij}$  και της απόστασης δέσμης. Εάν η μικρότερη απόσταση που βρέθηκε για το σωματίδιο  $i$  με ένα σωματίδιο  $j$  είναι  $d_{ij}$ , τα σωματίδια συνδυάζονται για να σχηματίσουν μια ενιαία οντότητα. Από την άλλη πλευρά, εάν η μικρότερη απόσταση ικανοποιεί  $d_{iB} < \min(d_{ij})$ , το σωματίδιο  $i$  θεωρείται πίδακας και αφαιρείται από τη λίστα των οντοτήτων. Το προαναφερθέν βήμα επαναλαμβάνεται μέχρι το σημείο που δεν έχουν μείνει άλλα σωματίδια. Ο αλγόριθμος  $anti-k_T$  έχει καλύτερη απόδοση από παρόμοιους αλγόριθμους όπως οι αλγόριθμοι  $k_T$ , Cambridge/Aachen και επαναληπτικού κώνου. Αυτός ο αλγόριθμος είναι ο τυπικός αλγόριθμος ομαδοποίησης ενεργειακών πιδάκων που χρησιμοποιείται στο πείραμα CMS και καταλήγει σε σωματίδια με υψηλό  $p_T$  αφού τα σωματίδια με χαμηλό  $p_T$  θεωρούνται υποσωματίδια ενός πίδακα.

Η απόσταση  $R$  επιλέγεται με βάση τις περιπτώσεις χρήσης της συλλογής πίδακα που προκύπτει. Η προεπιλεγμένη συλλογή πίδακα στο CMS αποτελείται από πίδακες με μικρή ακτίνα  $R$  0,4 και αναφέρεται ως συλλογή πίδακα "AK4". Αυτοί οι τύποι αντιστοιχούν τυπικά σε ένα μόνο κουάρκ ή ένα γλουόνιο που προέρχεται από τη διαδικασία σκληρής σκέδασης. Οι συλλογές αεριωθούμενων που χρησιμοποιούν μεγαλύτερες ακτίνες, όπως πίδακες 0,8 («AK8») χρησιμοποιούνται για την ανασύσταση σωματιδίων που διασπώνται με "σκληρό" τρόπο (hard scattering) με μεγάλη ώθηση Lorentz, όπως τα top κουάρκ με υψηλό  $p_T$ . Τα προϊόντα αποσύνθεσης σωματιδίων υψηλού  $p_T$  μπορούν να ευθυγραμμιστούν και έτσι μπορούν να ανακατασκευαστούν σε πίδακα single-R καθιστώντας ακόμη πιο δύσκολη την περαιτέρω διάκριση των σωματιδίων που βρίσκονται εντός της μεγάλης ακτίνας πίδακα AK8.

### 9.2.2 Soft Drop

Όταν ομαδοποιούνται σωματίδια σε πίδακες, οι αλγόριθμοι δεν μπορούν να υπολογίσουν πλήρως από πού προέρχεται ένα ένα σωματίδιο. Οι τεχνικές υποδομής πίδακα (Jet substructure techniques) εφαρμόζονται περαιτέρω για την απομάκρυνση της μαλακής ακτινοβολίας (grooming) προκειμένου να ληφθεί ένας πίδακας που στην καλύτερη περίπτωση περιλαμβάνει μόνο ενέργεια από σωματίδια της διαδικασίας σκληρής σκέδασης. Ο μεγαλύτερος αντίκτυπος στις μετρήσεις μάζας έχει η ακτινοβολία ευρείας γωνίας (wide angle radiation) που εξακολουθεί να ανακατασκευάζεται στον ίδιο πίδακα. Αυτό έχει ως αποτέλεσμα επιπλέον σωματίδια όπως ελαφρά κουάρκ ή γκλουόνια που συγκεντρώνονται σε έναν πίδακα που μπορεί να προέρχεται από διάσπαση top κουάρκ και ως συνέπεια προσθέτει περισσότερη μάζα όσο μεγαλύτερη είναι η απόσταση μεταξύ top κουάρκ και ενός ελαφρού παρτονίου/γκλουονίου. Το πείραμα CMS χρησιμοποιεί τον αλγόριθμο SoftDrop για να μετριάσει αυτό το αποτέλεσμα. Αυτή η μέθοδος αναιρεί το τελευταίο βήμα ομαδοποίησης, σπάζοντας τον τελικό πίδακα σε δύο συστατικά. Μετά από κάθε βήμα αποομαδοποίησης, ελέγχεται το κριτήριο SoftDrop :

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \left( \frac{\Delta R_{12}}{R_0} \right)^\beta \quad (9.3)$$

Εάν η συνθήκη δεν ισχύει, αφαιρείται το συστατικό με χαμηλότερη εγκάρσια ορμή, ενώ σε κάθε άλλη περίπτωση ο τελικός πίδακας παραμένει αμετάβλητος. Τα  $p_{T1}, p_{T2}$  δηλώνουν την εγκάρσια ορμή των δύο συστατικών ενώ η  $R_{12}$  δηλώνει την απόσταση μεταξύ των συστατικών στο επίπεδο  $y - \phi$ . Η παράμετρος  $R_0$  είναι η ακτίνα του τελικού πίδακα. Μια τιμή κατωφλίου  $z_{cut}$  σε συνδυασμό με έναν εκθέτη  $\beta$  καθορίζουν τη δύναμη του καλλωπισμού. Έχουν γίνει μελέτες απόδοσης στο πείραμα CMS με αποτέλεσμα τιμές  $z_{cut} = 0, 1$  και  $\beta = 0$ .

### 9.2.3 PUPPI

Ένας επιπλέον παράγοντας που μπορεί να μολύνει τους ανακατασκευασμένους πίδακες είναι η πρόσθετη ενέργεια που αποδίδεται σε πίδακες που προέρχεται από τις αλληλεπιδράσεις πρωτονίου-πρωτονίου εκτός από το συμβάν σκληρής σκέδασης στην κύρια κορυφή (primary vertex). Αυτή η πρόσθετη ενέργεια από αυτά τα συμβάντα ονομάζεται *pileup* και το CMS έχει αναπτύξει τους δικούς του αλγόριθμους για να την αντισταθμίσει. Εάν ένα φορτισμένο αδρόνιο ανακατασκευαστεί στον ανιχνευτή και αναγνωριστεί ότι προέρχεται από μια κορυφή συσσώρευσης (*pileup vertex*), αφαιρείται

από τη συλλογή σωματιδίων που χρησιμοποιούνται για το σχηματισμό αντικειμένων φυσικής. Αυτή η διαδικασία χρησιμοποιείται ευρέως στην ανακατασκευή πίδακα και αναφέρεται στην αφαίρεση φορτισμένου αδρονίου (charged hadron subtraction - CSH) .

Ένας εναλλακτικός τρόπος για να μετριάσει η συσσώρευση είναι το *pileup per particle identification* γνωστό ως *PUPPI* . Σε αυτό το πεδίο εφαρμογής, σε κάθε ανακατασκευασμένο σωματίδιο εφαρμόζεται ένα βάρος που αναφέρεται στην πιθανότητα αυτό το σωματίδιο να προέρχεται από αλληλεπιδράσεις σσιόβαξης (*pileup* ). Η πιθανότητα βασίζεται στην τοπική κατανομή ενέργειας γύρω από το σωματίδιο, τις πληροφορίες παρακολούθησης και τις ιδιότητες συσσώρευσης συμβάντων. Χρησιμοποιεί αυτήν την εκτίμηση για να κλιμακώσει την τετραορμή των υποψηφίων σωματιδίων PF πριν ομαδοποιηθεί σε έναν πίδακα.

## Κεφάλαιο 10

# Μελέτη Διαφορικού Ενεργού Διατομής Παραγωγής Ζεύγους $t\bar{t}$ με τον ανιχνευτή CMS

Το top κουάρκ ολοκληρώνει την τρίτη γενιά κουάρκ στο Καθιερωμένο Πρότυπο (SM) και η ακριβής γνώση των ιδιοτήτων του είναι κρίσιμη για την πλήρη κατανόηση της θεωρίας. Οι μετρήσεις της διαφορικής ενεργού διατομής παραγωγής του ζεύγους κουάρκ τοπ-αντιτοπ ( $t\bar{t}$ ) αντιμετωπίζουν τις προβλέψεις από τη κβαντικό χρωμοδυναμική (QCD) και έχουν τη δυνατότητα να περιορίζουν τις παραμέτρους QCD, ενώ είναι ευαίσθητοι στη φυσική πέραν του SM. Επίσης, η παραγωγή  $t\bar{t}$  είναι ένα κυρίαρχο υπόβαθρο SM για την αναζήτηση νέων φαινομένων φυσικής και επομένως η ακριβής γνώση της είναι απαραίτητη για νέες ανακαλύψεις.

Η μεγάλη απόδοση  $t\bar{t}$  που αναμένεται σε συγκρούσεις πρωτονίου-πρωτονίου (pp) στο CERN LHC επιτρέπει την εκτέλεση μετρήσεων του ρυθμού παραγωγής  $t\bar{t}$  σε ένα μεγάλο φασικό χώρο, διαφορετικά, ως συνάρτηση των κινηματικών ιδιοτήτων των  $t\bar{t}$ . Τέτοιες μετρήσεις έχουν πραγματοποιηθεί από το ATLAS [64, 65, 66, 67, 68, 69] και CMS [137, 70, 71, 72, 73, 74, 75, 76] οι συνεργασίες επιβεβαιώνουν ότι είναι εφικτό να πραγματοποιηθούν ακριβείς διαφορικές μετρήσεις της παραγωγής  $t\bar{t}$  και έχουν επίσης δείξει ενδιαφέρουσες αποκλίσεις από τη θεωρία. Σε αυτή το κεφάλαιο παρουσιάζεται η μέτρηση διαφορικής ενεργού διατομής του ζεύγους  $t\bar{t}$  στην αδρονική τελική κατάσταση, χρησιμοποιώντας συγκρούσεις  $pp$  σε  $\sqrt{s} = 13\text{ TeV}$  που καταγράφηκαν με τον ανιχνευτή CMS κατά την πλήρη εκτέλεση II του LHC 2016, 2017 και 2018 και αντιστοιχεί σε συνολική φωτεινότητα  $35,9\text{ fb}^{-1}$ ,  $41,5\text{ fb}^{-1}$  και  $59,7\text{ fb}^{-1}$ , αντίστοιχα και αντιστοιχούν σε συνολική φωτεινότητα  $137,1\text{ fb}^{-1}$ . Στο κανάλι αδρονικής διάσπασης, κάθε μποζόνιο  $W$  που προκύπτει από το τοπ κουάρκ διασπάται σε ένα ζεύγος ελαφρ κουάρκ. Ως αποτέλεσμα, η τελική κατάσταση αποτελείται από τουλάχιστον έξι παρτόνια (περισσότερα είναι πιθανά λόγω της ακτινοβολίας αρχικής και τελικής κατάστασης), δύο από τα οποία είναι  $b$  κουάρκ. Λόγω της υψηλής ώθησης (boost) που λαμβάνεται υπόψη σε αυτή τη μέτρηση ( $p_T > 450\text{ GeV}$ ), τα τοπ κουάρκ ανακατασκευάζονται ως πίδακες μεγάλης ακτίνας και η τελική κατάσταση αποτελείται από τουλάχιστον δύο τέτοιους ενεργειακούς πίδακες.

Επίσης στο ίδιο κεφάλαιο παρουσιάζεται ένας συνδυασμός αναζητήσεων για βαρείς συντονισμούς που διασπώνται σε ζεύγη τοπ-αντιτοπ κουάρκ, χρησιμοποιώντας το πλήρες σύνολο δεδομένων σύγκρουσης πρωτονίου-πρωτονίου που συλλέγονται σε ενέργεια κέντρου μάζας  $\sqrt{s} = 13\text{ TeV}$  στο Full Run II, που αντιστοιχεί σε μια ολοκληρωμένη φωτεινότητα  $137\text{ fb}^{-1}$ . Ο συνδυασμός περιλαμβάνει συνεισφορές από τα πλήρως αδρονικά κουάρκ τελικής κατάστασης τοπ-αντιτοπ που έχουν υψηλή εγκάρσια ορμή ( $p_T > 450\text{ GeV}$ ). Τα αποτελέσματα που παρουσιάζονται εδώ αυξάνουν περαιτέρω τα όρια μαζικού αποκλεισμού για πολλά μοντέλα νέας φυσικής.

## 10.1 Ανάλυση Ζεύγους τοπ-αντιτοπ κουαρκ

### 10.1.1 Στρατηγική Ανάλυσης

Η επιλογή βασικής γραμμής, συνοψίζεται στον Πίνακα 6.4

είναι κοινή για όλες τις περιοχές που χρησιμοποιούνται στην ανάλυση. Απαιτούνται τουλάχιστον δύο πίδακες μεγάλου  $P$  ( $R = 0, 8$ ) στην περίπτωση με τοπ πίδακα  $p_T > 450\text{ GeV}$  και δεύτερο τοπ πίδακα με  $p_T > 400\text{ GeV}$  και τιμές μάζας softdrop εντός του εύρους (50, 300) GeV. Επιπλέον, εφαρμόζεται βέτο στα λεπτόνια προκειμένου να ελαχιστοποιηθεί η πιθανότητα επιλέξετε λεπτονικές διασπάσεις κορυφής.

Πάνω από την επιλογή γραμμής βάσης, ορίζουμε τέσσερις συγκεκριμένες περιοχές με βάση την έξοδο του BDT (Boosted Decision Tree) που διακρίνει το σήμα  $t\bar{t}$  από το φόντο  $QCD$ , τις μάζες  $softdrop$  των πίδακα και τον αριθμό των με ετικέτα  $b$   $subjet$  σε κάθε πίδακα, που εξυπηρετούν διαφορετικούς σκοπούς ανάλυσης. Η περιοχή σήματος ( $SR$ ) είναι όπου εκτελούμε τις διαφορικές μετρήσεις και απαιτεί και τα δύο πίδακες για να έχουν ένα  $subjet$  με ετικέτα  $b$  και μια πιο περιορισμένη επιλογή της τιμής  $softdrop$  μάζας πίδακα  $120 - 220 GeV$ . Η καθαρή περιοχή ελέγχου  $QCD$  ( $CR$ ) έχει τα ίδια κριτήρια επιλογής με την προαναφερθείσα περιοχή σήματος, με τη μόνη διαφορά ότι η απαίτηση  $b$ -ετικέτας είναι ανεστραμμένο (οι πίδακες δεν πρέπει να περιέχουν  $subjet$  με ετικέτα  $b$ ). Αυτή η περιοχή χρησιμοποιείται για τη λήψη του σχήματος του φόντου  $QCD$  από τα δεδομένα (μέθοδος βάσει δεδομένων) για το καθένα μεταβλητή ενδιαφέροντος. Τέλος, ορίζεται η περιοχή σήματος  $A$  ( $SR_A$ ). Αυτή είναι η περιοχή όπου εκτελείται η προσαρμογή και χρησιμοποιείται για τον προσδιορισμό της κανονικοποίησης του φόντου  $QCD$ . Το  $SR_A$  είναι ένα υπερσύνολο της περιοχής σήματος που έχει ακριβώς τις ίδιες περικοπές αλλά πιο χαλαρή μάζα  $[50 - 300] GeV$ .

### 10.1.2 $t\bar{t}$ Εξαγωγή Σήματος

Το κυρίαρχο υπόβαθρο σε αυτήν την ανάλυση είναι η παραγωγή πιδάκων  $QCD$ . Υπάρχει μια πεπερασμένη πιθανότητα ότι οι συνηθισμένοι πίδακες, που προέρχονται από ακτινοβολία ενός παρτονίου, μιμούνται την τοπολογική υποδομή ενός τοπ κουαρκ. Η παραγωγή  $QCD$  multijet καταστέλλεται σημαντικά χρησιμοποιώντας το ανεπτυγμένο Boosted Decision Tree, το οποίο χρησιμοποιεί πληροφορίες υποδομής πιδάκων για να διακρίνει το σήμα από το υπόβαθρο. Επίσης, οι απαιτήσεις του  $b$ -tagging παίζουν τεράστιο ρόλο στην καταστολή της παραγωγής  $QCD$  multijet.

Προκειμένου να εκτιμηθεί η υπολειπόμενη συνεισφορά της παραγωγής  $QCD$  multijet, χρησιμοποιούμε μια τεχνική που βασίζεται στην υπόθεση ότι εάν απαιτήσουμε στο σύνολο των δεδομένων ανάλυσης να υπάρχουν 0  $b$ -tagged αντικείμενα, λαμβάνουμε ένα καθαρό δείγμα χωρίς  $QCD$  multijet. Ο φασικός χώρος υποβάθρου Control Region (CR), όπου υπάρχουν μόνο  $QCD$  multijet αποτελείται από δεδομένα  $QCD$  multijet αλλά και από λιγότερο ενεργητικά υπόβαθρα όπως για παράδειγμα συνεισφορά από  $W$ +jets.

### 10.1.3 Συστηματικές αβεβαιότητες

Οι συστηματικές αβεβαιότητες που εξετάζονται σε αυτήν την ανάλυση χωρίζονται σε δύο κατηγορίες: πειραματικές και θεωρητικές. Οι πρώτες περιλαμβάνουν όλες τις αβεβαιότητες που σχετίζονται με τις διαφορές στην απόδοση του αντικειμένου μεταξύ δεδομένων και προσομοίωσης. Οι θεωρητικές σχετίζονται με την ίδια την προσομοίωση και επηρεάζουν πρωτίστως τα τελικά αποτελέσματα στον παρτονικό φασικό χώρο κυρίως λόγω του  $migration$  matrix. Η παρακάτω λίστα περιγράφει εν συντομία αυτές τις αβεβαιότητες και τον τρόπο με τον οποίο έχουν αντιμετωπιστεί σε αυτή την ανάλυση. Θα πρέπει να σημειωθεί ότι για κάθε συστηματική μεταβολή οι διαφορικές διατομές επαναμετρώνται και η διαφορά ως προς το ονομαστικό αποτέλεσμα λαμβάνεται ως αποτέλεσμα αυτής της μεταβολής στη μέτρηση. Αξίζει να σημειωθεί ότι η ίδια διαδικασία που χρησιμοποιείται για την εφαρμογή των συστηματικών αβεβαιοτήτων στις υπολογιζόμενες τιμές διατομής εφαρμόζεται και στις θεωρητικές τιμές. Αυτό ισχύει μόνο για τις Θεωρητικές αβεβαιότητες που περιγράφονται παρακάτω.

#### Πειραματικές αβεβαιότητες

- Πρόβλεψη υποβάθρου  $QCD$
- *Κλίμακα αεριοθούμενων ενέργειας (JES)*
- *Ανάλυση ενέργειας πίδακα*
- *Αποτελεσματικότητα προσθήκης ετικετών υποκειμένου  $b$  (αδρονική)*
- *Pileup*
- *Trigger*

- Φωτεινότητα

### Θεωρητικές αβεβαιότητες

Οι θεωρητικές αβεβαιότητες χωρίζονται σε δύο υποκατηγορίες: σε αυτές που σχετίζονται με το στοιχείο μήτρας της σκληρής διαδικασίας και σε αυτές που σχετίζονται με τη μοντελοποίηση του ντους parton και του υποκείμενου συμβάντος. Πρακτικά, η πρώτη κατηγορία (αποτελούμενη από τις τρεις πρώτες πηγές παρακάτω) αξιολογείται με παραλλαγές των βαρών συμβάντων LHE που αποθηκεύονται στην ονομαστική προσομοίωση MC, ενώ η δεύτερη κατηγορία αξιολογείται με αποκλειστικά, εναλλακτικά δείγματα MC.

- *Συναρτήσεις διανομής Παρτον:*
- *Κλίμακες επανακανονικοποίησης και παραγοντοποίησης*
- *Ισχυρή σταθερά σύζευξης ( $\alpha_S$ )*
- *Ακτινοβολία αρχικής και τελικής κατάστασης (ISR, FSR)*
- *Στοιχείο μήτρας - αντιστοίχιση parton shower*
- *Συντονισμός υποκείμενου συμβάντος*

#### 10.1.4 Συνδυασμός διαφορετικών ετών

Ο στόχος αυτής της ανάλυσης είναι να συνδυάσει τις μετρήσεις κάθε μεμονωμένου έτους σε μια ενιαία «συνδυασμένη» μέτρηση. Ο συνδυασμός γίνεται με τον εξής τρόπο: Για κάθε έτος και για κάθε αβεβαιότητα, υπολογίζεται χωριστά το σήμα ( $D(x)$ ). Ταυτόχρονα υπολογίζονται οι τιμές efficiency και acceptance. Στη συνέχεια, συνδυάζουμε τις ετήσιες τιμές για κάθε πηγή. Με αυτόν τον τρόπο καταλήγουμε σε μία ονομαστική τιμή που είναι το άθροισμα των τεσσάρων μεμονωμένων κεντρικών τιμών και μία τιμή για κάθε αβεβαιότητα που είναι επίσης το άθροισμα των τεσσάρων μεμονωμένων τιμών. Κατά την προσθήκη των πηγών αβεβαιότητας λαμβάνεται υπόψη η συσχέτιση μεταξύ των ετήσιων μετρήσεων λογαριασμό όπως φαίνεται στον Πίνακα 5.22.

#### 10.1.5 Unfolding

Η μέτρηση μεταφέρεται μέσω της διαδικασίας του unfolding από τον φασικό χώρο των ανακατασκευασμένων αντικειμένων, στους φασικούς χώρους των παρτονίων και των ανιχνευτών. Η “ξεδιπλωμένη” διαφορική ενεργός διατομή σε κάθε φασικό χώρο αντίστοιχα, ορίζεται ως:

$$\frac{d\sigma_i^{\text{uv}\Phi}}{dx} = \frac{1}{\mathcal{L} \cdot \Delta x_i} \cdot \frac{1}{f_{2,i}} \cdot \sum_j \left( R_{ij}^{-1} \cdot f_{1,j} \cdot S_j \right), \quad (10.1)$$

όπου  $\mathcal{L}$  είναι η συνολική φωτεινότητα για όλα τα έτη και  $\Delta x_i$  είναι το πλάτος του  $i$ -ου bin της αντίστοιχης μεταβλητής  $x$ . Η ποσότητα  $f_{1,j}$ , είναι το κλάσμα των ανακατασκευασμένων γεγονότων στο  $j$ -ο bin που έχουν ένα κοινό συμβάν στο ξεδιπλωμένο επίπεδο, ενώ η ποσότητα  $f_{2,i}$ , ορίζεται ως το κλάσμα των γεγονότων στο unfolded επίπεδο που έχουν ένα κοινό ανακατασκευασμένο συμβάν. Τα σχήματα 5.62 και 5.63 δείχνουν τα κλάσματα που καθορίστηκαν προηγουμένως. Η ποσότητα  $R_{ij}^{-1}$  ορίζεται ως ο αντίστροφος πίνακας “μετανάστευσης” ή αλλιώς μιγατρικόν ματριξ, ο οποίος απεικονίζει το σύνολο των γεγονότων από δύο διαφορετικούς φασικούς χώρους που συμβαδίζουν, καθώς και τον αριθμό των γεγονότων των οποίων η μέτρηση στον παρτονικό ή στον χώρο των ανιχνευτών αντιστοιχεί σε αρκετά διαφορετική τιμή στον χώρο ανακατασκευής γεγονότων. Ο πίνακας είναι μη διαγώνιος, λόγω της πεπερασμένης ανάλυσης του ανιχνευτή όπως απεικονίζεται στα σχήματα 5.64 και 5.65.

Τα αποτελέσματα περιλαμβάνουν επίσης σύγκριση με τα θεωρητικά μοντέλα. Χρησιμοποιούνται δύο θεωρητικά μοντέλα, το POWHEG με RYTHIA 8 καθώς και το aMC@NLO[FxFx] με το RYTHIA 8. Η σύγκριση παρουσιάζεται με τη μορφή τιμών  $\chi^2$  που υπολογίζονται χρησιμοποιώντας τον πίνακα συνδιακύμανσης (covariance matrix). Χρησιμοποιώντας τις τιμές  $\chi^2$  και τον αριθμό βαθμών ελευθερίας

(NDOF ) υπολογίζονται επίσης τα p-values . Οι αβεβαιότητες λαμβάνονται επίσης υπόψη στον υπολογισμό των τιμών  $\chi^2$  καθώς ενσωματώνονται στον πίνακα συνδιακύμανσης. Οι τιμές  $\chi^2$  υπολογίζονται χρησιμοποιώντας

$$\chi^2 = V_{N_b}^T \cdot C_{N_b}^{-1} \cdot V_{N_b} \quad (10.2)$$

όπου  $V_{N_b}$  είναι το διάνυσμα των διαφορών μεταξύ των μετρούμενων τιμών και των αντίστοιχων θεωρητικών προβλέψεων. Το  $N_b$  είναι το μήκος του διανύσματος και αντιστοιχεί στον αριθμό των bins για κάθε μέτρηση. Το  $C_{N_b}^{-1}$  είναι ο αντίστροφος του πίνακα συνδιακύμανσης. Ο πίνακας συνδιακύμανσης υπολογίζεται χρησιμοποιώντας τον ακόλουθο τύπο

$$C = C_{stat} + C_{syst} \quad (10.3)$$

όπου  $C_{stat}$  είναι ο πίνακας συνδιακύμανσης για τις στατιστικές αβεβαιότητες και  $C_{syst}$  είναι ο πίνακας συνδιακύμανσης για τα θεωρητικά και συστηματικά σφάλματα. Ο πίνακας συνδιακύμανσης για τις συστηματικές αβεβαιότητες υπολογίζεται ως:

$$C_{ij}^{syst} = \sum_{k,l} \frac{1}{N_k} C_{j,k,l} C_{i,k,l}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N, \quad (10.4)$$

όπου  $C_{i,k,l}$  είναι η συστηματική αβεβαιότητα στο  $i$  bin της Ιπαραλλαγής της κπηγής και  $N_k$  είναι ο αριθμός των παραλλαγών για την πηγή  $k$  . Το άθροισμα καλύπτει όλες τις πηγές συστηματικών αβεβαιοτήτων και τις παραλλαγές τους.

### 10.1.6 Τελικά αποτελέσματα

Τα αποτελέσματα της ξεδιπλωμένης μέτρησης σε επίπεδο παρτον φαίνονται στο Εικ. 5.68- 5.76. Η σύγκριση με τις θεωρητικές προβλέψεις ακολουθεί τις τάσεις που παρατηρούνται στην πιστοληπτική μέτρηση. Δηλαδή, υπάρχει 20-40% χαμηλότερη συμπεριλαμβανομένης διατομής, ενώ τα σχήματα των διαφορικών κατανομών αναπαράγονται εύλογα από όλα τα μοντέλα.

Τα αποτελέσματα της ξεδιπλωμένης μέτρησης σε επίπεδο σωματιδίων φαίνονται στο Εικ. 5.77- 5.85. Η σύγκριση με τις θεωρητικές προβλέψεις ακολουθεί τις τάσεις που παρατηρούνται στην πιστοληπτική μέτρηση. Δηλαδή, υπάρχει 20-40% χαμηλότερη συμπεριλαμβανομένης διατομής, ενώ τα σχήματα των διαφορικών κατανομών αναπαράγονται εύλογα από όλα τα μοντέλα.

## 10.2 Ζ'

Το τοπ κουάρκ ολοκληρώνει την τρίτη γενιά κουάρκ στο Καθιερωμένου μοντέλο (SM) και η ακριβής γνώση των ιδιοτήτων του είναι κρίσιμη για τη συνολική κατανόηση της θεωρίας. Οι μετρήσεις της διατομής παραγωγής του ζεύγους τοπ-αντιτοπ κουάρκ ( $t\bar{t}$ ) επιβεβαιώνουν τις προβλέψεις της κβαντική χρωμοδυναμικής θεωρίας ( $X^{\Delta}$ ) και έχουν τη δυνατότητα να περιορίσουν τις  $X^{\Delta}$  παραμέτρους, ενώ τέτοιες μετρήσεις είναι ευαίσθητες στη φυσική πέρα από το SM. Επίσης, η διαδικασία παραγωγής  $t\bar{t}$  είναι ένα κυρίαρχο υπόβαθρο SM για την αναζήτηση νέων φαινομένων φυσικής και επομένως Η ακριβής γνώση είναι απαραίτητη για νέες ανακαλύψεις.

Πολυάριθμες επεκτάσεις του Καθιερωμένου μοντέλου (SM) προβλέπουν την ύπαρξη αλληλεπιδράσεων με κουάρκ τρίτης γενιάς, ειδικά το τοπ κουάρκ. Το σχετικό νέο σωματίδιο που περιέχεται σε αυτές τις θεωρίες θα μπορούσε να παρατηρηθεί ως συντονισμός του  $t\bar{t}$  σε πειράματα στο  $LHC$ . Παραδείγματα τέτοιων συντονισμών είναι έγχρωμα μονόχρωμα μποζόνια τύπου  $Z$  ( $Z'$ ) σε extended gauge θεωρίες , colorons και axigluons σε μοντέλα με εκτεταμένους ισχυρούς τομείς αλληλεπίδρασης, βαρύτερα αδέρφια Higgs σε μοντέλα με εκτεταμένους τομείς Higgs και διεγέρσεις Kaluza-Klein (KK) γκλουονίων , electroweak gauge bosons και γκραβιτόνια σε διάφορες επεκτάσεις του μοντέλου Randall-Sundrum (RS) . Αυτά τα μοντέλα προβλέπουν την ύπαρξη συντονισμών κλίμακας  $TeV$  με διατομές παραγωγής μερικών πισοβαρής σε  $\sqrt{s} = 13 TeV$ . Σε όλα αυτά τα παραδείγματα, η παραγωγή συντονισμού  $t\bar{t}$  θα μπορούσε να παρατηρηθεί στον εκθέτη της διαφοράς απόλυτης ταχύτητας μεταξύ των δύο κορυφαίων τοπ κουαρκ, την κινηματική μεταβλητή  $\chi$ .



Η ανάλυση αυτή βασίζεται στα κριτήρια επιλογής και στην ανάλυση που περιγράφεται στο προηγούμενο κεφάλαιο. Με βάση τα προηγούμενα, στην ανάλυση αυτή που ενδιαφέρει για αυτήν την αναζήτηση είναι η μεταβλητή  $\chi$  dijet . Επιλέγουμε αυτή τη μεταβλητή επειδή οι κατανομές της σχετίζονται με τις τελικές καταστάσεις που παράγονται μέσω αλληλεπιδράσεων  $X\Delta$  είναι σχετικά επίπεδες σε σύγκριση με τις κατανομές των μοντέλων Beyond Standard Model (BSM) ή των νέων σωματιδίων, που συνήθως κορυφώνονται σε χαμηλές τιμές  $\chi$ . Η επιλογή αυτής της μεταβλητής, πηγάζει από το γεγονός ότι η μάζα dijet του ζεύγους  $t\bar{t}$  υποφέρει από αβεβαιότητες της Ενεργειακής Κλίμακας, ενώ η μεταβλητή γωνιακής dijet  $\chi$  δεν επηρεάζεται.

Η μεταβλητή υπολογίζεται με τη μέτρηση της διαφοράς των rapidities των δύο τοπ κουάρκ στο πλαίσιο μηδενικής ορμής:

$$y^* = \frac{1}{2}(y_1 - y_2) \quad (10.5)$$

όπου  $y_{1,2}$  είναι οι rapidities του κύριου και δευτερεύοντος τοπ κουάρκ αντίστοιχα. Η μεταβλητή  $\chi$  ορίζεται ως:

$$\chi = e^{|2y^*|} = e^{|y_1 - y_2|}. \quad (10.6)$$

### 10.2.1 Στρατηγική Ανάλυσης

Προκειμένου να ενισχυθεί περαιτέρω η ύπαρξη του σήματος  $Z$  εντός της προτεινόμενης περιοχής σήματος, εφαρμόστηκαν διαφορετικά σεντ επιλογής μάζας dijet για κάθε μάζα  $Z$  αντίστοιχα. Παρατηρήσαμε ότι σε χαμηλότερες μάζες, η γωνιακή κατανομή του  $\chi$  είναι πιο ανταγωνιστική σε σχέση με την dijet κατανομή μάζας  $m_{JJ}$ . Για να ποσοτικοποιηθεί η βέλτιστη επιλογή μάζας dijet , υπολογίζουμε τη Significance ανά  $Z$  μάζα. Ορίζουμε ως Significance την ακόλουθη σχέση :

$$\text{Significance} = \frac{\text{Signal}}{\sqrt{\text{Bkg} + \text{Signal}}} \quad (10.7)$$

Με βάση την επιλογή του κατάλληλου  $m_{JJ}$  cut που καταλήγει σε μέγιστο Significance, επιβάλλονται οι περιορισμοί dijet μάζας όπως φαίνεται στον Πίνακα 10.1:

Mass (GeV)	Width (1%) (GeV)	$m_{JJ}$ Cut (GeV)
1200	12	1000
1400	14	1200
1600	16	1600
1800	18	1600
2000	20	1600
2500	25	2000
3000	30	2000
4000	40	2000
4500	45	2000

Πίνακας 10.1: Μεταβλητό  $m_{JJ}$  cut για όλες τις μάζες των διάφορων  $Z$ .

### 10.2.2 Ανάλυση Ευαισθησίας

Σε αυτήν την ενότητα, εκτελείται μια ανάλυση ευαισθησίας (Sensitivity Analysis ) προκειμένου να ανιχνευθούν διαφορές στο σχήμα μεταξύ των κατανομών  $t\bar{t}$  και  $Z$  σε σχέση με την μεταβλητή  $\chi$ . Και τα δύο αρχεία που απεικονίζονται στις κατανομές, προέρχονται από συμβάντα που δημιουργούνται από Μόντε Κάρλο προσομοιώσεις.

### 10.2.3 Αποτελέσματα

Σε αυτή την ενότητα, παρέχονται τα συνδυασμένα αποτελέσματα. Ο άξονας  $y$  αντιπροσωπεύει τα επίπεδα εμπιστοσύνης για την τιμή του λόγου της μετρούμενης διατομής έναντι της προβλεπόμενης τιμής που προβλέπεται από το Καθιερωμένο Πρότυπο.

Η μέθοδος Asymptotic Limits επιτρέπει τον γρήγορο υπολογισμό μιας εκτίμησης των παρατηρούμενων και αναμενόμενων ορίων, η οποία είναι αρκετά ακριβής όταν τα συνολικά γεγονότα δεν είναι πολύ μικρά και οι συστηματικές αβεβαιότητες δεν παίζουν σημαντικό ρόλο στο αποτέλεσμα. Ο υπολογισμός του ορίου βασίζεται σε μια ασυμπτωτική προσέγγιση των κατανομών της στατιστικής δοκιμής LHC, η οποία βασίζεται στον υπολογισμό δύο τιμών  $p, p_\mu, p_b$  και επομένως  $CL_s = \frac{p_\mu}{(1-p_b)}$ , όπου  $CL_s$  είναι η ασυμπτωτική προσέγγιση των ορίων υπολογισμού.

Για αυτήν την ανάλυση, χρησιμοποιείται το στατιστικό εργαλείο *Combine* [παραπομπή εδώ] που παρέχεται από το CMS. Αυτό το πλαίσιο παρέχει στο χρήστη τη δυνατότητα να τρέξει μια ανάλυση σε ασυμπτωτικά όρια. Το πρόγραμμα εκτυπώνει το όριο στην ισχύ του σήματος  $\rho$  που ορίζεται ως (αριθμός συμβάντων σήματος / αριθμός αναμενόμενων συμβάντων σήματος) π.χ. Παρατηρούμενο όριο:  $r < 1,6297, 95\%CL$ , το μέσο αναμενόμενο όριο Αναμενόμενο 50,0% :  $r < 2,3111$  και άκρα του εύρους 68% και 95% για τα αναμενόμενα όρια. Από προεπιλογή, τα όρια υπολογίζονται χρησιμοποιώντας τη συνταγή CLs, η οποία παίρνει την αναλογία των τιμών  $\rho$  κάτω από την υπόθεση σήματος συμπεριλαμβανομένου του υποβάθρου και μόνο συμπεριλαμβανομένου του υποβάθρου.

Η ανάλυση γίνεται αρχικά ξεχωριστά για κάθε έτος. Για κάθε μεταβολή μάζας  $Z$  συντίθεται μια *datacard*. Η *datacard* περιλαμβάνει όλες τις πληροφορίες που χρειάζεται το εργαλείο *Combine* για την ανάλυση, όπως αρχείο δεδομένων, αρχεία σήματος, αρχεία υποβάθρου καθώς και παραλλαγές αυτών. Σε αυτήν την ανάλυση, το αρχείο *Signal* αναφέρεται στα δεδομένα  $Z$ , ενώ το  $t\bar{t}$ , το QCD και το Subdominant Bkg (Single top,  $tW$ +jets) ορίζονται ως υπόβαθρο. Όσον αφορά τις παραλλαγές, λαμβάνονται υπόψη μόνο οι παραλλαγές του  $t\bar{t}$  καθώς αποτελούν το μεγαλύτερο υπόβαθρο σε αυτόν τον φασικό χώρο.

Για κάθε διακύμανση μάζας  $Z$  η ανάλυση εκτελείται με βάση την επιλεγμένη αντιστοιχισμένη μάζα που εμφανίζεται στο 10.1. Όλα τα αποτελέσματα συνδυάζονται σε ένα γράφημα. Κάθε γράφημα περιλαμβάνει τα παρατηρούμενα και αναμενόμενα ανώτερα όρια, συμπεριλαμβανομένων των συστηματικών παραλλαγών.

Προκειμένου να αναλυθούν τα αποτελέσματα όλων των ετών μαζί, είναι απαραίτητο να συνδυαστούν οι κάρτες δεδομένων (*datacard*) τους σε μία ενιαία. Αυτό γίνεται λαμβάνοντας υπόψη το αρχείο *combineCards.py* που επιτρέπει τον συνδυασμό πολλών *datacards*. Οι συνδυασμένες γραφικές παραστάσεις (Brazilian plots) εμφανίζονται στο ::

## 10.3 Συμπεράσματα

Σε αυτή την ανάλυση παρουσιάστηκε μια αναζήτηση για την παραγωγή βαρέων συντονισμών spin -1 ή spin -2 που διασπώνται σε  $t\bar{t}$  ως τελικές καταστάσεις. Τα δεδομένα αντιστοιχούν σε φωτεινότητα  $137,1 fb^{-1}$  και συλλέγονται με τον ανιχνευτή CMS σε συγκρούσεις πρωτονίου-πρωτονίου σε  $\sqrt{s} = 13 TeV$  στον LHC.

Η ανάλυση έχει σχεδιαστεί για να έχει υψηλή ευαισθησία σε μάζες συντονισμού άνω του 1 TeV, όπου τα προϊόντα διάσπασης τελικής κατάστασης ευθυγραμμίζονται λόγω των μεγάλων ωθήσεων Lorentz των τοπ κουάρκ. Δεν βρέθηκαν γεγονότα που δείχνουν συντονισμούς σωματιδίων με μάζα που διασπώνται σε  $t\bar{t}$ . Τα όρια στο 95%  $\Lambda$  ορίζονται στη διατομή παραγωγής νέων σωματιδίων spin -1 που διασπώνται σε  $t\bar{t}$  με σχετικά πλάτη διάσπασης που είναι είτε στενά είτε μεγάλα σε σύγκριση με την ανάλυση του ανιχνευτή.

Επιπλέον, τίθενται όρια στην παραγωγή σωματιδίων σε μοντέλα αναφοράς πέρα από το Καθιερωμένο Πρότυπο. Τα μποζόνια  $Z$  με σχετικά πλάτη  $\Gamma/M$  1% εξαιρούνται για εύρη μάζας 1,0–4,0 TeV, αντίστοιχα. Αυτή η αναζήτηση παρουσιάζει όρια στα μποζόνια  $Z$  ως συνάρτηση του σχετικού πλάτους του συντονισμού στο εύρος 10%.

# Bibliography

- [1] Mary K. Gaillard, Paul D. Grannis, and Frank J. Sciulli. The standard model of particle physics. *Reviews of Modern Physics*, 71(2):S96-S111, mar 1999.
- [2] Samoil M. Bilenky and J. Hosek. GLASHOW-WEINBERG-SALAM THEORY OF ELECTROWEAK INTERACTIONS AND THE NEUTRAL CURRENTS. *Phys. Rept.*, 90:73-157, 1982.
- [3] S. Chatrchyan et al +. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30-61, 2012.
- [4] G. Aad et al. Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1-29, sep 2012.
- [5] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.
- [6] TIM GERSHON. Overview of the cabibbo-kobayashi-maskawa matrix†. *Pramana*, 79(5):1091-1108, nov 2012.
- [7] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.
- [8] Axel Maas. Brout-englert-higgs physics: From foundations to phenomenology. *Progress in Particle and Nuclear Physics*, 106:132-209, may 2019.
- [9] S. Abachi et al. Observation of the top quark. *Physical Review Letters*, 74(14):2632-2637, apr 1995.
- [10] Haim Harari. (t/e) three generations of quarks and leptons, 1977.
- [11] CDF Collaboration. Observation of top quark production in  $\bar{p}p$  collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74:2626-2631, Apr 1995.
- [12] CMS Collaboration. Search for resonant  $t\bar{t}$  production in proton-proton collisions at  $\sqrt{s} = 8$  tev. *Physical Review D*, 93(1), jan 2016.
- [13] CMS Collaboration. Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at  $\sqrt{s}=13$  TeV. *The European Physical Journal C*, 80(2), jan 2020.
- [14] Hermann Kolanoski and Norbert Wermes. Introduction. In *Particle Detectors: Fundamentals and Applications*. Oxford University Press, 06 2020.
- [15] Jiri Prochazka. Elastic proton-proton collisions at high energies, Feb 2018. Presented 25 Sep 2018.
- [16] W. Xiong, A. Gasparian, H. Gao, D. Dutta, M. Khandaker, N. Liyanage, E. Pasyuk, C. Peng, X. Bai, L. Ye, K. Gnanvo, C. Gu, M. Levillain, X. Yan, D. W. Higinbotham, M. Meziane, Z. Ye, K. Adhikari, B. Aljawrneh, H. Bhatt, D. Bhetuwal, J. Brock, V. Burkert, C. Carlin, A. Deur, D. Di, J. Dunne, P. Ekanayaka, L. El-Fassi, B. Emmich, L. Gan, O. Glamazdin, M. L. Kabir, A. Karki, C. Keith, S. Kowalski, V. Lagerquist, I. Larin, T. Liu, A. Liyanage,

## BIBLIOGRAPHY

- J. Maxwell, D. Meekins, S. J. Nazeer, V. Nelyubin, H. Nguyen, R. Pedroni, C. Perdrisat, J. Pierce, V. Punjabi, M. Shabestari, A. Shahinyan, R. Silwal, S. Stepanyan, A. Subedi, V. V. Tarasov, N. Ton, Y. Zhang, and Z. W. Zhao. A small proton charge radius from an electron-proton scattering experiment. *Nature*, 575(7781):147–150, Nov 2019.
- [17] CMS Collaboration. Pileup mitigation at CMS in 13 TeV data. *Journal of Instrumentation*, 15(09):P09018–P09018, sep 2020.
- [18] Luigi Di Lella and Carlo Rubbia. The Discovery of the W and Z Particles. *Adv. Ser. Direct. High Energy Phys.*, 23:137–163, 2015.
- [19] Paul Langacker. The Physics of Heavy  $Z'$  Gauge Bosons. *Rev. Mod. Phys.*, 81:1199–1228, 2009.
- [20] Daniel Hayden, Raymond Brock, and Christopher Willis. *Z prime: A story*, 2013.
- [21] CMS Collaboration. Search for  $t\bar{t}$  resonances in highly boosted leptonjets and fully hadronic final states in proton-proton collisions at  $\sqrt{s}=13$  TeV. *Journal of High Energy Physics*, 2017(7), jul 2017.
- [22] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [23] O. Bruning, H. Burkhardt, and S. Myers. The Large Hadron Collider. *Prog. Part. Nucl. Phys.*, 67:705–734, 2012.
- [24] David Alesini. *Linac*, 2021.
- [25] J. B. ADAMS. The cern proton synchrotron. *Nature*, 185(4713):568–572, Feb 1960.
- [26] The Super Proton Synchrotron. 2012.
- [27] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.
- [28] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.
- [29] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.
- [30] V Veszpremi. Operation and performance of the CMS tracker. *Journal of Instrumentation*, 9(03):C03005–C03005, mar 2014.
- [31] V. Karimäki. The CMS tracker system project: Technical Design Report. 1997.
- [32] Aaron Dominguez. The cms pixel detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 581(1):343–346, 2007. VCI 2007.
- [33] CMS Collaboration. The cms phase-1 pixel detector upgrade. *JINST*, 16:P02027. 84 p, Feb 2021.
- [34] J.-L. Agram. Cms silicon strip tracker performance. *Physics Procedia*, 37:844–850, 2012. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).
- [35] M Friedl. The CMS silicon strip tracker and its electronic readout, 2001.
- [36] CMS Collaboration. Performance and operation of the CMS electromagnetic calorimeter. *Journal of Instrumentation*, 5(03):T03010–T03010, mar 2010.
- [37] D. Petyt. *ELECTROMAGNETIC CALORIMETER (ECAL)*. 2013.

## BIBLIOGRAPHY

- [38] Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at  $\sqrt{s}=7$  TeV. *Journal of Instrumentation*, 8(09):P09009–P09009, sep 2013.
- [39] Kerem Cankocak, Pawel de Barbaro, Dima Vishnevskiy, Yasar Onel, and CMS-HCAL Collaboration. CMS HCAL installation and commissioning. *Journal of Physics: Conference Series*, 160:012055, apr 2009.
- [40] CMS Collaboration. Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data. *Journal of Instrumentation*, 5(03):T03012–T03012, mar 2010.
- [41] M. C. Fouz. The CMS muon system. *Nucl. Instrum. Meth. A*, 446:366–372, 2000.
- [42] M. Shopova, A. Aleksandrov, R. Hadjiiska, P. Iaydjiev, G. Sultanov, M. Rodozov, S. Stoykova, Y. Assran, A. Sayed, A. Radi, S. Aly, G. Singh, M. Abbrescia, G. Iaselli, M. Maggi, G. Pugliese, P. Verwilligen, W. Van Doninck, S. Colafranceschi, A. Sharma, L. Benussi, S. Bianco, D. Piccolo, F. Primavera, A. Cimmino, S. Crucy, A. A. O. Rios, M. Tytgat, N. Zaganidis, M. Gul, A. Fagot, V. Bhatnagar, J. Singh, R. Kumari, A. Mehta, A. Ahmad, I. M. Awan, H. Shahzad, H. Hoorani, M. I. Asghar, S. Muhammad, W. Ahmed, M. A. Shah, S. W. Cho, S. Y. Choi, B. Hong, M. H. Kang, K. S. Lee, J. H. Lim, S. K. Park, M. S. Kim, I. B. Laktineh, F. Lagarde, M. Gouzevitch, G. Grenier, I. Pedraza, S. Carpineteyro Bernardino, C. Uribe Estrada, S. Carrillo Moreno, F. Vazquez Valencia, L. M. Pant, S. Buontempo, N. Cavallo, F. Fabozzi, I. Orso, L. Lista, S. Meola, M. Merola, P. Paolucci, F. Thyssen, G. Lanza, M. Esposito, A. Braghieri, A. Magnani, C. Riccardi, P. Salvini, I. Vai, P. Vitulo, P. Montagna, Y. Ban, S. J. Qian, M. Choi, Y. Choi, J. Goh, D. Kim, A. Dimitrov, L. Litov, P. Petkov, B. Pavlov, I. Bagaturia, D. Lomidze, C. Avila, A. Cabrera, J. C. Sanabria, I. Crotty, and J. Vaitkus. Performance of resistive plate chambers installed during the first long shutdown of the cms experiment, 2016.
- [43] D Acosta, G Apollinari, J Blomquist, R Breedon, N Bondar, Yu Bonushkin, E Borissov, A Bujak, B Bylsma, N Chester, D Chrisman, D Cline, S Dolinsky, S Durkin, D Eartly, T Ferguson, F Feyzi, J Gilmore, L Gorn, W Gorn, J Gu, L Gutay, J Hauser, S Hershman II, J Hoftiezer, O Kisselev, W Ko, J Korienek, A Korytov, J Kubic, J Layter, T.Y Ling, R Lovell, S Lusin, C Matthey, M Matveev, S Medved, G Mitselmakher, S Otwinowski, P Padley, F Petriello, Yu Pischalnikov, O Prokofiev, V Razmyslovich, D Reeder, J Roberts, P Robl, C Rush, J.-C Santiard, S Sedov, B Smith, S Sobolev, V Soulimov, B Tannenbaum, N Terentyev, A Vorobyov, and V Yarba. Large cms cathode strip chambers: design and performance. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 453(1):182–187, 2000. Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.
- [44] CMS Collaboration. The CMS trigger system. *Journal of Instrumentation*, 12(01):P01020–P01020, jan 2017.
- [45] V. Khachatryan et al. The CMS trigger system. *Journal of Instrumentation*, 12(01):P01020–P01020, jan 2017.
- [46] Valentina Gori. The CMS high level trigger. *International Journal of Modern Physics: Conference Series*, 31:1460297, jan 2014.
- [47] Tejinder Virdee, Achille Petrilli, and Austin Ball. CMS High Level Trigger. Technical report, CERN, Geneva, 2007. revised version submitted on 2007-10-19 16:57:09.
- [48] A Dominguez, D Abbaneo, K Arndt, N Bacchetta, A Ball, E Bartz, W Bertl, G M Bilei, G Bolla, H W K Cheung, M Chertok, S Costa, N Demaria, Daniel Dominguez Vazquez, K Ecklund, W Erdmann, K Gill, G Hall, K Harder, F Hartmann, R Horisberger, W Johns, H C Kaestli, K Klein, D Kotlinski, S Kwan, M Pesaresi, H Postema, T Rohe, C Schäfer, A Starodumov, S Streuli, A Tricomi, P Tropea, J Troska, F Vasey, and W Zeuner. CMS Technical Design Report for the Pixel Detector Upgrade. Technical report, 2012. Additional contacts: Jeffrey

## BIBLIOGRAPHY

- Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.
- [49] D Contardo, M Klute, J Mans, L Silvestris, and J Butler. Technical Proposal for the Phase-II Upgrade of the CMS Detector. Technical report, Geneva, 2015. Upgrade Project Leader Deputies: Lucia Silvestris (INFN-Bari), Jeremy Mans (University of Minnesota) Additional contacts: Lucia.Silvestris@cern.ch, Jeremy.Mans@cern.ch.
- [50] D Contardo, M Klute, J Mans, L Silvestris, and J Butler. Technical Proposal for the Phase-II Upgrade of the CMS Detector. Technical report, Geneva, 2015. Upgrade Project Leader Deputies: Lucia Silvestris (INFN-Bari), Jeremy Mans (University of Minnesota) Additional contacts: Lucia.Silvestris@cern.ch, Jeremy.Mans@cern.ch.
- [51] Florian Beaudette. The cms particle flow algorithm. 2014.
- [52] Ryan Atkin. Review of jet reconstruction algorithms. *Journal of Physics: Conference Series*, 645:012008, oct 2015.
- [53] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *The European Physical Journal C*, 72(3), mar 2012.
- [54] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ik/isubit/i/subjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, apr 2008.
- [55] Andrew J. Larkoski, Simone Marzani, Gregory Soyez, and Jesse Thaler. Soft Drop. *JHEP*, 05:146, 2014.
- [56] Simone Marzani, Lais Schunk, and Gregory Soyez. The jet mass distribution after soft drop. *The European Physical Journal C*, 78(2):96, Feb 2018.
- [57] Pileup Removal Algorithms. Technical report, CERN, Geneva, 2014.
- [58] CMS Collaboration. Pileup mitigation at CMS in 13 TeV data. *Journal of Instrumentation*, 15(09):P09018–P09018, sep 2020.
- [59] Benjamin Kreis. Particle flow and puppi in the level-1 trigger at cms for the hl-lhc, 2018.
- [60] C. Weiser. A combined secondary vertex based B-tagging algorithm in CMS. 1 2006.
- [61] Christian Weiser. A Combined Secondary Vertex Based B-Tagging Algorithm in CMS. Technical report, CERN, Geneva, 2006.
- [62] Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical report, CERN, Geneva, 2016.
- [63] Markus Stoye and. Deep learning in jet reconstruction at CMS. *Journal of Physics: Conference Series*, 1085:042029, sep 2018.
- [64] M. Aaboud et al. Measurements of top-quark pair differential cross-sections in the lepton+jets channel in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector. *JHEP*, 11:191, 2017.
- [65] Morad Aaboud et al. Measurements of top-quark pair differential cross-sections in the  $e\mu$  channel in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector. *Eur. Phys. J.*, C77(5):292, 2017.
- [66] M. Aaboud et al. Measurement of lepton differential distributions and the top quark mass in  $t\bar{t}$  production in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector. *Eur. Phys. J.*, C77(11):804, 2017.

## BIBLIOGRAPHY

- [67] Morad Aaboud et al. Measurement of top quark pair differential cross-sections in the dilepton channel in  $pp$  collisions at  $\sqrt{s} = 7$  and 8 TeV with ATLAS. *Phys. Rev.*, D94(9):092003, 2016.
- [68] Georges Aad et al. Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using pp collisions at  $\sqrt{s} = 7$  TeV in the ATLAS detector. *JHEP*, 06:100, 2015.
- [69] Georges Aad et al. Measurements of top quark pair relative differential cross-sections with ATLAS in  $pp$  collisions at  $\sqrt{s} = 7$  TeV. *Eur. Phys. J.*, C73(1):2261, 2013.
- [70] A. M. Sirunyan et al. Measurement of normalized differential  $t\bar{t}$  cross sections in the dilepton channel from pp collisions at  $\sqrt{s} = 13$  TeV. *JHEP*, 04:060, 2018.
- [71] Measurements of normalised multi-differential cross sections for top quark pair production in pp collisions at  $\sqrt{s} = 13$  TeV and simultaneous determination of the strong coupling strength, top quark pole mass and parton distribution functions. Technical Report CMS-PAS-TOP-18-004, CERN, Geneva, 2018.
- [72] Albert M Sirunyan et al. Measurements of  $t\bar{t}$  differential cross sections in proton-proton collisions at  $\sqrt{s} = 13$  TeV using events containing two leptons. *Submitted to: JHEP*, 2018.
- [73] Vardan Khachatryan et al. Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV. *Phys. Rev.*, D95(9):092001, 2017.
- [74] A. M. Sirunyan et al. Measurements of differential cross sections of top quark pair production as a function of kinematic event variables in proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JHEP*, 06:002, 2018.
- [75] Albert M Sirunyan et al. Measurement of double-differential cross sections for top quark pair production in pp collisions at  $\sqrt{s} = 8$  TeV and impact on parton distribution functions. *Eur. Phys. J.*, C77(7):459, 2017.
- [76] Vardan Khachatryan et al. Measurement of the differential cross section for top quark pair production in pp collisions at  $\sqrt{s} = 8$  TeV. *Eur. Phys. J.*, C75(11):542, 2015.
- [77] Vardan Khachatryan et al. Measurement of the differential cross sections for top quark pair production as a function of kinematic event variables in pp collisions at  $\sqrt{s}=7$  and 8 TeV. *Phys. Rev.*, D94(5):052006, 2016.
- [78] Serguei Chatrchyan et al. Measurement of differential top-quark pair production cross sections in  $pp$  collisions at  $\sqrt{s} = 7$  TeV. *Eur. Phys. J.*, C73(3):2339, 2013.
- [79] Morad Aaboud et al. Measurements of  $t\bar{t}$  differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector. *Phys. Rev.*, D98(1):012003, 2018.
- [80] Georges Aad et al. Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in  $\sqrt{s} = 8$  TeV proton-proton collisions using the ATLAS detector. *Phys. Rev.*, D93(3):032009, 2016.
- [81] Measurement of the  $t\bar{t}$  production cross section at 13 TeV in the all-jets final state. Technical Report CMS-PAS-TOP-16-013, CERN, Geneva, 2016.
- [82] Vardan Khachatryan et al. Measurement of the integrated and differential  $t\bar{t}$  production cross sections for high- $p_t$  top quarks in  $pp$  collisions at  $\sqrt{s} = 8$  TeV. *Phys. Rev.*, D94(7):072002, 2016.

## BIBLIOGRAPHY

- [83] Measurement of the differential  $t\bar{t}$  cross section with high- $p_T$  top-quark jets in the all-hadronic channel at  $\sqrt{s} = 8$  TeV. Technical Report CMS-PAS-TOP-16-018, CERN, Geneva, 2017.
- [84] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [85] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007.
- [86] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [87] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [88] Simone Alioli, S. O. Moch, and P. Uwer. Hadronic top-quark pair-production with one jet and parton showering. *JHEP*, 01:137, 2012.
- [89] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO single-top production matched with shower in POWHEG:  $s$ - and  $t$ -channel contributions. *JHEP*, 09:111, 2009. [Erratum: 10.1007/JHEP02(2010)011].
- [90] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [91] Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008.
- [92] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.
- [93] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, jun 2015.
- [94] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [95] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma Slade, and Maria Ubiali. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10), oct 2017.
- [96] CMS Collaboration. Extraction and validation of a new set of CMS pythia8 tunes from underlying-event measurements. *The European Physical Journal C*, 80(1), jan 2020.
- [97] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [98] Michal Czakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Comput. Phys. Commun.*, 185:2930, 2014.
- [99] Ye Li and Frank Petriello. Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code. *Phys. Rev. D*, 86:094034, 2012.



## BIBLIOGRAPHY

- [100] *Top Quark Production*, 2014.
- [101] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *Journal of High Energy Physics*, 2012(12), dec 2012.
- [102] CMS Collaboration. Measurement of differential  $t\bar{t}$  production cross sections using top quarks at large transverse momenta in pp collisions at  $\sqrt{s} = 13$  tev. *Physical Review D*, 103(5), mar 2021.
- [103] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. [,186(2003)].
- [104] Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12:P02014, 2017.
- [105] Albert M Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13:P05011, 2018.
- [106] M. Aaboud et al. Measurement of the Inelastic Proton-Proton Cross Section at  $\sqrt{s} = 13$  TeV with the ATLAS Detector at the LHC. *Phys. Rev. Lett.*, 117(18):182002, 2016.
- [107] Albert M Sirunyan et al. Precision luminosity measurement in proton-proton collisions at  $\sqrt{s} = 13$  TeV in 2015 and 2016 at CMS. *Eur. Phys. J. C*, 81(9):800, 2021.
- [108] A. M. Sirunyan et al. Precision luminosity measurement in proton-proton collisions at  $\sqrt{s} = 13$  TeV in 2015 and 2016 at CMS. Submitted to *Eur. Phys. J. C*, 2021.
- [109] CMS Collaboration. CMS luminosity measurement for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV. CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.
- [110] CMS Collaboration. CMS luminosity measurement for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV. CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.
- [111] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation>.
- [112] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources>.
- [113] R. Aldrovandi. Extended gauge theories. *J. Math. Phys.*, 32:2503–2512, 1991.
- [114] J Lorenzo Dí az-Cruz, Javier M Hernández-López, and Javier A Orduz-Ducuara. An extraiz/i gauge boson as a source of higgs particles. *Journal of Physics G: Nuclear and Particle Physics*, 40(12):125002, oct 2013.
- [115] E. H. Simmons. Colorons: Theory and phenomenology, 1997.
- [116] Oscar Antuñ ano, Johann H. Kühn, and Germán Rodrigo. Top quarks, axiguons, and charge asymmetries at hadron colliders. *Physical Review D*, 77(1), jan 2008.
- [117] Benjamin C. Allanach, Farvah Mahmoudi, Jordan P. Skittrall, and K. Sridhar. Gluon-initiated production of a kaluza-klein gluon in a bulk randall-sundrum model. *Journal of High Energy Physics*, 2010(3), mar 2010.
- [118] Klaus Moenig. Physics of electroweak gauge boson, 2003.
- [119] Lisa Randall and Raman Sundrum. Large mass hierarchy from a small extra dimension. *Physical Review Letters*, 83(17):3370–3373, oct 1999.
- [120] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026–026, may 2006.

## BIBLIOGRAPHY

- [121] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, jun 2015.
- [122] Michelangelo L Mangano, Mauro Moretti, Fulvio Piccinini, and Michele Treccani. Matching matrix elements and shower evolution for top-pair production in hadronic collisions. *Journal of High Energy Physics*, 2007(01):013–013, jan 2007.
- [123] Paolo Nason. A new method for combining NLO QCD with shower monte carlo algorithms. *Journal of High Energy Physics*, 2004(11):040–040, nov 2004.
- [124] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070–070, nov 2007.
- [125] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6), jun 2010.
- [126] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014.
- [127] Michał Czakon and Alexander Mitov. Top: A program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications*, 185(11):2930–2938, nov 2014.
- [128] Ye Li and Frank Petriello. Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code. *Physical Review D*, 86(9), nov 2012.
- [129] John M. Campbell and R.K. Ellis. MCFM for the tevatron and the LHC. *Nuclear Physics B - Proceedings Supplements*, 205-206:10–15, aug 2010.
- [130] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), feb 2011.
- [131] Geeta Yadav and Kolin Paul. Architecture and security of scada systems: A review, 2020.
- [132] Simatic scada systems.
- [133] Oliver Holme, Manuel Gonzalez-Berges, P. Golonka, and Sascha Schmeling. The jcop framework. *Conf. Proc. C051010:WE*, 2, 01 2005.
- [134] Clara Gaspar. PVSS amp; SMI++ : Tools for the Automation of large distributed control systems. 2005.
- [135] Clara Gaspar and Boda Franek. Tools for the automation of large control systems. Technical report, CERN, Geneva, 2005.
- [136] D. Harrington, R. Presuhn, and B. Wijnen. An architecture for describing simple network management protocol (SNMP) management frameworks. Technical report, December 2002.
- [137] A. M. Sirunyan et al. Measurement of differential  $t\bar{t}$  production cross sections using top quarks at large transverse momenta in pp collisions at  $\sqrt{s} = 13$  TeV. *American Physical Society (APS)*, 5:103, 2021.

## Appendix A

# Top Tagger Efficiency

In this section we present the validation of the in-house developed top tagger for this analysis.

The top tagger is validated using the tag and probe method. To evaluate this methodology, we further divide the phase space into sub-regions.

- Probe or  $SR_B$ : Baseline selection + tight Mass Cut (120, 220) GeV, no top-tagger selection
- Tight Region: Baseline selection + tight Mass Cut (120, 220) GeV, tight top-tagger cut (0.8)

The top tagger has been developed for all years that are studied in this analysis. The calculated values are presented for the UL files for all respective years (2016\_preVFP, 2016\_postVFP, 2017 and 2018).

We compute the efficiency as follows: We calculate the numerator, using jets that fulfill the following requirements: [1 jet passing ALL SR requirements (+btagging) + Tight TopTagger Cut (0.8)] & [1 jet passing ALL SR requirements (+btagging) with loose topTaggerCut (eg.  $>0.2$  for 2016)] . The denominator is calculated as follows: [1 jet passing ALL SR requirements (+btagging) + Tight TopTagger Cut (0.8)] & [1 jet passing ALL SR requirements (+btagging)]. The equation of the efficiency is the following:

$$\text{efficiency} = \frac{\text{Tight \& SR}}{\text{Tight \& Probe}} = \frac{(1 \text{ Jet pass Tight region} + 1 \text{ Jet pass SR})}{(1 \text{ Tight} + 1 \text{ Jet pass } SR_B)} \quad (\text{A.1})$$

In order to avoid any  $p_T$  bias, we implement jet randomization. Also except from a inclusive top-tagger efficiency Scale Factor, we calculate the efficiency of our top-tagger for each of the divided  $p_T$  regions to ensure that no  $p_T$  bias is imposed. Specifically:

- $p_T$ : [400, 600] GeV
- $p_T$ : [600, 800] GeV
- $p_T > 800$  GeV

Moreover, to be aligned with typical JMAR tagging evaluation methodologies, we have also divided the phase space into the  $p_T$  regions:

- $p_T$ : [400, 500] GeV
- $p_T$ : [500, 600] GeV
- $p_T > 600$  GeV

The Top Tagger Efficiency and SF results are provided in Figures [A.9](#), [A.11](#) and [A.10](#), [A.12](#) using JMAR proposed  $p_T$  regions respectively.

For each defined analysis region, it is crucial that we estimate the QCD contribution. We extract the QCD using a data-driven technique. For that reason, we perform a fit on the SoftDropMass in both regions defined in the efficiency equation (Tight & SR, Tight & Probe). The shape of the QCD distribution is estimated from the data distributions while inverting the b-tagging requirement as performed in the analysis section. The number of the QCD events are calculated in each region

from a fit performed on the SoftDropMass of the jet. To scale the  $t\bar{t}$  signal obtained from simulation, we calculate from the fit, the  $t\bar{t}$  signal strength ( $r_{t\bar{t}} = \frac{N_{observed}}{N_{expected}}$ ). Finally, for the calculation of the Signal distribution in each analysis region from data, we use the following equation:

$$\forall \text{regions} : S(x) = D(x) - N_{\text{QCD}} \times d_0(x) - B(x) \quad (\text{A.2})$$

where  $x = p_T^{t1,2}, |y^{t1,2}|, m^{t\bar{t}}, p_T^{t\bar{t}}, y^{t\bar{t}}$ ,  $S(x)$  is the signal,  $D(x)$  is the measured distribution in data in each region,  $d_0(x)$  is the QCD shape (taken from the data control region),  $B(x)$  is the subdominant backgrounds' contribution (both the shape and normalization are taken from simulation). The  $N_{\text{QCD}}$  is the fitted number of QCD events in each region.

We have measured the fraction of events used in the cross section measurement that are also used for the SF measurement to be of the order of 35% (due to phase-space differences and randomization in the SF measurement). We consider that this fraction is reasonably small.

Table A.1: Top Tagger efficiency Values for 2016 preVFP.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.757 \pm 0.058$	$0.791 \pm 0.009$	$0.791 \pm 0.01$
$p_T$ [400, 600]GeV	$0.742 \pm 0.067$	$0.793 \pm 0.011$	$0.793 \pm 0.021$
$p_T$ [600, 800]GeV	$0.774 \pm 0.134$	$0.79 \pm 0.016$	$0.79 \pm 0.025$
$p_T > 800$ GeV	$0.824 \pm 0.198$	$0.777 \pm 0.037$	$0.777 \pm 0.047$

 Table A.2: Top Tagger efficiency Values for 2016 preVFP using JMAR proposed  $p_T$  regions.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.757 \pm 0.058$	$0.791 \pm 0.009$	$0.791 \pm 0.01$
$p_T$ [400, 500]GeV	$0.806 \pm 0.136$	$0.792 \pm 0.021$	$0.792 \pm 0.031$
$p_T$ [500, 600]GeV	$0.721 \pm 0.076$	$0.793 \pm 0.013$	$0.793 \pm 0.022$
$p_T > 600$ GeV	$0.785 \pm 0.114$	$0.787 \pm 0.014$	$0.787 \pm 0.024$

Table A.3: Top Tagger efficiency Values for 2016 postVFP.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.79 \pm 0.052$	$0.786 \pm 0.008$	$0.786 \pm 0.011$
$p_T$ [400, 600]GeV	$0.776 \pm 0.061$	$0.79 \pm 0.01$	$0.79 \pm 0.021$
$p_T$ [600, 800]GeV	$0.81 \pm 0.104$	$0.781 \pm 0.015$	$0.781 \pm 0.024$
$p_T > 800$ GeV	$0.861 \pm 0.259$	$0.77 \pm 0.035$	$0.77 \pm 0.046$

In order to be sure that all statistical limitations are included in the  $p_T$  dependent Top Tagger SF values, we are using the  $p_T$  dependent error found in the Top Tagger SF values per  $p_T$  region. This is included only in the Data vs MC figures (A.1 - A.8), while we also provide the Top Tagger  $p_T$  inclusive values are provided in the following table.

APPENDIX A. TOP TAGGER EFFICIENCY

Table A.4: Top Tagger efficiency Values for 2016 postVFP using JMAR proposed  $p_T$  regions.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.79 \pm 0.052$	$0.786 \pm 0.008$	$0.786 \pm 0.011$
$p_T$ [400, 500]GeV	$0.782 \pm 0.1$	$0.773 \pm 0.018$	$0.773 \pm 0.029$
$p_T$ [500, 600]GeV	$0.774 \pm 0.076$	$0.8 \pm 0.012$	$0.8 \pm 0.02$
$p_T > 600$ GeV	$0.817 \pm 0.097$	$0.779 \pm 0.013$	$0.779 \pm 0.025$

Table A.5: Top Tagger efficiency Values for 2017.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.814 \pm 0.032$	$0.868 \pm 0.006$	$0.868 \pm 0.009$
$p_T$ [400, 600]GeV	$0.81 \pm 0.04$	$0.867 \pm 0.008$	$0.867 \pm 0.017$
$p_T$ [600, 800]GeV	$0.827 \pm 0.063$	$0.871 \pm 0.012$	$0.871 \pm 0.021$
$p_T > 800$ GeV	$0.793 \pm 0.132$	$0.869 \pm 0.029$	$0.869 \pm 0.037$

Table A.6: Top Tagger efficiency Values for 2017 using JMAR proposed  $p_T$  regions.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.814 \pm 0.032$	$0.868 \pm 0.006$	$0.868 \pm 0.009$
$p_T$ [400, 500]GeV	$0.808 \pm 0.069$	$0.854 \pm 0.014$	$0.854 \pm 0.023$
$p_T$ [500, 600]GeV	$0.812 \pm 0.047$	$0.872 \pm 0.009$	$0.872 \pm 0.018$
$p_T > 600$ GeV	$0.822 \pm 0.058$	$0.870 \pm 0.011$	$0.870 \pm 0.019$

Table A.7: Top Tagger efficiency Values for 2018.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.792 \pm 0.03$	$0.827 \pm 0.005$	$0.827 \pm 0.008$
$p_T$ [400, 600]GeV	$0.789 \pm 0.039$	$0.825 \pm 0.006$	$0.825 \pm 0.014$
$p_T$ [600, 800]GeV	$0.805 \pm 0.051$	$0.833 \pm 0.01$	$0.833 \pm 0.02$
$p_T > 800$ GeV	$0.752 \pm 0.104$	$0.822 \pm 0.024$	$0.822 \pm 0.037$

Table A.8: Top Tagger efficiency Values for 2018 using JMAR proposed  $p_T$  regions.

Eff. Type	Eff. Data (stat)	Eff. $t\bar{t}$ (stat)	Eff. $t\bar{t}$ (stat + systematic)
Inclusive	$0.792 \pm 0.03$	$0.827 \pm 0.005$	$0.827 \pm 0.008$
$p_T$ [400, 500]GeV	$0.739 \pm 0.074$	$0.811 \pm 0.011$	$0.811 \pm 0.019$
$p_T$ [500, 600]GeV	$0.807 \pm 0.045$	$0.832 \pm 0.007$	$0.832 \pm 0.018$
$p_T > 600$ GeV	$0.797 \pm 0.046$	$0.832 \pm 0.009$	$0.832 \pm 0.021$

Table A.9: Top Tagger SF Values for 2016 preVFP.

SF Type	Value $\pm$ error
Inclusive	$0.957 \pm 0.074$
$p_T$ [400, 600]GeV	$0.937 \pm 0.085$
$p_T$ [600, 800]GeV	$0.981 \pm 0.17$
$p_T > 800$ GeV	$1.06 \pm 0.26$

Table A.10: Top Tagger SF Values for 2016 preVFP using JMAR proposed  $p_T$  regions.

SF Type	Value $\pm$ error
Inclusive	$0.957 \pm 0.074$
$p_T$ [400, 500]GeV	$1.02 \pm 0.173$
$p_T$ [500, 600]GeV	$0.91 \pm 0.097$
$p_T > 600$ GeV	$0.997 \pm 0.15$

APPENDIX A. TOP TAGGER EFFICIENCY

Table A.11: Top Tagger SF Values for 2016 postVFP.

SF Type	Value $\pm$ error
Inclusive	1.01 $\pm$ 0.067
$p_T$ [400, 600]GeV	0.983 $\pm$ 0.078
$p_T$ [600, 800]GeV	1.04 $\pm$ 0.135
$p_T > 800$ GeV	1.12 $\pm$ 0.34

Table A.12: Top Tagger SF Values for 2016 postVFP using JMAR proposed  $p_T$  regions.

SF Type	Value $\pm$ error
Inclusive	1.01 $\pm$ 0.067
$p_T$ [400, 500]GeV	1.01 $\pm$ 0.132
$p_T$ [500, 600]GeV	0.971 $\pm$ 0.097
$p_T > 600$ GeV	1.05 $\pm$ 0.13

Table A.13: Top Tagger SF Values for 2017.

SF Type	Value $\pm$ error
Inclusive	0.938 $\pm$ 0.038
$p_T$ [400, 600]GeV	0.935 $\pm$ 0.046
$p_T$ [600, 800]GeV	0.95 $\pm$ 0.059
$p_T > 800$ GeV	0.912 $\pm$ 0.155

Table A.14: Top Tagger SF Values for 2017 using JMAR proposed  $p_T$  regions.

SF Type	Value $\pm$ error
Inclusive	0.938 $\pm$ 0.038
$p_T$ [400, 500]GeV	0.946 $\pm$ 0.082
$p_T$ [500, 600]GeV	0.931 $\pm$ 0.055
$p_T > 600$ GeV	0.945 $\pm$ 0.068

Table A.15: Top Tagger SF Values for 2018.

SF Type	Value $\pm$ error
Inclusive	0.958 $\pm$ 0.037
$p_T$ [400, 600]GeV	0.956 $\pm$ 0.048
$p_T$ [600, 800]GeV	0.967 $\pm$ 0.062
$p_T > 800$ GeV	0.914 $\pm$ 0.13

Table A.16: Top Tagger SF Values for 2018 using JMAR proposed  $p_T$  regions.

SF Type	Value $\pm$ error
Inclusive	0.958 $\pm$ 0.037
$p_T$ [400, 500]GeV	0.912 $\pm$ 0.093
$p_T$ [500, 600]GeV	0.971 $\pm$ 0.055
$p_T > 600$ GeV	0.959 $\pm$ 0.056

Table A.17: Top Tagger SF Values for all years.

Inclusive SF	Value $\pm$ error
2016 preVFP	0.957 $\pm$ 0.26
2016 postVFP	1.01 $\pm$ 0.3
2017	0.938 $\pm$ 0.16
2018	0.958 $\pm$ 0.13

## APPENDIX A. TOP TAGGER EFFICIENCY

The following pictures (Fig: A.1 - A.8) show the distributions of the probe jet and the jets used for the measurement in the related regions. The plots are normalized to the same area in order to compare the distributions, while from the distributions it can be seen that the probed phase space is very similar.

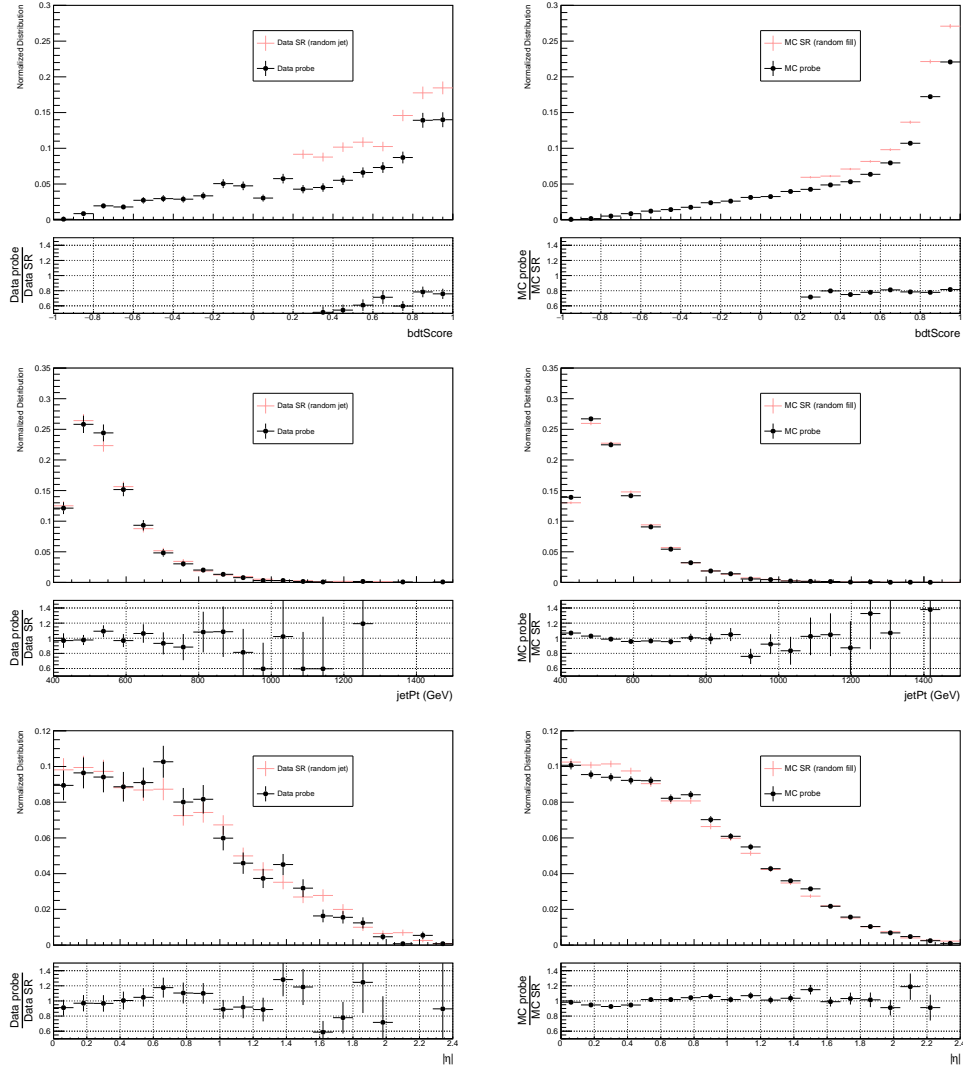


Figure A.1: Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 preVFP using a random jet.

In the following figures, we present the Data vs MC plots with and without the top tagger SF applied (Fig: A.13- A.21).

APPENDIX A. TOP TAGGER EFFICIENCY

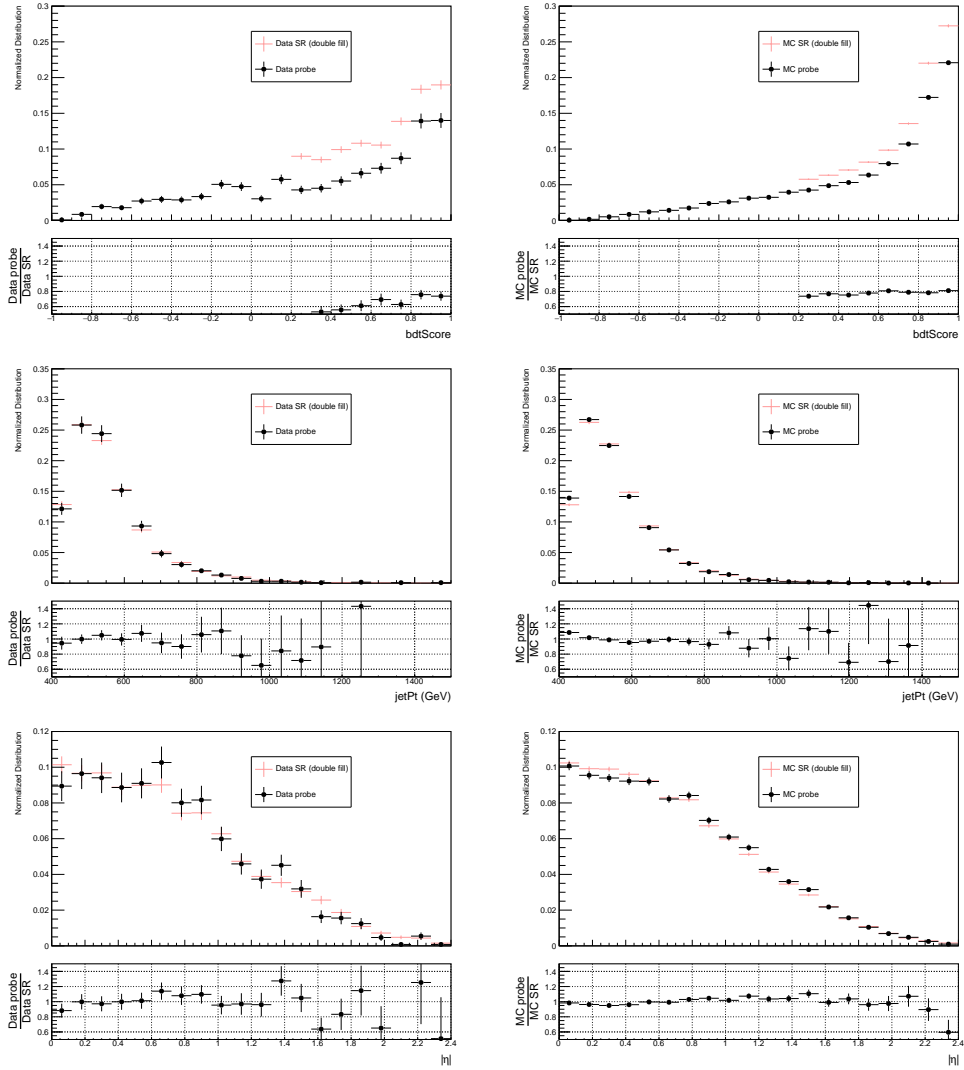


Figure A.2: Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 preVFP using both jets.



APPENDIX A. TOP Tagger Efficiency

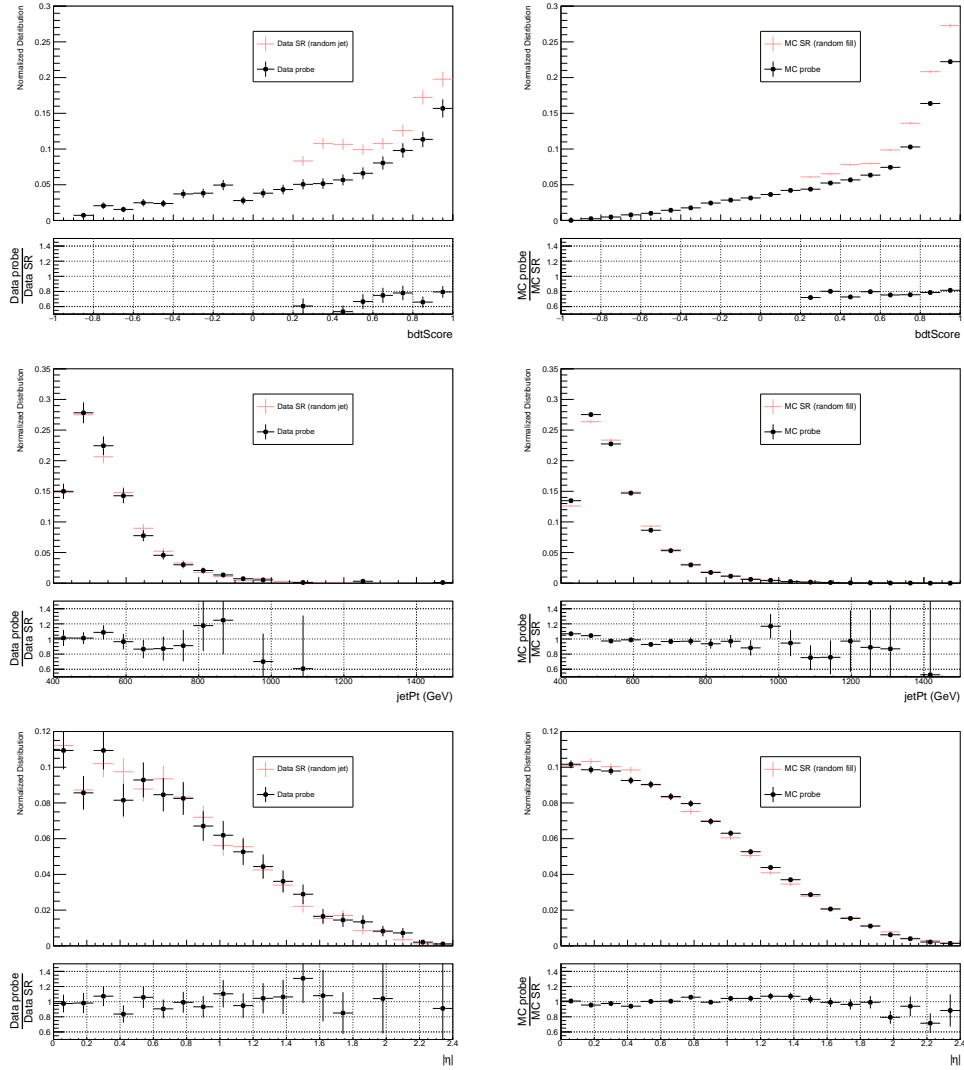


Figure A.3: Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 postVFP using a random jet.

APPENDIX A. TOP TAGGER EFFICIENCY

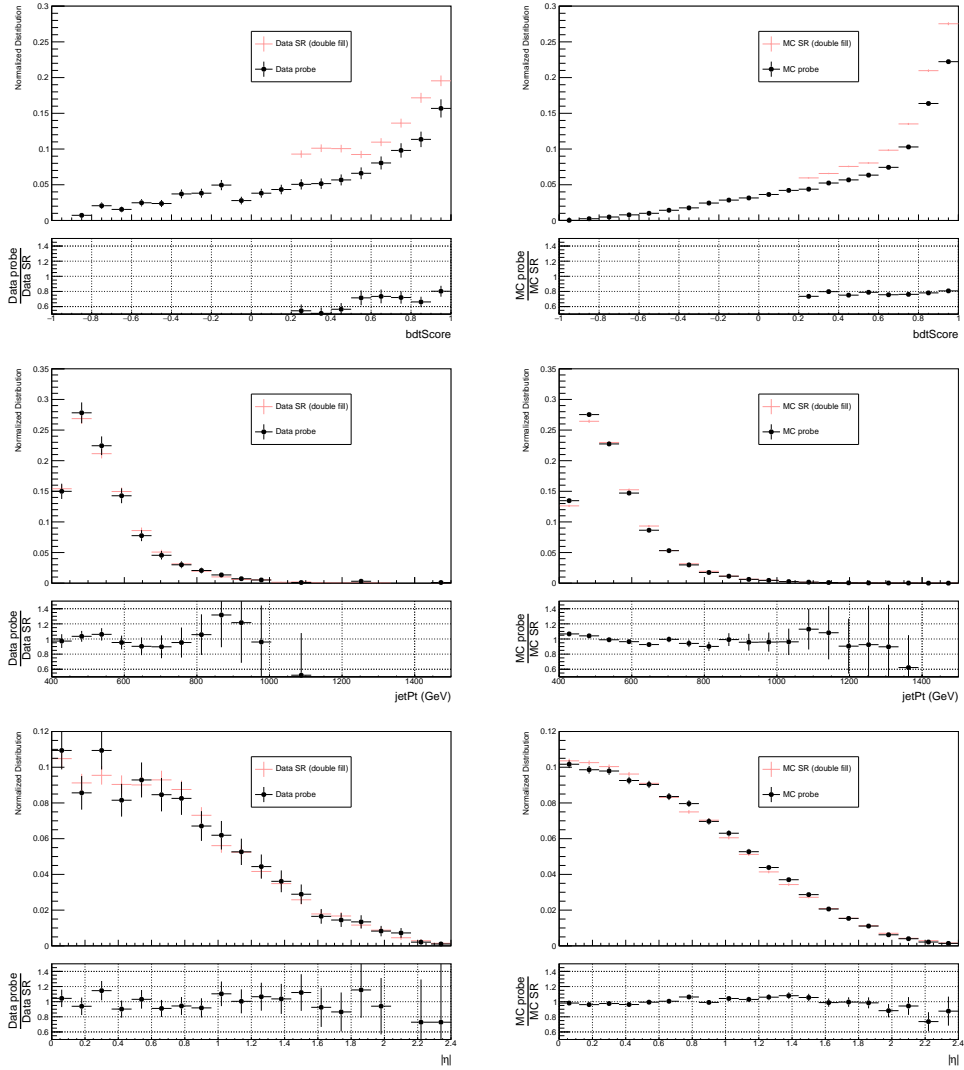


Figure A.4: Distributions of the probe jet and the jets used for the measurement in the related regions for 2016 postVFP using both jets.

APPENDIX A. TOP Tagger Efficiency

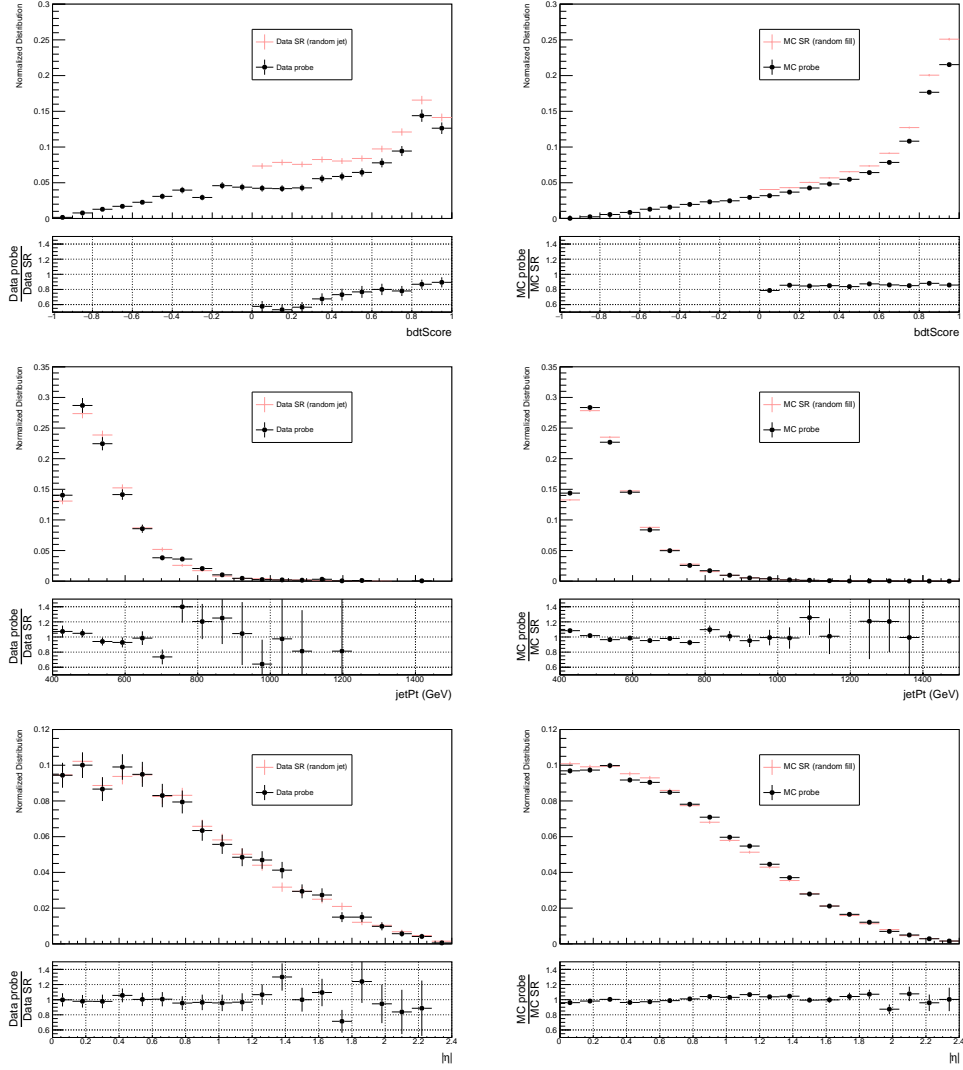


Figure A.5: Distributions of the probe jet and the jets used for the measurement in the related regions for 2017 using a random jet.

APPENDIX A. TOP TAGGER EFFICIENCY

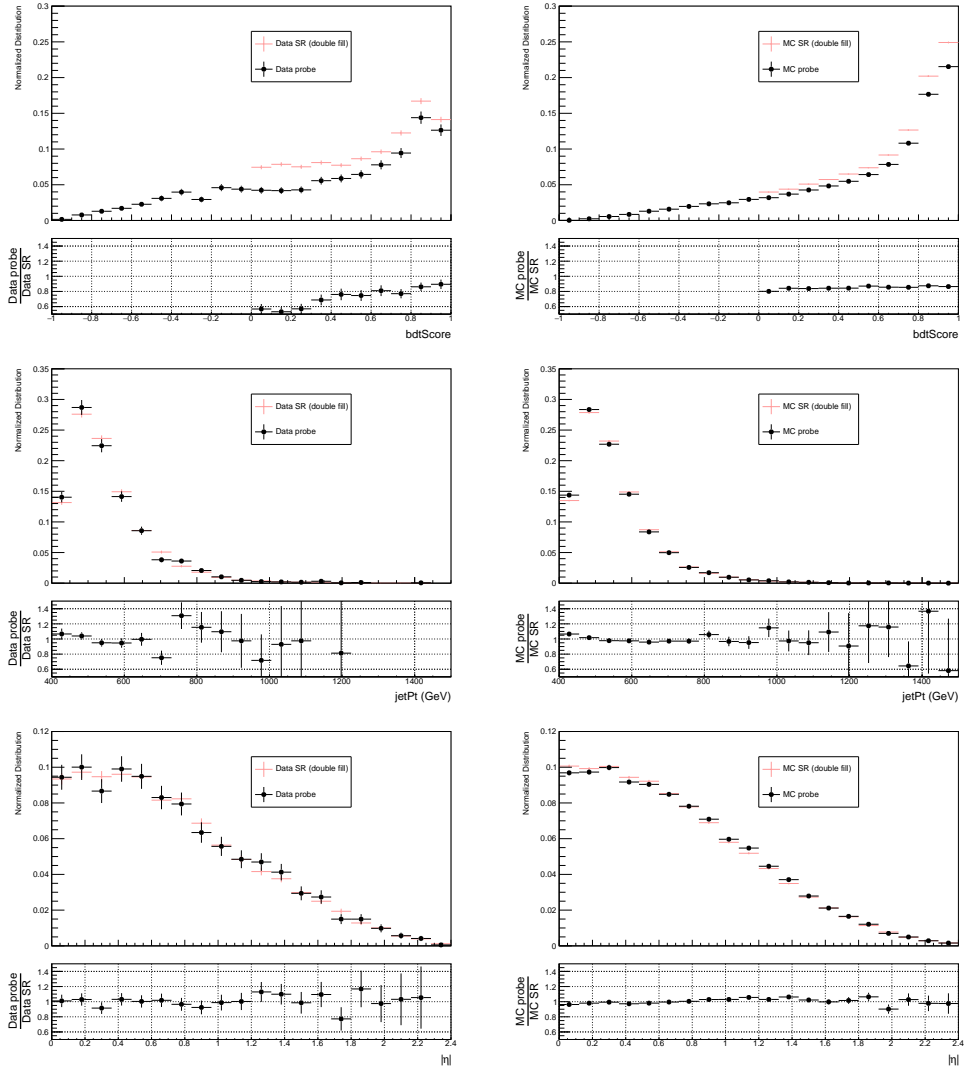


Figure A.6: Distributions of the probe jet and the jets used for the measurement in the related regions for 2017 using both jets.

APPENDIX A. TOP Tagger EFFICIENCY

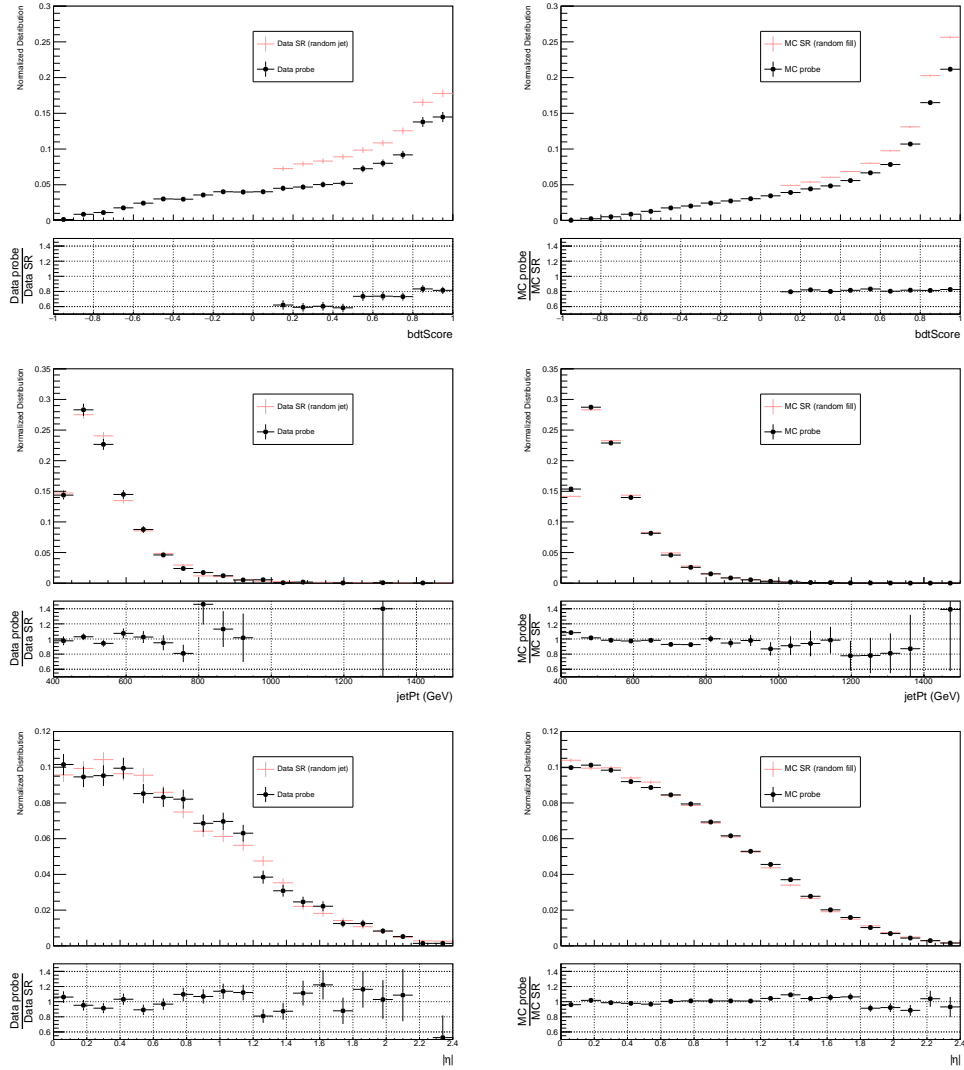


Figure A.7: Distributions of the probe jet and the jets used for the measurement in the related regions for 2018 using a random jet.

APPENDIX A. TOP TAGGER EFFICIENCY

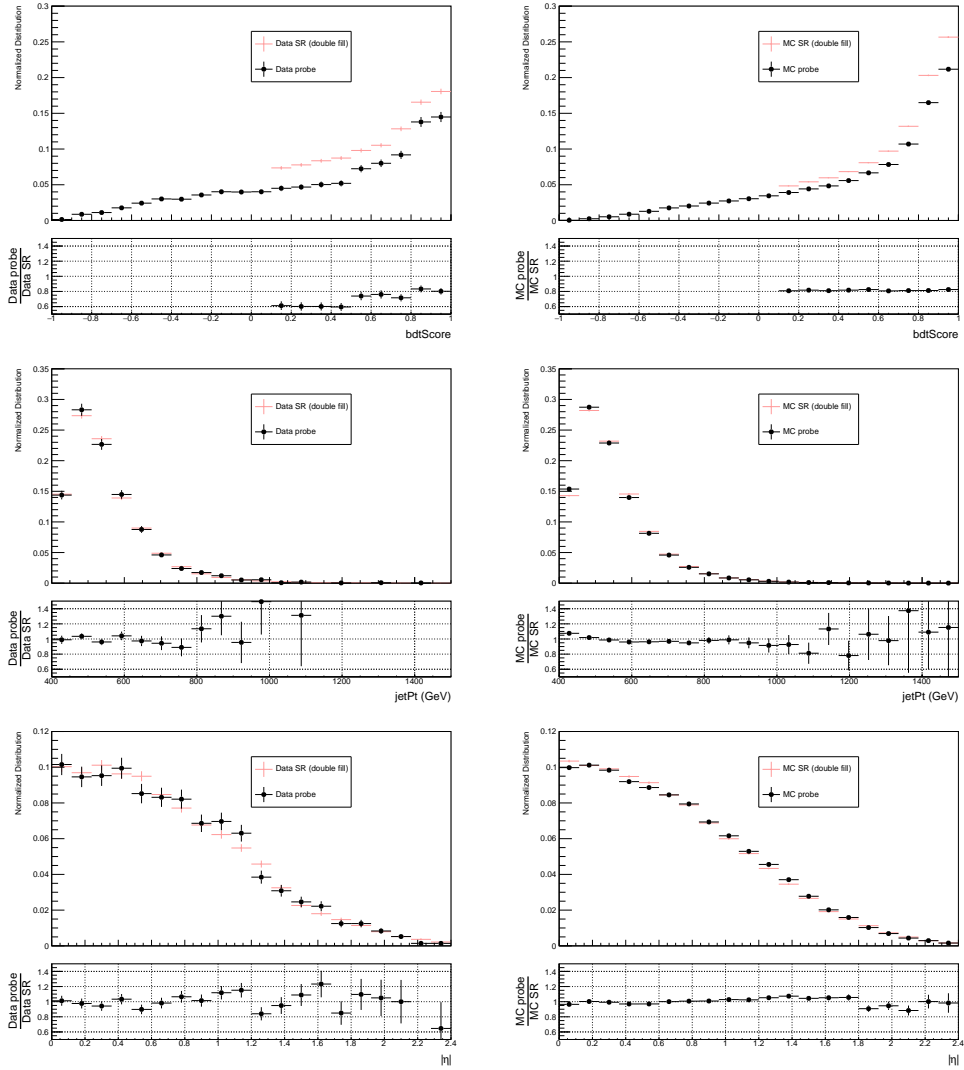


Figure A.8: Distributions of the probe jet and the jets used for the measurement in the related regions for 2018 using both jets.

APPENDIX A. TOP TAGGER EFFICIENCY

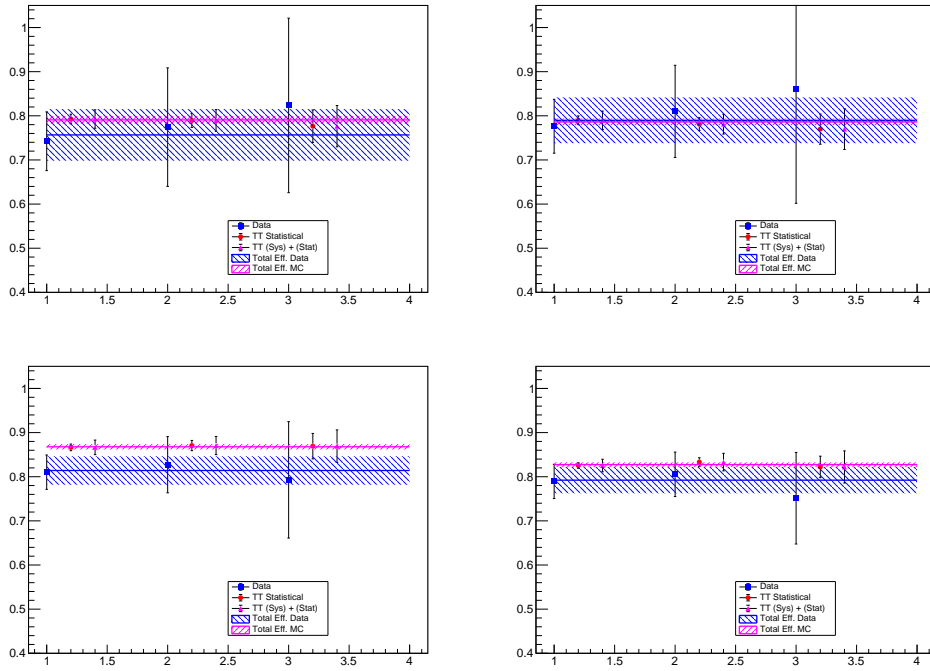


Figure A.9: Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per  $p_T$  region and inclusive with statistical and systematic uncertainties

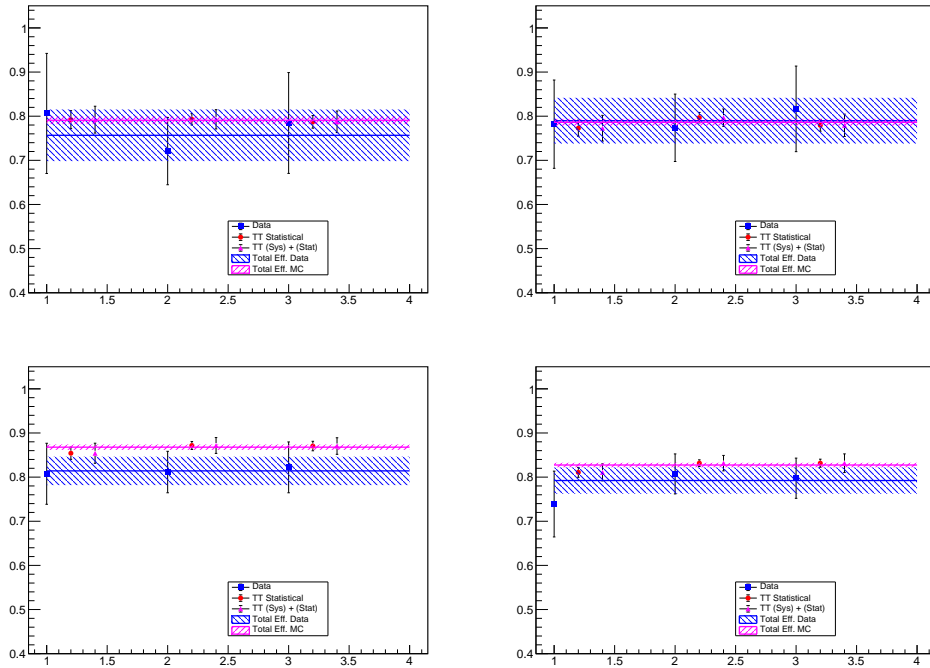


Figure A.10: Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per JMAR proposed  $p_T$  region and inclusive with statistical and systematic uncertainties

## APPENDIX A. TOP TAGGER EFFICIENCY

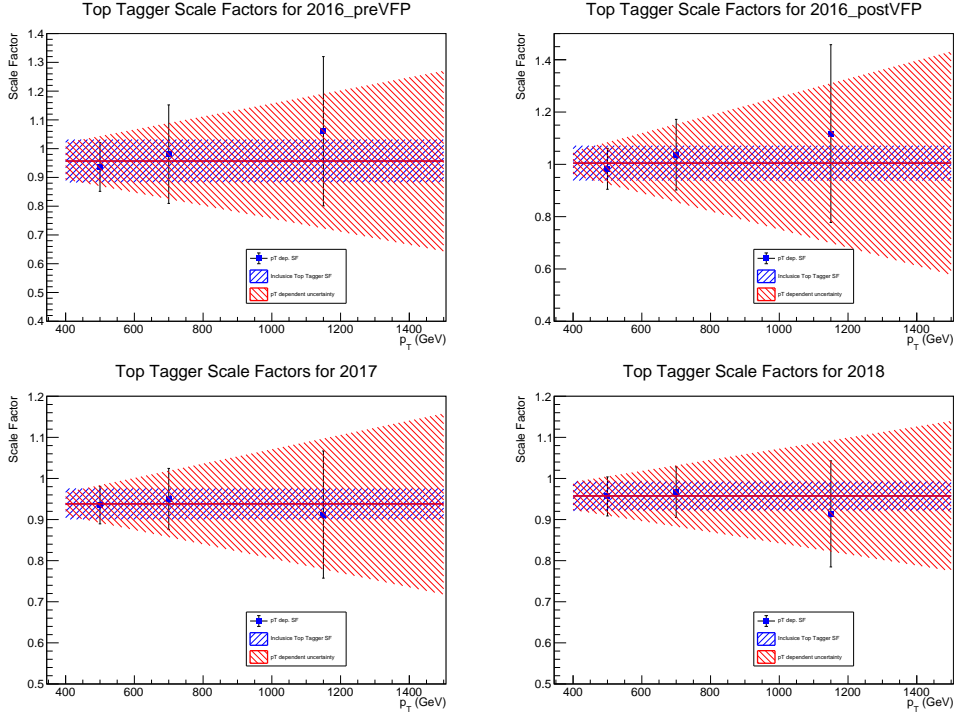


Figure A.11: Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per  $p_T$  region and inclusive with statistical and systematic uncertainties

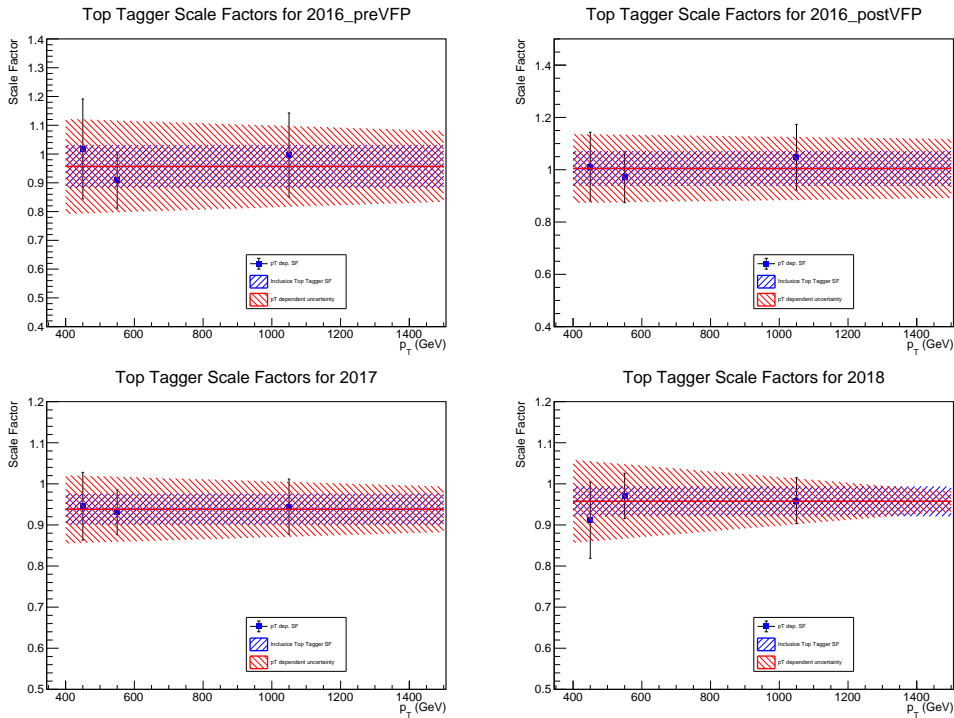


Figure A.12: Extracted top tagger efficiency per year (2016 preVFP top left, 2016 postVFP top right, 2017 bottom left, 2018 bottom right) per JMAR proposed  $p_T$  region and inclusive with statistical and systematic uncertainties



APPENDIX A. TOP TAGGER EFFICIENCY

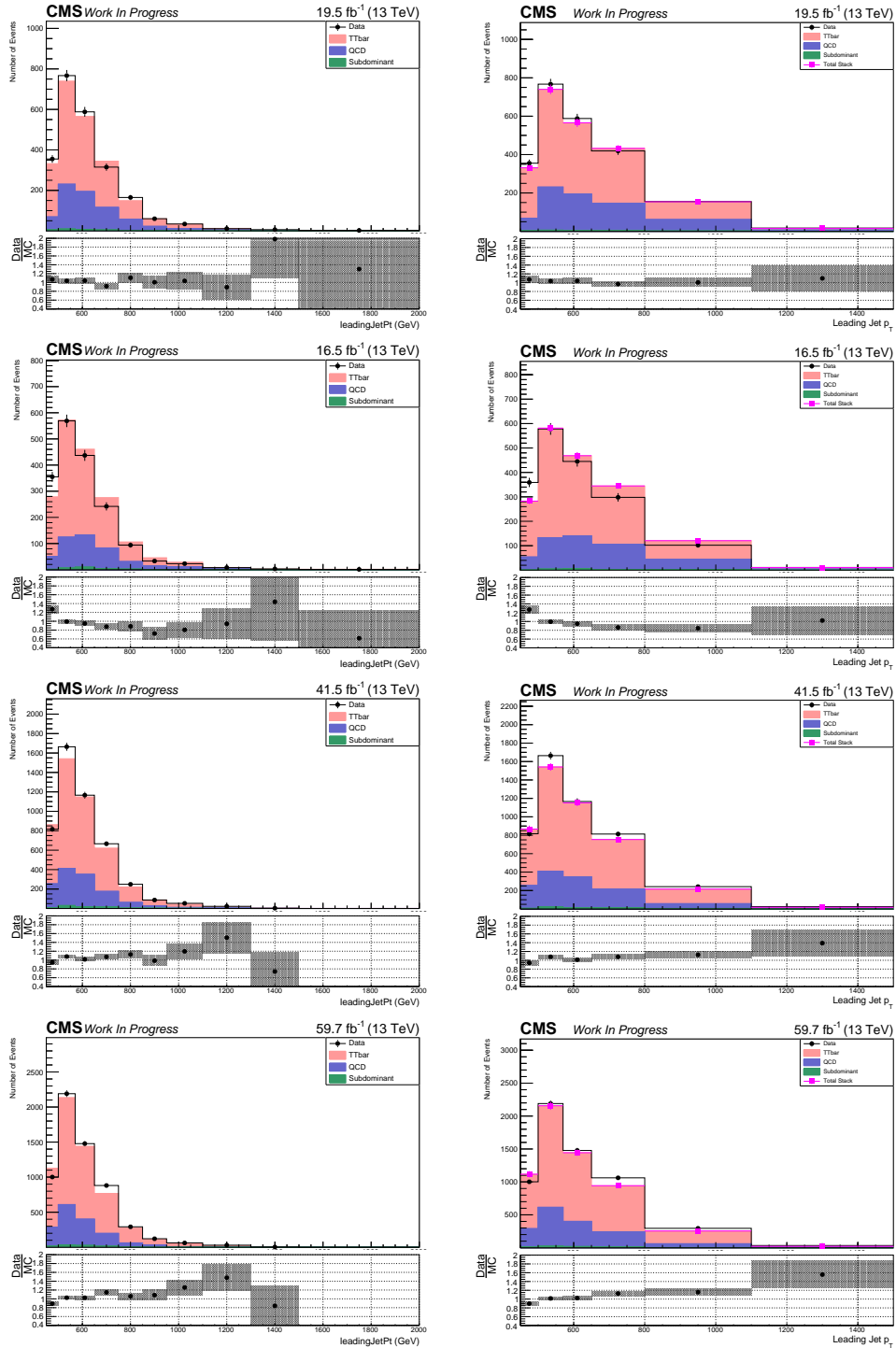


Figure A.13: Leading jet  $p_T$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

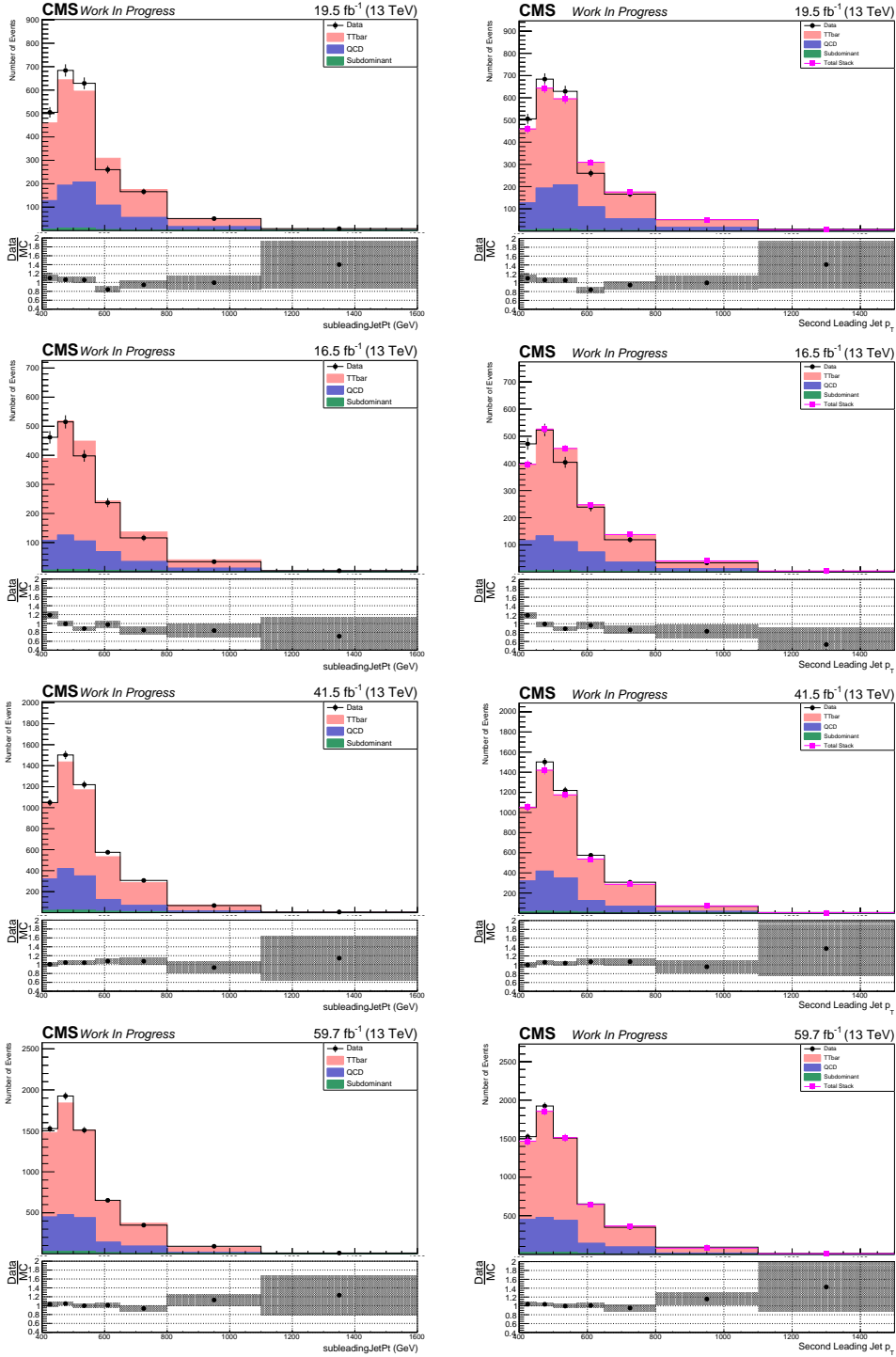


Figure A.14: Second leading jet  $p_T$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied). 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

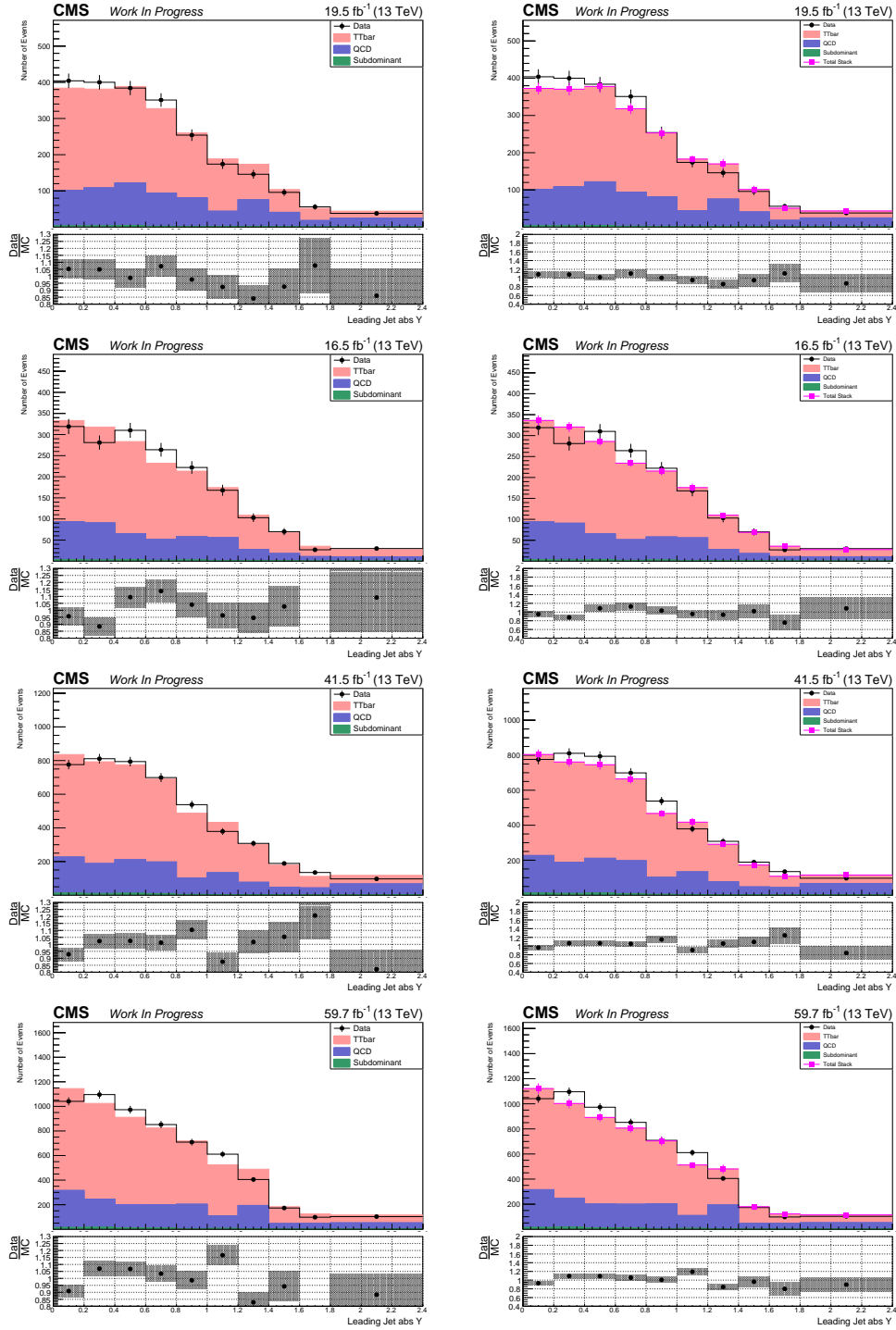


Figure A.15: Leading jet rapidity distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

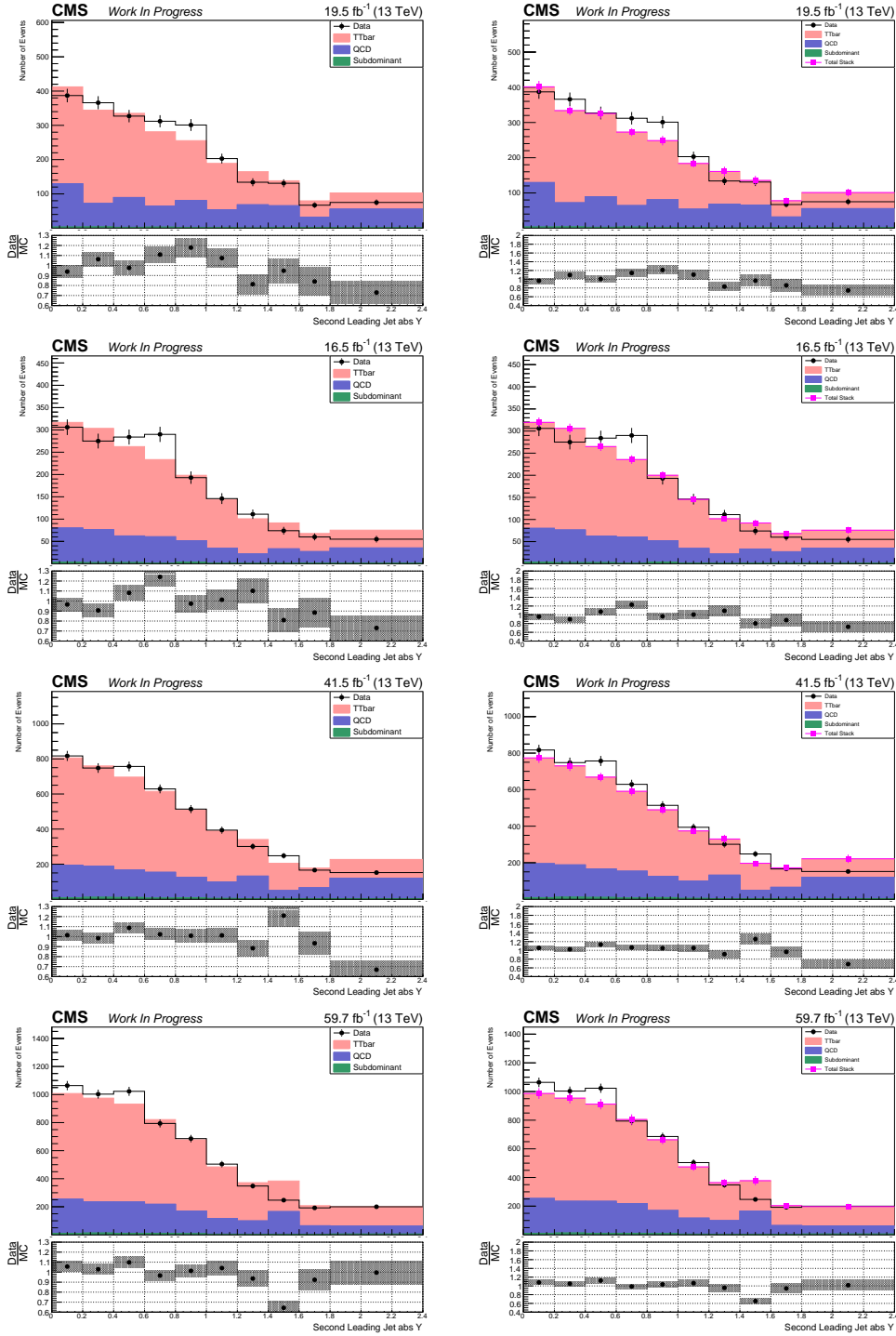


Figure A.16: Second leading jet rapidity distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied). 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

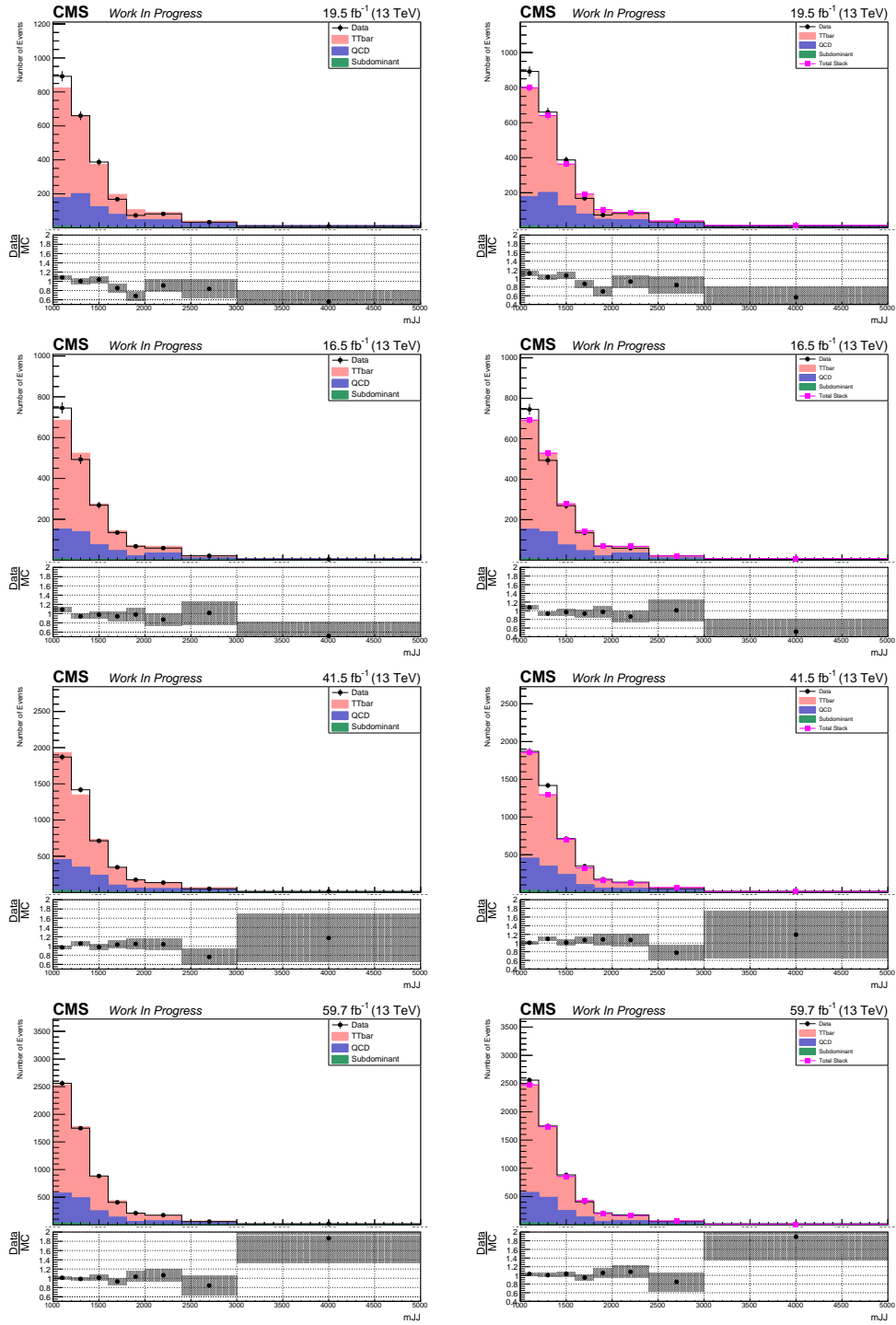


Figure A.17: Dijet mass (m<sub>JJ</sub>) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

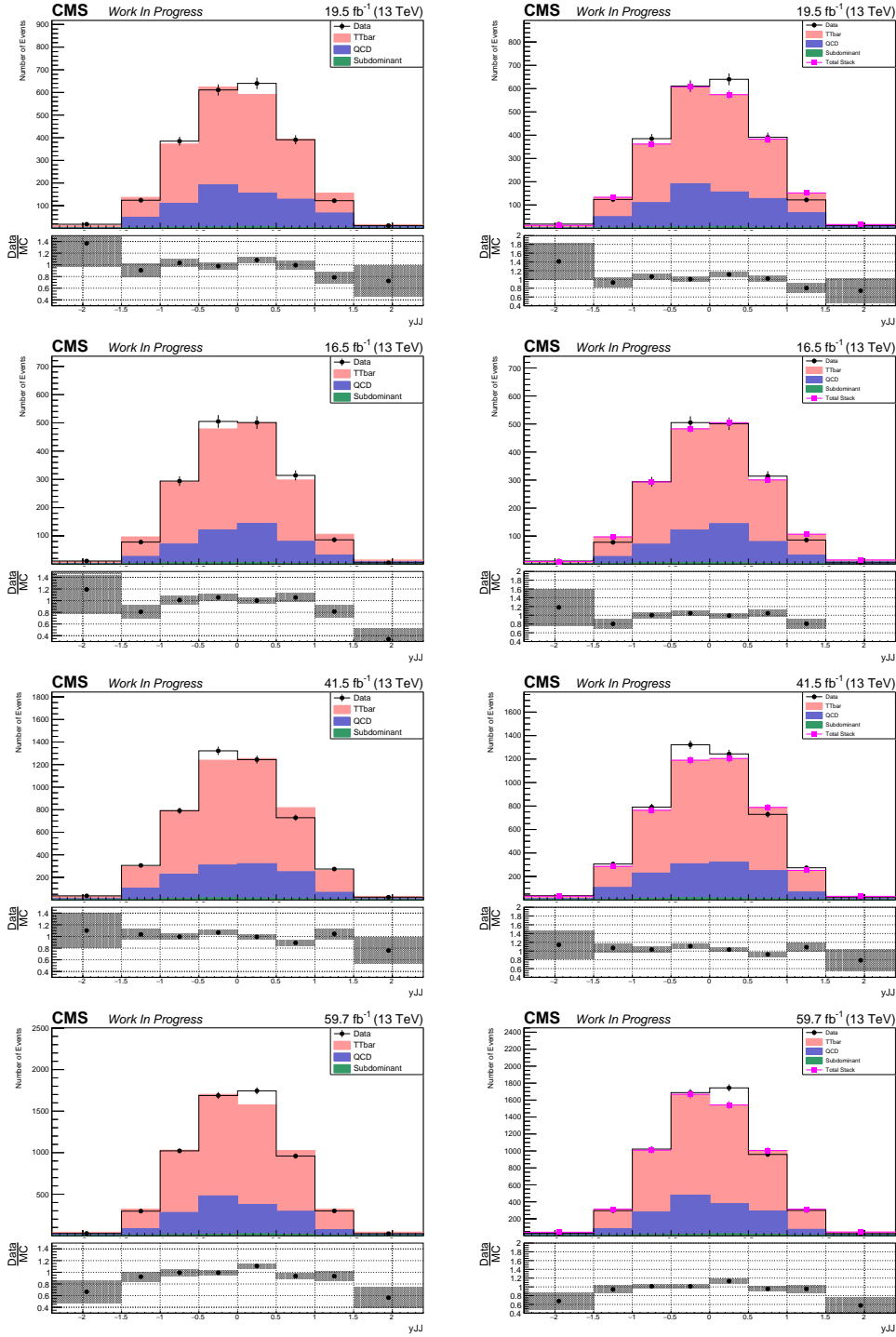


Figure A.18: Dijet rapidity ( $y_{JJ}$ ) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied). 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

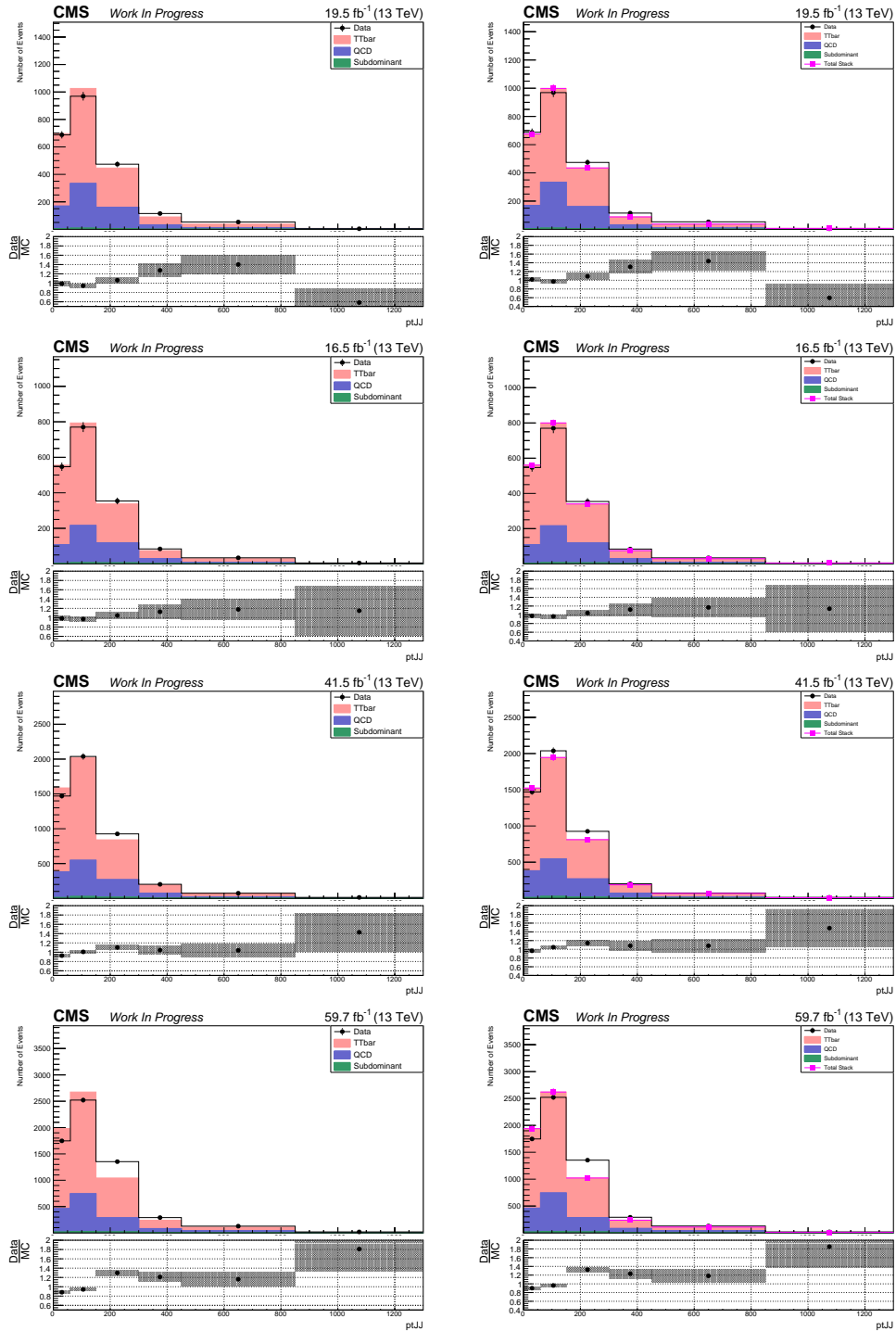


Figure A.19: Dijet  $p_T$  ( $pt_{JJ}$ ) distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied). 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).

APPENDIX A. TOP TAGGER EFFICIENCY

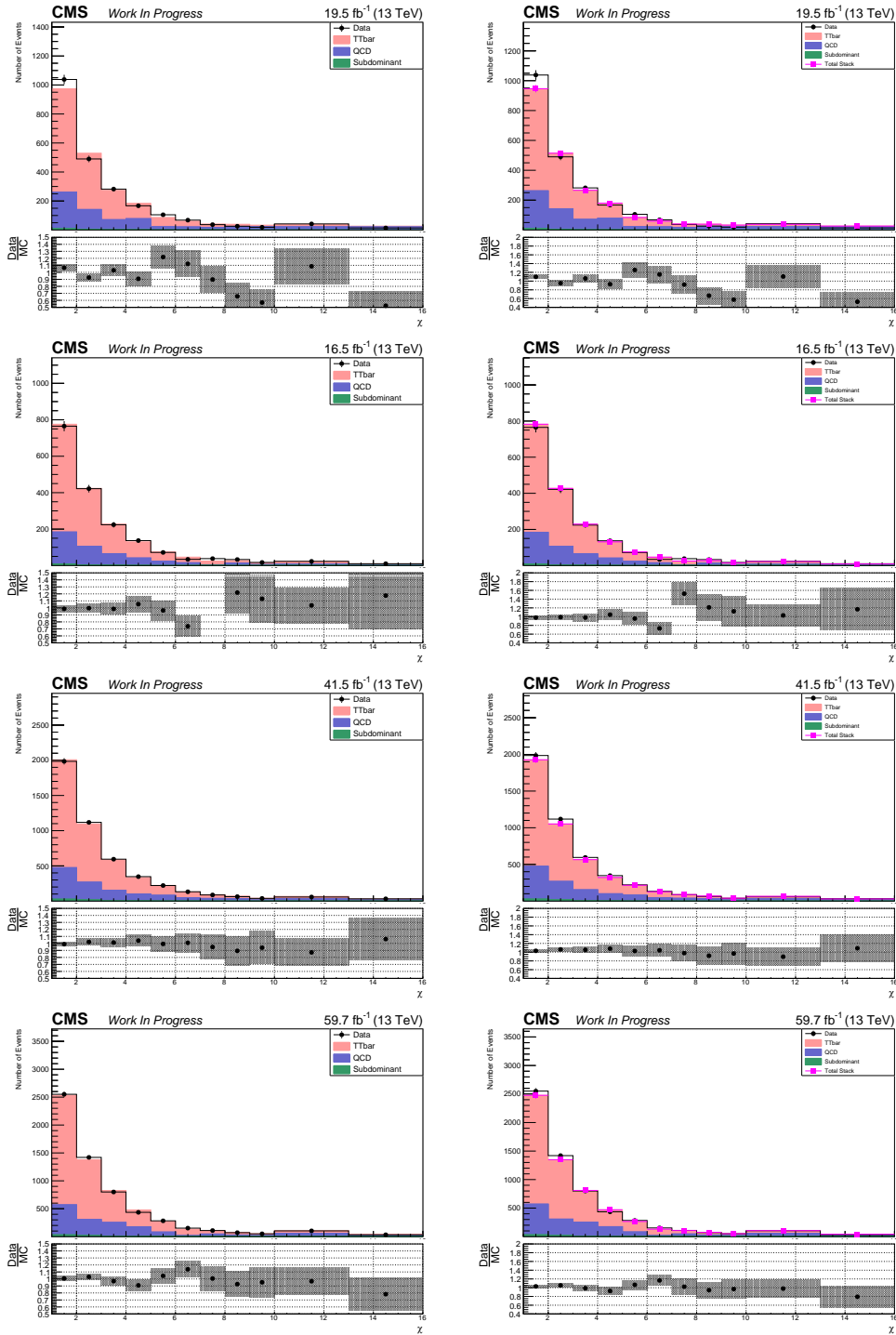


Figure A.20: Dijet  $\chi$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied. 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).



APPENDIX A. TOP TAGGER EFFICIENCY

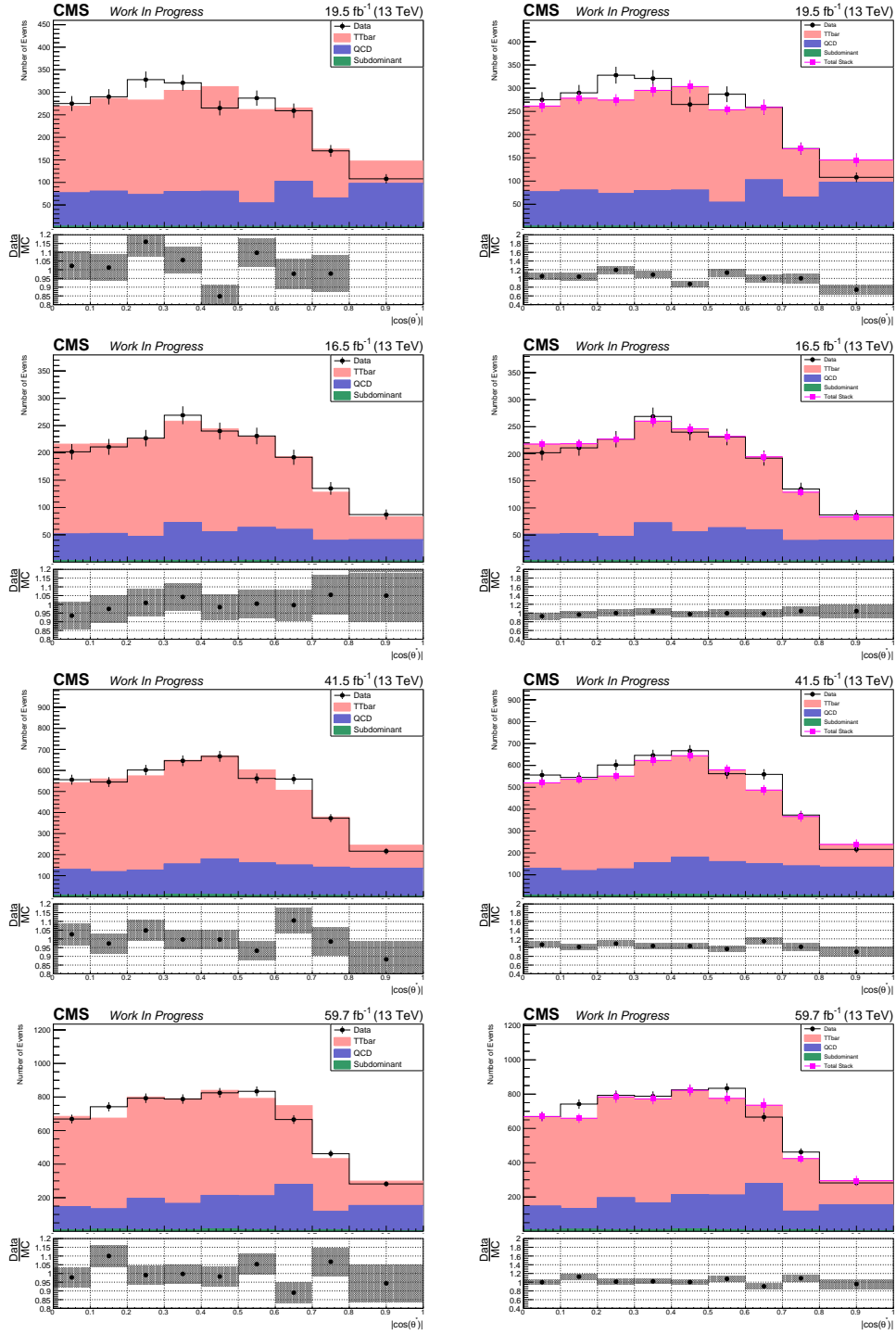


Figure A.21:  $\cos(\theta^*)$  distributions in the signal region for all years (left without Top Tagger SF and right with Top Tagger SF applied). 2016 preVFP on 1st row, 2016 postVFP on 2nd row, 2017 on 3rd row and 2018 on bottom row).