



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ  
ΠΟΛΥΤΕΧΝΕΙΟ**  
**ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ  
ΜΑΘΗΜΑΤΙΚΩΝ  
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**  
**ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ  
ΜΗΧΑΝΙΚΩΝ**

**ΕΚΕΦΕ «ΔΗΜΟΚΡΙΤΟΣ»**  
**ΙΝΣΤΙΤΟΥΤΟ ΝΑΝΟΕΠΙΣΤΗΜΗΣ  
ΚΑΙ ΝΑΝΟΤΕΧΝΟΛΟΓΙΑΣ**  
**ΙΝΣΤΙΤΟΥΤΟ ΠΥΡΗΝΙΚΗΣ ΚΑΙ  
ΣΩΜΑΤΙΔΙΑΚΗΣ ΦΥΣΙΚΗΣ**



**Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών**  
**«Φυσική και Τεχνολογικές Εφαρμογές»**

**Μέτρηση της ενεργού διατομής παραγωγής top  
quark σε συγκρούσεις pp σε  $\sqrt{s} = 13\text{TeV}$  στο  
πείραμα CMS**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**  
**του Θεόδωρου Χατζησταύρου**

**Επιβλέπων: Κωνσταντίνος Κουσουρής**

**Επίκουρος Καθηγητής ΣΕΜΦΕ**

**Αθήνα, Ιούνιος, 2020**

# Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Κουσουρή, για τον χρόνο που αφιέρωσε και τις γνώσεις που μου μετέδωσε κατά την διεκπόνηση της διπλωματικής. Επιπλέον ευχαριστώ τον κ. Τσιπολίτη για την συνεπίβλεψη και τις συμβουλές του, και όλη την ομάδα του ΉΜΣ για την καλή μας συνεργασία.



# Περίληψη

Αντικείμενο της διπλωματικής εργασίας είναι η μέτρηση της ενεργού διατομής παραγωγής τοπ κουάρκ μέσω της ηλεκτρασθενούς αλληλεπίδρασης, και συγκεκριμένα στο κανάλι  $t$  –  $channel$ . Το κανάλι αυτό έχει την μεγαλύτερη ενεργό διατομή σε σχέση με τα υπόλοιπα, και επιπλέον έχει χαρακτηριστικά που το βοηθούν να ξεχωρίσει από άλλες διαδικασίες.

Ο στόχος ουσιαστικά είναι η εξαγωγή ενός παράγοντα κλίμακας της θεωρητικής ενεργού διατομής, μέσα από προσαρμογή με την μέθοδο μέγιστης πιθανοφάνειας σε δεδομένα. Τα δεδομένα που χρησιμοποιήθηκαν προέκυψαν από συγκρούσεις πρωτονίων στον LHC κατά το έτος 2016 με συνολική φωτεινότητα  $L = 35.9fb^{-1}$ .

Για την προσαρμογή στα δεδομένα έγινε μελέτη διαφόρων πιθανών μεταβλητών που προσφέρουν διακριτική ικανότητα του σήματος από τις διαδικασίες υποβάθρου. Η διαχωριστική ικανότητα αυτών εκτιμήθηκε ποσοτικά. Η τελική επιλογή της μεταβλητής που θα γίνει η προσαρμογή έγινε μέσα από μελέτη Monte Carlo toys, όπου μπορούν να μετρηθούν τα χαρακτηριστικά της προσαρμογής ως προς την αντικειμενικότητα των εκτιμώμενων τιμών και το σφάλμα τους, και των συσχετίσεων που εμφανίζονται μεταξύ τους.

Μετά από την μελέτη αυτή έγιναν προσαρμογές σε περιοχές ελέγχου του φασικού χώρου που είναι πλούσιες σε μερικές από τις διαδικασίες υποβάθρου, καθώς και μία ταυτόχρονη προσαρμογή, που χρησιμοποιεί τις εκτιμήσεις αυτές σαν περιορισμούς, στην κύρια περιοχή του σήματος και στην περιοχή ελέγχου του κύριου υποβάθρου. Με τον τρόπο αυτό υπήρχε καλύτερος περιορισμός των διαδικασιών υποβάθρου που βελτιώνει την εκτίμηση του σήματος. Το αποτέλεσμα της ενεργού διατομής που μετρήθηκε είναι συμβατό με την θεωρητικά προβλεπόμενη τιμή στο Καθιερωμένο Πρότυπο και προηγούμενες μετρήσεις.



# Abstract

The object of this thesis is the measurement of single top quark events production cross section, and specifically the t-channel process. This channel has the larger single top cross section and has characteristics that aid in discriminating such events from background processes.

The goal is essentially to extract a scale factor for the theoretical cross section, with a maximum likelihood fit on data. The data used in this thesis come from  $pp$  collisions at LHC collected during the year 2016 with an integrated luminosity of  $L = 35.9 fb^{-1}$ .

For the fit on data a study took place concerning the different possible variables which are discriminating the signal from the background. The separability power of these variables was estimated quantitatively. The final choice of the variable for the fit was done after a Toy Monte Carlo study, where they can be studied the characteristics of the fit regarding the bias of estimations of parameters and their errors, as well as the correlatings among them.

After this study they were performed fits on control regions of phase space rich in some of background processes, and a simultaneous fit, which uses the previous as constraints, in the signal region and the dominant background control region. In this way the background processes were constrained more efficiently and therefore there was better estimation for the signal estimation. The result of the measured cross section was in agreement with the theoretical Standard Model prediction and previous measurements.



# Περιεχόμενα

Περίληψη	1
Abstract	3
Περιεχόμενα	6
<b>1 Θεωρητικό Υπόβαθρο</b>	<b>7</b>
1.1 Εισαγωγή	7
1.2 Το Καθιερωμένο Πρότυπο	8
1.2.1 Η Κβαντική Χρωμοδυναμική	9
1.2.2 Οι ηλεκτρασθενείς αλληλεπιδράσεις	12
1.3 Φυσική Συγκρούσεων Πρωτονίων	17
1.3.1 Pile-up	19
1.3.2 Η δομή του πρωτονίου	19
1.3.3 Σκληρή σκέδαση παρτονίων	21
1.3.4 Underlying Event	22
1.3.5 Κινηματικές Μεταβλητές	22
1.4 Το Top Quark	25
1.4.1 Διάσπαση top quark	25
1.4.2 Παραγωγή top quark	26
<b>2 Ο ανιχνευτής Compact Muon Solenoid (CMS)</b>	<b>31</b>
2.1 Εισαγωγή	31
2.2 Ανιχνευτής τροχιών (Tracker)	32
2.3 Ηλεκτρομαγνητικό Θερμιδόμετρο (ECAL)	33
2.4 Αδρονικό Θερμιδόμετρο (ECAL)	34
2.5 Σύστημα Μιονίων (Muon System)	35
<b>3 Analysis</b>	<b>39</b>
3.1 Introduction	39
3.2 Physical Processes	39
3.3 Monte Carlo Samples and Data Set	43



3.4	Event Selection . . . . .	46
3.5	Data Driven QCD Multijet Background . . . . .	47
3.6	Top Quark Reconstruction . . . . .	53
3.7	Fit Variable Study . . . . .	57
3.7.1	Discriminating Variables . . . . .	58
3.7.2	Toy MC fit validation . . . . .	62
3.8	Measurement of cross section . . . . .	69
3.8.1	Leptons efficiencies Scale Factors . . . . .	69
3.8.2	Constraining $t\bar{t}$ and $DY$ backgrounds . . . . .	72
3.8.3	Fit with electrons . . . . .	74
3.8.4	Fit with muons . . . . .	78
3.9	Results . . . . .	80
3.10	Conclusions . . . . .	86
<b>A</b>	<b>Maximum Likelihood fit</b>	<b>87</b>
A.1	Definition of method . . . . .	87
A.2	Extended Likelihood . . . . .	88
A.3	Binned Maximum Likelihood Fit . . . . .	89
A.4	Simultaneous fitting . . . . .	90
<b>B</b>	<b>Multivariate Analysis</b>	<b>91</b>
B.1	Deifinitions . . . . .	91
B.2	Boosted Decision Trees (BDT) . . . . .	92

# Κεφάλαιο 1

## Θεωρητικό Υπόβαθρο

### 1.1 Εισαγωγή

Η ύπαρξη του top κουάρκ ήταν αναμενόμενη από το 1977, όταν έγινε η ανακάλυψη του  $b$  κουάρκ στο Fermilab, καθώς είχε ήδη βρεθεί πως τα κουάρκ έρχονται σε διπλέτες ισοσπίν. Για πολλά χρόνια γίνονταν έρευνα για να βρεθεί το σωματίδιο μέχρι που επίσημα το 1995 ανακαλύφθηκε από τα πειράματα CDF και DØ στον  $p\bar{p}$  συγκρουστήρα(collider) Tevatron, στο Fermilab. Στα πειράματα αυτά βρέθηκε η παραγωγή ζεύγους  $t\bar{t}$  που όπως θα δούμε παρακάτω είναι η πιο πιθανή διαδικασία παραγωγής του. Ο λόγος που ήταν ήταν αδύνατη η εύρεση του top σε προηγούμενα πειράματα ήταν η μεγάλη μάζα του (173GeV), που απαιτούσε μεγάλη ενέργεια στο κέντρο μάζας. Ο Tevatron είχε ενέργεια  $\sqrt{s} = 1.8\text{TeV}$ (και αργότερα 1.96 TeV) που τον καθιστούσε μοναδικό για την μελέτη του top. Από τότε και ύστερα το top κουάρκ μελετήθηκε συστηματικά. Η νέα εποχή στην ακρίβεια μετρήσεων με top, ξεκίνησε μόλις άρχισε να τρέχει ο Large Hadron Collider (LHC) στο CERN, όπου μέχρι σήμερα παράχθηκαν εκατοντάδες εκατομμύρια top. Για αυτό το λόγο ο LHC συχνά χαρακτηρίζεται σαν “top quark factory”.

Όσον αφορά τη σημασία της μελέτης του top κουάρκ, αυτή κυρίως έχει να κάνει με την μάζα του. Επειδή είναι πολύ βαρύ προλαβαίνει να διασπαστεί πριν αδρονοποιηθεί. Έτσι είναι το μοναδικό “γυμνό” κουάρκ του οποίου τις ιδιότητες μπορούμε να μετρήσουμε άμεσα από τα προϊόντα της διάσπασης του που φτάνουν στους ανιχνευτές. Μάλιστα είναι το πιο βαρύ σωματίδιο που έχει ανιχνευθεί, και με μεγάλη διαφορά μάζας από τα υπόλοιπα σωματίδια, με αποτέλεσμα να έχει πολλαπλά κανάλια διάσπασης με πλούσιο φασικό χώρο. Επιπλέον λόγω της μάζας, όπως θα φανεί παρακάτω έχει την πιο ισχυρή σύζευξη με το μποζόνιο Higgs, και έχει μεγάλη συνεισφορά στις κβαντικές διορθώσεις από loops. Τέλος, στο πλαίσιο όπου το Καθιερωμένο Πρότυπο θεωρείται ως effective θεωρία, μίας πιο γενικής σε υψηλότερες ενέργειες, οι διορθώσεις στις ιδιότητες του top γίνονται αρκετά σημαντικές. Έτσι οι ακριβείς μετρήσεις των ιδιοτήτων του αποτελούν σημαντικό εργαλείο στην μελέτη της ισχύος του Καθιερωμένου Προτύπου, και τον αποκλεισμό ενδεχόμενων επεκτάσεων του.

Στο κεφάλαιο αυτό υπάρχουνε τρεις κύριες θεματικές ενότητες. Η πρώτη είναι το Καθιερωμένο Πρότυπο της σωματιδιακής φυσικής, που είναι το πλαίσιο στο οποίο περιγράφονται

Όνομα(Σύμβολο)	Μάζα	Ηλ. Φορτίο [e]	Σπίν
electron (e)	0.511 MeV	-1	1/2
muon ( $\mu$ )	105.7 MeV	-1	1/2
tau ( $\tau$ )	1777 MeV	-1	1/2
electron neutrino ( $\nu_e$ )	$\approx 0^*$	0	1/2
muon neutrino ( $\nu_\mu$ )	$\approx 0^*$	0	1/2
tau neutrino ( $\nu_\tau$ )	$\approx 0^*$	0	1/2
up (u)	$2.2^{+0.5}_{-0.4}$ MeV	2/3	1/2
charm (c)	$1.275^{+0.025}_{-0.035}$ GeV	2/3	1/2
top (t)	$173 \pm 0.4$ GeV	2/3	1/2
down (d)	$4.7^{+0.5}_{-0.3}$ MeV	-1/3	1/2
strange (s)	$95^{+9}_{-3}$ MeV	-1/3	1/2
bottom (b)	$4.18^{+0.04}_{-0.03}$ GeV	-1/3	1/2
photon ( $\gamma$ )	0	0	1
gluon (g)	0	0	1
W bosons ( $W^\pm$ )	80.4 GeV	$\pm 1$	1
Z boson ( $Z^0$ )	91.2 GeV	0	1
Higgs (h)	125 GeV	0	0

\*Για τα νετρίνα γνωρίζουμε πως τουλάχιστον 2 εξ'αυτών έχουνε μάζα από την εύρεση ταλαντώσεων. Ωστόσο για τις ενέργειες που μελετάμε είναι αμελητέες (εκτιμάται  $< 2eV$  από πειράματα όπως β-διάσπαση σε τρίτιο.)

Πίνακας 1.1: Τα σωματίδια του Καθιερωμένου Προτύπου και οι ιδιότητες τους.

οι αλληλεπιδράσεις των σωματιδίων που είναι γνωστά. Έπειτα, στην δεύτερη ενότητα, αναφέρονται τα βασικά χαρακτηριστικά που εμφανίζονται στις συγκρούσεις πρωτονίων, καθώς αποτελούν τον τρόπο με τον οποίο παράγονται τα σωματίδια στον Large Hadron Collider (LHC). Τέλος έχοντας το γενικό υπόβαθρο από τις δύο προηγούμενες ενότητες, συζητάται το κομμάτι της φυσικής που εστιάζει η παρούσα διπλωματική, δηλαδή η φυσική του top κουάρκ.

## 1.2 Το Καθιερωμένο Πρότυπο

Το Καθιερωμένο Πρότυπο (Κ.Π.) της σωματιδιακής φυσικής είναι, μέχρι σήμερα, η καλύτερη θεωρία που υπάρχει για την περιγραφή των σωματιδίων που γνωρίζουμε, και των αλληλεπιδράσεων αυτών. Οι θεμελιώδεις δυνάμεις που έχουν βρεθεί είναι οι ισχυρές, οι ασθενείς και οι ηλεκτρομαγνητικές. Έτσι το Κ.Π. αποτελείται από την Κβαντική Χρωμοδυναμική που περιγράφει τις ισχυρές αλληλεπιδράσεις, και το μοντέλο των Glashow-Weinberg-Salam που περιγράφει τις ηλεκτροασθενείς. Τα σωματίδια του Κ.Π. βρίσκονται στον Πίνακα 1.1. Από αυτά τα φερμιόνια χωρίζονται σε δύο κατηγορίες, τα λεπτόνια και τα κουάρκ. Τα λε-

πτόνια έχουνε 3 “γεύσεις”,  $e, \mu, \tau$ , που σε κάθε μία ανήκει το αντίστοιχο φορτισμένο λεπτόνιο (ηλεκτρόνιο, μιονίο, τάφ), και ένα νεutrino (ηλεκτρονίου, μιονίου, ταφ). Τα φορτισμένα λεπτόνια αλληλεπιδρούνε μέσω των ηλεκτρομαγνητικών και των ασθενών αλληλεπιδράσεων, και εάν εξαιρεθεί η μάζα τους έχουνε πανομοιότυπη συμπεριφορά. Τα νεutrino είναι ηλεκτρικά ουδέτερα και αλληλεπιδρούν μόνο μέσω των ασθενών αλληλεπιδράσεων. Η δεύτερη κατηγορία φερμιονίων, τα κουάρκ, είναι 6 συνολικά σε αριθμό, 3 πάνω τύπου (up-type) με ηλεκτρικό φορτίο  $+2/3$  και 3 κάτω τύπου (down-type) με  $-1/3$ . Τα κουάρκ αλληλεπιδρούνε μέσω των ασθενών αλληλεπιδράσεων, όπως και τα λεπτόνια, αλλά επιπλέον μέσω της ισχυρής αλληλεπίδρασης. Αυτή η διαφοροποίηση είναι σημαντική καθώς λόγω αυτής όλα τα κουάρκ εκτός του top δεν μπορούν να βρεθούν ελεύθερα αλλά “παγιδεύονται” μέσα σε συντονισμούς (αδρόνια).

Η θεωρία του Κ.Π. βασίζεται στο γενικό πλαίσιο των θεωριών με συμμετρία βαθμίδας, προκειμένου να περιγράψει τα φορτία και τους διατηρούμενους κβαντικούς. Μέσω των συμμετριών αυτών γίνεται η εισαγωγή των διανυσματικών μποζονίων με σπίν 1, τα οποία λειτουργούν ως φορείς των αλληλεπιδράσεων. Το φωτόνιο είναι ο φορέας της ηλεκτρομαγνητικής αλληλεπίδρασης, τα γκλουόνια της ισχυρής, και τα  $W^\pm, Z^0$  της ασθενούς. Συγκεκριμένα η ομάδα συμμετρίας του Κ.Π. είναι η σύνθεση  $SU(3)_C \times SU(2)_L \times U(1)_Y$ . Επιπλέον, είναι απαραίτητη η ύπαρξη ενός μηχανισμού για να αποδοθεί μάζα στα  $W^\pm, Z^0$  και στα φερμιόνια. Αυτός είναι ο μηχανισμός Higgs, στον οποίο συμμετέχει το τελευταίο σωματίδιο, το μποζόνιο Higgs. Το Higgs είναι αφόρτιστο και δεν έχει σπίν. Παρακάτω θα δούμε πώς τα δύο βασικά συστατικά που αναφέρθηκαν, η συμμετρία και ο μηχανισμός Higgs, δομούνε το Κ.Π. και τι συνέπειες έχουνε.

### 1.2.1 Η Κβαντική Χρωμοδυναμική

Το κομμάτι  $SU(3)_C$  στην ομάδα συμμετρίας του Κ.Π. αποτελεί την Κβαντική Χρωμοδυναμική (ΚΧΔ ή QCD) που περιγράφει την αλληλεπίδραση των κουάρκ με τα γκλουόνια, που φέρουν ισχυρό φορτίο ή χρώμα (color). Για αυτό ο μετασχηματισμός της συμμετρίας δρά μόνο πάνω στα πεδία κουάρκ, ενώ των λεπτονίων μένουν όπως έχουν

$$\begin{aligned} \psi_q(x) &\rightarrow e^{i\theta^a(x)T^a} \psi_q(x) & q = u, d, s, c, b, t \\ \psi_l(x) &\rightarrow \psi_l(x) & l = e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau \end{aligned} \quad (1.1)$$

όπου  $T^a$  είναι οι γεννήτορες της  $SU(3)$ , που μπορούν να απαρασταθούν με τους 8 πίνακες Gell-Mann,  $T^a = \lambda_a/2$ . Δηλαδή τα πεδία των κουάρκ είναι τριπλέτες χρώματος (color triplets), ενώ των λεπτονίων απλέτες (colour singlets). Ο αριθμός των χρωμάτων  $N_c = 3$  προέρχεται από μετρήσεις όπως ο λόγος  $R = \sigma(e^+e^- \rightarrow \mu^+\mu^-)/\sigma(e^+e^- \rightarrow \text{hadrons}) \sim N_c$ , το πλάτος διάσπασης του  $\pi_0$  σε δύο φωτόνια  $\Gamma(\pi_0 \rightarrow \gamma\gamma) \sim N_c^2$  και άλλα.

Τα πεδία των γκλουονίων  $G_\mu^a$  εισάγονται μέσω της συναλλοίωτης παραγώγου (covariant derivative)

$$D_\mu = \partial_\mu + ig_s T^a G_\mu^a(x) \quad (1.2)$$

όπου  $g_s$  η ισχυρή σύζευξη (strong coupling) η οποία χρειάζεται για να έχουμε  $SU(3)$  αναλ-

λοίωτη  $\mathcal{L}$  ως προς τους κινητικούς όρους των κουάρκ

$$\mathcal{L}_{quark}^{kin} = \sum_q i\bar{\psi}_q \gamma^\mu D_\mu \psi_q \quad (1.3)$$

Προκειμένου να ικανοποιείται η  $SU(3)$  αναλλοιώτητα τα γκλουόνια μετασχηματίζονται ως

$$G_\mu^a(x) \rightarrow G_\mu^a(x) + 1/g_s \partial_\mu \theta^a(x) - f^{abc} \theta^b(x) G_\mu^c(x) \quad (1.4)$$

όπου  $f^{abc}$  είναι οι σταθερές δομής (structure constants), που προκύπτουν από την  $[T^a, T^b] = i f^{abc} T^c$ . Από τον νόμο μετασχηματισμού είναι προφανές πως εάν υπήρχαν όροι μάζας  $\sim (G^\mu)^2$  θα παραβιάζονταν η συμμετρία. Για αυτό τα γκλουόνια είναι άμαζα. Για τον κινητικό όρο των γκλουονίων χρειάζεται ένας τανυστής αναλλοίωτος ως προς την  $SU(3)$ . Ο τανυστής αυτός μπορεί να ορισθεί μέσω της σχέσης

$$\begin{aligned} G_{\mu\nu}^a T^a &\equiv \frac{i}{g_s} [D_\mu, D_\nu] \\ \Rightarrow G_{\mu\nu}^a &= \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c \end{aligned} \quad (1.5)$$

και ονομάζεται “field strength”. Ο τελευταίος όρος είναι που διαφοροποιεί την Μη Αβελιανή από την Αβελιανή περίπτωση ( $f^{abc} = 0$ ), όπως είναι η  $U(1)$  του ηλεκτρομαγνητισμού. Λόγω αυτού δημιουργούνται αλληλεπιδράσεις μεταξύ των γκλουονίων, πράγμα που δεν ισχύει για τα φωτόνια. Οι αλληλεπιδράσεις αυτές μπαίνουν στην Λαγκραντζιανή από τον κινητικό όρο  $-\frac{1}{4} G_{\mu\nu}^a G^{a,\mu\nu}$  και αντιστοιχούν σε αλληλεπιδράσεις 3 και 4 γκλουονίων, με τα αντίστοιχα διαγράμματα Feynman



Το γεγονός πως τα γλουόνια φέρουν ισχυρό φορτίο και αλληλεπιδρούν, είναι άμεσα συσχετισμένο με την ιδιότητα της σταθεράς σύζευξης της QCD να μειώνεται με την ενέργεια, την λεγόμενη “ασυμπτωτική ελευθερία” (asymptotic freedom). Η ιδιότητα αυτή προκύπτει λόγω της εξάρτησης της σταθεράς σύζευξης από την ενέργεια κλίμακα επανακανονικοποίησης  $\mu$ . Η εξάρτηση αυτή περιγράφεται από την εξίσωση ομάδας επανακανονικοποίησης

$$\mu \frac{dg_s}{d\mu} = \beta(g_s) = \frac{g_s^3}{(4\pi)^2} b_0 + \frac{g_s^5}{(4\pi)^4} b_1 + \dots \quad (1.6)$$

που σε 1-loop

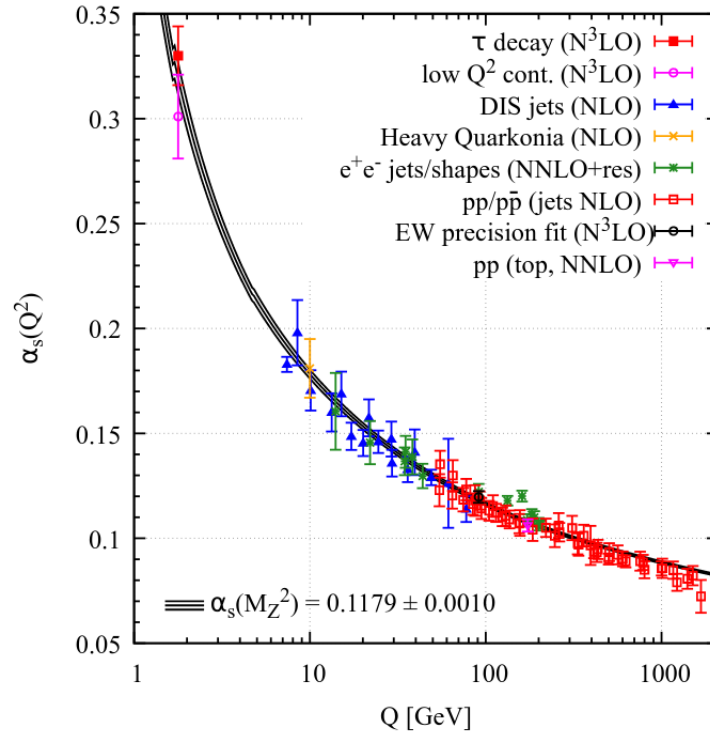
$$b_0 = -\frac{11}{3} C_A + \frac{4}{3} T_R n_f = -11 + \frac{2}{3} n_f \quad (1.7)$$

όπου  $C_A = 3$  είναι ο παράγοντας χρώματος (color factor) για εκπομπή γκλουονίου από γλουόνιο, ενώ το  $T_R = 1/2$  ο παράγοντας χρώματος για μετατροπή ενός γκλουονίου σε ζεύγος  $q\bar{q}$

(gluon splitting). Το  $n_f$  είναι ο αριθμός των ενεργών γεύσεων κουάρκς, δηλαδή συμπεριλαμβάνονται τα κουάρκς που έχουν  $m < \mu$  καθώς τα υπόλοιπα έχουν μικρή συνεισφορά. Για  $n_f = (3, 4, 5, 6)$  έχουμε  $b_0 = (-9, -25/3, -23/3, -7)$ . Έτσι σε 1-loop η σταθερά σύζευξης  $\alpha_s = g_s^2/4\pi$  έχει την εξάρτηση

$$\alpha_s(\mu) = \frac{\alpha_s(\mu_0)}{1 - \alpha_s(\mu_0) \frac{b_0}{2\pi} \log\left(\frac{\mu}{\mu_0}\right)} \quad (1.8)$$

Η συνάρτηση  $\beta(g_s)$  λοιπόν είναι αρνητική και αυτό συνεπάγεται την ασυμπτωτική ελευθερία στις υψηλές ενέργειες. Επιπλέον σημαίνει πως έχουμε πολύ μεγάλη σύζευξη σε χαμηλές ενέργειες, ιδιότητα που όπως θα δούμε συμβάλλει στην λεγόμενη “αδρονιοποίηση” των κουάρκς. Η εξάρτηση της  $\alpha_s$  από θεωρητικούς υπολογισμούς και η σύγκριση με πειραματικά δεδομένα φαίνεται στο Σχήμα 1.1. Βλέπουμε πως στην περιοχή ενεργειών 100-1000 GeV η τιμή της  $\alpha_s$  είναι 0.8-0.12, που είναι αρκετά μεγαλύτερη συγκριτικά με τις υπόλοιπες αλληλεπιδράσεις. Επειδή οι ενεργές διατομές, είναι ανάλογες της σταθεράς σύζευξης, οι διαδικασίες της QCD έχουν μεγαλύτερη πιθανότητα εμφάνισης. Μία σημαντική παρατήρηση από την (1.7), είναι



Σχήμα 1.1: Εξάρτηση της σταθεράς ζεύξης  $\alpha_s$  από την ενεργειακή κλίμακα και πειραματικά δεδομένα. Σαν  $\mu = Q$  δηλαδή η μεταφερόμενη τετραορμή στην διαδικασία που μελετάται. Ο παγκόσμιος μέσος όρος για κλίμακα  $M_Z$  είναι  $\alpha_s(M_Z^2) = 0.1179 \pm 0.0010$ .

πως υπάρχει ένας “ανταγωνισμός” των αλληλεπιδράσεων μεταξύ των γκλουονίων που τείνουν να μειώσουν το ισχυρό φορτίο σε μεγάλες ενέργειες, και των ζευγών  $q\bar{q}$  που τείνουν να το αυξήσουν. Αυτή είναι η διαφορά με την ηλεκτροδυναμική, όπου τα φωτόνια δεν αλληλεπιδρούν μεταξύ τους, και έτσι με την αύξηση της ενέργειας έχουμε αύξηση της σύζευξης  $\alpha_{em}$ .

Επιπλέον από την ίδια σχέση φαίνεται πως η σύζευξη απειρίζεται σε κάποια ενεργειακή κλίμακα  $\Lambda_{QCD}$  που είναι της τάξης 200MeV (πόλος Landau). Η κλίμακα αυτή είναι εκείνη όπου σταματάει η ισχύς της θεωρίας διαταραχών και δίνει την χαρακτηριστική χωρική και χρονική κλίμακα της αδρονοποίησης  $\tau \sim 1/\Lambda_{QCD} = 1\text{fm} = 10^{-23}\text{s}$ .

### 1.2.2 Οι ηλεκτρασθενείς αλληλεπιδράσεις

Έχοντας λοιπόν αναφέρει την φυσική μέσω της  $SU(3)$  ομάδας, μένει το υπόλοιπο μέρος της ομάδας συμμετρίας του Κ.Π., δηλαδή η  $SU(2)_L \times U(1)_Y$  η οποία περιγράφει τις ηλεκτρασθενείς αλληλεπιδράσεις (EW)<sup>1</sup>. Ο δείκτης  $Y$  της  $U(1)$  αντιστοιχεί στο υπερφορτίο (hypercharge) που είναι ο γεννήτορας της συμμετρίας. Δηλαδή ένας μετασχηματισμός  $U(1)_Y$  είναι της μορφής

$$\psi(x) \rightarrow e^{i\theta(x)Y} \psi(x) \quad , \quad Y = \text{αριθμός} \quad (1.9)$$

Ο δείκτης  $L$  στην ομάδα  $SU(2)$  δείχνει πως οι μετασχηματισμοί αυτής δρύνε μη τετριμένα μόνο στις Left συνιστώσες των  $\psi(x)$ . Δηλαδή αν έχουμε δύο φερμιονικά πεδία  $\psi_1(x), \psi_2(x)$  που αλληλεπιδρούνε μέσω της ασθενούς τότε

$$\begin{pmatrix} \psi_{1L} \\ \psi_{2L} \end{pmatrix} \rightarrow e^{i\theta^a(x)T^a} \begin{pmatrix} \psi_{1L} \\ \psi_{2L} \end{pmatrix} \quad , \quad T^a = \frac{\tau_a}{2} \quad \text{πίνακες Pauli } a=1,2,3$$

$$\psi_{1R} \rightarrow \psi_{1R} \quad , \quad \psi_{2R} \rightarrow \psi_{2R} \quad (1.10)$$

Έτσι τα κομμάτια αριστερής χειραλικότητας βρίσκονται σε διπλέτα  $SU(2)$  ενώ τα δεξιούς χειραλικότητας είναι απλέτες. Η συμπεριφορά αυτή της ασθενούς αλληλεπίδρασης προέρχεται από την παραβίαση της ομοτιμίας, που χρειάζεται για να εξηγήσει ασυμμετρίες γωνιακών κατανομών στα πειράματα (όπως στο πείραμα Wu με  $\beta$ -διάσπαση  $Co$  σε  $Ni$ ), τον λόγο διάσπασης πιονίων σε μόνια κ.λ.π. .

Από τον τρόπο που συμμετέχουν στις ασθενείς αλληλεπιδράσεις, τα πεδία λεπτονίων και των κουάρκ αναπαρίστανται, ως προς την  $SU(2)_L \times U(1)_Y$ , ως

<sup>1</sup>Σε όσα θα αναφερθούν εννοείται η ύπαρξη των γκλουονίων και του χρώματος των κουάρκ, αλλά η προσοχή εστιάζεται στο EW κομμάτι οπότε παραλείπονται

*Quarks :*

$$SU(2) \text{ doublet} \quad Q_i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix} \text{ με } Y_{Q_i} = +\frac{1}{6} \quad , \quad SU(2) \text{ singlets} \quad u_R^i \text{ με } Y_{u_R^i} = +\frac{2}{3} \quad d_R^i \text{ με } Y_{d_R^i} = -\frac{1}{3}$$

$$i = 1, 2, 3$$

up-type q's:

$$u^1 = u \quad u^2 = c \quad u^3 = t$$

down-type q's:

$$d^1 = d \quad d^2 = s \quad d^3 = b$$

*Leptons :*

$$SU(2) \text{ doublet} \quad L_i = \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix} \text{ με } Y_{L_i} = -\frac{1}{2} \quad , \quad SU(2) \text{ singlet} \quad e_R^i \text{ με } Y_{e_R^i} = +\frac{2}{3}$$

$$i = 1, 2, 3$$

flavors:

$$1 = e \quad 2 = \mu \quad 3 = \tau$$

Τα δεξιόστροφα μέρη των νετρίνων δεν συμπεριλαμβάνονται στα παραπάνω καθώς δεν αλληλεπιδρούν με κανέναν τρόπο. Τα νετρίνα δηλαδή που παρατηρούνται στην φύση είναι καθαρά αριστερόστροφα. Οι τιμές των υπερφορτίων είναι τέτοιες ώστε, όπως θα φανεί, τα ηλεκτρικά φορτία των σωματιδίων να είναι σωστά. Ήδη μπορεί να παρατηρήσει κανείς πως στις απλές  $SU(2)$ , δηλαδή τα  $R$  κομμάτια, έχει δωθεί ως υπερφορτίο το ηλεκτρικό φορτίο. Επιπλέον τα πεδία βαθμίδας της ομάδας συμμετρίας είναι 3 από την  $SU(2)$ , τα  $W_1^\mu, W_2^\mu, W_3^\mu$  και 1 για την  $U(1)$  το  $B^\mu$ . Τα πεδία αυτά είναι τόσα όσα χρειάζονται για να εισαχθούν τα διανυσματικά μποζόνια που γνωρίζουμε από τα πειράματα  $W^\pm, Z^0$  για ασθενείς και το  $\gamma$  του ηλεκτρομαγνητισμού.

Ωστόσο είναι γνωστό πειραματικά πως οι ασθενείς αλληλεπιδράσεις έχουν μικρή εμβέλεια, πράγμα που ισοδυναμεί με την ύπαρξη μάζας για τα  $W^\pm, Z^0$ , ενώ η ηλεκτρομαγνητική αλληλεπίδραση άπειρη που είναι συμβατό με την  $m_\gamma = 0$ .<sup>2</sup> Εν γένει οι όροι μάζας σπάνε την συμμετρία βαθμίδας στην Λαγκραντζιανή, όπως είδαμε και στα γκλουόνια που πρέπει να είναι άμαζα. Επιπλέον ακόμα και αν κανείς βάλει τους όρους μάζας άμεσα στην θεωρία, τότε ο διαδότης των  $W$  και  $Z$  μποζονίων βαθμίδας παίρνει τη μορφή

$$D_{\mu\nu}(p) = \frac{-i}{p^2 - m^2 + i\epsilon} \left( g_{\mu\nu} - \frac{p_\mu p_\nu}{m^2} \right) \quad (1.11)$$

όπου ο δεύτερος όρος στην παρένθεση εισέρχεται λόγω της επιπλέον πόλωσης και σχετίζεται με την ύπαρξη μάζας. Αυτός ο όρος δημιουργεί απειρισμούς σε χβαντικές διορθώσεις, που για να αφαιρεθούν χρειάζεται εισαγωγή νέων τελεστών στην θεωρία που είναι μη επανακανονικοποιησιμοι. Το γεγονός αυτό πρακτικά σημαίνει πως θα μείωνε την προβλεπτική ικανότητα της θεωρίας καθώς θα χρειάζονταν άπειρες παραμέτρους που θα έπρεπε να προσδιοριστούν από τα δεδομένα<sup>3</sup>.

<sup>2</sup>Στο μέσο δυναμικό της αλληλεπίδρασης δύο σωματιδίων μέσω ενδιάμεσης διάδοσης σωματιδίου μάζας  $m$ ,  $V(r)$ , υπάρχει ένας παράγοντας  $e^{-mr}$ .

<sup>3</sup>Στην πράξη δεν θα χρειάζονταν όλες, αλλά αυτές που θα είχαν κύρια συνεισφορά. Ωστόσο και πάλι θα χρειάζονταν επιπλέον παράμετροι.



Αυτό που μπορεί να κάνει κανείς με στόχο να λύσει τα προβλήματα που αναφέρθηκαν, είναι να ξεκινήσει έχοντας την συμμετρία της ηλεκτρασθενούς, η οποία να σπάσει με δυναμικό τρόπο. Για να συμβεί κάτι τέτοιο εισάγονται βαθμωτά πεδία, των οποίων οι όροι στην  $\mathcal{L}$  σέβονται την συμμετρία της θεωρίας. Για τα πεδία αυτά εισάγεται δυναμικό το οποίο διαθέτει εκφυλισμένες καταστάσεις κενού, λόγω της συμμετρίας,  $|\Omega\rangle$  που όταν το σύστημα “επιλέξει” αναγκαστικά μία εξ’αυτών τότε παύει να ισχύει η συμμετρία. Δηλαδή η ιδέα είναι πως αν και η  $\mathcal{L}$  διαθέτει την συμμετρία, το κενό της θεωρίας την παραβιάζει, με αποτέλεσμα να σπάει. Μία φυσική διεργασία με αυτήν την ιδιότητα υπόκεινται σε “αυθόρμητο σπάσιμο συμμετρίας”. Η διαδικασία αυτή έχει δύο πολύ χρήσιμα χαρακτηριστικά.

1. Μπορεί κανείς να μελετήσει απευθείας τις αλληλεπιδράσεις γύρω από το κενό της θεωρίας, θεωρώντας παραμετροποίηση των βαθμωτών πεδίων γύρω από το αυτό, δηλαδή δεν χρειάζεται η αναλυτική μελέτη της δυναμικής μέχρι να φτάσει το σύστημα στην ισορροπία.
2. Στην μελέτη γύρω από το σημείο ισορροπίας όπου η συμμετρία έχει σπάσει, συνεχίζει να υπάρχει η πληροφορία της σπασμένης συμμετρίας μέσα από σχέση των παραμέτρων, πράγμα που ενισχύει την ικανότητα προβλέψεων της θεωρίας καθώς η ισχύς τέτοιων σχέσεων μπορεί να μελετηθεί μέσω πειραματικών δεδομένων.

Επομένως στο Κ.Π. χρησιμοποιείται για να αποδωθούν οι μάζες των μποζονίων ο μηχανισμός Higgs, που είναι μία διαδικασία σπασίματος συμμετρίας. Αρχικά πρέπει να εισάγει κανείς τα βαθμωτά πεδία με τέτοιο τρόπο που να ικανοποιείται η συμμετρία στην  $\mathcal{L}$ , και να υπάρχουν αρκετοί βαθμοί ελευθερίας για αποδωθούν οι μάζες, καθώς τα μποζόνια  $W^\pm, Z$  με την απόκτηση μάζας θα έχουνε μία επιπλέον πόλωση. Το σπάσιμο της συμμετρίας πρέπει να γίνει με τέτοιο τρόπο ώστε τελικά να παραμείνει η  $U(1)_Q$ . Επιπλέον επειδή η  $SU(2)_L$  είναι χειραλική συμμετρία, δεν μπορούνε να υπάρχουνε όροι μάζας για τα φερμιόνια  $m\bar{\psi}\psi$  και ο πιο απλός τρόπος να αντιμετωπιστεί το πρόβλημα είναι η χρήση όρων Yukawa, δηλαδή όροι της μορφής  $y\bar{\psi}\psi$ . Για να ικανοποιηθούν οι απαιτήσεις που αναφέρθηκαν η πιο απλή λύση είναι η προσθήκη μίας  $SU(2)$  διπλέτας με μιγαδικά βαθμωτά

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.12)$$

η οποία έχει  $Y = 1/2$ . Για την διπλέτα αυτή η  $SU(2)_L \times U(1)_Y$  αναλλοίωτη Λαγκρατζιανή,  $\mathcal{L}_{\text{Higgs}}$  είναι

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi^\dagger \Phi) \quad (1.13)$$

όπου  $D_\mu$  η συνναλοίωτη παράγωγος

$$D_\mu = \partial_\mu - igW_\mu^a T^a - ig' Y B_\mu \quad (1.14)$$

με  $g, g'$  οι συζεύξεις για  $SU(2)_L$  και  $U(1)_Y$  αντίστοιχα. Το δυναμικό είναι

$$V(\Phi^\dagger \Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.15)$$

με  $\mu^2, \lambda > 0$ . Το δυναμικό αυτό ελαχιστοποιείται όταν  $|\Phi^\dagger \Phi| = \frac{\mu^2}{2\lambda}$ , που είναι η συνθήκη που ικανοποιούν οι καταστάσεις κενού. Έτσι το δυναμικό Higgs επάγει μία αναμενόμενη τιμή του κενού (vacuum expectation value - vev) διάφορη του μηδενός η οποία χωρίς βλάβη γενικότητας μπορεί να θεωρηθεί πως είναι πραγματική και βρίσκεται στην κάτω συνιστώσα της διπλέτας  $\Phi$ . Τότε μία ανάπτυξη γύρω από την κατάσταση κενού που επιλέχθηκε μπορεί να γραφεί ως

$$\Phi = e^{i\xi^a(x)\tau^a/v} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} + \frac{h(x)}{\sqrt{2}} \end{pmatrix} \quad (1.16)$$

όπου  $v = \sqrt{\frac{\mu^2}{\lambda}}$ , το  $h(x)$  είναι το πεδίο του μποζονίου Higgs, ενώ τα  $\xi^a(x)$  αντιστοιχούν στα λεγόμενα “Goldstone bosons” τα οποία είναι πάντοτε όσα οι γεννήτορες των σπασμένων συμμετριών. Ο λόγος που χρησιμοποιήθηκε η παραμετροποίηση αυτή για το  $\Phi$  γύρω από το κενό, είναι πως φαίνεται άμεσα πως τα μποζόνια Goldstone μπορούν να εξαφανισθούν μέσω ενός μετασχηματισμού βαθμίδας

$$\Phi \rightarrow e^{-i\xi^a(x)\tau^a/v} \Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.17)$$

Η βαθμίδα αυτή λέγεται “unitary gauge”. Στην βαθμίδα αυτή οι βαθμοί ελευθερίας των μποζονίων Goldstone που εξαλείφονται, γίνονται οι πολώσεις των διανυσματικών μποζονίων που αντιστοιχούν στους γεννήτορες των σπασμένων συμμετριών. Οι μάζες των διανυσματικών μποζονίων φαίνονται από την ανάπτυξη του κινητικού όρου

$$\begin{aligned} (D_\mu \Phi)^\dagger (D^\mu \Phi) &= \left| \begin{pmatrix} \partial_\mu + \frac{i}{2}(gW_\mu^3 + g'B_\mu) & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & \partial_\mu + \frac{i}{2}(-gW_\mu^3 + g'B_\mu) \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \right|^2 \\ &= \frac{1}{2}(\partial_\mu h)^2 + m_W^2 W_\mu^+ W^{-\mu} \left(1 + \frac{h}{v}\right)^2 + \frac{1}{2}m_Z^2 Z_\mu Z^\mu \left(1 + \frac{h}{v}\right)^2 \end{aligned} \quad (1.18)$$

που στο πέρασμα στην δεύτερη σειρά η έκφραση γράφηκε στις καταστάσεις μάζας, δηλαδή στα φυσικά μποζόνια που είναι τα

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2) \quad (1.19)$$

$$Z_\mu = \cos \theta_w W_\mu^3 - \sin \theta_w B_\mu \quad (1.20)$$

$$A_\mu = \sin \theta_w W_\mu^3 + \cos \theta_w B_\mu \quad (1.21)$$

με  $\tan \theta_w = \frac{g'}{g}$ . Έτσι οι μάζες είναι

$$m_W = \frac{gv}{2} \quad m_Z = \frac{gv}{2 \cos \theta_w} \quad m_A = 0 \quad (1.22)$$

Προκύπτει το επιθυμητό φάσμα για τις μάζες που είναι τα δύο  $W$  μποζόνια να έχουνε ίση μάζα και το  $Z$  με μάζα μεγαλύτερη, εφόσον  $\cos \theta_w < 1$ . Το πεδίο  $A_\mu$  είναι το φωτόνιο και δεν έχει μάζα, ενώ αντιστοιχεί στον γεννήτορα της παραμένουσας συμμετρίας που είναι ο τελεστής του

ηλεκτρικού φορτίου. Για να βρεθεί ο τελεστής του φορτίου γράφει κανείς την συναλλοίωτη παράγωγο  $D_\mu$  ως προς τα φυσικά πεδία  $Z, A$

$$D_\mu = \dots - igW_3^\mu T^3 - ig'YB_\mu = -iA_\mu(g \sin \theta_w T^3 + g' \cos \theta_w Y) - iZ_\mu(g \cos \theta_w T^3 - g' \sin \theta_w Y) \quad (1.23)$$

Επομένως εφόσον  $\tan \theta_w = g'/g \Rightarrow g \sin \theta_w = g' \cos \theta_w \equiv e$ , ο όρος με το φωτόνιο  $A_\mu$  γράφεται σαν  $-ieA_\mu Q$  με

$$Q \equiv T^3 + Y \quad (1.24)$$

Έτσι οι τιμές που δίνονταν για το υπερφορτίο είναι τέτοιες που να προκύπτουν οι σωστές τιμές του ηλεκτρικού φορτίου μετά το σπάσιμο της συμμετρίας.

Η μάζα του Higgs και οι self-interactions αυτού, προκύπτουν από την ανάπτυξη του δυναμικού γύρω από το κενό

$$\begin{aligned} V(h) &= -\frac{\mu^2}{2}(v+h)^2 + \frac{\lambda}{4}(v+h)^4 \\ &= \frac{\lambda}{4}h^4 + \lambda v h^3 + \lambda v^2 h^2 \end{aligned} \quad (1.25)$$

δηλαδή η μάζα του Higgs είναι  $m_h = \sqrt{2\lambda}v = \sqrt{2}\mu$ , και οι υπάρχουσες αλληλεπιδράσεις μεταξύ τριών και τεσσάρων  $h$ .

Πηγαίνοντας στο κομμάτι των φερμιονίων αρχικά οι κινητικοί όροι που είναι αναλλοίωτοι ως προς την συμμετρία είναι

$$\begin{aligned} \mathcal{L}_{\text{fermion}} &= i\bar{L}_i \gamma^\mu D_\mu L_i + i\bar{e}_{Ri} \gamma^\mu D_\mu e_{Ri} \\ &+ i\bar{Q}_i \gamma^\mu D_\mu Q_i + i\bar{u}_{Ri} \gamma^\mu D_\mu u_{Ri} + i\bar{d}_{Ri} \gamma^\mu D_\mu d_{Ri} \end{aligned}$$

Τώρα όσον αφορά την εισαγωγή των όρων μάζας των φερμιονίων με τρόπο που να σέβεται την συμμετρία της θεωρίας, αυτό επιτυγχνάζεται με την εισαγωγή όρων Yukawa

$$\mathcal{L}_{\text{Yukawa}} = -y_e^i \bar{L}^i \Phi e_R^i - Y_d^{ij} \bar{Q}^i \Phi d_R^j - Y_u^{ij} \bar{Q}^i \bar{\Phi} u_R^j + h.c. \quad (1.26)$$

όπου το  $\bar{\Phi}$  ορίζεται ως  $\bar{\Phi} = i\tau_2 \Phi^*$ , που έχει μη μηδενικό το άνω μέρος της διπλέτας, ώστε μετά το σπάσιμο συμμετρίας να αποδίδονται σωστά οι μάζες για τα up-type κουάρκ. Η εμφάνιση των μαζών μετά το σπάσιμο συμμετρίας εμφανίζεται μέσω της μη μηδενικής αναμενόμενης τιμής του κενού. Για παράδειγμα για τα λεπτόνια

$$-y_e \bar{L} \Phi e_R = -\frac{y_e}{\sqrt{2}} \begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \begin{pmatrix} 0 \\ v+h \end{pmatrix} = -m_e \bar{e} e + \dots \quad (1.27)$$

όπου  $m_e = \frac{y_e v}{\sqrt{2}}$ .

Πολύ σημαντικό είναι πως για τα κουάρκς υπάρχουν πίνακες με Yukawa couplings, καθώς από τις βασικές αρχές της θεωρίας δεν αποκλείονται οι μη διαγώνιοι όροι. Αυτό οδηγεί στο γεγονός καταστάσεις μάζας των down-type κουάρκ  $d'_i$  να είναι διαφορετικές από εκείνες των γεύσεων  $d_i$ , και συνδέονται μέσω ενός unitary πίνακα που ονομάζεται Cabbibo-Kobayashi-Maskawa (CKM)

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.28)$$

Τα μέτρα των στοιχείων του πίνακα αυτού κατά απόλυτο τιμή έχουν μετρηθεί να είναι

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.230 & 0.975 & 0.041 \\ 0.008 & 0.040 & 0.999 \end{pmatrix} \quad (1.29)$$

Για τα λεπτόνια κάτι τέτοιο δεν συμβαίνει στο Κ.Π. επειδή τα νετρίνα θεωρούνται άμαζα. Ο πίνακας CKM εφόσον είναι unitary έχει 9 ανεξάρτητες παραμέτρους. Από αυτές οι 5 σχετικές φάσεις των κουάρκ μπορούν να απορροφηθούν με επαναορισμό των πεδίων, και μένουν 4 ανεξάρτητες παράμετροι, που είναι 3 γωνίες Euler και μία μιγαδική φάση  $e^{i\delta}$  που ευθύνεται για τα φαινόμενα παραβίασης της συμμετρίας CP στις αλληλεπιδράσεις με αλλαγή γεύσης. Επιπλέον εφόσον τα πλάτη μετάβασης από το κουάρκ  $i$  στο  $j$  είναι ανάλογα του  $|V_{ij}|^2$ , λόγω της unitarity που ουσιαστικά εκφράζει την διατήρηση της συνολικής πιθανότητας, ισχύει ότι

$$\sum_i |V_{ij}|^2 = \sum_j |V_{ij}|^2 = 1 \quad (1.30)$$

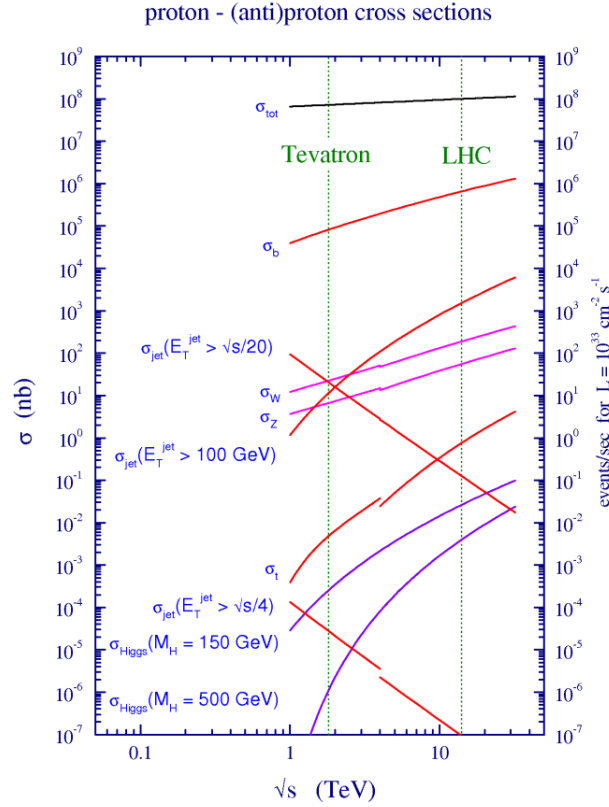
Συνοψίζοντας, μέσα από τα παραπάνω φαίνεται πως η θεωρία των ηλεκτρασθενών μέσω της χρήσης του σπασίματος συμμετρίας, είναι μία θεωρία με ισχυρή προβλεπτική ικανότητα καθώς με σχετικά λίγες παραμέτρους περιγράφεται όλη η δυναμική των σωματιδίων που είναι γνωστά. Με την εισαγωγή του μηχανισμού Higgs υπάρχουν σχέσεις μεταξύ των παραμέτρων αυτών που έχουν ελεγχθεί πειραματικά και μέχρι σήμερα βρίσκονται σε καλή συμφωνία με τα δεδομένα.

### 1.3 Φυσική Συγκρούσεων Πρωτονίων

Στον Large Hadron Collider (LHC) τα σωματίδια παράγονται μέσω συγκρούσεων πρωτονίων που έχουν επιταχυνθεί σε πολύ υψηλές ενέργειες. Οι δύο δέσμες των πρωτονίων έχουν ίσες ενέργειες  $E_b$  και, όσο το δυνατόν, αντίθετες διευθύνσεις ώστε να μεγιστοποιείται η ενέργεια του κέντρου μάζας  $\sqrt{s}$  που είναι η οφέλιμη ενέργεια για την παραγωγή σωματιδίων. Η διεύθυνση που κινούνται οι δέσμες θεωρείται ο άξονας  $z$ . Δηλαδή για τα δύο πρωτόνια οι τετραορμές είναι

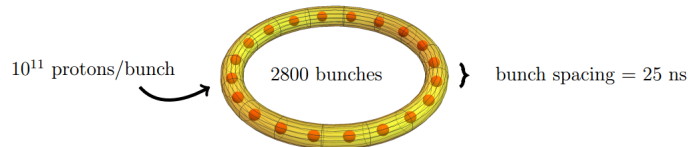
$$P_1 = (E_b, 0, 0, E_b) \quad P_2 = (E_b, 0, 0, -E_b) \quad (1.31)$$

με αποτέλεσμα το σύστημα του εργαστηρίου και του κέντρου μάζας να ταυτίζονται. Η ενεργός διατομή για την αλληλεπίδραση δύο πρωτονίων είναι  $\sim 100mb$ , όπως φαίνεται στο Σχήμα 1.2. Από την συνολική ενεργό διατομή ένα μέρος είναι ενδιαφέρον για εμάς, καθώς θέλουμε τις ανελαστικές σκεδάσεις όπου τα δύο πρωτόνια σπάνε και αλληλεπιδρούν τα παρτόνια τους, και όχι τις ελαστικές ή diffractive. Η ενεργός διατομή για τέτοιες σκεδάσεις είναι περίπου τα  $70mb$  της συνολικής. Αν και αυτή η ενεργός διατομή είναι μεγάλη, από το ίδιο Σχήμα βλέπει κανείς πως οι ενεργές διατομές για διαδικασίες όπως η παραγωγή top ή Higgs είναι πολύ μικρή. Έτσι για να παράγονται πολλά γεγονότα με ενδιαφέρουσα φυσική χρειάζεται οι συγκρουόμενες δέσμες να αποτελούνται από πακέτα (bunches) με πολλά πρωτόνια. Στον LHC



Σχήμα 1.2: Ενεργές διατομές για κάποιες τυπικές διαδικασίες σε συγκρούσεις πρωτονίων συναρτήσει της ενέργειας κέντρου μάζας.

η κάθε δέσμη αποτελείται από περίπου  $N_b \approx 2800$  πακέτα με  $10^{11}$  πρωτόνια έκαστη, οι οποίες συγκρούονται με συχνότητα  $f \sim 25 \text{ ns}$ .



Τα χαρακτηριστικά της δέσμης καθορίζουν την φωτεινότητα (luminosity) που δίνεται από την σχέση

$$\mathcal{L} = F \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} f N_b \quad (1.32)$$

όπου  $\sigma_x, \sigma_y$  είναι τα εύρη του κατακόρυφου γκαουσιανού προφίλ της δέσμης στις διευθύνσεις  $x$  και  $y$  αντίστοιχα, και είναι της τάξης των  $10 \mu\text{m}$ . Ο παράγοντας  $F$  είναι ένας γεωμετρικός διορθωτικός παράγοντας. Η ποσότητα αυτή δείχνει πόσοι πιθανοί συνδυασμοί συγκρούσεων υπάρχουνε ανά μονάδα επιφάνειας της δέσμης ανά μονάδα χρόνου. Επομένως ο ρυθμός (rate)  $R$  παραγωγής γεγονότων ενεργού διατομής  $\sigma$  δίνεται από την σχέση

$$R = \sigma L \quad (1.33)$$

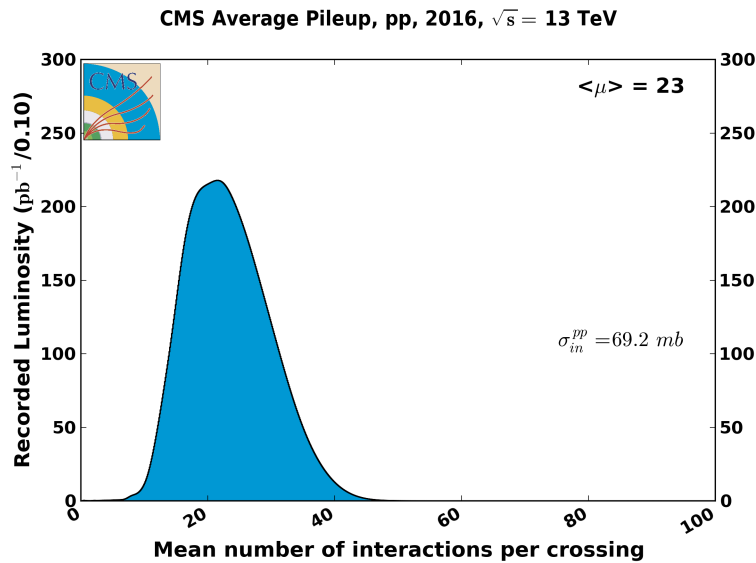
Η Luminosity ορίζεται ανά μονάδα χρόνου, για αυτό ονομάζεται και “instantaneous luminosity”. Εάν κανείς θέλει να βρεί τον αριθμό των γεγονότων(events) για δεδομένο χρονικό διάστημα τότε ολοκληρώνει ως προς  $t$

$$N = \int R dt = \sigma \int \mathcal{L} dt \equiv \sigma \mathcal{L}_{int} \quad (1.34)$$

όπου  $\mathcal{L}_{int}$  είναι η integrated luminosity.

### 1.3.1 Pile-up

Η αύξηση της instantaneous luminosity οδηγεί σε αύξηση ενός φαινομένου που ονομάζεται “pile-up”. Γενικά ο όρος “pile-up” αναφέρεται στην κατάσταση όπου ο ανιχνευτής καταγράφει περισσότερα από ένα γεγονότα ταυτόχρονα. Στις συγκρούσεις πρωτονίων του LHC η κατάσταση αυτή περιλαμβάνει δύο επιμέρους κομμάτια, το “in time” και το “out of time” pile-up. Το in time pile-up αναφέρεται στην κατάσταση όπου στην σύγκρουση δύο bunches από πρωτόνια γίνονται παραπάνω από μία σχεδιάσεις πρωτονίων. Ο αριθμός τέτοιων συγκρούσεων ακολουθεί κατανομή Poisson με μέση τιμή  $\mu = \mathcal{L}/N_b \sigma_{inel} \sim 23$  για  $\sigma_{inel} \approx 70 mb$ . Το out of time pile-up είναι ο επιπλέον αριθμός συγκρούσεων που καταγράφονται ταυτόχρονα όταν ο χρόνος απόκρισης του ανιχνευτή είναι μεγαλύτερος από τον χρόνο μεταξύ δύο διαδοχικών bunches. Στον ανιχνευτή CMS για το έτος 2016 ο μέσος αριθμός αλληλεπιδράσεων ανά bunch crossing φαίνεται στο Σχήμα 1.3

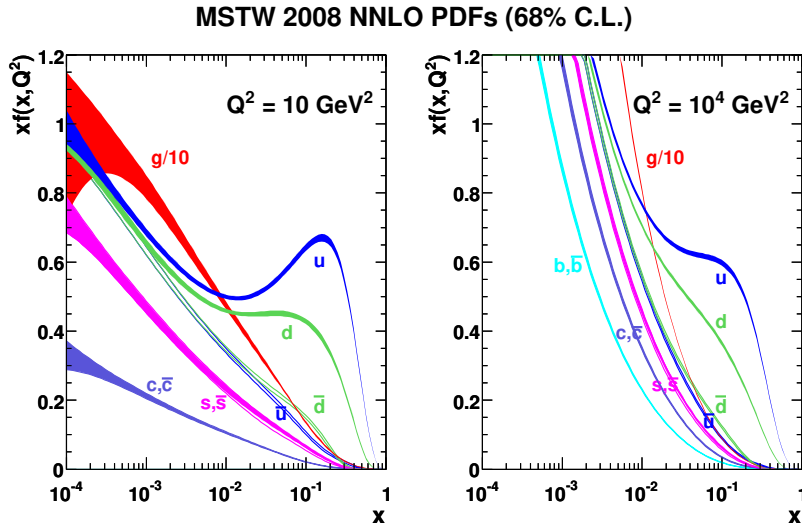


Σχήμα 1.3: Μέσος αριθμός αλληλεπιδράσεων ανά bunch crossing για συγκρούσεις  $pp$  στα  $13 TeV$  στον ανιχνευτή CMS το έτος 2016.

### 1.3.2 Η δομή του πρωτονίου

Μεγάλης σημασίας για τις συγκρούσεις πρωτονίων είναι η δομή του πρωτονίου. Το πρωτόνιο είναι μία περίπλοκη δέσμια κατάσταση τριών κουάρκ  $uud$ , που ονομάζονται κουάρκ

σθένους (valence quarks). Η περιπλοκότητα έγκειται στο γεγονός πως μέσα σε αυτήν την αλληλεπιδρούσα κατάσταση δημιουργούνται γκλουόνια και ζεύγη  $q\bar{q}$  από gluon splittings που λέγονται κουάρκ θάλασσας(sea quarks). Έτσι μέσα στο πρωτόνιο υπάρχει ένας τεράστιος αριθμός από  $g, q, \bar{q}$  που ονομάζονται παρτόνια(partons). Η δομή του πρωτονίου αναφέρεται στον τρόπο που είναι κατανομημένα τα παρτόνια μέσα στα πρωτόνια.



Σχήμα 1.4: Οι κατανομές  $PDFs$  των παρτονίων στο πρωτόνιο σε ακρίβεια NNLO από το group MSTW. Περιλαμβάνονται οι μάντες αβεβαιότητας ενός  $\sigma$  (68%)

Το κύριο χαρακτηριστικό που είναι σχετικό με τις συγκρούσεις πρωτονίων, είναι η κατανομή της τετραορμής του πρωτονίου στα διάφορα παρτόνια. Για παρτόνιο τύπου  $i = g, u, \bar{u}, d, \bar{d}, \dots$ , η πυκνότητα πιθανότητας να μεταφέρει κλάσμα  $x$  της ορμής του πρωτονίου είναι  $f_i(x, \mu_F^2)$ , όπου  $\mu_F$  είναι η κλίμακα παραγοντοποίησης(factorization scale). Η κλίμακα  $\mu_F$  διαχωρίζει το κομμάτι της σκληρής σκέδασης των παρτονίων, που μπορεί να υπολογιστεί με διαταρακτικές μεθόδους, από το κομμάτι χαμηλής ενέργειας όπου παύει να ισχύει η θεωρία διαταραχών και περιγράφεται από τις  $PDFs$ . Κατέχει δηλαδή ανάλογο ρόλο με την  $\mu_R$  της επανακανονικοποίησης, με την διαφορά ότι αναφέρεται σε απειρισμούς χαμηλών ενεργειών infra-red IR, και όχι ultra-violet UV που αφορά την επανακανονικοποίηση. Πολλές φορές τίθεται  $\mu_F = \mu_R = Q$  όπου  $Q$  η χαρακτηριστική κλίμακα της διαδικασίας που μελετάται.

Για τον προσδιορισμό των  $PDFs$  γίνεται προσαρμογή σε δεδομένα, κυρίως από  $ep$  συγκρούσεις, σε χαμηλά  $Q^2$  και έπειτα αυτές εξελίσσονται σε υψηλότερες  $Q^2$  μέσω των εξισώσεων DGLAP οι οποίες προκύπτουν από το γεγονός πως τα παρατηρούμενα μεγέθη δεν πρέπει να εξαρτώνται από την αυθαίρετη κλίμακα  $\mu_F$ . Κάθε ομάδα που ασχολείται με την εξαγωγή των  $PDFs$  χρησιμοποιεί διαφορετικές παραμετροποιήσεις, και δίνει διαφορετικά αποτελέσματα που πρέπει ωστόσο να είναι στατιστικά συμβατές. Στο Σχήμα 1.4 φαίνονται οι  $PDFs$  για τα διάφορα είδη παρτονίων για  $Q^2 = 10 GeV^2$  και  $Q^2 = 10^4 GeV^2$  αντίστοιχα. Όπως παρατηρεί κανείς τα valence quarks κυριαρχούν στα μεγάλα  $x$  ενώ στα χαμηλά τα γκλουόνια, καθώς είναι πιο πιθανή η εμφάνιση soft γκλουονίων. Επιπλέον εφόσον τα γκλουόνια έχουνε μεγα-

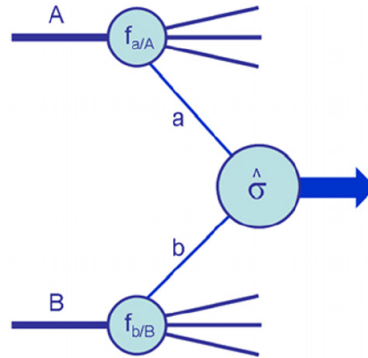
λύτερη πιθανότητα εμφάνισης από τα ζεύγη  $q\bar{q}$  καθώς ο color factor τους είναι μεγαλύτερος. Τα πιο “βαριά” κουάρκ έχουν μικρότερη πιθανότητα εμφάνισης για αυτό οι κατανομές τους διαφέρουν. Στις υψηλότερες ενέργειες οι διαφορές αυτές μικραίνουν και επιπλέον είναι πιο πιθανή η εμφάνιση ζευγών  $q\bar{q}$  και γκλουονίων.

### 1.3.3 Σκληρή σκέδαση παρτονίων

Όταν συγκρούονται ανελαστικά δύο πρωτόνια A και B, με τετραορμές  $P_A$  και  $P_B$  τότε στην πραγματικότητα συγκρούονται δύο παρτόνια των πρωτονίων  $a$  και  $b$ , με τετραορμές  $p_a = x_a P_A$ ,  $p_b = x_b P_B$ . Οι εγκάρσιες συνιστώσες των παρτονίων θεωρούνται αμελητέες. Σύμφωνα λοιπόν με το λεγόμενο factorization theorem η ενεργός διατομή για μία διαδικασία  $p_A p_B \rightarrow X$ , μπορεί να αναλυθεί μέσω της ενεργού διατομής των παρτονίων  $\hat{\sigma}_{ab \rightarrow X}$ <sup>4</sup> και των PDFs ως εξής

$$\sigma_{p_A p_B \rightarrow X} = \sum_{a,b} \int dx_a f_{a,p_A}(x_a, \mu_F^2) \int dx_b f_{b,p_B}(x_b, \mu_F^2) \hat{\sigma}_{ab \rightarrow X}(x_a P_A, x_b P_B, \alpha_S, \mu_F, \mu_R) \quad (1.35)$$

που αν και διαισθητικά μοιάζει προφανής δεν είναι εύκολο να αποδειχθεί γενικά<sup>5</sup>. Το κομμάτι της σκληρής σκέδασης των παρτονίων υπολογίζεται στο πλαίσιο της θεωρίας διαταραχών. Εικονικά αυτή η σχέση αναπαρίστανται στο Σχήμα 1.5.



Σχήμα 1.5: Σκληρή σκέδαση μεταξύ δύο παρτονίων από πρωτόνια.

Η σχέση αυτή είναι πολύ σημαντική και για εκτιμήσεις. Για τα παρτόνια με  $x_a, x_b$  η ενέργεια του κέντρου μάζας δέν είναι εκείνη των πρωτονίων  $S = (13TeV)^2$  αλλά η

$$\hat{s} = x_a x_b S \quad (1.36)$$

έτσι εάν θέλει κανείς ένα top στην τελική κατάσταση πρέπει

$$\hat{s} > m_t^2 \Rightarrow x_a x_b > 1.8 \times 10^{-4} \quad (1.37)$$

οπότε μπορεί να δει κανείς για παράδειγμα τι είναι πιο πιθανό σε αυτήν την κινηματική περιοχή ως προς τα αρχικά παρτόνια.

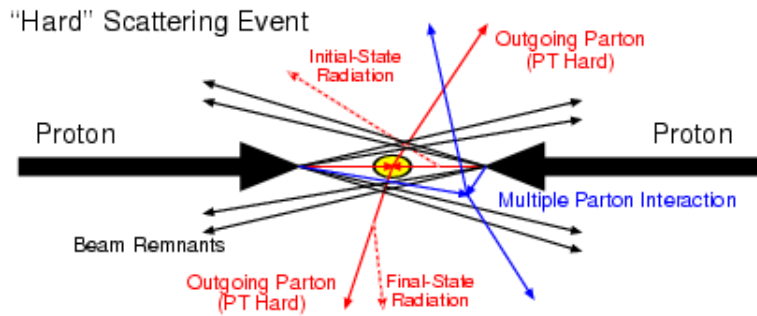
<sup>4</sup>Εν γένει το καπέλο  $\hat{\sigma}$  αναφέρεται σε παρτονικές ποσότητες.

<sup>5</sup>Έχει αποδειχθεί μόνο για DIS και Drell Yan



### 1.3.4 Underlying Event

Κατά την σύγκρουση δύο πρωτονίων πέρα της σκληρής σκέδασης μπορεί να γίνουν επιπλέον συγκρούσεις παρτονίων (multiple parton interactions MPI). Επιπλέον τόσο πριν από την σκληρή σύγκρουση των παρτονίων όσο μετά, μπορεί να εκπέμφθει ακτινοβολία που αναφέρεται ως initial-state και final-state radiation (ISR και FSR). Τέλος υπάρχουν και μέρη των πρωτονίων που δεν αλληλεπιδράσαν που λέγονται beam-beam remnants (BBR). Όλα αυτά τα παρτόνια θα εμφανισθούν στην τελική κατάσταση και πρέπει να ληφθούν υπόψη στις προσομοιώσεις. Τέτοια γεγονότα αναφέρονται με τον όρο underlying event. Στο Σχήμα 1.6 φαίνονται όλες οι συνιστώσες του underlying event που εμφανίζονται σε μία σύγκρουση πρωτονίων.



Σχήμα 1.6: Σχηματική αναπαράσταση ανελαστικής σκέδασης δύο πρωτονίων.

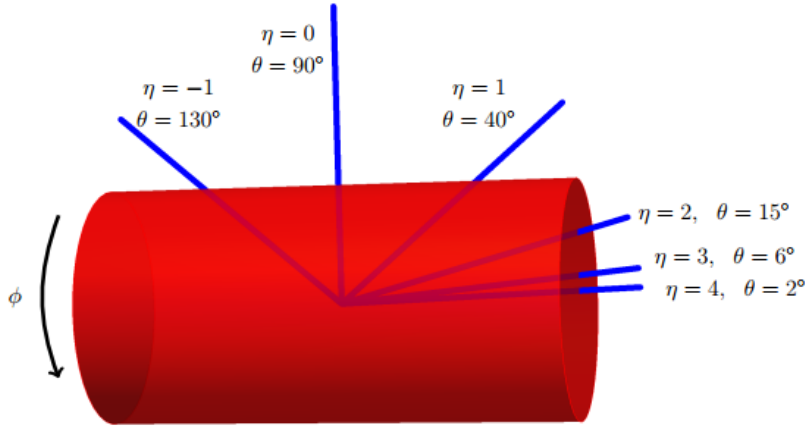
### 1.3.5 Κινηματικές Μεταβλητές

Όπως αναφέρθηκε στην αρχή του υποκεφαλαίου στις συγκρούσεις πρωτονίων του LHC ο άξονας στον οποίο κινούνται οι συγκρουόμενες δέσμες είναι ο  $z$  (beam axis) και το εγκάρσιο επίπεδο στην σύγκρουση είναι το επίπεδο  $xy$  (transverse plane). Έτσι ως προς τα συγκρουόμενα πρωτόνια το σύστημα αναφοράς του εργαστηρίου και εκείνο του κέντρου μάζας τους ταυτίζονται. Ωστόσο, όπως αναφέρθηκε, η πραγματική σύγκρουση γίνεται μεταξύ παρτονίων των πρωτονίων που φέρουν ένα τυχαίο κλάσμα ορμής του πρωτονίου, με αποτέλεσμα για μια παρτονική σκληρή σκέδαση τα προϊόντα έχουνε μία άγνωστη ορμή στον  $z$  άξονα τυχαίου χαρακτήρα. Για αυτόν τον λόγο χρειαζόμαστε ποσότητες που να είναι longitudinal Lorentz boost invariant.

Αρχικά οι ορμές του εγκάρσιου επιπέδου έχουνε την ιδιότητα αυτή και μέσω αυτών ορίζεται η εγκάρσια ορμή και το μέτρο της

$$\vec{p}_T \equiv (p_x, p_y) = (p_T \cos \phi, p_T \sin \phi), \quad p_T \equiv |\vec{p}_T| \quad (1.38)$$

όπου  $\phi$  η αζιμουθιακή γωνία.



Σχήμα 1.7: Γωνιακές μεταβλητές σε συγκρούσεις πρωτονίων.

Μία σημαντική κινηματική ποσότητα είναι η ωχύτητα (rapidity)  $y$  η οποία ορίζεται ως

$$y \equiv \frac{1}{2} \log \frac{E + p_z}{E - p_z} \quad (1.39)$$

η οποία έχει την ιδιότητα πως σε ένα longitudinal boost αλλάζει μόνο κατά μία σταθερά, και έτσι οι διαφορές  $\Delta y$  είναι αναλλιώτες. Η  $y$  σχετίζεται με την γωνία ως προς τον άξονα της σύγκρουσης. Αυτή η συσχέτιση φαίνεται άμεσα στην περίπτωση άμαζων σωματιδίων όπου

$$y = \log \cot \frac{\theta}{2}, \quad m = 0 \quad (1.40)$$

Η rapidity για  $m = 0$  ονομάζεται pseudorapidity  $\eta$ . Στο Σχήμα 1.7 φαίνεται η σχέση με της με την γωνία  $\theta$  σε μερικά παραδείγματα.

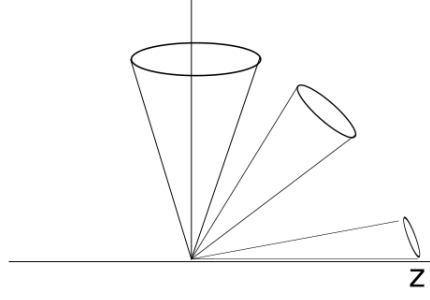
Εάν κανείς γράψει την γωνιακή μετρική για σωματίδια με  $m = 0$  με την βοήθεια της  $\eta$  προκύπτει ότι

$$(d\Omega)^2 = (d\theta)^2 + \sin^2 \theta (d\phi)^2 = \frac{1}{\cosh^2 \eta} [(d\eta)^2 + (d\phi)^2] \quad (1.41)$$

δηλαδή στον χώρο  $\eta - \phi$  είναι μία ευκλείδια μετρική. Έτσι δίνεται το κίνητρο να ορίσει κανείς για τις γωνιακές αποστάσεις έναντι του  $\Delta\Omega$  το

$$\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta \phi)^2} \quad (1.42)$$

Από την σχέση 1.41 φαίνεται πως η σταθερή  $\Delta R$  συνεπάγεται μικρότερο γωνιακό άνοιγμα σε μεγαλύτερες  $\eta$ , όπως φαίνεται στο Σχήμα 1.41.



Σχήμα 1.8: Κώνοι ίδιου  $\Delta R$ . για διαφορετικές τιμές της pseudorapidity.

Το γεγονός αυτό είναι σημαντικό καθώς εφόσον τα φαινόμενα του underlying event και pile-up τείνουν να δώσουν μεγαλύτερη contamination στην forward περιοχή, δηλαδή εκείνη με μεγάλα  $\eta$ , από ότι στην central, με αποτέλεσμα στο επίπεδο  $\eta - \phi$  αυτή να είναι περίπου ομοιόμορφη.

Μέσω των μεταβλητών που αναφέρθηκαν ορίζονται και άλλες χρήσιμες μεταβλητές. Μία από αυτές είναι η ελλειπούσα εγκάρσια ορμή (missing transverse momentum) η οποία ορίζεται ως

$$\vec{p}_T^{\text{miss}} \equiv - \sum_i \vec{p}_{T,i} \quad (1.43)$$

όπου  $\vec{p}_{T,i}$  οι επιμέρους εγκάρσιες ορμές σε ένα γεγονός. Ονομάζεται ελείπουσα καθώς κανονικά οι εγκάρσιες ορμές των τελικών σωματιδίων πρέπει να αθροίζονται στο μηδέν, λόγω διατήρησης της ορμής. Ωστόσο για σωματίδια που δεν ανιχνεύονται όπως τα νετρίνα δεν μπορεί να βρεθεί η ορμή τους και εκτιμάται έμμεσα μέσω αυτού του μεγέθους.

Ένα ακόμα σημαντικό μέγεθος είναι η αναλλοίωτη μάζα. Η αναλλοίωτη μάζα κάποιων σωματιδίων με τετραορμές  $p_i^\mu$  ορίζεται ως

$$m^2 = \left( \sum_i p_i^\mu \right)^2 \quad (1.44)$$

και χρησιμοποιείται σε περιπτώσεις όπου ένα σωματίδιο διασπάται σε κάποια άλλα σωματίδια που ανιχνεύουμε, για παράδειγμα το  $Z \rightarrow l^+l^-$ . Τότε η αναλλοίωτη μάζα ισούνται με την μάζα του σωματιδίου. Για δύο σωματίδια στην τελική κατάσταση

$$m = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2} \quad (1.45)$$

Υπάρχουν ωστόσο περιπτώσεις που ένα σωματίδιο διασπάται σε δύο σωματίδια που το ένα είναι μη ανιχνεύσιμο, όπως γίνεται με τα νετρίνα, για παράδειγμα  $W \rightarrow l\nu_l$ . Τότε δεν μπορεί να υπολογισθεί η αναλλοίωτη μάζα, αλλά υπάρχει ένα αντίστοιχο μέγεθος που είναι η εγκάρσια μάζα (transverse mass) η οποία υπολογίζεται μόνο από τις κάθετες συνιστώσες. Ο ορισμός της είναι

$$m_T = \sqrt{(E_T^1 + E_T^2)^2 - (\vec{p}_T^1 + \vec{p}_T^2)^2} \quad (1.46)$$

Γενικά  $0 \leq m_T \leq m$ , και συγκεκριμένα τα σωματίδια  $p_1, p_2$  βρίσκονται στο εγκάρσιο επίπεδο  $m_T = m$ , ενώ εάν είναι παράλληλα στην δέσμη  $m_T = 0$ .

## 1.4 Το Top Quark

### 1.4.1 Διάσπαση top quark

Το top κουάρκ διασπάται στο Καθιερωμένο Πρότυπο μέσω της ασθενούς αλληλεπίδρασης σε ένα  $W$  μποζόνιο και ένα down-type κουάρκ  $q = d, s, b$ . Το πλάτος διάσπασης έχει την εξής μορφή

$$\Gamma(t \rightarrow W^+q) = \frac{|V_{tq}|^2 m_t^3}{16\pi v^2} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \quad (1.47)$$

Τα ίδια με αντίστροφα πρόσημα ισχύουν για το anti-top. Ωστόσο όπως το στοιχείο έχει μετρηθεί  $|V_{tb}| \approx 0.99$  επομένως το  $BR(t \rightarrow W^+b) = |V_{tb}|^2 / \sum_q |V_{tq}|^2$  είναι σχεδόν μονάδα. Δηλαδή το top με πολύ καλή προσέγγιση μπορεί να θεωρηθεί πως διασπάται πάντα προς παραγωγή bottom.

Από την εξίσωση 1.47 βλέπει κανείς πως το πλάτος διάσπασης του top είναι ανάλογο του λόγου  $m_t^3/16\pi v^2 \sim 1.7\text{GeV}$ , δηλαδή λόγω της μεγάλης μάζας του υπάρχει αρκετά μεγάλο πλάτος διάσπασης. Το πλάτος αυτό είναι πολύ μεγαλύτερο της χαρακτηριστικής κλίμακας της αδρονποίησης  $\Lambda_{QCD} \sim 200\text{MeV}$  και για αυτό το top διασπάται πριν να αδρονποιηθεί. Αυτή η ιδιότητα είναι εξαιρετικά σημαντική καθώς είναι το μόνο κουάρκ που μπορεί να παρατηρηθεί απευθείας από την διάσπαση του.

Μετά την διάσπαση του top το b κουάρκ θα δώσει ένα jet, και το  $W$  ανιχνεύεται μέσα από τα προϊόντα διάσπασης του. Έτσι οι διασπάσεις του top κατατάσσονται με βάση την διάσπαση του  $W$ , οι οποίες χωρίζονται στις αδρονικές

$$W^+ \rightarrow \{u\bar{d}, c\bar{s}\} \quad (1.48)$$

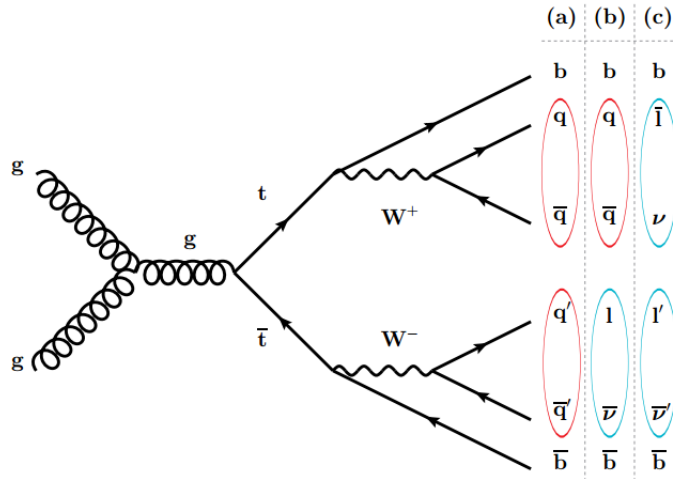
και στις λεπτονικές

$$W^+ \rightarrow \{e^+\nu_e, \mu^+\nu_\mu, \tau^+\nu_\tau\} \quad (1.49)$$

Επειδή τα κουάρκ έχουν και επιπλέον βαθμούς ελευθερίας λόγω του χρώματος είναι πιο πιθανές οι αδρονικές διασπάσεις από τις λεπτονικές. Συγκραχόμενα

$$\frac{BR(W \rightarrow \text{hadronic})}{BR(W \rightarrow \text{leptonic})} = \frac{(ud, cs) \times color}{(e\nu_e, \mu\nu_\mu, \tau\nu_\tau)} = \frac{6}{3} \quad (1.50)$$

Έτσι για διαδικασίες παραγωγής top με ένα  $W$  στην τελική κατάσταση τα 2/3 είναι αδρονικές(hadronic) και το 1/3 έχουν ένα λεπτόνιο(leptonic). Αντίστοιχα για διαδικασίες παραγωγής top με δύο  $W$  στην τελική κατάσταση, όπως  $gg \rightarrow t\bar{t}$  που φαίνεται στο Σχήμα 1.9, υπάρχουν τα ενδεχόμενα η τελική κατάσταση να είναι πλήρως αδρονική(full hadronic) (4/9), ημιλεπτονική δηλαδή να έχει ένα λεπτόνιο και jets (4/9), και διλεπτονική δηλαδή να έχει δύο λεπτόνια (1/9). Έτσι

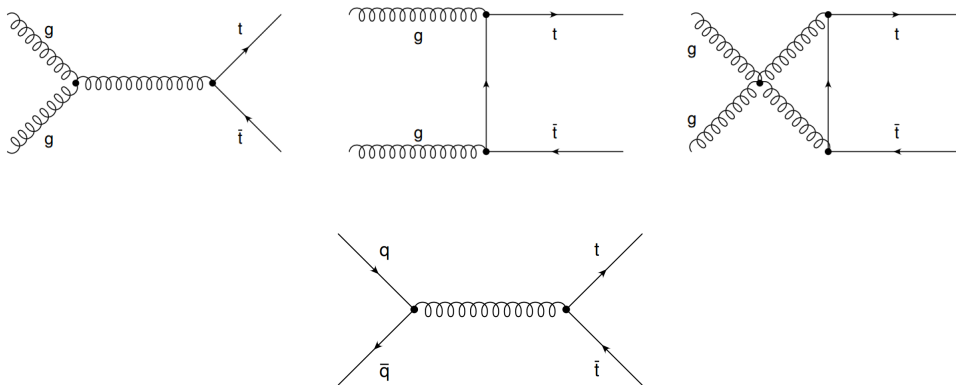


Σχήμα 1.9: Τελικές καταστάσεις για την διαδικασία  $gg \rightarrow t\bar{t}$ .

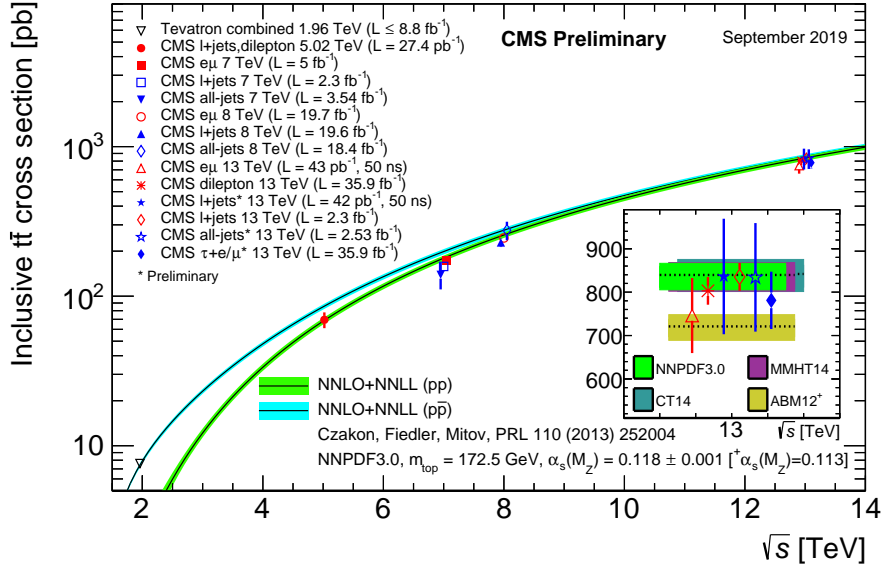
Επιπλέον όσον αφορά τις λεπτονικές διασπάσεις, η διάσπαση σε  $\tau$  δεν μπορεί να ανιχνευθεί άμεσα, αλλά μέσα από τα προϊόντα διάσπασης του  $\tau$  που είναι ένα ζεύγος από jets (64.8%) ή σε ένα ηλεκτρόνιο (17.8%) και μιονίο (17.4%) μαζί με τα αντίστοιχα νετρίνα και ένα  $\nu_\tau$ . Έτσι οι διασπάσεις αυτές συνεισφέρουν τελικά κατά μέρος στο αδρονικό και κατά άλλο μέρος στο λεπτονικό κομμάτι.

### 1.4.2 Παραγωγή top quark

Στις συγκρούσεις πρωτονίων του LHC η παραγωγή των top γίνεται κυρίως μέσα από τις ισχυρές αλληλεπιδράσεις των παρτονίων. Η παραγωγή αυτή οδηγεί σε ζεύγη  $t\bar{t}$ , καθώς παράγονται μέσω gluon splittings, και γίνεται μέσω gluon fusion και αλληλεπίδραση  $q\bar{q}$  προς γκλουόνιο. Τα αντίστοιχα διαγράμματα σε tree level φαίνονται στο Σχήμα 1.10. Επειδή όπως αναφέρθηκε και στο κομμάτι που συζητήθηκαν οι PDFs τα γκλουόνια βρίσκονται σε μεγαλύτερη αφθονία μέσα στα πρωτόνια, τα τρία διαγράμματα της gluon fusion αποτελούν το 90% της ενεργού διατομής παραγωγής  $t\bar{t}$ .



Σχήμα 1.10: Διαγράμματα Feynman για παραγωγή  $t\bar{t}$  στον LHC σε tree level.



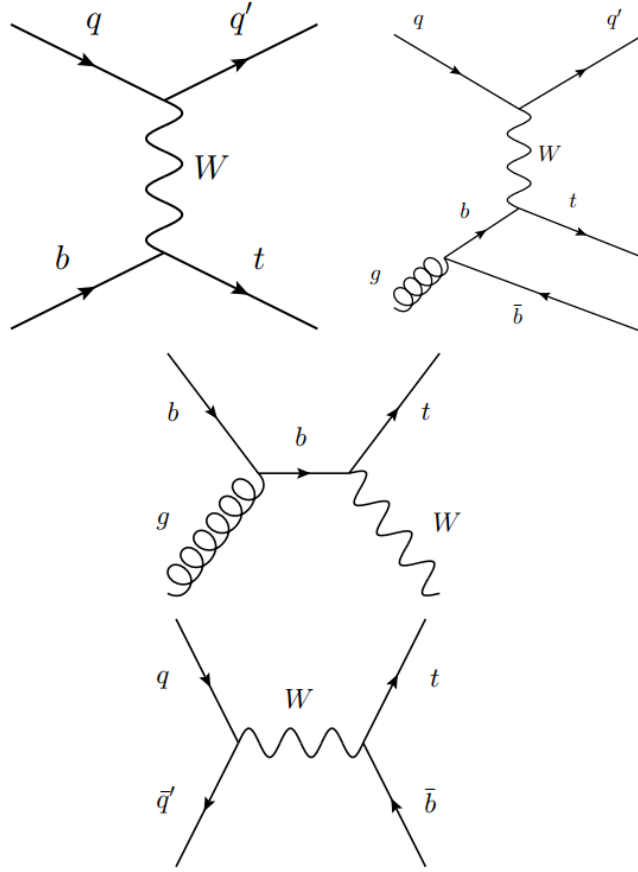
Σχήμα 1.11: Μετρήσεις ενεργού διατομής παραγωγής  $t\bar{t}$  στο πείραμα CMS σε σύγκριση με θεωρητικούς υπολογισμούς σε ακρίβεια NNLO+NNLL.

Για διαφορετικές ενέργειες οι μετρήσεις της ενεργού διατομής  $t\bar{t}$  στο πείραμα CMS φαίνονται στο Σχήμα 1.11. Στα  $\sqrt{s} = 13\text{TeV}$  που θα γίνει η ανάλυση, και για μάζα top  $m_t = 172.5\text{GeV}$  η τιμή της είναι

$$\sigma_{t\bar{t}}(\sqrt{s} = 13\text{TeV}, m_t = 172.5\text{GeV}) = 832.8^{+19.8+35.1}_{-29.2-35.1}\text{pb} \quad (1.51)$$

δηλαδή είναι μετρημένη με ακρίβεια περίπου 7%.

Πέραν της παραγωγής του top μέσω ισχυρών αλληλεπιδράσεων, υπάρχουν οι διαδικασίες μέσω ηλεκτρασθενών. Στις διαδικασίες αυτές παράγεται ένα top στην τελική κατάσταση, για αυτό αναφέρονται ως “single top”. Η παραγωγή γίνεται μέσω της κορυφής  $Wtb$ , δηλαδή την ίδια με την οποία γίνεται η διάσπαση. Αυτό είναι σημαντικό καθώς αυτές οι διαδικασίες προσφέρονται για δυνατότητα μελέτης της δομής της κορυφής αυτής. Τα διαγράμματα Feynman σε tree level των διαδικασιών “single top” βρίσκονται στο Σχήμα 1.12.



Σχήμα 1.12: Διαγράμματα Feynman για παραγωγή “single top” στον LHC σε tree level.

Ο κύριος τρόπος παραγωγής γεγονότων single top είναι μέσω ενός κουάρκ  $q$  και ενός  $b$  κουάρκ, που είναι μία διαδικασία t-channel. Το  $b$  μπορεί είτε να έρθει απευθείας από το πρωτόνιο, είτε έμμεσα μέσω ενός αρχικού γκλουονίου. Στην δεύτερη περίπτωση θα υπάρχουνε δύο jets από  $b$  κουάρκ στην τελική κατάσταση. Το αρχικό  $q$  θα πρέπει να είναι up-type κουάρκ για παραγωγή top και down-type κουάρκ για παραγωγή anti-top. Για το πρωτόνιο ωστόσο υπάρχουνε πιο πολλά  $u$  από ότι  $d$ , πράγμα που οδηγεί στην ασύμμετρη παραγωγή top και anti-top μέσω της διαδικασίας αυτής. Έτσι οι ενεργές διατομές που προκύπτουν για  $\sqrt{s} = 13\text{TeV}$  είναι

$$\sigma_{t\text{-chan.}}^{\text{top}}(\sqrt{s} = 13\text{TeV}) = 130 \pm 1(\text{stat.}) \pm 19(\text{syst.}) \quad (1.52)$$

και

$$\sigma_{t\text{-chan.}}^{\text{anti-top}}(\sqrt{s} = 13\text{TeV}) = 77 \pm 1(\text{stat.}) \pm 12(\text{syst.}) \quad (1.53)$$

ενώ η συνολική ενεργός διατομή είναι

$$\sigma_{t\text{-chan.}}^{t+\bar{t}}(\sqrt{s} = 13\text{TeV}) = 207 \pm 2(\text{stat.}) \pm 31(\text{syst.}) \quad (1.54)$$

όπου φαίνεται πως η κύρια πηγή αβεβαιότητας είναι τα συστηματικά σφάλματα.

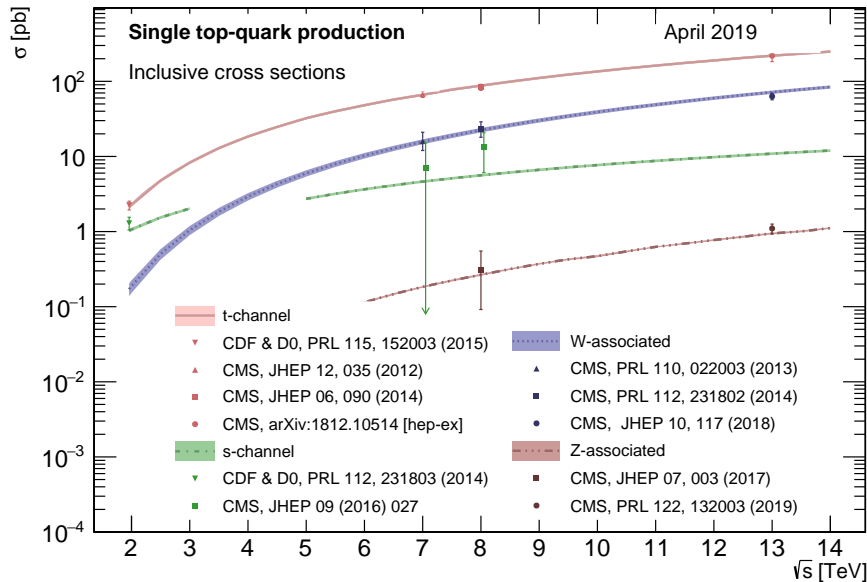
Η δεύτερη κύρια συνεισφορά σε single top γεγονότα είναι η λεγόμενη “W associated production”. Αυτή προκύπτει από την αλληλεπίδραση ενός  $b$  κουάρκ και ενός γκλουονίου.

Το  $b$  κουάρκ που παράγεται έπειτα διασπάται σε  $top$  και  $W$ . Είναι η μοναδική διαδικασία  $single\ top$  με δύο  $W$  μποζόνια στην τελική κατάσταση. Η ενεργός διατομή είναι αρκετά μικρότερη της  $t$ -channel διαδικασίας αλλά όχι αμεληταία, και έχει μετρηθεί ως

$$\sigma_{Wt} = 94 \pm 10(stat.)_{-22}^{+28}(syst.) \pm 2(lumi)pb \quad (1.55)$$

Τέλος, η πιο μικρή συνεισφορά για παραγωγή  $single\ top$ , είναι η  $s$ -channel διαδικασία μέσω  $q$  και  $\bar{q}$  προς  $W$ , το οποίο έπειτα διασπάται σε  $t\bar{b}$ . Κινηματικά είναι αρκετά δύσκολη η παραγωγή τέτοιων γεγονότων για αυτό η ενεργός διατομή είναι πολύ μικρότερη από τις διαδικασίες που προαναφέρθηκαν. Η τιμή της έχει μετρηθεί δεν έχει μετρηθεί ακόμα στα  $13\text{TeV}$  αλλά αναμένεται περίπου στα  $10pb$ . Η διαδικασία αυτή πρέπει να έχει την ασύμμετρη παραγωγή  $t$  και  $\bar{t}$  για τον ίδιο λόγο με το  $t$ -channel. Επιπλέον για  $single\ top$  υπάρχει η διαδικασία “ $Z$  associated” η οποία έχει πολύ μικρή ενεργό διατομή, η οποία θα θεωρηθεί αμεληταία.

Συνολικά οι μετρήσεις στο CMS για τις ενεργούς διατομές  $single\ top$  διαδικασιών σε σύγκριση με τις θεωρητικές καμπύλες βρίσκονται στο Σχήμα 1.13.



Σχήμα 1.13: Μετρήσεις ενεργού διατομής παραγωγής  $single\ top$  στο πείραμα CMS σε σύγκριση με θεωρητικούς υπολογισμούς.



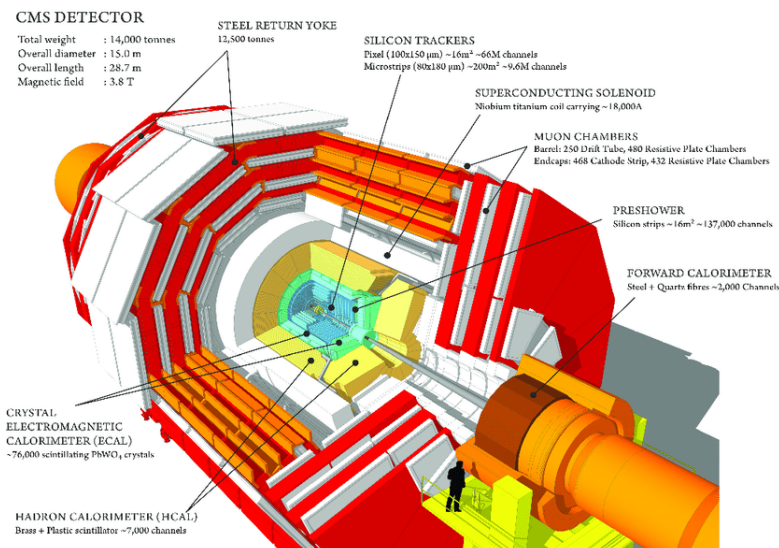


## Κεφάλαιο 2

# Ο ανιχνευτής Compact Muon Solenoid (CMS)

### 2.1 Εισαγωγή

Ο ανιχνευτής CMS είναι ένα σύμπλεγμα από επιμέρους ανιχνευτικά συστήματα ώστε να μπορεί να μετρηθούν με ακρίβεια οι τροχιές των σωματιδίων (tracking), οι ενέργειες τους (calorimetry) και να γίνεται η ταυτοποίησή τους (identification). Για τον λόγο αυτό οι ανιχνευτές τέτοιου τύπου λέγονται multi-purpose detectors. Το όνομα του CMS προέρχεται από το μικρό μέγεθος (διάμετρος 15m) του συγκριτικά με τον στόχο που επιτυγχάνει (Compact), το ανεπτυγμένο σύστημα μιονίων που του επιτρέπει να έχει καλή απόδοση στην ταυτοποίηση των μιονίων και την ακρίβεια της ορμής τους (Muon) και τον σωληνοειδή υπεραγωγίμο μαγνήτη (Solenoid).



Σχήμα 2.1: Σχηματική αναπαράσταση των στοιχείων του ανιχνευτή CMS.

Η εικόνα του ανιχνευτή φαίνεται στο Σχήμα 2.1. Όπως φαίνεται ο ανιχνευτής αποτελείται

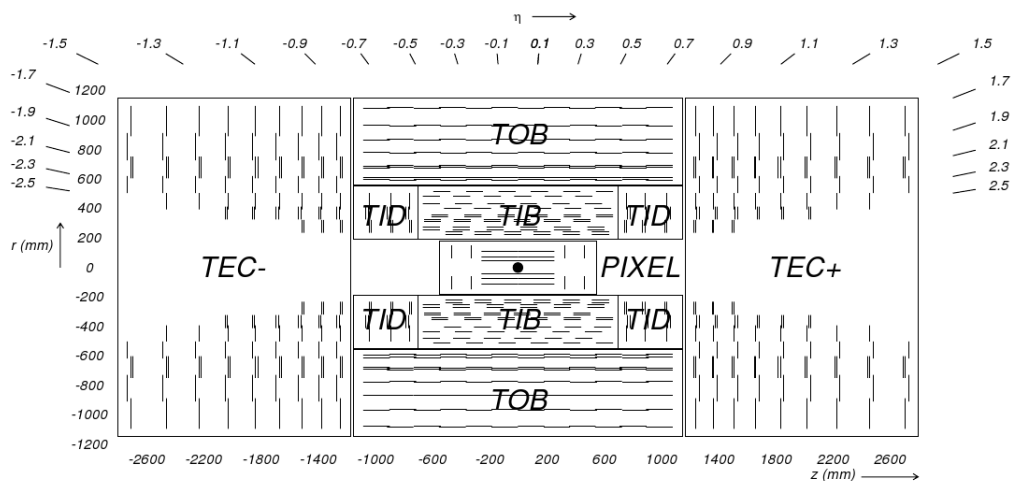
από στρώσεις υποανιχνευτών οι οποίοι βρίσκονται ο ένας μετά τον άλλον σε έναν κύλινδρο, παράλληλα στον άξονα που συγκρούονται οι δέσμες, που λέγεται barrel και κλείνει στο άνω και κάτω μέρος και δύο κομμάτια που λέγονται endcaps. Το κέντρο του συστήματος συντεταγμένων του ανιχνευτή είναι το σημείο της σύγκρουσης των δεσμών πρωτονίων. Το εγκάρσιο επίπεδο ορίζεται έτσι ώστε ο άξονας  $x$  να εκτείνεται ακτινικά προς το κέντρο του LHC.

Στις παρακάτω παραγράφους θα περιγραφούν τα επιμέρους στοιχεία του ανιχνευτή, ξεκινώντας από το εσωτερικό και πηγαίνοντας σταδιακά προς τα έξω. Δηλαδή η σειρά είναι η ακόλουθη

1. Ανιχνευτής τροχιών (Tracker)
2. Ηλεκτρομαγνητικό θερμιδόμετρο (electromagnetic calorimeter-ECAL)
3. Αδρονικό θερμιδόμετρο (hadronic calorimeter-HCAL)
4. Σύστημα μιονίων (Muon System)

## 2.2 Ανιχνευτής τροχιών (Tracker)

Ο ανιχνευτής τροχιών έχει ως στόχο την ανακατασκευή της τροχιάς των φορτισμένων σωματιδίων που παράγονται από τις συγκρούσεις πρωτονίων. Ο ισχυρός κυλινδρικός μαγνήτης του ανιχνευτή (3.8 Tesla) καμπυλώνει τις τροχιές των σωματιδίων αυτών. Μέσω των τροχιών μπορεί να εκτιμηθεί η ορμή τους με μεγάλη ακρίβεια και να γίνει έτσι η ανακατασκευή διαφόρων κορυφών αλληλεπίδρασης από τροχιές προϊόντων διασπάσεων.



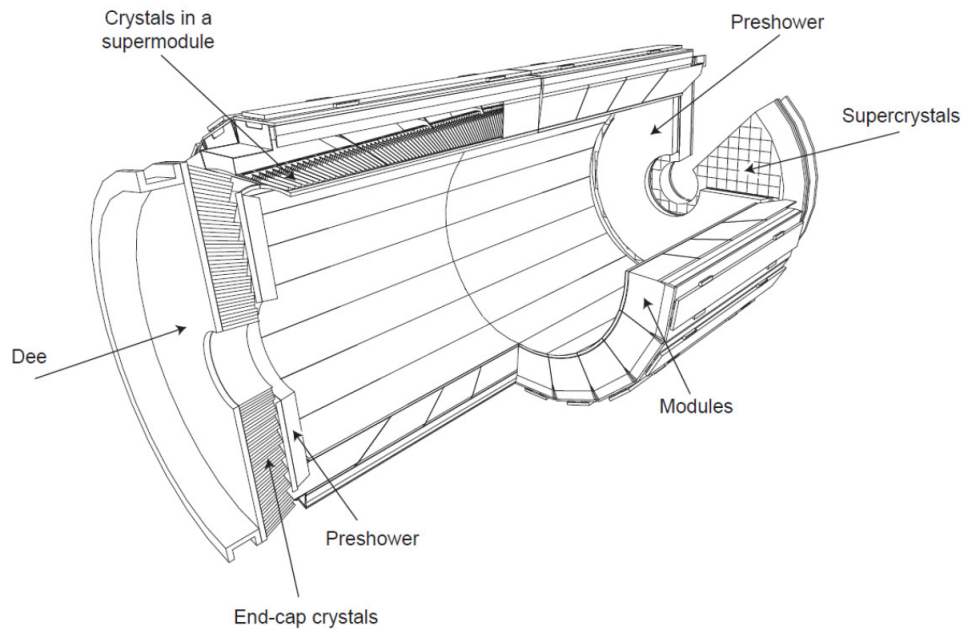
Σχήμα 2.2: Σχηματική αναπαράσταση του ανιχνευτή τροχιών.

Η γεωμετρία του ανιχνευτή αυτού απεικονίζεται στο Σχήμα 2.2. Ο ανιχνευτής έχει μήκος 5.2 μ και καλύπτει περιοχή  $r < 1.2m$  και  $|\eta| < 2.5$ . Λόγω της πολύ μεγάλης φωτεινότητας του LHC ο ανιχνευτής πρέπει να έχει πολύ καλή διακριτική ικανότητα και αντοχή στην ακτινοβολία. Επιπλέον πρέπει να μπορεί να επιτευχθεί αυτή η καλή διακριτική ικανότητα έχοντας όσο το

δυνατόν λιγότερο υλικό ώστε ο ανιχνευτής αυτός να μην δημιουργεί σημαντικές απώλειες στα σωματίδια που τον διαπερνούν. Ο ανιχνευτής αποτελείται από μεγάλους strip ανιχνευτές πυριτίου που εσωτερικά έχουν ένα μικρό ανιχνευτή πυριτίου με pixel μεγάλης ακρίβειας ( $\sim 10\mu\text{m}$ ) που είναι κοντά στο σημείο της σύγκρουσης. Και στις δύο περιπτώσεις υπάρχουν στρώσεις παράλληλα στο barrel αλλά και στα endcaps.

## 2.3 Ηλεκτρομαγνητικό Θερμιδόμετρο (ECAL)

Το ηλεκτρομαγνητικό θερμιδόμετρο βρίσκεται στο εσωτερικό του υπεραγωγίου μαγνήτη του CMS. Αποτελείται από κρυστάλλους σπινθηρισμού  $\text{PbWO}_4$  γρήγορης απόκρισης, σε διάταξη υψηλής διακριτικής ικανότητας. Το υλικό αυτό επιλέχθηκε γιατί τα σωματίδια διανύουν μικρό μήκος μέχρι να εναποθέσουν την ενέργεια του στον κρύσταλλο (μήκος ακτινοβολίας  $X_0 = 0.89\text{cm}$ ) και ο κατεγισμός που δημιουργείται είναι μικρού εύρους (ακτίνα Moliere  $R_M = 2.5\text{cm}$ ).



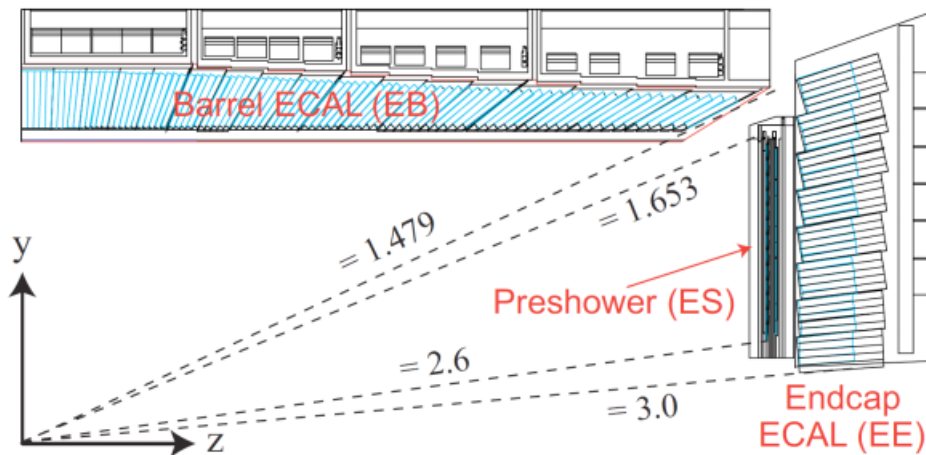
Σχήμα 2.3: Σχηματική αναπαράσταση του ηλεκτρομαγνητικού θερμιδόμετρου του CMS.

Τα επιμέρους κομμάτια του θερμιδόμετρου είναι

- **ECAL Barrel (EB)**: Στο κομμάτι αυτό οι κρύσταλλοι είναι ελαφρώς γυρισμένοι σε διάταξη ώστε να δείχνουν προς σημείο σύγκρουσης. Αποτελείται από 61200 κρυστάλλους σχήματος κομμένης πυραμίδας με  $23\text{cm}$  μήκος και  $2.2 \times 2.2\text{cm}^2$  εμβαδό πλευράς, που ορίζει την διακριτική του ικανότητα. Το γωνιακό εύρος που καλύπτει αυτό το κομμάτι είναι  $|\eta| < 1.48$ .
- **ECAL Endcaps (EE)**: Σε κάθε endcap υπάρχουν αντίστοιχοι με το EB κομμάτι κρύσταλλοι με διαφορετικές διαστάσεις (μήκος  $22\text{cm}$  και εμβαδό  $2.86 \times 2.86\text{cm}^2$ ). Ο

αριθμός τους είναι 3662 στο κάθε endcap και εκτείνουν την κάλυψη του θερμιδόμετρου μέχρι  $|\eta| < 3$ .

- **ECAL Preshower (ES)**: Το κομμάτι αυτό βρίσκεται μπροστά από τα endcap κομμάτια και καλύπτει μία περιοχή  $1.65 < |\eta| < 2.6$ . Αποτελείται από απορροφητές μολύβδου εξοπλισμένους με silicon strip ανιχνευτές, που έχουν ακρίβεια εμβαδού  $1.90 \times 61 \text{mm}^2$ . Το μέρος αυτό έχει μικρό πάχος που αντιστοιχεί σε  $\sim 3X_0$  και ο ρόλος του είναι για να μπορεί να διαχωριστεί η διαδικασία  $\pi^0 \rightarrow \gamma\gamma$ , λόγω της πολύ μικρής γωνίας μεταξύ των δύο φωτονίων σε μεγάλα  $\eta$ .



Σχήμα 2.4: Πλάγια όψη του ηλεκτρομαγνητικού θερμιδόμετρου.

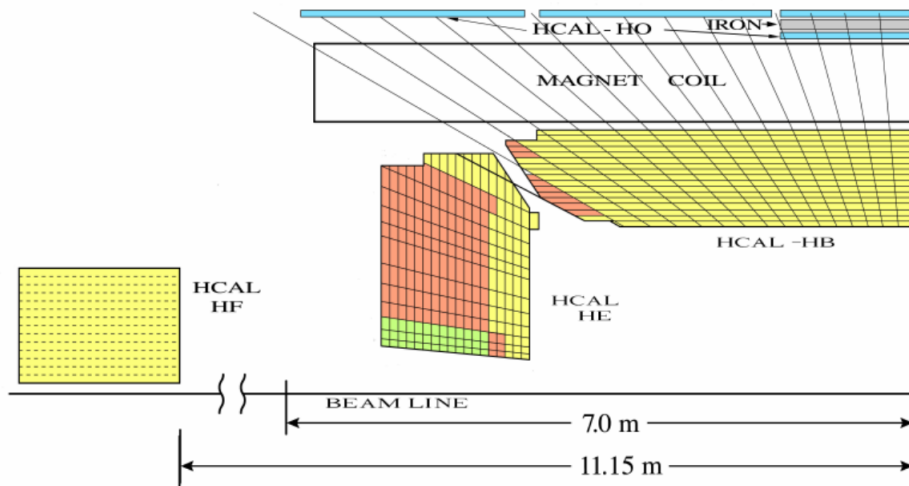
## 2.4 Αδρονικό Θερμιδόμετρο (ECAL)

Το αδρονικό θερμιδόμετρο (HCAL) μετράει την ενέργεια αδρονίων που φτάνουν στον ανιχνευτή. Χρειάζεται για να μπορούν να ανακατασκευασθούν τα jets. Αποτελείται από σπινθηριστές παρεμβολόμενους με απορροφητές.

Έχει τέσσερα κύρια μέρη όπως φαίνεται στο Σχήμα 2.5, τα

- **ECAL Barrel (HB)**
- **ECAL Endcap (HE)**
- **ECAL Outer (HO)**
- **ECAL Forward (HF)**

Όπως και στο ηλεκτρομαγνητικό καλορίμετρο τα κομμάτια barrel και endcap καλύπτουν  $|\eta| < 3.0$ . Αυτά α κομμάτια καλύπτουν το ηλεκτρομαγνητικό καλορίμετρο και το κομμάτι του barrel φτάνει μέχρι τον μαγνήτη. Το HB αποτελείται από 36 αζιμουθιακές σφήνες που σχηματίζουν δύο half-barrels. Κάθε σφήνα αποτελείται από επίπεδους χάλκινους απορροφητές



Σχήμα 2.5: Πλάγια όψη του αδρονικού θερμοδόμετρου.

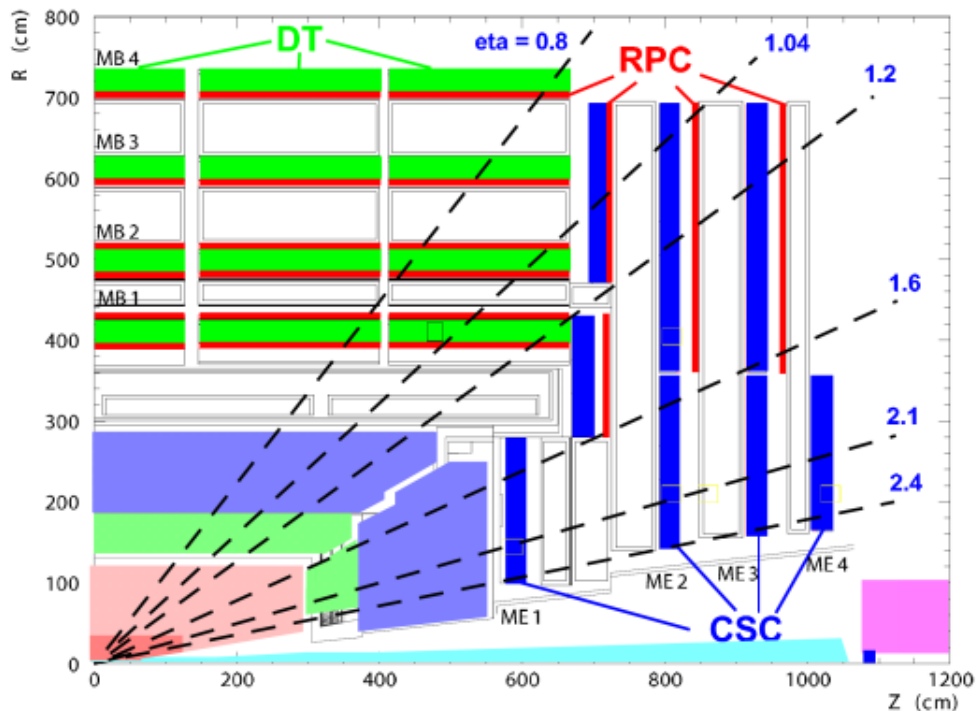
παράλληλους στον άξονα  $z$ . Ο πρώτος και ο τελευταίος απορροφητής είναι από ατσάλι. Συνολικά 17 πλαστικοί σπινθηριστές, εισέρχονται μεταξύ των πλακών. Παρόμοια πράγματα ισχύουν για τα HE κομμάτια. Οι σπινθηριστές έχουν διακριτική ικανότητα στον χώρο  $\eta - \phi$ , περίπου  $\Delta\eta \times \Delta\Phi = 0.087 \times 0.087$  για  $|\eta| < 1.6$  και  $0.17 \times 0.17$  για  $|\eta| > 1.6$ . Επειδή τα θερμοδόμετρα δεν καταγράφουν πάντα όλη την ενέργεια των αδρονικών καταγισμών υπάρχουν μετά τον μαγνήτη τα HO κομμάτια. Έτσι αυξάνεται ενεργά το πάχος του αδρονικού καλοριμέτρου. Επιπλέον για να υπάρχει μεγαλύτερη γωνιακή κάλυψη το θερμοδόμετρο συμπληρώνεται μέσω του HF. Αυτό το κομμάτι δέχεται πολύ μεγάλη ακτινοβολία επειδή βρίσκεται σε μικρή γωνία ως προς τον άξονα της δέσμης και για αυτό χρησιμοποιούνται ανθεκτικές οπτικές ίνες, που διανύουν παράλληλα μέσα στον απορροφητή και συλλέγουν την ακτινοβολία Cherenkov. Η κάλυψη που προσφέρει είναι  $2.9 < |\eta| < 5.2$ . Παίζει σημαντικό ρόλο τόσο στην ανακατασκευή forward jets αλλά και σε καλύτερες μετρήσεις της  $E_T^{miss}$ .

## 2.5 Σύστημα Μιονίων (Muon System)

Το σύστημα μιονίων του CMS αποτελείται από τέσσερις σταθμούς κάθε ένας από τους οποίους αποτελείται από ανιχνευτές ιονισμού αερίου, έξω από το σωληνοειδές που βρίσκονται μεταξύ επιφανειών από ατσάλι. Ο λόγος ύπαρξης του συστήματος αυτού είναι η καλύτερη ανακατασκευή των μιονίων, τα οποία είναι τα πιο διεισδυτικά και μπορούν να περάσουν από όλα τα υπόλοιπα συστήματα του ανιχνευτή.

Το σύστημα μιονίων συνολικά έχει κάλυψη  $|\eta| < 2.4$  και αποτελείται από συνδυασμό ανιχνευτών τριών διαφορετικών ειδών

- Θάλλαμοι **Drift Tube (DT)** στην περιοχή του barrel με κάλυψη  $|\eta| < 2.1$ . Το σύστημα των DTs αποτελείται από 5 κυλίνδρους ομόκεντρους ομοαξονικούς με τον ανιχνευτή. Κάθε δίσκος έχει 4 σταθμούς διαφορετικής ακτίνας. Κάθε σωλήνας έχει διάμετρο 4



Σχήμα 2.6: Πλάγια όψη του συστήματος μιονίων.

εκατοστά και στο κέντρο του υπάρχει ένα σύρμα και μείγμα αερίων αργού και διοξειδίου του άνθρακα. Όταν ένα μιονιο περάσει από μέσα από το αέριο ιονίζει τα μόρια, και τα ηλεκτρόνια που δημιουργούνται κατευθύνονται στο σύρμα λόγω ηλεκτρικού πεδίου που υπάρχει στο εσωτερικό του σωλήνα. Ανιχνεύοντας το σημείο που τα ηλεκτρόνια έπεσαν πάνω στο σύρμα και τον χρόνο ολίσθησης υπολογίζεται η τροχιά του μιονίου.

- **Cathode Strip Chamber (CSC)** για να διαχειρίζονται τους πιο υψηλούς ρυθμούς σωματιδίων λόγω του μικρού χρόνου απόκρισης, και την ανομοιογένεια του μαγνητικού πεδίου στην περιοχή  $0.9 < |\eta| < 2.4$ . Αυτοί διαθέτουν θετικά φορτισμένα σύρματα (άνοδοι) και αρνητικά φορτισμένα (κάθοδοι) κάθετα τοποθετημένα. Το πλέγμα των συρμάτων βρίσκεται μέσα σε μείγμα αερίου. Όταν ένα μιονιο περάσει ιονίζει τα μόρια του αερίου και τα ηλεκτρόνια που παράγονται οδηγούνται προς τις καθόδους δημιουργώντας χιονοστιβάδα. Το σήμα που παράγεται από αυτήν δίνει πληροφορία για την τροχιά του σωματιδίου που πέρασε.
- **Resistive Plate Chambers (RPC)** που βρίσκονται και στο barrel και στο endcap και καλύπτουν εύρος  $|\eta| < 2.1$ . Αυτοί λειτουργούν συμπληρωματικά ως προς τους προηγούμενους, παρέχοντας πληροφορία για σκανδαλισμό. Αποτελούνται από δύο πλάκες, μια θετικά φορτισμένη που λειτουργεί σαν άνοδος και μια αρνητικά φορτισμένη που έχει το ρόλο της καθόδου και στο ενδιάμεσο τους υπάρχει αέριο. Όταν ένα μιονιο περάσει από το αέριο προκαλεί ιονισμό. Λόγω των φορτισμένων πλακών στην περιοχή του αερίου υπάρχει ηλεκτρικό πεδίο. Έτσι τα ηλεκτρόνια που παράγονται από τον ιονισμό επιταχύνονται και δημιουργούν χιονοστιβάδα. Τα προϊόντα της χιονοστιβάδας

οδηγούνται προς την άνοδο όπου και συλλέγονται δίνοντας έτσι πληροφορία για το μιονίο που πέρασε. Το ηλεκτρικό πεδίο είναι τέτοιο ώστε το σήμα να έρχεται σε πολύ μικρότερο χρόνο από την συχνότητα των συγκρούσεων.





## Κεφάλαιο 3

# Analysis

### 3.1 Introduction

This chapter is the main part of this thesis, which is the analysis of data from the CMS detector in order to estimate a scale factor for the cross section of single top quark production in the t-channel. First the samples used and the data set from the detector events are given. After that a synopsis of the analysis strategy, that is to be performed, is mentioned. As we will see in this part the core element for our strategy estimation of the cross section is a fit performed on data to extract the corresponding scale factor. In order to do that, one needs good modeling of the theoretical distributions for the different processes involved, and an appropriate variable used for the fit. The first part is related to the Monte Carlo samples, and the extraction of a template for the QCD Multijet background from data, since this process cannot be described well from the Monte Carlo approach. Later on, the second part of the choice of a fitting variable is studied. As we will see, the requirements for this choice are the discrimination power (Signal Vs Background) that is offered from the shape of a variable, and most importantly the behavior of a maximum likelihood fit on it. This whole procedure is performed using just simulated samples, in order to select an estimation strategy in an unbiased and validated manner. After the selection of the fit procedure is made, the real data are being used to make the measurement of the signal cross section. The chapter ends with some conclusions and remarks on the procedure and the results.

### 3.2 Physical Processes

In this section they will be mentioned the signal and backgrounds characteristics that define the events signatures in the detector, and therefore are important for the analysis.

#### Single Top t-channel

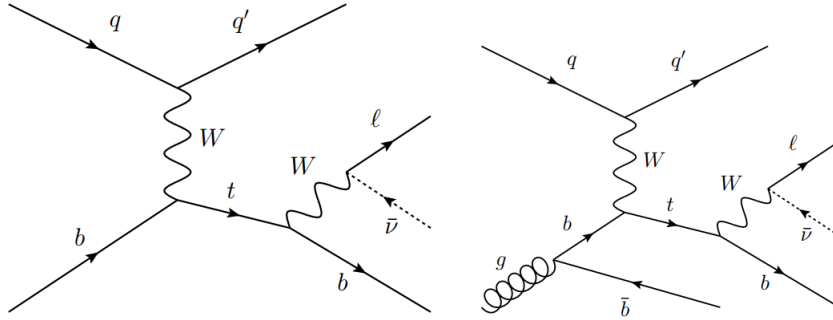


Figure 3.1: Feynman diagrams for t-channel that show the expected final state.

The single top t-channel is the signal process, and therefore it defines the wanted final state. In this analysis the case of the leptonic  $W$  boson, coming from the (anti)top quark decay is considered. As it is seen from the leading tree level Feynman diagrams of the process in Figure 3.1, the final state consists of one lepton, one light flavored quark, and one or two b-tagged jets, depending on the case of production through a b-quark originating from proton or gluon splitting. In the second case the b quark originating from the top decay will yield a central jet, while the b originating from gluon will be forward and with low  $p_T$ . This means that this extra b-jet will not be able to be reconstructed and pass the needed requirements. For this reason as final state it is considered to have one b-tagged jet, and of course one lepton and a light-jet. Also from the kinematics the light-jet will be forward which is a characteristic discriminating property of the ST t-channel, from the backgrounds. An other property is related with the fact that in t-channel production the top is highly polarized, and this is transferred to its decay products, yielding specific spin configuration for leptons which is translated into angular kinematic distributions.

### Other Single Top processes

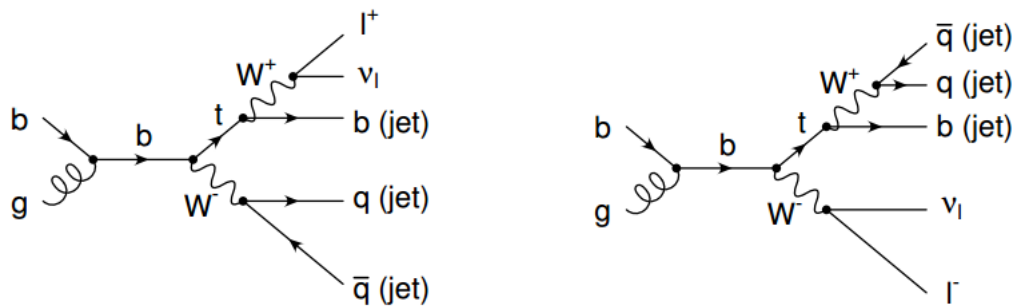


Figure 3.2: Examples of Feynman diagrams for the  $Wt$  process.

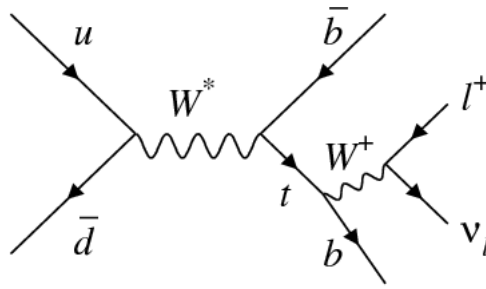
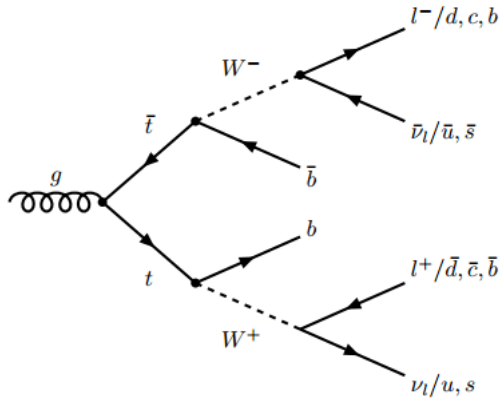


Figure 3.3: Feynman diagram for the s-channel process.

There are other two processes for a single-top final state production. The first is the  $Wt$  production, where the top quark is produced in association with a  $W$  boson. Therefore there are two  $W$  bosons, in the final state for this process. For the one lepton requirement the one of them should decay leptonically while the other hadronically. If the lepton comes from the  $W$  of top decay then the top can be reconstructed correctly, otherwise not. An other feature is that these events prefer three jets in the final state with one b-tagged, while it will be required to have two, and are therefore suppressed. The second single-top process is the s-channel production, where in the final state there are two b-tagged jets, one coming from top decay. This background has small cross section and the selection of one b-tagged jet can further reduce it, especially big b-tagging efficiency.

#### top anti-top pair production

Figure 3.4: Feynman diagram for the  $t\bar{t}$  possible final states.

The  $t\bar{t}$  production is the main process of producing top quarks in the LHC. The top quarks are produced via strong interaction from gluon splitting, therefore they are top and anti-top pair. The final state of interest is this where one of the  $W$  bosons, produced from top quarks, decays leptonically and the other one hadronically. This background has big cross section  $\sim 832pb$  in total, but it is suppressed from low jet multiplicity cuts, and

the one b-tag jet requirement especially if the b-tagging efficiency is high<sup>1</sup>. Also although there are real top quark decays in these events the lepton may be reconstructed with the b-jet of the other top leading to more wide distribution in top mass.

### W+Jets

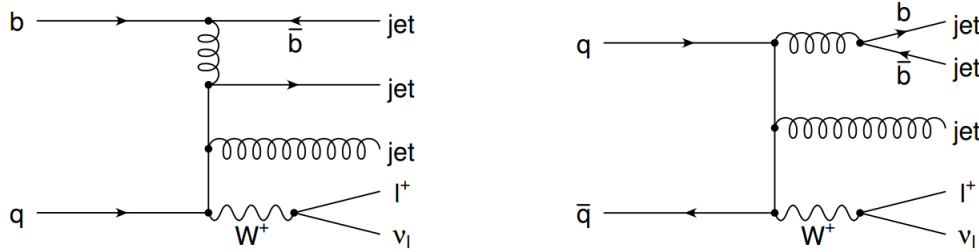


Figure 3.5: Examples of Feynman diagrams for  $W + jets$  with b-quarks in final state ( $Wb, Wb\bar{b}$ ).

The process  $W + jets$  refers to events in which a W boson is produced with additional jets in the final state. For the wanted final state the case of W decaying to one lepton and the corresponding neutrino is considered. The total cross section of these events is of order of  $\sim 10nb$  per lepton flavor. The main cases of interest for such events are these which contain one or two b-jets in the final state. Although the selection of the signal will be focussing on having one b-tagged jet, in the second case the one b-jet may not be reconstructed. These events can be discriminated from the signal because of the fact that the b-jet and the lepton did not originate from a top decay, leading to different kinematic characteristics and also not giving the correct invariant mass.

### Z+Jets

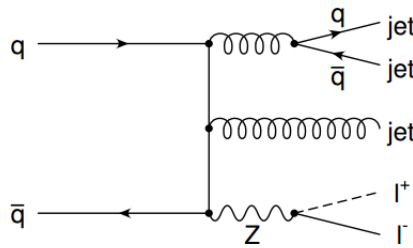


Figure 3.6: Examples of a Feynman diagram for  $Z + jets$ .

A similar process to  $W + jets$  is the  $Z + jets$  where instead of a W boson there is a Z. The total cross sections for such events is a bit lower than for W and is of order of  $\sim 1nb$  per lepton flavor. Although these events contribute much less, because these events contain two leptons from the  $Z \rightarrow l^+l^-$  decay, and it is more unlikely to miss the one lepton from the detection. That is the reason that these events are also called Drell-Yan.

### Diboson

<sup>1</sup>The expected yield has a factor depending on b-tagging efficiency  $\epsilon_b$  which is  $\epsilon_b(1 - \epsilon_b)$

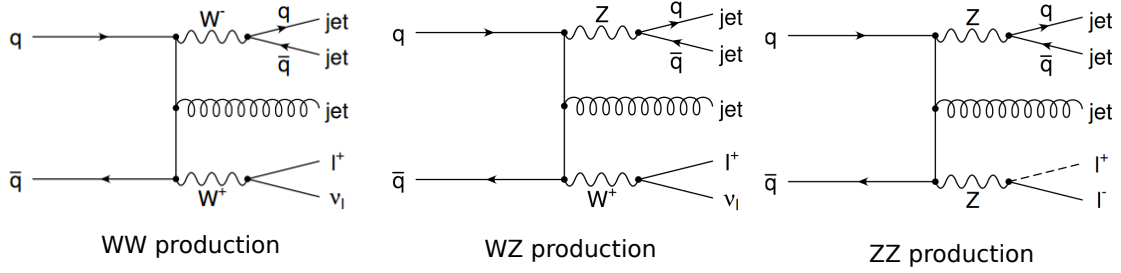


Figure 3.7: Examples of Feynman diagrams for the Diboson background.

The Diboson process refers to the production of two vector bosons ( $WW, WZ, ZZ$ ). In this process, the events of interest are those where the one of the two bosons decays leptonically, and the other to a pair of quarks. These events have a total cross section of order  $\sim 100pb$  and yield an almost negligible contribution, for the final state of interest.

#### QCD Multijet production

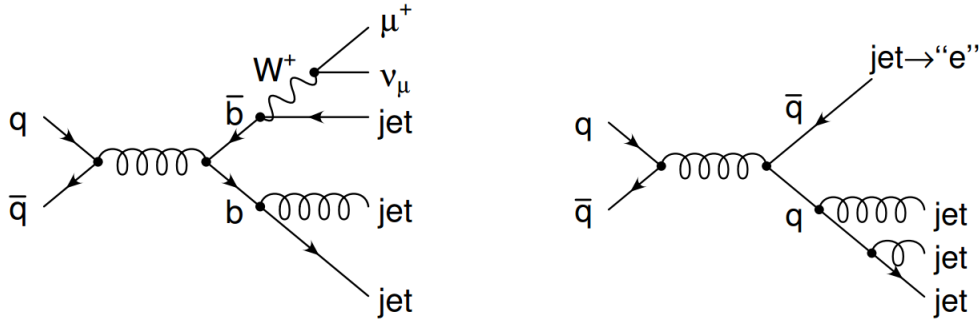


Figure 3.8: Examples of Feynman diagrams for the QCD Multijet background.

The QCD Multijet background originates from parton  $2 \rightarrow 2$  interactions, which give two jets and extra jets coming from initial or final state radiation. For some events a lepton (electron or muon) is produced through decays of heavy flavored quarks, or from produced hadrons. Therefore these events may contain such leptons which are called “non-prompt”, and mimic the final state of signal. Also the existence of a lepton can be reproduced by misidentification of a jet in the detector as an electron, which is called “fake”. Although these phenomena for this process are very rare, compared to the total, the result still is significant due to the large cross section which is of order  $\sim mb$ .

### 3.3 Monte Carlo Samples and Data Set

As mentioned in the previous section, the signal is the single top (ST) t-channel production, and there are several other background processes, which can be subdivided into categories (rest of ST,  $t\bar{t}$ , W+Jets, Drell-Yan, Diboson). For both the signal and the background processes one needs theoretical modeling of the distributions for the different variables that can be measured at the detector ( $p_T, E, m_T, \dots$ ) to compare them with

Process	Generator	Cross Section[pb](Precision)
Single Top (ST)		
t-channel top	PowhegV2	136.02(NLO)
t-channel anti-top	Powheg	80.95(NLO)
tW top	—''—	35.6(NNLL)
tW anti-top	—''—	35.6(NNLL)
s-channel	MadGraph@NLO	10.32(NLO)
Top-Anti-top (TT)		
$t\bar{t}$	Powheg	832(NNLO)
W+Jets		
0 Jets	Madgraph@NLO	49670(NLO)
1 Jet	—''—	8264(NLO)
2 Jets	—''—	2628(NLO)
Drell-Yan(DY) (or Z+Jets)		
$H_T \in [70, 100]$ GeV	Madgraph	169.9(LO)
$H_T \in [100, 200]$ GeV	—''—	147.4(LO)
$H_T \in [200, 400]$ GeV	—''—	40.99(LO)
$H_T \in [400, 600]$ GeV	—''—	5.678(LO)
$H_T \in [600, 800]$ GeV	—''—	1.367(LO)
$H_T \in [800, 1200]$ GeV	—''—	0.6304(LO)
$H_T \in [1200, 2500]$ GeV	—''—	0.1514(LO)
$H_T \in [2500, +\infty)$ GeV	—''—	0.003565(LO)
Diboson		
WW	Pythia8	115.00(NLO)
WZ	Pythia8	47.13(NLO)
ZZ	Pythia8	16.523(NLO)

Table 3.1: Monte Carlo event generator samples used in this analysis, and their corresponding cross sections with their precisions.

the data and extract parameters estimations. Thus it is needed the shape of these distributions and an initial estimate of the corresponding number events for each process, therefore the cross section. For this purpose Monte Carlo event generators are used. The Monte Carlo samples that are used in this study, and the corresponding cross sections, are shown in Table 3.1.

Each of these samples contains a number of generated events. Beyond the precision of the cross sections and the MC characteristics (shower evolution, hadronization model, detector etc.) the samples carry, due to their random nature, statistical fluctuations. The bigger the number of the generated events, the less statistical fluctuations the simulation has, therefore yields a more accurate shape. In practice, one wants the statistical fluctuations of the simulation to be smaller from these of the data. Thus the number of generated events is needed to be bigger enough from the number of events that will be measured in the real data. More precisely this should be the case in the part of phase space that is of interest. A process may be abundant in the total phase space but very suppressed in the area of interest. Then even if someone generates many events, very few may be useful. For example for the DY process the differential cross-section falls rapidly with the energy. To solve problems like these, one splits the phase space into subsets with respect to some variable, and generates events separately on each subset. In the example of DY the variable is the  $H_T$ , as seen in Table 3.1.

Now after the selection of events inside the phase space of interest has been made, it is needed to estimate the expected number of events i.e. the expected yield  $N_{exp.}$ . For this reason the integral of the distributions, i.e. the number of events that pass the selection  $N_{pass}$ , gets rescaled using the cross section and the integrated luminosity of the detector for the period that the data were recorded, to match the expected yield. For a process with index  $j$  e.g.  $j = W+jets$  the expected yield is calculated as follows

$$N_{exp.}^j = \sum_{i \in j} \sigma_i \mathcal{L}_{int} \frac{N_{pass}^i}{N_{gen.}^i} \quad (3.1)$$

where  $N_{gen.}$  is the total number of generated events, and  $i$  runs on all samples that belong to the process. This expression essentially tells that the fraction of events in the phase space of interest, estimated from monte carlo, is multiplied with the total expected yield  $\sigma_i \mathcal{L}_{int}$ , to give the expected yield in that region.

As for the data that are used in this analysis, they are recorded with the CMS detector in 2016 year. As seen in Figure 3.9 the integrated luminosity, for this period, delivered is about  $40.99 fb^{-1}$  and the recorded  $37.80 fb^{-1}$ . But for our analysis it is needed to have a lepton in the final state, electron or muon<sup>2</sup>, and therefore the datasets used come from events where specific triggers are fired. Thus the integrated luminosity is not that of the whole data recorded data set, but was measured about  $35.9 fb^{-1}$  with 2.5% uncertainty.

Of course the basic selection, coming from the triggers is not enough and one needs to consider in which area(s) of phase space is the most profitable to work. This discussion

<sup>2</sup>The data sets names are SingleElectron, SingleMuon.



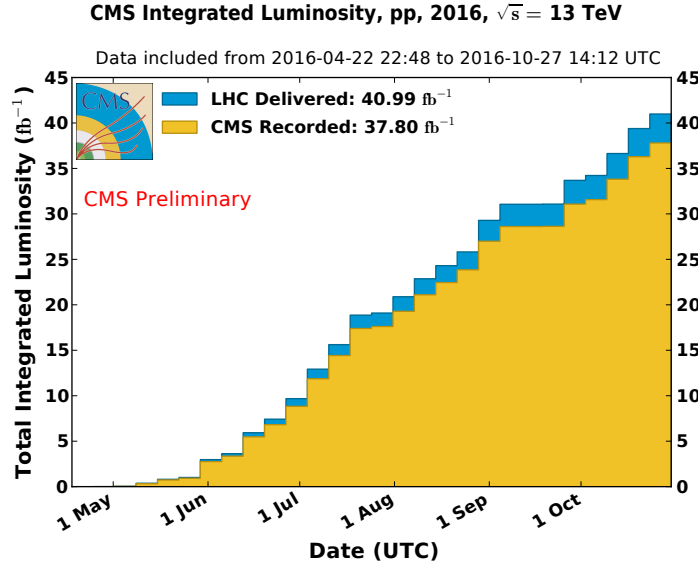


Figure 3.9: Cumulative distribution of luminosity measured with the CMS detector per week for 2016 period.

takes place in the next subsection.

### 3.4 Event Selection

From the total data they are considered for the analysis some parts of the phase space. These regions are defined through cuts on different variables. First of all there is signal region i.e. the region which contains the more signal events, which is required to have 2 jets, 1 of which is b-tagged and 1 isolated lepton (2J1T region). For the lepton it is required to have  $p_T > 35\text{GeV}$  and  $|\eta| < 2.1$ , if it is electron, and  $p_T > 30\text{GeV}$  and  $|\eta| < 2.4$  if it is muon. As for the jets  $p_T > 40\text{GeV}$  for all, and  $|\eta| < 2.4$  for b-tagged and  $|\eta| < 4.7$  for non b-tagged. The leptons are required to fire the HL triggers of tight isolation, which for electrons are

- HLT\_Ele27\_eta2p1\_WPTight\_Gsf\_v
- HLT\_Ele32\_eta2p1\_WPTight\_Gsf\_v

and for muons

- HLT\_IsoMu27\_v
- HLT\_IsoTkMu24\_v

The goal of this analysis is to find the ratio of the signal cross section from the nominal one, using a maximum likelihood fit. The fit will use the signal region since it contains the signal, but also it is needed to have other regions in order to have better estimations of the different backgrounds. Thus they are also used three control regions (CRs), each of which

will be used to constraint a corresponding background. The CR for W+Jets is the 1J1T region, which contains mostly W+Jets and QCD Multijet backgrounds. For  $t\bar{t}$  the CR is the 2J2T with 2 leptons of opposite sign and flavor i.e.  $(e^\pm\mu^\mp)$ , which is almost pure with  $t\bar{t}$ , it just has some small ST  $tW$  contribution. For  $DY$ , or Z+jets, the region 2J1T with 2 leptons of same flavor and opposite sign is selected. This region contains mainly  $DY$  and  $t\bar{t}$  events, and small  $t\bar{t}$  and Diboson backgrounds. But from the invariant mass of two leptons which only for  $DY$  is a peak at 90GeV, this background can be constrained.

### 3.5 Data Driven QCD Multijet Background

The simulation of QCD Multijet background process is very challenging. This comes from the fact that although the total cross section is very high, of order of mb, we are only interested in a phase space where events that can mimic the existence of an isolated lepton may pass the selection<sup>3</sup>. Therefore a very small fraction of total QCD multijet events can do that, but due to the large cross section these are enough to give a significant amount of background. The tiny selection efficiency leads immediately to some important problems for a simulation. One basic problem is that to generate events that pass the selection criteria, a much larger total number of events need to be generated leading to computationally intense production of such samples. Someone could argue that this is not a problem, since you may split your phase space and do what was described before at section 3.3. But in order to do that it is needed very good precision at the differential cross section. This is related to one other problem that comes up, which is that the higher order corrections become very important for events that can pass the selection.

For the reasons mentioned, QCD shape and normalization are estimated through a data-driven approach. The procedure for a region e.g. 2J1T is the following

- Divide the data of the region depending to the lepton isolation property. That means that there is one region with an isolated lepton(Iso region), meaning far from jets, that is defined through the cuts:
  - 1) 1 Lepton
  - 2) 0 Non Isolated Leptons
  - 3) Triggers for isolated lepton fired (different for electrons and muons)
 which is the region that we will work with for the analysis, and a side-band region(or Non Iso region), that is orthogonal to the previous, with a non-isolated lepton defined again through the cuts:
  - 1) 0 Leptons
  - 2) 1 Non Isolated Leptons
  - 3) Triggers for non-isolated lepton fired (different for electrons and muons)

---

<sup>3</sup>These events contain “non-prompt” leptons i.e. leptons not coming from the primary vertex but from produced hadrons or heavy quark(b,c) decays, or “fake” leptons which is the case that a hadronic jet signature is identified as a lepton.

- Assuming the side-band region has high purity in QCD Multijet events, for a variable  $X$  the shape of data is taken as the template of the multijet background distribution  $QCD(X)$ .
- Then it is assumed that the shape  $QCD(X)$  from the Non Iso region, is approximately the same in the Iso region, since physically these events have the same nature but they are misinterpreted by our limited efficiencies.
- Then for a rough estimate of QCD yield in the Iso region, it is calculated the difference of events in data and the total MC estimations of the other processes

$$N_{QCD} = N_{data} - N_{MC} \quad , N_{MC} = N_{t\text{-chan.}} + N_{t\bar{t}} + \dots \quad (3.2)$$

this estimation is of course not taken to be precise, it is just used for having a nominal value for this background.

The regions of interest which are expected to contain QCD Multijet events are the signal region 2J1T, and W+jets control region 1J1T. For these two regions the yields for events with electrons and muons in the final state are shown in Table 3.2.

Region	<b>2J1T</b>		<b>1J1T</b>	
<b>process</b>	$e$	$\mu$	$e$	$\mu$
t-channel	30612	58412	19527	38189
tW	13983	23985	5665	9731
s-channel	1058	2134	914	1816
$t\bar{t}$	84312	151108	22236	39335
W+Jets	112866	207928	287593	531593
Z+Jets(DY)	10870	14784	10961	20109
Diboson	2261	3968	2116	4552
QCD Multijet	60571	160963	158634	523859
Data	316532	623181	507747	1169180

Table 3.2: Nominal yields for different processes, and the total from data, for 2J1T and 1J1T regions.

For the 2J1T region, the QCD multijet yield is comparable with the  $t\bar{t}$  background, and consists one of the three dominant backgrounds. As for the 1J1T it is the dominant background of W+jets with respect to other processes. For these regions the basic kinematic variables are shown in Figures 3.10-3.13, which are the lepton, and jet(s)  $p_T, |\eta|$ , and the  $m_T^W, p_T^{miss}$ .

As seen, the QCD Multijet leptons tend to be more soft i.e. have low  $p_T$ , and the same holds for the  $E_T^{miss}$ . The electrons have an increasing distribution with  $|\eta|$ , while the muons decreasing. This happens since for the electrons the jets that yield “fake” electrons tend to be more forward. For the electrons there is a fall in  $1.5 < |\eta| < 1.7$  which is due

to the space between the barrel and the endcap in detector. Also, as expected, the  $m_T^W$  distribution does not have a peak, and the distribution for electrons falls more smoothly than muons. For the light-jet the  $|\eta|$  distribution has a sudden increase from 2.5-3.0 and this is due to pile-up. Then it has a sudden fall since there the Endcap of the detector ends.

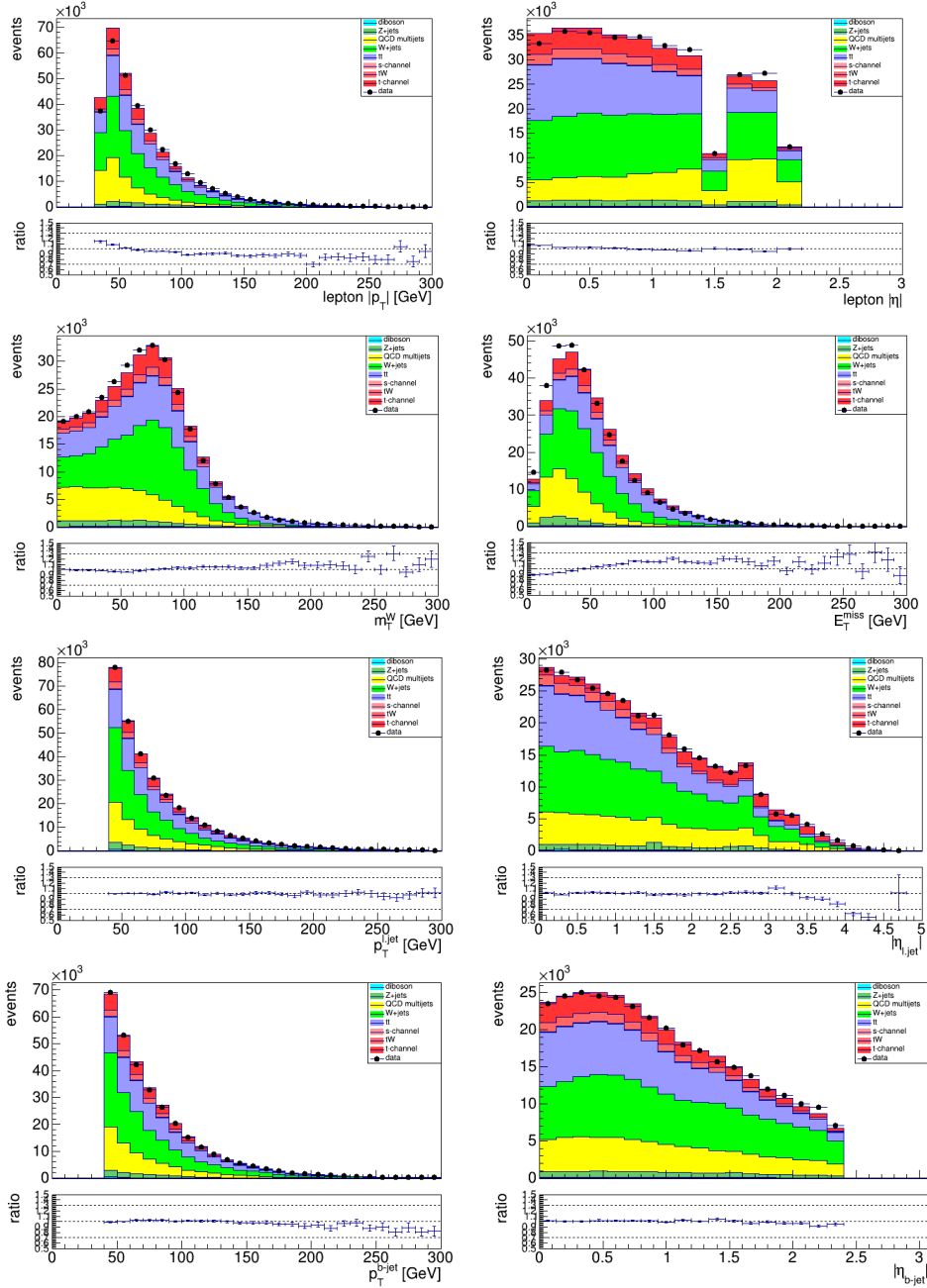


Figure 3.10: Stacked histograms of distributions of different processes for kinematic variables in 2J1T 1 electron category.

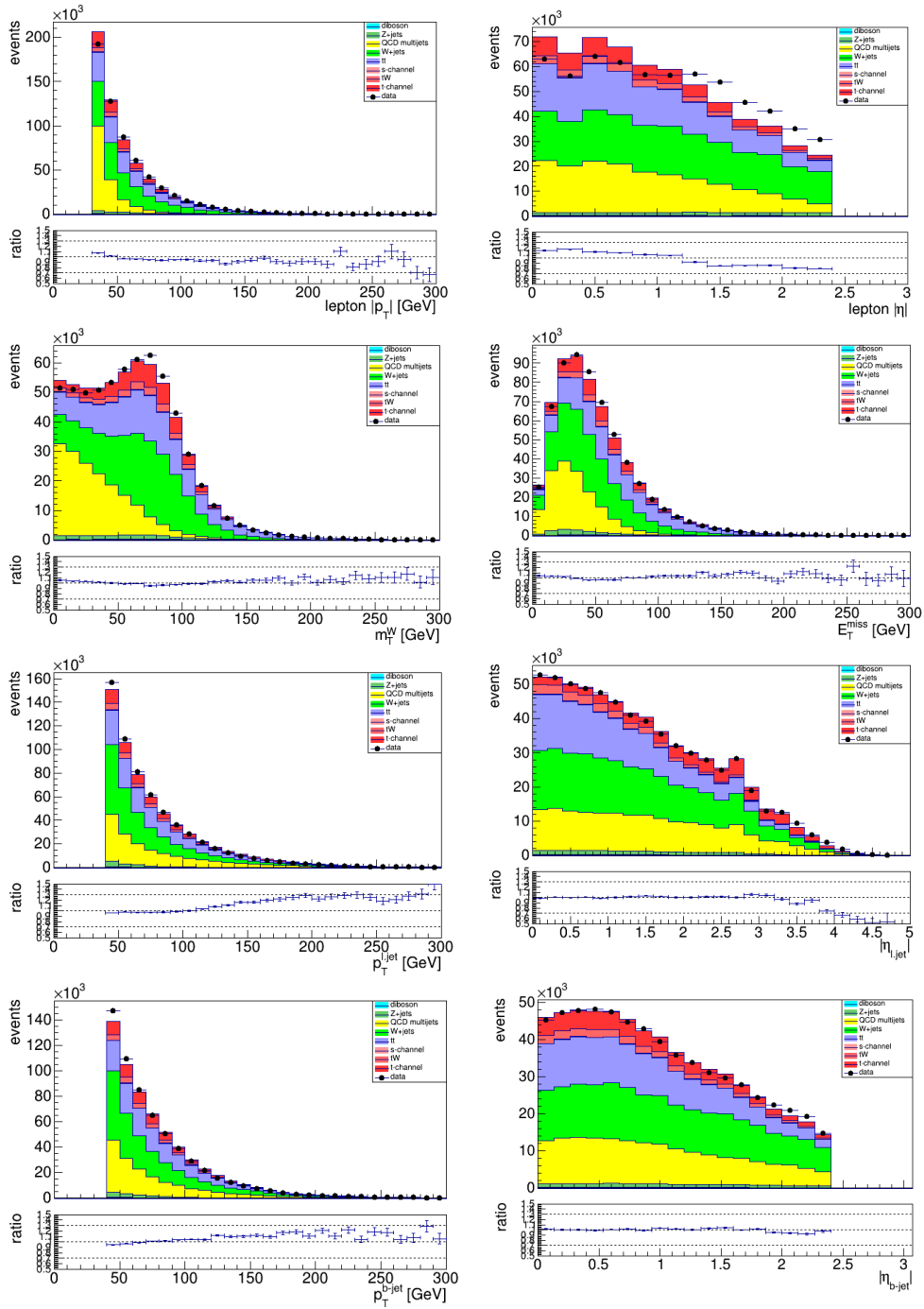


Figure 3.11: Stacked histograms of distributions of different processes for kinematic variables in 2J1T 1 muon category.

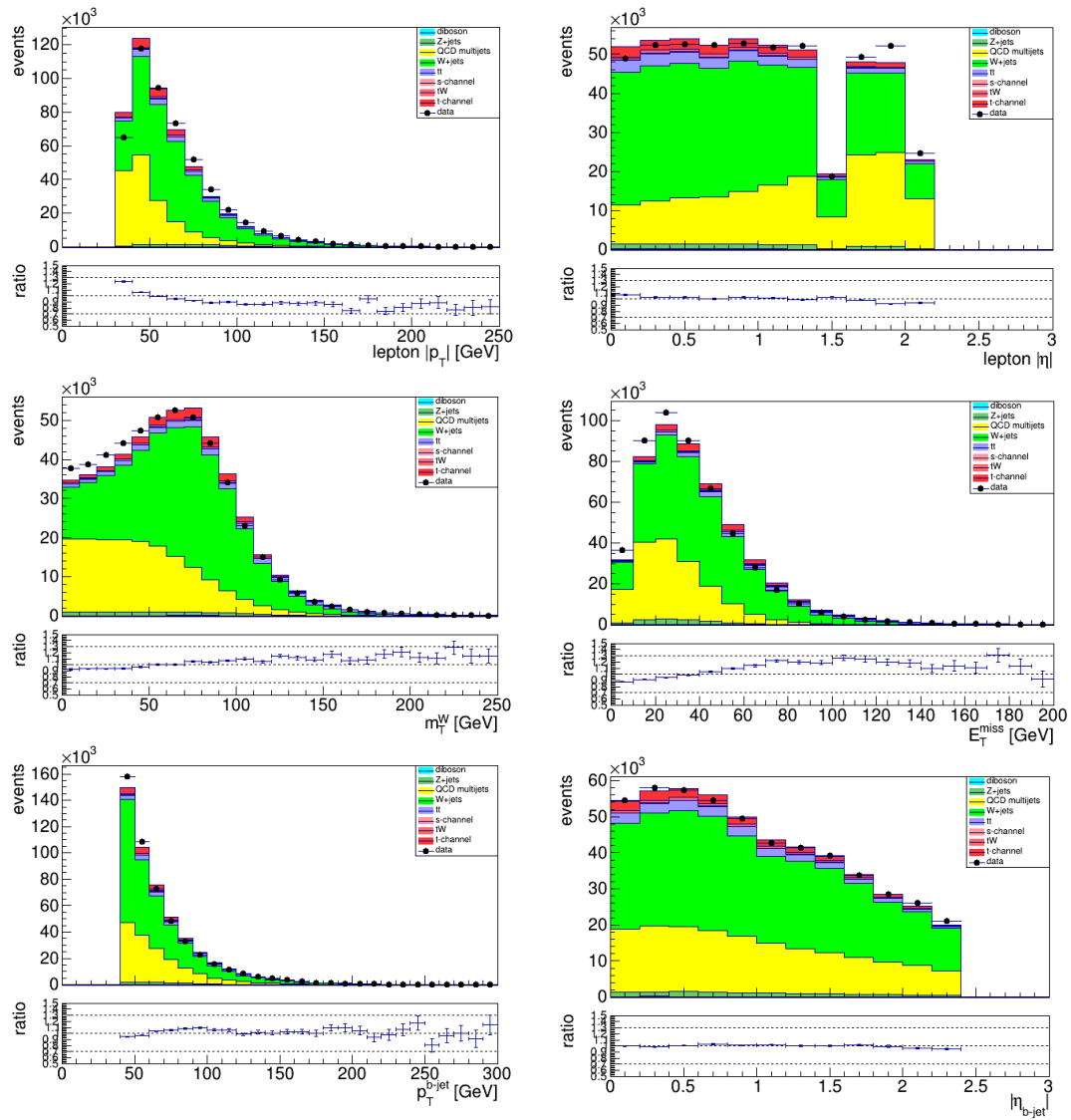


Figure 3.12: Stacked histograms of distributions of different processes for kinematic variables in 1J1T 1 electron category.

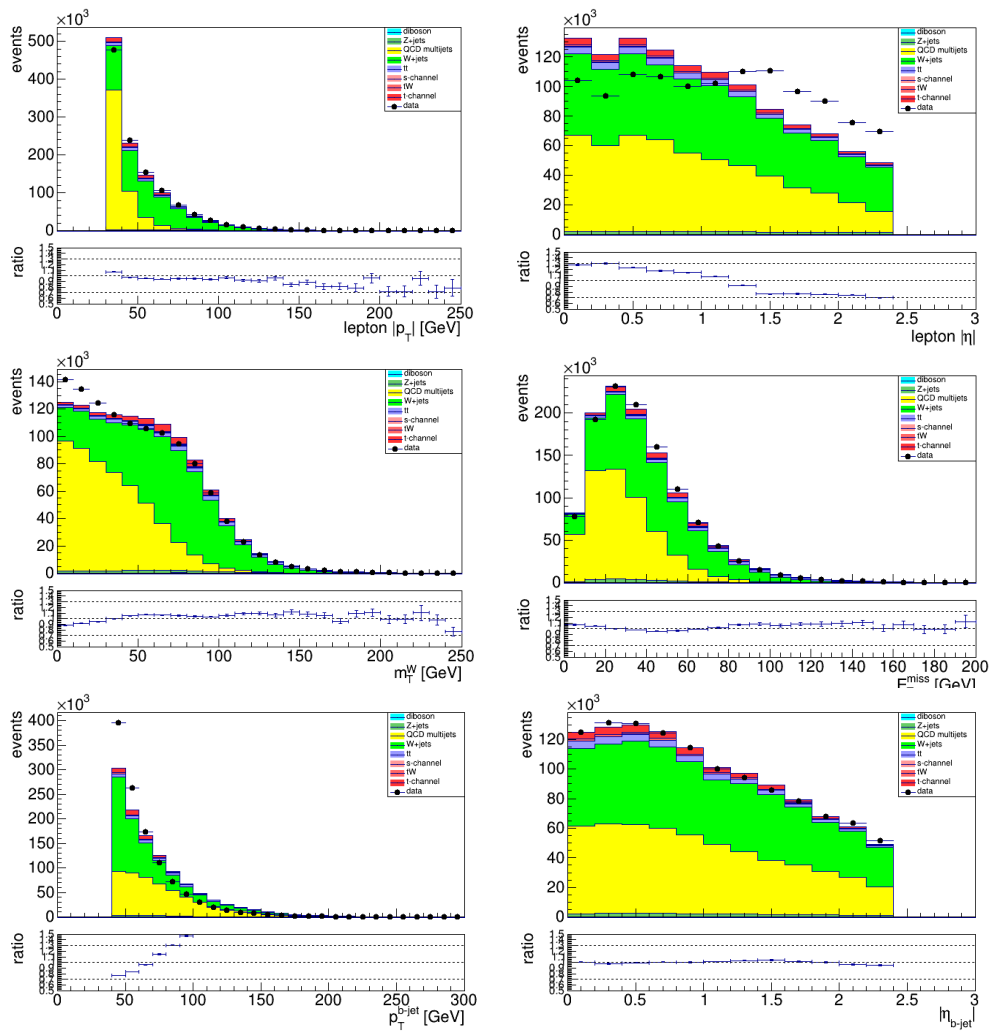


Figure 3.13: Stacked histograms of distributions of different processes for kinematic variables in 1J1T 1 muon category.

### 3.6 Top Quark Reconstruction

Apart from the kinematic variables that can be calculated directly from the detector information, single top quark events are offered for reconstructing the top quark and thus allowing more possibilities for variables to be explored. This happens since the final state of single top is pretty clear and consists of a lepton and a neutrino, coming from the  $W$  leptonic decay, and a b-tagged jet. In order to calculate the 4-vector of top quark  $p_{top}$  the 4-momentum conservation at  $tWb$  vertex is used

$$p_{top} = p_w + p_{b-jet} \quad (3.3)$$

For the b-tagged jet, the information needed i.e.  $\vec{p}_T, E$ , invariant mass  $m_{b-jet}$ , is available from the jet reconstruction. Now, for  $W$  there is again a conservation of 4-momentum at  $Wl\nu$  vertex

$$p_W = p_l + p_\nu \quad (3.4)$$

Again the  $p_l$  comes from lepton reconstruction, but for the neutrino which is not observable in the detector, the only information comes from  $\vec{p}_T^{miss}$ . Thus, even if someone accepts the  $\vec{p}_T^{miss}$  as the neutrino  $\vec{p}_T^\nu$ , the longitudinal component  $p_z^\nu$  is missing. In order to fix the missing degree of freedom, the  $W$  boson is required to be on-shell i.e.  $p_W^2 = (M_W)^2$ , with  $M_W = 80.4\text{GeV}$ .<sup>4</sup> Then

$$(p_l + p_\nu)^2 = 2p_l \cdot p_\nu = M_W^2 \quad (3.5)$$

This is a second order equation with respect to  $p_z^\nu$ , which is the only unknown. Solving this equation the solutions are

$$p_z^\nu = \frac{\Lambda p_z^l}{p_T^{l2}} \left[ 1 \pm \sqrt{D} \right] \quad (3.6)$$

where

$$D = \sqrt{1 - \frac{p_T^{l2}}{\Lambda^2} (p_l^2 p_T^{\nu 2} - \Lambda^2)}, \quad \Lambda = \frac{m_w^2}{2} + \vec{p}_T^l \cdot \vec{p}_T^\nu \quad (3.7)$$

The sign of value of the determinant has physical meaning namely

- $m_T^W < M_W$  then  $D > 0$
- $m_T^W = M_W$  then  $D = 0$
- $m_T^W > M_W$  then  $D < 0$

The  $D = 0$  case in practice never happens exactly but approximately, and in this case the  $p_z^l$  is almost uniquely defined (there are 2 solutions very close to each other). There is the case  $D > 0$  where there are two solutions for  $p_z^\nu$ . From these the convention used here is that the smaller one is chosen. Finally for the  $D < 0$  case, which yields imaginary solutions which are not physically acceptable, there must be a choice of treatment of



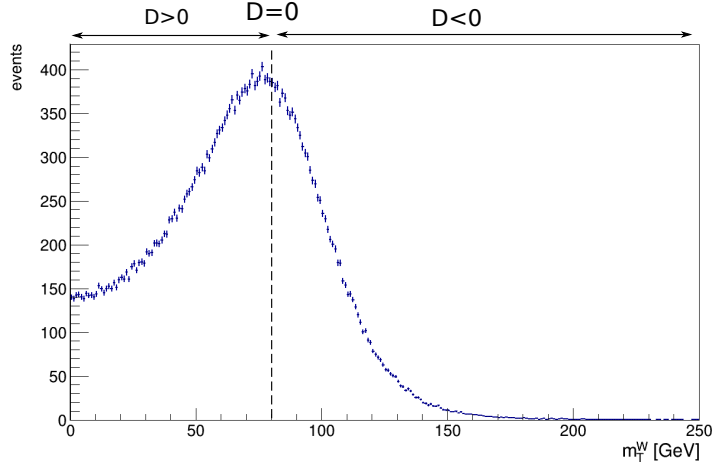


Figure 3.14: Distribution of  $m_T^W$  for the ST t-channel process.

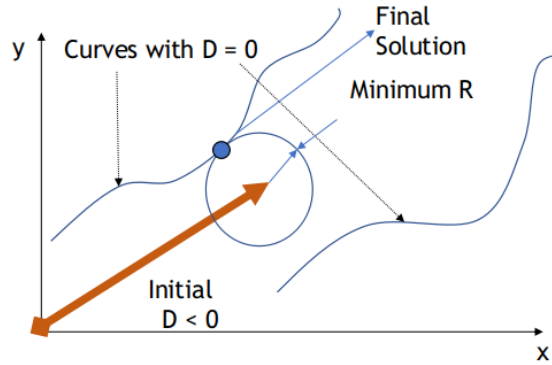


Figure 3.15: Representation of the CMS algorithm for cases of  $m_T^W > M_W$ .

these cases. As shown in Figure 3.14 for many events  $m_T^W > m_W$  holds and therefore these events cannot be simply rejected, but it is needed a way to cope with the problem.

This way is arbitrary and there are different procedures in literature. Here it is followed the standard CMS algorithm, in which the missing  $\vec{p}_T$  is modified, which is reasonable since it has the lower resolution.<sup>5</sup> In this procedure, which is graphically shown in Figure 3.15, for every  $p_x$  component it is found the corresponding  $p_y$  which satisfies  $D = 0$  or equivalently  $m_T^W = M_W$ . This yields two curves on  $p_x - p_y$  plane, which all are candidates for the  $\vec{p}_T$ . From those the closest solution to the initial measured  $E_T^{miss}$  vector is chosen.

From the top reconstruction two main variables were extracted, for the discrimination of signal versus the background distributions. The first is the reconstructed top quark mass  $m_{top}$ . The probability distributions of the t-channel and the main background processes are shown in Figure 3.16. The t-channel and the  $t\bar{t}$  distributions have a peak close to the

<sup>4</sup>Of course this is not always the case, and the mass follows a Breit-Wigner distribution with mean value  $M_W$  and width  $\Gamma_{W \rightarrow l\nu}$ . But the ratio  $\Gamma_{W \rightarrow l\nu}/M_W \sim 0.01$ , therefore it is a reasonable approximation.

<sup>5</sup>Of course different procedures should not give incompatible answers. A validity check was done using an alternative method used in ATLAS collaboration.

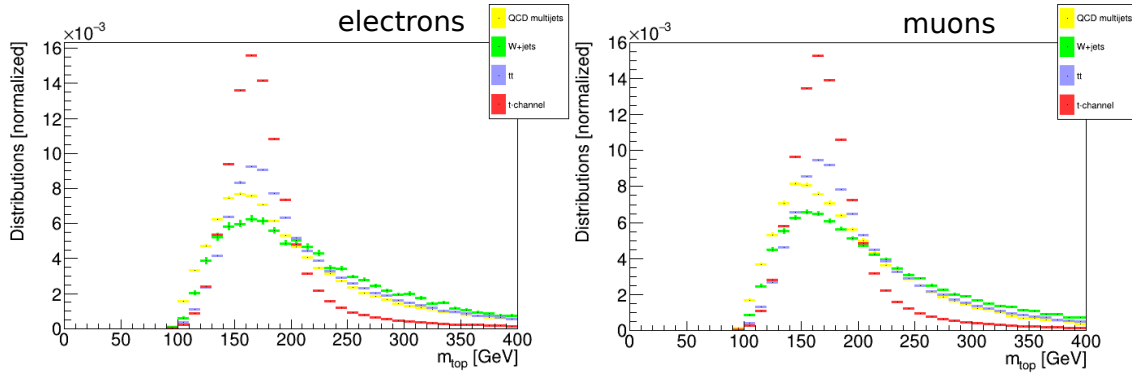


Figure 3.16: Probability distributions of top mass for the signal and the main background processes, for electrons (left) and muons (right).

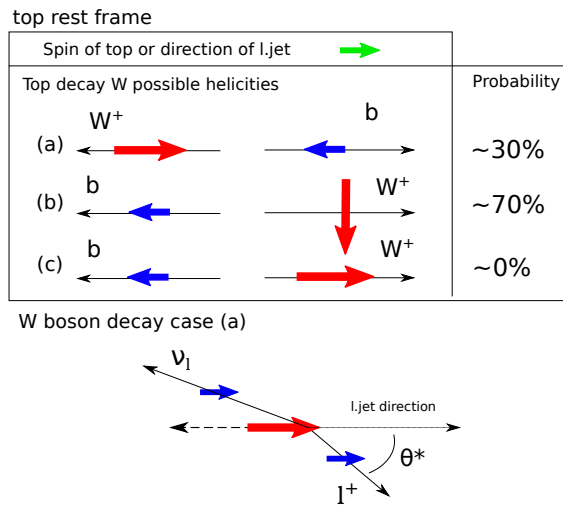


Figure 3.17

value of top mass, while the QCD and W+jets backgrounds near to  $\sim 150\text{GeV}$  since they do not contain top quarks. The  $t\bar{t}$  peak has bigger width from the t-channel, which is caused from events where the lepton and the b-tagged jet were not originating from the same top quark.

Another important discriminating variable is the  $\cos(\theta^*)$ , which is the angle between the lepton and the light quark jet direction, in the top quark rest frame. This variable exploits the feature of t-channel to have the top quark highly polarized in the light-jet direction. Due to this fact the spin of W and therefore the lepton and neutrino have a preference which leads to having bigger probability of  $\cos(\theta^*) > 0$ , as shown in Figure 3.17. The probability distributions for this variable are shown in Figure 3.18, which shows there is an asymmetry of order 30% for the t-channel distribution, while the other distributions are more flat.

In addition to the reconstructed top mass they were defined two alternatives using the transverse components of the decay products, and will be referred to as “transverse top

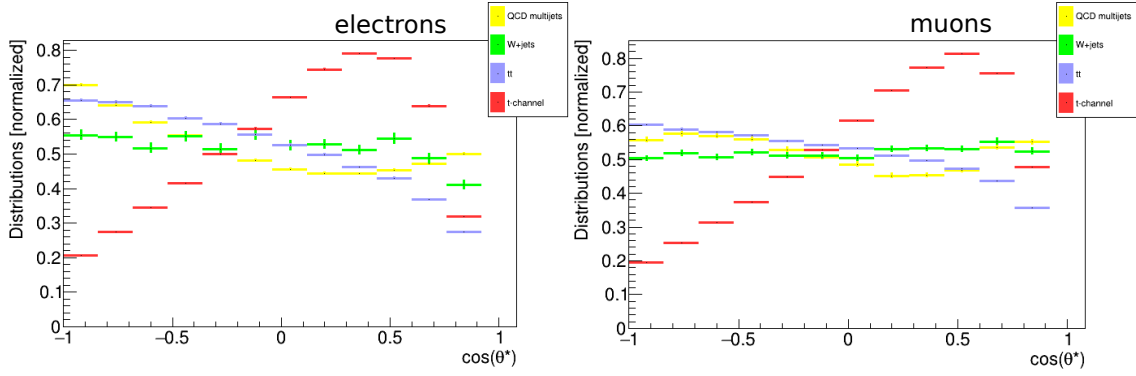


Figure 3.18: Probability distributions of  $\cos(\theta^*)$  for the signal and the main background processes, for electrons (left) and muons (right).

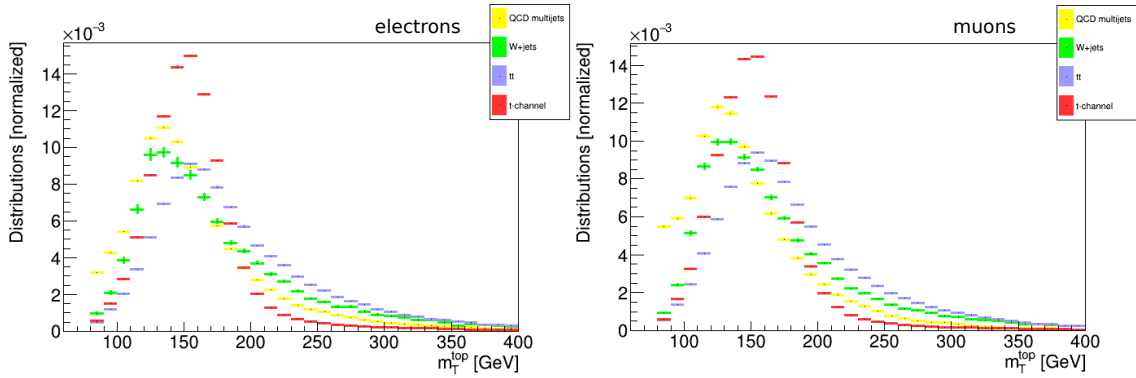


Figure 3.19: Probability distributions of  $m_T^{top}$  for the signal and the main background processes, for electrons (left) and muons (right).

mass". The first way is motivated by the definition of mass through the 4-vectors

$$p_{top} = p_W + p_{b-jet} \quad \Rightarrow \quad m_{top}^2 = m_W^2 + m_{b-jet}^2 + 2p_W \cdot p_{b-jet} \quad (3.8)$$

and taking the  $m_W = 80.4 GeV$  fixed, the invariant mass of b-jet calculated from its 4-vector, and the last term with the interior product calculated from transverse 4-vectors  $p_T^W, p_T^{b-jet}$  the definition is taken to be

$$m_T^{top} \equiv \sqrt{m_W^2 + m_{b-jet}^2 + 2p_T^W \cdot p_T^{b-jet}} \quad (3.9)$$

The probability distributions for this variable are shown in Figure 3.19. It is observed that in contrast with  $m_{top}$  coming from the reconstruction this variable has more discriminating shapes for the different processes. The width of t-channel is again smaller than this of  $t\bar{t}$ , but they have a peak difference. Also the peaks of QCD, W+Jets are more far, from the distributions with top quarks.

The second definition that was used was that of  $m_T^{top(2)}$  defined purely from transverse components

$$\left(m_T^{top(2)}\right)^2 = \left(p_T^W + p_T^{b-jet}\right)^2 = (m_T^W)^2 + (m_T^{b-jet})^2 + 2p_T^W \cdot p_T^{b-jet} \quad (3.10)$$

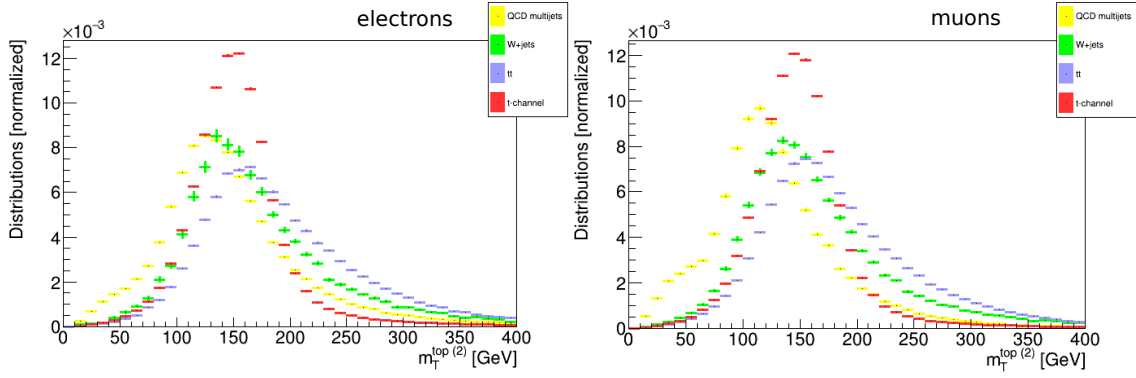


Figure 3.20: Probability distributions of  $m_T^{top(2)}$  for the signal and the main background processes, for electrons (left) and muons (right).

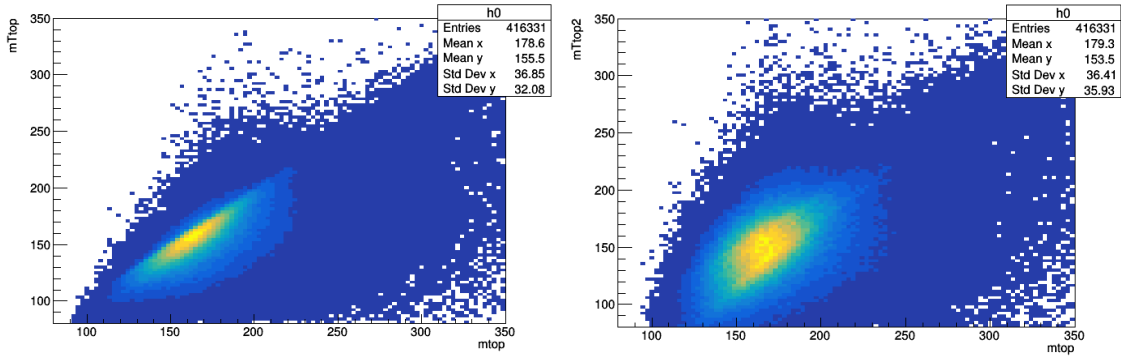


Figure 3.21: Correlation plots of reconstructed  $m_{top}$  and  $m_T^{top(2)}$  (left), or  $m_T^{top(2)}$  (right) for simulated events of signal (t-channel).

This definition has more broad distributions than  $m_T^{top}$ , as expected since it has the  $m_T^W$  term instead of  $m_W$  which is a fixed number. Other than that it has similar characteristics to the other definition.

From the two definitions the first is more close to the reconstructed top mass, as it is shown in Figure 3.21.

### 3.7 Fit Variable Study

In this section it is studied the choice of a variable at which a Maximum Likelihood (ML) fit<sup>6</sup> will be performed, in order to extract the signal cross section scale factor  $r_{t-ch}$ , where for a process  $i$  the

$$r_i \equiv N_i^{fit} / N_i^{nom}. \quad (3.11)$$

where  $N_i^{fit}$  is the fitted yield, and  $N_i^{nom}$  is the nominal value of events for process  $i$ , estimated from MC as it was seen before. These  $r_i$  parameters get in the fitted template

<sup>6</sup>The theory about ML fitting is contained in Appendix A.

distribution  $F(X)$ , of variable  $X$ , in the following way

$$F(X) = \sum_i r_i N_i^{nom.} PDF_i(x) \quad (3.12)$$

For this choice there are two main aspects that will be taken under consideration. The first one is that the shape variable used for the fit should be discriminating for the signal (ST t-channel) with respect to dominant background processes ( $t\bar{t}, W+jets$ ). This is a minimum requirement for the fit to be meaningful. If that doesn't happen then strong (anti)correlations will be created between the signal and the background, and the problem of maximizing the Likelihood becomes ill-defined. The second is that in addition the ML fit should be well behaved, meaning it should be robust to statistical fluctuations and give an unbiased estimation for the value and the error. This is checked through the Toy Monte Carlo methodology, that is generating multiple sets of pseudodata and performing the fit on them in order to have an ensemble of results. From this ensemble it may be checked that the fit is healthy, regarding the characteristics that were mentioned.

For the following work they will be used the MC templates for the electron channel, in the Signal Region i.e. 2J1T.

### 3.7.1 Discriminating Variables

In order to quantify the idea of discrimination power for a variable it is needed to have some quantity that measures it. The problem of discriminating the signal from the main backgrounds is essentially a classification problem with three classes denoted as  $\omega_\alpha$  with an index  $\alpha = t - chan., t\bar{t}, W + jets$  labeling the processes. Thus they will be used some class separability measures for the quantification. Two measures will be used for the candidate variables that will be checked for consistency at the implementation. These measures are

- Divergence (D) :  $D_{\alpha\beta} = \int [p(x|\omega_\alpha) - p(x|\omega_\beta)] \log \left( \frac{p(x|\omega_\alpha)}{p(x|\omega_\beta)} \right) dx$ . This is a positive definite measure and the bigger its value, the most discriminant is the variable between the two classes. It is also easy to see that in case that the distributions are exactly the same  $D_{\alpha\beta} = 0$ , and that it satisfies the property  $D_{\alpha\beta} = D_{\beta\alpha}$ .<sup>7</sup>
- Area Under Curve (AUC): AUC is the area under the ROC curve of  $\omega_\alpha$  and  $\omega_\beta$ .<sup>8</sup>. Its value is 0.5 when the two classes are unseparable with respect to the variable, and the deviation from this value is a measure of the discrimination power.

The candidate fit variables that will be checked are

- a Boosted Decision Tree (BDT)

<sup>7</sup>For the motivation of definition and more about this measure see "Pattern Recognition" by S. Theodoridis, K. Koutroumbas. Section 5.6

<sup>8</sup>Curve of rejection,  $1 - \epsilon_\beta$ , of class  $\omega_\beta$  with respect to selection efficiency  $\epsilon_\alpha$  of class  $\omega_\alpha$  for all possible cuts of the variable (see Appendix B)

- $\Delta R(\text{lepton}, \text{b-jet})$
- $|\eta_{l,jet}|$
- $m_T^W$
- $m_T^{top}$
- $m_T^{top(2)}$

The BDT is a Multivariate Discriminant, that is a method which combines the information from multiple variables in order to optimize the discrimination between certain classes. For details about BDTs one may read the Appendix B. The distributions of input variables used for the classifier are shown in Figure 3.22. Their descriptions are the following

Variable	Description
lJetEta	Absolute value of pseudorapidity for light jet
mtop	Reconstructed top quark mass
DijetMass	Invariant mass of the two jets
dRlBMin	Distance $\Delta R$ between lepton and b-jet
c_theta_star	$\cos(\theta^*)$ with $\theta^*$ being the angle of lepton and light jet in top quark rest frame
PtSumJets	Sum of the two jets transverse momenta
mtw	W boson transverse mass
met	missing transverse momentum
dRJets	Distance $\Delta R$ between the two jets
lepEta	Absolute value of pseudorapidity of lepton
WbosonEta	Absolute value of pseudorapidity of reconstructed W boson
lJetMass	Invariant mass of light jet
mTtop	The definition of transverse-like mass of top quark

The BDT uses 400 Trees for the forest, with Gradient Boosting method. The Shrinkage parameter was set to 0.2 and the Trees have maximum depth of 2. For the separation of nodes the Gini index measure was used. These parameters were chosen by giving the best ROC curve, after different trials of combinations. The Training samples output of the classifier for signal and background distributions are plotted in Figure 3.23. At the same plot they are shown the distributions for independent test samples, that confirm there is no overtraining. The BDT ROC curve is shown in Figure 3.24, together with the ROC for a Fisher discriminant for comparison. The BDT is more discriminating as it is expected, since it is a non-linear classifier in contradiction to Fisher.

Table 3.3 shows the Divergence and  $|\text{AUC} - 0.5|$  for the variables under consideration. It is observed that the BDT is the best for discriminating the signal from each of the main backgrounds. The other promising variables are  $|\eta_{l,jet}|$  which is the most discriminating

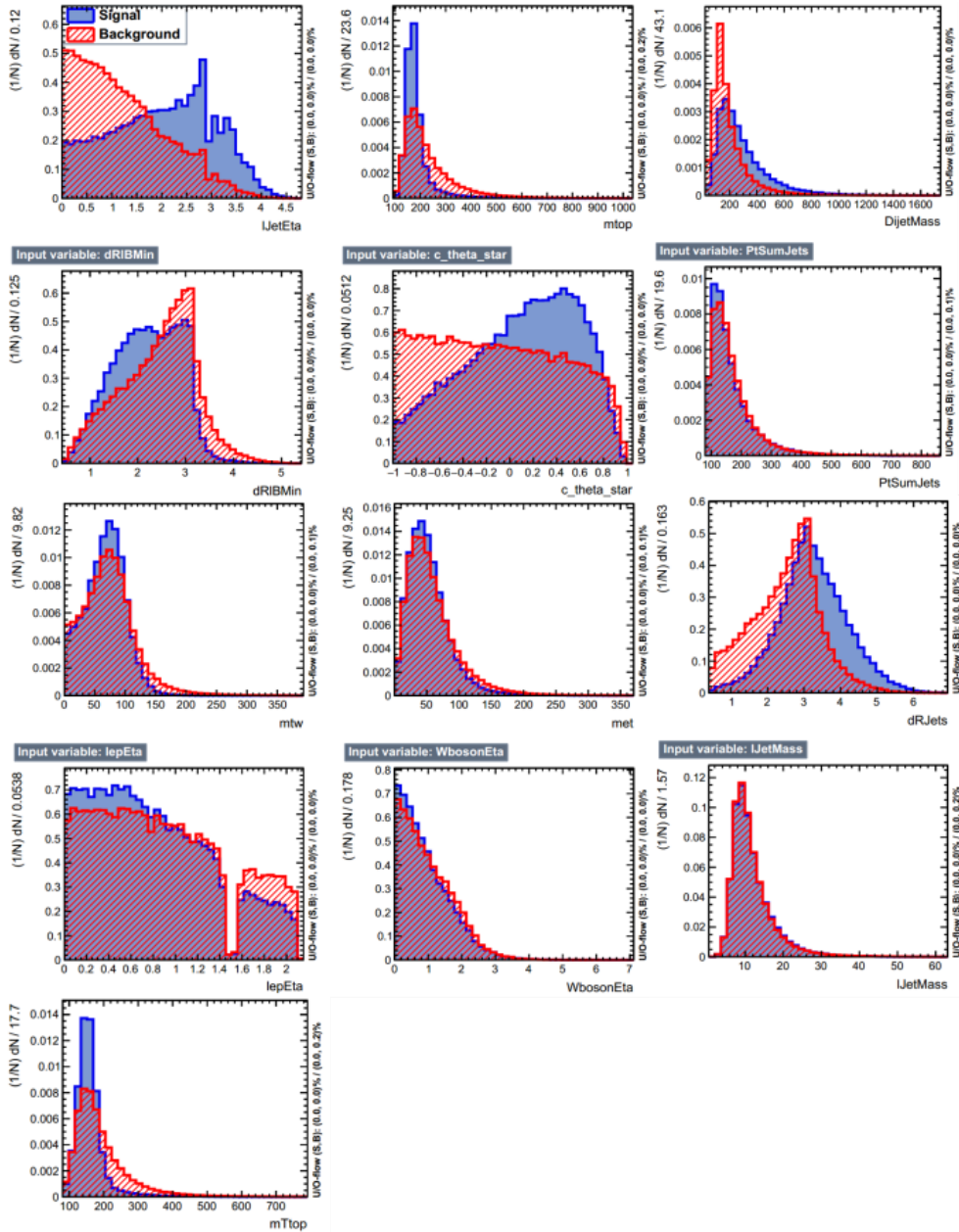


Figure 3.22: Variables distributions for signal(red) and background(blue) used in the BDT classifier.

single variable, due to the property of the light jet to be forward in ST t-channel events, and the  $m_T^{top}$ ,  $m_T^{top(2)}$ . The last two have less discriminating power between the signal and the two backgrounds than BDT or  $|\eta_{jet}|$ , although they have some significant difference for the two backgrounds.

There is a remark about the comparison of D and the AUC measures. As seen they generally agree, confirming the validity of this picture, but for some cases, like the  $m_T^{top}$  t-channel/W+jets, they may seem incompatible. This disagreement is not the case, it is rather a drawback of AUC measure. At these cases the difference stated by the D measure

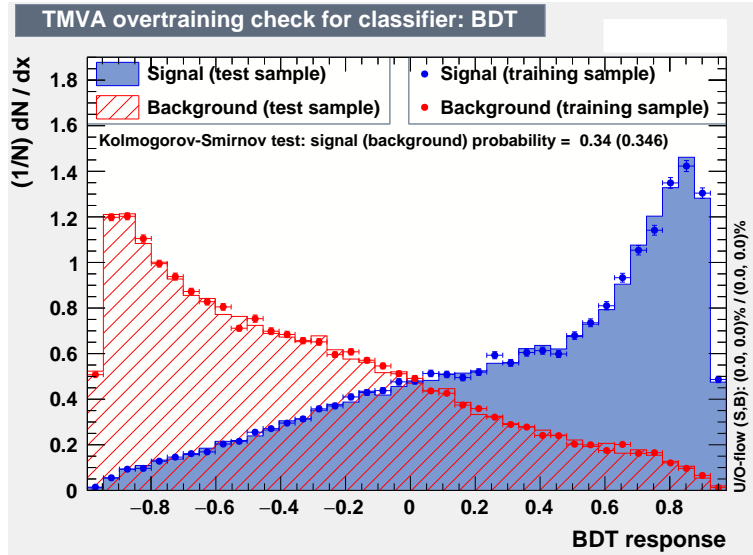


Figure 3.23: BDT response distribution for signal and background, for training and test sample.

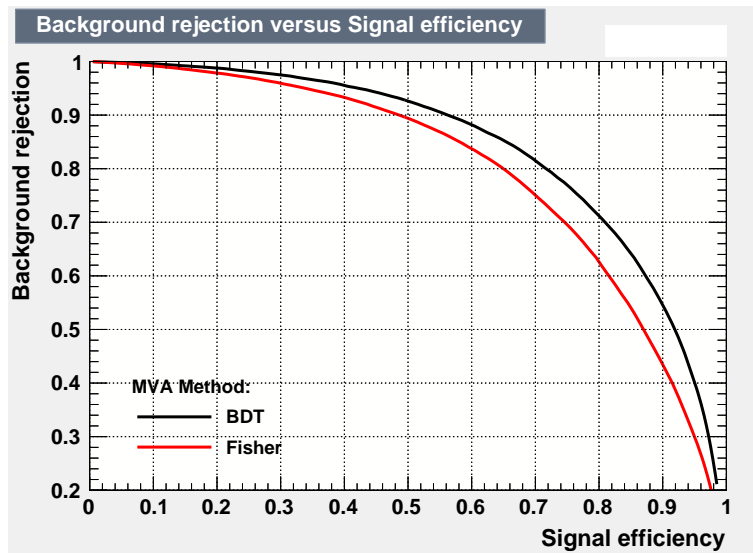


Figure 3.24: Receiver Operator Curve for the BDT and a Fisher discriminant.

is correct, and it manifests on the ROC curve (Figure 3.25a). It is just that the integral has cancellations if some part of the curve is smaller than  $y(x) = 1 - x$ , and other bigger<sup>9</sup>. When that does not happen, e.g. for the same variable  $t$ -channel/ $t\bar{t}$  the measures agree (Figure 3.25b). Thus the D is a better measure, since it measures better the local differences of distributions which are also reflected on the ROC curve but cannot be seen by AUC.

Now, having as candidates for the fit variable the BDT,  $|\eta_{jet}|$ ,  $m_T^{top}$ ,  $m_T^{top(2)}$ , they will be further checked about the behavior at a ML fit. This is the objective of the next section.

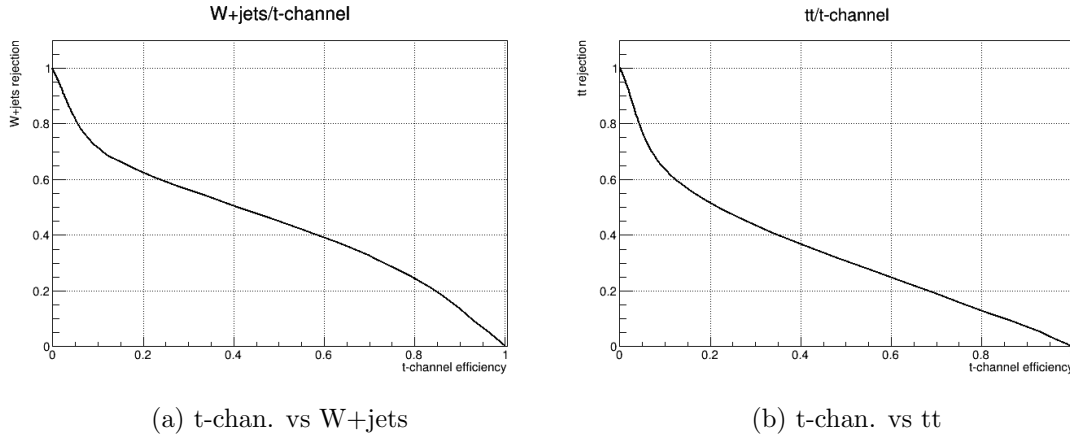
<sup>9</sup>Maybe a better measure would be  $\int |ROC(x) - (1 - x)| dx$



Variable:	BDT		dRlBMin		lJetEta	
$\omega_\alpha/\omega_\beta$	D	AUC - 0.5	D	AUC - 0.5	D	AUC - 0.5
t-channel / W+jets	1.75	0.32	0.35	0.14	0.42	0.18
t-channel / $t\bar{t}$	1.94	0.33	0.07	0.04	0.83	0.24
$t\bar{t}$ / W+jets	0.03	0.01	0.14	0.09	0.07	0.06

Variable:	mTw		mTtop		mTtop2	
$\omega_\alpha/\omega_\beta$	D	AUC - 0.5	D	AUC - 0.5	D	AUC - 0.5
t-channel / W+jets	0.03	0.02	0.34	0.05	0.27	0.07
t-channel / $t\bar{t}$	0.29	0.05	0.51	0.26	0.59	0.18
$t\bar{t}$ / W+jets	0.14	0.07	0.13	0.18	0.11	0.19

Table 3.3: The class separability measures values for the discriminating variables.

Figure 3.25: ROC curves for t-channel versus  $t\bar{t}$  and W+jets backgrounds, for  $m_T^{top}$  variable.

### 3.7.2 Toy MC fit validation

The goal of this section is to check the properties of a ML fit on the discriminating variables candidates. This is done by using pseudodata, that the fit is performed on them. More specifically, the algorithm that is implemented is the following

1. Choose a specific signal strength  $r_{t\text{-chan.}} = N_{t\text{-chan.}}^{in} / N_{t\text{-chan.}}^{nom.}$ .
2. For each background process  $i$ , generate a yield  $N_i^{in.} \sim \text{Poisson}(N_i^{nom.})$ .
3. Generate  $N_i^{in}$  events for the process  $i$  (background or signal) that follow the distribution of the corresponding template. After add all these events and make the generated events dataset (pseudodata).
4. Perform the ML fit on the pseudodata set.
5. Repeat the above multiple times.

The template for the fit on the pseudodata sets is

$$f(x) = N_{t-ch}PDF_{t-ch}(x) + N_{tt}PDF_{tt}(x) + \dots \quad (3.13)$$

with Gaussian constraints on all backgrounds yields, except of QCD. The mean values of the Gaussians are the nominal estimations from MC  $N_i^{nom}$ , and the standard deviations are  $10\%N_{tt}^{nom}$  for the  $t\bar{t}$ , and  $20\%$  for all the rest backgrounds. These values are arbitrary, but they are for sure bigger than the known precision for these values so the results will be reliable. For the  $t\bar{t}$  a smaller sigma was given, because its cross section is of higher theoretical precision, than rest of processes.

After the procedure mentioned the output quantities are histograms of

1. Fitted yield minus the inserted yield  $N_{fit} - N_{in}$ .
2. Error  $\delta N_{fit}$ (or  $\sigma_{fit}$ ).
3. Pulls for  $N$  i.e.  $g = \frac{N_{fit} - N_{in}}{\sigma_{fit}}$
4. Correlation coefficients of the signal yield t-channel with the yield of  $t\bar{t}$ , and the yield of  $W + jets$ .

From these distributions the main requirements, for a fit to be well behaved are

- pull distribution to be a standard Gaussian, with mean value  $\langle pull \rangle = 0$ , which means unbiased estimation of yield, and  $\sigma_{pull} = 1$ , which means unbiased yield error estimation. Of course these values will not be exactly 0 or 1 respectively, but compatible with them within their error. They will be closer to these, for larger number of toys which drops the errors.
- Small (anti)correlations of signal with dominant backgrounds, which means that the estimated value of the signal yield is independent of that of background processes yields. This is connected with the discrimination power studied in the previous section. For example if two processes are almost identical, then the fit, in order to describe some feature of data, may raise the one process and drop the other, which leads to an anti-correlation.

It was observed that at many cases the error extracted with the parabolic approximation, implemented with HESSE, gave two peaks at its distribution. In terms of Likelihood shape, which changes for each dataset, this means the approximation of the paraboloid around maximum is not always valid. An example is shown in Figure 3.26 which comes from signal strength=1, fits on  $m_T^{top}$  variable. In the same Figure the exact asymmetric errors  $\sigma_{lo}, \sigma_{up}$  are shown, implemented with MINOS. It is observed that for roughly half of the fits the parabolic approximation is good, but for the rest the error is overestimated.

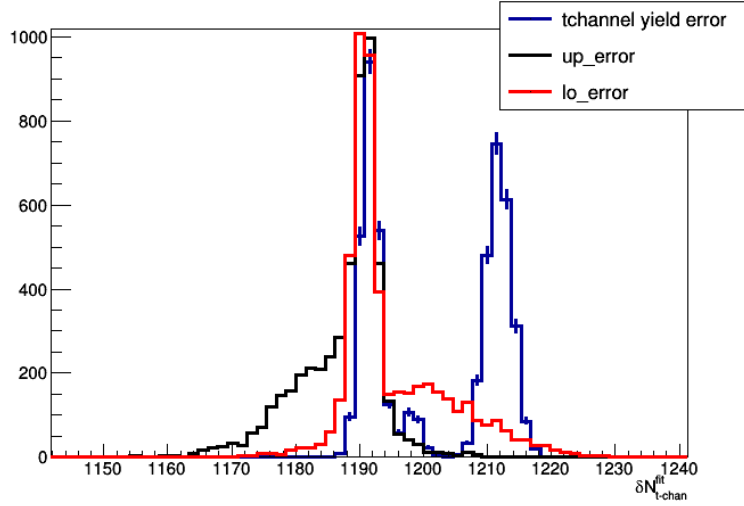
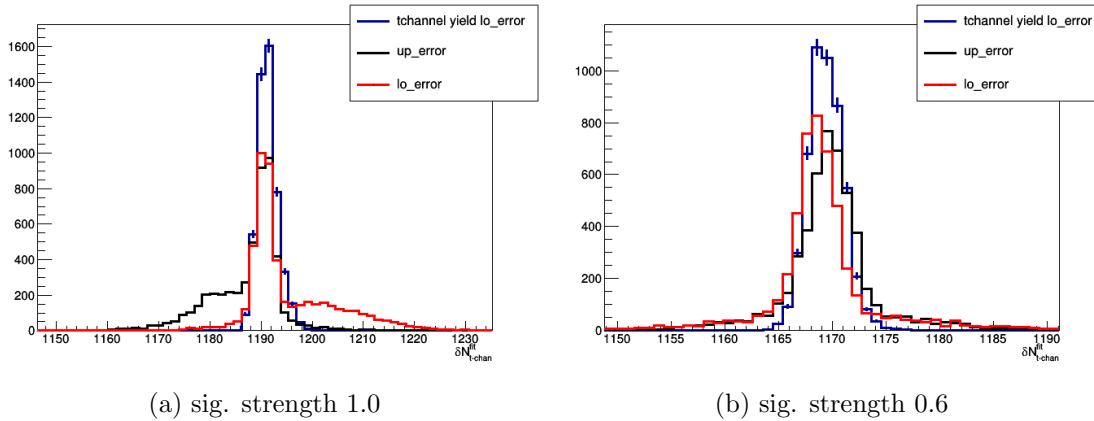


Figure 3.26: Distributions of parabolic approximation error  $\sigma$  (blue) and asymmetric errors  $\sigma_{lo}$  (red),  $\sigma_{up}$  (black), for signal strength 1.0. The fit variable is the  $m_T^{top}$

In order to cope with this problem the definition of error chosen for a fit is the symmetrized error

$$\sigma_{sym} \equiv \frac{\sigma_{lo} + \sigma_{up}}{2} \quad (3.14)$$

which was found to have a good behavior. For the example used the corresponding plot is shown in Figure 3.27a. Also, for the errors the differences between  $\sigma_{lo}, \sigma_{up}$  depend on the signal strength as seen in 3.27b.



(a) sig. strength 1.0

(b) sig. strength 0.6

Figure 3.27: Distributions of symmetrized error  $\sigma_{sym}$  (blue) and asymmetric errors  $\sigma_{lo}$  (red),  $\sigma_{up}$  (black), for different signal strengths. The fit variable is the  $m_T^{top}$

The symmetrized error is the sigma used for the pull definition. In order to check that for the distribution of this pull, there will be the same coverage between  $-\sigma_{pull}$  and  $+\sigma_{pull}$  as that of the interval from MINOS, it is also calculated an “asymmetric pull”  $g_{asym}$ . which

is defined as

$$g_{\text{asym.}} = \begin{cases} \frac{N_{\text{fit}} - N_{\text{in}}}{\sigma_{\text{lo}}}, & \text{if } N_{\text{fit}} > N_{\text{in}} \\ \frac{N_{\text{fit}} - N_{\text{in}}}{\sigma_{\text{up}}}, & \text{if } N_{\text{fit}} \leq N_{\text{in}} \end{cases} \quad (3.15)$$

The ‘‘asymmetric’’ part comes from the treatment of error. By its definition there is no guarantee this pull’s distribution is Gaussian, although if it is, it has  $\sigma$  with the correct coverage. In our case where  $\sigma_{\text{lo}}$  and  $\sigma_{\text{up}}$  are close, it is Gaussian<sup>10</sup>. To prove this, if  $\alpha$  the coverage of the MINOS interval for one estimation  $N_{\text{fit}}$ ,  $[-\sigma_{\text{lo}}, \sigma_{\text{up}}]$  then by definition

$$\begin{aligned} \alpha &= \Pr(-\sigma_{\text{lo}} < N_{\text{in}} - N_{\text{fit}} < \sigma_{\text{up}}) \\ &= \Pr(-\sigma_{\text{lo}} < N_{\text{in}} - N_{\text{fit}} < 0) + \Pr(0 \leq N_{\text{in}} - N_{\text{fit}} < \sigma_{\text{up}}) \\ &= \Pr\left(-1 < \frac{N_{\text{in}} - N_{\text{fit}}}{\sigma_{\text{lo}}} < 0\right) + \Pr\left(0 \leq \frac{N_{\text{in}} - N_{\text{fit}}}{\sigma_{\text{up}}} < 1\right) \\ &= \Pr\left(0 < \frac{N_{\text{fit}} - N_{\text{in}}}{\sigma_{\text{lo}}} < 1\right) + \Pr\left(-1 < \frac{N_{\text{fit}} - N_{\text{in}}}{\sigma_{\text{up}}} \leq 0\right) \end{aligned}$$

With the comments mentioned under consideration, the fit validation distributions for the different variable candidates were found. The results for 1000 generated pseudodata sets and signal strength 1.0 are shown in Figures (3.28 - 3.31).

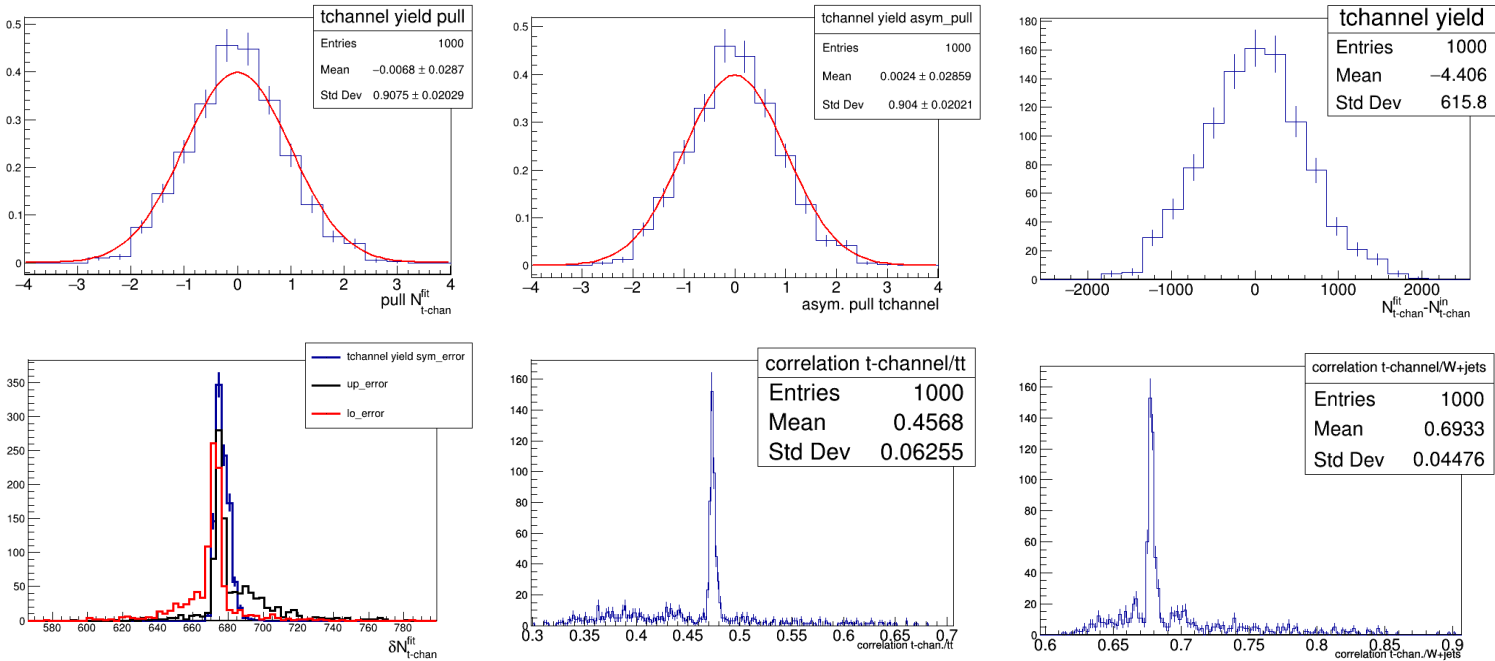


Figure 3.28: Results from Toy Monte Carlos for the BDT variable, for 1000 toys and sig. strength = 1.0.

<sup>10</sup>There is also a small bias in the mean value of this Gaussian which of order  $\frac{\sigma_{\text{lo}} - \sigma_{\text{up}}}{\sigma_{\text{lo}}}$

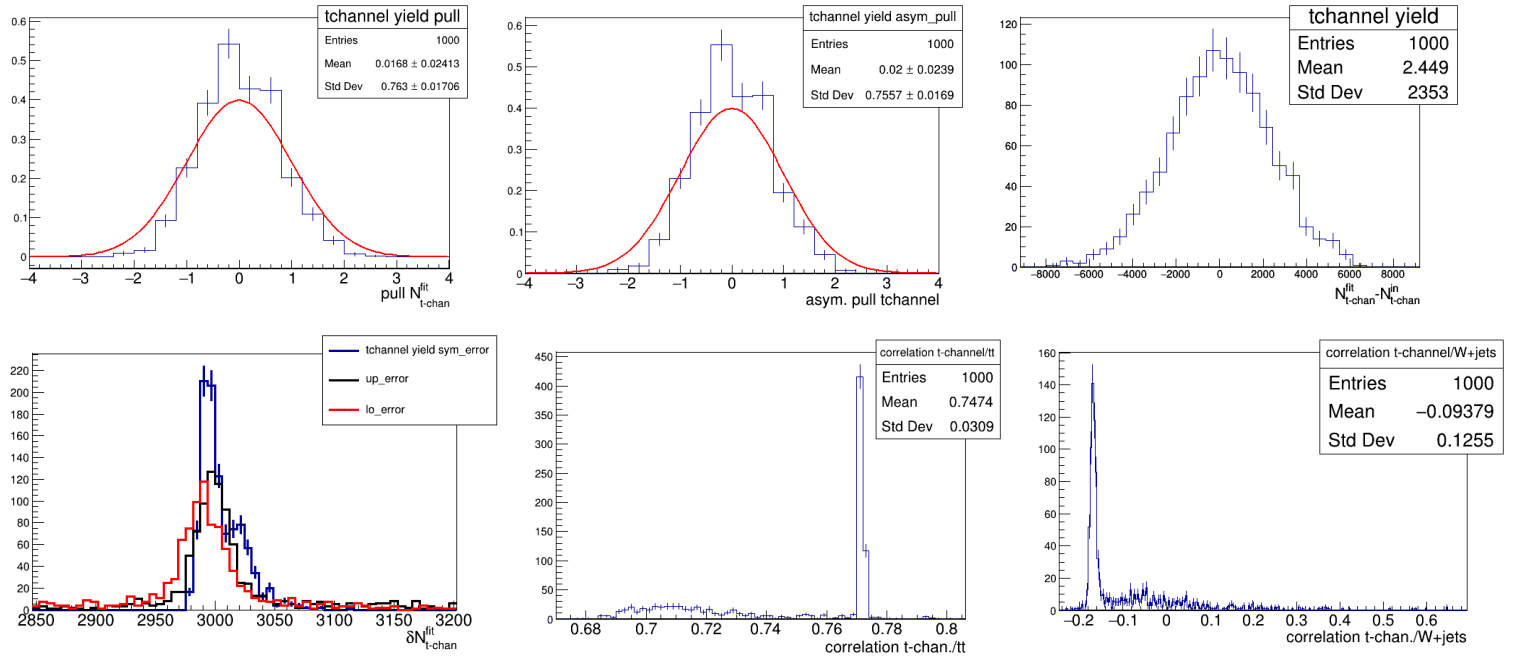


Figure 3.29: Results from Toy Monte Carlos for the  $|\eta_{jet}|$  variable, for 1000 toys and sig. strength = 1.0.

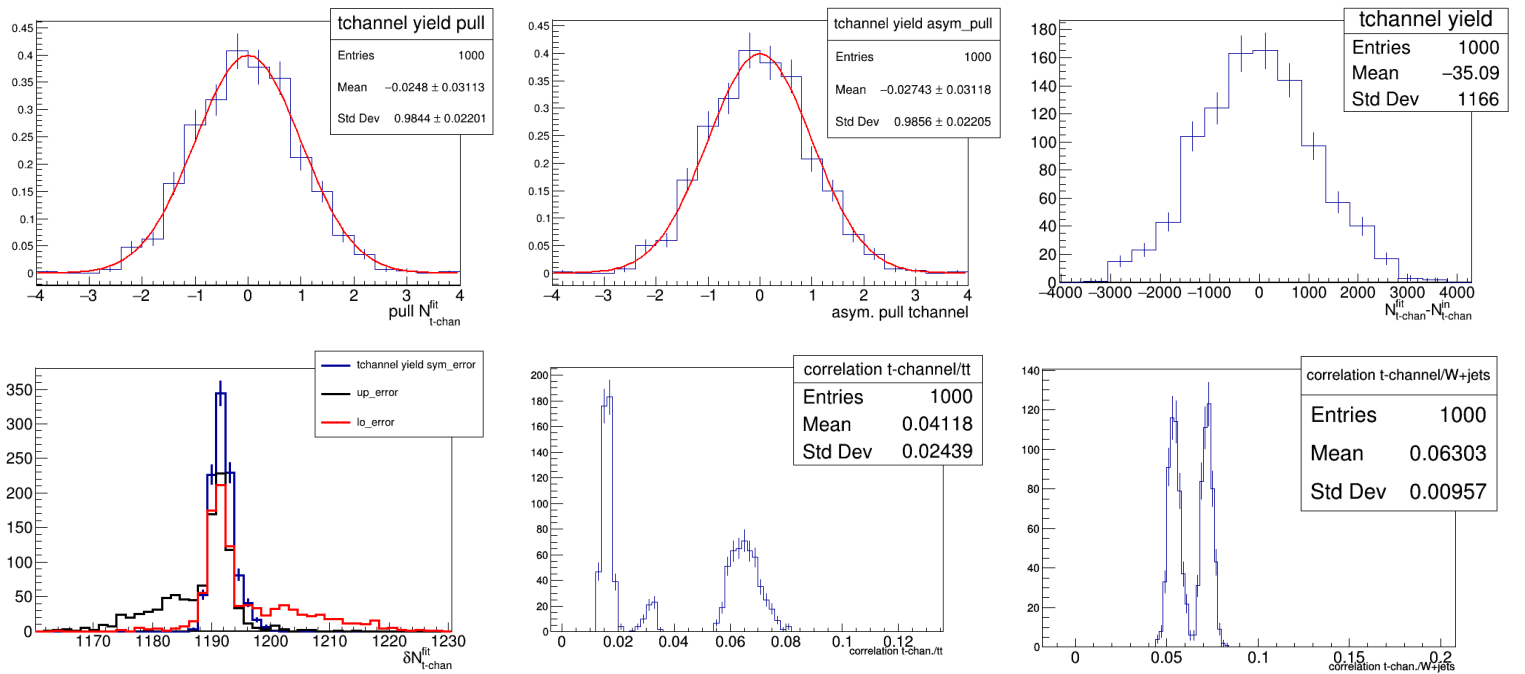


Figure 3.30: Results from Toy Monte Carlos for the  $m_T^{top}$  variable, for 1000 toys and sig. strength = 1.0.

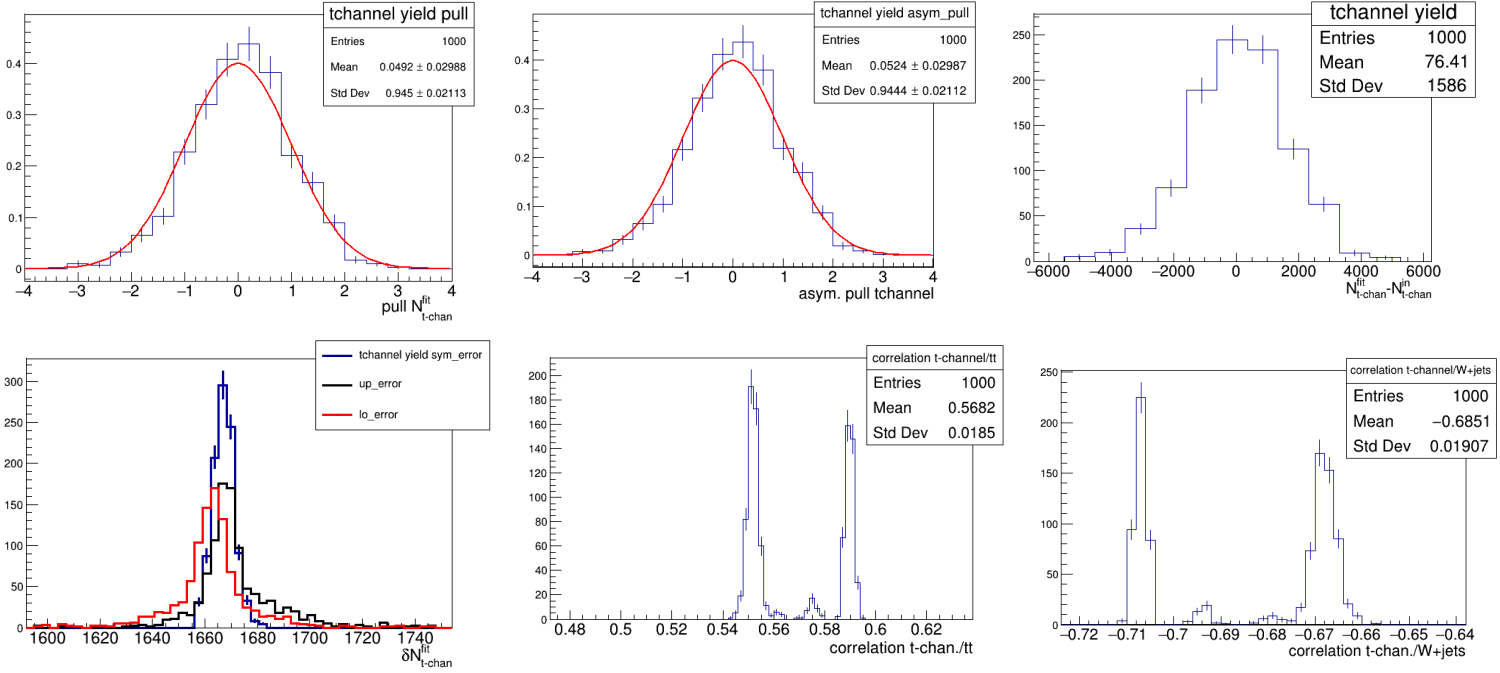


Figure 3.31: Results from Toy Monte Carlos for the  $m_T^{top, (2)}$  variable, for 1000 toys and sig. strength = 1.0.

The mean values of the pull distributions of signal yield, for all variables are compatible with zero, which means there is no bias in the estimated value. As for the error of the pull, which has the correct coverage validated by the  $g_{asym.}$ , the only variable which is unbiased is the  $m_T^{top}$  variable. These biases can be seen also, from the difference between the error of the  $N_{t-channel}^{fit} - N_{t-channel}^{in}$  distribution and the peak of  $\delta N_{t-channel}^{fit}$ . More specifically for each variable some remarks are the following

- BDT: There is 10% bias for the error. Also the correlations of signal with  $t\bar{t}$  and W+Jets are big enough and the distributions 2 peaks, a sharp and a broad one. This might be related with the 2 different kind of errors for the parabolic approximation error, since the covariance matrix is calculated using this approximation.
- $|\eta_{l,jet}|$ : There is bias 25% for the error. Again there are two peaks one sharp and one broad. There is very big correlation of signal with  $t\bar{t}$  always, and for most of cases small with W+Jets.
- $m_T^{top}$ : The error is unbiased, within the precision of estimated  $\sigma_{pull}$ . There are some peaks for the correlations of signal with the two main backgrounds. These are all sharp, and show there is small correlation.
- $m_T^{top, (2)}$ : The error has some small bias. Like  $m_T^{top}$  are some sharp peaks at the correlations distributions, but the difference is that they appear significant correlations.

Taking these remarks under consideration, it is concluded that the variable for the fit should be  $m_T^{top}$ . For the  $m_T^{top}$  variable the  $\langle pull \rangle$ ,  $\sigma_{pull}$ , and mean correlation with tt and W+jets coming from 5000 pseudodata sets, for different signal strengths were calculated and they are shown in Figure 3.32. The results show that the fit value is unbiased, the error has some small underestimation of order  $\sim 3\%$  and the correlations rise for signal with  $t\bar{t}$  and fall with W+Jets, but they are always small. This check validates the good behavior of the fit, for different possible values of signal yield.

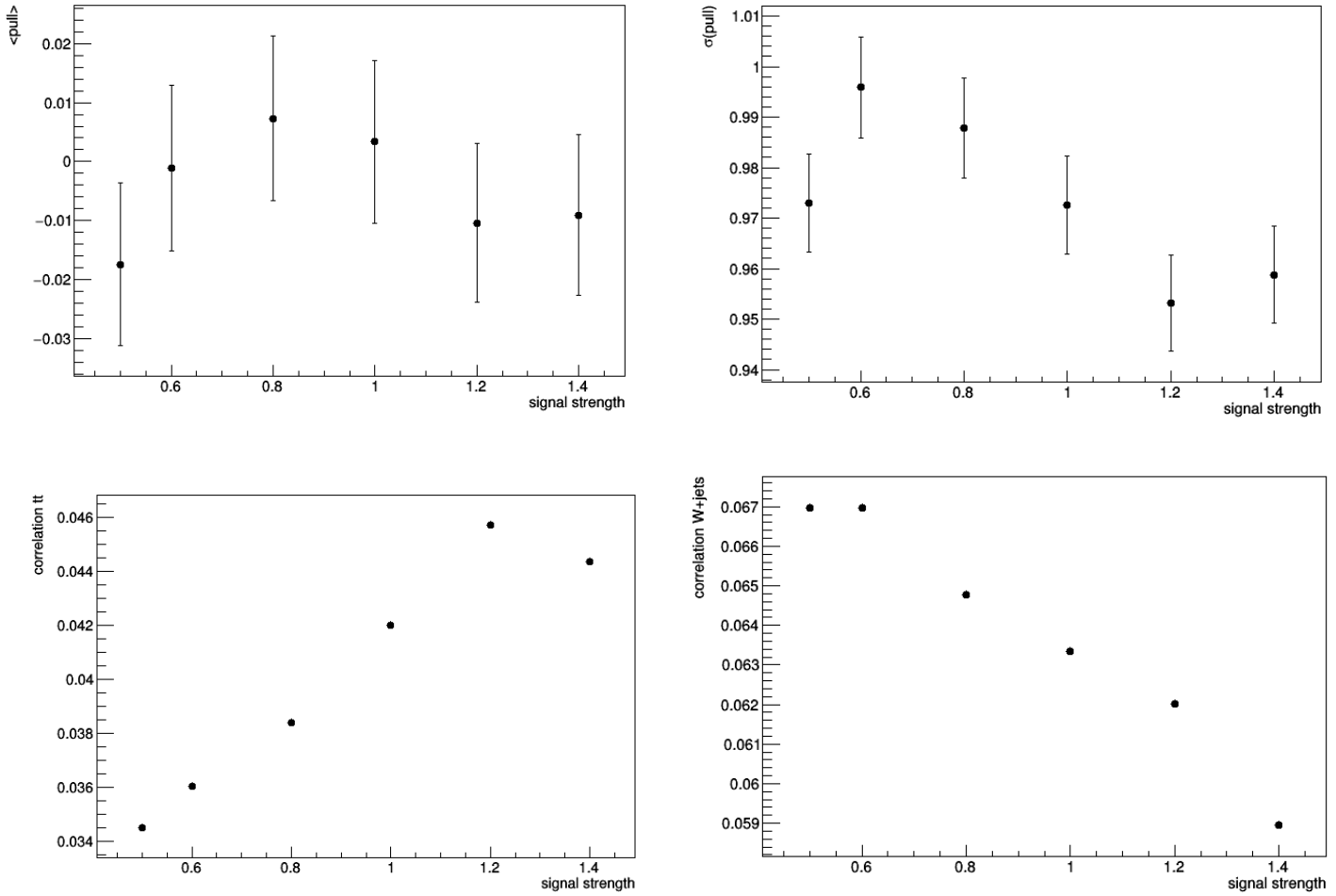


Figure 3.32: Results from Toy MCs of 5000 pseudodatasets with different signal strengths, for  $\langle pull \rangle$ ,  $\sigma(pull)$ , correlations of t-channel with tt and W+jets.

It is also important for the following, to mention that through the toy MC procedure, the fits were found to have no sensitivity regarding the the smallest processes of diboson and ST s-channel. In particular the relative error assigned to them by the fit was the one given by the Gaussian constraint, unless the later was set very big e.g. 100% where multiple values all bigger than 40% where found. These remarks can be seen from Figure 3.33.

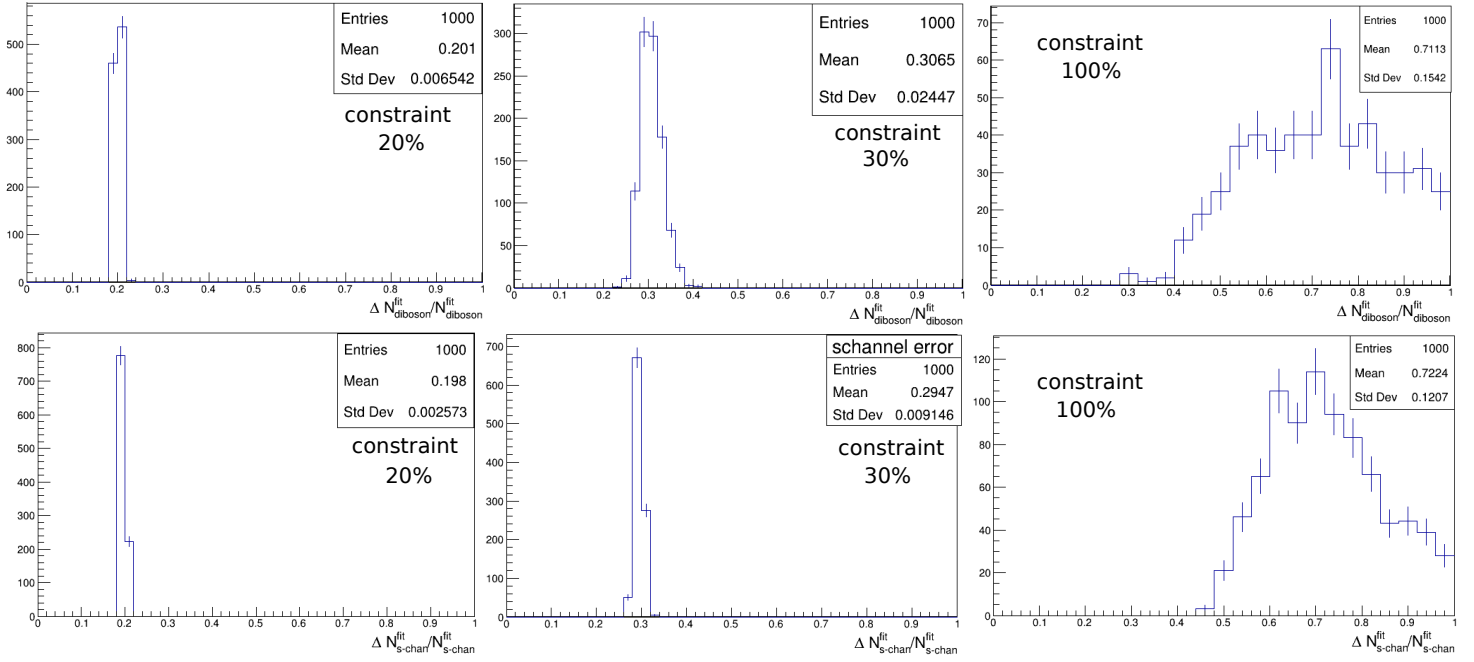


Figure 3.33: Diboson and ST s-channel processes relative errors distributions, from Toy MCs using different values for the Gaussian constraints sigmas.

### 3.8 Measurement of cross section

In this section the maximum likelihood fit for the extraction of the signal cross section scale factor is performed on data. The first step of the strategy that will be followed is to find some estimations for the  $t\bar{t}$  and  $DY$  backgrounds from separate fits on the corresponding control regions, that will be later used as external constraints. Then a simultaneous fit will be performed using the signal region and the  $W + jets$  control region. The reason for not implementing a separate fit for  $W + jets$  background also, is that its control region is not so pure like the  $t\bar{t}$ ,  $DY$  and also contains significant amount of signal events. Therefore a separate fit in this region would not be statistically independent (uncorrelated) from the signal process.

#### 3.8.1 Leptons efficiencies Scale Factors

For the better description of the processes with the simulations, scale factors (SFs) were applied for taking into account the differences of lepton efficiencies between data and MC. The main reason was that they were found significant deviations for the  $t\bar{t}$  distributions at the CR. These scale factors depend on  $p_T$  and  $\eta$  of the lepton and they are for electrons

1. Trigger efficiency  $\epsilon_{trig}$
2. Reconstruction efficiency  $\epsilon_{reco}$



### 3. Identification efficiency $\epsilon_{ID}$

and for muons

1. Trigger efficiency  $\epsilon_{trig}$
2. Identification efficiency  $\epsilon_{ID}$

The estimation of these efficiencies is done using the tag-and-probe technique. In this technique a high purity sample of events with a Z boson decaying into two leptons  $l^+l^-$  is considered ( $e$  or  $\mu$ ). So for example, for the muon trigger efficiencies, the one reconstructed muon which fired the trigger is the tagged lepton. Then the other muon, which is the probe lepton, may or may not be firing the trigger. The efficiency of the muon trigger is then estimated by the ratio of the number of probe leptons that fired the trigger to the total number of probe leptons.

With this technique the estimation of the efficiencies is clear and unbiased, although the efficiencies can be measured only at the phase space in which the process of Z decay to leptons is abundant. This is why there is some  $p_T$  restriction on these efficiencies. For the rest  $p_T$  values scale factors were not applied.

The results for the  $t\bar{t}$  CR for the variables  $m_T^W, p_T^e, |\eta_e|, E_T^{miss}$ , are shown in Figures 3.34-3.37. It is observed that the ratios of simulation to the data are falling by approximately 10% and they also become more flat, which indicates shape corrections from these factors fit better the data, and thus the corrected distributions differ mainly on the scale. For the DY as seen in Figure 3.38 the ratio again becomes more flat and the scale difference between MC and data becomes bigger. These results suggest that after the fits to the corresponding regions the expectation is for  $t\bar{t}$  to get a SF of order 0.9, and DY of order 1.5. The scale factor needed for the DY can be explained by the fact that the MC for this process is of LO precision.

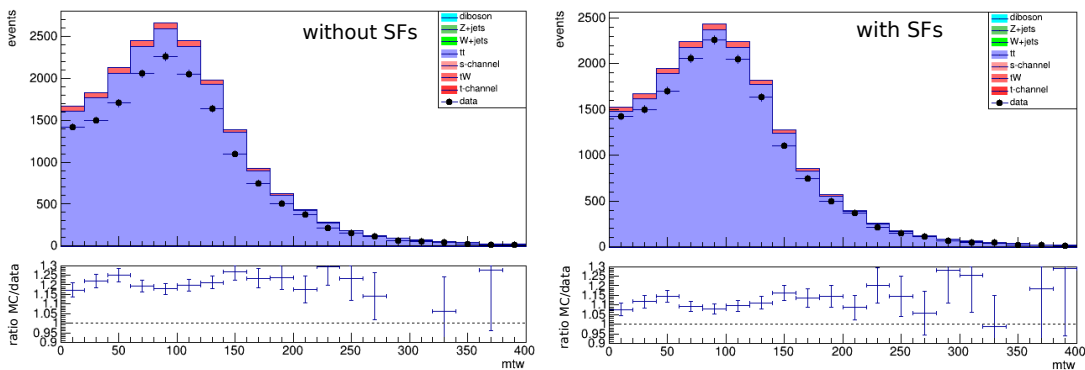


Figure 3.34: Comparison of  $m_T^W$  distributions for the  $t\bar{t}$  CR, before and after the application of lepton efficiencies SFs.

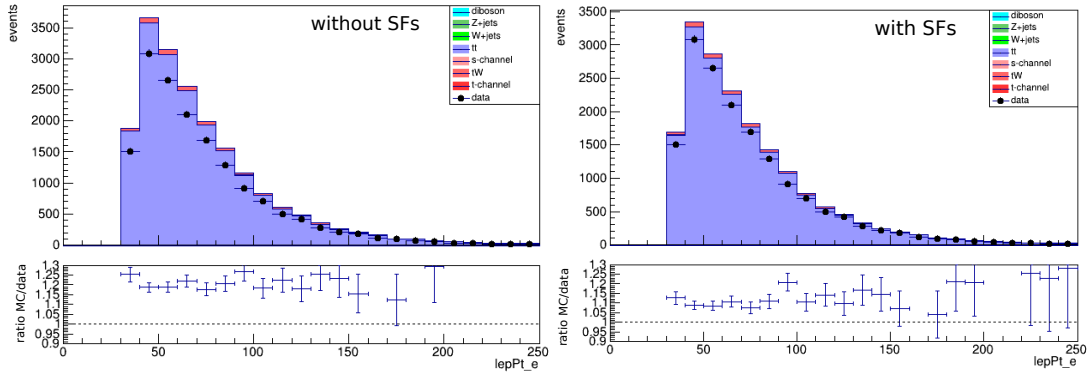


Figure 3.35: Comparison of  $p_T^e$  distributions for the  $t\bar{t}$  CR, before and after the application of lepton efficiencies SFs.

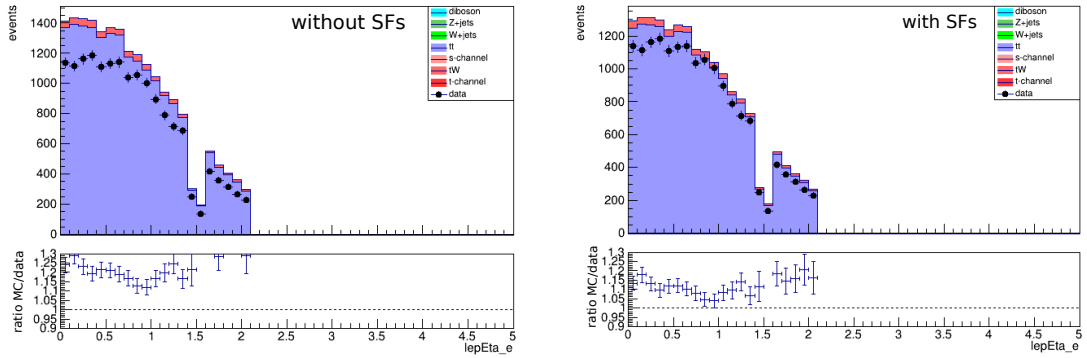


Figure 3.36: Comparison of  $|\eta_e|$  distributions for the  $t\bar{t}$  CR, before and after the application of lepton efficiencies SFs.

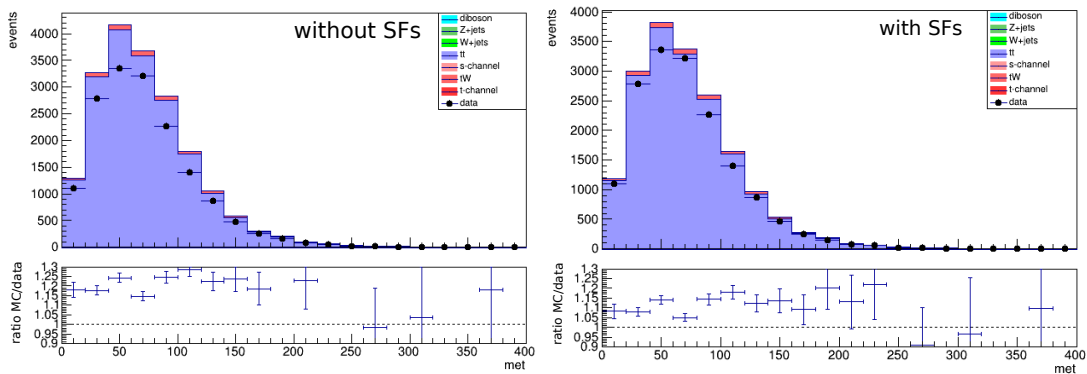


Figure 3.37: Comparison of distributions for the  $t\bar{t}$  CR, before and after the application of lepton efficiencies SFs.

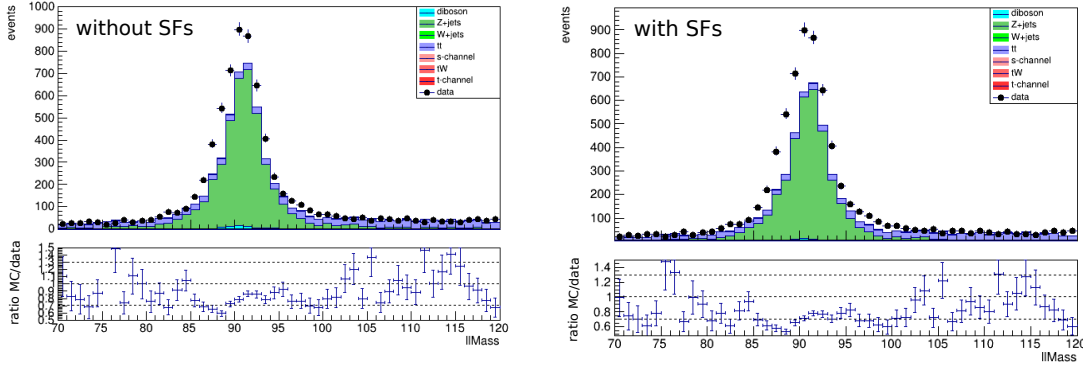


Figure 3.38: Comparison of  $m_{\ell\ell}$  distributions for the Drell Yan CR, before and after the application of lepton efficiencies SFs.

### 3.8.2 Constraining $t\bar{t}$ and $DY$ backgrounds

For the extraction of constraints for the  $t\bar{t}$  and  $DY$  background processes, ML fits on the corresponding CRs were used. The  $t\bar{t}$  fit was performed on  $m_T^W$  variable, using only the templates of  $t\bar{t}$  and  $tW$ , with the later having a tight Gaussian constraint of 10%. The reason for this is that the ST  $tW$  has very similar distributions with the  $t\bar{t}$ ,<sup>11</sup> and since this background is small relatively to  $t\bar{t}$ , the true deviations at its yield would not be statistically important. The fitted templates are shown in Figure 3.39 and its results for the parameters are

$r_{t\bar{t}}$	$0.9 \pm 0.01$
$r_{tW}$	$1.0 \pm 0.1$

The constraint of  $r_{t\bar{t}}$  that will be used from now on, is a Gaussian with mean value 0.9 and standard deviation 0.05. This conservative approach is better than using the error of the fit in order to take into account factors that did not get involved in the fit, like the real  $tW$  difference mentioned before, shape corrections etc.

For the  $DY$  process scale factor estimation the fit is performed on the invariant mass  $m_{\ell\ell}$  of two electrons, which is very discriminating since only for  $DY$  it has a peak at  $M_Z = 90.2 GeV$ . For the  $t\bar{t}$  background the Gaussian constraint mentioned was imposed, and for the small contributions of  $tW$  and Diboson they were set 10% errors. The results of the fit are

$r_{DY}$	$1.484 \pm 0.022$
$r_{t\bar{t}}$	$0.93 \pm 0.03$
$r_{tW}$	$1.0 \pm 0.1$
$r_{Diboson}$	$1.0 \pm 0.1$

<sup>11</sup>Even with use of MVA methods the fits were still no sensitive enough to discriminate the two backgrounds importantly

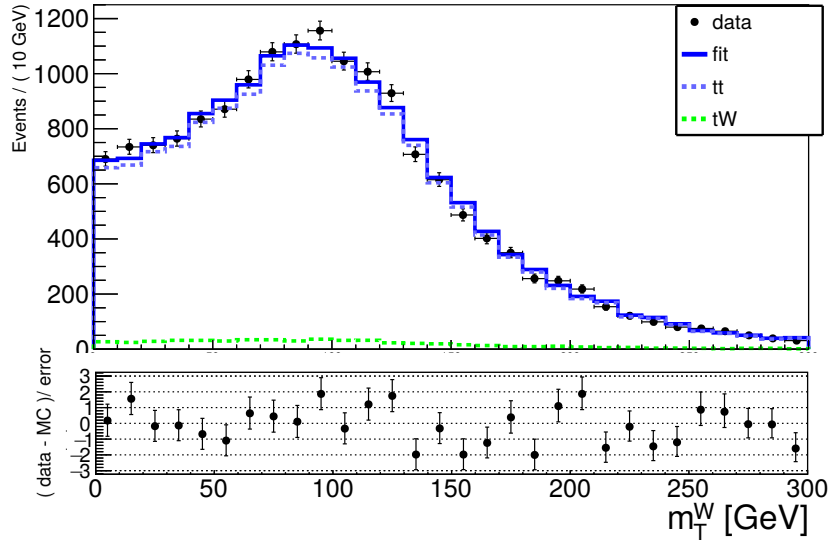


Figure 3.39: Fit on  $m_T^W$  in 2J2T 2 leptons opposite flavor and sign region.

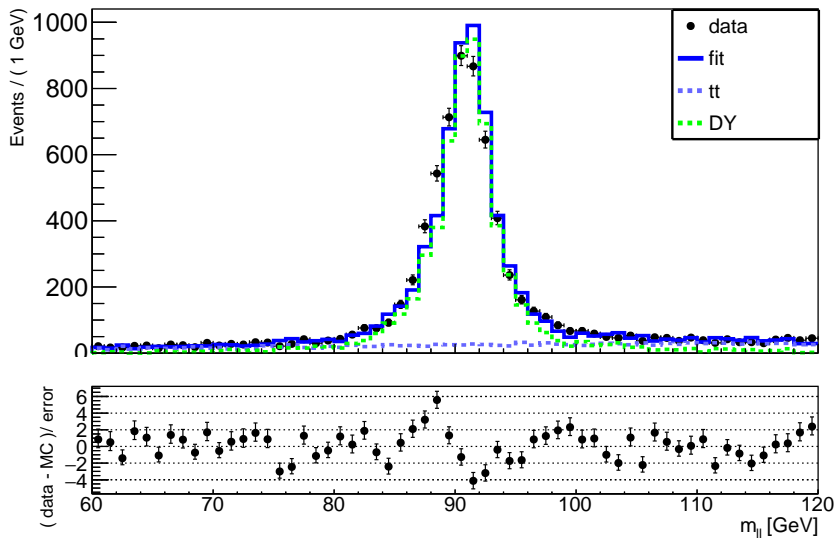


Figure 3.40: Fit on  $m_{ll}$  in 2J1T  $e^+e^-$  region.

The estimated scale factor for the DY is approximately at 1.5 as it was expected. The yield of  $t\bar{t}$  is compatible with the constraint imposed, and the two small backgrounds are practically set by their constraints. From this result, the  $DY$  constraint for  $r_{DY}$  which will be used at the next fitting procedures will be a Gaussian with mean value 1.48 and standard deviation 0.03.

### 3.8.3 Fit with electrons

For the signal extraction a ML fit will be performed simultaneously, for electrons only, on the signal region 2J1T 1lepton and the control region 1J1T 1lepton of W+Jets for this background to be constrained. Regarding the sensitivity issues for the small backgrounds namely the diboson and ST s-channel, the former will not be considered in the fit while the later will be absorbed in the ST  $tW$  as one “rest ST” process. Also this “rest ST” which is almost pure  $tW$  will share a common scale factor with the  $t\bar{t}$ , because they have almost identical shapes. The QCD Multijet backgrounds for each region, they will be set different parameters for the fit in each region. Thus the templates that will be fitted are

$$\begin{aligned}
2J1T(x) = & f_{t-ch} N_{t-ch}^{2J1T(nom)} PDF_{t-ch}^{2J1T}(x) \\
& + f_{t\bar{t}+restST} \left( 0.9 \times N_{t\bar{t}}^{2J1T(nom)} + N_{restST}^{2J1T(nom)} \right) PDF_{t\bar{t}+restST}^{2J1T}(x) \\
& + f_{W+jets} N_{W+jets}^{2J1T(nom)} PDF_{W+jets}^{2J1T}(x) + f_{DY} 1.48 \times N_{DY}^{2J1T(nom)} PDF_{DY}^{2J1T}(x) \\
& + f_{QCD}^{2J1T} N_{QCD}^{2J1T(nom)} PDF_{QCD}^{2J1T}(x)
\end{aligned}$$

and

$$\begin{aligned}
1J1T(x) = & f_{t-ch} N_{t-ch}^{1J1T(nom)} PDF_{t-ch}^{1J1T}(x) \\
& + f_{t\bar{t}+restST} \left( 0.9 \times N_{t\bar{t}}^{1J1T(nom)} + N_{restST}^{1J1T(nom)} \right) PDF_{t\bar{t}+restST}^{1J1T}(x) \\
& + f_{W+jets} N_{W+jets}^{1J1T(nom)} PDF_{W+jets}^{1J1T}(x) + f_{DY} 1.48 \times N_{DY}^{1J1T(nom)} PDF_{DY}^{1J1T}(x) \\
& + f_{QCD}^{1J1T} N_{QCD}^{1J1T(nom)} PDF_{QCD}^{1J1T}(x)
\end{aligned}$$

where  $x$  is the variable of the fit  $m_T^{top}$ , and the estimated fit parameters are the factors  $f$ . On some of these factors Gaussian constraints are imposed. Specifically for  $f_{W+jets}$  a Gaussian(1.,0.3), for  $f_{t\bar{t}+restST}$  Gaussian(1.,0.05) and for  $f_{DY}$  the Gaussian(1.,0.03). For the  $t\bar{t} + restST$  the same error as stated for the  $t\bar{t}$  was set since the  $restST$  process has much smaller yield, and a bigger error for this process would be covered.

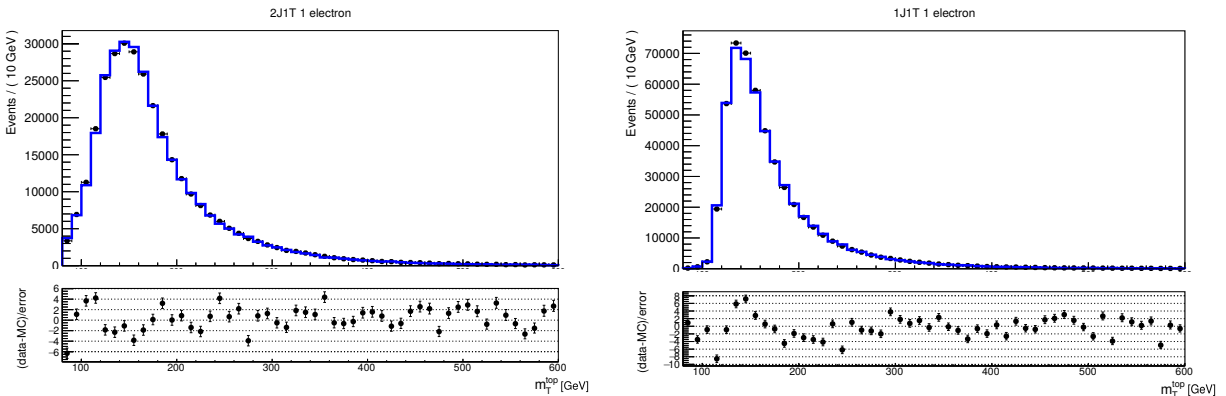


Figure 3.41: Simultaneous fit in 2J1T+1 electron and 1J1T+1 electron regions.

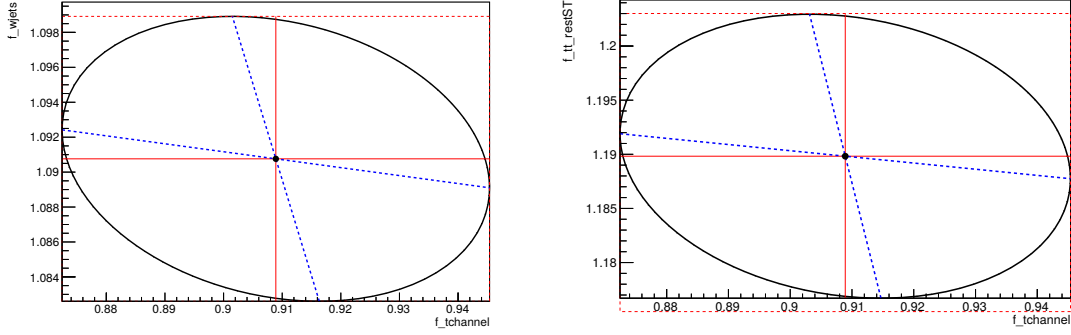


Figure 3.42: Correlations of signal scale factor  $f_{tchannel}$ , with fit parameters  $f_{W+jets}$  and  $f_{\bar{t}\bar{t}+restST}$  from simultaneous fit in 2J1T+1 electron and 1J1T+1 electron regions.

The fit of the two regions is shown in Figure 3.41, and the correlation plots of signal with the two dominant backgrounds in Figure 3.42. The  $r$ 's results are shown in Table 3.4.

The correlation coefficients from the covariance matrix are  $-0.12$  for  $\bar{t}\bar{t} + restST$  and  $-0.16$  for  $W + jets$ , so there are small anti-correlations. These small anti-correlations are a little different from the Toy MC procedure about the  $m_T^{top}$  variable, which predicted small correlations, because of

1. The simultaneous fit has a different likelihood since it contains the 1J1T region, and the more tight Gaussian external constraints.
2. The  $restST$  is absorbed into the  $\bar{t}\bar{t}$ .

$r_{t-ch.}$	$0.91 \pm 0.04$
$r_{W+jets}$	$1.091 \pm 0.008$
$r_{DY}$	$1.37 \pm 0.012$
$r_{\bar{t}\bar{t}}$	$1.071 \pm 0.012$
$r_{restST}$	$1.190 \pm 0.013$
$r_{QCD}^{2j1t}$	$0.738 \pm 0.016$
$r_{QCD}^{1j1t}$	$0.84 \pm 0.01$

Table 3.4: Results for the scale factors  $r_i$  of the different processes, from simultaneous fit in 2J1T+1 electron and 1J1T+1 electron regions.

Although it is assumed that the transfer from one fit to the other does not ruin the unbiased character of the procedure.<sup>12</sup> The argument is that the effect of the differences is only to constrain better the backgrounds. To justify this argument the results of a fit in signal region only, according to the exact Toy MC characteristics which will be called

<sup>12</sup>To validate this assumption completely a new set of Toy MC would be needed for the simultaneous fit, but its beyond the scope of this thesis.

“loose constraints fit”, and one again in signal region having the exact characteristics of the simultaneous fit as constraints which will be called “tight constraints fit”. For the loose constraints fit the results are shown in Figure 3.43, and for the tight constraints in Figure 3.44. Indeed for the first one there are small correlations as predicted by the Toy MC. For the second, the correlation of signal  $t\bar{t} + restST$  is bigger than this with  $t\bar{t}$  in loose constraints fit, due to the change of shape when including the rest of ST processes in  $t\bar{t}$  template, while the  $W + jets$  remains very few correlated.

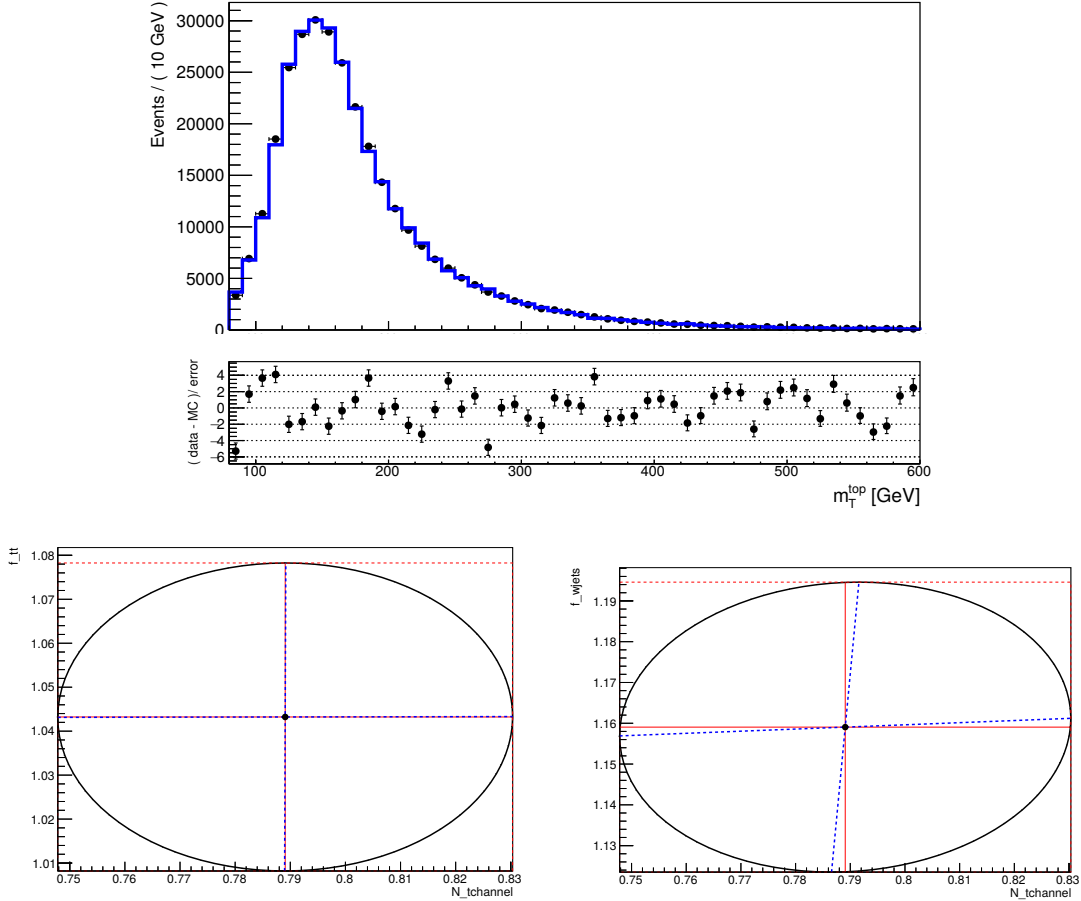


Figure 3.43: Fit and correlations for the “loose constraints fit” in signal region.

In Figure 3.45 they are shown the results for the  $r$ 's estimated from the simultaneous fit, and the “loose”/“tight” constraints fits. It is observed that the external constraints on the backgrounds give more precise estimations for the yields. This especially for the DY process is crucial since the “loose” fit was not sensitive on it. Also for the  $restST$  process, which also had big error in the “loose” case, adding it to the  $t\bar{t}$  background created a much better estimation. These are advantages of both “tight” fit in signal region and the simultaneous fit. But the second in addition due to high statistics of  $W+jets$  in 1J1T region, it highly constraints this background. In addition it constraints better the QCD Multijet since it is anticorrelated with the  $W+jets$  due to their similar shapes. Therefore

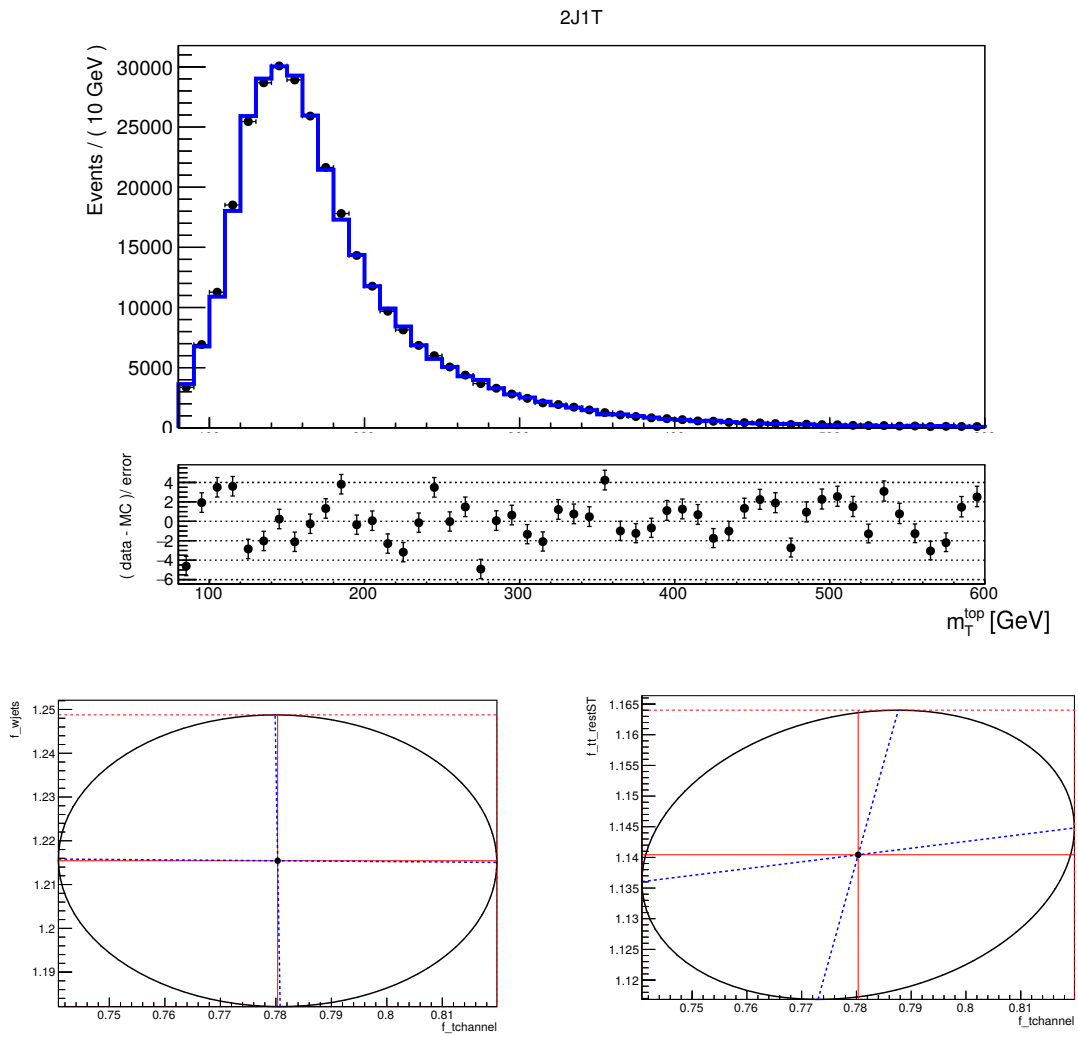


Figure 3.44: Fit and correlations for the “tight constraints fit” in signal region.



having better constrained backgrounds leads to better estimation of signal. Although the result of the higher statistics of the simultaneous fit, improves poorly the precision of the signal, since the main contribution of it is contained in the signal region. In order to improve this precision the dataset of muons would be required. The results of all the fits, for  $r_{t\text{-chan.}}$  are statistically compatible.

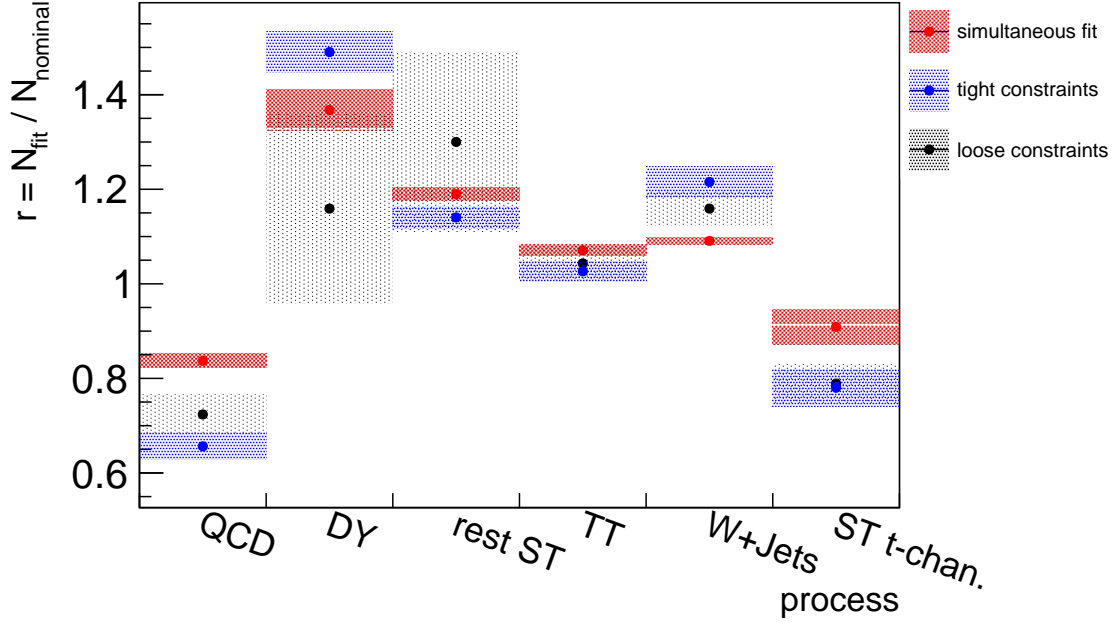


Figure 3.45: Comparison of  $r$ 's estimated using the two fits of signal region, and the simultaneous fit in signal and W+jets control region.

### 3.8.4 Fit with muons

The muons data sample will not be used for the extraction of the resulting scale factor, because of QCD Multijet background have shown indications of bad modeling. The reason behind this is not clarified and it is to be further studied in the future. Therefore the role of this dataset will be validating the results of the electrons data. In order to do that, the scale factors for all the backgrounds except QCD, estimated from the simultaneous fit on electrons channel, will be applied as external constraints. Then the fit will give an estimation for  $r_{t\text{-chan.}}$  which will be compared with the electrons result. The fit and the correlations of signal with main backgrounds are shown in Figure 3.46. The  $r_{t\text{-chan.}}$  has high anti-correlation with  $t\bar{t}$  and some low anti-correlation with W+jets. The result of the fit is

$$r_{t\text{-chan.}} = 0.837 \pm 0.027$$

which is compatible with the  $0.91 \pm 0.04$  found before. The smaller error for the muons comes from the higher statistics.

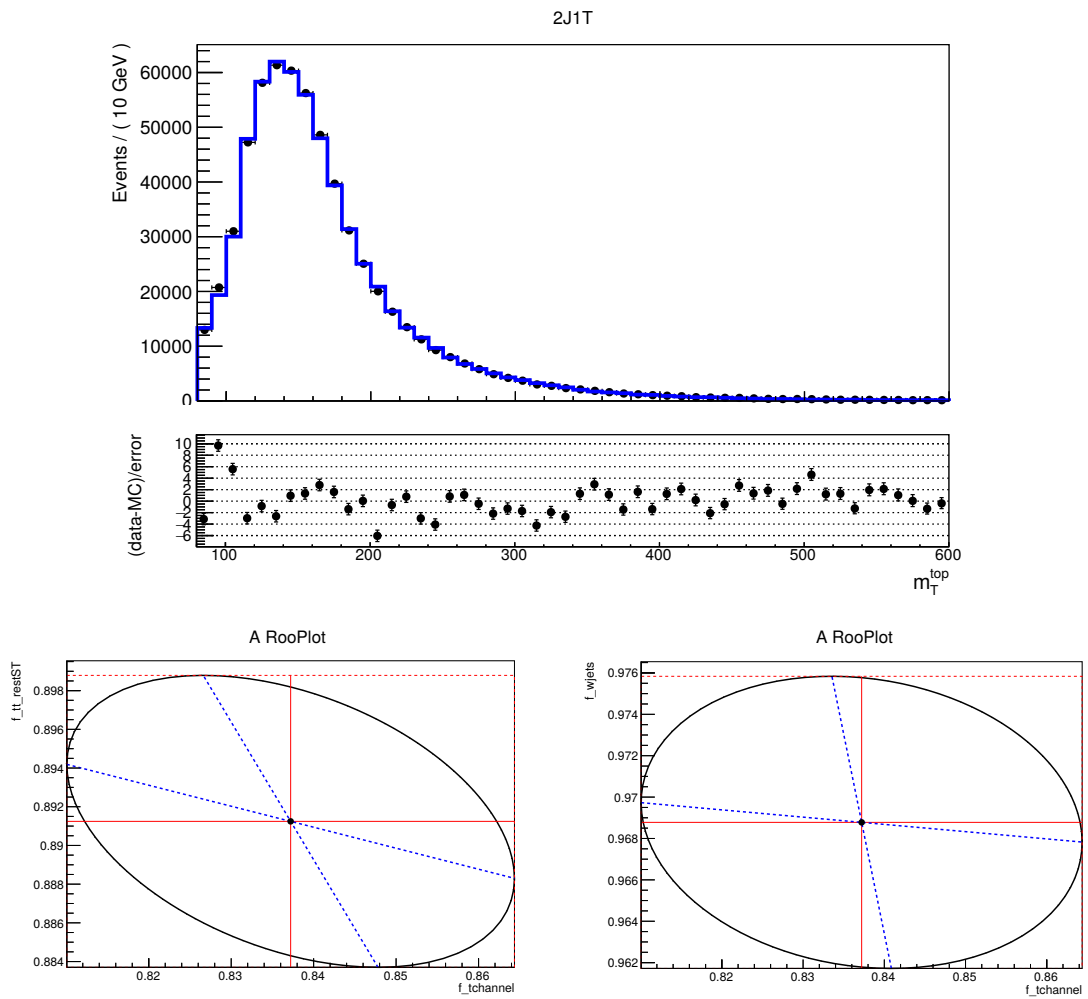


Figure 3.46: Fit in signal region for muons and the correlations plots of signal with two dominant backgrounds.

### 3.9 Results

The extracted scale factor for the signal from the fit is

$$r_{t\text{-}chan.} = 0.909 \pm 0.036(stat.) \quad (3.16)$$

which corresponds to cross section

$$\sigma_{t\text{-}ch.} = 197.2 \pm 7.8(stat.)pb \quad (3.17)$$

The real error is bigger since in this result they are included no systematic errors. A recent measurement on the same luminosity and energy, in CMS experiment found

$$\sigma_{t\text{-}ch} = 219.0 \pm 1.5(stat.) \pm 33.0(syst)pb \quad (3.18)$$

and therefore the results are in statistical agreement. The statistics from this result is better since they are used more regions, and also a ML fit is performed for both muons and electrons. Also, the systematics is the main contribution in the error of this analysis.

The distributions for some variables before and after applying the scale factors found from the fit are shown in Figures 3.48 - 3.58.

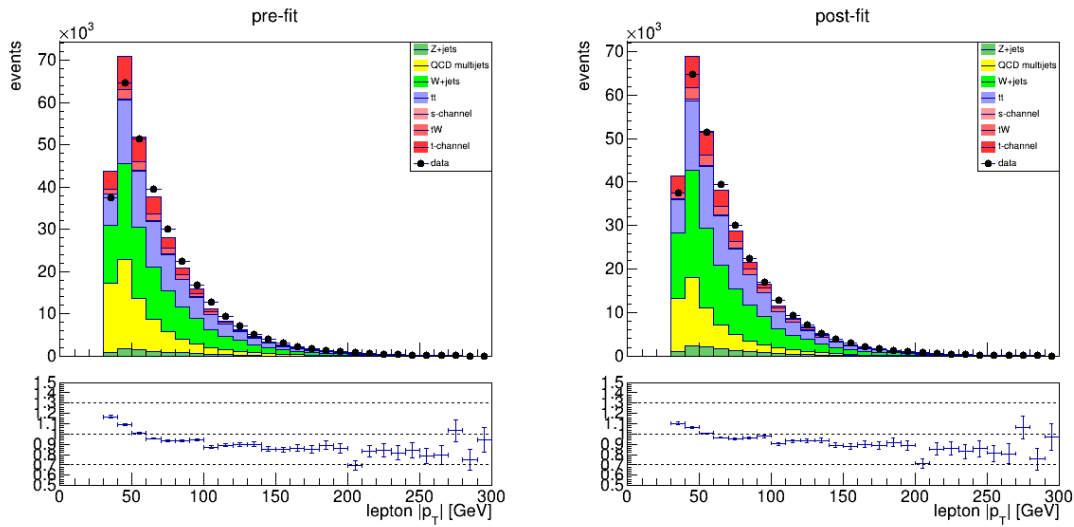


Figure 3.47: Pre-fit and post-fit stacked distributions of processes compared to data for the electron  $p_T$ .

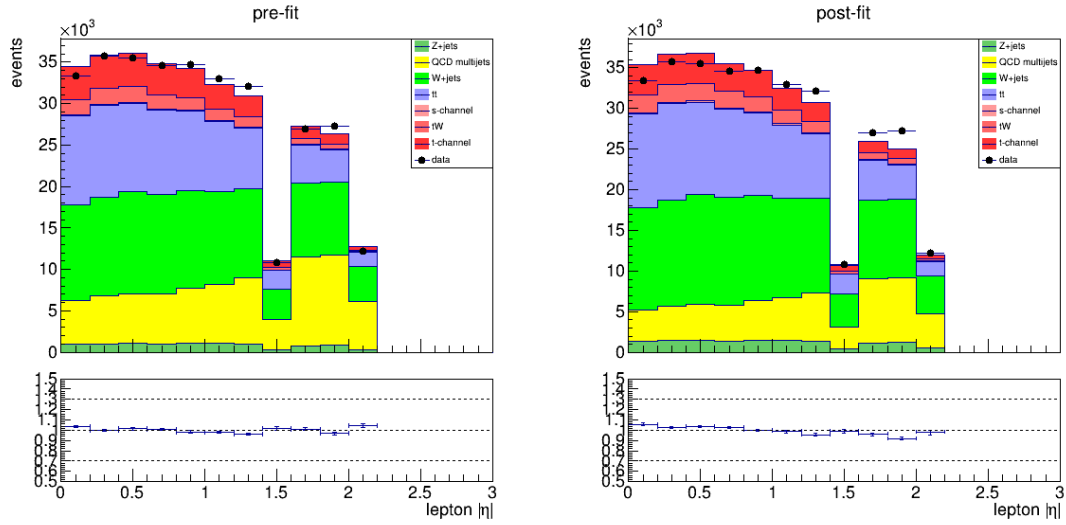


Figure 3.48: Pre-fit and post-fit stacked distributions of processes compared to data for the electron  $|\eta|$ .

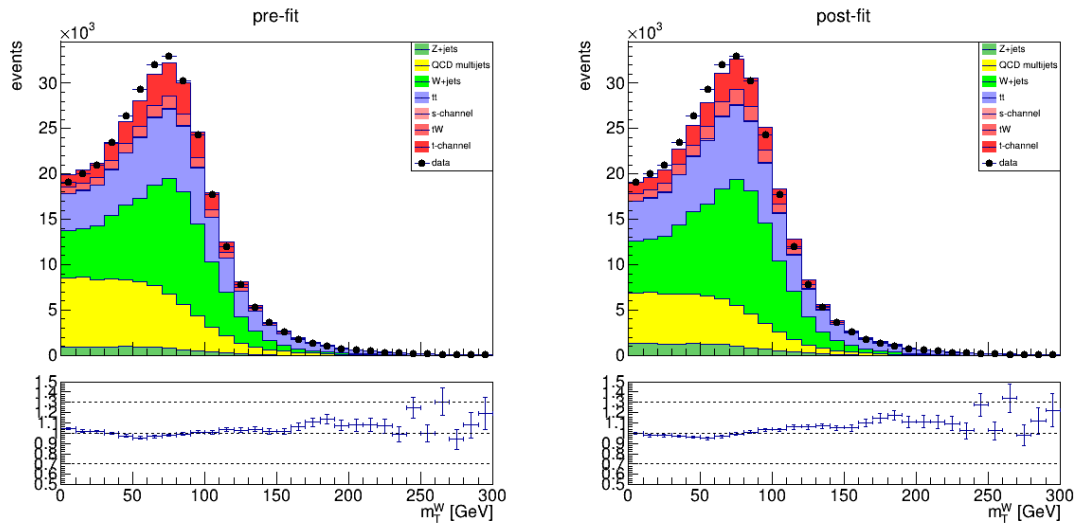


Figure 3.49: Pre-fit and post-fit stacked distributions of processes compared to data for the  $m_T^W$ .

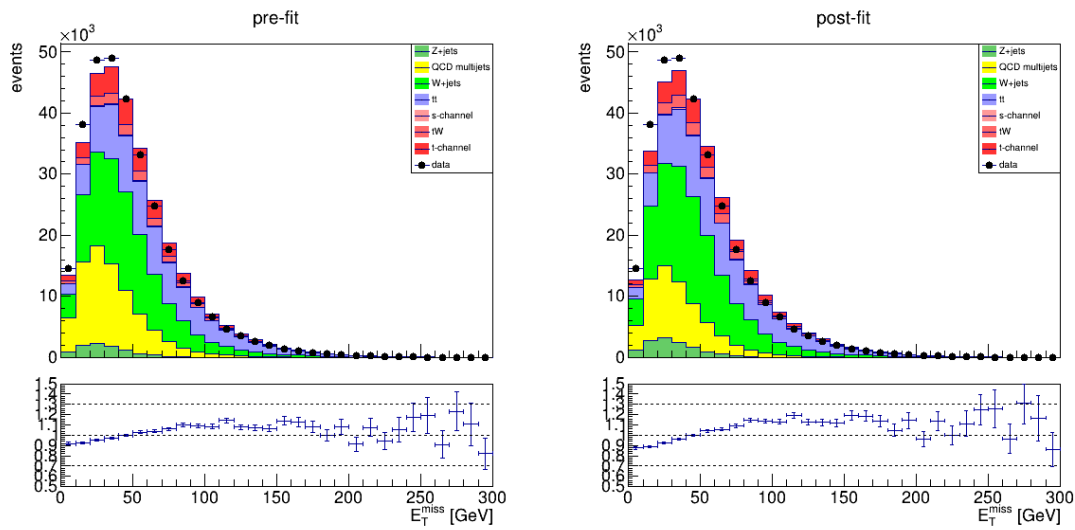


Figure 3.50: Pre-fit and post-fit stacked distributions of processes compared to data for the  $E_T^{miss}$ .

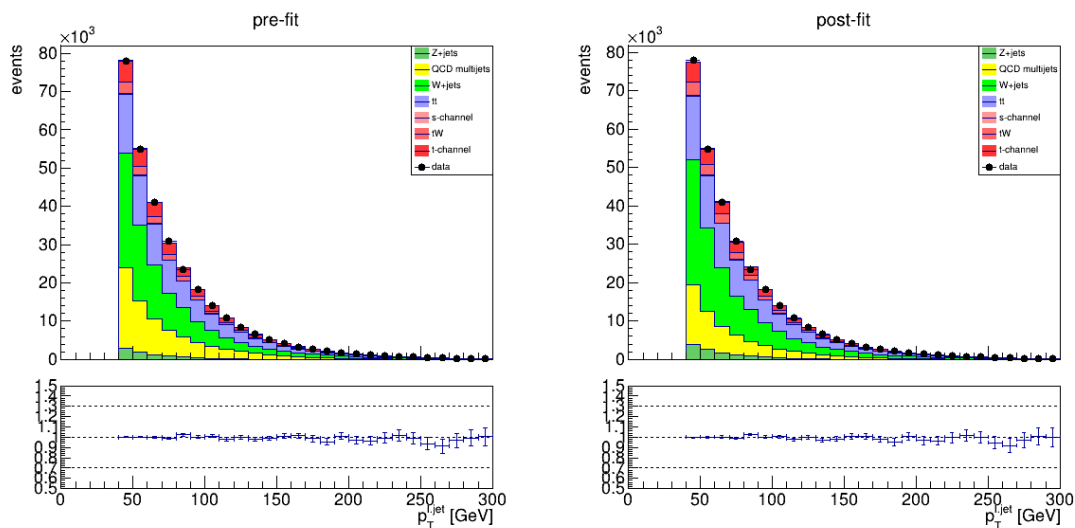


Figure 3.51: Pre-fit and post-fit stacked distributions of processes compared to data for the light jet  $p_T$ .

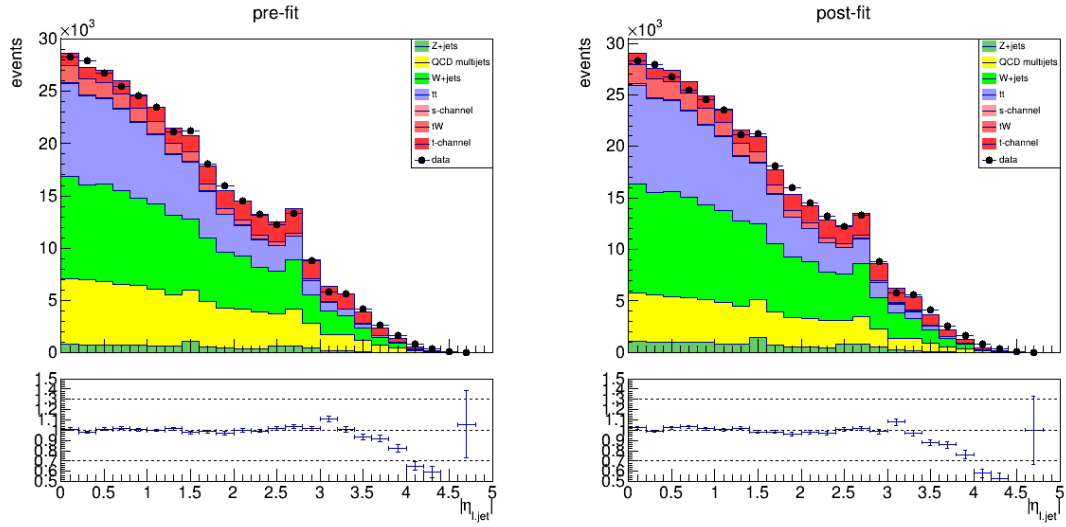


Figure 3.52: Pre-fit and post-fit stacked distributions of processes compared to data for the light jet  $|\eta|$ .

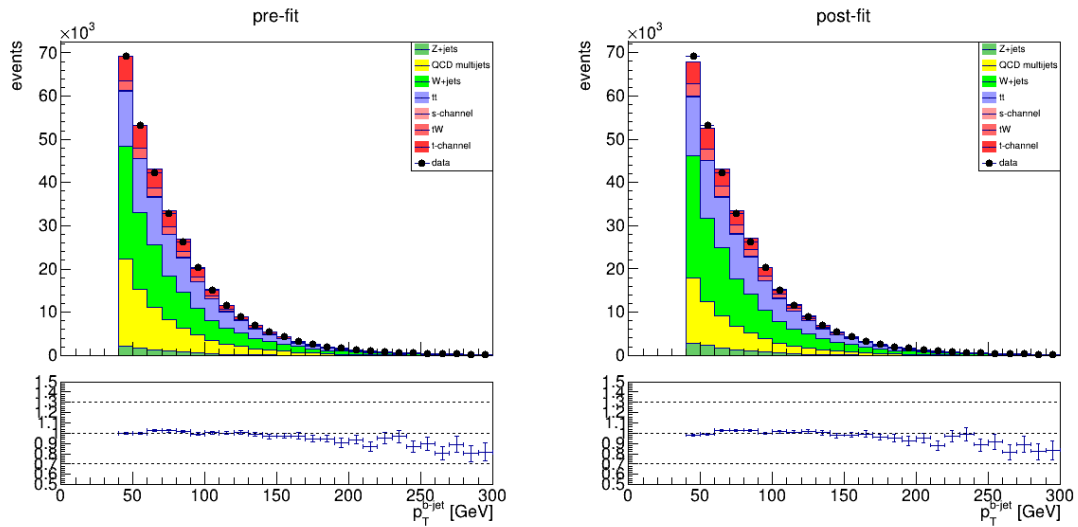


Figure 3.53: Pre-fit and post-fit stacked distributions of processes compared to data for the b-tagged jet  $p_T$ .

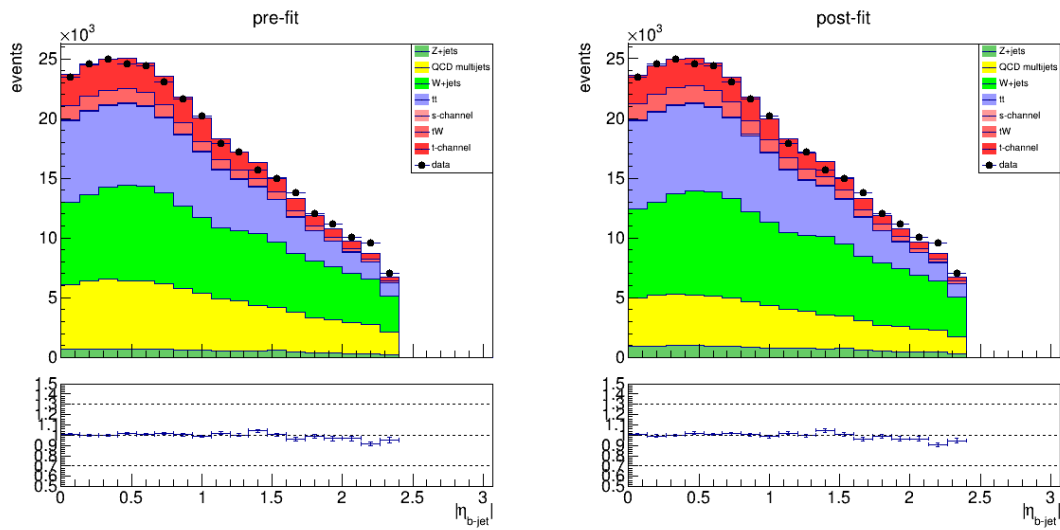


Figure 3.54: Pre-fit and post-fit stacked distributions of processes compared to data for the b-tagged jet  $|\eta|$ .

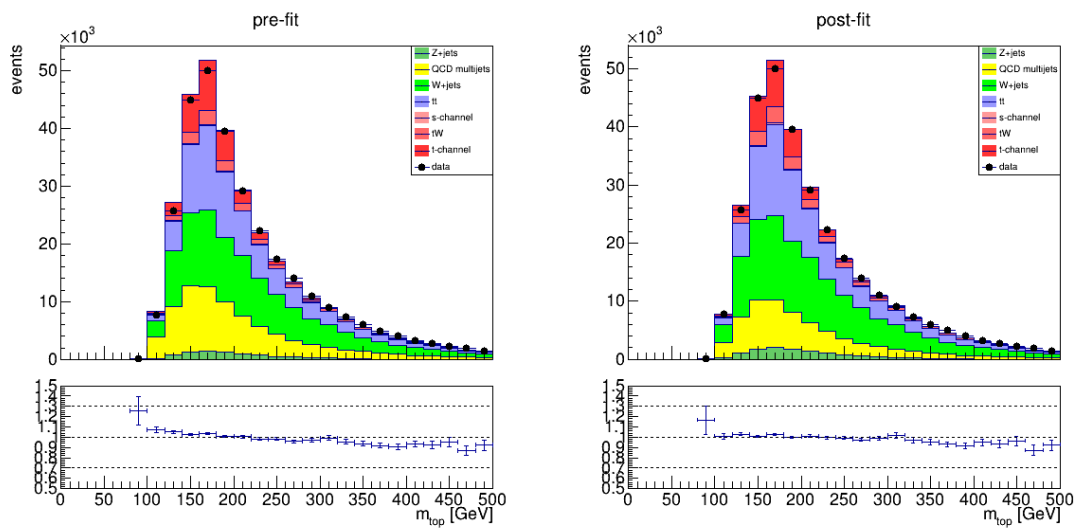


Figure 3.55: Pre-fit and post-fit stacked distributions of processes compared to data for the reconstructed top mass  $m_{top}$ .

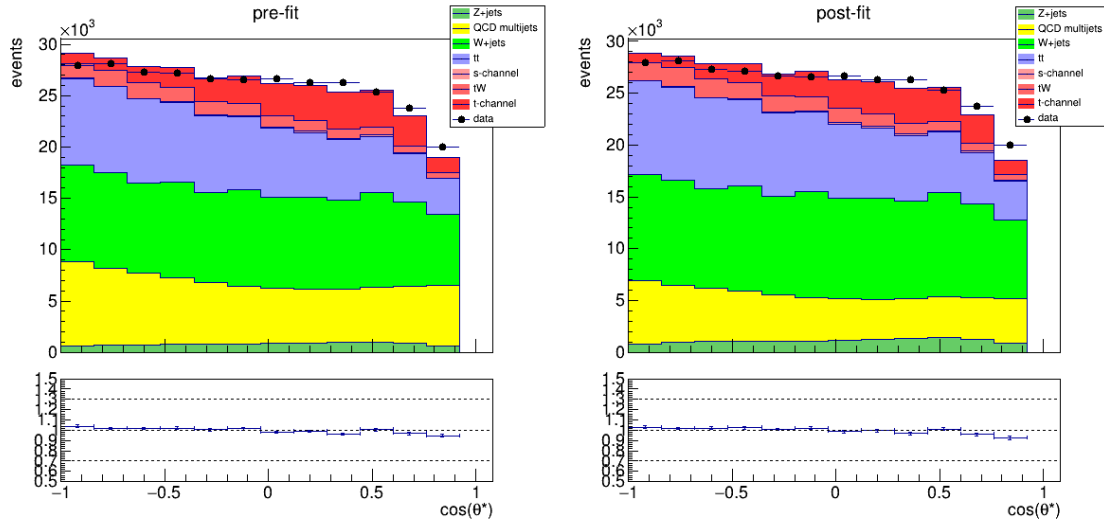


Figure 3.56: Pre-fit and post-fit stacked distributions of processes compared to data for the variable  $\cos(\theta^*)$ .

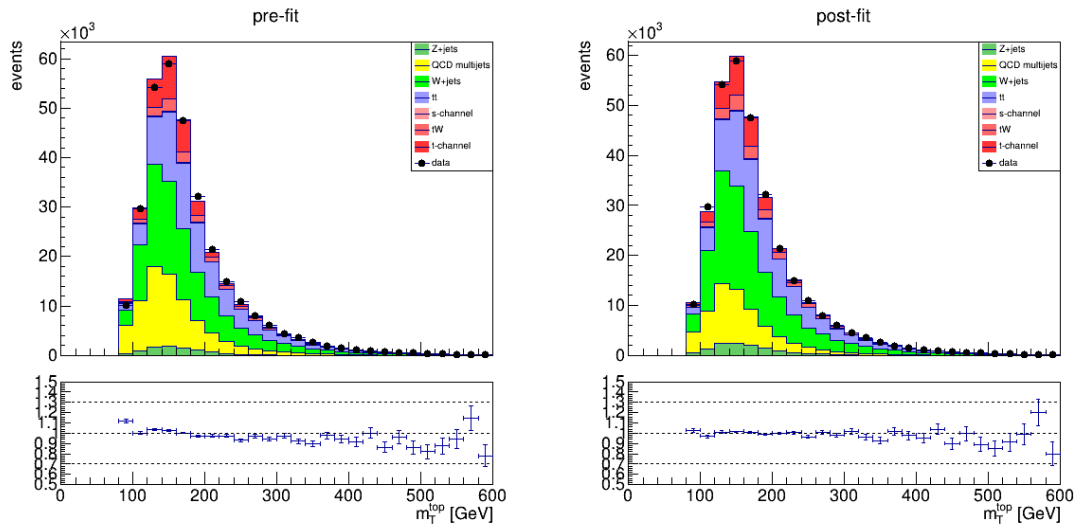


Figure 3.57: Pre-fit and post-fit stacked distributions of processes compared to data for the  $m_T^{\text{top}}$ .



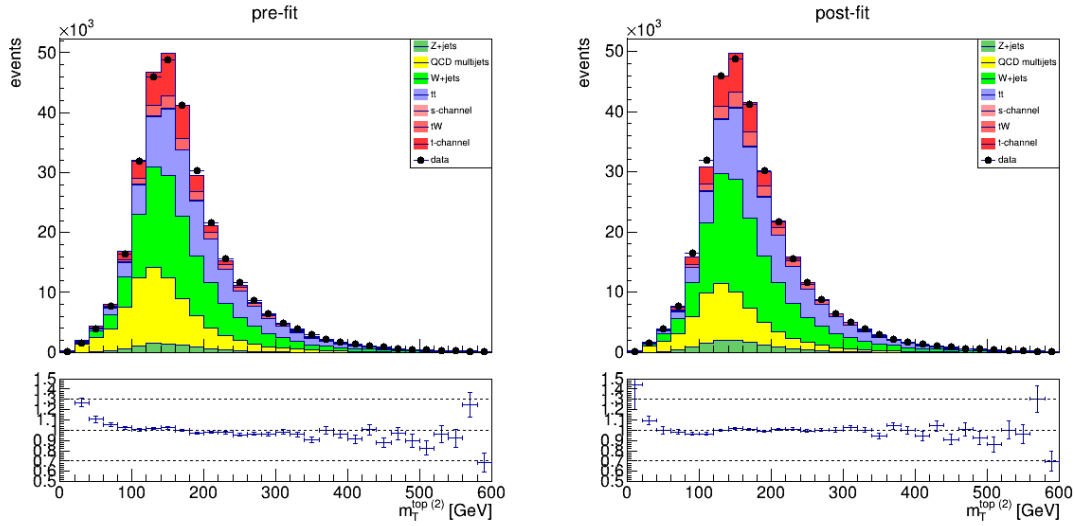


Figure 3.58: Pre-fit and post-fit stacked distributions of processes compared to data for the  $m_T^{top(2)}$ .

### 3.10 Conclusions

In this thesis the scale factor for the cross section of the single top t-channel process was measured. The result is compatible with the standard model prediction as well a measurement from a recent analysis. No systematic uncertainties were included in the result. The statistics would be significantly improved if the muons dataset was used for the estimation. For this a better modeling for QCD Multijet background is needed which is left for future study. For this analysis they were used the tools of maximum likelihood fitting, and MVA methods. The importance of validating the characteristics of the fit through Toy MC was shown, since it provides a framework that the choice of the way to extract the parameter of interest is unbiased and quantitatively checked to be well behaved.

# Appendix A

## Maximum Likelihood fit

### A.1 Definition of method

The maximum likelihood (ML) technique, refers to the problem of estimating the values of the parameters of a model  $\theta_j$ ,  $j = 1, \dots, m$  given a dataset of  $N$  measurements  $x_1, \dots, x_N$ . This is called the fit procedure. Each measurement in general is a vector with  $n$  components (variables) and the fit is therefore  $n$ -dimensional. For the method the likelihood function is defined, which uses the probability density of observing  $x$  for a given set of parameters  $f(x; \theta)$ . For  $N$  independent measurements of the same experiment this is

$$L(\{x_i\}; \theta) = \prod_{i=1}^N f(x_i^1, \dots, x_i^n; \theta_1, \dots, \theta_m) \quad (\text{A.1})$$

The estimator of maximum likelihood finds the estimated parameters  $\hat{\theta}$  by maximizing this quantity

$$\left. \frac{\partial L(x; \theta)}{\partial \theta_i} \right|_{\theta = \hat{\theta}} = 0 \quad \forall \quad \theta_i \quad (\text{A.2})$$

or equivalently minimizing the  $-\log L(x; \theta)$ .

The estimate  $\theta$  is dependent of the given dataset, and therefore has statistical fluctuations. So the estimation must be accompanied by an error.

One method is approximating the likelihood function as a Gaussian, or equivalently the  $-\log L$  as parabolic ( $n$  dimensional paraboloid in general). For large number of measurements this is a good approximation. In this case the covariance matrix  $C$  is defined as the inverse of the Hessian matrix, which is the matrix of second derivatives.

$$C_{ij} = \left[ - \frac{\partial^2 \log L(\{x_i\}, \theta_1, \dots, \theta_m)}{\partial \theta_i \partial \theta_j} \right]_{\theta_{i,j} = \hat{\theta}_{i,j}}^{-1} \quad (\text{A.3})$$

This method yields always symmetric errors for the estimations  $\hat{\theta}_i \pm \sigma_i$ , where  $\sigma_i = \sqrt{C_{ii}}$ . Also the correlation coefficients of the different parameters are found by the non-diagonal elements of  $C$ , namely  $\rho_{ij} = \sqrt{C_{ij}} / \sigma_i \sigma_j$ .

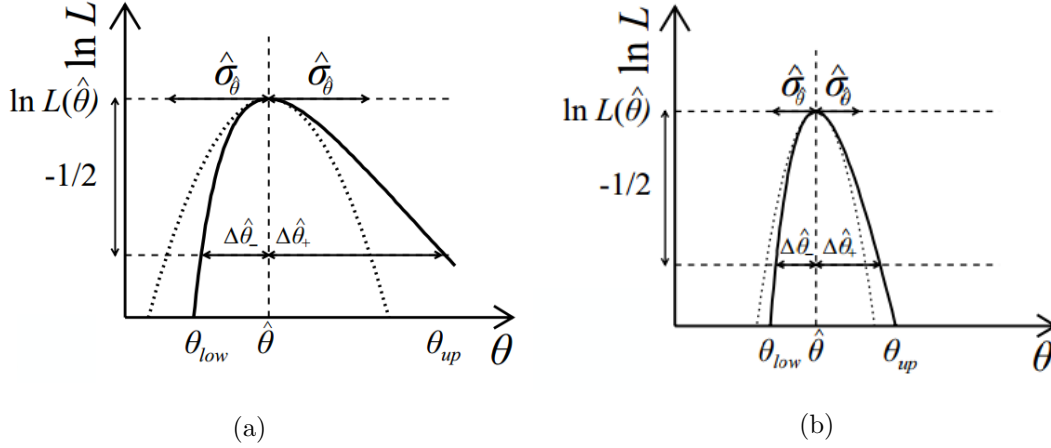


Figure A.1: Cases of  $-\log L$  shapes with (a) Parabolic approximation is not valid (b) Valid parabolic approximation.

In general, the shape of Likelihood is not Gaussian. A way to estimate the uncertainties then, which is numerically equivalent for the Gaussian case, is to use the contour defined by values  $\theta'$  such that

$$-\log L(\theta') = -\log L_{max} + \frac{1}{2} \quad (\text{A.4})$$

where  $L_{max} = L(\{x_i\}; \hat{\theta})$ , and it is called 1-sigma contour. The projection of this curve on a parameter  $\theta_i$  defines the error interval of the estimation  $[\theta_i - \Delta\theta_-, \theta_i + \Delta\theta_+]$  which is in general asymmetric. This definition is exact, and in the case that the error of Gaussian approximation is not consistent with the  $\Delta\theta_{\pm}$ , the approximation is invalid. In Figure A.1 they are shown two cases, one which the parabolic approximation is valid and one which is not.

## A.2 Extended Likelihood

In high energy physics it is usually the case where the fit has as parameters the yield of signal events  $s$  which follow a probability distribution function  $f_s(x; \theta_s)$ , and the yield of the background events  $b$  which follow  $f_b(x; \theta_b)$ . The  $\theta_{s,b}$  are other parameters defining the shapes of the distributions. There could be more than one backgrounds but it is considered this case without any loss of generality. Then the total events number  $N$  is depending on  $s, b$  since  $N = s + b$ , and follows a Poisson distribution with mean  $s + b$

$$N \sim \text{Poisson}(N; s + b) = \frac{(s + b)^N}{N!} e^{-(s+b)} \quad (\text{A.5})$$

Thus the Likelihood is written as

$$L(\{x_i\}; s, b, \theta_s, \theta_b) = \frac{(s + b)^N e^{-(s+b)}}{N!} \prod_{i=1}^N [s f_s(x_i; \theta_s) + b f_b(x_i; \theta_b)] \quad (\text{A.6})$$

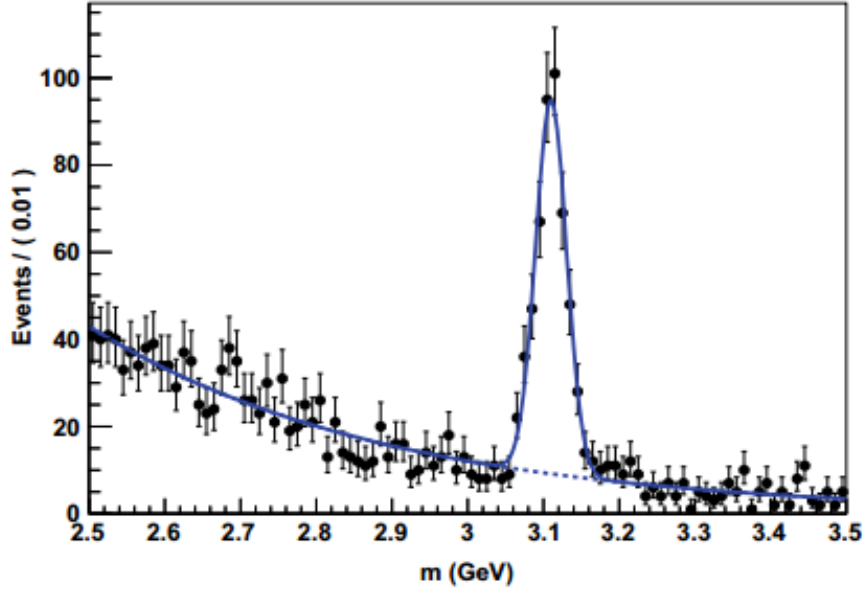


Figure A.2: Example of an extended ML fit on an invariant mass distribution. Data are shown as a binned histogram but the individual values of the random variable  $m$  are used in the likelihood function.

For example one may have a peak in an invariant mass distribution that is fitted with a Gaussian distribution for signal, and an exponential background as in Figure A.2. In this example the parameters of the fit are the yields  $s, b$  of the two processes, and the shape parameters mean  $\mu$  and  $\sigma$  of Gaussian and  $\lambda$  decay parameter for exponential.

This type of likelihood functions where the number of data  $N$  depends on parameters of distribution is called “extended likelihood function”.

### A.3 Binned Maximum Likelihood Fit

In the previous sections the method is defined in a way where each event enters separately into the likelihood. For this reason it is called “unbinned” ML fit. This is statistically optimal to do, but its CPU consumption increases linearly with number of events  $N$ . For this reason high energy physics the most popular implementation is the “binned” ML, where the events are grouped in bins of the observale(s)  $x$  and the counted event numbers in the bins are used in the likelihood function. In addition when the theoretical model template cannot be a continuous PDF but it comes from a binned histogram of generated events with Monte Carlo. If one wants to fit this template then binned ML fit is the only way to do so.

Here it will be considered there is an one dimensional fit i.e. one fitted variable  $x$ , although the same hold for multidimensional fits. The variable  $x$  is partitioned into  $N$  bins. Then if  $n$  is the total number of independent events and  $n_i$  the number of events in bin  $i$ , then each bin is a Poissonian experiment with observed number of events  $n_i$ . In case of

a continuous PDF  $f(x; \theta)$ , in order to find the expected number  $\mu_i$  of events in each bin one integrates the corresponding PDF and multiplies by  $n$

$$\mu_i = \mu_i(\theta) = \int_{\text{bin } i} f(x; \theta_1, \dots, \theta_m) dx \quad (\text{A.7})$$

Therefore each bin has a distribution  $\text{Poisson}(n_i; \mu_i(\theta))$ . The likelihood function is the written

$$L = \prod_{i=1}^N \text{Poisson}(n_i, \mu_i(\theta)) = \prod_{i=1}^N \frac{\mu_i(\theta)^{n_i}}{n_i!} e^{-\mu_i(\theta)} \quad (\text{A.8})$$

In the case where the number of events per bin is big enough the Poisson distributions can be approximated by Gaussians with mean value  $\mu_i$  and sigma  $\sqrt{\mu_i}$ . Then taking the  $-2 \log L$  and dropping the terms which do not contain fit parameters one gets the  $\chi^2$  minimization method i.e. minimization of

$$\sum_{i=1}^N \frac{(n_i - \mu_i(\theta))^2}{\mu_i(\theta)} \quad (\text{A.9})$$

## A.4 Simultaneous fitting

In a physics analysis it is usually the case where the dataset can be separated into regions, regarding their content in signal or background events. The region which has most of signal events is called “signal region” (SR), while a region which is dominant by some background processes, or even better only one of them, are called “control regions” (CR). Let’s assume for simplicity there are two regions, a signal and a control. The dataset in the signal(control) region will be denoted as  $\{x_{sr}\}$  ( $\{y_{cr}\}$ ) where  $x, y$  can be the same or different variables. Then if  $\theta_c$  are common parameters in the two regions while  $\theta_{sr(cr)}$  refer only to one of them, the likelihood function can be written as

$$L(\{x_{sr}\} \cup \{y_{cr}\}; \theta_c, \theta_{sr}, \theta_{cr}) = L_{sr}(\{x_{sr}\}; \theta_c, \theta_{sr}) L_{cr}(\{y_{cr}\}; \theta_c, \theta_{cr}) \quad (\text{A.10})$$

where the  $L_{sr(cr)}$  are the corresponding regions likelihoods. Therefore for the common parameters more events will be used and a more precise estimation will be extracted. Such a fit where two or more datasets are fitted with common parameters is called “simultaneous”. A usual example is that a CR is defined which is almost pure to a background process, and the simultaneous fit constraints its yield. Also there may be regions which contain signal events and a combined simultaneous fit will give a better estimation. Simultaneous fitting is a method that allows exploiting better the available data, and is therefore popular in High Energy Physics analyses.

# Appendix B

## Multivariate Analysis

### B.1 Definitions

Multivariate analysis(MVA) techniques focus on the problems of “classification” and “regression”. Here the focus is on the first one. Classification in general is about separating classes from each other. In case of HEP it is usually defined as having two classes the signal and the background, which is a binary classification problem. In other cases it is multiclass problem.

In order to solve that problem of separating events, one may use variables defined in the phase space of the data. These variables are called input “features”  $x$ . In MVA for every feature vector  $x$  it is assigned a scalar value  $y$ , which is the classifier method output. The result is that there is a probability distribution for a classifier output  $y$  for signal events  $p(y|x \in Sig)$  and one for background events  $p(y|x \in Bkg)$ . If these two are separate enough then  $y$  can be used as discriminant variable between the two classes. The picture of a classifier resulting distributions is shown in Figure B.1.

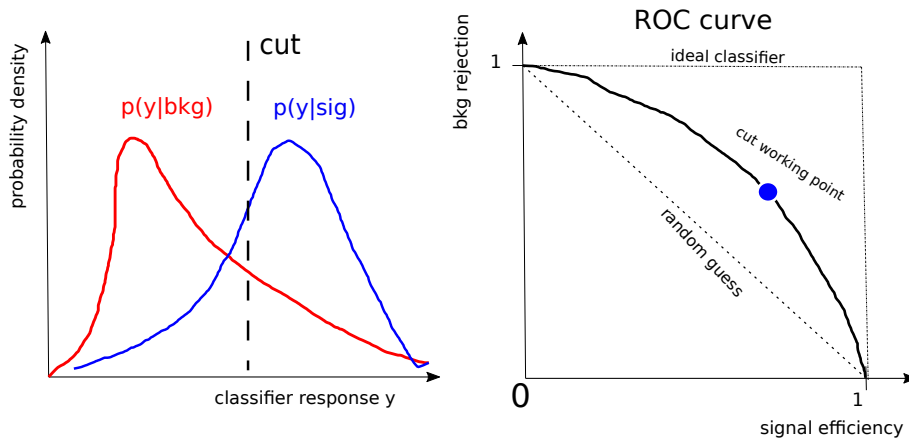


Figure B.1: The response probability distributions of a classifier for the signal and background classes (left) and the corresponding ROC curve, with the cut working point.

The quality of the class separability offered can be seen from the Receiver Operator

Curver (ROC). This curve is produced by considering several cuts on the output  $y$  and calculating for each the signal efficiency and the background rejection. In the case where the two distributions are uniform and the guess is completely random this curve is the  $y = -x$  curve. Ideally one wants to reject all background and have only signal. But in realistic cases something in the middle happens. These all cases are shown in Figure B.1. For a specific cut the point corresponding on the ROC curve is called working point.

Therefore as seen the goal of the classifier is to get the two distributions  $p(y|sig), p(y|bkg)$  in an optimal way for these two to be as separate as possible. For each classifier this procedure is unique, although the general concept is that it is given a sample of data which are labeled, meaning it is known to the classifier which events are signal and which background. Then for these the classifier is trained, through an optimization procedure to separate the classes.

Something that one should be careful when performing the training of a classifier is the “overtraining”. This means that the classifier should not be trained so much that its output is using the statistical fluctuations of the training dataset. Then for a different sample the classifier will give statistically different distributions. For this reason it is always used a test dataset on which the classifier method is applied and it is checked the output distributions for it are compatible with those of the training dataset.

There are many types of classifiers linear and non-linear, like Fisher, Neural Networks, Boosted Decision Trees etc. In this thesis the Boosted Decision Trees(BDT) is used therefore there is some reference here on the function of this classifier.

## B.2 Boosted Decision Trees (BDT)

A decision tree is a binary classifier, which is based on the logic of separating the phase space by applying multiple optimal cuts iteratively.

This means that one starts with the full phase space of input variables  $x_i$  and at first step an optimal cut is applied to the best discriminating variable available. Then two subspaces are created. For each of them the same procedure is performed, and so on. This procedure is shown schematically in Figure B.2. When a split is happening it is created a “node”. In case a split does not happen and the subspace stays as it is this is a “leaf”. In the end all subspaces correspond to leaves.

Now, in order to choose which split to perform, and if a split should be done or stop, a quantity is defined that shows how much is the purity of classes in the resulting spaces. If it is higher than the initial subspace, a split is performed. If not then the subspace corresponds to a leaf. Such quantities are the Gini index  $I = p(1 - p)$  where  $p$  is the probability for selecting one process, and the cross entropy  $I = -p \log p - (1 - p) \log(1 - p)$ . The split which creates the bigger  $\Delta I > 0$  is performed.

The decision tree classifier is good because it is very intuitive and it does not get influenced by input variables which low separability power. The main drawback of this method is that it is highly sensitive to statistical fluctuations of the training sample. For

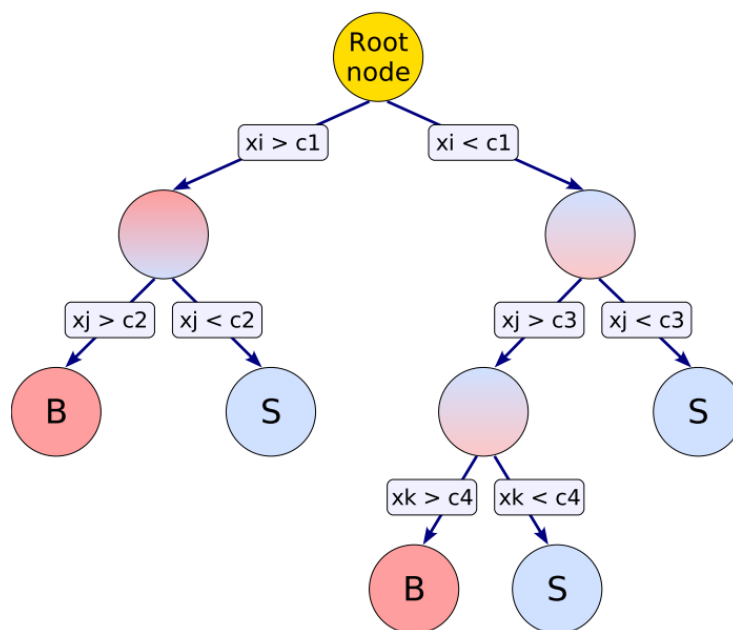


Figure B.2: Schematic representation of a decision tree.

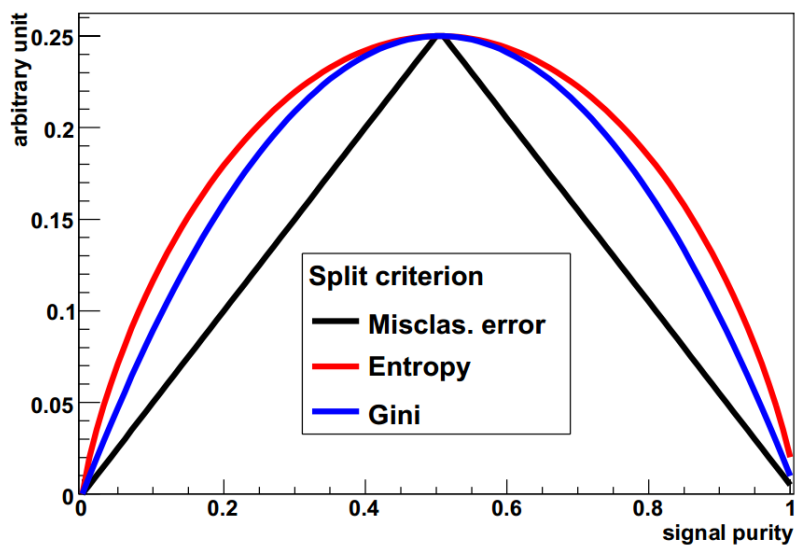


Figure B.3: Different measures of signal purity in a subspace.



this reason a boosting method is used. The idea of boosting is that instead of using one classifier, namely a decision tree, it will be used the “weighted average” of an ensemble of such classifiers, which each time get a weight regarding how good results they give. Boosting stabilizes the response of the decision trees with respect to the fluctuations of the training sample and in addition it enhances the performance compared to using a single tree. Boosting methods can be used in general for weak classifiers.

Formally for  $M$  weak classifiers the function  $F(x)$  which is the final classifier response is written as

$$F(x; \{\beta_m, a_m\}) = \sum_{m=0}^M \beta_m f(x; a_m) \quad (\text{B.1})$$

which is the weighted sum of parametrised base functions  $f(x; a_m)$  which correspond to the individual weak classifiers. The weights  $\beta_m$  and the parameters  $a_m$  of the classifiers are adjusted by minimizing the deviation of the predicted  $F(x)$  and the true value  $y$  in the training sample. This minimization corresponds to the minimization of a loss function  $L(F, y)$ . The loss function fully determines the boosting procedure. A popular choice is the  $L(F, y) = e^{-F(x)y}$  which can be minimized by implementing the so called AdaBoost algorithm. In this thesis the boosting procedure that was used is that of GradientBoost. This method uses the binomial log-likelihood loss

$$L(F, y) = \log \left( 1 + e^{-2F(x)y} \right) \quad (\text{B.2})$$

This loss function cannot be minimized by an algorithm like the AdaBoost case but the gradient(steepest) decent minimization needs to be used. In general it works better for trees with small depth 2-4, and for its robustness it should be used small learning rate (which is the step of gradient decent), through a parameter called “shrinkage”. For small shrinkage 0.1-0.3 many trees should be grown.

# Bibliography

- [1] CMS collaboration, Measurement of the inclusive cross section of single top quark production in the t-channel at  $\sqrt{s} = 13TeV$ .  
CDS server. CMS PAS TOP-16-003.
- [2] Ulrich Husemann, Top-Quark Physics: Status and Prospects.  
arXiv:1704.01356v2
- [3] Particle Data Group 2018, Top quark review
- [4] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss  
TMVA Toolkit for Multivariate Data Analysis with ROOT Users Guide
- [5] W. Verkerke, D. Kirkby, RooFit Users Manual v2.91
- [6] Luca Lista, Statistical Methods for Data Analysis in Particle Physics, 2nd ed. Springer
- [7] Ilya Narsky, Frank C. Porter, Statistical Analysis Techniques in Particle Physics *Fits, Density Estimation and Supervised Learning*, Wiley
- [8] Konstantinos Koutroumpas, Sergios Theodoridis, Pattern Recognition  
Academic Press 2008
- [9] Olaf Behnke, Kevin Kroeninger, Gregory Schott, and Thomas Schoerner-Sadenius,  
Data Analysis in High Energy Physics “A Practical Guide to Statistical Methods”.  
Wiley 2013
- [10] C. Grupen, B. Shwartz , Particle Detectors  
Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology
- [11] CMS Collaboration, CMS Physics
- [12] CMS Collaboration, Detector Drawings, CDS
- [13] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker” 2014
- [14] CMS Collaboration, The CMS Experiment at the CERN LHC

- 
- [15] C. Biino, “The CMS Electromagnetic Calorimeter: overview, lessons learned, 2015 during Run 1 and future projections”
  - [16] CMS Collaboration, “Precision crystal calorimetry in LHC Run II with the CMS ECAL, 2017”
  - [17] CMS Collaboration, “Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter, 2010”
  - [18] CMS Collaboration, “CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter”
  - [19] F. Gasparini, “The CMS muon detector: From the first thoughts to the final design”
  - [20] CMS Collaboration, “The performance of the CMS muon detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV at the LHC”
  - [21] CMS Collaboration, “CMS reconstruction improvement for the muon tracking by the RPC chambers”
  - [22] Matthew D. Schwartz, “Quantum Field Theory and the Standard Model”, Cambridge University Press 2014.
  - [23] Matthew D. Schwartz, “Collider Physics”, TASI
  - [24] R.K. Ellis, W.J. Stirling, B.R. Webber, “QCD and Collider Physics”, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology
  - [25] A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, “Parton distributions for the LHC”, arXiv:0901.0002v3 [hep-ph] 7 Jul 2009
  - [26] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”.

