

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΦΥΣΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΠΕΙΡΑΜΑΤΙΚΗΣ ΦΥΣΙΚΗΣ ΥΨΗΛΩΝ
ΕΝΕΡΓΕΙΩΝ



Μελέτη Διαφορικής Ενεργού Διατομής της
παραγωγής ζεύγους top-antitop quark στον
ανιχνευτή CMS

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΤΟΥ
ΓΕΩΡΓΙΟΥ ΜΠΑΚΑ

Επιβλέπων: Κωνσταντίνος Κουσουρής
Επίκουρος Καθηγητής ΕΜΠ

ΑΘΗΝΑ, Ιανουάριος 2018

Abstract

This thesis aims to study the $t\bar{t}$ quark production in proton-proton collisions recorded with the CMS detector at CERN at 13 TeV centre-of-mass energy. The differential cross sections, as a function of various observables, are measured and unfolded to the particle and parton levels. Lastly, this thesis will investigate the impact of the theoretical uncertainties on the final differential cross section results.

Περίληψη

Η διπλωματική εργασία έχει σκοπό την μελέτη της παραγωγής ζεύγους $t\bar{t}$ κουαρκ από την σύγκρουση πρωτονίων που ανιχνεύθηκαν στο πείραμα CMS στο CERN, σε 13 TeV ενέργεια στο κέντρο μάζας. Πιο συγκεκριμένα, η ενεργός διατομή εκφράζεται διαφορικά, συναρτήσει διαφόρων κινηματικών μεταβλητών, διορθωμένη ως προς την πεπερασμένη διακριτική ικανότητα του ανιχνευτή. Τα πειραματικά αποτελέσματα συγκρίνονται με τις προβλέψεις θεωρητικών μοντέλων.

Περιεχόμενα

1	Επισκόπηση Φυσικής Στοιχειωδών Σωματιδίων	3
1.1	Το καθιερωμένο πρότυπο	3
1.2	Κβαντική Ηλεκτροδυναμική- Quantum Electrodynamics QED	5
1.3	Κβαντική Χρωμοδυναμική - Quantum Chromodynamics QCD	6
1.4	Ασθενής Αλληλεπίδραση	6
1.5	Ηλεκτρασθενής Αλληλεπίδραση	7
1.6	Ο μηχανισμός Higgs και η αυθόρμητη ρήξη συμμετρίας	8
2	Top Quark	9
2.1	Εισαγωγή	9
2.2	Παραγωγή Top Κουαρκ	9
2.3	Παραγωγή $t\bar{t}$	10
3	Ο Ανιχνευτής Compact Muon Solenoid	14
3.1	The Large Hadron Collider	14
3.1.1	Το επιταχυντικό σύστημα του LHC	14
3.1.2	Ανιχνευτικές Διατάξεις του LHC	16
3.2	Compact Muon Solenoid	18
3.2.1	Ανιχνευτές Τροχιάς - Tracker Detector	20
3.2.2	Crystal Electromagnetic Calorimeter (ECAL)	23
3.2.3	Hadron Calorimeter (HCAL)	25
3.2.4	Ο Μαγνήτης	28
3.2.5	Ανιχνευτές Μιονίων	28
3.2.6	Σύστημα Σκανδαλισμού (Trigger)	30
4	Ανακατασκευή Φυσικών Αντικειμένων	32
4.1	Ανακατασκευή Αδρονικών Πιδάκων	32
4.1.1	Ανακατασκευή Jet	34
4.1.2	Ο αλγόριθμος anti- k_T	34
4.1.3	Ενεργειακή Βαθμονόμηση Jet	36
4.1.4	Ενεργειακή Διόρθωση Jet	37

4.1.5	Αναγνωρίζοντας Jet από b κουαρκ	38
4.2	Ο Αλγόριθμος Ροής Σωματιδίων- Particle Flow Algorithm	41
4.3	Η Φυσική των συγκρούσεων Πρωτονίων-Πρωτονίων	42
4.4	Boosted Jets	43
4.5	Ταυτοποίηση B-Jets	44
5	Analysis	46
5.1	Trigger Efficiency	46
5.1.1	Trigger	46
5.1.2	Trigger Efficiency	47
5.1.3	Mathematics of the Efficiency	47
5.1.4	Trigger Efficiency Analysis	48
5.2	Unfolding	50
5.2.1	Inverse Problems	51
5.2.2	Discretisation and Linear Solution	52
5.2.3	Unfolding Using the Least Squares Method	54
5.2.4	Norm Regularisation Method	54
5.2.5	Implementation of the method	57
5.2.6	Regularisation Conditions	57
5.2.7	TUnfold Class	58
5.3	Monte Carlo Datasets	58
5.3.1	General	59
5.3.2	Variables	59
5.3.3	Response Matrices-Selection	60
5.3.4	Closure Tests	65
5.3.5	Correction Factors	71
5.4	Signal Extraction -Fiducial Cross Section	74
5.5	Uncertainties	75
5.5.1	Estimation of Theoretical Uncertainties	76
5.6	Reconstruction of Parton and Particle Quantities	78
5.6.1	Differential Cross Section	78
5.6.2	Extrapolation to Parton Phase Space	79
5.6.3	Extrapolation to Particle Phase Space	81
5.7	Conclusions	86
	Appendices	87
	A Dynamic Unfold Code	88

Κεφάλαιο 1

Επισκόπηση Φυσικής Στοιχειωδών Σωματιδίων

Σκοπός της φυσικής στοιχειωδών σωματιδίων είναι η ερμηνεία του σύμπαντος που παρατηρούμε με βασικά δομικά στοιχεία και θεμελιώδεις αλληλεπιδράσεις μεταξύ των στοιχείων αυτών. Ξεκινώντας από την κβαντομηχανική περίπου στην αρχή του 20ου αιώνα, οι φυσικοί έχουν αναπτύξει μία θεωρία η οποία απαντά ερωτήματα για την εξέλιξη του σύμπαντος. Μέχρι και σήμερα, η υποκείμενη θεωρία, κυρίως δηλαδή το Καθιερωμένο πρότυπο (Standard Model), είναι η επικρατέστερη και ακριβέστερη θεωρία στη Φυσική. Στις επόμενες σελίδες θα περιγραφούν τα βασικά στοιχεία της θεωρίας αυτής. Θα γίνει όμως μία πιο λεπτομερής αναφορά όσον αφορά το στοιχειώδες σωματίδιο-τοπ κουαρκ (t quark), καθώς αυτό παίζει και τον σημαντικότερο ρόλο στην ανάλυση αυτή.

1.1 Το καθιερωμένο πρότυπο

Το επικρατέστερο και πληρέστερο μοντέλο που υπάρχει για την περιγραφή και μελέτη των στοιχειωδών σωματιδίων είναι το καθιερωμένο πρότυπο (Standard Model, SM) [1]. Μέχρι στιγμής το πρότυπο αυτό είναι το μόνο που περιγράφει με ικανοποιητική ακρίβεια όλα τα πειραματικά αποτελέσματα που έχουν παρατηρηθεί. Σύμφωνα με το SM τα στοιχειώδη σωματίδια χωρίζονται σε 2 κατηγορίες όπως φαίνεται και στην εικόνα 1.1. Οι δύο κατηγορίες είναι τα μποζόνια και τα φερμιόνια. Τα φερμιόνια είναι σωματίδια από τα οποία αποτελείται η ύλη και χωρίζονται σε δύο κατηγορίες. Τα λεπτόνια (leptons) και τα κουάρκ (quark). Η κατηγορία των λεπτονίων αποτελείται από το ηλεκτρόνιο, το μιονίο και το ταυ. Τα σωματίδια αυτά μπορούν να βρεθούν ελεύθερα στη φύση. Επιπλέον σε κάθε ένα από τα σωματίδια αυτά, αντιστοιχεί ένα νετρίνο. Δηλαδή υπάρχουν τα νετρίνα ηλεκτρονίου, μιονίου και ταυ. Όσον αφορά τα κουάρκ, είναι 6 και δεν μπορούν να βρεθούν ελεύθερα παρά μόνο σε συνδυασμούς

των 2 ή 3 δημιουργώντας άλλα σωματίδια, τα αδρόνια.

Πέρα όμως από την ύπαρξη των στοιχειωδών αυτών σωματιδίων, το SM περιγράφει και τον τρόπο με τον οποίο τα στοιχειώδη σωματίδια αλληλεπιδρούν. Στη φύση υπάρχουν 4 στοιχειώδης δυνάμεις, οι οποίες διέπουν τις αλληλεπιδράσεις των σωματιδίων. Οι 4 αυτές δυνάμεις είναι η ισχυρή, η ασθενής, η ηλεκτρομαγνητική και η βαρύτητα. Η βαρύτητα ωστόσο είναι πολύ ασθενής σε σχέση με τις υπόλοιπες, με αποτέλεσμα να μην επηρεάζει τα στοιχειώδη σωματίδια και συνεπώς να μην περιλαμβάνεται στο καθιερωμένο πρότυπο. Για κάθε μία από αυτές τις δυνάμεις, το SM προβλέπει την ύπαρξη σωματιδίων (μποζονίων) μέσα από οποία "εκφράζεται" η δύναμη.

Drei Generationen
der Materie (Fermionen)

	I	II	III	
Masse →	2,3 MeV	1,275 GeV	173,07 GeV	125,9 GeV
Ladung →	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0
Spin →	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0
Name →	u up	c charm	t top	H Higgs Boson
				q e/p-Quant
Quarks	d down	s strange	b bottom	g Gluon
	ν_e Elektron- Neutrino	ν_μ Myon- Neutrino	ν_τ Tau- Neutrino	Z⁰ Z Boson
Leptonen	e Elektron	μ Myon	τ Tau	W[±] W Boson
				Eichbosonen

Figure 1.1: Ο πίνακας των στοιχειωδών σωματιδίων σύμφωνα με το καθιερωμένο πρότυπο

Το μοντέλο αυτό προβλέπει και την ύπαρξη ενός ακόμα σωματιδίου, το οποίο ευθύνεται για τη μάζα των υπολοίπων. Το σωματίδιο αυτό είναι το σωματίδιο Higgs και παρατηρήθηκε το 2012 όταν και ανακοινώθηκε από τα πειράματα του LHC η παρατήρηση και ανακάλυψη του.

Παρά τις σημαντικές ιδιότητες του καθιερωμένου προτύπου, αναφέρθηκε και πριν η μη δυνατότητα περιγραφής της βαρύτητας. Γενικά το καθιερωμένο πρότυπο ως μοντέλο είναι ελλιπές καθώς δεν περιλαμβάνει όλες τις αλληλεπιδράσεις που συμβαίνουν στη φύση. Ένας από τους στόχους του LHC είναι να παρατηρηθούν οι φυσικές διεργασίες που δεν περιγράφονται από το καθιερωμένο πρότυπο. Αυτό σημαίνει δηλαδή να γίνει αναζήτηση ενδείξεων μίας νέας φυσικής, πέρα από το καθιερωμένο πρότυπο. Μία από τις πιο υποσχόμενες θεωρίες της οποίας η ύπαρξη θα μελετηθεί και ελεγχθεί από τον LHC είναι αυτή της υπερσυμμετρίας (Supersymmetry).

1.2 Κβαντική Ηλεκτροδυναμική- Quantum Electrodynamics QED

Ως προς την ιστορία του καθιερωμένου προτύπου, η QED είναι ο πρώτος μηχανισμός που περιέγραφε γωνιστές αλληλεπιδράσεις σε μία κβαντική θεωρία πεδίου (quantum field theory)[1]. Το φορτίο της QED ονομάζεται ηλεκτρομαγνητικό φορτίο και συνήθως ανάγεται σε μονάδες του θεμελιώδους φορτίου του ηλεκτρονίου, e . Ο μηχανισμός κατασκευής μίας κβαντικής θεωρίας πεδίου ξεκινά παίρνοντας την Λαγκρανζιανή συνάρτηση για τα φερμιόνια, η οποία περιγράφεται από τους σπίνορες του Dirac και απαιτεί να είναι αναλλοίωτη κάτω από μετασχηματισμούς βαθμίδας. Η κβαντική ηλεκτροδυναμική απαιτεί ότι η Λαγκρανζιανή είναι αναλλοίωτη κάτω από μετασχηματισμούς βαθμίδας μίας $U(1)$ ομάδας συμμετρίας. Η λαγκρανζιανή της εξίσωσης Dirac δίνεται παρακάτω:

$$L = \bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi$$

όπου ο πρώτος όρος είναι ο κινηματικός όρος και ο δεύτερος όρος είναι ο όρος της μάζας. Το ψ είναι ο σπίνορας Dirac και m η μάζα που αντιστοιχεί στον σπίνορα αυτό. Ένα διάνυσμα γ^μ από πίνακες χρησιμοποιείται για να γραφτεί η Λαγκρανζιανή στον χώρο των τετραδιανυσμάτων. Ο όρος m αναφέρεται στις συνιστώσες των τετρανυσμάτων. Ένας μετασχηματισμός $U(1)$ ορίζεται ως:

$$\psi \rightarrow \psi' = \exp^{iq\alpha} \psi$$

όπου η παράμετρος q αντιστοιχίζεται σε μία δύναμη σύζευξης και το α αναπαριστά μία φάση που εξαρτάται από τις χωροχρονικές συντεταγμένες. Ενώ ο όρος της μάζας είναι αναλλοίωτος κάτω από τον παραπάνω μετασχηματισμό, ο όρος $\partial_\mu\psi$ δεν είναι. Για να κατασκευάσουμε μία Λαγκρανζιανή αναλλοίωτη κάτω από μετασχηματισμούς $U(1)$, εισάγουμε ένα νέο διανυσματικό πεδίο $A_\mu(x)$ καθώς και ορίζουμε τη συναλλοίωτη παράγωγο D_μ :

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

Με την εισαγωγή των παραπάνω, η νέα Λαγκρανζιανή γράφεται ως εξής:

$$L = \bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma^\mu\psi A_\mu(x)$$

όπου ο πρώτος όρος είναι ο κινηματικός όρος, ο δεύτερος όρος είναι ο όρος μάζας και ο τρίτος όρος είναι ο όρος αλληλεπίδρασης. Έτσι με την απαίτηση να είναι η Λαγκρανζιανή αναλλοίωτη κάτω από μετασχηματισμούς $U(1)$, προβλέπεται η ύπαρξη ενός διανυσματικού πεδίου A_μ . Αυτό το διανυσματικό πεδίο συνδέεται με το φωτόνιο. Αφού δεν υπάρχει όρος μάζας για το πεδίο του φωτονίου, τότε το φωτόνιο δεν έχει μάζα.

1.3 Κβαντική Χρωμοδυναμική - Quantum Chromodynamics QCD

Η θεωρία πίσω από την QCD εξελίσσεται όπως και στην QED. Το φορτίο στην QCD είναι το χρώμα και υπάρχει σε τρεις μορφές: πράσινο, μπλε και κόκκινο. Κάθε χρώμα έχει και το αντι-χρώμα του και υπάρχει στα αντίστοιχα αντισωματίδια [1]. Συνεπώς πρέπει κανείς να λάβει υπ όψη περιστροφές στον τρισδιάστατο χώρο των χρωμάτων, που περιγράφεται από την SU(3) συμμετρία. Οι γεννήτορες T_α , όπου $\alpha = 1$ έως 8, χρησιμοποιούνται για να εκφράσουν τις περιστροφές αυτές. Όπως και στην QED, ορίζεται μία συναλοιώτη παράγωγος για να κατασκευαστεί μία συμμετρική Λαγκρανζιανή. Η συναλλοιώτη παράγωγος στην QCD είναι:

$$D_\mu = \partial_\mu + ig_s T^\alpha G_\mu^\alpha$$

όπου g_s είναι ο όρος που περιγράφει τη σύζευξη των κουαρκ με την ισχυρή δύναμη. Τα διανυσματικά πεδία G_μ^α συνδέονται με την ύπαρξη των γκλουονίων. Επειδή οι γεννήτορες της SU(3) δεν μετατίθενται μεταξύ τους, επιτρέπονται και αλληλεπιδράσεις μεταξύ δύο γκλουονίων. Η τελική Λαγκρανζιανή δίνεται ως εξής:

$$L = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - g_s \bar{\psi}(\gamma^\mu T^\alpha G_\mu^\alpha)\psi - \frac{1}{4}G_{\mu\nu}^\alpha G^{\alpha\mu\nu}$$

όπου ο πρώτος όρος είναι ο κινηματικός και όρος μάζας, ο δεύτερος όρος είναι η αλληλεπίδραση μεταξύ κουαρκ και γκλουονίων και ο τρίτος όρος ο κινηματικός όρος και ο όρος αλληλεπίδρασης γκλουονίου-γκλουονίου.

Οι αλληλεπιδράσεις γκλουονίου-γκλουονίου οδηγούν σε μία σημαντική επίδραση στην QCD. Η πυκνότητα ενέργειας μεταξύ δύο χρωματικών φορτίων, αυξάνεται όσο αυξάνεται και η απόστασή τους. Συνεπώς, η ενέργεια αυξάνεται και η ενέργεια που αποθηκεύεται στο πεδίο που τα δύο φορτία σχηματίζουν. Σε κάποιο σημείο η ενέργεια είναι αρκετά υψηλή έτσι ώστε να δημιουργήσει νέα coloured charged σωματίδια. Γι' αυτό το λόγο, τα κουαρκ δεν μπορούν να βρισκονται ελεύθερα αλλά μόνο σε ουδέτερες (χρωματικά) δεσμευμένες καταστάσεις τα οποία ονομάζονται αδρόνια.

1.4 Ασθενής Αλληλεπίδραση

Η τρίτη δύναμη που περιγράφεται από το καθιερωμένο πρότυπο είναι η ασθενής δύναμη[1]. Το φορτίο της ισχυρής δύναμης καλείται ισοσπίν και το συναντά κανείς στα μποζόνια W^\pm, Z . Επειδή όλα τα μποζόνια της ασθενούς αλληλεπίδρασης έχουν μάζα, η ασθενής αλληλεπίδραση έχει πολύ μικρό εύρος. Η ομάδα συμμετρίας που περιγράφει την ασθενή αλληλεπίδραση είναι η SU(2)_L όπου L δείχνει την προτίμηση ζεύξης με αριστερόστροφα σωματίδια και δεξιόστροφα αντισωματίδια. Τα μποζόνια

W^\pm κάνουν σύζευξη μόνο με αριστερόστροφα σωματίδια σε αντίθεση με το μποζόνιο Z που κάνει σύζευξη και με δεξιόστροφα. Μέσω των μποζονίων W , τα φερμιόνια μπορούν να αλλάξουν γεύση. Τα αριστερόστροφα κουαρκ και λεπτόνια είναι διατεταγμένα σε διπλέτες (doublets) ενώ τα δεξιόστροφα σωματίδια αναπαριστώνται από singlets. Επειδή δεν υπάρχουν δεξιόστροφα νετρίνο, το περιεχόμενο των σωματιδίων περιγράφεται ως:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L, \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, u_R, d_R, c_R, s_R, b_R, t_R, e_R, \mu_R, \tau_R.$$

Η θεωρία επιτρέπει μόνο αλλαγή γεύσης μόνο μέσα στις διπλέτες αυτές. Ωστόσο έχουν παρατηρηθεί διασπάσεις από μία γεννιά σε μία άλλη. Η αλλαγή γεύσης μεταξύ των κουαρκ περιγράφεται από τον πίνακα μίξης των Cabibbo-Kobayash-Maskawa (CKM) και είναι ο εξής:

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

Η πιθανότητα ότι ένα κουαρκ τύπου i θα μετατραπεί σε ένα κουαρκ τύπου j όταν εκλύει ένα W μποζόνιο είναι ανάλογη της ποσότητας $|V_{ij}|$. Λόγω των καταστάσεων μίξης, παρατηρείται όχι μόνο παραβίαση της ομοτιμίας αλλά και παραβίαση του φορτίου (CP-violation).

1.5 Ηλεκτρασθενής Αλληλεπίδραση

Για την καλύτερη κατανόησξ των δυνάμεων του καθιερωμένου προτύπου, θα ήταν πολύ χρήσιμο να μπορέσουμε να ενοποιήσουμε όλες τις δυνάμεις σε έναν ενιαίο μηχανισμό[1]. Το πρώτο βήμα για την ενοποίηση αυτή μπορεί να γίνει ενώνοντας την QED με την ασθενή αλληλεπίδραση. Το φορτίο Q και το ασθενές ισοσπίν T_3 συνδυάζονται με τέτοιο τρόπο έτσι ώστε να σχηματιστεί το υπερφορτίο Y :

$$Y = 2(Q - I_3)$$

Η Λαγκρανζιανή πρέπει να είναι πάλι αναλλοίωτη κάτω από μετασχηματισμούς των ομάδων $SU(2)_L \times U(1)_Y$. Τα παραπάνω έχουν ως αποτέλεσμα τέσσερα άμαζα μποζόνια W_1, W_2, W_3 και B .

1.6 Ο μηχανισμός Higgs και η αυθόρμητη ρήξη συμμετρίας

Τα μποζόνια W_1, W_2, W_3 και B , με βάση τα παραπάνω, είναι άμαζα. Αντίθετα, οι μάζες των W και Z μποζονίων μετρώνται να είναι στα 80 GeV και 91 GeV αντίστοιχα. Επομένως, αναπτύχθηκε μία θεωρία η οποία λύνει το πρόβλημα αυτό, μέσω αυθόρμητου σπάσιμου συμμετρίας [1].

Εισάγεται ένα νέο πεδίο όπου η πυκνότητα Lagrange για τοπικές μεταβολές δεν είναι συμμετρική, ενώ η Lagrange του πεδίου είναι. Αυτό συνεπάγεται όρους μάζας για ηλεκτρασθενή μποζόνια καθώς και ένα νέο μποζόνιο, το μποζόνιο Higgs, στη πυκνότητα Lagrange του καθιερωμένου προτύπου. Μετά από το αυθόρμητο σπάσιμο συμμετρίας, τα W_1 και W_2 συνδυάζονται με τέτοιο τρόπο έτσι ώστε να πάρουμε τα W^\pm μποζόνια της ασθενούς αλληλεπίδρασης μέσω

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

Τα άλλα δύο πεδία W_3 και B συνδυάζονται και δίνουν ως αποτέλεσμα το μποζόνιο Z από την ασθενή και το φωτόνιο γ από την ηλεκτρομαγνητική αλληλεπίδραση:

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

Εδώ εισήχθη και η γωνία ασθενούς μίξης (weak mixing angle) θ_W , η οποία επίσης συσχετίζει τις μάζες M_W και M_Z των W^\pm και Z μποζονίων μέσω της:

$$M_Z = \frac{M_W}{\cos \theta_W}$$

Σε αυτό το μηχανισμό μπορούν να περιλαμβάνονται και μάζες φερμιονίων. Επομένως, εισάγεται μία σύζευξη Yukawa με το πεδίο Higgs ανάλογη με τις μάζες. Μετά το σπάσιμο της συμμετρίας, το Καθιερωμένο Πρότυπο μπορεί να γραφτεί με μία Λαγκρανζιανή αναπαράσταση. Γι'αυτό, η πυκνότητα Lagrange QCD πρέπει να συνδυαστεί με την συνολική Πυκνότητα Lagrange της ηλεκτρασθενούς αλληλεπίδρασης μετά το σπάσιμο της συμμετρίας. Η συνολική πυκνότητα Lagrange περιέχει επίσης όλους τους κινητικούς όρους των φερμιονίων, όρους μάζας των ηλεκτρασθενών μποζονίων και του μποζονίου Higgs, όρους σύζευξης των φερμιονίων σε ηλεκτρασθενή βαθμωτά μποζόνια (gauge bosons), όρους αλληλεπίδρασης μεταξύ ηλεκτρασθενών μποζονίων και του Higgs μποζονίου καθώς και όρους που εκφράζουν την σύζευξη Yukawa των φερμιονίων στο πεδίο Higgs [2].

Κεφάλαιο 2

Top Quark

2.1 Εισαγωγή

Το τοπ κουαρκ είναι μέλος της διπλέτας weak-isospin που περιέχει το bottom κουαρκ με φορτίο $Q = 2/3$ και $T_3 = +1/2$ [3]. Επειδή το τοπ κουαρκ είναι πιο βαρύ από το W μποζόνιο, είναι το μόνο κουαρκ που διασπάται ημι-ασθενώς σε ένα W μποζόνιο και ένα b κουαρκ. Επομένως, έχει μια πολύ μικρή διάρκεια ζωής και διασπάται πριν μπορέσει να συμβεί η αδρονοποίηση. Είναι το μοναδικό κουάρκ της οποίας η τάξη της σύζευξης του Yukawa με το μποζόνιο Higgs είναι τάξης μονάδας. Το top κουάρκ παίζει έναν ειδικό ρόλο στο καθιερωμένο πρότυπο (SM) και στις θεωρίες πέρα από το καθιερωμένο πρότυπο. Η φαινομενολογία του Καθιερωμένου προτύπου παρέχει ένα μοναδικό εργαστήριο όπου μπορεί να εξεταστεί η κατανόηση της ισχυρής αλληλεπίδρασης, τόσο διαταραχών όσο και μη διαταρακτικών καταστάσεων. Μια ακριβής γνώση των κορυφαίων ιδιοτήτων του top κουάρκ, όπως η μάζα, η σύζευξη, η ενεργός διατομή, το branching ratio της διάσπασης κ.λπ., μπορούν να συναγάγουν μεγάλης σημασίας βασικές πληροφορίες σχετικά με τις θεμελιώδεις αλληλεπιδράσεις.

2.2 Παραγωγή Τοπ Κουαρκ

Στις αδρονικές συγκρούσεις, τα τοπ κουαρκ παράγονται κυρίως σε ζεύγη μέσω της διαδικασίας $q\bar{q} \rightarrow t\bar{t}$ και $gg \rightarrow t\bar{t}$ σε leading order στην QCD. Στις ενέργειες του LHC, περίπου 90% της παραγωγής είναι από πλευρικές διαδικασίες στα $\sqrt{s} = 14\text{TeV}$ [4].

Στο LHC υποθέτοντας ότι η κορυφαία μάζα κουάρκ είναι $173.2\text{GeV}/c^2$ οι διατομές είναι: $\sigma_{t \text{ bart}} = 173.6_{-5.9}^{4.5}\text{pb}$ στο $\sqrt{s} = 7\text{TeV}$ και $\sigma_{t\bar{t}} = 815.96_{-28.61}^{+19.37}\text{pb}$ στο $\sqrt{s} = 13\text{TeV}$. Υποθέτοντας ότι $|V_{tb}|$ Οι διατομές για μια ενιαία κορυφαία παραγωγή είναι ανάλογες με το $|V_{tb}^2|$, πράγμα που σημαίνει ότι δεν χρειάζεται καμία επιπλέον υπόθεση για τον αριθμό των οικογενειών κουάρκ ή στην μοναδικότητα του πίνακα

CKM στην εξαγωγή $|V_{tb}|$. Χωριστές μετρήσεις των διαδικασιών στα κανάλια s και t παρέχουν σημαντικές πληροφορίες για τη φυσική πέρα από το πρότυπο μοντέλο.

Με μια μάζα πάνω από το όριο Wb και $|V_{tb}| \gg |V_{td}|, |V_{ts}|$, το πλάτος διάσπασης του τοπ κουάρκ αναμένεται να κυριαρχείται από το κανάλι των δύο σωμάτων (2 body channel): $t \rightarrow Wb$. Το πλάτος που προβλέπεται στο SM είναι:

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2a_s}{3\pi} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right]$$

παραλείποντας τους όρους τάξης $m_b^2/m_t^2, a_s^2$, όπου m_t αναφέρεται στην pole mass του τοπ κουάρκ. Το πλάτος για μια τιμή $m_t = 173.3 \text{ GeV}/c^2$ είναι $1.35 \text{ GeV}/c^2$ και αυξάνεται όσο αυξάνεται η μάζα. Με μια διάρκεια ζωής περίπου 0.5^{-24} s, το τοπ κουάρκ αναμένεται να διασπαστεί πριν από την εμφάνιση ανώτερων top-flavoured αδρονίων των $t\bar{t}$ -quarkonium-bound καταστάσεις μπορούν να σχηματιστούν. Οι διορθώσεις a_s^2 QCD για το Γ_t είναι επίσης διαθέσιμες βελτιώνοντας τη συνολική θεωρητική ακρίβεια περισσότερο από 1 %.

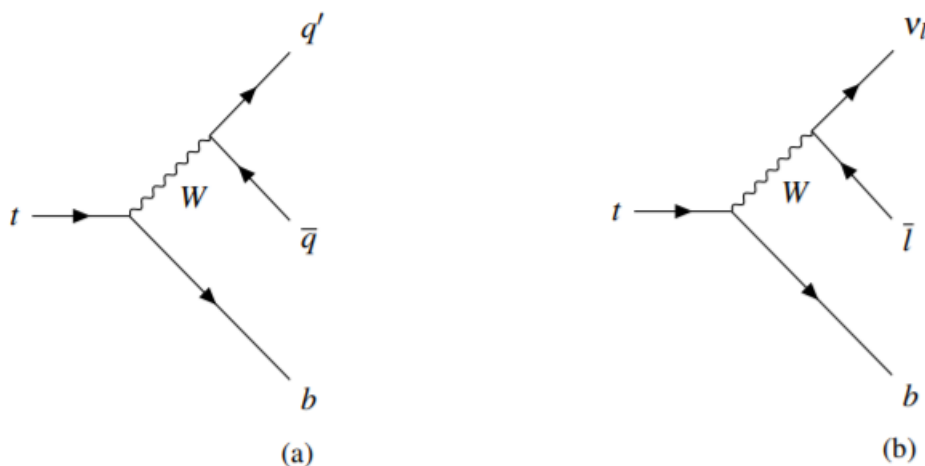


Figure 2.1: Feynman diagrams of a fully hadronic (a) and lepton + jet (b) decay of the top quark. Decays of anti top quarks are constructed similarly

2.3 Παραγωγή $t\bar{t}$

Οι τελικές καταστάσεις για την βασική διαδικασία παραγωγής ζευγών μπορεί να χωριστεί σε 3 κατηγορίες [4]

A.

$$t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow q\bar{q}' b q'' \bar{q}'' \bar{b}$$

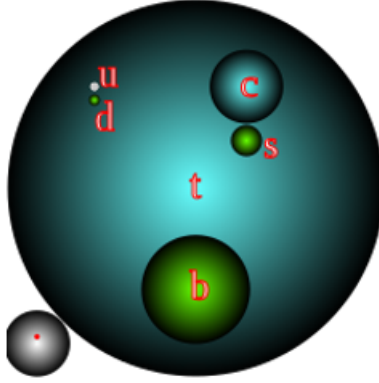


Figure 2.2: Top quark mass with relation to the masses of the other quarks

(45.7%)

B.

$$t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow qq'bl^-\bar{\nu}_l\bar{b} + l^+\nu_l bq''\bar{q}''\bar{q}$$

(43.8%)

C.

$$t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+\nu_l b l'^-\bar{\nu}_l\bar{b}$$

(10.5%)

Τα κουάρκ στις τελικές καταστάσεις καταλήγουν σε jets από αδρόνια. Οι τρεις παραπάνω κατηγορίες διαδικασιών παραγωγής ζεύγους τοπ κουαρκ αναφέρονται ως κανάλια τύπου "all-jets" (A), lepton + jets (B) και dilepton (II) (C). Οι σχετικές συνεισφορές τους, συμπεριλαμβανομένων των αδρονικών διορθώσεων, δίνονται στις παρενθέσεις. Το I στις παραπάνω διαδικασίες αναφέρεται στα e, μ ή τ. Εκτός από τα κουάρκ που προκύπτουν από την αποσύνθεση του τοπ κουάρκ, η επιπλέον ακτινοβολία QCD από τα coloured σωματίδια στα γεγονότα μπορεί να οδηγήσει σε επιπλέον jets. Ο αριθμός των jets που ανακατασκευάζονται στους ανιχνευτές εξαρτάται από την κινηματική της διάσπασης, καθώς και από τον αλγόριθμο ανακατασκευής των jets που χρησιμοποιούνται από την ανάλυση. Πληροφορίες σχετικά με την εγκάρσια ορμή των νετρίνων λαμβάνεται από την ανισοροπία στην εγκάρσια ορμή που μετράται σε κάθε γεγονός.

Είναι ευκολότερο να αναγνωρίσει κανείς τοπ κουαρκ στο QCD $t\bar{t}$ κανάλι σε αντίθεση με το ηλεκτρασθενές single τοπ κουαρκ κανάλι λόγω μίας λιγότερο χαρακτηριστικής επίδρασης καθώς και μεγαλύτερου υποβάθρου κυρίως λόγω της παραγωγής των $t\bar{t}$ and W +jets.

Τα πλήρως αδρονικά και lepton+jets κανάλια κυριαρχούν και συμβαίνουν το 45,7% και 43.8 % του χρόνου αντίστοιχα. 10,5 % όλων των γεγονότων $t\bar{t}$, καταλή-

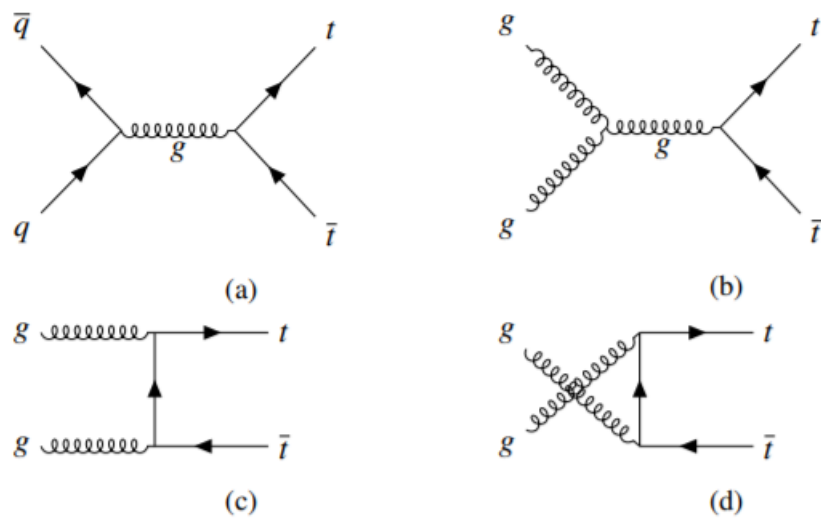


Figure 2.3: Feynman diagrams showing the production of top quark pairs. Displayed are quark anti-quark annihilation (a), and gluon-gluon fusion in the s-channel (b), t-channel (c) and u-channel (d)

γουν σε 2 λεπτόνια στην τελική τους κατάσταση. Η ανάλυση αυτή θα επικεντρωθεί στην πλήρως αδρονική διάσπαση.

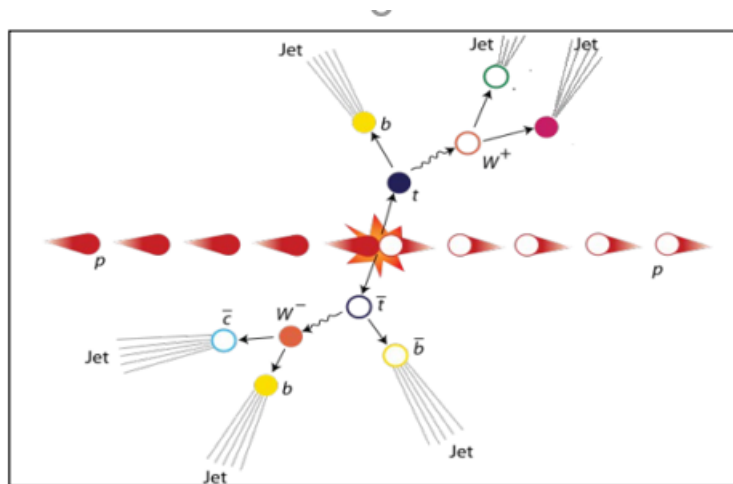


Figure 2.4: Fully Hadronic $t\bar{t}$ decay

Pile Up

Υψηλός ρυθμός σύγκρουσης απαιτείται στον LHC για τη συλλογή μεγάλου όγκου δεδομένων και τη δυνατότητα να παρατηρηθούν ακόμη και πολύ σπάνιες διαδικασίες. Αλλά ο υψηλός ρυθμός έχει επίσης μειονεκτήματα. Δεδομένου ότι οι αλληλεπιδράσεις πρωτονίων-πρωτονίων πραγματοποιούνται ανά bunch crossing, δεν είναι μόνο ένα ενδιαφέρον αλλά και πολλά άλλα ενδιαφέροντα γεγονότα που προκύπτουν, κυρίως soft διαδικασίες QCD, και παρατηρούνται ταυτόχρονα στον ανιχνευτή. Αυτή η επίδραση ονομάζεται pile up. Αν δεν διορθωθεί, οι μετρήσεις της ενέργειας δεν περιλαμβάνουν πάντοτε τα σωματίδια που προέρχονται από τη διαδικασία του hard scattering που μας ενδιαφέρει. Για να μειώσουμε τα αποτελέσματα pile up, είναι σημαντικό να είναι σε θέση κανείς να προσδιορίσει επιπρόσθετη ενέργεια καθώς και να διακρίνει μεταξύ της "σκληρής" σκέδασης και κορυφών από πρόσθετες αλληλεπιδράσεις.

Underlying Event

Εφόσον τα πρωτόνια είναι σύνθετα σωματίδια, μπορούν να προκύψουν πολλαπλές αλληλεπιδράσεις παρτονίων σε μία σύγκρουση πρωτονίων. Η παραπάνω διαδικασία ονομάζεται underlying Event και αφορά γεγονότα με παρτόνια που αλληλεπιδρούν μεταξύ τους και προκύπτουν από τα πρωτόνια που συγκρούονται μεταξύ τους. Αφού και οι δύο αλληλεπιδράσεις έχουν κοινό σημείο σύγκρουσης, δεν μπορούν να διαχωριστούν. Ως εκ τούτου, το underlying event πρέπει να μπορεί να γίνει κατανοητό και να περιλαμβάνεται στο στις προσομοιώσεις των γεγονότων.

Κεφάλαιο 3

Ο Ανιχνευτής Compact Muon Solenoid

Στο κεφάλαιο αυτό θα γίνει μια σύντομη αναφορά στον μεγάλο επιταχυντή αδρονίων (Large Hadron Collider, LHC) καθώς και στον ανιχνευτή Compact Muon Solenoid (CMS).

3.1 The Large Hadron Collider

Ο μεγάλος επιταχυντής αδρονίων (Large Hadron Collider, LHC) είναι ο μεγαλύτερος επιταχυντής σωματιδίων που έχει κατασκευαστεί ποτέ. Ο LHC βρίσκεται και λειτουργεί στις εγκαταστάσεις του ευρωπαϊκού κέντρου έρευνας στην πυρηνική φυσική (CERN). Ο επιταχυντής τέθηκε για πρώτη φορά σε λειτουργία το 2008. Είναι κυκλικού σχήματος με περίμετρο 27 χιλιόμετρα και βρίσκεται 100 μέτρα κάτω από τη γη στα Γαλλο-Ελβετικά σύνορα. Η λειτουργία του επιταχυντή βασίζεται στην επιτάχυνση 2 δεσμών πρωτονίων που κινούνται σε αντίθετες κατευθύνσεις. Οι δέσμες αυτές, συγκρούονται σε 4 προκαθορισμένα σημεία στα οποία έχουν τοποθετηθεί ανιχνευτές σωματιδίων. Ο επιταχυντής αυτός είναι φτιαγμένος, έτσι ώστε να μπορεί να επιτυγχάνει συγκρούσεις με ενέργεια ίση με 14 TeV στο σύστημα κέντρου μάζας. Αυτή η ικανότητα του επιταχυντή τον κάνει τον ισχυρότερο επιταχυντή που έχει κατασκευαστεί ποτέ. Στόχος της κατασκευής του είναι η μελέτη της δομής και της συμπεριφοράς των στοιχειωδών σωματιδίων καθώς και των αλληλεπιδράσεών τους [5].

3.1.1 Το επιταχυντικό σύστημα του LHC

Το επιταχυντικό σύστημα του LHC βρίσκεται, όπως αναφέρθηκε και προηγουμένως, 100 μέτρα κάτω από τη γη. Κατασκευάστηκε στο τούνελ στο οποίο λειτουργούσε

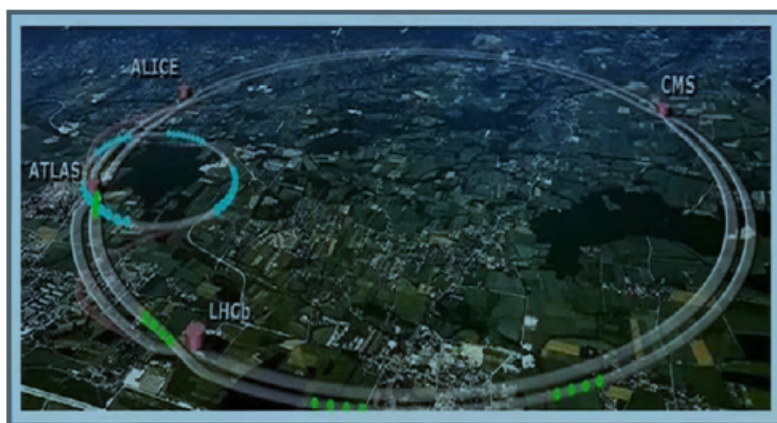


Figure 3.1: Αεροφωτογραφία του ευρωπαϊκού κέντρου έρευνας στη φυσική στοιχειωδών σωματιδίων (CERN). Στη φωτογραφία απεικονίζονται σχηματικά τα τούνελ στα οποία επιταχύνονται οι δέσμες καθώς και τα τέσσερα σημεία στα οποία βρίσκονται οι ανιχνευτές. Όλες οι διατάξεις βρίσκονται 100 μέτρα κάτω από την επιφάνεια της γης.

ο επιταχυντής LEP (Large Electron-Positron Collider). Ο LHC αποτελείται από ένα σύνολο υπεραγωγίων μαγνητών που χρησιμοποιούνται για την καμπύλωση της δέσμης, και την εσίασή της στα σημεία που πραγματοποιούνται οι συγκρούσεις. Για την επιτάχυνση/καμπύλωση της δέσμης χρησιμοποιούνται περίπου 1200 μαγνήτες ενώ για την εστίαση περίπου 400. Οι μαγνήτες διατηρούνται σε θερμοκρασία 1.9K κατά τη διάρκεια λειτουργίας του ανιχνευτή και για την ψύξη τους χρησιμοποιείται υγρό ήλιο. Στο εσωτερικό του κάθε μαγνήτη βρίσκονται 2 θάλαμοι κενού μέσα στους οποίους κινείται η δέσμη, όπως φαίνεται και στην εικόνα 3.2.

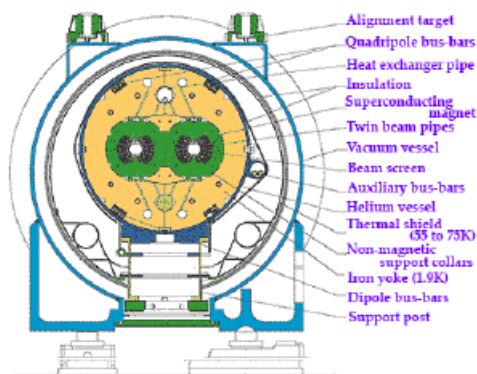


Figure 3.2: Εγκάρσια τομή ενός από τους μαγνήτες του LHC

Για την επιτάχυνση της δέσμης σωματιδίων, το CERN διαθέτει ένα σύστημα

μικρότερων επιταχυντικών διατάξεων που έχουν σκοπό την μερική επιτάχυνση της δέσμης πριν εκείνη εισέλθει στον κεντρικό επιταχυντή, τον LHC. Η δέσμη περνά διαδοχικά καθέναν από τους επιταχυντές πριν εισέλθει στον LHC για το τελικό στάδιο της επιτάχυνσης και τη σύγκρουση των δύο δεσμών. Τα πρωτόνια προέρχονται από μία φιάλη με αέριο υδρογόνου το οποίο εισέρχεται σε ισχυρό ηλεκτρικό πεδίο, όπου και αφαιρούνται τα ηλεκτρόνια από τους πυρήνες με αποτέλεσμα να προκύπτουν πρωτόνια. Το πρώτο στάδιο της επιτάχυνσης των πρωτονίων γίνεται στον επιταχυντή LINAC 2 ο οποίος είναι γραμμικός επιταχυντής που επιταχύνει τα πρωτόνια σε ενέργεια 50 MeV. Στη συνέχεια η δέσμη εισέρχεται στο Proton Synchrotron Booster (PSB) και έπειτα στο Proton Synchrotron όπου τα πρωτόνια επιταχύνονται σε ενέργεια 26 GeV. Το επόμενο στάδιο επιτάχυνσης επιτυγχάνεται στο Super Proton Synchrotron όπου και τα πρωτόνια αποκτούν ενέργεια 450 GeV. Τέλος η δέσμη εισέρχεται στον LHC για το τελικό στάδιο επιτάχυνσης.

Οι δύο παράμετροι που καθορίζουν την λειτουργία του LHC είναι οι ενέργειες των 2 συγκρουόμενων δεσμών και η φωτεινότητά τους (Luminosity). Ο ρυθμός των γεγονότων (συγκρούσεων) που προκύπτουν από τη λειτουργία του επιταχυντή δίνεται από τη σχέση

$$R = \sigma L$$

όπου σ είναι η ενεργός διατομή της φυσικής διεργασίας και L η φωτεινότητα. Για να διατηρηθεί η αποδοτικότητα του πειράματος για μία δεδομένη ενέργεια E , η φωτεινότητα της διάταξης πρέπει να αυξάνεται σε αναλογία με το E^2 , καθώς η ενεργός διατομή είναι αντιστρόφως ανάλογη με την E^2 , δηλαδή $\sigma \propto 1/E^2$. Η φωτεινότητα ενός επιταχυντή στον οποίο συγκρούονται δύο πακέτα με αριθμό σωματιδίων n_1 και n_2 με συχνότητα f είναι:

$$L = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$$

όπου σ_x, σ_y είναι τα εγκάρσια προφίλ της δέσμης. Η προβλεπόμενη φωτεινότητα του LHC είναι $L = 10^{34} \text{cm}^{-2}\text{s}^{-1}$. Για την επίτευξη τόσο μεγάλης φωτεινότητας, θα πρέπει η συχνότητα των συγκρούσεων καθώς και ο αριθμός των πρωτονίων σε κάθε πακέτο να είναι αρκετά μεγάλος. Γι'αυτό το λόγο κάθε δέσμη πρωτονίων αποτελείται από 2808 πακέτα πολύ κοντά μεταξύ τους στα οποία περιέχονται περίπου 10^{11} πρωτόνια που συγκρούονται κάθε 25ns. Αυτό δίνει συχνότητα συγκρούσεων ίση με 40 MHz.

3.1.2 Ανιχνευτικές Διατάξεις του LHC

Όπως αναφέρθηκε και προηγουμένως οι συγκρούσεις των δεσμών πρωτονίων παρατηρούνται σε 4 προκαθορισμένα σημεία πάνω στο δακτύλιο του επιταχυντή, όπου εκεί

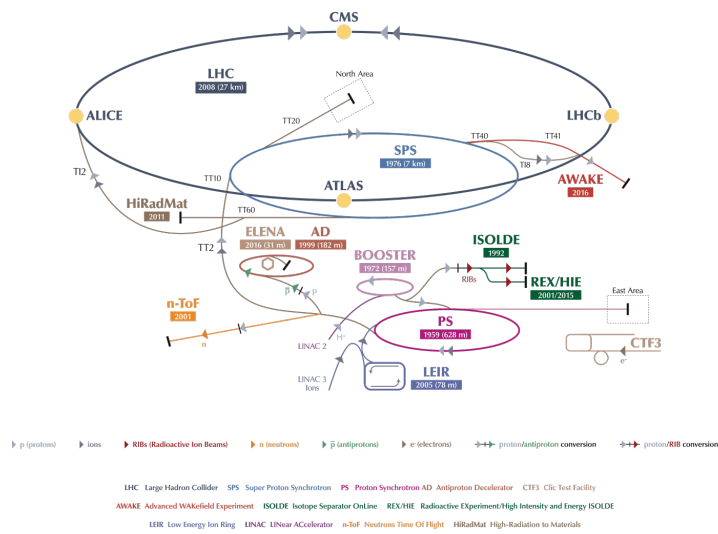


Figure 3.3: Το σύμπλεγμα επιταχυντών και πειραμάτων του LHC

είναι τοποθετημένοι ανιχνευτές που σκοπεύουν στην παρατήρηση και μελέτη των συγκρούσεων. Συνολικά ο LHC έχει 8 σημεία πρόσβασης στα 4 από τα οποία είναι τοποθετημένοι ανιχνευτές. Οι ανιχνευτές είναι οι εξής:

- A Toroidal LHC Apparatus (ATLAS), point 1
- Compact Muon Solenoid (CMS), point 5
- LHC Beauty (LHCb), point 2
- A Large Ion Collider Experiment (ALICE), point 8

Οι δύο πρώτοι ανιχνευτές CMS και ATLAS, είναι πειράματα γενικού ενδιαφέροντος στη φυσική υψηλών ενεργειών. Αντίθετα, οι LHCb και ALICE είναι κατασκευασμένοι για τη μελέτη συγκεκριμένων φαινομένων. Πιο συγκεκριμένα το LHCb έχει σαν σκοπό την παρατήρηση του b (beauty) κουάρκ καθώς και τη μέτρηση της παραβίασης της συμμετρίας CP (CP violation). Το πείραμα ALICE είναι σχεδιασμένο για να μελετήσει τις συγκρούσεις βαρέων ιόντων αλλά και βαρέων ιόντων - πρωτονίων οι οποίες θα πραγματοποιούνται εκτός από αυτές των πρωτονίων.

Πέρα από αυτά τα 4 πειράματα υπάρχουν και άλλα μικρότερα όπως το TOTEM και το CASTOR τα οποία μοιράζονται το ίδιο σημείο σύγκρουσης με εκείνο του CMS.

3.2 Compact Muon Solenoid

Ο ανιχνευτής CMS είναι ένας ανιχνευτής γενικού ενδιαφέροντος. Είναι σχεδιασμένος και κατασκευασμένος έτσι ώστε να μπορεί να ανιχνεύει όλα τα στοιχειώδη σωματίδια. Το σύνολο των ανιχνευτικών διατάξεων που βρίσκονται στο CMS σχηματίζουν έναν κύλινδρο με διάφορα επίπεδα ομοαξονικά στον άξονα της δέσμης σαν ένα βαρέλι (barrel). Όμως η ανάγκη ανίχνευσης όλων των σωματιδίων που παράγονται κατά τη σύγκρουση των πρωτονίων στο κέντρο του CMS επιβάλλει την επιπλέον προσαρμογή δύο ακόμα δίσκων κάθετα στον άξονα της δέσμης στα άκρα του βαρελιού έτσι ώστε ο ανιχνευτής να είναι ερμητικά κλειστός (endcaps). Ο ανιχνευτής έχει μήκος 28,7m και ακτίνα 15m και ζυγίζει περίπου 14.000 τόνους [5].

Ένα από τα σωματίδια τα οποία έχουν το μεγαλύτερο ενδιαφέρον ως προς την ανίχνευσή τους είναι τα μόνια. Η ύπαρξη και ξεκάθαρη μελέτη μιονίων είναι αναγκαία για την πραγματοποίηση διεργασιών οι οποίες είναι πολύ σημαντικές για τη λειτουργία του πειράματος. Για το λόγο αυτό και το CMS είναι κατασκευασμένο έτσι ώστε να ανιχνεύει μόνια με μεγάλη ακρίβεια. Από εκεί μάλιστα προκύπτει και η λέξη muon στο όνομα του ανιχνευτή.

Πολύ σημαντικό κομμάτι του ανιχνευτή είναι τέλος ο μαγνήτης, ο οποίος επιτρέπει την ακριβή μέτρηση της ορμής των σωματιδίων που παράγονται. Στο CMS επιλέχθηκε η χρησιμοποίηση ενός σωληνοειδούς (solenoid), ο οποίος παράγει μαγνητικό πεδίο έντασης 4T παράλληλα στον άξονα της δέσμης με αποτέλεσμα να καμπυλώνει τις τροχιές των σωματιδίων κάθετα στον άξονα αυτής.

Γνωρίζουμε πως η ανακατασκευή της ορμής ενός μιονίου από έναν ανιχνευτή, έχει άμεση εξάρτηση από την ισχύ του μαγνητικού πεδίου αλλά και το μήκος του ανιχνευτή. Πιο συγκεκριμένα:

$$\frac{\sigma_{P_T}}{P_T} = \frac{\sigma_s}{s} = \sigma_s \frac{8P_T}{0.3BL^2}$$

Όπου P_T η εγκάρσια συνιστώσα της ορμής του σωματιδίου, s η καμπύλωση της τροχιάς από το μαγνητικό πεδίο του ανιχνευτή με ένταση μαγνητικού πεδίου B , και L το μήκος του ανιχνευτή. Το CMS έχει μικρό μέγεθος και μεάλο μαγνητικό πεδίο σε αντίθεση με τον ATLAS Ο οποίος έχει μεγαλύτερο μέγεθος και μικρότερο μαγνητικό πεδίο.

Ο ανιχνευτής CMS αποτελείται από τις παρακάτω επιμέρους ανιχνευτικές διατάξεις:

- Σύστημα ανίχνευσης τροχιών που δίνει πληροφορίες για το είδος και την ορμή των σωματιδίων που προέρχονται από την σύγκρουση των δεσμών
- Ένα ηλεκτρομαγνητικό θερμιδόμετρο, για την ανίχνευση και μελέτη φωτονίων και ηλεκτρονίων

- Ένα αδρονικό θερμιδόμετρο για την ανίχνευση και μελέτη ουδέτερων αλλά και φορτισμένων αδρονίων
- Ένα μιονικό σύστημα υψηλής απόδοσης για την μελέτη και ανίχνευση μιονίων

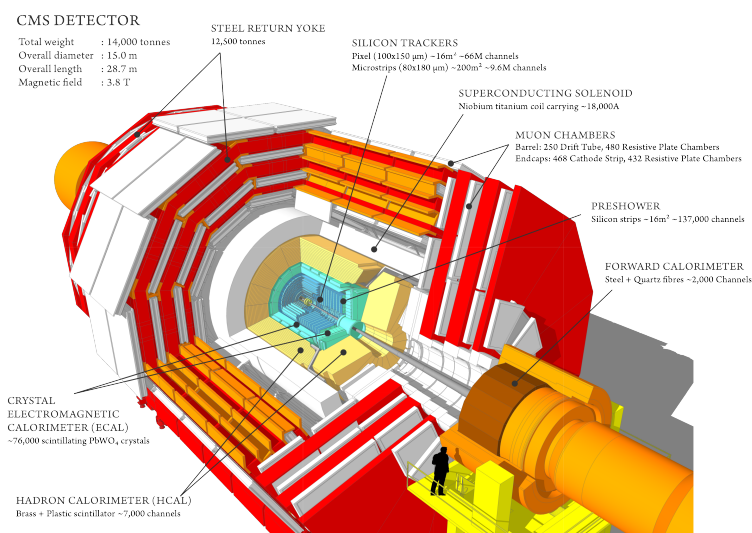


Figure 3.4: Ο ανιχνευτής Compact Muon Solenoid. Ο ανιχνευτής αποτελείται από 4 επιμέρους υποανιχνευτές, τον Tracker, το ηλεκτρομαγνητικό θερμιδόμετρο, το αδρονικό θερμιδόμετρο και τους ανιχνευτές μιονίων

Το CMS έχει σχεδιαστεί σε διάφορα επίπεδα καθένα από τα οποία έχει έναν ξεχωριστό ρόλο όπως αναφέρθηκε και προηγουμένως. Κάθε επίπεδο αποτελεί έναν υποανιχνευτή (subdetector) ο οποίος είναι υπεύθυνος για την συλλογή ενός συγκεκριμένου κομματιού πληροφορίας. Όλα τα κομμάτια συλλέγονται και ο συνδυασμός τους τελικά δίνει την δυνατότητα ταυτοποίησης των σωματιδίων που δημιουργήθηκαν κατά της σύγκρουση των δύο δεσμών πρωτονίων. Όπως αναφέρθηκε και προηγουμένως, το CMS αποτελείται από 4 υποανιχνευτές. Τον ανιχνευτή τροχιάς (Tracker), το ηλεκτρομαγνητικό θερμιδόμετρο (Electromagnetic Calorimeter, ECAL), το αδρονικό θερμιδόμετρο (Hadronic Calorimeter, HCAL), και το ανιχνευτικό σύστημα μιονίων (Muon detectors).

Η διάταξη των υποανιχνευτών στο CMS είναι η εξής. Τα 3 ανιχνευτικά συστήματα Tracker, ECAL, HCAL βρίσκονται με αυτή τη σειρά (ξεκινώντας από το σημείο σύγκρουσης των δεσμών) εσωτερικά του μαγνήτη. Αντίθετα, το μιονικό σύστημα βρίσκεται έξω από το μαγνήτη. Κάθε ένα από τα κομμάτια αυτά διαθέτει και ένα κομμάτι του στα καπάκια του ανιχνευτή (endcaps).

Στις επόμενες ενότητες θα γίνει πλήρης περιγραφή του κάθε ανιχνευτή που χρησιμοποιείται στο CMS ξεχωριστά. Είναι σημαντικό όμως να γίνει μία αναφορά

στο σύστημα συντεταγμένων που χρησιμοποιείται στα πειράματα του LHC. Για την περιγραφή του ανιχνευτή μπορεί να χρησιμοποιηθεί ένα καρτεσιανό σύστημα συντεταγμένων σύμφωνα με το οποίο ο άξονας x δείχνει προς το κέντρο του δακτυλίου του LHC, ο άξονας y δείχνει προς τα πάνω κάθετα στον άξονα της δέσμης και ο άξονας z είναι παράλληλος με τον άξονα της δέσμης. Ωστόσο λόγω του κυλινδρικού σχήματος του ανιχνευτή, η περιγραφή γίνεται πιο εύκολη χρησιμοποιώντας ένα άλλο σύστημα αναφοράς το οποίο χαρακτηρίζεται από τις ποσότητες (r, φ, η) . Πιο συγκεκριμένα, r είναι η απόσταση από τον άξονα z, φ είναι η αζιμουθιακή γωνία από τον άξονα x και η (pseudo-rapidity) μία ποσότητα η οποία δίνεται από τη σχέση $\eta = -\ln(\tan(\frac{\theta}{2}))$, όπου θ η γωνία από το θετικό ημιάξονα z. Στην εικόνα 3.5 δίνεται η pseudo-rapidity για διάφορες τιμές της γωνίας θ .

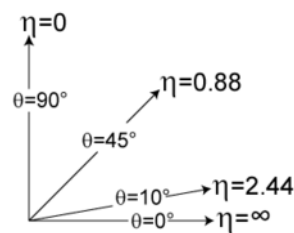


Figure 3.5: Η pseudorapidity ως συνάρτηση της γωνίας θ

3.2.1 Ανιχνευτές Τροχιάς - Tracker Detector

Ο ανιχνευτής τροχιάς χρησιμοποιείται για να ανιχνεύσει την ορμή ενός σωματιδίου και τη διαδρομή που εκείνο ακολούθησε μέσα στο μαγνητικό πεδίο. Όσο περισσότερο κυρτή είναι η διαδρομή που ακολουθεί το σωματίδιο, τόσο μικρότερη είναι και η ορμή του σωματιδίου. Ο ανιχνευτής τροχιάς στο πείραμα CMS καταγράφει τη διαδρομή των φορτισμένων σωματιδίων, υπολογίζοντας τις συντεταγμένες τους σε ορισμένα σημεία "κλειδιά" (key points).

Ο ανιχνευτής τροχιάς μπορεί να ανακατασκευάσει τη διαδρομή μιονίων και ηλεκτρονίων υψηλής ενέργειας καθώς και τη διαδρομή αδρονίων. Ο ανιχνευτής τροχιάς μπορεί να ανακατασκευάσει ακόμα και διαδρομές από σωματίδια που προέρχονται από σωματίδια με πολύ μικρό χρόνο ζωής (πχ b quark).

Κάθε μέτρηση έχει χωρική διακριτική ικανότητα 10 μm . Η τελική σχεδίαση του ανιχνευτή τροχιάς αποτελείται εξ ολοκλήρου από Πυρίτιο. Είναι φτιαγμένο από 75 εκατομύρια ηλεκτρονικές συσκευές ανάγνωσης. Οι δύο ανιχνευτές τροχιάς που χρησιμοποιούνται στο πείραμα είναι οι: Silicon Pixel και Silicon Microstrip ανιχνευτές.

Πίξελ Πυριτίου - Silicon Pixels

Οι ανιχνευτές πίξελ πυριτίου είναι ανιχνευτές στερεάς κατάστασης. Έχουν μέγεθος όσο ένα κουτί παπουτσιών. Το κάθε κουτί περιέχει 65 εκατομύρια πίξελ που χρησιμοποιούνται για ανίχνευση. Αυτό επιτρέπει στους ανιχνευτές να ανιχνεύουν τη διαδρομή σωματιδίων που προκύπτουν από τις συγκρούσεις στον επιταχυντή. Πιο συγκεκριμένα, οι ανιχνευτές πίξελ πυριτίου μπορούν να ανιχνεύσουν $\frac{10 \times 10^6 \text{ particles}}{\text{cm}^2 \text{ s}}$.

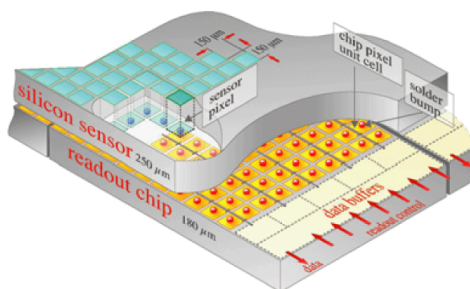


Figure 3.6: Silicon Pixel Detector

Κάθε στρώμα είναι χωρισμένο σε τμήματα. Κάθε τμήμα είναι εξοπλισμένο με έναν αισθητήρα πυριτίου (100-150μm). Όταν ένα φορτισμένο σωματίδιο διαπερνά τον ανιχνευτή, έχει αρκετή ενέργεια έτσι ώστε ηλεκτρόνια αποβάλλονται από τα άτομα του πυριτίου δημιουργώντας ζεύγη ηλεκτρονίων και οπών. Κάθε πίξελ χρησιμοποιεί ηλεκτρικό ρεύμα για να συλλέξει τα παραπάνω φορτία στην επιφάνεια ενός μικρού ηλεκτρικού σήματος. Η ενίσχυση του σήματος γίνεται μέσω της τεχνικής bump-bonding technique.

Ανιχνευτές Στριπ Πυριτίου - Silicon Strips Detectors

Οι ανιχνευτές αυτοί έχουν σχεδόν την ίδια λειτουργία με εκείνους των silicon strips. Ακριβώς μετά τους ανιχνευτές πίξελ, τα σωματίδια περνούν μέσα από 10 στρώματα από ανιχνευτές πυριτίου λωρίδας (silicon strip detectors). Το πείραμα CMS αποτελείται από τέσσερα στρώματα στα εσωτερικά βαρέλια (inner barrels TIB) που είναι συναρμολογημένα σε φλοιούς με δύο εσωτερικά καπάκια (endcaps), καθένα αποτεύεται από τρεις μικρούς δίσκους. Το εξωτερικό βαρέλι (Outer barrel TOB) αποτελείται από έξι ομόκεντρα στρώματα. Τέλος δύο καπάκια (TEC) συμπληρώνουν τον ανιχνευτή. Κάθε τμήμα πυριτίου είναι σχεδιασμένο διαφορετικά για τη θέση που έχει στον ανιχνευτή.

Υπάρχουν 15.200 πολύ ευαίσθητες ενότητες. Υπάρχουν 10 εκατομύρια λωρίδες (strips) που διαβάζονται από 80.000 μικροηλεκτρονικά τσιπς. Κάθε τμήμα αποτελείται από αισθητήρες, μία μηχανική δομή υποστήριξης και ηλεκτρονικά συστήματα ανάγνωσης.

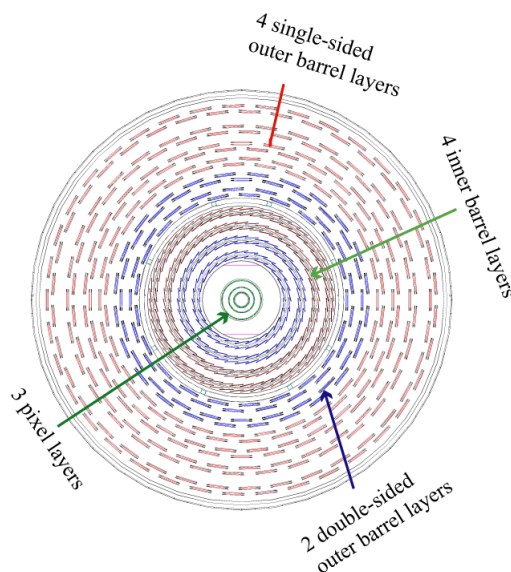


Figure 3.7: Position of Silicon Strip Detectors in CMS experiment

Οι ανιχνευτές πυριτίου έχουν σε γενικές γραμμές γρήγορη απόκριση και πολύ καλή χωρική διακριτική ικανότητα 10-20 μm . Ο φυσικός μηχανισμός με τον οποίο λειτουργούν οι ανιχνευτές στριπ πυριτίου. Όταν ένα φορτισμένο σωματίδιο περνά από τον ανιχνευτή, έχει αρκετή ενέργεια έτσι ώστε να διώξει ηλεκτρόνια από τα άτομα του πυριτίου, δημιουργώντας έτσι ζεύγη ηλεκτρονίων και οπών. Τα ηλεκτρόνια που εξάγονται από τα άτομα πυριτίου, παράγουν ένα μικρό παλμό ο οποίος ενισχύεται από τσιπς APV25 δίνοντας "χτυπήματα" (hits) όταν ένα σωματίδιο περνάει, επιτρέποντας έτσι στον χρήστη να ανακατασκευάσει τη διαδρομή του σωματιδίου.

Για την ελαχιστοποίηση της διαταραχής του πυριτίου λόγω της υψηλής ακτινοβολίας που παράγεται από τις συγκρούσεις, το συγκεκριμένο σημείο του ανιχνευτή κρατιέται σε σταθερή θερμοκρασία $-20C^{\circ}$ έτσι ώστε να σταματήσουν τυχόν ζημιές και να αποτρέψουν τη διαίωσή τους.

Ο APV25 αντιστοιχεί στον Analogue Pipeline Voltage Chip. Υπάρχουν 4 ή 6 μέσα σε ένα "hybrid", το οποίο περιέχει ηλεκτρονικά έτσι ώστε να υπάρχει δυνατότητα παρακολούθησης και άλλων παραμέτρων (πχ θερμοκρασία) και παρέχει χρονικές πληροφορίες που χρησιμοποιούνται για την αντιστοίχιση "hits" και συγκρούσεων. Ο APV25 αποθηκεύει τα σήματα στην μνήμη για μερικά μs και μετά τα επεξεργάζεται και τα στέλνει σε ένα laser, έτσι ώστε τα σήματα να μετατρέπονται σε υπέρυθροι παλμοί. Οι παλμοί μεταφέρονται στη συνέχεια με ένα καλώδιο οπτικής ίνας 100m για περαιτέρω ανάλυση σε ένα περιβάλλον χωρίς ακτινοβολία.

3.2.2 Crystal Electromagnetic Calorimeter (ECAL)

Τα θερμιδόμετρα είναι σε γενικές γραμμές ανιχνευτές που χρησιμοποιούνται για την ανίχνευση ενεργειακής ροής, φορτισμένων σωματιδίων και ουδέτερων σωματιδίων. Αυτοί οι ανιχνευτές μπορούν να μετατρέψουν την ενέργεια ενός σωματιδίου που περνά από τον ανιχνευτή σε απόκριση του ανιχνευτή. Για να μετρηθεί η ολική ενέργεια του φορτισμένου σωματιδίου ή του φωτονίου, το σωματίδιο πρέπει να σταματηθεί πλήρως μέσα στον ανιχνευτή. Η διαδικασία αυτή ξεκινά με το αρχικό σωματίδιο να εισέρχεται στον ανιχνευτή και έπεται το λεγόμενο particle shower (Ηλεκτρομαγνητικός καταγισμός).

Ο ηλεκτρομαγνητικός καταγισμός σωματιδίων ξεκινά όταν ένα φορτισμένο σωματίδιο ή ένα σωματίδιο γ , φτάνει στον ανιχνευτή. Το αρχικό σωματίδιο παράγει ακτινοβολία και θερμότητα. Γενικά, είναι πολύ δύσκολη η αποσύνθεση ενός ηλεκτρομαγνητικού καταγισμού σωματιδίων. Είναι αρκετά ενδιαφέρον ο τρόπος με τον οποίο ο ηλεκτρομαγνητικός καταγισμός δημιουργείται.

Το αρχικό φωτόνιο εισέρχεται στον ανιχνευτή έχοντας αρχική ενέργεια E_0 . Αν το αρχικό σωματίδιο είναι φωτόνιο, παράγεται ένα ηλεκτρόνιο και ένα ποσιτρόνιο μέσω δίδυμης γέννησης. Η πιθανότητα το φωτόνιο να παράξει ένα ηλεκτρόνιο και ένα ποσιτρόνιο είναι 64% μέσα σε ένα διάστημα Q_0 μήκος ακτινοβολίας. Το μήκος ακτινοβολίας (radiation length) X_0 είναι το μήκος εκείνο στο οποίο η ενέργεια του σωματιδίου μειώνεται κατά έναν παράγοντα $1/e$ λόγω εκπομπής ακτινοβολίας. Η απώλεια ενέργειας από τα παραχθέντα ηλεκτρόνια και ποζιτρόνια είναι συνάρτηση κυρίως του ατομικού αριθμού (Z) του ανιχνευτή. Η κρίσιμη ενέργεια του φορτισμένου σωματιδίου είναι η ενέργεια στην οποία οι απώλειες ενέργειας λόγω ιονισμού είναι ακριβώς ίδια με την απώλεια ενέργειας λόγω ακτινοβολίας πέδησης (Brehmstrahlung). Η μέση ενέργεια για τα παραχθέντα ηλεκτρόνια και ποσιτρόνια είναι $E = E_0/2$ αντίστοιχα. Αν η μέση ενέργεια του εισερχόμενου φορτισμένου σωματιδίου είναι μεγαλύτερη από την κρίσιμη ενέργεια, τότε παράγεται ακτινοβολία πέδησης. Μέσα στο επόμενο Q_0 , η ενέργεια του σωματιδίου θα είναι $\frac{E_0}{2e}$. Ο μέσος αριθμός σωματιδίων μετά από ένα μήκος $2X_0$ είναι 4. Μετά από αυτό το σημείο συνεχίζεται η παραγωγή γ σωματιδίων. Μετά από t γενιές, περίπου 2^t σωματίδια έχουν παραχθεί. Ο ηλεκτρομαγνητικός καταγισμός σωματιδίων σταματά μόλις η ενέργεια των ηλεκτρονίων που βρίσκονται μέσα στο ηλεκτρομαγνητικό καλορίμετρο είναι $E(e^-) = E_c = \frac{E_0}{2^t}$.

Με σκοπό να εξηγηθεί ο φυσικός μηχανισμός του ανιχνευτή, θα ήταν σκόπιμο να γίνει μία εξήγηση του όρου σπινθηρισμός. Όταν ένα άτομο διεγείρεται, δηλαδή του δίνεται ενέργεια, ένα ηλεκτρόνιο πηγαίνει σε υψηλότερη ενεργειακή στοιβάδα και μόλις αποδιεγερθεί και επιστρέψει στην βασική κατάσταση, ελευθερώνει ενέργεια μέσω ενός φωτονίου. Στο CMS, όταν ένα ηλεκτρόνιο υψηλής ενέργειας ή ένα φωτόνιο συγκρούεται με έναν πυρήνα του ECAL χρυσάλλου, παράγεται ηλεκτρομαγνητικός καταγισμός φωτονίων, ποζιτρονίων και ηλεκτρονίων έτσι ώστε τα άτομα στο υλικό

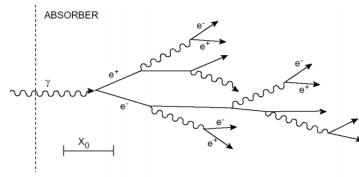


Figure 3.8: Electromagnetic Particle Shower

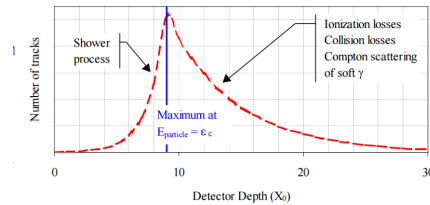


Figure 3.9: Number of tracks as a function of detector depth

του ανιχνευτή λαμβάνουν ενέργεια από τα σωματίδια που περνούν και διεγείρουν και άλλα ηλεκτόνια. Τα διεγερμένα ηλεκτρόνια γρήγορα επιστρέφουν στην βασική τους κατάσταση εκπέμποντας την παραπάνω ενέργεια που τους προσφέρθηκε με τη μορφή ενός φωτονίου μπλε φωτός. Μία συσκευή φωτοπολλαπλασιαστή PMT λαμβάνει τα φωτόνια που παράγονται από σπινθηρισμό και έτσι η συνολική ποσότητα που παράγεται από αυτή τη διαδικασία είναι ανάλογη προς την ενέργεια που εναποτέθηκε από το αρχικά εισερχόμενο σωματίδιο στον κρύσταλλο. Από όλη την παραπάνω διαδικασία, μπορεί να υπολογιστεί η συνολική ενέργεια του εισερχόμενου φωτονίου ή ηλεκτρονίου.

Ο ανιχνευτής φαίνεται στην 3.10. Το φορτισμένο σωματίδιο ή το φωτόνιο εισέρχεται στον ανιχνευτή προκαλώντας ηλεκτρομαγνητικό καταιγισμό. Το σήμα στη συνέχεια ενισχύεται χρησιμοποιώντας έναν φωτοπολλαπλασιαστή PMT (Photon Multiplier Tube).

Στο πείραμα CMS χρησιμοποιείται κρύσταλλος μολύβδου/βολφραμίου $PbWO_4$ για τους 80.000 κρυστάλλους που χρησιμοποιούνται. Πρόκειται για ένα υλικό με πολύ υψηλή πυκνότητα που παράγει σπινθηρισμό ο οποίος οδηγεί σε γρήγορους, μικρούς

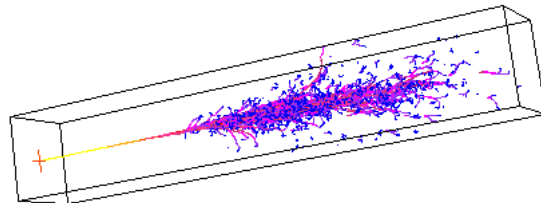


Figure 3.10: Electromagnetic Calorimeter

και καλά ορισμένους ηλεκτρομαγνητικούς καταγισμούς σωματιδίων. Το υλικό αυτό επιλέχθηκε καθώς είναι πολύ πυκνό ($8.28g = cm^3$) και διαθέτει μικρό μήκος ακτινοβολίας ($X_0 = 0.89cm$) και ακτίνα Moliere ($R_m = 0.89cm$) (Η ακτίνα Moliere είναι ένα μέγεθος το οποίο χαρακτηρίζει τον ηλεκτρομαγνητικό καταγισμό που δημιουργείται από ένα φωτόνιο ή ηλεκτρόνιο. Ορίζεται ως η ακτίνα ενός κυλίνδρου μέσα στον οποίο περιέχεται το 90% της ενέργειας του καταγισμού). Το ηλεκτρομαγνητικό καλορίμετρο έχει πολύ καλή ενεργειακή διακριτική ικανότητα ($\sigma/E = 3\% \sqrt[3]{E}$). Ο κρύσταλλος μολύβδου/βολφραμίου είναι επίσης εύκολο να παραχθεί από "raw" υλικά. Μερικές πληροφορίες για τον κρύσταλλο που χρησιμοποιείται είναι:

- Οι κρύσταλλοι μετρούν $2,2 \times 2,2 \times 23$ cm στο βαρέλι και $3 \times 3 \times 22$ cm στα καπάκια
- Υπάρχουν ακριβώς 75.848 κρύσταλλοι στο ECAL
- Η πυκνότητα του κρυστάλλου είναι $8,3 \frac{g}{cm^3}$

Το υλικό που χρησιμοποιείται έχει όμως και μειονεκτήματα. Η απόδοση φωτός εξαρτάται κυρίως από τη θερμοκρασία λόγω της θερμότητας που προέρχεται από τα ηλεκτρονικά. Για να λυθεί το παραπάνω πρόβλημα της ψύξης, ένα σύστημα ψύξης κρατά τη θερμοκρασία των 100 τόνων κρυστάλλων σε ένα εύρος $0.1C^O$. Ένα ακόμα πρόβλημα είναι ότι η απόδοση φωτός που προέρχεται από σπινθηρισμό είναι σχετικά μικρή και έτσι πρέπει να χρησιμοποιηθούν αισθητήρες φωτός, οι οποίοι λαμβάνουν το φως και το μετατρέπουν σε ηλεκτρικό σήμα το οποίο στη συνέχεια ενισχύεται. Το ενισχυμένο ηλεκτρικό σήμα μεταφέρεται μετά για περαιτέρω ανάλυση μέσω οπτικών ινών.

Όπως όλοι οι υποανιχνευτές του CMS έτσι και το ECAL χωρίζεται σε 2 περιοχές. Το κομμάτι που βρίσκεται στο barrel (ECAL Barrel, EB) και αυτό που βρίσκεται στα endcaps (EE). Συνολικά το ECAL καλύπτει την περιοχή $1.2m < r < 1.8m$, $|\eta| < 3$ Το EB καλύπτει την περιοχή $|\eta| < 0.1479$ και αποτελείται από κομμάτια τα οποία ονομάζονται supermodules. Κάθε supermodule αποτελείται από 4 modules, στα οποία το πρώτο περιέχει 500 κρυστάλλους και τα υπόλοιπα 400 κρυστάλλους. Συνολικά υπάρχουν 26 supermodules και 61200 κρυστάλλοι στο EB. Τέλος το κομμάτι του ανιχνευτή στο endcap καλύπτει την περιοχή $0.1479 < |\eta| < 3$. Το EE αποτελείται από 2 ημικύκλια η δομική μονάδα του καθενός, περιέχει 5×5 κρυστάλλους που ονομάζονται "supercrystals". Έτσι κάθε EE περιέχει 7324 κρυστάλλους.

3.2.3 Hadron Calorimeter (HCAL)

Τα αδρονικά καλορίμετρα χρησιμοποιούνται για να μετρήσουν την ενέργεια των αδρονίων, δηλαδή σωματιδίων που αποτελούνται από κουάρκ και γλουόνια. Επίσης, παρέχει μία έμμεση μέτρηση για την ύπαρξη ή μη από μη αλληλεπιδρώντα αφόρτιστα σωματίδια όπως για παράδειγμα τα νετρίνα.

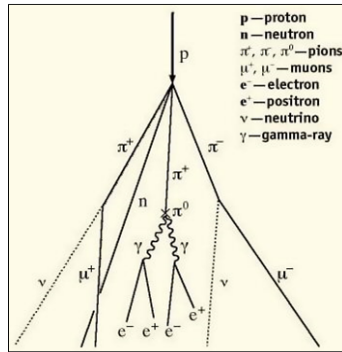


Figure 3.11: Hadron Shower

Το αδρονικό καλορίμετρο "ενεργοποιείται" όταν ένα σωματίδιο αλληλεπιδρά αδρονικά με ενέργεια μεγαλύτερη από 5GeV. Η αλληλεπίδραση ενός αδρονίου με έναν πυρήνα είναι

- Ελαστική $p+N \rightarrow p + N$ (σ_e)
- Ανελαστική $p+N \rightarrow X$ (σ_{incl})

Η φυσική διαδικασία η οποία είναι υπεύθυνη για τη διάδοση του αδρονικού καταιγισμού είναι αρκετά διαφορετική σε σχέση με εκείνη του ηλεκτρομαγνητικού καταιγισμού. Περίπου η μισή ενέργεια από την ενέργεια των αδρονίων του γεγονότος περνά σε δευτερεύοντα σωματίδια. Το υπόλοιπο της ενέργειας καταναλώνεται σε παραγωγή διαφόρων σωματιδίων από αργά πιόνια. Τα φαινόμενα τα οποία καθορίζουν την ανάπτυξη ενός αδρονικού καταιγισμού είναι: παραγωγή αδρονίων, πυρηνική αποδιέγερση και η αποσύνθεση πιονίων και μιονίων. Τα ουδέτερα πιόνια αποτελούν το 1/3 των συνολικά παραχθέντων πιονίων και η ενέργειά τους διασκορπίζεται σε μορφή ηλεκτρομαγνητικού καταιγισμού. Ένα επιπλέον σημαντικό χαρακτηριστικό του αδρονικού καταιγισμού είναι ότι χρειάζεται μεγαλύτερο χρόνο για να εξελιχθεί σε σχέση με τον ηλεκτρομαγνητικό καταιγισμό. Αυτό μπορεί να γίνει διακριτό, συγκρίνοντας τον αριθμό των σωματιδίων σε σχέση με το βάθος διείσδυσης τους για τα πιόνια και τα ηλεκτρόνια στον εκάστοτε αδρονικό και ηλεκτρομαγνητικό καταιγισμό. Η διαμήκης εξέλιξη του αδρονικού καταιγισμού συνδέεται με την πυρηνική απορρόφηση (ή μήκος αλληλεπίδρασης) μέσω της σχέσης: $\lambda = \frac{A}{N_A \sigma_{abs}}$.

Το HCAL είναι ένα δειγματοληπτικό καλορίμετρο, που σημαίνει ότι μπορεί να προσδιορίσει τη θέση, την ενέργεια και την χρονική στιγμή που έφτασε ένα σωματίδιο. Αυτή η μέτρηση πραγματοποιείται χρησιμοποιώντας εναλλασσόμενα στρώματα από απορροφητές και σπινθηριστές φθορισμού, δηλαδή υλικά τα οποία παράγουν έναν ταχύ παλμό φωτός όταν ένα σωματίδιο τα διαπερνά. Ειδικές οπτικές ίνες μαζεύουν το φως αυτό και το μεταφέρουν στα ηλεκτρονικά ανάγνωσης όπου εκεί υπάρχουν φωτοανιχνευτές που ενισχύουν το σήμα. Το ποσό του φωτός σε μία συγκεκριμένη

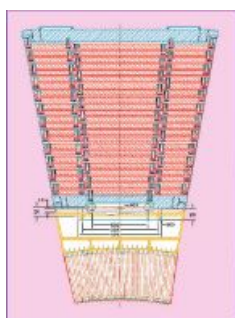


Figure 3.12: HCAL module showing sampling layers

περιοχή αθροίζεται πάνω στα διάφορα στρώματα από όλες τις περιοχές βάθους του ανιχνευτή, σχηματίζοντας έτσι "πύργο" (tower). Το συνολικό ποσό που ανιχνεύεται είναι μία μέτρηση για την ενέργεια του εισερχόμενου σωματιδίου.

Καθώς τα σωματίδια διασπώνται, μπορεί να παραχθούν νέα σωματίδια τα οποία δεν αφήνουν κανένα ίχνος για την ύπαρξη τους μέσα στον ανιχνευτή του CMS. Για να ανιχνευθούν αυτά τα σωματίδια, πρέπει ο HCAL να μπορεί να "αιχμαλωτίσει" κάθε σωματίδιο που προέρχεται από τις συγκρούσεις. Με αυτό τον τρόπο, αν δούμε σωματίδια να βγαίνουν από τη μία πλευρά του ανιχνευτή και όχι από την αντιδιαμετρική, τότε υπάρχει ανισοροπία στην ενέργεια και την ορμή και συμπαιρένεται ότι παράγονται "αόρατα" σωματίδια.

Για να μπορέσει ο ανιχνευτής HCAL να συγκρατήσει και να μετρήσει όλα τα σωματίδια που περνάνε από εκείνον, πρέπει να υπάρχει τουλάχιστον 1 m ενεργό υλικό. Για να επιτευχθεί αυτό το χαρακτηριστικό, ο HCAL είναι σχεδιασμένος σε τρία κομμάτια: Βαρέλι (Barrel HB, HO), καπάκια (endcaps HE) και εμπρόσθιο (forward HF). Υπάρχουν 3 "σφήνες" στο βαρέλι, η καθεμία από τις οποίες ζυγίζει 26 τόνους. Αυτές σχηματίζουν το τελευταίο στρώμα από ανιχνευτή μέσα στο μέσα στο μαγνήτη. Η έξω πλευρά του βαρελιού (HO) βρίσκεται έξω από το μαγνήτη, εξασφαλίζοντας ότι δεν θα υπάρξουν διαρροές ενέργειας. Με έναν παρόμοιο τρόπο, 3 "σφήνες" (wedges) στα καπάκια μετρούν την ενέργεια των σωματιδίων καθώς αυτά περνάνε μέσα από τον σοληνωειδή μαγνήτη.

Το βαρέλι και τα καπάκια του CMS είναι καλορίμετρα δειγματοληπτικά και είναι φτιαγμένα από επαναλαμβανόμενα στρώματα από πυκνό απορροφητή και πλαστικούς σπινθηριστές. Όταν ένα αδρόνιο χτυπήσει μία πλάκα απορροφητή (ορείχαλκος ή ατσάλι), μία αλληλεπίδραση μπορεί να προκύψει, παράγοντας έτσι διάφορα δευτερεύοντα σωματίδια που οδηγούν στη δημιουργία αδρονικού καταιγισμού. Καθώς ο καταιγισμός αυτός εξελίσσεται μέσα στον ανιχνευτή, τα σωματίδια περνούν μέσα από τα διάφορα στρώματα του ενεργού σπινθηριστή, και έτσι οι σπινθηριστές εκπέμπουν μπλε-ιώδες φως. Μέσα σε κάθε στρώση του ανιχνευτή, υπάρχουν οπτικές ίνες με διάμετρο μικρότερης του 1mm, οι οποίες απορροφούν αυτό το φως. Αυτές μετατρέπουν το

μπλε-ιώδες φως σε πράσινο φως, και νέα διαυγή οπτικά καλώδια μεταφέρουν το πράσινο φως στα ηλεκτρονικά ανάγνωσης που βρίσκονται σε συγκεκριμένα σημεία μέσα στον όγκο του HCAL.

Τα οπτικά σήματα καταφθάνουν στα ηλεκτρονικά ανάγνωσης από τα Megatiles. Megatile είναι ένα στρώμα από μικρότερα στρώματα (απορροφητές και σπινθηριστές) και το μέγεθός τους εξαρτάται τόσο από τη θέση τους στο χώρο καθώς και από τον προσανατολισμό τους σε σχέση με την αρχική σύγκρουση. Στα ηλεκτρονικά ανάγνωσης, φθάνουν όλα τα σήματα το ένα πίσω από το άλλο, και αθροίζονται οπτικά έτσι ώστε να δημιουργήσουν τους ονομαζόμενους πύργους "tower". Η ολοκληρωμένη άθροιση πάνω στα οπτικά σήματα, δίνει τελικά το μονοπάτι που ακολουθεί το σωματίδιο μέσα στον HCAL και επίσης μέσα από την οπτική άθροιση προκύπτει μία μέτρηση για την ενέργεια και το είδος του σωματιδίου.

3.2.4 Ο Μαγνήτης

Ένα από τα πιο δύσκολα και σημαντικότερα θέματα στην κατασκευή του ανιχνευτή CMS ήταν η επιλογή του μαγνήτη, το είδος καθώς και η ισχύς του. Τελικώς στο CMS επιλέχθηκε η χρήση ενός υπεραγωγίμου σωληνοειδούς ($L = 13\text{m}$) παράλληλα στον άξονα της δέσμης ο οποίος παράγει μαγνητικό πεδίο 4T. Ο μαγνήτης είναι τοποθετημένος έτσι, ώστε τα 3 ανιχνευτικά συστήματα του CMS είναι τοποθετημένα στο εσωτερικό του. Ο ρόλος του μαγνήτη είναι η καμπύλωση των τροχιών των σωματιδίων έτσι ώστε να μετρηθεί με ακρίβεια η ορμή τους. Το σωληνοειδές του CMS καλύπτει μία περιοχή $|\eta| < 1.5$

3.2.5 Ανιχνευτές Μιονίων

Το πείραμα CMS είναι εξοπλισμένο με ανιχνευτές μιονίων τόσο στο βαρέλι καθώς και στα καπάκια της διάταξης. Μέσα στο βαρέλι υπάρχουν 250 Drift Tubes και 480 resistive plate chambers. Στα καπάκια του ανιχνευτή υπάρχουν 432 Resistive Plate Chambers και 468 Cathode Strips. Το μιονικό σύστημα, αποτελείται από 4 σταθμούς και καλύπτει συνολικά μια έκταση 25000 m^2 ενεργής περιοχής. Το μιονικό σύστημα καλύπτει μία περιοχή $4 < r < 7.4$ και $|\eta| < 2.4$

Η τεχνολογία λειτουργίας πίσω από κάθε ανιχνευτή είναι απλή. Όλοι οι ανιχνευτές που χρησιμοποιούνται για την ανίχνευση μιονίων, είναι ανιχνευτές αερίου. Σε έναν ανιχνευτή αερίου, όταν ένα φορτισμένο σωματίδιο περνά (εδώ μιονίο), έχει αρκετή ενέργεια έτσι ώστε να ιονίσει τα άτομα του αερίου. Αυτό σημαίνει ότι ενέργεια εναποτίθεται στα ηλεκτρόνια σθένους των ατόμων του αερίου και συνεπώς εκείνα ξεφεύγουν από το άτομο. Τα ηλεκτρόνια που ξεφεύγουν από τα άτομα του αερίου έχουν αρκετή ενέργεια έτσι ώστε να δώσουν περαιτέρω ενέργεια και σε άλλα ηλεκτρόνια σθένους διαφορετικών ατόμων. Έτσι δημιουργείται μία χιονοστιβάδα ηλεκτρονίων. Χρησιμοποιώντας μία άνοδο (θετικά φορτισμένο καλώδιο) και μία κάθοδο

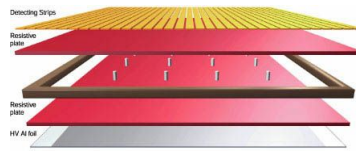


Figure 3.13: Resistive Plate Chambers Layers

(αρνητικά φορτισμένο καλώδιο) η χιονοστιβάδα ηλεκτρονίων που έχουν φύγει από τα άτομα του αερίου επιταχύνεται προς την άνοδο (λόγω του ηλεκτρικού πεδίου που δημιουργείται από άνοδο και κάθοδο).

Resistive Plate Chambers

Οι Resistive plate chambers είναι ανιχνευτές αερίου που χρησιμοποιούνται κυρίως για σκανδαλισμό. Οι ανιχνευτές αυτοί έχουν καλή χωρική διακριτική ικανότητα $\frac{\sigma}{p_T} = 1\% @ 50 GeV$ και $10\% @ 1 TeV$ χρονική διακριτική ικανότητα της τάξης των ns.

Κάθε μονάδα του ανιχνευτή αποτελείται από δύο παράλληλες πλάκες. Κάθε πλάκα είναι φορτισμένη θετικά και αρνητικά αντίστοιχα. Και οι δύο πλάκες είναι κατασκευασμένες από υψηλής αντίστασης πλαστικό υλικό. Οι πλάκες είναι χωρισμένες από όγκο που είναι γεμάτος το αέριο που χρησιμοποιεί ο ανιχνευτής.

Όταν ένα μίονιο περνά, δίνει αρκετή ενέργεια στα ηλεκτρόνια σθένους με αποτέλεσμα τα ηλεκτρόνια να ελευθερώνονται από τα άτομα του αερίου. Τα ηλεκτρόνια αυτά χτυπούν με τη σειρά τους και άλλα άτομα κάνοντας ακριβώς την ίδια διαδικασία αφού έχουν αρκετή ενέργεια για περεταίρω ιονισμό. Έτσι δημιουργείται μία χιονοστιβάδα ηλεκτρονίων που επιταχύνεται προς την άνοδο. Το σήμα συλλέγεται από μικρά μεταλλικά στριπς μετά από μία γνωστή χρονική καθυστέρηση. Ο ανιχνευτής χρησιμοποιείται επίσης σαν σκανδαλιστής δίνοντας πληροφορίες για το εάν ένα σήμα είναι χρήσιμο ή όχι.

Muon Drift Tubes

Οι ανιχνευτές Drift Tubes είναι τοποθετημένοι στο κύριο σώμα του CMS, στο barrel, όπου η ροή σωματιδίων είναι χαμηλή και το μαγνητικό πεδίο ομογενές και όχι τόσο ισχυρό. Το σύστημα των drift tubes αποτελείται από 5 κυλίνδρους ομόκεντρους με τον άξονα της δέσμης που καλύπτουν όλο το barrel. Κάθε δίσκος αποτελείται από 4 stations διαφορετική ακτίνας το καθένα (MB1 - MB4). Κάθε σωλήνας (tube) έχει διάμετρο 4 εκατοστά και στο κέντρο του υπάρχει ένα σύρμα και μείγμα αερίων αργού και διοξειδίου του άνθρακα. Όταν ένα μίονιο περάσει μέσα από το αέριο, ιονίζει τα μόρια του αερίου. Τα ηλεκτρόνια που δημιουργούνται από τον ιονισμό κατευθύνονται προς το σύρμα λόγω του ηλεκτρικού πεδίου που υπάρχει στο εσωτερικό του σωλήνα.

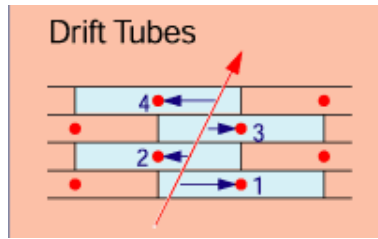


Figure 3.14: Muon Drift Tubes

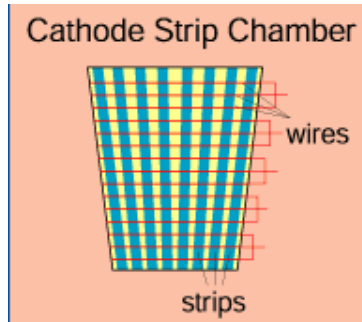


Figure 3.15: Cathode Strip Chambers

Cathode Strip Chambers

Οι ανιχνευτές είναι τοποθετημένοι στα καπάκια του CMS, όπου το μαγνητικό πεδίο είναι ανομοιογενές και η ροή των σωματιδίων είναι αρκετά υψηλή. Οι CSC έχουν πολύ μικρό χρόνο απόκρισης και έτσι μπορούν να χρησιμοποιηθούν σε περιβάλλον με τόσο μεγάλη ακτινοβολία. Επίσης έχουν τη δυνατότητα να παρέχουν πληροφορίες σχετικά με το χρόνο διέλευσης του σωματιδίου

Η αρχή λειτουργίας τους βασίζεται σε ένα πλέγμα συρμάτων. Διαθέτουν θετικά φορτισμένα σύρματα τα οποία λειτουργούν σαν ανόδοι και αρνητικά φορτισμένα κάθεται στις ανόδους που λειτουργούν σαν κάθοδοι. Το πλέγμα βρίσκεται μέσα σε ένα μείγμα αερίου. Όταν ένα φορτισμένο σωματίδιο περάσει μέσα από την ενεργή περιοχή του ανιχνευτή τα μόρια του αερίου ιονίζονται και τα ηλεκτρόνια που παράγονται από τον ιονισμό οδηγούνται προς τις καθόδους δημιουργώντας χιονοστιβάδα. Το σήμα που παράγεται από την χιονοστιβάδα δίνει πληροφορία για την τροχιά του σωματιδίου που πέρασε.

3.2.6 Σύστημα Σκανδαλισμού (Trigger)

Κατά τη διάρκεια λειτουργίας του LHC οι ανιχνευτές καταγράφουν γεγονότα με συχνότητα 10^9 Hz. Αυτό σημαίνει ότι κατά τη διαδικασία αυτή κάθε ανιχνευτής θα παράγει περίπου 100 TeraByte δεδομένων κάθε δευτερόλεπτο. Η συνολική ποσότητα

αυτής της πληροφορίας είναι αδύνατο να αποθηκευτεί διότι δεν υπάρχουν τα μέσα για την αποθήκευση τόσο μεγάλου όγκου δεδομένων αλλά και γιατί ένα αρκετά μικρό κομμάτι των γεγονότων αυτών έχουν φυσικό ενδιαφέρον έτσι ώστε να χρησιμοποιηθούν σε περαιτέρω ανάλυση. Για τους παραπάνω λόγους έχει αναπτυχθεί ένα σύστημα σκανδαλισμού έτσι ώστε να μειωθεί ο όγκος των δεδομένων που λαμβάνονται και να συγκρατούνται μόνο δεδομένα με φυσικό ενδιαφέρον. Το σύστημα σκανδαλισμού που χρησιμοποιείται στο CMS ονομάζεται **Trigger and Data Acquisition System (TRIDAS)**. Στο CMS η διαδικασία αυτή γίνεται σε 2 στάδια. Πρώτα τα δεδομένα περνούν από τον Level 1 Trigger και στη συνέχεια από τον High Level Trigger (HLT).

Ο Level 1 Trigger είναι το πρώτο στάδιο του συστήματος σκανδαλισμού του CMS. Ο συνολικός χρόνος που απαιτείται για συλλογή και ανάλυση των δεδομένων είναι περίπου 3.2μs. Τα δεδομένα αποθηκεύονται σε buffers μέχρι να αποφασίσει ο trigger για το αν είναι χρήσιμα ή όχι. Ο χρόνος που απαιτείται ώστε ο trigger να "αποφασίσει" αν τα δεδομένα είναι ενδιαφέροντα ή όχι είναι 1μs. Ο Level 1 Trigger αποτελείται αποκλειστικά από hardware το οποίο βρίσκεται τοποθετημένο όσο το δυνατό πιο κοντά στον ανιχνευτή έτσι ώστε να μην υπάρχει καθυστέρηση της πληροφορίας από τον ανιχνευτή στον σκανδαλιστή. Εξαιτίας του πολύ μικρού χρόνου που έχει στη διάθεση του το σύστημα αυτό, τα δεδομένα τα οποία χρησιμοποιεί προέρχονται αποκλειστικά από τα θερμιδόμετρα και το σύστημα μιονίων. Μετά το στάδιο αυτό ο συνολικός όγκος των δεδομένων έχει μειωθεί σε 50 kHz και τα δεδομένα που διαλέγονται μεταφέρονται στον HLT για περαιτέρω ανάλυση.

Ο High Level Trigger αποτελείται αποκλειστικά από software. Το σύστημα αυτό χρησιμοποιεί αλγόριθμους οι οποίοι φιλτράρουν τα γεγονότα σε 3 στάδια. Στο πρώτο στάδιο, χρησιμοποιούνται δεδομένα μόνο από τα θερμιδόμετρα και το μιονικό σύστημα ανάλογα με τον Level 1 Trigger αλλά με περισσότερη λεπτομέρεια. Στο δεύτερο στάδιο χρησιμοποιούνται δεδομένα και από τον tracker. Στο τρίτο και τελευταίο στάδιο, χρησιμοποιείται όλη η πληροφορία και γίνεται ανακτασσκευή ολόκληρου του γεγονότος. Μόλις η διαδικασία του triggering έχει ολοκληρωθεί τα δεδομένα θα έχουν μειωθεί σε 100 Hz. Τα γεγονότα που έχουν περάσει και τα 2 στάδια Triggering χρησιμοποιούνται έπειτα για περαιτέρω offline ανάλυση.

Κεφάλαιο 4

Ανακατασκευή Φυσικών Αντικειμένων

Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, το πείραμα του CMS παράγει έναν πολύ μεγάλο αριθμό δεδομένων που προέρχεται από τους υποανιχνευτές του. Ο μεγάλος αυτός αριθμός από "χτυπήματα" (hits) και εναποθέσεις ενέργειας σε πολλαπλά στρώματα χρησιμοποιώντας κάθε φορά διαφορετική τεχνολογία για την ταυτοποίηση και ανίχνευση των σωματιδίων, πρέπει να μεταφραστεί σε αντικείμενα χρήσιμα για την περαιτέρω φυσική ανάλυση. Η διαδικασία επανακατασκευής των δεδομένων εξελίσσεται σε 3 βήματα:

- **Τοπική Ανακατασκευή (Local Reconstruction):** Εδώ τα πρωταρχικά δεδομένα όπως οι εναποθέσεις ενέργειας σε ένα θερμιδόμετρο περνούν από προεπεξεργασία μέσα από βασικά clusters.
- **Περιφερειακή Ανακατασκευή (Regional Reconstruction):** Σε αυτό το βήμα συνδυάζονται πληροφορίες από διάφορα subdetectors του ίδιου είδους έτσι ώστε να δημιουργηθούν super clusters μέσα στο θερμιδόμετρο.
- **Καθολική Ανακατασκευή (Global Reconstruction):** Συνδυάζοντας πληροφορίες από όλους τους διαθέσιμους υποανιχνευτές όπως τον ανιχνευτή τροχιών και το θερμιδόμετρο, η διαδικασία καθολικής επανακατασκευής τελικά επιστρέφει τα φυσικά αντικείμενα που έχουν ενδιαφέρον.

4.1 Ανακατασκευή Αδρονικών Πιδάκων

Η ανακατασκευή και ταυτοποίηση των σωματιδίων των οποίων οι τελικές καταστάσεις περιέχουν κουάρκ παραμένει ακόμα και σήμερα ένα πολύ δύσκολο έργο της φυσικής των υψηλών ενεργειών.

Για την φυσική του top κουάρκ στο ημιλεπτονικό κανάλι, η ανακατασκευή των αδρονίων είναι πολύ σημαντική αφού ένα από τα top κουάρκ σε αυτό το κανάλι διασπάται σε ένα b κουάρκ και ένα W μποζόνιο, το οποίο με τη σειρά του διασπάται σε 2 κουάρκ. Τα 3 αυτά κουάρκ χρησιμοποιούνται για την ταυτοποίηση της μάζας του top κουάρκ. Συνεπώς, η ενέργεια του κάθε jet καθώς και η χωρική μέτρηση είναι ένα από τα σημαντικότερα προβλήματα που εμποδίζουν την ακρίβεια με την οποία μπορεί να μετρηθεί η μάζα του top Κουάρκ στο κανάλι αυτό.

Τα κουάρκ που παράγονται από ανελαστική σκέδαση δεν μπορούν να υπάρξουν ως ελεύθερα σωματίδια λόγω περιορισμών QCD. Αυτό που μπορεί να μετρηθεί σε έναν ανιχνευτή είναι το αποτέλεσμα από μία διαδικασία μετατροπής coloured partons σε ένα καταιγισμό από colourless partons.

Λόγω του γεγονότος ότι το σύνολο των σωματιδίων δεν είναι συγγραμικά, το "αποτύπωμα" των κουάρκ και των γλουονίων ονομάζεται jet. Θα μπορούσαμε να πούμε ότι, τα jet είναι τοπικά clusters ενέργειας συνοδευόμενα από τροχιές σωματιδίων. Ένα jet περιέχει κυρίως σωματίδια όπως πιόνια, καόνια, ακόμα και πρωτόνια, νετρόνια.

Τα jet χωρίζονται σε δύο κατηγορίες:

- **Prompt:** δηλαδή jet Που προέρχονται από την κύρια σύγκρουση
- **Pile Up:** δηλαδή jet Που προέρχονται από δευτερεύουσες συγκρούσεις (ίδιο Bunch crossing)

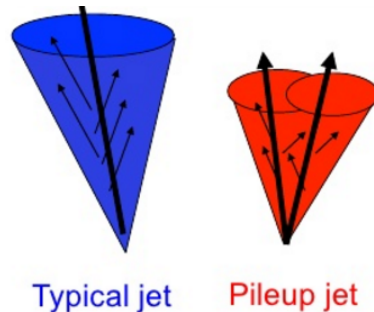


Figure 4.1: Κατηγορίες Jet

Η ανακατασκευή των jet είναι ένα πολύ κρίσιμο κομμάτι της ανάλυσης των αδρονικών τελικών καταστάσεων (hadronic final states). Λόγω της διαδικασίας καταιγισμού, μικρές αλλαγές κατά την αρχή της αδρανοποίησης (hadronization process) μπορούν να οδηγήσουν σε αρκετά μεγάλες αποκλίσεις με αυτό που θα μετρηθεί στο τέλος. Επίσης τα τελικά αποτελέσματα επηρεάζονται και λόγω ηλεκτρονικού θορύβου κάνοντας τη διαδικασία ταυτοποίησης ακόμα δυσκολότερη. Για την αντιμετώπιση των προβλημάτων αυτών, έχουν δημιουργηθεί διάφοροι αλγόριθμοι για τη ταυτοποίηση και μελέτη των Jet.

4.1.1 Ανακατασκευή Jet

Η ανακατασκευή των jet ξεκινά διαβάζοντας αρχικά τα δεδομένα που έρχονται από τα θερμιδόμετρα (ECAL και HCAL). Και τα δύο θερμιδόμετρα, δίνουν πληροφορίες που χρησιμοποιούνται ως είσοδος στην διαδικασία ανακατασκευής των jet. Η διαδικασία εξελίσσεται σε τέσσερα στάδια:

1. Πρώτα όρια καταστολής μηδενισμού (First zero suppression thresholds) εφαρμόζονται σε κάθε κανάλι από κάθε θερμιδόμετρο. Για το HCal το όριο αυτό είναι από 2-4 ADC "χτυπήματα" με μετρούμενο θόρυβο στα 0.75 ADC "χτυπήματα" (counts). Για το ECal, η ελάχιστη ενέργεια που απαιτείται είναι 90MeV στο βαρέλι και 450MeV στα καπάκια του ανιχνευτή. Αυτή η ενέργεια αντιστοιχεί σε 2-3σ πάνω από το μετρούμενο θόρυβο.
2. Οι ξεχωριστές πληροφορίες που λαμβάνονται από τα θερμιδόμετρα ECal και HCal, συνδυάζονται σε ένα projective ECalPlusHCalTowers. Οι "πύργοι" αυτοί είναι ένας συνδυασμός από 1 HCal και 3x3 ECal κελιά. Εδώ επιβάλλεται ένα cut στην διαμήκη ενέργεια του κάθε "πύργου" στα $E_T = 0.5\text{GeV}$ το οποίο βοηθά στην καταστολή συλλογής ενέργειας από Pile Up και Underlying events στο θερμιδόμετρο.
3. Στο τρίτο βήμα, η λίστα των ECalPlusHCalTowers χρησιμοποιείται ως είσοδος στον πραγματικό αλγόριθμο για την ανακατασκευή των jet, ο οποίος τελικά συγκεντρώνει όλη την ενέργεια που βρίσκεται στα jet, χρησιμοποιώντας μία συγκεκριμένη φόρμουλα. Έπειτα ένα cut στη διαμήκη ενέργεια κοντά στα 10GeV εφαρμόζεται στα Jet καθώς τα jet με χαμηλές ενέργειες δεν είναι καλά ορισμένα.
4. Τελικά τα ανακατασκευασμένα jet βαθμονομούνται με τη βοήθεια καλά ορισμένων φυσικών γεγονότων. Εδώ διαφορετικές πιθανότητες προκύπτουν για να λάβει κανείς την ακριβή ενέργεια του jet. Κάποιος διακρίνει μεταξύ εξωτερικών (πχ test beams) και εσωτερικών (πχ In situ) διαδικασιών βαθμονόμησης με δεδομένα. Η in situ βαθμονόμηση εφαρμόζεται ανακατασκευάζοντας δύο από τα τρία jet που προέρχονται από τη διάσπαση του W με τέτοιο τρόπο ώστε η συνολική ενέργεια που μετράται στα 2 jet να είναι ίσης με αυτή της γνωστής αναλλοίωτης μάζας του W ($80.22\text{GeV}/c^2$).

4.1.2 Ο αλγόριθμος anti- k_T

Για την ανακατασκευή των jet που παράγονται κατά τη διάρκεια συγκρούσεων πρωτονίων-πρωτονίων χρησιμοποιούνται δύο κατηγορίες αλγορίθμων. Η πρώτη κατηγορία

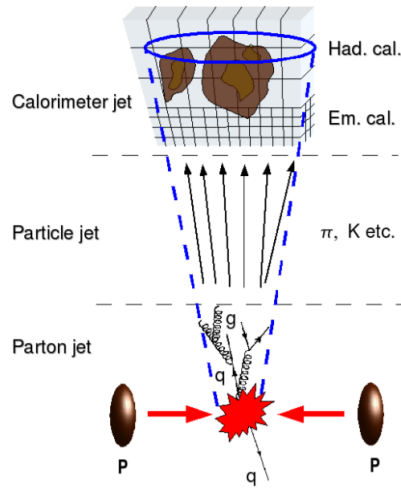


Figure 4.2

είναι οι αλγόριθμοι Cone Type, ενώ η δεύτερη κατηγορία είναι οι αλγόριθμοι Sequential Clustering. Σε αυτό το υποκεφάλαιο θα γίνει μία εισαγωγή στην δεύτερη κατηγορία και πιο συγκεκριμένα θα γίνει εστίαση στον αλγόριθμο anti- k_T [6]. Γενικά υπάρχουν άλλοι δύο αλγόριθμοι που ανήκουν στην κατηγορία αυτή: Ο αλγόριθμος k_T και ο αλγόριθμος Cambridge/Aachen.

Οι αλγόριθμοι αυτοί έχουν κάποιες ιδιότητες που τους χαρακτηρίζουν. Αρχικά οι αλγόριθμοι αυτοί ξεκινούν με την υπόθεση ότι τα σωματίδια μέσα στα jet έχουν μικρές διαφορές στη διαμήκη ορμή τους (transverse momentum) και τα σωματίδια ομαδοποιούνται με βάση τον χώρο των ορμών. Αυτό έχει σαν αποτέλεσμα τη δημιουργία περιοχών jet οι οποίες έχουν κυμαινόμενη επιφάνεια. Ενώ αρχικά οι αλγόριθμοι αυτοί είχαν αρκετά αργή απόδοση, ήταν όμως αποτελεσματικοί. Πλέον χρησιμοποιείται το πρόγραμμα FastJet μέσω του οποίου ο υπολογισμός των jet γίνεται πολύ γρηγορότερα. Τέλος οι αλγόριθμοι Sequential Clustering είναι όλοι IRC Safe. Με τον όρο IRC Safe, ορίζεται ότι ένα γεγονός θα είναι infrared και collinear safe εάν στο όριο μίας διάσπασης ή κατά την εκπομπή ενός soft σωματιδίου, το παρατηρούμενο γεγονός παραμένει αμετάβλητο. Μέσω της ιδιότητας αυτής, ο αλγόριθμος εγγυάται ακύρωση πραγματικών και εικονικών αποκλίσεων σε υπολογισμούς υψηλότερων όρων (higher order calculations). Για τον υπολογισμό ενός jet σε διαταραχτική (perturbative) QCD ο αλγόριθμος που πρέπει να χρησιμοποιηθεί πρέπει να είναι IRC Safe αφού εκπομπή soft σωματιδίων καθώς και collinear διασπάσεις δεν πρέπει να αλλάζουν τα hard jets.

- $O(X; p_1, \dots, p_n, p_{n+1} \rightarrow 0) \rightarrow O(X; p_1, \dots, p_n)$
- $O(X; p_1, \dots, p_n || p_{n+1}) \rightarrow O(X; p_1, \dots, p_n + p_{n+1})$

Γενικά όλοι οι αλγόριθμοι αυτού του τύπου λειτουργούν με παρόμοια μέθοδο. Αρχικά ορίζουμε ως μεταβλητή πρώτης απόστασης (first distance variable):

$$d_{ij} = \min(p_{ti}^a, p_{tj}^a) \times \frac{R_{ij}^2}{R}$$

όπου a είναι ένας εκθέτης που αντιστοιχεί σε έναν αλγόριθμο, $R_{ij} = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ και είναι η απόσταση των σωματιδίων i, j στον (η, ϕ) χώρο, R είναι η παράμετρος ακτίνας η οποία καθορίζει το τελικό μέγεθος του jet (συνήθως 0.4-0.7). Ορίζουμε ως μεταβλητή δεύτερης απόστασης (second distance variable):

$$d_{iB} = p_{ti}^a$$

είναι δηλαδή η απόσταση στον χώρο των ορμών μεταξύ της δέσμης και του σωματιδίου που ανιχνεύθηκε.

Οι αλγόριθμοι Sequential Clustering λειτουργούν βρίσκοντας πρώτα την ελάχιστη τιμή του ζεύγους $[d_{ij}, d_{iB}]$. Αν το d_{ij} είναι το ελάχιστο από τα δύο, τότε τα i, j σωματίδια συγχωνεύονται σε 1 σωματίδιο χρησιμοποιώντας πρόσθεση των τετραορμών τους. Αν όμως το ελάχιστο είναι το d_{iB} , τότε το i σωματίδιο τίθεται ως τελικό jet και αφαιρείται από την λίστα των σωματιδίων. Η διαδικασία αυτή επαναλαμβάνεται είτε μέχρι όλα τα σωματίδια να είναι μέρος ενός jet (Inclusive clustering) είτε μέχρι να έχει βρεθεί μία επιθυμητή ποσότητα από jets (Exclusive clustering).

Ο αλγόριθμος που θα χρησιμοποιηθεί στην διπλωματική αυτή είναι της κατηγορίας Sequential Clustering και πιο συγκεκριμένα ο αλγόριθμος *anti* - k_T . Ο αλγόριθμος *anti*- k_T θέτει την παράμετρο $a = -2$. Αυτό έχει ως αποτέλεσμα ότι οι μεταβλητή πρώτης και δεύτερης απόστασης θα γίνουν αντίστοιχα:

$$d_{ij} = \min\left(\frac{1}{p_{ti}^2}, \frac{1}{p_{tj}^2}\right) \times \frac{R_{ij}^2}{R}$$

$$d_{iB} = \frac{1}{p_{ti}^2}$$

Η μεταβλητή πρώτης απόστασης κυριαρχείται από υψηλά p_T και ο αλγόριθμος προτιμά να συλλέγει (cluster) hard σωματίδια. Έτσι η περιοχή του jet έχει μικρές διακυμάνσεις και ο αλγόριθμος δεν είναι ευαίσθητος σε UE και PU. Ο αλγόριθμος είναι ο καλύτερος για την ανάλυση των Jet αλλά όχι για την υποδομή τους.

4.1.3 Ενεργειακή Βαθμονόμηση Jet

Η ενεργειακή βαθμονόμηση των jet είναι ένα πολύ σημαντικό σημείο κάθε ανάλυσης που χρησιμοποιεί τελικές αδρονικές καταστάσεις για την ανακατασκευή κινηματικών ιδιοτήτων. Η απόκριση του HCal ανιχνευτή είναι ανάλογη της ενέργειας που εναποτίθεται από τα σωματίδια σύμπτωσης. Όμως η συσχέτιση μεταξύ της ενέργειας

του σωματιδίου και του σήματος που επιστρέφει ο ανιχνευτής πρέπει να ρυθμιστεί πειραματικά. Αρχικά πραγματοποιούνται δέσμες δοκιμής (test beams) με γνωστά σωματίδια και γνωστές ενέργειες για την βαθμονόμηση των ανιχνευτών. Αργότερα κατά τη διάρκεια του τρέχοντος πειράματος, χρησιμοποιούνται αρκετά μελετημένες και γνωστές φυσικές διεργασίες για τη βαθμονόμηση των ανιχνευτών. Για παράδειγμα για την βαθμονόμηση του ECal χρησιμοποιείται η διάσπαση $Z \rightarrow e^+e^-$ ενώ για τη βαθμονόμηση του HCal χρησιμοποιούνται γ +jets γεγονότα. Η ιδέα της γ +jet βαθμονόμησης είναι η παρακάτω:

Στην leading order εικόνα ενός γ +jet γεγονότος, υπάρχουν δύο παρτόνια (partons) (qg ή $q\bar{q}$) τα οποία συγκρούονται μεταξύ τους. Αυτή η διαδικασία μπορεί να παράξει ένα φωτόνιο και ένα παρτόνιο (στη συνέχεια jet) όπως φαίνεται παρακάτω.

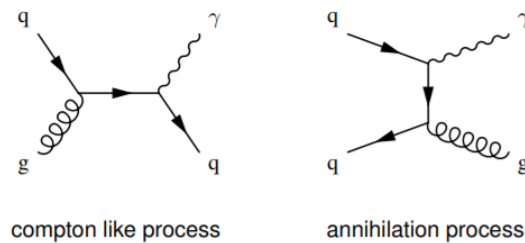


Figure 4.3: Διαγράμματα Feynman για γ +jet διαδικασίες

Η ιδέα για χρησιμοποίηση των γεγονότων αυτών για τη βαθμονόμηση του αδρονικού θερμιδόμετρου βασίζεται στο γεγονός, ότι εκτός από την αρχική κατάσταση εγκάρσια ορμής, η εγκάρσια ορμή (p_T) του φωτονίου που παράγεται κατά τη διαδικασία είναι κατά βάση αντίθετη στο ανατρέπον (recoiling) jet και έτσι:

$$\vec{p}_T^{\gamma+jet} = \vec{p}_T^{\gamma} + \vec{p}_T^{jet} \approx \vec{0}$$

Λόγω της καλύτερης ενεργειακής διακριτικής ικανότητας του ECal (1%) και λόγω της ακριβούς απόλυτης ενεργειακής κλίμακας, η ενέργεια του jet μπορεί να διορθωθεί γνωρίζοντας την ενέργεια που εναποθέτησε το φωτόνιο μέσα στο ECal. Στο τέλος, το ανακατασκευασμένο jet διορθώνεται. Έχει παρατηρηθεί ότι με αυτή τη μέθοδο, η ενέργεια του jet μπορεί να βρεθεί με ένα σφάλμα της τάξης του 3-10% αναλόγως με την εγκάρσια ενέργεια του jet.

4.1.4 Ενεργειακή Διόρθωση Jet

Το CMS έχει αναπτύξει ένα σύστημα διόρθωσης των jet πολλαπλών επιπέδων. Οι διορθωτικοί παράγοντες λαμβάνονται από προσομοιώσεις βασισμένες σε δεδομένα από test beams, από δεδομένα που έρχονται απευθείας από συγκρούσεις και από

προσομοιώσεις συντονισμένες με δεδομένα συγκρούσεων. Οι διορθώσεις αυτές πρέπει να εφαρμοστούν με τον ακόλουθο τρόπο:

- **Offset:** Διόρθωση λόγω pile-up και ηλεκτρονικού θορύβου
- **Σχετική η:** Απαιτούμενη διόρθωση ώστε να γίνει η απόκριση του jet επίπεδη σε συνάρτηση του η_{jet} για μία συγκεκριμένη p_T .
- **Απόλυτη p_T :** Απαιτούμενη διόρθωση για να γίνει η απόκριση του jet επίπεδη σε συνάρτηση του jet p_T

4.1.5 Αναγνωρίζοντας Jet από b κουαρκ

Επειδή το top (t) κουαρκ διασπάται σχεδόν μόνο σε ένα b κουαρκ και ένα W μποζόνιο, η ταυτοποίηση των b jet παίζει πολύ σημαντικό ρόλο στην διάκριση μεταξύ σήματος και υποβάθρου. Επίσης, η αποτελεσματική ταυτοποίηση των b jet βοηθά στον συνδυασμό των 3 jet για την ταυτοποίηση της διάσπασης του top κουαρκ.

Μία σημαντική ιδιότητα των b jets είναι ότι ένα μέρος αυτών περιέχουν μη απομονωμένα λεπτόνια τα οποία προέρχονται από την λεπτονική διάσπαση του W μποζονίου.

$$W \rightarrow l^- \bar{\nu}_l$$

Τα λεπτόνια αυτά που βρίσκονται μέσα στα b-jets μπορούν να προέλθουν από τρεις συγκεκριμένες διαδικασίες. Οι διαδικασίες αυτές και οι ρυθμοί διάσπασής τους (branching ratio, BR) δίνονται από:

- Η άμεση $b \rightarrow l$ διάσπαση: $b \rightarrow W^- X, W^- \rightarrow l^- \bar{\nu}_l$, BR $\approx 10,7\%$
- Η αλληλουχία $b \rightarrow c \rightarrow l$: $b \rightarrow W^{*-} c, c \rightarrow l^+ \nu_l X$, BR $\approx 8.0\%$
- Η αλληλουχία "λανθασμένου" προσήμου $b \rightarrow \bar{c} \rightarrow l$: $b \rightarrow W^{*-} X, W^{*-} \rightarrow q \bar{c}, \bar{c} \rightarrow l^- \bar{\nu}_l X$, BR $\approx 1.6\%$

Όλες αυτές οι διαδικασίες μαζί δίνουν έναν συνολικό ρυθμό διάσπασης για τη διάσπαση τουλάχιστον ενός φορτισμένου λεπτονίου περίπου ίση με 19.3 % για κάθε τύπο λεπτονίου.

Για το διαχωρισμό των b jets από άλλα jets που προέρχονται από ελαφριά κουαρκ ή γλουόνια, χρησιμοποιείται μία πολύ σημαντική παράμετρος των b αδρονίων. Λόγω των μικρής τιμής των συντελεστών V_{ub} και V_{cb} του πίνακα Cabibbo-Kobayashi-Maskawa, ο χρόνος ζωής των b αδρονίων είναι αντιστοιχεί σε περίπου 1.5ps και αυτό αντιστοιχεί σε μία εμβέλεια $\beta\gamma c\tau \approx 450\mu m$. Αυτή η τιμή είναι μία τάξη μεγέθους μεγαλύτερη από την χωρική διακριτική ικανότητα του ανιχνευτή τροχιάς.

Αυτό δίνει τη δυνατότητα για διαχβρισμό μεταξύ της κύριας κορυφής (primary vertex) από ανελαστική σκέδαση και της δευτερεύουσας κορυφής (secondary vertex) που προέρχεται από διάσπαση του b.

Οι τροχιές που χρησιμοποιούνται για τον b tagging αλγόριθμο είναι συνδυαστικές τροχιές (combinatorial tracks). Αυτό σημαίνει ότι χρησιμοποιούνται πληροφορίες και από τον (silicon) ανιχνευτή πυριτίου τροχιών καθώς και από τον ανιχνευτή τροχιών pixel. Μόνο τροχιές με $\Delta R_i < 0.3$ χρησιμοποιούνται και οι τροχιές πρέπει να αποτελούνται από τουλάχιστον 8 χτυπήματα (hits), δύο εκ των οποίων πρέπει να προέρχονται από τον ανιχνευτή pixel. Επιπρόσθετα κάθε τροχιά πρέπει να έχει εγκάρσια ορμή περίπου 1GeV/c. Τέλος οι χ^2 βαθμοί ελευθερίας του fit της τροχιάς πρέπει να είναι κάτω από 10 και η παράμετρος εγκάρσιας κρούσης (transverse impact parameter)¹ πρέπει να είναι μικρότερη από 2mm.

Ένα πρόβλημα που προκύπτει είναι η ανακατασκευή των κορυφών. Με τις τροχιές να περνάνε ένα πρώτο επίπεδο επιλογής, οι κορυφές ανακατασκευάζονται χρησιμοποιώντας ένα φίλτρο Kalman. Στο επόμενο βήμα, οι ανακατασκευασμένες κορυφές και οι επιλεγμένες κύριες κορυφές συγκρίνονται εφαρμόζοντας cuts στην απόστασή τους και την αναλλοίωτη μάζα των σωματιδίων που σχετίζονται με τις κορυφές. Κορυφές με αντίθετα φορτισμένες τροχιές και μία αναλλοίωτη μάζα κοντάς στο K_S^0 παραλείπονται. Για τις υπόλοιπες υποψήφιες δευτερεύουσες κορυφές υπολογίζονται διάφορες παράμετροι, οι οποίες όλες χρησιμοποιούνται σαν είσοδος για τον τελικό διευκρινιστή του b (b tagging discriminator).

1. Η αναλλοίωτη μάζα του φορτισμένου σωματιδίου που αντιστοιχεί στην κορυφή
2. Η πολυπλοκότητα των φορτισμένων τροχιών που σχετίζονται με την κορυφή
3. Η απόσταση της κύριας από την δευτερεύουσα κορυφή στο επίπεδο κάθετο στη δέσμη διαιρεμένο με το σφάλμα της
4. Η ενέργεια των φορτισμένων σωματιδίων από τις δευτερεύουσες κορυφές διαιρεμένη με την ενέργεια όλων των φορτισμένων σωματιδίων του jet
5. Οι rapidities των τροχιών των φορτισμένων σωματιδίων σε σχέση με την κατεύθυνση του Jet

Όλες οι παραπάνω μεταβλητές στη συνέχεια συνδυάζονται μέσα σε ένα Likelihood ratio όπως φαίνεται παρακάτω:

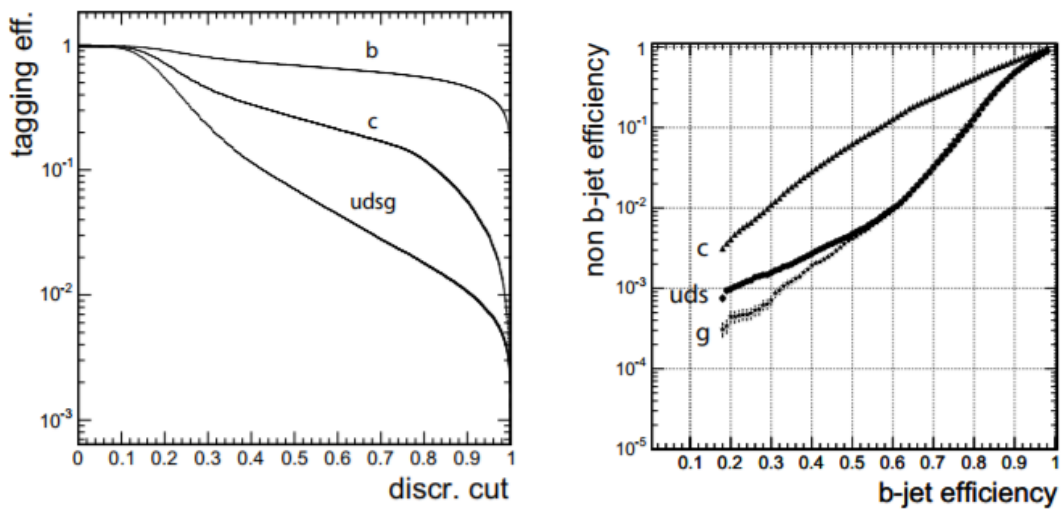
$$y = f_{BG}(c) \frac{L^b}{L^b + L^c} + f_{BG}(q) \frac{L^b}{L^b + L^q}$$

¹Ως παράμετρος εγκάρσιας κρούσης (transverse impact parameter) ορίζεται η απόσταση της κύριας κορυφής (primary vertex) από το σημείο κοντινότερης απόστασης στο εγκάρσιο επίπεδο

όπου $f_{BG}(c, q)$ αντιστοιχεί σε οποιαδήποτε αρχική πληροφορία για τη μη κατασκευή του μη b jet ενώ $L(b, c, q)$ σχετίζεται με την πιθανότητα να προέρχεται ένα jet από b, c, ...κουαρκ ή γλουόνιο και υπολογίζεται χρησιμοποιώντας τις παραπάνω μεταβλητές. Κατά συνέπεια, τα b jets έχουν υψηλότερες τιμές διευκρινιστή ενώ τα non b jets χαμηλότερες. Η αποδοτικότητα ταυτοποίησης (tagging efficiency) ϵ_q ορίζεται ως εξής:

$$\epsilon_q = \frac{\# \text{ jets γεύσης } q \text{ που έχει ταυτοποιηθεί σαν b jet}}{\# \text{ jets με γεύση } q}$$

Αυτή η αποδοτικότητα είναι η b tagging αποδοτικότητα για b jets και η αποδοτικότητα λανθασμένης αναγνώρισης για non b jets. Η b tagging αποδοτικότητα για διάφορες γεύσεις ως συνάρτηση των cut του διευκρινιστή φαίνεται στο σχήμα 4.4a. Στο σχήμα 4.4b, jet από ελαφριά κουαρκ μπορούν να κατασταλούν κατά δύο τάξεις μεγέθους με b tagging αποδοτικότητα 60% στην περιοχή του βαρελιού, ενώ με 50% στα καπάκια.



(a) b tagging αποδοτικότητα σε συνάρτηση του cut από τον διευκρινιστή για διάφορες γεύσεις jet αναγνώρισης σε ένα δείγμα με QCD δεδομένα με $30\text{GeV}/c$ και $-1 < \eta < 1.4$ χρησιμοποιώντας ημιλεπτονικά top ζεύγη με $p_T > 200\text{GeV}/c$ και $-1 < \eta < 1.4$
 (b) Non b jet αποδοτικότητα λανθασμένης αναγνώρισης σε συνάρτηση b jet tagging αποδοτικότητα για διάφορες γεύσεις jet χρησιμοποιώντας ημιλεπτονικά top ζεύγη με $p_T > 30\text{GeV}/c$ και $-1 < \eta < 1.4$

4.2 Ο Αλγόριθμος Ροής Σωματιδίων- Particle Flow Algorithm

Ο αλγόριθμος ανακατασκευής ροής σωματιδίων χρησιμοποιείται σχεδόν σε όλες τις αναλύσεις στο CMS. Έχει σκοπό την αναγνώριση και την ανακατασκευή κάθε σωματιδίου που προκύπτει από τις συγκρούσεις πρωτονίων στο κέντρο του ανιχνευτή, συνδυάζοντας όλες τις πληροφορίες που προέρχονται από τους υποανιχνευτές. Το αποτέλεσμα που προκύπτει από τον αλγόριθμο ροής σωματιδίων οδηγεί σε μία βελτιωμένη απόδοση στην ανακατασκευή των jet αλλά και στην ταυτοποίηση ηλεκτρονίων, μιονίων και ταυ σωματιδίων. Στη συνέχεια θα παρουσιαστεί ο τρόπος με τον οποίο λειτουργεί ο αλγόριθμος [7].

Ο Αλγόριθμος Ροής Σωματιδίων έχει σκοπό την αναγνώριση και ανακατασκευή όλων των σωματιδίων που προκύπτουν από τις συγκρούσεις στο κέντρο του ανιχνευτή. Αυτή η διαδικασία πραγματοποιείται χρησιμοποιώντας όλες τις πληροφορίες που προέρχονται από όλους τους υποανιχνευτές του πειράματος. Η αποτελεσματικότητα του αλγορίθμου εξαρτάται άμεσα από κάποιες συγκεκριμένες αρχές:

- Η μεγιστοποίηση της απόστασης μεταξύ φορτισμένων και ουδέτερων αδρονίων
- Θερμιδόμετρα με μεγάλη ακρίβεια

Επίσης σημαντικό για τη διαδικασία είναι ένας αποτελεσματικός ανιχνευτής τροχιάς, αλλά όχι τόσο σημαντικός όσο η ακρίβεια του θερμιδομέτρου. Το πείραμα του CMS πληρεί όλες τις παραπάνω αρχές στο έπακρο. Πιο συγκεκριμένα το field integral είναι δύο φορές μεγαλύτερο σε σχέση με άλλα πειράματα και το ηλεκτρομαγνητικό θερμιδόμετρο έχει εξαιρετική απόδοση και λεπτομέρεια.

Ο αλγόριθμος ροής σωματιδίων βασίζεται σε μία αποτελεσματική και καθαρή ανακατασκευή τροχιάς, σε ένα αλγόριθμο συμπλέγματος (cluster) που έχει την ικανότητα να αποσυναρμολογεί επικαλυπτόμενους καταγισμούς σωματιδίων (particle showers) και σε έναν αποτελεσματική διαδικασία σύνδεσης που μπορεί να συνδέσει τις εναποθέσεις ενέργειας του κάθε σωματιδίων σε κάθε υποανιχνευτή. Με λίγα λόγια ο αλγόριθμος μπορεί να περιγραφεί ως εξής; Οι τροχιές προκύπτουν μέσω των θερμιδομέτρων, αν βρεθούν μέσα στα όρια ενός από τα πολλά clusters, τα οποία σχετίζονται με την εκάστοτε τροχιά. Το σύνολο των τροχιών και των clusters αποτελούν ένα φορτισμένο αδρόνια και τα στοιχεία κατασκευής του φορτισμένου αδρονίου παύουν πλέον να χρησιμοποιούνται στον αλγόριθμο. Τα μίονια αναγνωρίζονται πιο πριν έτσι ώστε οι τροχιές τους να μη μπορούν να μπερδευτούν με ένα φορτισμένο αδρόνιο. Τα ηλεκτρόνια είναι πιο δύσκολο να τα διαχειριστεί κανείς. Λόγω της ακτινοβολίας πέδησης που εκπέμπουν τα ηλεκτρόνια, χρησιμοποιείται μία συγκεκριμένη ανακατασκευή της τροχιάς τους καθώς και μία συγκεκριμένη διαδικασία για τη σύνδεση των clusters φωτονίων με τα ηλεκτρόνια έτσι ώστε να αποφευχθεί η μέτρηση της ενέργειας δύο φορές. Όταν όλες οι παραπάνω τροχιές έχουν

ανακατασκευαστεί, τα εναπομείναντα cluster αντιστοιχούνται σε φωτόνια όσον αφορά το ηλεκτρομαγνητικό θερμιδόμετρο (ECAL) και σε ουδέτερα αδρόνια όσον αφορά το αδρονικό θερμιδόμετρο (HCAL).

Όταν όλες οι εναποθέσεις ενός σωματιδίου έχουν συσχετιστεί, μπορεί να εκτιμηθεί η φύση του, και οι πληροφορίες από τους υποανιχνευτές συνδυάζονται έτσι ώστε να υπολογιστεί άριστα η τετραορμή του. Σε περίπτωση που η βαθμονομημένη ενέργεια των θερμιδομέτρων των cluster που σχετίζονται με μία τροχιά υπερβεί την ορμή της τροχιάς για περισσότερο από ένα σίγμα (σ), τότε η διαφορά αυτή αποδίδεται σε ένα επικαλυπτόμενο ουδέτερο σωματίδιο (φωτόνιο ή αδρόνιο) που έχει ενέργεια που ανταποκρίνεται στη διαφορά των δύο μετρήσεων. Η τελική λίστα των σωματιδίων αποτελείται από φορτισμένα αδρόνια, φωτόνια, ουδέτερα αδρόνια, ηλεκτρόνια και μόνια και χρησιμοποιείται για την ανακατασκευή των jet και της ελλείπουσας εγκάρσιας ενέργειας (E_T^{miss}).

4.3 Η Φυσική των συγκρούσεων Πρωτονίων-Πρωτονίων

Στα πειράματα συγκρούσεων (collision experiments), η ενέργεια κέντρου μάζας \sqrt{s} ορίζει την ενέργεια που είναι διαθέσιμη για την παραγωγή σωματιδίων καθώς και για να αποκτήσουν τα σωματίδια αυτά κινητική ενέργεια. Στον LHC, s , καθορίζεται από την ορμή των συγκρουόμενων πρωτονίων μέσω της σχέσης:

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

όπου p_1 και p_2 είναι οι τετραορμές των δύο συγκρουόμενων πρωτονίων με ενέργειες E και τριορμή $\approx p$. Επειδή οι δύο δέσμες πρωτονίων έχουν την ίδια ενέργεια και αντίθετη φορά, η υπόθεση $E_1 = E_2 = E_{proton}$ και $p_1 \approx -p_2$ ισχύει. Με αυτό, η ενέργεια κέντρου μάζας γίνεται :

$$\sqrt{s} = 2E_{proton}.$$

Επειδή τα πρωτόνια είναι σύνθετα σωματίδια, η ακριβής διασπορά περιλαμβάνει κουαρκς και γλουόνια, δηλαδή παρτόνια (partons), που έχουν μόνο ένα πολύ μικρό ποσοστό της αρχικής ορμής της δέσμης $x p_{proton}$. Η ενέργεια κέντρου μάζας της διαδικασίας της σκληρής διασποράς $\sqrt{s'}$ (hard scattering) είναι μόνο ένα μικρό ποσοστό από τα 13TeV. Αν υποθέσει κανείς ότι οι μάζες των παρτονίων είναι μικρές σε σχέση με την ορμή τους, η ενέργεια κέντρου μάζας της διαδικασίας αυτής είναι:

$$\sqrt{s'} = \sqrt{x_1 x_2} \sqrt{s}$$

Εδώ x_1 και x_2 προσδιορίζουν το ποσοστό της ορμής που έχουν τα δύο συγκρουόμενα πρωτόνια. Επειδή αυτό το ποσοστό δεν είναι γνωστό, η ενέργεια κέντρου μάζας στη "σκληρή" διαδικασία δεν είναι επίσης γνωστή. Επομένως για να κάνει κανείς ανάλυση δεδομένων, πρέπει να χρησιμοποιήσει μεταβλητές που δεν εξαρτώνται από την αρχική ορμή στην κατεύθυνση της διαδρομής των πρωτονίων. Στους ανιχνευτές συγκρουόμενων αδρονίων, η χρήση μίας συνάρτησης πυκνότητας παρτονίων (PDF) είναι πολύ σύνηθες. Οι συναρτήσεις αυτές επιστρέφουν την πιθανότητα να βρεθεί ένα παρτόνιο μέσα σε ένα συγκεκριμένο διάστημα ορμής σε ένα πρωτόνιο. Μετρώνται σε πειράματα συγκρούσεων ηλεκτρονίων-πρωτονίων, και είναι πολύ σημαντικά για την πρόβλεψη ενεργών διατομών σε συγκρούσεις πρωτονίων-πρωτονίων.

4.4 Boosted Jets

Ως jet ορίζεται ένας διαχωρισμένος κώνος αδρονίων και άλλων σωματιδίων που παράγονται από την αδρονιοποίηση (hadronization) ενός κουάρκ ή ενός γκλουονίου.² Τα boosted jets είναι jets τα οποία έχουν εγκάρσια ορμή p_T μεγαλύτερη από 400 GeV, όπου τα προϊόντα από το top quark συγχωνεύονται σε ένα μεγάλο jet. Σκοπός της μελέτης των Boosted jets είναι η ανακατασκευή δύο μεγάλων jet τα οποία περιέχουν τα προϊόντα διάσπασης των top quarks. Τα boosted jet δίνουν πληροφορίες για τις θεωρητικές προβλέψεις σε υψηλά p_T καθώς και η μελέτη τους συμβάλει στην διερεύνηση νέων θεωρητικών μοντέλων. Για την ανακατασκευή των boosted jets χρησιμοποιούνται ειδικές τεχνικές ανακατασκευής.

Μπορούμε να χωρίσουμε τα Boosted jets σε δύο κατηγορίες. Η πρώτη κατηγορία είναι τα λεπτονικά top boosted jets. Στην κατηγορία αυτή ανήκουν τα jet, στα οποία το λεπτόνιο που προκύπτει από την διάσπαση ενός W μποζονίου βρίσκεται κοντά στο b-jet. Η δεύτερη κατηγορία jet είναι τα αδρονικά top boosted jets. Στην κατηγορία αυτή ανήκουν jet, στα οποία όλα τα προϊόντα διάσπασης του top βρίσκονται μέσα σε ένα cluster. Για τον σαφή προσδιορισμό των jet αυτών, πρέπει να μελετηθεί η υποδομή του jet.

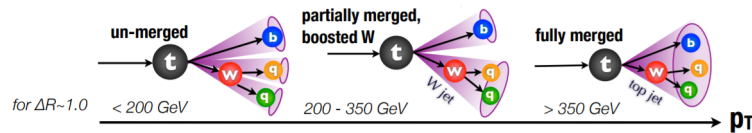


Figure 4.4: Boosted Hadronic Top

²Hadronization(Αδρονιοποίηση): Η διαδικασία σχηματισμού αδρονίων από κουάρκ και γκλουόνια

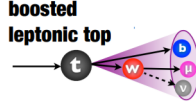


Figure 4.5: Boosted Hadronic Top

Όταν συσσωρεύονται σωματίδια μέσα σε jets, οι αλγόριθμοι δεν διαφέρουν αν τα σωματίδια προέρχονται από hard process ή από pile up. Η τεχνικές υποδομής χρησιμοποιούνται για την αφαίρεση του soft radiation (grooming) έτσι ώστε να παίρνουμε ένα jet το οποίο στην καλύτερη περίπτωση περιέχει μόνο ενέργεια από σωματίδια από την διαδικασία hard scattering. Την μεγαλύτερη επίδραση στις μετρήσεις μάζας των jet έχει η ακτινοβολία ευρείας γωνίας (wide angle radiation) η οποία ανακατασκευάζεται μέσα στο ίδιο jet. Αν ένα ελαφρύ κουαρκ ή γλουόνιο συσσωρεύεται μέσα σε ένα jet που προέρχεται από διάσπαση ενός top κουαρκ, προστίθεται περισσότερη μάζα όσο μεγαλώνει η απόσταση του top κουαρκ με το αλαφρύ παρτόνιο. Μία από τις τεχνικές υποδομής που μετριάζει το φαινόμενο αυτό είναι ο αλγόριθμος **Soft Drop** [8]. Η μέθοδος αυτή χωρίζει το τελικό jet σε δύο συστατικά. Μετά από κάθε βήμα ανακατασκευής, το κριτήριο της soft drop διαδικασίας ελέγχεται:

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

Αν η προϋπόθεση αυτή δεν ικανοποιείται, το κομμάτι με την μικρότερη εγκάρσια ορμή αφαιρείται από το τελικό boosted jet. Εδώ οι όροι p_{T1}, p_{T2} περιγράφουν την εγκάρσια ορμή των δύο συστατικών που βρίσκονται και ελέγχονται κάθε φορά στο Jet, και ΔR_{12} η μεταξύ τους απόσταση. Η παράμετρος R_0 είναι η ακτίνα του τελικού jet. Η παράμετρος z_{cut} είναι μία παράμετρος υποβάθρου. Η παράμετροι z_{cut} μαζί με την παράμετρο β καθορίζουν το πόσο ισχυρή είναι η διαδικασία grooming. Στο CMS χρησιμοποιούνται οι τιμές $z_{cut} = 0.1$ και $\beta = 0$.

4.5 Ταυτοποίηση B-Jets

Σε αναλύσεις που ασχολούνται με την μελέτη των top quarks, είναι χρήσιμο να γίνεται η ταυτοποίηση των bottom quarks αφού βρίσκονται σχεδόν πάντοτε σαν προϊόν διάσπασης ενός top quark. Οι ιδιότητες των b κουαρκ καθιστούν δυνατή την εύρεση και ταυτοποίηση jet που προέρχεται από ένα bottom quark. Αφού τα b αδρόνια έχουν αρκετά υψηλό χρόνο ζωής (1.5ps) σε συνδυασμό με την ταχύτητά τους, ταξιδεύουν μερικά mm μέσα στον ανιχνευτή μέχρι να διασπαστούν. Αυτό οδηγεί σε δευτερεύουσες κορυφές (secondary vertices) σε σημεία που διασπάται το b quark. Το CMS κατέχει ένα σύστημα ανίχνευσης με πολύ υψηλή ακρίβεια που

κρίνει δυνατή την ανίχνευση τόσο κύριων όσο και δευτερευόντων κορυφών. Τα χαρακτηριστικά αυτά προέρχονται από τον αλγόριθμο CSV (Combined Secondary Vertex). Επίσης το κλάσμα της ενέργειας από φορτισμένα αδρόνια σε ένα jet και η αναλοίωτη μάζα όλων των σωματιδίων που έχουν ανατεθεί σε δευτερεύουσα κορυφή λαμβάνονται υπ'όψην έτσι ώστε να γίνεται η διάκριση τους από non b tagged jets, δηλαδή jet που δεν περιέχουν προϊόντα από διάσπαση b quark. Ο CSV αλγόριθμος δίνει μία συνεχή έξοδο, μεταξύ 1 και 0 υποδεικνύοντας αν ένα jet είναι b-tagged ή όχι αντίστοιχα [9].

Κεφάλαιο 5

Analysis

5.1 Trigger Efficiency

5.1.1 Trigger

As stated before, the CMS uses a trigger system in order to limit the amount of the data received only to the data that is useful to the further analysis. During the LHC operation, the detectors record data with 10^9 Hz frequency. As a consequence, every detector produces about 100TeraByte of data every second. This amount of data is impossible to be stored and for that reason the trigger is used. A trigger is a system that decides which collision is interesting and whether it should be stored or not for further processing and physic analysis. The trigger system used in CMS is called **Trigger and Data Acquisition System (TRIDAS)**. Its process is separated into two stages. The so called stages are the Level 1 Trigger and the High Level Trigger (HLT).

The Level 1 Trigger is the 1st level of the trigger system. It is a hardware system that reduces the data rate to 50kHz. The subdetectors provide coarse information which is processed by fast electronic systems that decide whether the information is worth keeping or not. Very basic algorithms are used to reconstruct physical objects. The next stage of the trigger System is the High Level Trigger. It is a software system that reduces the data rate to 100Hz. It uses the information coming from all of the detectors. It is a system of combined simplified algorithms that run in a dedicated computer farm in order to reconstruct physical objects. The combination of a Level 1 trigger seed and a set of High Level Trigger conditions form a trigger path.

There are three types of trigger paths. The first type of trigger path is the Signal Trigger. The signal trigger is a path that aims to collect interesting signal

events and is always unprescaled.¹ The signal trigger type path consumes most of the available bandwidth. The second type of trigger path is the Control Triggers. These are paths that aim to collect events for studying the backgrounds. Lastly there are the Backup Trigger paths or the Backup Triggers. These are paths that come to life in extreme conditions in order to avoid prescaling. Their aim is to collect signal events with lower rate (stricter conditions applied than in normal paths).

5.1.2 Trigger Efficiency

Trigger efficiency is the fraction of events satisfying the offline selection criteria and are also selected by a specific trigger path, over the events that are only selected by a specific trigger path.

$$\text{efficiency} = \frac{\text{events passing the trigger \& offline criteria}}{\text{events passing the offline criteria}}$$

The ideal value for the trigger efficiency should be 100%. This value could be achieved only if the offline reconstruction is identical to the online.

What happens in practice is that data never passes all the events that would pass the offline criteria. Only events that pass some trigger are collected. This means that the definition of the trigger efficiency is only applicable in the simulation. As a consequence an unbiased control sample of events is needed. The events are collected with a trigger path (reference trigger) with looser selection conditions and if possible with orthogonal conditions. Trigger Efficiency is redefined as:

$$\text{efficiency} = \frac{\text{events passing the trigger \& offline criteria \& and the reference trigger}}{\text{events passing the offline criteria \& the reference trigger}}$$

5.1.3 Mathematics of the Efficiency

The trigger efficiency is defined as e , where

$$e = \frac{N_{pass}}{N_{tot}}$$

and N_{pass} is a subset of N_{tot} . The trigger efficiency is bounded in the range [0,1]. The proper mathematical approach is the following. There is a series of N_{tot} Bernoulli's experiments (e.g. binomial distribution) with N_{pass} successes. The confidence interval for the success probability is the efficiency, that is calculated.

¹Prescaling: running a path every N events and allows access to an otherwise inaccessible phase space at the cost of lower effective luminosity

In order to calculate the trigger efficiency, Wilson Interval is used:

$$e = \frac{1}{1 + \frac{z^2}{n}} \left[\hat{p} + \frac{z^2}{2n} \pm z \sqrt{\frac{\hat{p}(1 - \hat{p})}{n} + \frac{z^2}{4n^2}} \right]$$

where $z : 1 - \frac{a}{2}$ quantile of the standard normal distribution, a is the error of the confidence interval, $n = N_{tot}$ and $\hat{p} = \frac{N_{pass}}{N_{tot}}$. The standard error is $a = 0.3173$ with $z = 1$. In extreme examples;

- $N_{pass} = 0$, $e = [0, \frac{1}{n+1}]$
- $N_{pass} = N_{tot}$, $e = [\frac{n}{n+1}, 1]$
- for $n \rightarrow \infty$ (normal approximation): $e \approx \hat{p} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

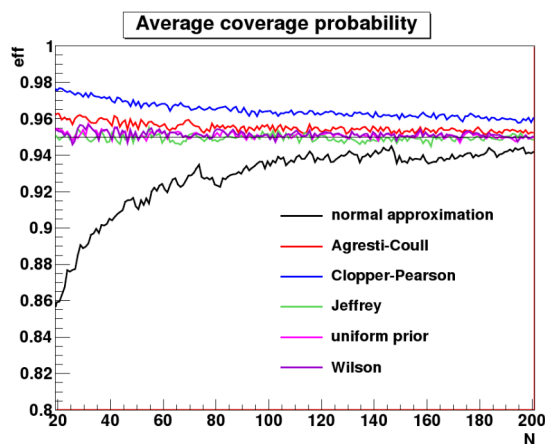


Figure 5.1: The average (over all possible true efficiencies) coverage probability for different number of total events is shown in the above picture.

5.1.4 Trigger Efficiency Analysis

The first step of the analysis was to calculate the trigger efficiency of a dataset different than the one that is going to be studied. The dataset is flatTree_SingleMuon.root. The goal was to calculate the trigger efficiency using specific trigger paths that were given in the dataset. As stated before the trigger efficiency is defined as:

$$\text{efficiency} = \frac{\text{events passing the trigger \& offline criteria \& and the reference trigger}}{\text{events passing the offline criteria \& the reference trigger}}$$

. The dataset consisted of not only the data but also a TBranch named TriggerBit, which is an array that consists of trigger paths. Every element of the array is a trigger path. Specifically, the element TriggerBit[0] was used as the signal trigger of the dataset. The reference trigger was the element TriggerBit[8] and the offline cuts that were imposed were:

- nJets ≥ 1 , number of jets more than one
- jetNBSub[0] ≥ 0 or jetNBSub[1] ≥ 0 , one jet is at least b-tagged
- mva ≥ 0.3 , Fisher Discriminant
- 140 GeV \leq jetMassSoftDrop[0],[1] \leq 200 GeV

The Trigger Efficiency is drawn with respect to the reconstructed level topPt quantity. Two reference triggers were studied. The first is the

HLT_SingleMuon27, which is the High Level Trigger that asks for an isolated muon with transverse momentum $p_T \geq 27$ GeV (380×10^3 events).

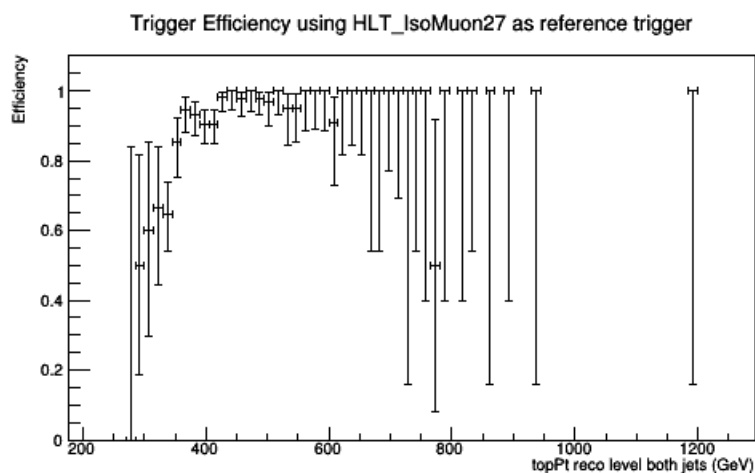


Figure 5.2: Trigger Efficiency as a function of jetPt(Transverse momentum for the topPt quantity on the reconstructed level) using HLT_SingleMuon27 as a reference trigger

The next histogram shows the Trigger Efficiency using a different reference Trigger, HLT_Mu50, which is a High Level Trigger that asks for a muon with transverse momentum $p_T \geq 50$ GeV (100×10^3 events).

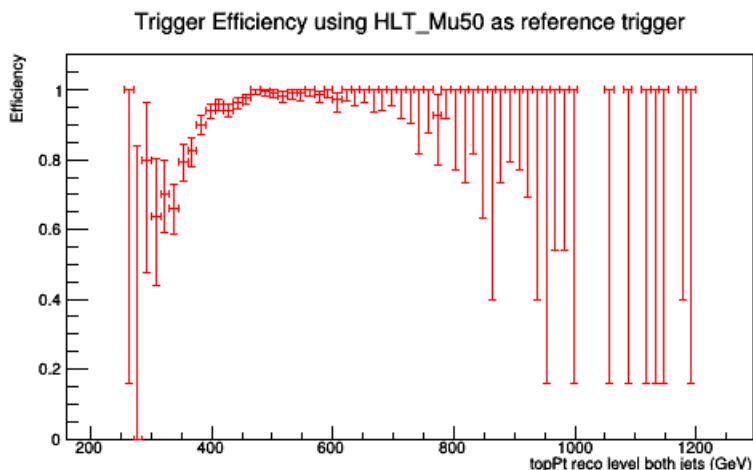


Figure 5.3: Trigger Efficiency as a function of jetPt(Transverse momentum for the topPt quantity on the reconstructed level) using HLT_Mu50 as a reference trigger

5.2 Unfolding

Unfolding is a technique used to decompose a measurement y into several sources x given the measurement uncertainties and a matrix of migrations A [10]. A short summary of the algorithm used, is given in the following lines. The best x matching the measurement y within error determined by minimizing the function $L_1 + L_2 + L_3$, where $L_1 = (y - Ax)^T V_y y^{(y-Ax)}$, $L_2 = \tau^2 (L(x - x_0))^T L(x - x_0)$, $L_3 = \lambda \sum_i (y_i - (Ax)_i)$. The term L_1 is connected with the least square minimisation. The term L_2 defines the regularisation (the smoothness condition of x) where the **parameter** τ^2 gives the strength of ten regularisation. The term L_3 is an optional area constraint with Lagrangian parameter lamda (λ), ensuring that the normalisation of x is consistent with the normalisation of y .

The input for the measurement is y , which is the vector of measured quantities with n_y dimensions. The covariance matrix of y is V_{yy} with $n_y \times n_y$ dimensions. In many cases V is diagonal and is calculated from the errors of the input parameter y . Matrix A is called the migration matrix and is the response matrix made using data x and y . It has dimensions $n_y \times n_x$ and it demonstrates the overlapping of the parameters x and y for each bin of their histograms. The result emerging from unfolding is vector x , which is the unknown underlying distribution with n_x dimensions. Also the Unfolding method calculates the error matrix of x , V_{xx} .

The regularisation of the method is strongly connected to the parameter **tau** τ . Tau is the parameter that defines the regularisation strength of the method. Another regularisation parameter is the matrix L . L is the matrix of the regularisation conditions with dimensions $n_l \times n_x$ depends on the structure of the input

data y . Last, the parameter x_0 is calculated from simulation.

The preservation of the area is succeeded using the following parameters. Lamda is the Lagrangian multiplier, y_i is one component of the vector y and $(Ax)_i$ is one component of the vector Ax . The determination of the unfolding result x :

- not constrained: the minimisation is performed as a function of x for fixed lamda = 0
- constrained: stationary point is found as a function of x and lamda

The constraint can be useful to reduce biases on the result x in cases where the vector y follows non-Gaussian probability densities.

5.2.1 Inverse Problems

The aim of experiments in particle physics is the estimation of probability distributions of variables like energy, particle mass or decay length from a data sample recorded from a detector and reconstructed from raw data by involved algorithms. The data is collected in the form of histograms representing an estimate of the distribution. The distribution $g(s)$ of the measured values s is related to the distribution $f(t)$ of the true variable t by migration, distortions and transformations. Using Monte Carlo (MC) methods the direct process from an assumption $f^{theory}(t)$ on the true distribution $f(t)$ is difficult and known to be ill posed, meaning that small changes in the measured distribution can cause large changes in the reconstructed true distribution.

In particle high energy physics, the inverse process is usually called unfolding.

Inverse Process (unfolding) \rightarrow measured distribution $g(s) \Rightarrow f(t)$ true distribution

There is also the direct process called folding which is described respectively,

Direct Process (MC) \rightarrow true/MC distribution $f(t) \Rightarrow g(s)$ measured Distribution

Both processes are described by the Fredholm integral equation of the first kind

$$\int_{\omega} K(s, t) f(t) dt + b(s) = g(s)$$

, with a Kernel function $K(s,t)$ describing the physical measurement process. The distribution $b(s)$ represents a potential contribution to the measured distribution $g(s)$ from background sources. The Unfolding process requires a determination

of the distribution $f(t)$ from the measured distribution, where the Kernel function is describing the detector response and is also called the response function and is usually known from a Monte Carlo simulation based on a f^{theory} .

The typical statistical and systematic effects, in a simulated measurement which are typical for particle physics experiments are

- In all the bins there are statistical fluctuations from the Poisson statistics
- There is a migration between the bins due to finite resolution (smearing)
- Limited acceptance and reduced efficiency means that potential entries in a bin are missing
- Due to a non linear detector response there is a shift in a certain direction

The effects described above, are typical for the detector response in particle physics. If the goal of the experiment is to test specific predictions of a theory, then the folding process is applied, using the predicted $f^{theory}(t)$ combined with the detector response in order to be compared with the measured distribution $g(s)$. On the other hand, the reconstruction of the true distribution $f(t)$ by unfolding allows a direct comparison with predictions of a theory. For a quantitative comparison a full covariance matrix is required for the reconstructed true distribution.

5.2.2 Discretisation and Linear Solution

The inverse problem is given by the Fredhold integral. In order for to allow a numerical solution the integral has to be discretised in the following manner:

$$y = Ax$$

where \mathbf{A} is a matrix, and \mathbf{x}, \mathbf{y} are vectors. Specifically,

- True Distribution $f(t) \rightarrow x$, n-vector of unknowns
- Measured Distribution $g(s) \rightarrow y$, m-vector of measured data
- Kernel $K(s, t) \rightarrow A$, rectangular m by n response matrix

In the following, the variables s, t and the vectors \mathbf{x}, \mathbf{y} are assumed to be one dimensional. There are several discretisation methods that can be applied. The most usual method is to represent the distribution as histograms integrating the distributions over a short interval, meaning the bin, given the grids for the true distribution and the grids for the measured distribution respectively. The bin limits are calculated as

$$y_i = \int_{s_{i-1}}^{s_i} ds g(s), i = 1, 2, ..m$$

The same method can be used for the discitisation of the Kernel $K(s,t)$ as well as the true distributiown $f(t)$, if the response is determined from a MC sample. The elemnts y_i are calculated according the product $y_i = A_i^T x$ where the vector A_i is defined as a column vector containing elements of matrix A . The elements of the response matrix are elements that are positive or zero. If the assumed theoretical distribution coming from the MC is a normalised probability density function, then each element of the response matrix can be interpreted as a conditional probability

$$A_{ij} = P(Observed\ in\ bin\ i | True\ Value\ in\ bin\ j)$$

The sum ϵ of the elements A_{ij} over all bins i of the observed distribution gives the detection efficiency of the measurement detector. This includes a geometrical acceptance and other factors. The determination and calculation of the response matrix \mathbf{A} is fundamental for the unfolding. The elements of the matrix have statistical errors, which can delimit the effective rank of the unfolding problem.

If the assumption is made that the response matrix A is accurate and the relation $Ax_{exact} = y_{exact}$ is valid, the measured distribution \mathbf{y} deviates from y_{exact} only by data errors due to statistical fluctuations. Assuming that \mathbf{e} represents the m -dimensional vector with the data errors,

$$y = y_{exact} + e = Ax_{exact} + e$$

In particle physics the statistical properties of the measurement are usually well known. The elements of the vector \mathbf{y} are often counts and thus follow the Poisson distribution, in which case the maximum likelihood solution is adequate.

The number n of parameters (for example the number of bins of the reconstructed true distribution) is usually small and it is essential to calculate the full covariance matrix V of the result, which might be used for further statistical analysis. If the Fredholm equation is solved for the estimate \hat{x} by a linear transformation of the data vector \mathbf{y} according to $\hat{x} = (A^\dagger y)$, the propagation of the data uncertainties due to the unfolding is straightforward:

$$V_x = A^\dagger V_y A^\dagger$$

In the case where $m=n$, it can be solved using the inverse matrix A^{-1} but there are often problems because \mathbf{A} matrix has a bad condition or is even singular. So the situation where $m=n$ should be avoided. sda

In the recommended case where $m \ll n$, the n -by- m matrix A^\dagger can be constructed and used to determine the estimate \hat{x} :

$$\hat{x} = A^\dagger y = A^\dagger y_{exact} + A^\dagger e = A^\dagger A x_{exact} + A^\dagger e$$

The pseudo-inverse satisfies the relation $A^\dagger A = I$ and is a generalisation of the inverse matrix. It allows the naive solution in the least-squares sense, derived from the requirement:

$$\min F(x) \text{ with } F(x) = (Ax - y)^T V_y^{-1} (Ax - y)$$

where the inverse of the data covariance matrix V_y is included to take into account the accuracies which differ between the elements of the data vector.

5.2.3 Unfolding Using the Least Squares Method

The least squares solution is expressed by using the pseudo-inverse A^\dagger . We will write this pseudo inverse matrix in terms of SVD matrices where:

$$A^\dagger = V \Sigma^{-1} U^T$$

The least squares estimate \hat{x} is given by:

$$\hat{x} = A^\dagger y = V \Sigma^{-1} (U^T y) = \sum_{j=1}^n \frac{1}{\sigma_j} (u_j^T y) v_j = \sum_{j=1}^n \frac{c_j}{\sigma_j} v_j$$

The data y are transformed by U^T to a n -vector $c = U^T y$, representing the transformed measurement. The elements $c_j = u_j^T y$ are called Fourier coefficients and are elements of the measurement. The coefficients tend to decrease fast for large values of j when the distribution of y is smooth. They are also statistically independent. From the above expression, it is understood that insignificant Fourier coefficients with small singular values σ_j will result in large fluctuations in the unfolding result \hat{x} and can make the result unacceptable.

The calculation of the uncertainty of the estimate \hat{x} is straightforward because of the linear transformation of the data y in the solution shown above.

5.2.4 Norm Regularisation Method

The singular values σ_j of the response matrix A decrease without a clear gap between large and small singular values. Some oscillations are generated if sharp cut-off between singular values of the same magnitude is used. These oscillations can be reduced by smooth cut-offs which are introduced with regularisation methods [11].

The Standard method for the solution of ill-posed problems is the regularisation method. The expression to be minimised with respect to the unfolding result included the least squares and the log-likelihood expressions. A second term of the type $\Omega(x) = \|Lx\|^2$ with a certain matrix L , ensures smoothness of the unfolding result and contributes a weight, given by the regularisation parameter τ :

$$\min(F(x) + \tau\|Lx\|^2).$$

In the least squares solution the matrix A^\dagger is replaced by the regularised matrix $A^\#$:

$$\hat{x} = A^\#y = [(A^T A + \tau L^T L)^{-1}]y$$

The regularisation term $\tau L^T L$ is added to the matrix $C = A^T A$ of the normal equations and as $y = Ax_{exact} + e$ one obtains:

$$\hat{x} = A^\# Ax_{exact} + A^\# e = x_{exact} + (A^\# A - I)x_{exact} + A^\# e$$

where the second term is the systematic error and the third term is the statistical error. The product $\Xi = A^\# A$ is called the resolution matrix. The resolution matrix is not equal to the unit matrix and thus the method has a systematic bias (second term). The unmeasurable and unnatural oscillations of the solution lead to this potential bias of the \hat{x} . The smoothing effect of the resolution matrix gives no or small systematic errors for smooth exact distributions and large systematic deviations for oscillating distributions. The measured distribution y has to be compared with the distribution \hat{y} which corresponds to the estimated reconstructed distribution \hat{x} :

$$\hat{y} = A\hat{x} = (AA^\#)y$$

Norm Regularisation

The simplest regularisation method is the norm regularisation, where $L=I$. For a given regularisation parameter τ , the estimate \hat{x} can be determined by standard methods of linear algebra (inverse matrix). However, the SVD solution is simpler because the regularisation matrix is diagonal and allows a clear understanding of the effects of the regularisation to be obtained. Using the SVD method, the solution of the estimate \hat{x} is written in the form:

$$\hat{x} = V[(\Sigma^2 + \tau I)^{-1}\Sigma^2]\Sigma^{-1}(U^T y) = (VFS^{-1}U^T)y$$

where the matrix $F = (\Sigma^2 + \tau I)^{-1} \Sigma^2$ is a diagonal matrix with elements equal to filter factors φ_j , and the term $U^T y$ is the coefficients c_j . The estimate \hat{x} is expressed by the following sum:

$$\hat{x} = \sum_{j=1}^n \frac{c_j}{\sigma_j} \phi_j \nu_j \text{ with } \phi_j = \frac{\sigma_j^2}{\sigma_j^2 + \tau}$$

In case of diagonalisation the values σ_j^2 are replaced with the eigenvalues λ_j . The effect of the regularisation is thus the introduction of the filter factor ϕ_j for each term with a strength which depends on the regularisation parameter τ .

The filter factor ϕ_j will appear also in the expression of the covariance matrix:

$$V_x = \sum_{j=1}^n \frac{1}{\sigma_j} \phi_j^2 \nu_j \nu_j^T$$

If the selected regularisation parameter τ is small enough to reduce the only the insignificant Fourier coefficients, there will be no bias introduced.

The norm regularisation corresponds to the original regularisation proposal by Tikhonov and Phillips. The parameter τ can be interpreted as the a priori measurement error $s_{reg} = \frac{1}{\sqrt{\tau}}$ for each component of the vector x . Individual values of s_{reg} can be introduced, corresponding to a regularisation term $\Omega(x) = \sum_{j=1}^n \frac{x_j^2}{s_j^2}$.

When the solutions are not smooth, and the data has a high relative precision there are two possibilities in order to change the regularisation term $\Omega(x) = \|Lx\|^2$. The first is to make the regularisation term into

$$\Omega(x) = \|L(x - x_0)\|^2$$

with the a priori assumption x_0 on the resulting vector x .

The second one is to amend the Monte Carlo simulation to include a realistic a priori assumption $f(t)^{model}$ about the function $f(t)$ and to already include the model function in the definition of the response matrix

$$\int_{\Omega} [K(s, t) f(t)^{model}] f(t)^{cor} dt = g(s).$$

In this way, only an almost constant correction function $f(t)^{cor}$ has to be determined with $f(t) = f(t)^{model} f(t)^{cor}$. This option is available in unfolding methods of particle physics. The elements of this redefined matrix \mathbf{A} are now integers, meaning the number of Monte Carlo events from bin j of \mathbf{x} measured in bin i of \mathbf{y} .

5.2.5 Implementation of the method

The result of the unfolding is calculated as stated in the following lines.

$$L^2 = L :: L \text{regularisation conditions squared}$$

$$\epsilon_j = \sum_i A_{ij} \text{ vector of efficiencies}$$

$$E^{-1} = ((A^T(V_{yy})^{-1}A) + \tau^2 L^2)$$

$$x = E(A^T(V_{yy})^{-1}y + \tau^2 L^2 x_0 + \frac{\lambda}{2} * \epsilon) \text{ is the result}$$

The derivatives $\frac{dx_k}{dA_{ij}}, \frac{dx_k}{dy_i}, \frac{dx_k}{d(\tau^2)}$ are also calculated for further usage. The covariance matrix V_{xx} is calculated as: $V_{xx} = \sum_{kl} \frac{dx_i}{dy_k} V_{yy} \frac{dx_j}{dy_l}$

The algorithm is based on "Standard" matrix inversion. This means that this method is limited due to numerical accuracy and computing cost problems when it comes to matrices with large dimensions. Thus the algorithm should not be used for large dimensions of x and y (n_x not larger than 200 and n_y not larger than 1000).

A very important factor for this algorithm to be successful is the choice of tau. The method implemented in the Library used (TUnfold) is the L-curve method. The L-Curve is a two dimensional curve where on the x-axis lies $\log(x^2)$ and on the y axis lies $\log(\text{regularisation condition})$. The so called L-curve has its name because in many cases the curve looks like an L. The best choice of tau is in the kink of the L curve. The L-curve method tests a given number of points in a predefined tau-range and searches for the maximum of the curvature in the L-curve (kink position). If no tau range is given, the range of the scan is determined automatically.

5.2.6 Regularisation Conditions

In order to avoid large oscillations and large correlations on the output bins, regularisation is needed for most unfolding problems. This means that some extra conditions are applied on the output bins.

If the regularisation of the unfolding is strong, i.e. large parameter of tau, then the distribution of the output x or its derivatives will look like the bias distribution. If the parameter tau is small, the distribution x is independent of the bias.

There are three basic types of regularisation in the TUnfold library:

Condition	Regularisation
kRegModeNone	none
kRegModeSize	minimize the size of (x-x0)
kRegModeDerivative	minimize the 1st derivative of (x-x0)
kRegModeCurvature	minimize the 2nd derivative of (x-x0)

The regularisation schemes kRegModeDerivative and kRegModeCurvature have the nice feature that they create correlations between x-bins, whereas the non-regularized unfolding tends to create negative correlations between bins. For these regularisation schemes the parameter tau could be tuned so that the correlations are smallest, as an alternative for the L-curve method.

5.2.7 TUnfold Class

The goal of this Master Thesis was to unfold Monte Carlo true data and compare them with reconstructed Monte Carlo Data. It refers to data before and after the detection. The equation that has to be solved is the following.

$$x_{true} = R(x_{true}, x_{reco}) * x_{reco}$$

where x_{true} is the events that happen at the moment of collision, x_{reco} is the vector containing the events that are received from the detectors and R is the response matrix that indicates the probability for the x_{true} to be measured as x_{reco} . The process that is followed is the process that was analysed before using x_{reco} as the know vector y and x_{true} as the unknown vector x.

—————Tichonov regularisation and stat error bars

The unfolding process is the only process that can be used so that the two measurements (reconstructed and true events) are compared to each other. Every detector has a response matrix, that differs from detector to detector. This process can be applied only on simulation level because only then is the x_{true} vector known (Monte Carlo).

5.3 Monte Carlo Datasets

This section aims to inform the reader, the process that was followed during the unfolding analysis. But firstly, it would be necessary that a brief introduction of the variables used in the analysis is made.

5.3.1 General

The process of unfolding, is going to be applied on several Monte Carlo Data root files acquired from CMS experiment. Each dataset, is produced using different parameters, so that the process of $t\bar{t}$ decay is simulated. In this analysis, the main purpose is the extrapolation of the data from the reconstructed level to the full particle phase space, which is the parton level space. The reader could imagine the three phase spaces as three sets. The reconstructed and the particle level set are subsets of the full partonic phase space. Also, the two sets are intersected. The intersection of the reconstructed and particle sets, will lead to the response matrix of the reconstructed and particle data.

The Monte Carlo Datasets used in this analysis are:

- TuneCUETP8M2T4_13TeV-powheg-pythia8 (nominal)
- TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8
- TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8
- TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8
- TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8
- TuneCUETP8M2T4down_13TeV-powheg-pythia8
- TuneCUETP8M2T4up_13TeV-powheg-pythia8

The process is described as followed: From the overlap of the reconstructed and the particle data the response matrix is acquired. Using the response matrix as the migration matrix, the Unfolding process is made, so that from the reconstructed data the particle level data is acquired. After the Unfolding the particle level data is extrapolated to the full particle phase space. A function that could describe the above process is the following:

$$f(x^{particle}) = a(x^{particle})R^{-1}(x^{reco}, x^{particle})[A(x^{reco})D(x^{reco})]$$

where $a(x^{particle})$ is a factor of extrapolation and $A(x^{reco})$ is a factor of the overlapping phase spaces.

5.3.2 Variables

The variables used in the analysis are divided in three categories. As shown in the picture below, the section where the collision happens, is called the "**Parton Level**". The section where the hadronization begins, is called the "**Particle**".

Level". The final section which is the detectors, is called the "**Detector Level**". In the following table, the physical quantity that is going to be analysed is connected with the levels mentioned above. Also the name of each variable in each level is given.

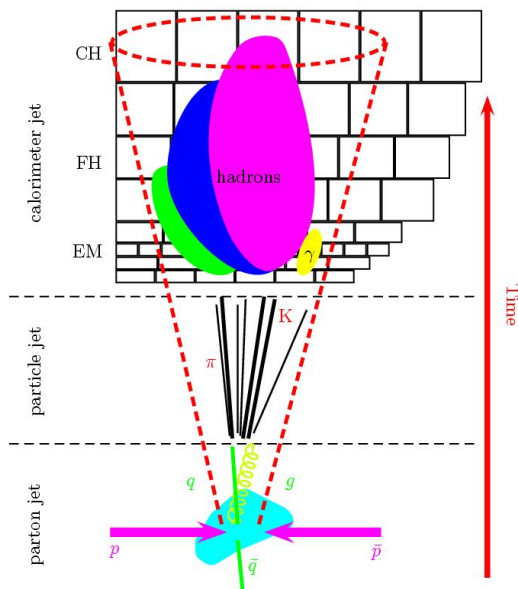


Figure 5.4: Sketch of the evolution from the hard-scatter parton to a jet in the calorimeter

Physical Quantity	Detector Level	Particle Level	Parton Level
Top Jet Transverse Momentum	jetPt	genjetPt	ptTopParton
Transverse Momentum of $t\bar{t}$	ptJJ	ptJJGen	ptTTbarParton
Rapidity y of $t\bar{t}$	yJJ	yJJGen	yTTbarParton
Mass of $t\bar{t}$ jets	mJJ	mJJGen	mTTbarParton

Table 5.1: Variables used for Unfolding

5.3.3 Response Matrices-Selection

A response matrix is a two dimensional histogram that shows the correlation between two variables. If the events of the central diagonal of the two dimensional histogram coincide, this means that the two variables are strongly correlated. In the following analysis, the response matrices are filled with events of the same

parameter in different "Levels". In particular for each variable mentioned in the above section, two response matrices were constructed. The x axis of both of the response matrices was filled with events from the Reconstructed-Detector Level. The y axis, was filled with events coming from the Particle Level or the Parton Level.

Before the events filled the two response matrices, specific cuts were applied for each axis. For the better understanding of the cuts, it is necessary that the variables used in the cuts are explained. The cuts applied on the events on the reconstructed level were the following:

- $n\text{Jets} > 1$
- $\text{jetPt}[1] > 400$
- $\text{jetNBSub}[0] > 0$ and $\text{jetNBSub}[1] > 0$
- $\text{triggerBit}[2] == \text{true}$, this is Signal Trigger
- $\text{mva} > 0.9$
- $\text{jetMassSoftDrop}[i] > 140$ and $\text{jetMassSoftDrop}[i] < 200$ for $i=0,1$

The cuts applied on the events of the particle level were the following:

- $n\text{JetsGen} > 1$
- $\text{genJetPt}[1] > 400$
- $\text{mvaGen} > 0.9$
- $\text{genJetMass}[i] > 140$ and $\text{genJetMass}[i] < 200$ for $i = 0,1$

The cuts applied on the events of the parton level were the following:

- Leading Jet $p_T \gtrsim 400\text{GeV}$

By demanding $n\text{Jets}$ (particle and reconstructed level) to be greater than 1 means that 2 jets are needed in order the top quarks decay. Also, the cut $\text{jetPt}[1]$ be greater than 400 GeV means that a big value of transverse momentum is needed so that the top decay products are collimated. The cuts $\text{jetNBSub}[0]$, $\text{jetNBSub}[1] > 0$ refer to the need that a btagged subject is inside the big jet as the top decays into a W boson and a b quark. The mva variable is a fisher discriminant and demands from the jets to have substructure given the fact that the top quarks decay into W bosons and b quarks. The mass cuts (genJetMass and jetMassSoftDrop) have to be compatible with that of the top quark mass.

Multivariate Analysis

It is a Set of statistical analysis methods that simultaneously analyze multiple measurements on the studied object. In our case it is used to separate signal from background. The variables are dependent and correlated in various ways.

We call Supervised Learning the process when the Multivariate method is trained over a sample where the result is known (for example a Monte Carlo Simulation of signal and background).

An example of MVA methods are:

- Neural Network
- Decision Tree
- Fisher Discriminant
- Rectangular Cut optimization

In this case Aim is the identification of boosted top-antitop pair and we the Neural Network is used. The cut is set to be $mva \gtrsim 0,9$ where the signal vs background is clear as shown in the following picture:

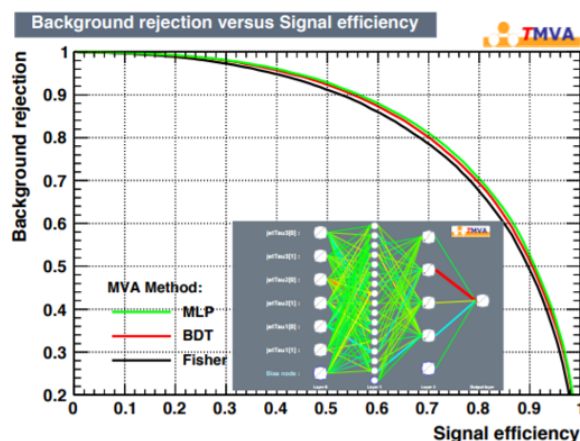


Figure 5.5: Performance of Multivariate analysis methods

The following response matrices were calculated using the nominal Monte Carlo Dataset (TuneCUETP8M2T4_13TeV-powheg-pythia8).

In the following pages, the response matrices of the variables topPt, pt ttbar, yttbar and ttbar mass will be presented respectively.

In the following response matrices, the binning on the x, y axes is variable. The variable binning is used because by raising the width of the bins, more data is found inside each bin and as a consequence the efficiency and purity grows bigger.

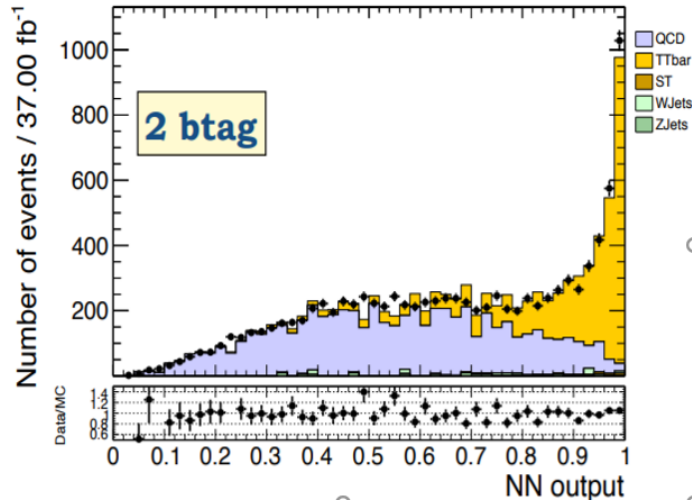


Figure 5.6: Discrimination of Signal and Background is clear over 0.9 for the Neural Network

Specifically, the bin width is chosen in such a way so that the purity² and the efficiency³ be greater than 50%.

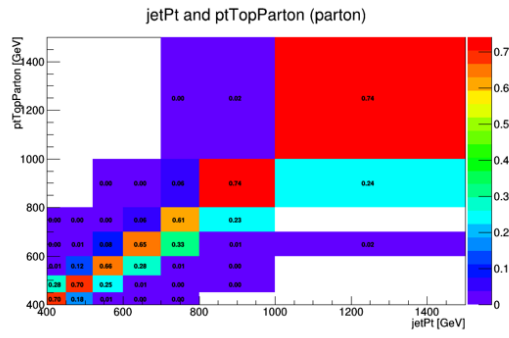
The response matrices are normalised to 1. The binning used on the following response matrices is:

- topPt variable: [400,450,520,600,700,800,1000,1500]GeV
- pt ttbar variable: [0,80,220,360,500,650,800,1000]GeV
- ttbar mass variable: [1000,1200,1400,1600,1800,2000,2400,2800, 3500]Gev
- ttbar rapidity variable: 15 bins with constant width in [-2.4,-1.6,-1.2,-0.8,-0.6,-0.4,-0.2,0,0.2,0.4,0.6,0.8,1.2,1.6,2.4]

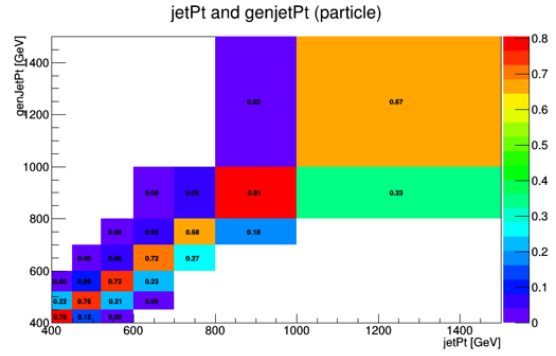
In the beginning of our analysis, the particle (gen) cuts included a cut that was related with the btagging in the particle level quantity. More specifically, the cut was applied on the events, whether the particle level is btagged or not. In the end, it seemed that the efficiency with respect to the reconstructed (reco) level quantity dropped. As a result, it was decided that the btagging for the leading and subleading jet would be ignored, in order for the efficiency to stabilize.

²Efficiency: if the x axis of the response matrix is the partonPt then the column of a normalised response matrix shows the efficiency

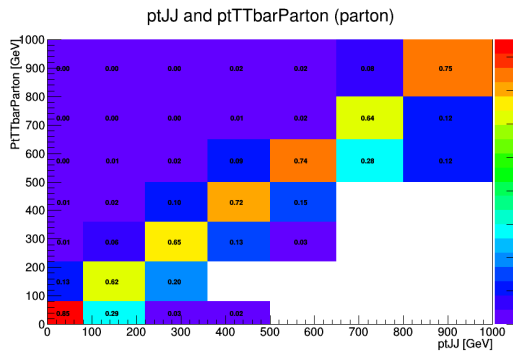
³Purity: If the y axis of the response matrix is the recoPt then the lines of a normalised response matrix corresponds to the purity



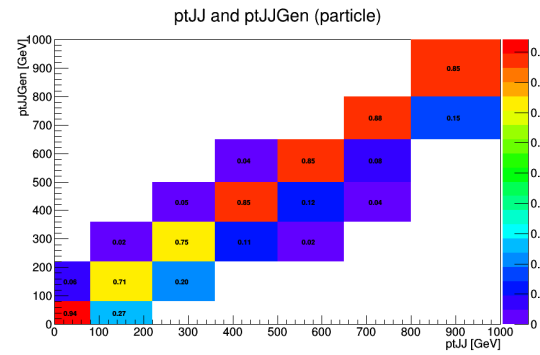
(a)



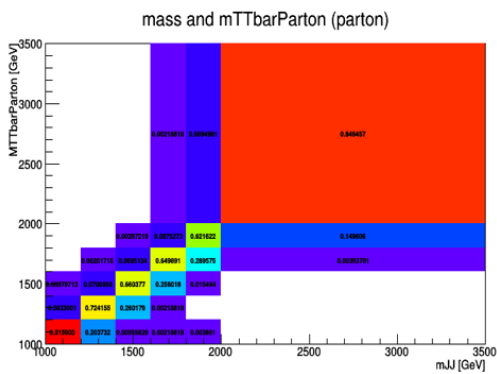
(b)



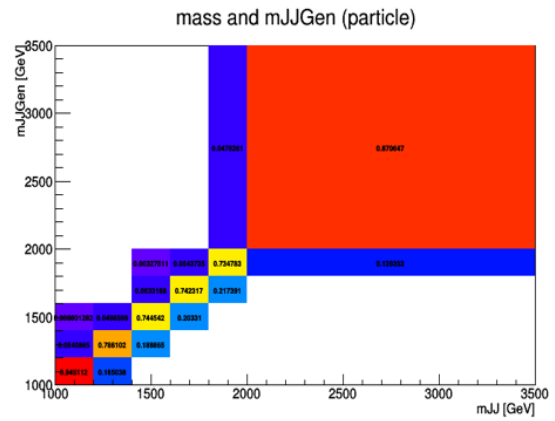
(a)



(b)

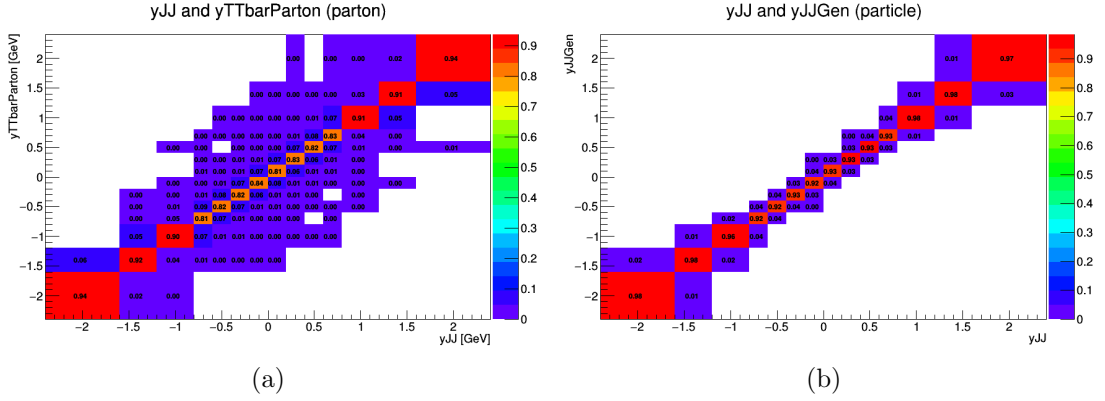


(a)

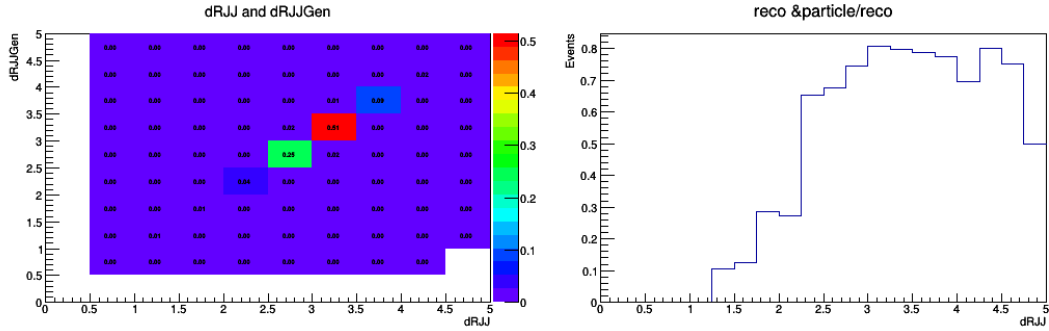


(b)

During the process it was found out that the for small δR in the particle level, the btagging of the genJetJets is really bad and "kills" many of the events needed.



That means that when for example the particle level transverse momentum (p_t is pt_{JJ} in particle level) is high the respective particle level δR (dR_{JJ}) is low. This creates a strong slope in the efficiency, meaning that the correction factor that is going to be needed for the calculation of the x_{gen} particle level quantity from the reco phase space extrapolating to the particle level phase space will be extremely high as the correction factor is defined as $\text{correction factor} = \frac{1}{\text{efficiency}}$. This will eventually result in the serious increase of the systematic uncertainty.



(a) Response Matrix for dR_{JJ} (reco dR) and dR_{JJGen} (particle) (b) Events Passing reco and particle cuts over reco cuts vs dR_{JJ}

5.3.4 Closure Tests

As it was stated in the unfolding section, unfolding is a process that solves the equation

$$y = Ax \rightarrow x = A^{-1}y$$

In this dataset the unfolding process was applied on the **Reconstructed** events which play the role of x , using as A^{-1} the response matrix for each variable and y

the **Generated** (particle level) events. Then the same process was applied where **Parton** events took the role of x as y is still the Reconstructed level quantities.

The unfolding process was applied on the four variables, topPt, pt ttbar, yttbar and ttbar mass and the process was tested for several values of tau (τ)⁴.

The values of tau that were used were the following: $\tau = 10^{-9}, 10^{-7}, 10^{-5}, 10^{-3}, 0.1$. The following histograms show the reconstructed data (hist parton) and the unfolded data compared with the particle level-generated data (his particle) for the several values of tau (hUnf1, hUnf2, hUnf3, hUnf4, hUnf5) respectively. Except from the fixed values of tau, TUnfold class has a function in which for a fixed number of scans, the best regularisation parameter tau is calculated. The respective unfolded histogram is the hUnfBest histogram.

It is clear that the ideal situation is when the unfolded histogram (hUnf) is identical with the particle or parton histogram respectively in each occasion.

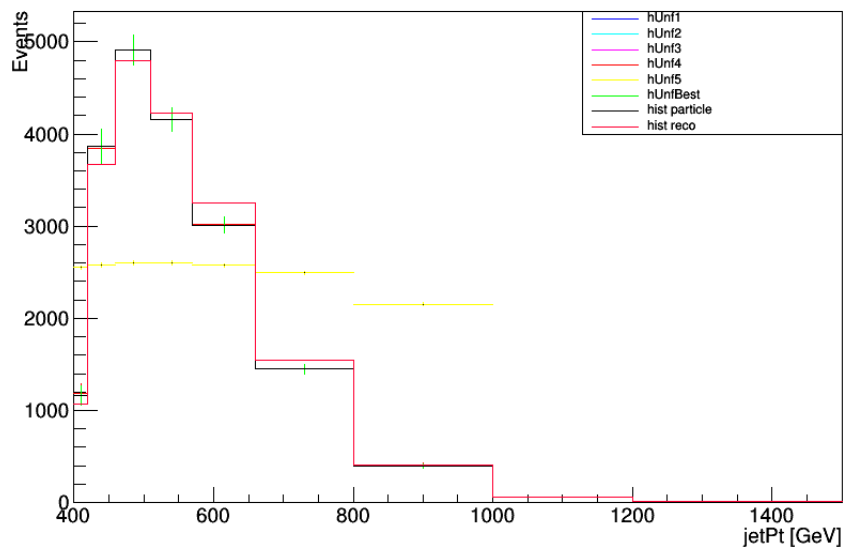


Figure 5.7: Unfolding on the variable topPt for several values of tau

⁴Tau: regularisation parameter used in unfolding

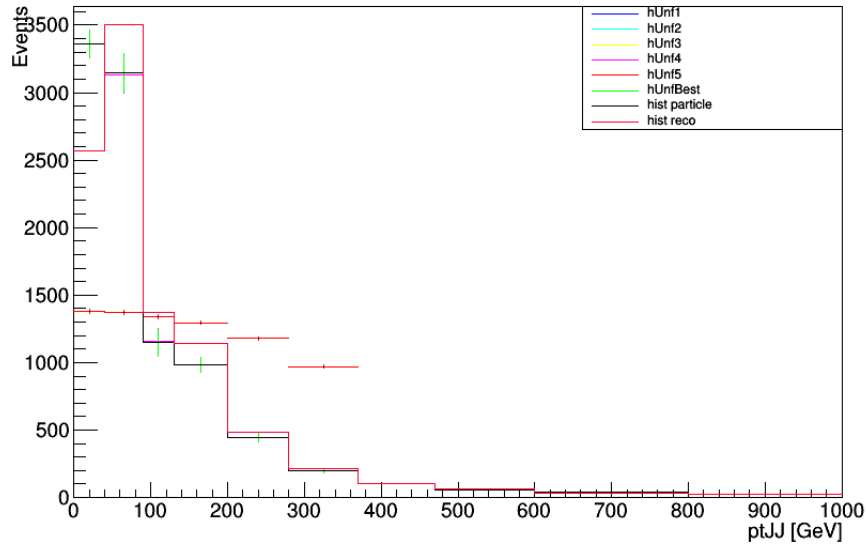


Figure 5.8: Unfolding on the variable $pt_{t\bar{t}}$ for several values of τ

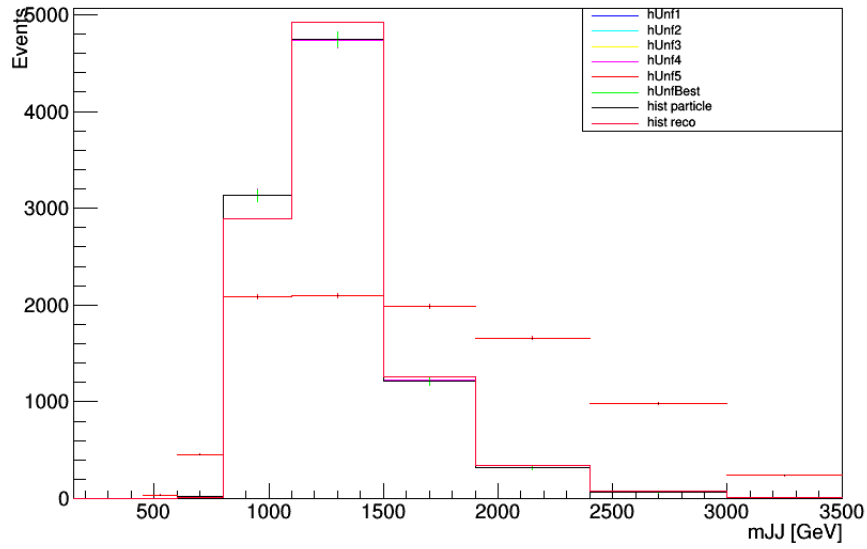


Figure 5.9: Unfolding on the variable $tt_{barmass}$ for several values of τ

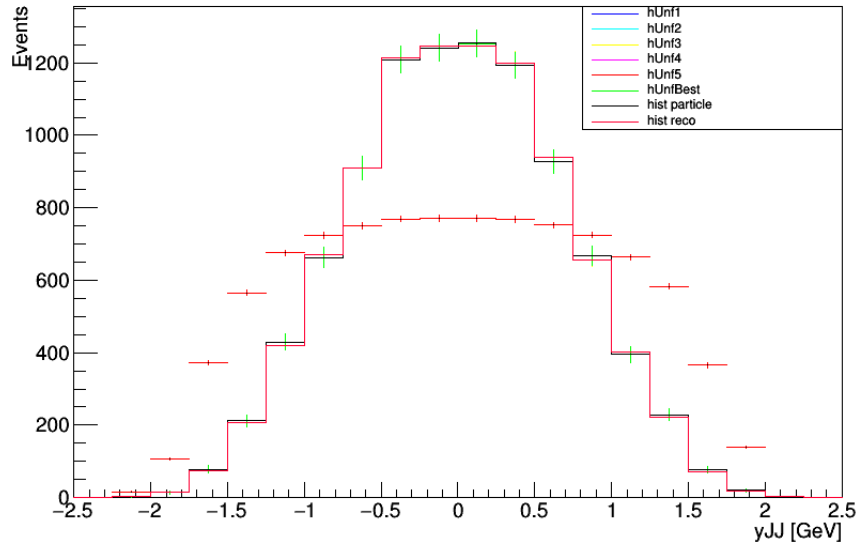


Figure 5.10: Unfolding on the variable $y_{t\bar{t}}$ (rapidity) for several values of τ

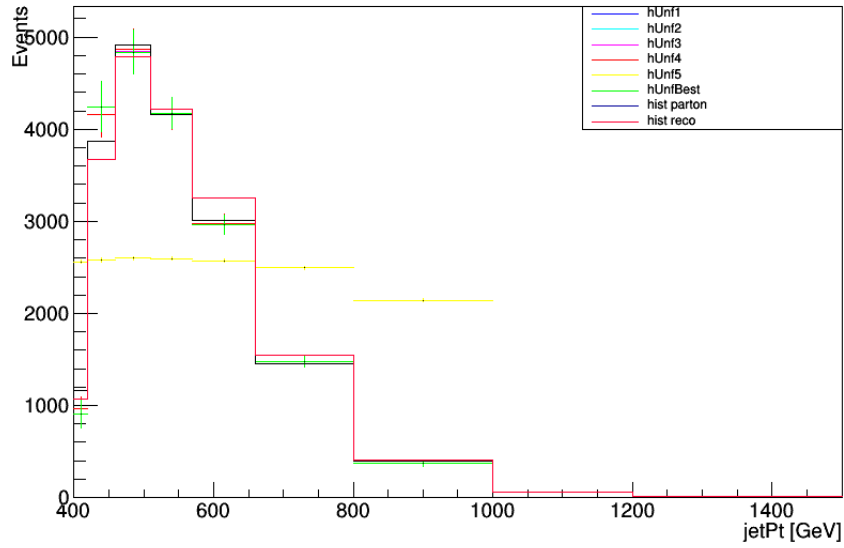


Figure 5.11: Unfolding on Parton the variable topPt (jet Transverse momentum) for several values of tau

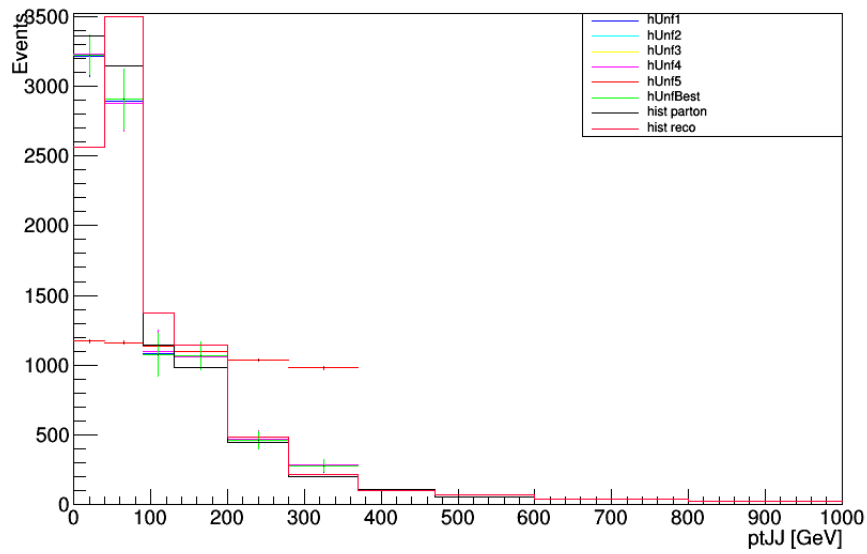


Figure 5.12: Unfolding on Parton vs the variable pt (jet Transverse momentum) for several values of tau

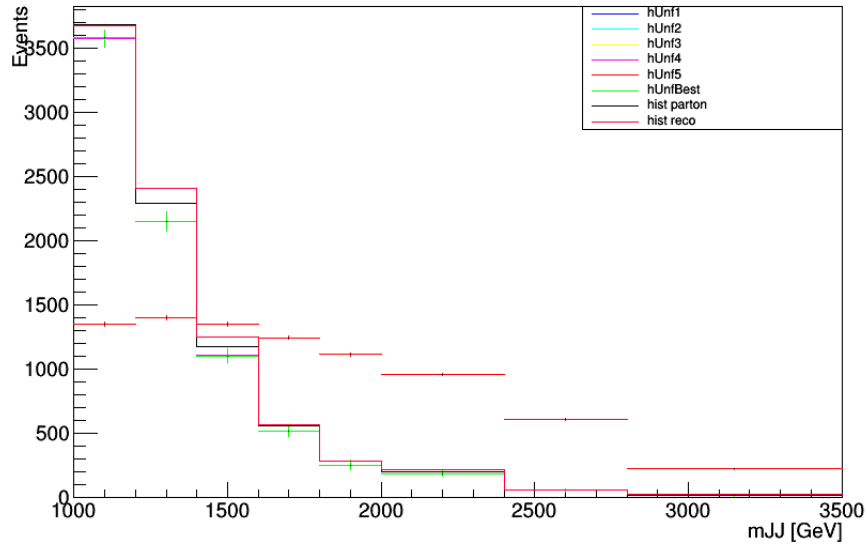


Figure 5.13: Unfolding on Parton vs the variable mass for several values of tau

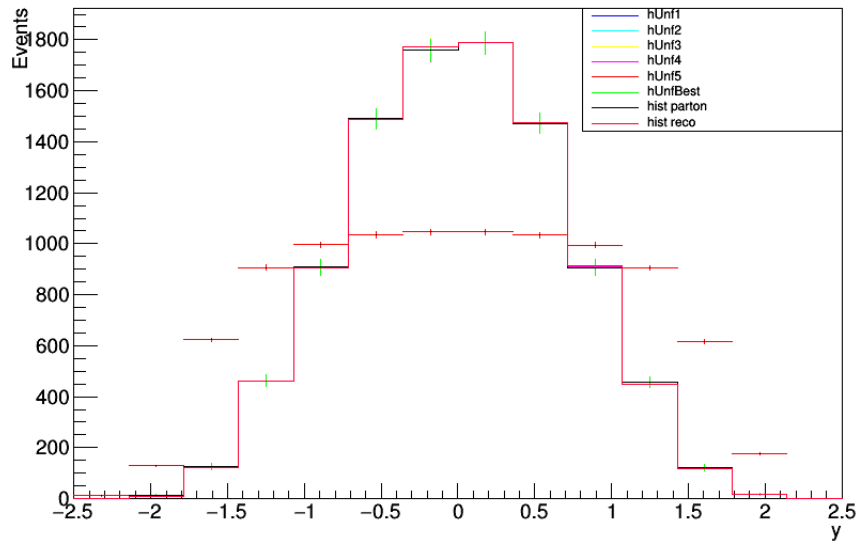


Figure 5.14: Unfolding on Parton vs the variable y (rapidity) for several values of tau

5.3.5 Correction Factors

In this section, the correction factor (or true efficiency) for the extrapolation from the parton level phase space to the reconstructed level phase space, the correction factor for the extrapolation to the particle/reconstructed overlapping phase space from the reconstructed phase space as well as the correction factor for the extrapolation from the overlapping reconstructed/particle phase space to the final particle phase space will be calculated using all the Monte Carlo Datasets.

The fraction that determines the correction factor for the extrapolation from the full parton phase space to the reconstructed phase space is defined as:

$$\text{Correction Factor (Parton Level Quantity)} = \frac{\text{events passing reco cuts and signal trigger}}{\text{eventCounter}}$$

where eventCounter is a histogram containing all of the events of the full partonic phase space. This fraction is filled and drawn with respect to the Parton Level quantity that is studied.

Respectively, the fraction that determines the correction factor for the extrapolation from the reconstructed level phase space to the common reconstructed/particle phase space is defined as:

$$\text{Correction Factor (Reco Level Quantity)} = \frac{\text{events passing reco and particle cuts}}{\text{events passing reco cuts}}$$

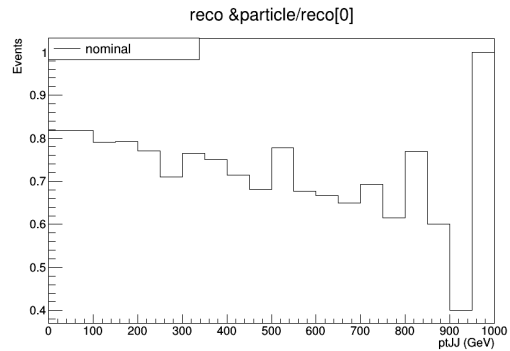
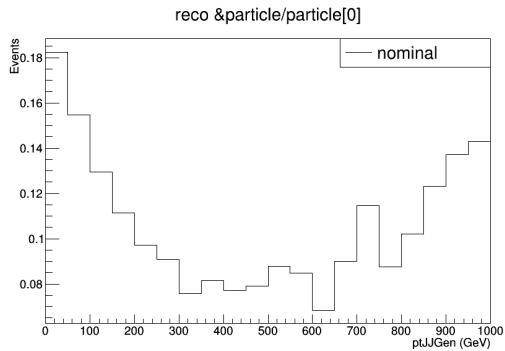
This fraction is filled and drawn with respect to the Reconstructed Level quantity that is studied.

The fraction that determines the correction factor for the extrapolation to the particle phase space from the common reconstructed/particle phase is defined as:

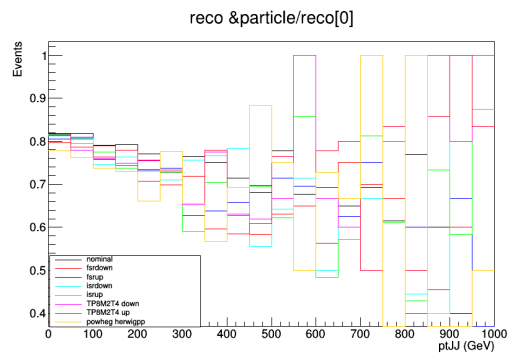
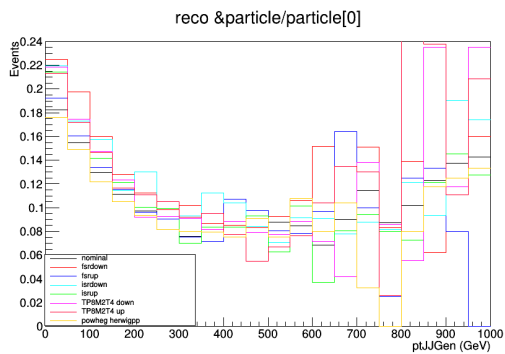
$$\text{Correction Factor (Particle Level Quantity)} = \frac{\text{events passing reco and particle cuts}}{\text{events passing particle cuts}}$$

This fraction is filled and drawn with respect to the Particle Level quantity that is studied.

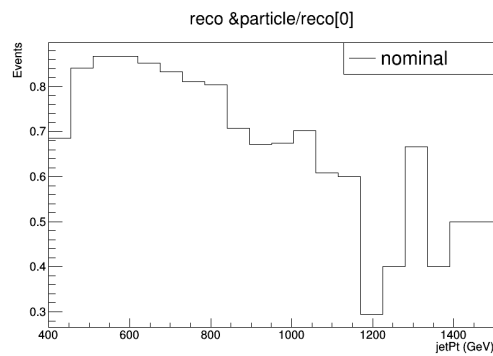
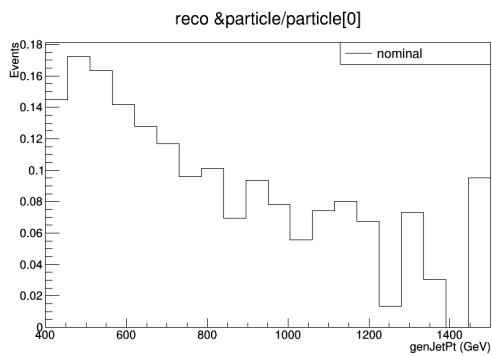
Let us begin with the efficiencies concerning the parton quantity, meaning the extrapolation from the full partonic phase space to the reconstructed phase space.



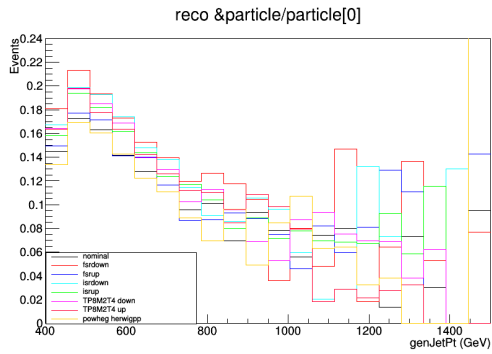
(a) Events Passing reco and particle cuts over particle cuts vs ptJJGen for nominal Monte Carlo
 (b) Events Passing reco and particle cuts over particle cuts vs ptJJ for nominal Monte Carlo



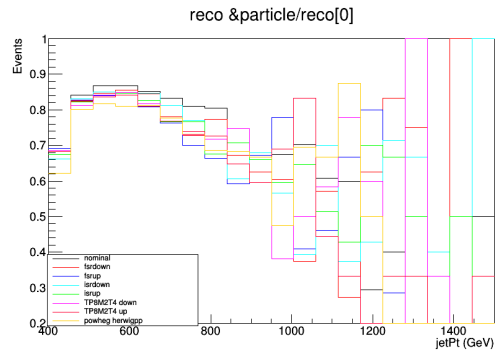
(a) Events Passing reco and particle cuts over particle cuts vs ptJJGen
 (b) Events Passing reco and particle cuts over particle cuts vs ptJJ



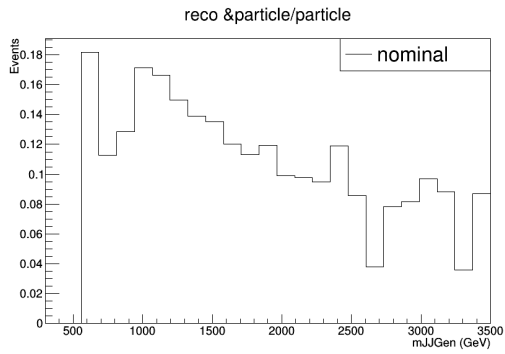
(a) Events Passing reco and particle cuts over particle cuts vs genJetPt for nominal Monte Carlo
 (b) Events Passing reco and particle cuts over particle cuts vs jetPt for nominal Monte Carlo



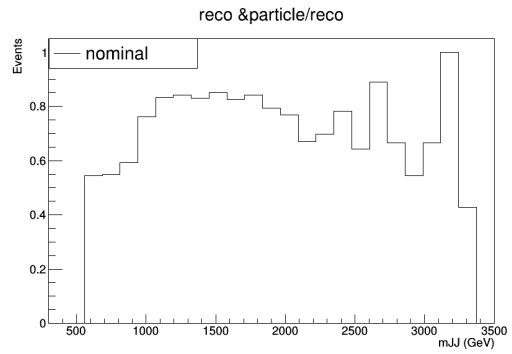
(a) Events Passing reco and particle cuts over particle cuts vs genPt



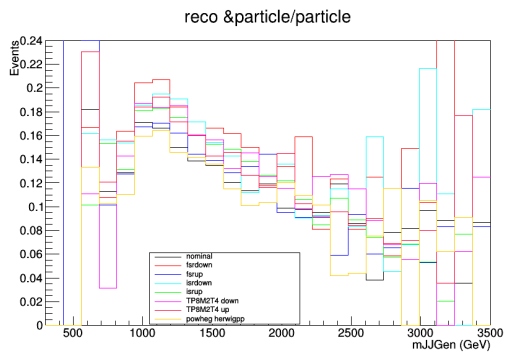
(b) Events Passing reco and particle cuts over particle cuts vs jetPt



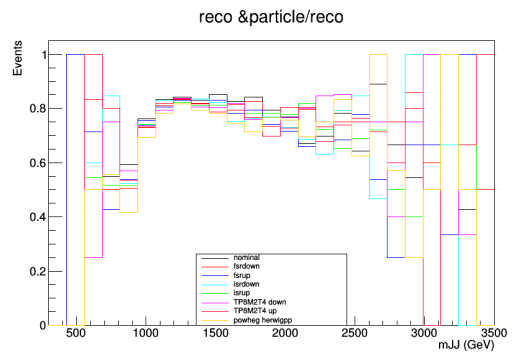
(a) Events Passing reco and particle cuts over particle cuts vs mJJGen for nominal Monte Carlo



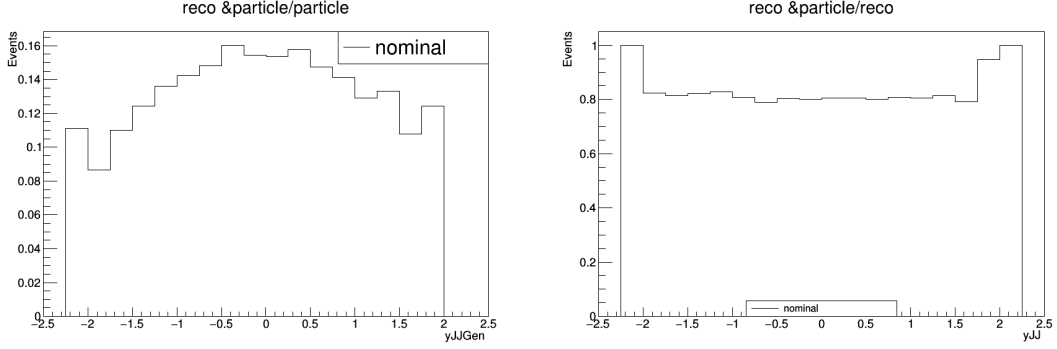
(b) Events Passing reco and particle cuts over particle cuts vs mJJ for nominal Monte Carlo



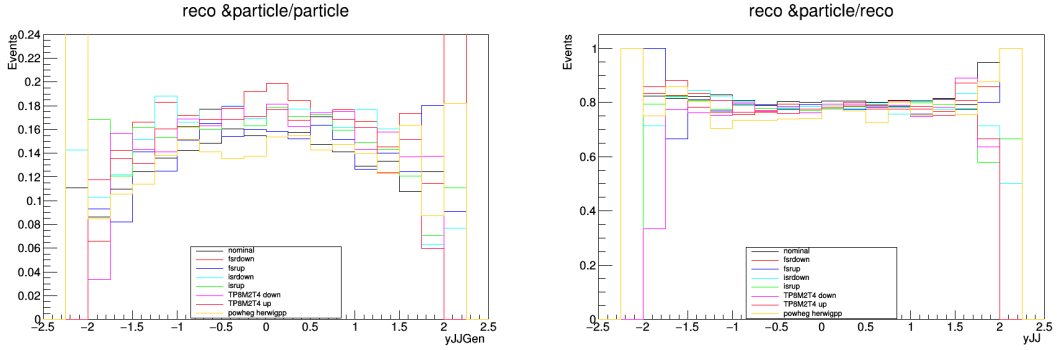
(a) Events Passing reco and particle cuts over particle cuts vs mJJGen



(b) Events Passing reco and particle cuts over particle cuts vs mJJ



(a) Events Passing reco and particle cuts over particle cuts vs yJJGen for nominal Monte Carlo
 (b) Events Passing reco and particle cuts over particle cuts vs yJJ for nominal Monte Carlo



(a) Events Passing reco and particle cuts over particle cuts vs yJJGen
 (b) Events Passing reco and particle cuts over particle cuts vs yJJ

As shown in the shown histograms, this procedure is carried out on 8 Monte Carlo Dataset including the nominal Monte Carlo. By carrying this procedure on all of the Monte Carlo datasets, one is given an error band for the correction factor that will be used finally in reproducing the requested particle level quantity.

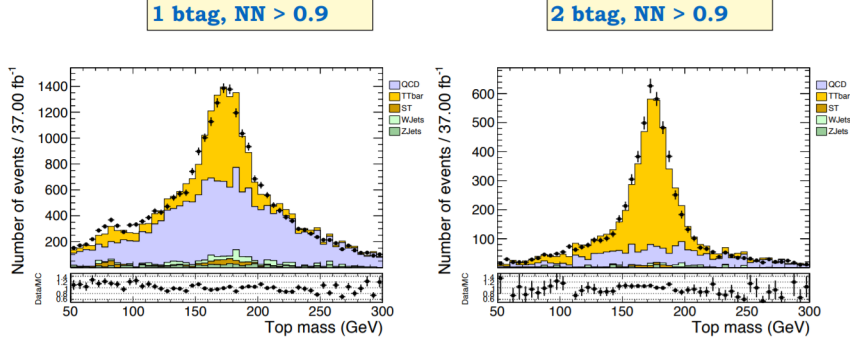
5.4 Signal Extraction -Fiducial Cross Section

The signal is extracted from data with a template fit on the top mass distribution

$$f_m(t) = N_S \times S(m_t) + N_{QCD} \times QCD(m_t) + N_{other} \times B(m_t)$$

where N_S, N_B, N_{other} are free parameters, and the last term is the contribution from the subdominant backgrounds.

The top mass estimator m_t is calculated using the SoftDrop Mass technique of the leading AK8 jet. The signal shape as well as the subdominant backgrounds



shape is taken from the simulation. The dominant background is the QCD multijet production. Its data driven estimate is calculated from the 0-btag sample with the exact kinematic cuts. A transition factor is applied, which is motivated by a MC dataset.

Top mass Estimator

The top mass estimator in our case is the SoftDrop mass of the leading jet. From the pictures below, it is evident that the top peak is visible in the 1-btag (exclusive) category too. The W peak is also visible in the 1-btag category. The combination of the 2-btag and the Neural Network requirements enhance the signal significantly.

The fiducial cross section is needed in order to calculate the theoretical differential cross section. The way that the fiducial cross section is calculated is using the signal:

$$S(x) = D(x) - R_{yield}N_{bkg}R_{shape}(x)B(x)$$

where $S(x)$ is the signal, $D(x)$ is the data, R_{yield} is the ratio between extended and reduced selection, N_{bkg} is the fitted background yield extended selection, R_{shape} is the background shape correction from non-closure and $B(x)$ is the background template taken from the 0-btag sample.

5.5 Uncertainties

This section covers which were the uncertainties in this analysis and the way that these uncertainties were calculated. The final calculations consist of two types of uncertainties. The first type is the statistical uncertainties. By statistical Uncertainties we mean the following:

- Statistics

- Errors from the Unfolding process
- Jet Energy Scale
- Jet Energy Radiation
- B tagging Efficiency

In this analysis, the statistical uncertainties were taken directly from the Root classes that were used for the extrapolation to the partonic and particle phase space.

The second type of uncertainties is the Theoretical Uncertainties. These uncertainties affect the extrapolation factors as well as the response matrices. The Theoretical uncertainties are strongly related to the type of Monte Carlo that was studied. Except from the nominal Monte Carlo Dataset, we saw earlier that several types of Monte Carlo Datasets were used in this analysis. The nominal Monte Carlo in this analysis was: TuneCUETP8M2T4_13TeV-powheg-pythia8. The rest of the Monte Carlo Datasets studied, had some of the options of the nominal monte carlo altered.

The way that the nominal Monte Carlo parameters were altered can be found from the Monte Carlo name in the sense that:

- Isr (Initial State Radiation) meaning that radiation is emitted at the beginning of the quark annihilation process (up or down quark) of the $t\bar{t}$ production.
- Fsr(Final State Radiation) meaning that radiation is emitted at the end of the process of the production of $t\bar{t}$ (b quark)
- "Tune" Meaning a set of Monte Carlo parameters for Pythia8 is altered ⁵

This analysis will focus on the Theoretical uncertainties.

5.5.1 Estimation of Theoretical Uncertainties

The sets of the different Monte Carlo datasets used in this analysis where:

- fsrUp ,1
- fsrDown , 2

⁵Powheg NLO process is the core Monte Carlo that is responsible for the simulation of the hard scatter and Pythia8 is the process responsible for the simulation of the hadronization and the multi-parton interactions

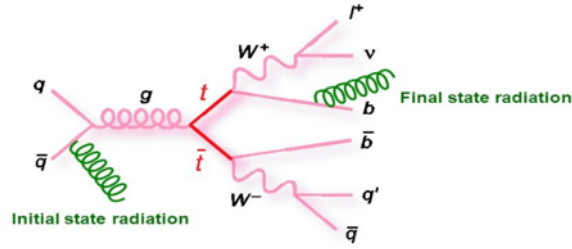


Figure 5.15: Isr and Fsr diversions from the Nominal Monte Carlo

- isrUp, 3
- isrDown, 4
- Tune Up, 5
- Tune Down, 6

For the final theoretical uncertainties of the analysis, they were calculated in the following way:

The method that we followed calculated the error bar in every bin of the value that was studied. The relative deviation from the nominal MC value was first calculated for each Monte Carlo.

$$e_i = |y_{nominal} - y_i|$$

where $i = 1..6$, $y_{nominal}$ is the calculated value of the nominal MC, and y_i is the calculated value for the respective MC studied.

Then for each type of the different MC (fsr, isr, tuned) we found the average of the values e_i .

$$e_{fsr} = \frac{1}{2}(e_1 + e_2)$$

$$e_{isr} = \frac{1}{2}(e_3 + e_4)$$

$$e_{tune} = \frac{1}{2}(e_5 + e_6)$$

The total error for each bin that was:

$$e_{total} = \sqrt{e_{fsr}^2 + e_{isr}^2 + e_{tune}^2}$$

In this way, the uncertainties are symmetrical over and under the calculated nominal value.

5.6 Reconstruction of Parton and Particle Quantities

The main purpose of this section is the calculation of the particle quantity and the parton quantity using the response matrices acquired from the Monte Carlo Nominal Dataset. The x_{gen} and x_{parton} are the particle quantities and the parton quantities that are going to be calculated respectively (topP, pt ttbar, y, ttbar mass). The x_{gen} is given by:

$$x_{particle} = \frac{1}{P_{gen}} \text{Response}(\text{reco, gen}) \frac{1}{P_{reco}} x_{reco}$$

And the x_{parton} :

$$x_{parton} = \frac{1}{P_{parton}} \text{Response}(\text{reco, parton}) x_{reco}$$

where x_{gen} refers to the particle level quantity, x_{reco} refers to the reco level quantity, x_{parton} refers to the parton quantity, $\text{Response}(\text{reco, gen})$ is the response matrix between the reco level and particle level quantity, $\text{Response}(\text{reco, parton})$ is the response matrix between the reco level and the particle level, $P_{gen} = \text{P}(\text{ pass gen \& reco cuts over gen cuts})$, $P_{reco} = \text{P}(\text{ pass gen \& reco cuts over reco cuts})$ and $P_{parton} = \text{P}(\text{pass Reco cuts and Signal Trigger / full partonic phase space})$.

5.6.1 Differential Cross Section

The differential cross section of the extrapolated data was calculated as followed: For the Particle Level

$$\frac{dN}{dX} \left[\frac{pb}{GeV} \right] = \frac{f(x_{particle})}{L\Delta X}$$

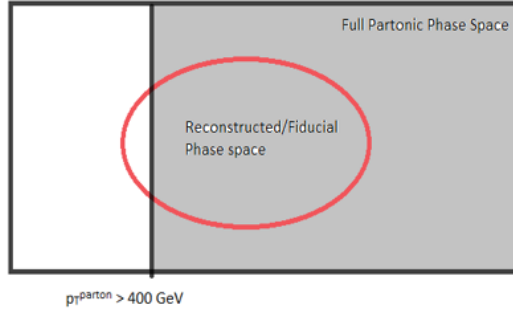
where X is the quantity studied each time and L is the total Luminosity (in this case 37000 pb^{-1}). In the same manner, the differential Cross section in the parton level of the extrapolated data

$$\frac{dN}{dX} \left[\frac{pb}{GeV} \right] = \frac{f(x_{parton})}{L\Delta X}$$

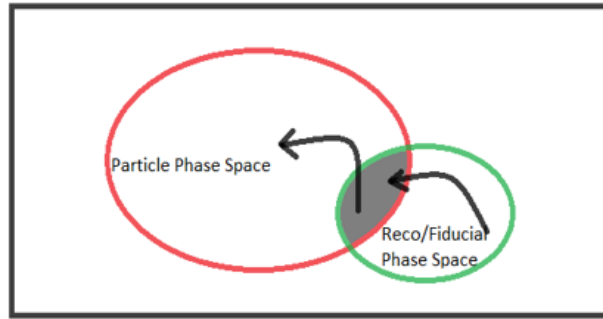
The theoretical differential Cross section for the particle and parton level was calculated as followed:

$$\left(\frac{dN}{dX} \right)_{theoretical}^{level} = \frac{1}{N_{total}} \frac{\sigma_{theory}}{\Delta X} x_{theory}^{level}$$

where level = particle or parton, N_{total} is the total number of events in each phase space, σ_{theory} is the fiducial total cross section that was calculated in section



(a) Extrapolation from reconstructed level to Parton Phase Space



(b) Extrapolation from reconstructed level to Particle Phase Space

5.4 and x_{theory}^{level} is the theoretical quantity studied taken from the nominal Monte Carlo for each level respectively.

5.6.2 Extrapolation to Parton Phase Space

The purpose of this subsection is the extrapolation from the reconstructed level phase space to the parton phase space. As it was written previously, the way to calculate the parton quantity is:

$$x_{parton} = \frac{1}{P_{parton}} \text{Response}(\text{reco, parton}) x_{reco}$$

The procedure for extrapolating to the partonic phase space consists of .. steps. The first step is to Unfold the x_{reco} quantity using the Response Matrix of reconstructed and partonic events. The x_{reco} quantity is produced from the new dataset for each systematic variation. Using the dynamic Unfolding code, the output of the function is the Unfolded histogram using the reco fiducial phase space and the

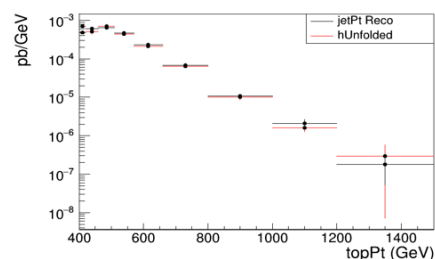
respective response matrix(reco, parton). The following step is to use the Unfolded Histogram and multiply each of its bin content with the respective extrapolation factor. This procedure is done in order to extrapolate from the reconstructed phase space to the full partonic phase space.

The final deliverable must be $\frac{d\sigma}{dX}$, where X is mass (mTTbar), rapidity (yT-Tbar), transverse momentum (ptTTbar) and top Transverse momentum (topPt) respectively. In order to transform the yield into sigma (σ), the $\frac{dN}{dX}$ quantity must be multiplied with the Luminosity. In this analysis, the Luminosity is $L = 37 \times 10^3 pb^{-1}$. As a next step, every bin must be divided with its bin width so it is scaled to 1. The True Partonic Phase space coming from the Nominal Monte Carlo must be multiplied with the total cross section σ and then be divided with the number of total events N_{total} , so that it is normalized to 1. The total cross section is $\sigma = 832 pb$ and the total events are $N_{total} \approx 70 \times 10^6$ events

The following histograms show

- 1. Unfolded Histogram with Reco quantity.
- 2. Extrapolation Factor (1 extrapolation factor for Parton level and 2 extrapolation factors for Particle Level)
- 3. Differential cross section after Final Extrapolation to the desired phase space (parton or particle) and theoretically calculated differential cross section
- 4. Fraction of Final Extrapolated Differential cross section over theoretical differential cross section.

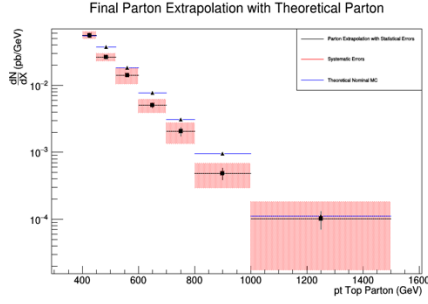
The histograms are shown with respect to all the physical quantities that are studied, meaning mass, transverse momentum (p_T), rapidity y and jet transverse momentum ($topP_T$).



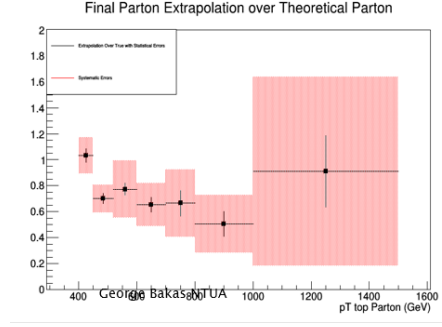
(a) Unfolded Histogram and Reco Quantity Histogram for top p_T in LogY scale



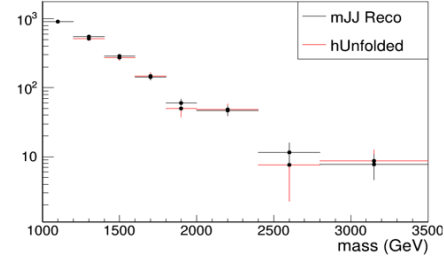
(b) Extrapolation factor for top p_T for parton level



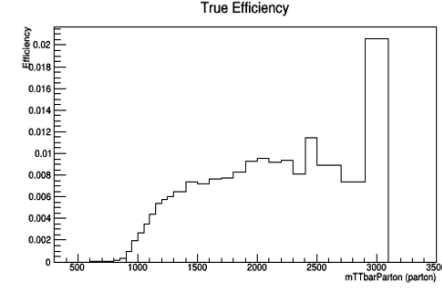
(a) Differential Cross Sections of extrapolated and theoretical (pb/GeV) for top p_T in LogY scale



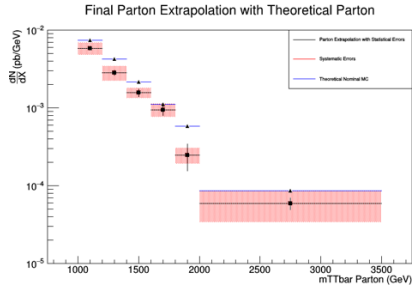
(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section



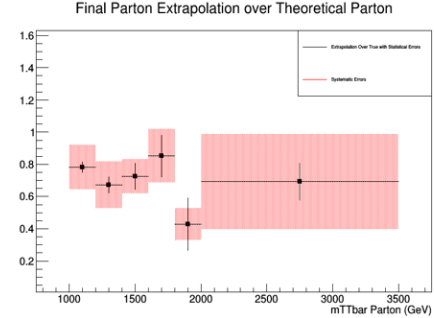
(a) Unfolded Histogram and Reco Quantity Histogram for mass of $t\bar{t}$ in LogY scale



(b) Extrapolation factor for mass of $t\bar{t}$ for parton level



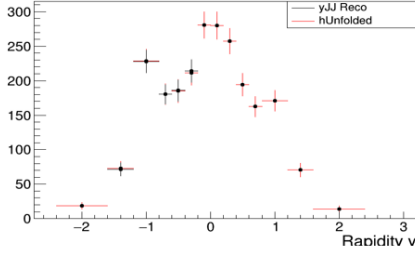
(a) Differential Cross Sections of extrapolated and theoretical (pb/GeV) for mass of $t\bar{t}$ in LogY scale



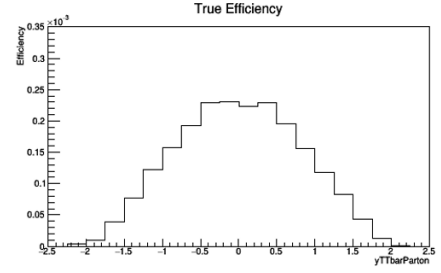
(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section for mass of $t\bar{t}$

5.6.3 Extrapolation to Particle Phase Space

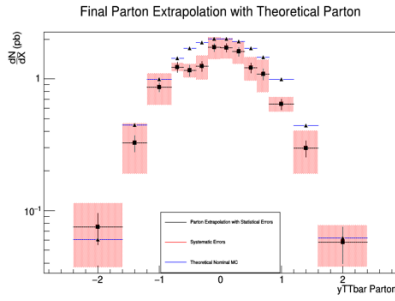
The purpose of this subsection is the extrapolation from the reconstructed level phase space to the particle level phase space. As it was written previously, the



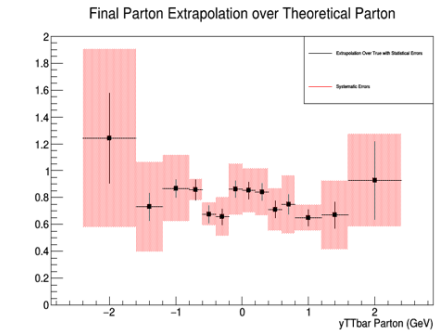
(a) Unfolded Histogram and Reco Quantity Histogram for Rapidity of $t\bar{t}$ in LogY scale



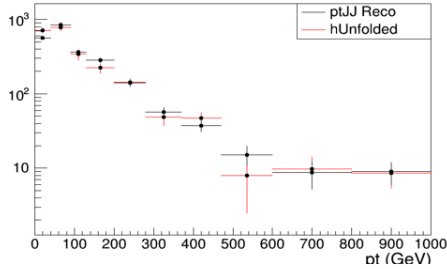
(b) Extrapolation factor for Rapidity of $t\bar{t}$ for parton level



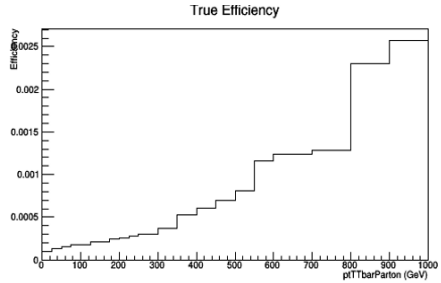
(a) Differential Cross Sections of extrapolated and theoretical (1/GeV) for Rapidity y of $t\bar{t}$ in LogY scale



(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section for Rapidity of $t\bar{t}$



(a) Unfolded Histogram and Reco Quantity Histogram for p_T of $t\bar{t}$ in LogY scale

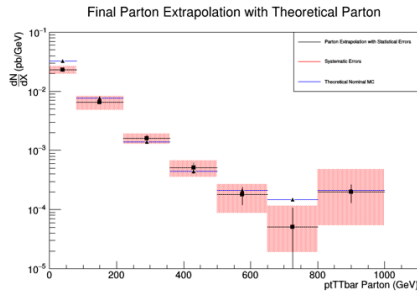


(b) Extrapolation factor for p_T of $t\bar{t}$ for parton level

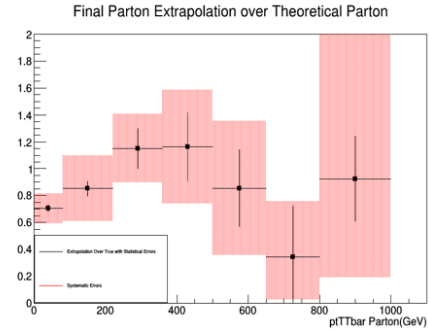
way to calculate the particle quantity is:

$$x_{particle} = \frac{1}{P_{gen}} \text{Response}(\text{reco, gen}) \frac{1}{P_{reco}} x_{reco}$$

The procedure of extrapolating to the particle phase space is the following;



(a) Differential Cross Sections of extrapolated and theoretical (pb/GeV) for p_T of $t\bar{t}$ in LogY scale

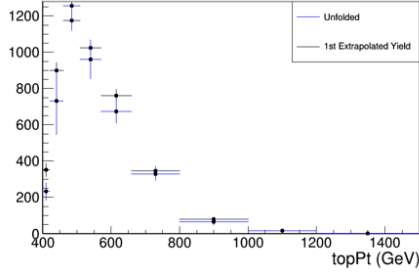


(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section for p_T of $t\bar{t}$

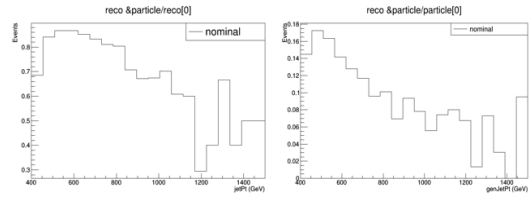
Firstly, the extrapolation from the reconstructed phase space to the common space between the reconstructed level and the particle level must be done. This is achieved by multiplying every bin content of the reconstructed histogram with the extrapolation factor (or correction factor) which is defined as the events passing the reconstructed and particle cuts, over the events passing only the reconstructed level cuts with respect to the reco level quantity. The second step of the procedure is to use the Dynamic Unfolding code so that the Unfolded histogram is produced using as input the Response Matrix of the reconstructed level and particle level events and the as reco histogram the new histogram produced by the multiplication of the correction factor with the fiducial reco level histogram. The Unfolded histogram is used lastly to extrapolate to the particle level phase space, by multiplying each of its bin content with the other extrapolation factor (or particle correction factor), which is defined as the fraction of events passing reconstructed and particle level cuts over events passing only particle cuts with respect to the particle level quantity. The final outcome of this process is the Histogram showing the full particle level phase space.

The following histograms show

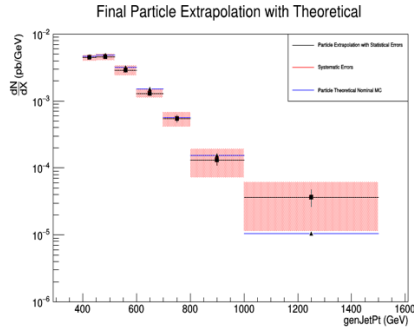
- 1. Unfolded Histogram with Reco quantity.
- 2. Errors of Unfolded Histogram with Errors of the Reco quantity histogram.
- 3. Extrapolated Histogram to the Particle Level phase space with the Re-binned Histograms from the Nominal Monte Carlo datasets showing the full Particle Level phase space.



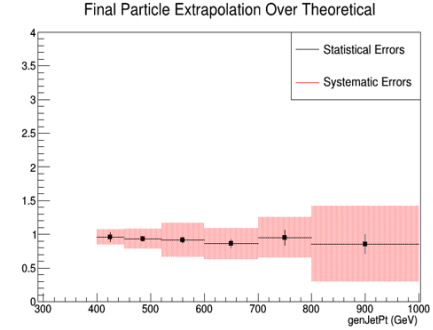
(a) Unfolded Histogram and Reco Quantity Histogram for top p_T in LogY scale



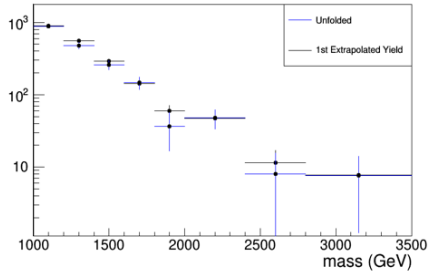
(b) Extrapolation factors for top p_T for particle level



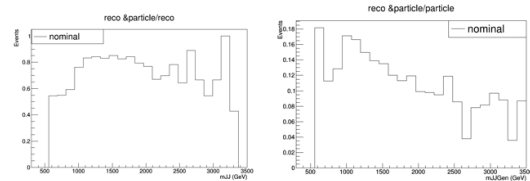
(a) Differential Cross Sections of extrapolated and theoretical (pb/GeV) for top p_T in LogY scale in Particle Level



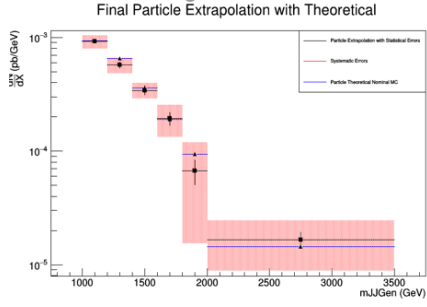
(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section in Particle Level



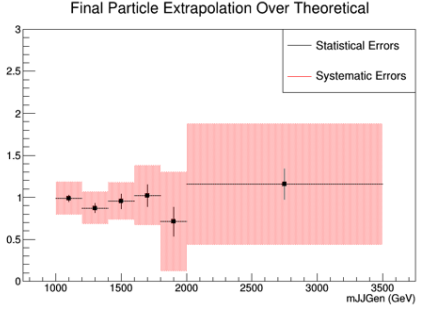
(a) Unfolded Histogram and Reco Quantity Histogram for mass of $t\bar{t}$ in LogY scale in Particle Level



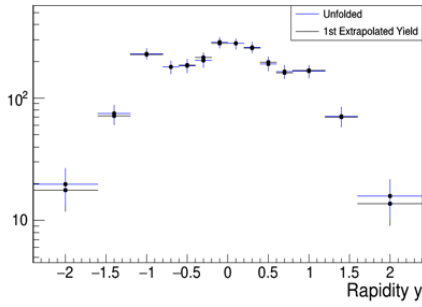
(b) Extrapolation factor for mass of $t\bar{t}$ for Particle Level



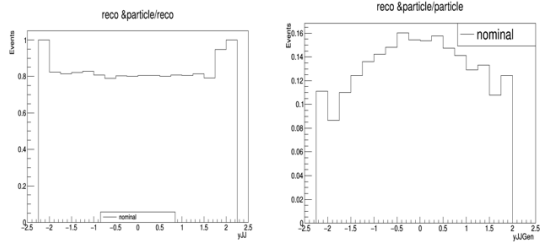
(a) Differential Cross Sections of extrapolated and theoretical (pb/GeV) for mass of $t\bar{t}$ in LogY scale in Particle Level



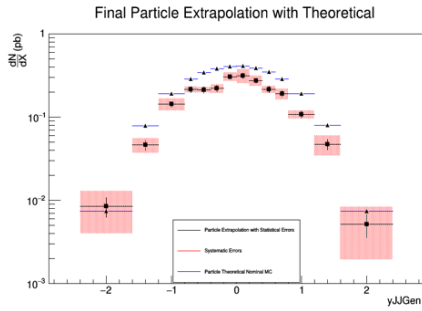
(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section for mass of $t\bar{t}$ in Particle Level



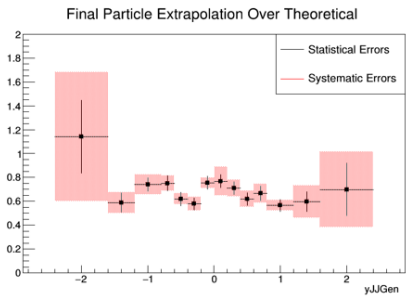
(a) Unfolded Histogram and Reco Quantity Histogram for Rapidity of $t\bar{t}$ in LogY scale in Particle Level



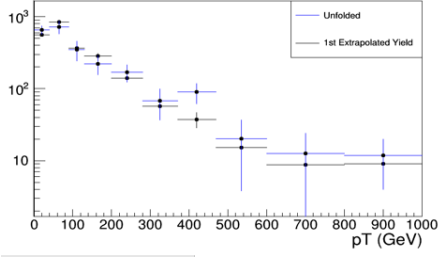
(b) Extrapolation factor for Rapidity of $t\bar{t}$ for Particle Level



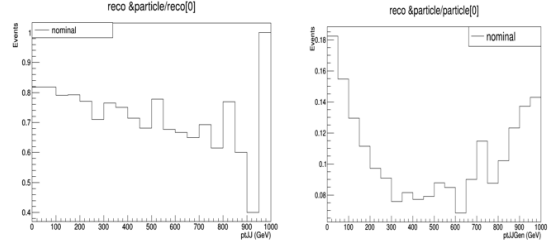
(a) Differential Cross Sections of extrapolated and theoretical (1/GeV) for Rapidity y of $t\bar{t}$ in LogY scale in Particle Level



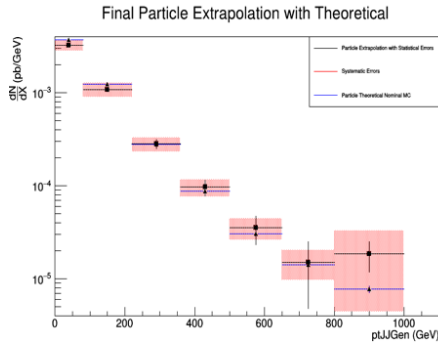
(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section for Rapidity of $t\bar{t}$ in Particle Level



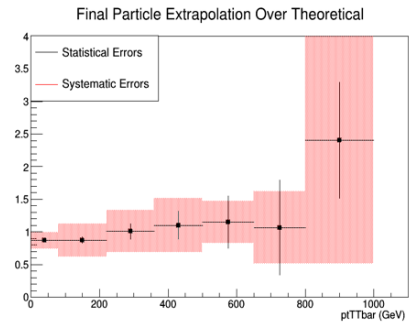
(a) Unfolded Histogram and Reco Quantity Histogram for p_T of $t\bar{t}$ in LogY scale in Particle Level



(b) Extrapolation factor for p_T of $t\bar{t}$ for Particle Level



(a) Differential Cross Sections of extrapolated and theoretical (pb/GeV) for p_T of $t\bar{t}$ in LogY scale in Particle Level



(b) Fraction of Extrapolated differential cross section over theoretical diff. cross section for p_T of $t\bar{t}$ in Particle Level

5.7 Conclusions

We have studied the $t\bar{t}$ production in proton proton (pp) collisions at 13 TeV centre of mass energy recorded by the CMS detector. Also we presented the differential cross sections at parton and particle levels.

In addition, the comparison with theory predictions show consistent shapes but systematically lower cross section in data. This is though a known effect, that was also reported by the ATLAS experiment and other CMS measurements [12] [13].

Finally, in this thesis we have focused on the unfolding of the measured cross sections to the parton and particle levels and we have investigated the impact of the theoretical (simulation) uncertainties on the final differential cross section results.

Appendices

Appendix A

Dynamic Unfold Code

This is the dynamic code used for Unfolding on particle or parton level. The code has 2 inputs. The 1st input is the response matrix used in the Unfolding method. The second input is the reconstructed quantity histogram. The output of this function is the desired Unfolded histogram.

Listing A.1: C++ code using listings

```
1 #include <TTree.h>
2 #include <TFile.h>
3 #include <TGraph.h>
4 #include <TLegend.h>
5 #include <TAttFill.h>
6
7 #include <iostream>
8 #include <TMatrixD.h>
9 #include <TMatrixDSparse.h>
10 #include <TMatrixDSym.h>
11 #include <TMatrixDSymEigen.h>
12 #include <TMath.h>
13 #include <TUnfold.h>
14
15
16 TH1F *dyn_unfold(TH2F *h_resp, TH1F *h_reco)
17 {
18     gStyle->SetOptStat(0);
19
20     int size_respx = h_resp->GetXaxis()->GetNbins();
21     int size_respy = h_resp->GetYaxis()->GetNbins();
22     int size_reco = h_reco->GetXaxis()->GetNbins();
```

```

23 cout<< size_resp_x<<endl;
24 cout<< size_resp_y<<endl;
25 cout<< size_reco<<endl;
26
27 //these is where the binning control is made
28 //checks if binning of the input histograms are correct
29 if (size_resp_x == size_resp_y && size_resp_x== size_reco )
30 {
31 int temp = size_resp_x;
32 for (int i =0; i<= temp; i++)
33 {
34     Double_t xlow_resp = h_resp->GetXaxis()->GetBinLowEdge(i);
35
36     Double_t ylow_resp = h_resp->GetYaxis()->GetBinLowEdge(i);
37
38     Double_t low_reco = h_resp->GetXaxis()->GetBinLowEdge(i);
39
40     Double_t low_2 = h_resp->GetXaxis()->GetBinLowEdge(i);
41
42     if(xlow_resp!=ylow_resp || xlow_resp!=low_reco ||
43         xlow_resp!=low_2)
44     {
45 cout<<"Number of bins is right. Variable binning is wrong!"<<endl;
46     }
47 }
48 }
49 else
50 {
51 cout<<"The number of bins on the histograms is wrong!"<<endl;
52
53 cout<<"Check and try again"<<endl;
54 abort();
55 }
56
57
58 //if everything is ok, begin the unfolding process
59
60 //1st step is to give the response matrix
61 TUnfold *unfold = new TUnfold(h_resp ,TUnfold::kHistMapOutputVert ,
62     TUnfold::kRegModeSize);

```

```

63
64 TCanvas *can1 = new TCanvas("1D","1D",900,600);
65 TLegend *leg = new TLegend(0.9,0.7,0.63,0.9);
66
67 //Unf Histogram must have same binning with Reco Hist
68 TH1F *hUnf;
69 hUnf = (TH1F*)h_reco -> Clone("hUnf");
70 hUnf->Clear();
71
72 float biasScale=0.0;
73 int size = hUnf->GetSize()-2;
74 cout<<size<<endl;
75
76 Int_t *binMap=new Int_t[size+2];
77 for(Int_t i=1;i<=size;i++) binMap[i]=i;
78 binMap[0]=-1;
79 binMap[size+1]=-1;
80
81 //The method starts with any parameter tau
82 Double_t tau = 10E-8;
83
84 unfold->DoUnfold(tau,h_reco,biasScale);
85
86 hUnf->SetLineColor(kBlue);
87 //the process for finding the best tau begins here
88 Int_t nScan=30; //this is number of scans
89
90 Double_t tauMin=1.0E-9;
91 Double_t tauMax=1.0;
92
93 Int_t iBest;
94 TSpline *logTauX,*logTauY;
95 TGraph *lCurve;
96
97 // this method scans the parameter tau
98 // finds the kink in the L curve
99 // finally, the unfolding is done for the "best" choice of tau
100
101 iBest=unfold->ScanLcurve(nScan,tauMin,tauMax,
102 &lCurve,&logTauX,&logTauY);

```

```
103
104
105 cout<<" Best tau: " <<iBest<<endl;
106 cout<<" tau=" <<unfold->GetTau()<<endl;
107 unfold->DoUnfold( unfold->GetTau() , h_reco , biasScale );
108 unfold->GetOutput( hUnf , binMap );
109
110 hUnf->Draw();
111
112 leg->AddEntry( hUnf , " hUnf" , " l" );
113 h_reco->SetLineColor( kRed );
114 h_reco->Draw( " same" );
115
116 leg->AddEntry( h_reco , " hist reco" , " l" );
117 leg->Draw();
118
119 return hUnf;
120 }
```

Βιβλιογραφία

- [1] D. Griffiths, “Introduction to Elementary Particles, Revised Edition”. WILEY-VCH, 2008.
- [2] CMS Collaboration, “Top Quark Mass Measurements”. <http://cms-results.web.cern.ch/cms-results/public-results/publications/TOP/index.html>. [Online; accessed July 10, 2017]
- [3] Particle Data Group *Review of Particle Physics 2016*, Chinese Physical Society, 1674-1137
- [4] CMS Collaboration, “Search for t t-bar resonances in highly boosted lepton+jets and fully hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV”, doi:10.1007/JHEP07(2017)001, arXiv:1704.03366
- [5] CMS Collaboration, “The CMS Experiment at the CERN LHC”, JINST 3 (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [6] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm”, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [7] F. Beaudette, “The CMS Particle Flow Algorithm”, arXiv:1401.8155
- [8] A. J. Larkoski, S. Marzani, G. Soyez et al., “Soft Drop”, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.
- [9] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, Journal of Instrumentation 8 (2013), no. 04, P04013
- [10] O. Behnke, K. Kröniger, G. Schott, T. Schörner-Sadenius *Data Analysis in High Energy Physics*. WILEY-VCH, 978-3-527-41058-3
- [11] Luca Lista *Statistical Methods for Data Analysis in Particle Physics*, Springer, 978-3-319-20175-7

- [12] Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s}=13$ TeV using the ATLAS detector
- [13] CMS Collaboration, “Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV”, doi:10.1103/PhysRevD.95.092001, arXiv:1610.04191
- [14] Top Quark physics with B-Jets containing muons with CMS
<http://www.iihe.ac.be/publications/MasterThesis-StijnBlyweert.pdf>
- [15] Measurement of the Jet Mass Distribution in Boosted Top Quark Decays
http://www.iexp.uni-hamburg.de/groups/pd/sites/default/files/filefield_paths/Ma
- [16] A. Pich, “The Standard Model of Electroweak Interactions”, 2005.
- [17] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235
- [18] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214