

**UNIVERSITÀ DEGLI STUDI DI NAPOLI
“FEDERICO II”**



Scuola Politecnica e delle Scienze di Base

Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Triennale in Fisica

**Ricostruzione di quark top ad alto impulso tramite
algoritmi di machine learning per la ricerca di una
nuova particella W’ con il rivelatore CMS ad LHC**

Relatori:

Alberto Orso Maria Iorio

Candidato:

Daniele Boccanfuso

Matr. N85001274

Anno Accademico 2019/2020

Alla mia famiglia

Indice

Introduzione	3
1 LHC e CMS	5
1.1 LHC	5
1.1.1 Caratteristiche e funzionamneto di LHC	5
1.2 L'esperimento CMS	7
1.2.1 Sistema di coordinate	8
1.2.2 I sottorivelatori	9
1.2.3 Il sisitema a Muoni	11
1.2.4 Il sistema di trigger e acquisizione dati	12
2 Il Modello Standard	13
2.1 Panoramica del Modello Standard	13
2.1.1 Le particelle elementari	14
2.1.2 L'interazione elettromagnetica	16
2.1.3 L'interazione forte	16
2.1.4 l'interazione debole	17
2.1.5 il Bosone di Higgs	18
2.2 Il quark top	18
2.2.1 Processi di produzione	19
2.2.2 Processi di decadimento	19
2.3 Fisica oltre il Modello Standard	20
2.3.1 Il bosone W'	20
2.3.2 Anomalie nel settore B	20
3 Ricostruzione degli oggetti fisici	21
3.1 Fenomeno in esame e campioni	21
3.1.1 Il decadimento di W'	21
3.1.2 I campioni di dati	22
3.2 Ricostruzione dei quark top standard	23
3.2.1 Muoni Tight	23

<i>INDICE</i>	2
3.2.2 Jet b	24
3.3 Machine Learning	24
3.3.1 Decision Tree	24
3.3.2 XGBoost	25
3.3.3 Variabili dell'allenamento	28
3.3.4 Efficienze	29
3.4 Ricostruzione W'	32
3.4.1 Ricostruzione della massa invariante	32
3.4.2 Confronto con l'analisi standard	34
Conclusione	35
Bibliografia	37

Introduzione

Il *Large Hadron Collider*, in breve LHC, è il più grande acceleratore di particelle al mondo. Grazie ai suoi 27 km di circonferenza è in grado di accelerare e far collidere fasci di protoni fino a energie nel centro di massa di 13 TeV e luminosità istantanea di $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Queste condizioni sono ideali per eseguire misure di precisione e ricerca di nuove particelle. Uno degli esperimenti che si trovano a LHC è il *Compact Muon Solenoid* (CMS), un grande rivelatore a indirizzo generale pensato per studiare al meglio i complessi risultati derivanti dalle collisioni.

La teoria di riferimento in fisica subnucleare è il Modello Standard, attualmente è senza dubbio la teoria che meglio descrive le particelle e le loro interazioni. Le sue predizioni sono quasi tutte in eccellente accordo con gli esperimenti, ed esso descrive in modo soddisfacente tre delle quattro forze fondamentali. Tuttavia il Modello Standard non può dirsi una teoria esaustiva, infatti non predice il valore di molti parametri al suo interno, non è in grado di spiegare la prevalenza della materia ordinaria sull'antimateria, non contempla assolutamente la forza gravitazionale e altro ancora. Per questo motivo numerosissimi sono i modelli alternativi o che estendono il Modello Standard, chiamati genericamente "*Beyond the Standard Model*" (BSM).

Molti modelli prevedono l'esistenza di un nuovo bosone carico, massivo chiamato W' , con caratteristiche simili alla sua controparte del Modello Standard. Esistono modelli che prevedono un rapporto privilegiato tra questo bosone e la terza famiglia di quark, particolarmente interessanti per le potenziali implicazioni legate alle anomalie nella fisica dei mesoni B recentemente scoperte. In questo lavoro ci concentreremo sul decadimento in quark top e quark bottom, $W' \rightarrow tb$.

I dati per questa analisi sono stati generati con un algoritmo di simulazione MonteCarlo ipotizzando una massa per W' di 4 TeV. Con una tale massa i prodotti di decadimento avranno necessariamente un alto impulso, questo fa sì che le particelle nello stato finale siano molto ravvicinate tra loro, il che può rendere difficile la ricostruzione.

Allo scopo di migliorare il riconoscimento dei quark top useremo l'algoritmo di machine learning supervisionato chiamato XGBoost. Tale algoritmo è usato per

classificare i quark top presenti in un evento a partire dalle informazioni ricavate dal detector. Questo approccio si propone di migliorare l'efficienza rispetto alla ricostruzione usuale, che avviene tramite l'applicazione di richieste in sequenza su singole variabili. Sono presentati i parametri e i risultati dell'allenamento, la ricostruzione del bosone W e il confronto con la tecnica di analisi standard.

La tesi ha la seguente struttura:

- il primo capitolo descrive l'acceleratore LHC e l'esperimento CMS;
- il secondo capitolo si concentra sul Modello Standard e possibili estensioni di esso;
- il terzo capitolo descrive la tecnica di analisi e i risultati.

Capitolo 1

LHC e CMS

1.1 LHC

Il *Large Hadron Collider*, abbreviato in LHC[10], è il più grande acceleratore di particelle al mondo, costruito dall' *Organizzazione europea per la ricerca nucleare* (CERN) tra il 1998 e il 2008 a Ginevra, con lo scopo di cercare risposta alle domande fondamentali della fisica subnucleare, come l'origine della massa e l'asimmetria tra materia e antimateria, o la ricerca della materia oscura e in generale tutti i possibili segnali di fisica oltre il Modello Standard, accessibili alle sue energie. L'acceleratore è situato in un tunnel di 27 km di circonferenza, precedentemente occupato dal *Large Electron-Positron Collider* (LEP), a 100 m di profondità. All'interno collidono fasci di adroni ad alta energia, ricreando condizioni simili a quelle che si trovavano negli istanti successivi al Big Bang. Grazie alle energie raggiunte e alla quantità di dati raccolta, nel 2012 è stato possibile osservare per la prima volta il bosone di Higgs, tassello fondamentale del Modello Standard, negli esperimenti CMS e ATLAS.

1.1.1 Caratteristiche e funzionamento di LHC

Il Large Hadron Collider accelera fasci di protoni o ioni di piombo, raggiungendo finora energie nel centro di massa di 13 TeV e 1,4 TeV rispettivamente. I protoni si ottengono facilmente ionizzando un gas di idrogeno. Prima di essere iniettate in LHC le particelle attraversano un lungo sistema di acceleratori, composto dalle seguenti macchine: il *Linear Accelerator 2* (LINAC2), che porta i protoni da fermi a 50 MeV, il *Proton Synchrotron Booster* (PSB), che porta l'energia a 1.4 GeV, il *Proton Synchrotron* (PS), che accelera i protoni fino a 25 GeV, e infine il *Super Proton Synchrotron* (SPS), che permette di raggiungere l'energia di 450 GeV. Ognuna di queste macchine era usata come acceleratore autonomo prima di

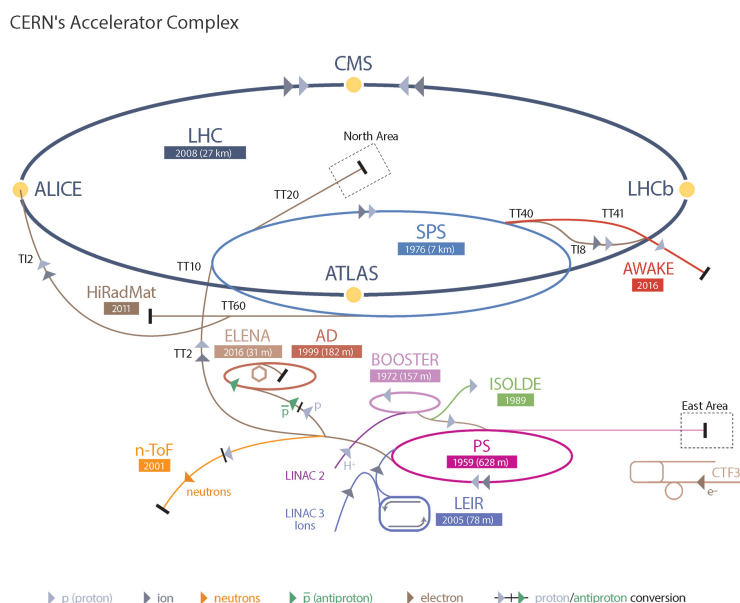


Figura 1.1: Il complesso di acceleratori al CERN

diventare un iniettore per l'acceleratore successivo, lo stesso LHC in futuro potrebbe essere usato come iniettore per una macchina ancora più grande: il *Future Circular Collider* (FCC), un acceleratore da 100 km di diametro. Per curvare la traiettoria di particelle che viaggiano ad altissima velocità sono necessari intensi campi magnetici, LHC è dotato di 1232 dipoli magnetici per curvare il fascio, ciascuno lungo 15 metri e pesante 35 tonnellate, e 392 quadrupoli magnetici per collimarli, oltre a esapoli, ottupoli e decapoli che correggono piccole imperfezioni nell'orbita dei protoni. Tali magneti sono superconduttori costituiti da una lega di Niobio e Titanio, in cui circolano correnti di 11kA che generano campi di 8,3 Tesla, ma per funzionare devono essere mantenuti a una temperatura di 1,9°K. Tale temperatura è mantenuta costante grazie a un imponente impianto di raffreddamento a Elio superfluido. Un fascio di LHC è formato da circa 2800 pacchetti (*bunches*), ciascuno contenente circa 1.15×10^{11} protoni, che collidono ogni 25 ns.

Ovviamente solo pochi protoni interagiscono tra di loro, il punto di collisione prende il nome di vertice primario. Da qui si generano le particelle oggetto di studio, quelle a vita media più lunga procedono fino ad arrivare al rivelatore, altre decadono in punti detti vertici secondari. I vertici sono rappresentati in figura 1.2.

I bunch collidono solo in 4 punti lungo l'anello, dove sono collocati i 4 esperimenti principali:

- ALICE (*A Large Ion Collider Experiment*): ottimizzato per lo studio di

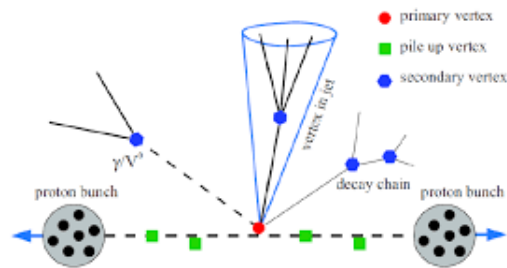


Figura 1.2: Vertici primari e secondari in una collisione

collisioni tra ioni di Piombo, ha lo scopo di osservare una fase della materia chiamata plasma di quark e gluoni (QGP), l'esistenza e le proprietà di questa fase sono indispensabili per la comprensione della cromodinamica quantistica.

- **ATLAS (*A Toroidal LHC ApparatuS*)**: è un rivelatore a indirizzo generale in grado di esplorare una vasta gamma di fenomeni, come quelli legati ai quark top e beauty o al bosone di Higgs, o processi al di fuori del Modello Standard.
- **CMS (*Compact Muon Solenoid*)**: rivelatore a indirizzo generale, caratterizzato dal suo design per la rivelazione dei muoni, verrà descritto in dettaglio nella prossima sezione.
- **LHCb**: un rivelatore pensato per lo studio del quark bottom, sfrutta 2 fasci di diversa energia, uno a 7 TeV e l'altro a 450 GeV, mira a trovare una spiegazione per l'asimmetria tra materia e antimateria.

1.2 L'esperimento CMS

Il Compact Muon Solenoid[15] è un rivelatore progettato per studiare il maggior numero possibile di risultati nelle collisioni di LHC, dalle misure di precisione sul Modello Standard alla ricerca di nuova fisica. Per questo motivo è costituito da diversi sottomoduli in grado di misurare traiettoria, massa, energia, impulso e carica di un'ampia gamma di particelle, difatti le uniche particelle che sfuggono alla rivelazione sono i neutrini. Dal momento che le particelle prodotte nelle collisioni sono dirette in tutte le direzioni, CMS è studiato in modo da coprire completamente l'angolo solido, eccetto per una regione attorno ai tubi di fascio, e ha una forma cilindrica per sfruttare al meglio la simmetria delle collisioni. Il corpo centrale prende il nome di *barrel* e i tappi laterali si chiamano *endcap*, è lungo

21 m, ha un diametro di 15 m e pesa 14000 tonnellate. L'elemento caratteristico di CMS è il grande solenoide superconduttore da cui prende il nome, esso ha un diametro di 7 m e crea un campo magnetico in grado di curvare la traiettoria delle particelle cariche ad alta energia e permettere accurate misure dell'impulso.

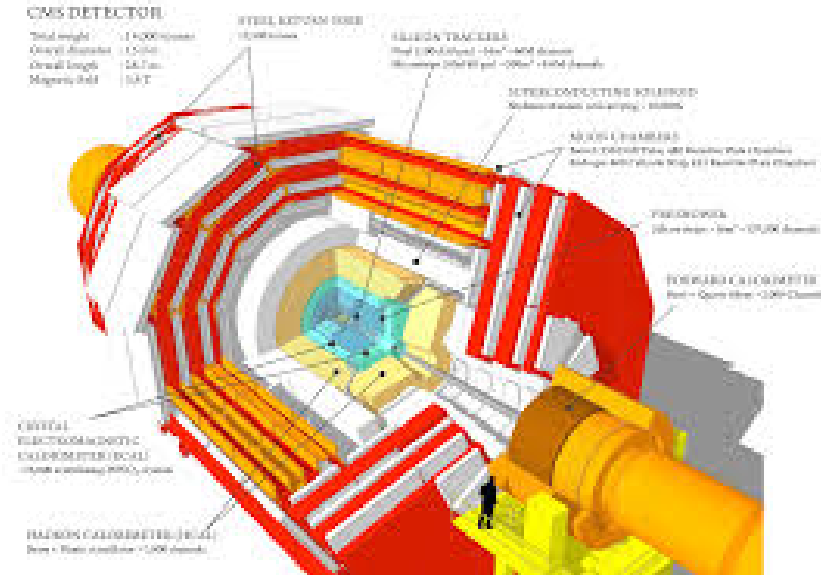


Figura 1.3: rappresentazione grafica di CMS

1.2.1 Sistema di coordinate

Per CMS si adotta un sistema di coordinate destrorso centrato nel punto di collisione nominale, l'asse x punta verso il centro dell'anello di LHC e funge da coordinata orizzontale, l'asse y punta verso l'alto e funge da coordinata verticale, l'asse z è diretto lungo il fascio di LHC che circola in senso antiorario. Tuttavia, data la simmetria della struttura, per gli oggetti fisici ricostruiti si utilizza un sistema di coordinate cilindriche, dove r è la coordinata radiale dall'asse z , l'angolo azimutale ϕ è l'angolo di rotazione attorno z , con origine sull'asse x e crescente in senso orario guardando lungo la direzione positiva di z , l'angolo polare θ è l'angolo di rotazione attorno x , con origine sull'asse z e crescente in senso orario guardando lungo la direzione positiva di x . In fisica subnucleare è conveniente usare, invece dell'angolo θ , la pseudorapidità η definita come:

$$\eta = -\ln \left[\tan \frac{\theta}{2} \right]$$

questa quantità si trasforma linearmente sotto boost di Lorentz lungo l'asse z , in più, in queste trasformazioni, la differenza $\Delta\eta$ tra due particelle relativistiche

è invariante. Come si può vedere dalla figura 1.4, mentre θ cresce da zero η

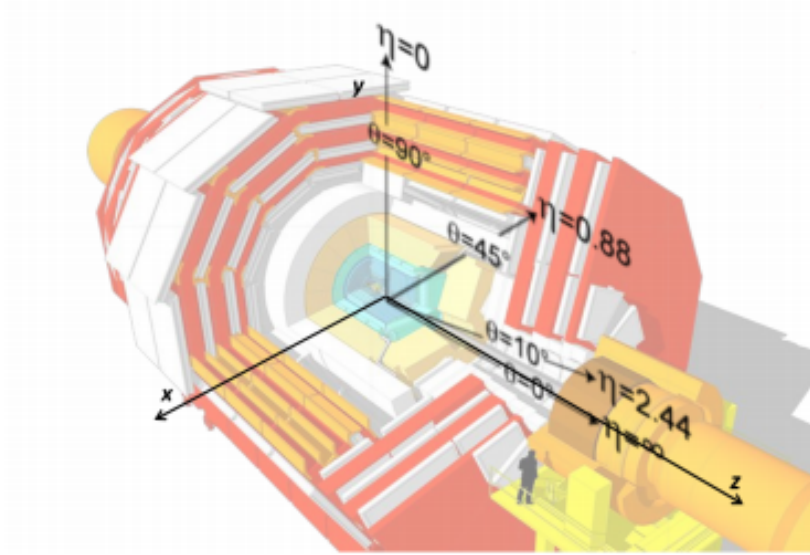


Figura 1.4: Il sistema di coordinate all'interno di CMS e la dipendenza funzionale di η da θ

decresce da infinito. In queste coordinate la distanza angolare ΔR tra due oggetti è un invariante per boost lungo l'asse z :

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

Altre variabili molto importanti, definite nel sistema cartesiano, sono il momento trasverso e l'energia trasversa:

$$\vec{p}_t = \sqrt{\vec{p}_x^2 + \vec{p}_y^2}$$

$$E_t = E \sin \theta$$

dove p_x e p_y sono le componenti x e y del momento di una particella, ed E è la sua energia.

1.2.2 I sottorivelatori

L'apparato sperimentale di CMS è formato da sottorivelatori, oltre a un sistema di trigger e acquisizione dati (DAQ). È fondamentale che ciascun sottorivelatore abbia una risposta molto rapida per evitare perdite di informazione durante i tempi morti. La figura 1.5 mostra chiaramente la loro disposizione, partendo dal centro abbiamo:

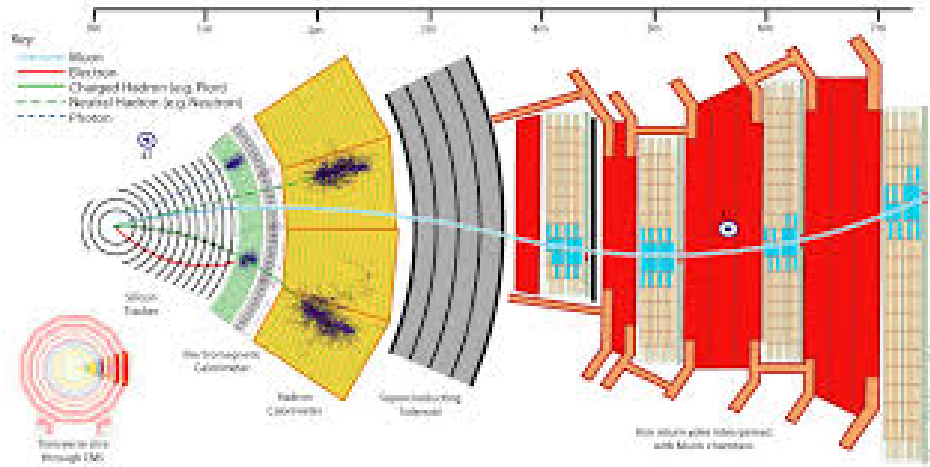


Figura 1.5: sezione trasversale di CMS

- Il *silicon tracker*[3], copre la regione con $|\eta| < 2.5$ e $r < 1.2$ m, ricostruisce accuratamente la posizione di muoni, elettroni e adroni per misurarne il momento e localizzare il vertice principale, essendo l'elemento più interno riceve il più alto volume di particelle ed è stato costruito in materiali resistenti alle radiazioni. È composto da due parti:
 - il rivelatore a pixel, più interno, fondamentale per la ricostruzione di particelle a brevissima vita media, misura la posizione con una precisione di $10 \mu m$ per le misure radiali e $20 \mu m$ per quelle trasversali, grazie ai suoi 65 milioni di pixel al silicio;
 - il rivelatore a microstrips, più esterno, ha una risoluzione tra i $35 \mu m$ e i $52 \mu m$ in direzione radiale, e i $530 \mu m$ in direzione trasversale, consiste in 4 strati di strip al silicio parallele al fascio;

il principio di funzionamento di questo rivelatore è basato sulle proprietà del silicio drogato: attraversando il mezzo le particelle formano in esso coppie elettrone-lacuna. Applicando un campo elettrico si muovono gli elettroni verso un sensore che li raccoglie sotto forma di impulso elettrico della durata di un nanosecondo. È importante che questo processo sottragga meno energia possibile alla particella incidente.

- Il *calorimetro elettromagnetico* (ECAL)[1], copre la regione con $|\eta| < 3$ e 1.25 m $< r < 1.8$ m, misura l'energia dei fotoni e degli elettroni. È formato da 75848 cristalli di uno scintillatore inorganico solido: il tungstenato di piombo ($PbWO_4$). Quando una particella penetra nel mezzo ne eccita gli elettroni, questi si diseccitano emettendo luce, che viene assorbita da

fotomoltiplicatori, che quindi restituiscono un segnale proporzionale all'energia del proiettile. Il $PbWO_4$ ha una lunghezza di radiazione di 0.89 cm e un raggio di Moliere di 2.2 cm il che permette a ECAL di assorbire completamente gli sciame elettromagnetici e distinguerli, inoltre emette l'85% della luce in 25 ns, proprio il tempo che intercorre tra la collisione di 2 bunch.

- Il *calorimetro adronico* (HCAL)[2], copre la regione con $|\eta| < 5$, misura l'energia degli adroni, ovvero protoni, neutroni, kaoni e pioni. È un calorimetro a campionamento in grado di contenere completamente gli sciame adronici: si alternano strati di materiale assorbente a strati di materiale attivo. Il materiale assorbente è l'ottone, che ha una lunghezza di radiazione di 1.49 cm e una lunghezza di interazione nucleare di 16.42 cm. Il calorimetro adronico gioca un ruolo cruciale nella determinazione dell'energia trasversa mancante, una misura indiretta dell'energia di particelle non interagenti (e.g. neutrini) o addirittura di nuove particelle.
- Il *solenioide*, è un elettromagnete composto da spire superconduttrici, genera al suo interno un campo uniforme di 3.8T, parallelo all'asse z . Il magnete è lo strumento attorno al quale è costruito l'intero esperimento, il suo compito è quello di curvare le particelle cariche. Nota la curvatura di una particella è infatti possibile determinare il suo impulso e la sua carica. Al di fuori del solenoide si trova una struttura in ferro chiamata *iron yoke*, lett. giogo metallico. L'*iron yoke* riduce gli effetti di bordo del solenoide, e permette di chiudere le linee di campo magnetico in maniera quasi uniforme anche all'esterno del solenoide, dove l'intensità del campo è di circa 2T. Questo consente una doppia misura del momento dei muoni.
- Il sistema a muoni, copre la regione con $|\eta| < 2.4$, usato per l'identificazione, la misura del momento e per il trigger dei muoni, verrà discusso in dettaglio nella prossima sezione.

1.2.3 Il sistema a Muoni

Come suggerisce il nome dell'esperimento, la rivelazione dei muoni è di centrale importanza per CMS. I muoni sono particelle in tutto e per tutto uguali agli elettroni, fatta eccezione per la loro massa che è circa 200 volte superiore, e per la loro vita media di 2.2×10^{-6} s. I muoni possono attraversare spessi strati di materiale e arrivare sino alla parte esterna di CMS. Sono un mezzo molto potente per distinguere processi fisici di interesse dal fondo. Il sistema a muoni copre una superficie di 25000 m^2 , utilizza tre tipi diversi di rivelatori a gas: camere a drift (Drift Tubes, DT), camere a strip (Cathode Strip Chambers, CSC), camere

a piatti resisitivi (Resistive Plate Chambers, RPC)[11]. La traiettoria è ricostruita combinando le tracce nel sistema a muoni e quelle nel silicon tracker.

- Le camere a drift sono posizionate nel barrel e coprono fino a una pseudo-rapidità di 1.2. Sono composta da 3 *superlayers* (SL) a loro volta composti da 4 strati. Nei 2 SL esterni i fili anodici sono paralleli alla direzione del fascio e misurano le coordinate $r - \phi$ con precisione di $100 \mu m$, nella camera interna i fili sono ortogonali al fascio e misurano la coordinata z . Gli strati sono divisi da celle di metà grandezza per evitare punti ciechi.
- Le camere a strip si trovano negli endcap e coprono la regione con $0.9 < |\eta| < 2.4$, sono 468 in totale. Le CSC devono avere un'alta risoluzione, risposta rapida e alta resistenza alla radiazione, poichè negli endcaps il rate dei muoni è più alto e il campo magnetico non è uniforme. Le CSC forniscono una misura di $r - \phi$ e di η .
- Le camere a piatti resistivi si trovano sia nel barrel che negli endcaps, hanno un'ottima risoluzione temporale ma una risoluzione spaziale più bassa delle DT e CSC. Le RPC sono in grado di identificare l'istante di una interazione in meno dei 25 ns che intercorrono tra le collisioni dei bunch. Quindi il sistema di trigger per muoni basato sulle RPC è in grado di stabilire la collisione a cui associare la traccia del muone anche in caso di rate molto elevati.

1.2.4 Il sistema di trigger e acquisizione dati

Il sistema di trigger e acquisizione dati (DAQ), serve a selezionare e immagazzinare i dati degli eventi per la successiva analisi. Tuttavia immagazzinare tutti i dati provenienti dalle collisioni, considerando l'altissima frequenza a cui avvengono, è impossibile con le attuali tecnologie. È quindi necessario selezionare gli eventi per permettere la scrittura sulla memoria che attualmente avviene con frequenza di 100 Hz. Il sistema di trigger seleziona gli eventi più significativi, è costituito da due fasi: il *Level-1 trigger* (L1) fornisce una selezione rapida e automatica, basata sui depositi di energia nei calorimetri e tracce dei muoni, la decisione finale sull'accettare o meno un evento spetta all'*High Level Trigger* (HLT), che esegue una selezione basata su un software.

Capitolo 2

Il Modello Standard

2.1 Panoramica del Modello Standard

Il Modello Standard (MS) è la teoria fisica che meglio descrive le particelle elementari e le interazioni tra di esse[13]. Ad oggi sono note quattro interazioni fondamentali: l'interazione forte, l'interazione elettromagnetica, l'interazione debole e l'interazione gravitazionale. Quest'ultima non è descritta dal Modello Standard, mentre quella elettromagnetica e quella debole sono unificate in una sola interazione detta elettrodebole; è previsto che ad alte energia anche l'interazione forte si possa unificare alle prime due. Il Modello Standard ha ricevuto numerose conferme sperimentali nel corso degli ultimi decenni, tra cui alcune delle più importanti sono: la scoperta dei bosoni W e Z nel 1984 al CERN, la scoperta del quark top nel 1995 al Fermilab, la scoperta del bosone di Higgs nel 2012 al CERN, grazie agli esperimenti ATLAS e CMS. Ciò nonostante il Modello Standard è lungi da potersi dire una teoria completa, alcuni dei suoi limiti sono:

- non è in grado di descrivere la forza gravitazionale;
- predice una massa nulla per i neutrini, che invece è noto siano dotati di una piccola massa dagli esperimenti astroparticellari;
- non predice l'esistenza della materia oscura, di cui si hanno prove da osservazioni astrofisiche;
- non predice il valore di molti suoi parametri, tra cui le masse delle particelle e le costanti di accoppiamento;
- ha bisogno di termini correttivi molto grandi per essere compatibile con la massa del bosone di Higgs misurata (circa 125 GeV).

Per risolvere questi problemi i fisici sono impegnati da anni nella ricerca di evidenze dirette di nuova fisica oltre il Modello Standard, i cui modelli sono indicati genericamente come *Beyond the Standard Model* (BMS).

2.1.1 Le particelle elementari

Le particelle elementari sono i costituenti fondamentali della materia, ovvero quelli che si ritiene non abbiano una struttura interna.

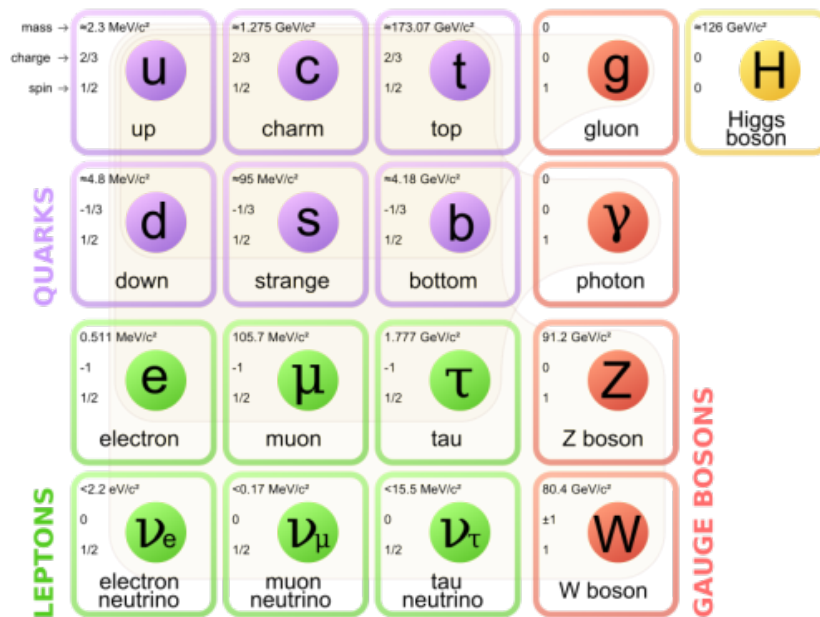


Figura 2.1: Le particelle del Modello Standard

Sono rappresentate in figura 2.1, per ognuna di esse esistono le rispettive antiparticelle, con uguale massa e spin, ma numeri quantici opposti. Si dividono in:

- fermioni: hanno spin semintero e di conseguenza seguono la statistica di Fermi-Dirac, tutta la materia ordinaria che osserviamo è costituita da fermioni;
- bosoni: hanno spin intero e di conseguenza seguono la statistica di Bose-Einstein, sono i mediatori delle interazioni fondamentali, più il bosone di Higgs;

I fermioni si dividono a loro volta in leptoni e quark.

I leptoni interagiscono attraverso l'interazione elettromagnetica, debole e gravitazionale. Si distinguono in leptoni carichi e a carica nulla. I leptoni carichi

sono l'elettrone (e), il muone (μ) e il tauone (τ), hanno carica $q = -1$ in unità naturali. A ognuna di queste particelle è associato un neutrino. Il leptone carico e il neutrino corrispondente costituiscono una famiglia leptonica. Si osserva che in ogni interazione la differenza tra numero di leptoni e numero di antileptoni è conservata. Pertanto si introduce la grandezza numero leptonico che vale $+1$ per i leptoni, -1 per gli antileptoni e 0 per le altre particelle, tale che la somma algebrica sia conservata in ogni processo. La conservazione del numero leptonico si applica separatamente per le famiglie. Solo il fenomeno dell'oscillazione dei neutrini, non previsto dal Modello Standard, permette ai neutrini di cambiare famiglia.

I quark interagiscono attraverso tutte e quattro le interazioni fondamentali. Sono le uniche particelle ad avere carica elettrica frazionaria, hanno una carica di colore che si scambiano nelle interazioni forti, e hanno 6 sapori: up (u) e down (d), charm (c) e strange (s), top (t) e bottom (b), a volte chiamati anche *truth* e *beauty*. Si legano tra di loro per formare particelle più pesanti chiamate barioni (formate da tre quark) e mesoni (formate da un quark e un antiquark), tutte queste si chiamano adroni. Analogamente al numero leptonico si introduce il numero barionico, una quantità conservata in ogni processo. Per un fenomeno chiamato confinamento del colore non è possibile osservare quark singoli. Il Modello Standard non esclude l'esistenza di stati chiamati tetraquark e pentaquark, ma non è stata ancora trovata evidenza sperimentale per questi. Data l'importanza del quark top nel processo studiato in questa tesi, verrà descritto in dettaglio successivamente.

Tutti i fermioni possono essere raggruppati in tre generazioni, ciascuna contenente 2 quark e 2 leptoni. I fermioni di generazioni superiori sono più massivi e meno stabili, per cui tendono a decadere in particelle della generazione precedente. Le ricerche di una quarta generazione di quark e leptoni non hanno prodotto risultati, e l'ipotesi di nuove famiglie che seguano la stessa struttura delle prime tre è fortemente sfavorita dalla scoperta dell'Higgs alla massa di 125.10 GeV.

Forza	Intensità Relativa	Raggio
Forte	1	10^{-15}
Elettromagnetica	10^{-2}	∞
Debole	10^{-13}	10^{-18}
Gravitazionale	10^{-38}	∞

Tabella 2.1: Intensità relativa delle interazioni fondamentali e il loro raggio, notare l'intensità estremamente bassa della forza gravitazionale.

2.1.2 L'interazione elettromagnetica

L'interazione elettromagnetica è descritta dalla teoria dell'elettrodinamica quantistica (QED). La particella mediatrice è il fotone (γ), un bosone di spin 1, neutro e di massa nulla, infatti il suo range è infinito e l'interazione elettromagnetica si manifesta anche a livello macroscopico. Richard Feynman, in particolare, sviluppò le regole per calcolare le ampiezze di transizione per i processi di scattering e decadimenti, e quindi le sezioni d'urto e le costanti di decadimento. Il metodo, basato sui diagrammi di Feynman, è stato poi utilizzato anche per le altre interazioni, ad oggi i diagrammi di Feynman sono uno strumento indispensabile in fisica delle particelle. L'interazione elettromagnetica non viola nessuna legge di conservazione.

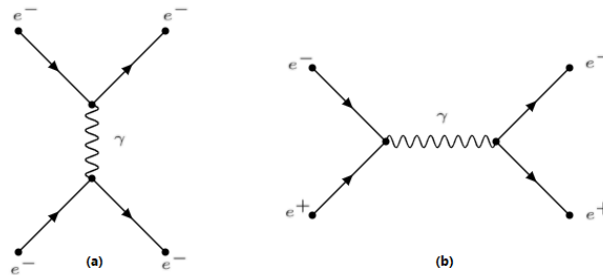


Figura 2.2: due semplici processi in QED (a): scattering e-e (b): annichilazione elettrone-positrone

2.1.3 L'interazione forte

L'interazione forte è descritta dalla teoria della cromodinamica quantistica (QCD). Introduce una nuova quantità chiamata carica di colore, sempre conservata. I quark possono avere colori *green* (G), *red* (R) e *blue* (B) o i rispettivi anticolori. Le particelle composite, nonché i leptoni, i bosoni mediatori delle altre forze e il bosone di Higgs sono complessivamente neutri. I quark si scambiano continuamente il colore tramite il gluone, un bosone di spin 1 e massa nulla. Poiché il colore si conserva nelle interazioni i gluoni devono essere bicolore. Questo permette, a differenza dell'elettromagnetismo, interazioni gluone-gluone. L'effetto di queste interazioni è che la forza ha un range di circa $10^{-15} m$, nonostante il range del gluone sia infinito. Inoltre spiega due fenomeni peculiari della QCD:

- Il confinamento del colore: separare 2 quark richiede un'energia infinita, non è possibile osservare 1 quark isolato. All'aumentare della distanza tra 2 quark cresce l'energia potenziale, finché il tubo di flusso di colore non si

spezza creando un'altra coppia di quark, che a loro volta porta a processi di rottura del tubo di flusso e produzione di altre coppie. Si creano dunque getti (jets) di adroni che conservano memoria dell'energia del quark da cui hanno avuto origine. Essi sono più collimati quanto più è alta l'energia del quark da cui hanno avuto origine.

- La libertà asintotica: l'interazione diventa sempre più debole al diminuire della distanza, questo implica che i quark si muovono all'interno degli adroni come se fossero liberi.

Anche l'interazione forte non viola alcuna legge di conservazione.

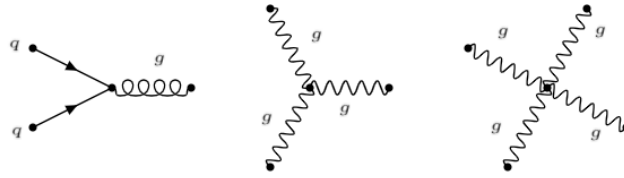


Figura 2.3: tipici vertici di interazione in QCD

2.1.4 l'interazione debole

L'interazione debole è così chiamata perché molto meno intensa delle precedenti. È mediata dai bosoni $W^\pm$, di spin 1, carica ± 1 e massa 80 GeV, e dal bosone Z^0 , di spin 1, carica 0 e massa 91 GeV. Data la massa elevata dei bosoni il range dell'interazione debole si aggira attorno ai 10^{-18} m. La debole intensità fa sì che i tempi di decadimento siano relativamente lunghi rispetto ai decadimenti elettromagnetici e forti, questo rende più facile identificare i processi che avvengono per interazione debole. La caratteristica più importante è che si tratta dell'unica interazione in grado di violare la simmetria C, la simmetria P e il loro prodotto CP, ed è l'unica che permette il cambiamento di sapore dei quark, anche tra quark di famiglie diverse. La probabilità di transizione attraverso i sapori delle varie famiglie è regolata dalla matrice Cabibbo-Kobayashi-Maskawa:

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

In una interazione debole, la probabilità che un quark q si trasformi in un quark q' è proporzionale a $|V_{qq'}|^2$. Gli elementi sulla diagonale principale sono quasi

uguali a uno, ciò significa che è più probabile per un quark trasformarsi in uno appartenente alla stessa generazione. L'interazione debole e l'interazione elettromagnetica, sebbene molto differenti, si unificano in un'unica interazione, detta elettrodebole, a energie dell'ordine delle centinaia di GeV. Questa è intimamente legata a un tassello fondamentale del Modello Standard: il campo di Higgs.

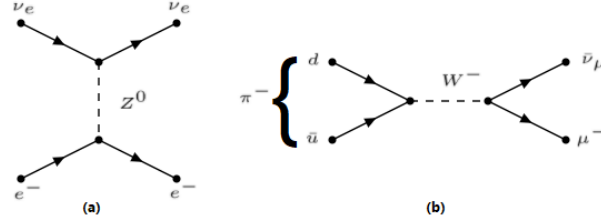


Figura 2.4: (a): scattering elettrone-neutrino, corrente neutra (b): decadimento del pione π^- , corrente carica

2.1.5 il Bosone di Higgs

Nella teoria elettrodebole i 4 bosoni $W^\pm$, Z^0 e γ risultano privi di massa. Introdurre nella lagrangiana i termini di massa “a mano”, eseguire cioè una rottura forzata della simmetria, porta la teoria ad essere non rinormalizzabile. Il problema è aggirabile introducendo una rottura spontanea di simmetria[9], ovvero supporre che lo stato a minore energia, lo stato di vuoto, sia degenere. La rottura spontanea di simmetria porta, tuttavia, a un'ulteriore difficoltà: si crea un bosone a massa nulla, non fisico, chiamato Bosone di Goldstone. La soluzione è data dal meccanismo di Higgs, che elimina il bosone di Goldstone e prevede l'esistenza di un campo scalare a cui è associato un bosone massivo, neutro e con spin 0, il bosone di Higgs appunto. Il meccanismo può essere esteso anche ai fermioni per spiegare la loro massa, anche se non è in grado di predirne il valore. Il bosone di Higgs è stato osservato per la prima volta nel 2012 dagli esperimenti ATLAS e CMS, che hanno misurato una massa di 125.10 GeV[18].

2.2 Il quark top

Il quark top è la particella elementare più massiva tra quelle del Modello Standard, con una massa di $173.0 \pm 0.4 \text{ GeV}$. Appartiene alla terza generazione e ha una carica di $+\frac{2}{3}$. Il quark top decade esclusivamente attraverso l'interazione debole con una vita media nell'ordine di 10^{-25} s , dunque non ha il tempo di adronizzare,

permettendo ai fisici di eseguire misure dirette della massa, dello spin e della carica.

2.2.1 Processi di produzione

Considerata l'elevata massa del quark top, occorrono energie molto elevate per produrlo. Attualmente, dopo lo spegnimento di Tevatron dove fu scoperto nel 1995, LHC è l'unica macchina in grado di raggiungere tali energie. I processi di produzione sono molteplici e dipendono dall'interazione coinvolta. Il principale processo di produzione, dovuto all'interazione forte, è quello in coppie di top quark-antiquark in seguito a collisioni tra gluoni molto energetici o coppie quark-antiquark. I processi di produzione attraverso interazione debole sono più rari, ma permettono la creazione di un quark top singolo, essi sono: s -channel, t -channel e produzione associata tW . I diagrammi di Feynman corrispondenti sono rappre-

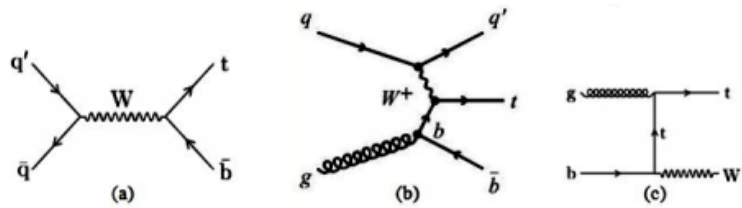


Figura 2.5: Diagrammi di Feynman caratteristici dei processi di produzione singola di top quark: (a) s -channel, (b) t -channel, (c) produzione associata tW

sentati in figura 2.5. L' s -channel consiste in un'annichilazione $q\bar{q}$ che, tramite un bosone W , produce un quark top e un antiquark bottom. Il t -channel consiste in un quark bottom che si trasforma in un quark top scambiando un bosone W con un altro quark, nelle collisioni p - p generalmente si tratta di un quark up o quark down. La produzione associata tW consiste nell'interazione di un gluone con un quark bottom che restituisce un quark top e un bosone W .

2.2.2 Processi di decadimento

Come già detto prima, il quark top decade solo per interazione debole, nella quasi totalità dei casi in un quark bottom e bosone W . Decadimenti del tipo $t \rightarrow Wd$ e $t \rightarrow Ws$ sono rarissimi. Ciò è dovuto al fatto che nella matrice CKM $|V_{tb}|^2 \gg$

$|V_{td}|^2, |V_{ts}|^2$. Il branching ratio di questo canale, definito come:

$$R = \frac{\lambda(t \rightarrow Wb)}{\lambda(t \rightarrow Wq)}$$

dove q indica la somma di λ su b, s e d . Tale rapporto è prossimo a 1.

2.3 Fisica oltre il Modello Standard

2.3.1 Il bosone W'

Come spiegato in precedenza, l'esistenza di fisica oltre il Modello Standard è necessaria per colmare le lacune dello stesso. Molti modelli alternativi predicono l'esistenza, entro il range energetico di LHC, di un nuovo bosone W' di spin 1 e carica ± 1 . Il bosone W' appare, per esempio, nel left-right symmetric model, nella teoria delle dimensioni extra di kaluza-klein, nei modelli little Higgs. Il modello top-flavour congettura un accoppiamento di W' con la terza generazione di fermioni, il che rende interessante la ricerca di un decadimento W' in quark top e bottom, più precisamente $W'^+ \rightarrow t\bar{b}$, $W'^- \rightarrow \bar{t}b$. Il modello ununified-SM predice un W' leptofobico, che decade esclusivamente in quark [17][4]. Sebbene la fenomenologia per i modelli Beyond Standard Model sia molto complicata, il particolare canale di ricerca in cui le risonanze decadono in quark t e b può costituire una prova per tutti quei modelli che predicono un accoppiamento ai quark.

2.3.2 Anomalie nel settore B

Non esistono al momento prove dirette dell'esistenza di W' . Recentemente, negli esperimenti BaBar, Belle e LHCb sono stati misurati decadimenti semi-leptonici di mesoni B in mesoni D e D^* e sono state trovate discrepanze significative con il Modello Standard [5]. In particolare nei branching ratio:

$$R(D) = \frac{\lambda(\bar{B} \rightarrow D\tau^-\bar{\nu}_\tau)}{\lambda(\bar{B} \rightarrow D\ell^-\bar{\nu}_\ell)}$$

$$R(D^*) = \frac{\lambda(\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau)}{\lambda(\bar{B} \rightarrow D^*\ell^-\bar{\nu}_\ell)}$$

dove $\ell = e, \mu$. Le misure hanno trovato un eccesso di decadimenti in disaccordo con i valori calcolati dal MS di circa 4σ . Il processo alla base è il decadimento via corrente carica $b \rightarrow c\tau\nu$, che è fortemente soppresso nel Modello Standard. Un bosone W' che si accoppia con la seconda e terza generazione di fermioni apporterebbe un nuovo contributo fisico al decadimento in grado di spiegare la discrepanza.

Capitolo 3

Ricostruzione degli oggetti fisici

3.1 Fenomeno in esame e campioni

3.1.1 Il decadimento di W'

Scopo di questa tesi è analizzare il decadimento di un bosone W' in un quark top e un quark bottom, in particolare considerando le catene di decadimento leptoniche del quark top[16]. Vengono considerati dunque i casi in cui un quark t decade debolmente in un quark b e in un bosone W , che a sua volta decade rapidamente in un leptone e il suo neutrino. In questa analisi abbiamo scelto come leptone il muone.

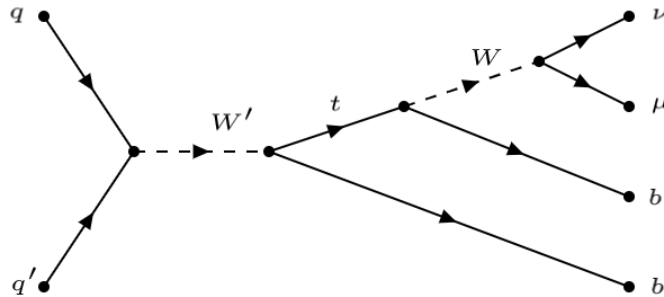


Figura 3.1: Diagramma di Feynman rappresentativo della produzione e decadimento considerato di W'

Lo stato finale dunque è dato da 2 quark b , un muone e un neutrino. Il processo è rappresentato in figura 3.1. Il neutrino non è rivelabile direttamente, il suo impulso viene ricavato tramite il cosiddetto momento trasverso mancante. Assumendo la chiusura cinematica dell'evento nel piano x - y , tale quantità corrisponde

alla somma vettoriale degli impulsi delle particelle nel piano x-y sfuggite alla rivelazione, che in media è nullo a meno della presenza di neutrini molto energetici. Naturalmente anche il quark b non è osservabile in maniera diretta, esso infatti adronizza in appena 10^{-24} s, dando vita a jet adronici.

3.1.2 I campioni di dati

Per il nostro studio utilizziamo 2 campioni simulati con un algoritmo MonteCarlo nelle condizioni di presa dati di LHC, in cui W' ha una massa di 4 TeV. I campioni contengono tutte le informazioni legate alla ricostruzione dei quark top, quindi le caratteristiche di jet e leptoni come vengono riprodotti all'interno del detector, in particolare i loro quadri-momenti e le caratteristiche degli algoritmi usati per individuarli all'interno di CMS. La ricostruzione di W' si effettua sommando il quadrimomento del jet b, proveniente da W' stesso, con il quadrimomento del quark top, quest'ultimo va a sua volta ricostruito dai suoi prodotti di decadimento ed è questo che pone maggiori difficoltà dal punto di vista sperimentale. Data la notevole massa del W' preso in considerazione, i quark b e t avranno un alto impulso, e quindi un alto boost di Lorentz nel sistema di riferimento di CMS, e i loro prodotti di decadimento saranno molto vicini tra loro. Per esempio, se una particella A di massa m_A e impulso trasverso p_{TA} decade nelle particelle B e C, $A \rightarrow BC$, la separazione angolare ΔR_{BC} tra B e C è data approssimativamente dalla seguente formula:

$$\Delta R_{BC} \approx \frac{2m_A}{p_{TA}}$$

Nella fattispecie del quark top, può accadere che il muone e il neutrino si trovino

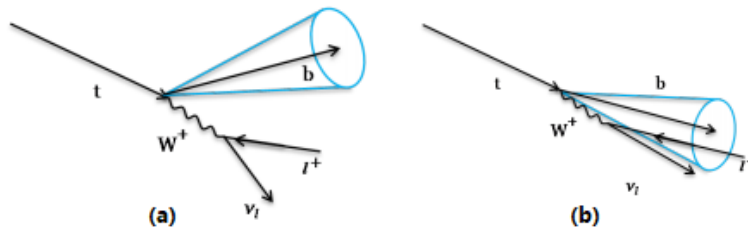


Figura 3.2: a: configurazione resolved b: configurazione merged

tanto vicini al jet b da essere contenuti in esso. Distinguiamo quindi il caso in cui i leptoni sono contenuti nel jet b, detto configurazione *merged*, dal caso in cui non lo sono, detto configurazione *resolved*. Per la precisione, questa discriminazione dipende dal raggio usato per la definizione del jet. Se la distanza angolare tra muone e jet nel piano $\eta - \phi$ è minore di 0.4, allora si parla di configurazione

merged, altrimenti resolved. Le due configurazioni sono rappresentate in figura 3.2.

Gli eventi sono distinti proprio dal tipo di jet che contengono al loro interno e raccolti in campioni differenti. Eseguiamo quindi l'analisi sia sul campione che contiene gli eventi merged che su quello contenente gli eventi resolved. Le informazioni contenute nel campione ci permettono di individuare con esattezza l'appartenenza di una particella ricostruita alla catena di decadimento di un quark top. Definiamo la variabile y detta *verità MonteCarlo* come un numero associato ai quark t uguale a 1 per i top veri, ossia quelli di segnale, e 0 per i quark falsi, che costituiscono il fondo.

3.2 Ricostruzione dei quark top standard

La ricostruzione standard del quark top (senza ML) è eseguita definendo una serie di variabili per i muoni e i jet, e poi effettuando dei tagli sulla base del processo fisico che si vuole osservare.

3.2.1 Muoni Tight

I muoni sono rivelati sia nel silicon tracker che nel sistema a muoni. Sono ricostruiti con una tecnica chiamata *filtro di Kalman*[12], e sono classificati diversamente a seconda delle tracce che lasciano. Se il muone è rivelato solo nel sistema a muoni prende il nome di *stand alone muon*. Un muone rivelato nel silicon tracker, e con almeno una traccia nel sistema a muoni si chiama *tracker muon*, l'algoritmo per identificarli è particolarmente utile per muoni a bassa energia. Se le tracce di un muone nel sistema a muoni sono compatibili con quelle nel silicon tracker si parla di *global muon*. Un global muon con ulteriori requisiti che ne accertano la qualità prende il nome di *tight muon*. Tra i requisiti ci sono, a esempio, una soglia superiore al χ^2 del fit della traccia e un parametro di impatto piccolo, ovvero la distanza minima dal punto di collisione. Sono i muoni tight maggiormente usati nell'analisi degli eventi. Un'altra variabile molto importante è l'*Isolamento*, questa restituisce una misura dell'energia depositata in un cono di raggio 0.4 nel piano $\eta - \phi$ attorno al muone da adroni carichi, neutri e fotoni; in questo modo è possibile capire se un muone è contenuto in un jet adronico. In questa analisi verrà utilizzato il *Mini-Isolamento*, dove il raggio del cono considerato è funzione dell'impulso del muone, tanto più piccolo quanto più grande è l'impulso[19]

3.2.2 Jet b

. I quark e i gluoni generati dalle collisioni p-p adronizzano prima di poter interagire direttamente con il rivelatore, ovvero creano sciame di adroni dalla forma conica chiamati jet. Esistono diversi algoritmi per ricostruire i jet, il più noto si chiama *anti - k_t* [7]. I jet con raggio $R = 0.4$ sono definiti AK4 (*narrow jet*), i jet con raggio $R = 0.8$ sono definiti AK8 (*fat jet*). CMS è anche in grado di distinguere il sapore del quark da cui ha avuto origine il jet, per il quark b si usa un algoritmo multivariato chiamato *Deep Combined Secondary Vertex* (JetDCSV)[14].

3.3 Machine Learning

3.3.1 Decision Tree

Ricorriamo all'uso del Machine Learning allo scopo di ottenere efficienze migliori del metodo standard. L'algoritmo utilizzato in questa analisi è *eXtreme Gradient Boosting* (XGBoost), si tratta di un algoritmo di Machine Learning supervisionato, spiegato in dettaglio nella sezione successiva. Preso un campione di dati $D = (x_i, y_i)$ con $(|D| = n, x \in \mathbb{R}^m, y \in \mathbb{R})$, dove n è il numero di dati e m il numero di variabili, il nostro obiettivo è predire y usando le informazioni racchiuse nel vettore x . Prima di iniziare l'allenamento il campione viene diviso in due parti (generalmente non uguali). Una parte verrà effettivamente utilizzata per l'allenamento, detto campione di *train*, l'altra parte verrà utilizzata per verificare la bontà dell'allenamento, detto campione di *test*. L'elemento cardine di XGBoost è il *decision tree*[6]. In generale, un decision tree fa una domanda e classifica il dato in base alla risposta. La domanda può riguardare una variabile discreta (e.g. "il muone è un tight muon?") oppure continua (e.g. "l'impulso del quark è maggiore di 50?"). Il punto di partenza del tree si chiama *radice*, gli step successivi si chiamano *nodi*, i punti di arrivo si chiamano *foglie*. L'aspetto di un decision tree generico è rappresentato in figura (3.3). L'applicazione di un requisito divide il campione in nodi, a ognuno dei quali corrisponde un certo numero di dati classificati come segnale o fondo. Naturalmente alcuni dati saranno classificati incorrettamente, un'informazione che è facilmente ricavabile dal confronto col campione originale. A ogni nodo quindi si associa la *purezza*, i tagli che portano a nodi con purezze più elevate sono i migliori. I decision tree sono in grado di classificare molto accuratamente i dati del campione di *train*, ma non sono altrettanto performanti per il campione di *test*. In altre parole, sono molto efficaci per classificare un campione specifico, ma mancano di generalità; si incorre cioè nel fenomeno dell'*overtraining*. Esistono diverse tecniche per superare questa limitazione dei decision tree che prendono il nome di Boosting.

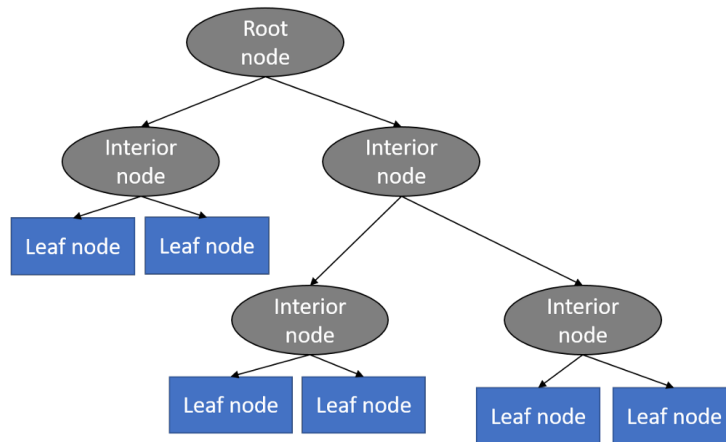


Figura 3.3: Rappresentazione schematica di un decision tree

3.3.2 XGBoost

XGBoost è un algoritmo di machine learning avanzato basato sui decision tree[8]. Può essere usato per classificare o per eseguire regressioni. È in grado di analizzare una grossa quantità di dati con un dispendio relativamente basso di risorse computazionali, inoltre permette di modificare molti parametri dell'allenamento. Queste qualità lo rendono un algoritmo estremamente versatile. Quando usato come classificatore XGBoost produce un output tra 0 e 1, generalmente i dati con output > 0.5 sono classificati come segnale e gli altri come fondo, ma è possibile cambiare questa soglia. D'ora in avanti chiameremo l'output *top-tagging*. XGBoost crea una sequenza di tree dalla lunghezza predeterminata, quindi calcola lo scarto tra il top-tagging e la verità, il tree successivo è costruito in modo da ridurre leggermente lo scarto del precedente. In dettaglio, l'algoritmo esegue i seguenti step:

1. Per prima cosa assegna un top tagging di 0.5 a tutti i dati¹, e ne calcola lo scarto rispetto alla classificazione corretta.
2. Crea un tree inizialmente formato da una sola foglia, contenente tutti gli scarti precedenti. Quindi assegna un punteggio s chiamato *similarity score*.
3. Applica i tagli e a ogni foglia creata assegna un similarity score. Poi assegna al taglio un punteggio chiamato *gain*, che è funzione di s . Più alto è il gain migliore è il taglio per la classificazione.

¹il valore di 0.5 per y predetta è di default sia per la classificazione che per la regressione. Può essere modificato manualmente, ma difatto non ha influenza sull'allenamento.

4. Per ogni taglio calcola la differenza tra il gain e il parametro di complessità γ (definito a priori). Se la differenza è negativa il taglio viene ripristinato, altrimenti viene mantenuto. Questa operazione è un primo argine all'overtraining.
5. Completato il tree, calcola l'output value w di ogni foglia, questi valori sono i nuovi scarti.
6. Il top-tagging è calcolato come il top-tagging iniziale più $w\epsilon$, dove ϵ è una costante compresa tra 0 e 1 chiamata *learning rate*. Il learning rate riduce la correzione del tree allo scarto per contrastare l'overtraining.
7. Il ciclo si ripete dal punto 2 fino a quando non si raggiunge il numero massimo di tree, oppure fin quando non si verifica la condizione di *early stopping*, ovvero una condizione che permette di interrompere il ciclo quando l'algoritmo non aumenta significativamente le performance, definita in seguito.

ogni ciclo prende il nome di epoca.

Le variabili che governano la costruzione del tree sono s , il similarity score, e w , l'output value. Questi sono ricavati da un principio di massima verosimiglianza. In particolare, l'espressione per w è quella che rende minima la quantità:

$$\sum_{i=1}^r L(y_i, p_i) + \gamma T + \frac{1}{2} \lambda w^2$$

dove y è il valore osservato (verità MC), p è il valore predetto (top-tagging), γ è il parametro di complessità, T è il numero di foglie nel tree, λ è il parametro di regolarizzazione, r è il numero di scarti in una foglia. Analizziamo più in dettaglio l'espressione.

La funzione $L(y_i, p_i)$ si chiama *loss function*. Quando XGBoost è usato per classificare $L(y_i, p_i)$ assume di solito la forma:

$$L(y_i, p_i) = -[y_i \log p_i + (1 - y_i) \log(1 - p_i)]$$

e prende il nome di *log-loss*. La loss function è una particolare metrica che ci dice quanto le nostre predizioni differiscono dalla verità. Notare che per piccoli w , $L(y_i, p_i)$ può essere riscritta come $L(y_i, p_i^0 + w)$, dove p_i^0 è la predizione al tree precedente. Il termine γT regola la *potatura*, l'operazione che ripristina un taglio effettuato, difatti accorciando un ramo del tree, allo scopo di evitare l'overtraining. Il termine $\frac{1}{2} \lambda w^2$ è il termine di regolarizzazione, che anche ha lo scopo di scongiurare l'overtraining. Più grande è λ più piccolo è il valore ottimale di w . Sfruttiamo anche la loss function per evitare l'overtraining introducendo una

condizione di early stopping: alla fine di ogni epoca calcoliamo la log-loss sul campione di test, se la log-loss non diminuisce significativamente dopo un certo numero di epoche l'allenamento viene interrotto. In questo studio abbiamo ri-

Tabella 3.1: Parametri dell'allenamento

	epoche	learning rate	early stopping
slow	260	0.001	35
medium	195	0.005	25
fast	130	0.01	15

petuto l'allenamento tre volte utilizzando velocità di apprendimento, numero di epoche e condizioni di early stopping differenti, i parametri sono esplicitati nella tabella 3.1.

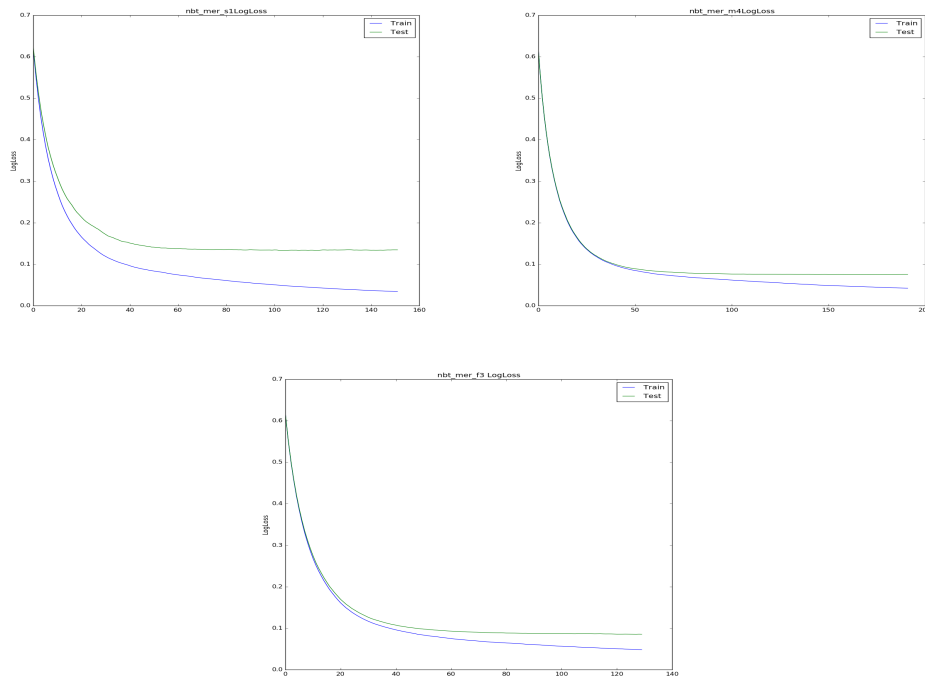


Figura 3.4: logloss per la configurazione nbt merged, slow (in alto a sinistra), medium (in alto a destra), fast (in basso)

In figura 3.4 sono rappresentate le log-loss per allenamenti a velocità diverse per la configurazione nbt merged. Notare che in tutti e tre i casi l'allenamento si interrompe prima di raggiungere il limite massimo di epoche, evidenziando l'utilità della condizione di early stopping per evitare l'overtraining.

3.3.3 Variabili dell'allenamento

Prima di iniziare l'allenamento si effettua una preselezione dei dati, per scartare quegli eventi che sono immediatamente riconducibili al fondo. Per esempio sono scartati gli eventi in cui il muone ha un parametro d'impatto $> |1|$ nel piano $x - y$ o ha Mini-Isolamento > 5 . Per ottimizzare l'algoritmo i campioni sono stati suddivisi in 4 fasce per impulso del quark top: la prima fascia ha eventi con $0 < p_t < 500$, la seconda ha $500 < p_t < 1000$, la terza ha $1000 < p_t < 2000$ e l'ultima ha $p_t > 2000$. Assumendo che W' sia creato a riposo, è ragionevole pensare che la sua energia si distribuisca equamente tra i prodotti di decadimento, il quark top e il quark bottom, e data la massa di 4 TeV, ci aspettiamo che l'impulso del quark t si aggiri attorno ai 2000 GeV. D'ora in avanti i grafici saranno riferiti alla fascia di impulso tra 1000 e 2000 (a meno che non sia specificato diversamente) perché è questa la più statisticamente significativa.

Come variabili che costituiscono il vettore x sono stati utilizzati 2 set leggermente diversi. Entrambi fanno uso di tutte le variabili usate per la ricostruzione standard, cioè:

- La condizione di global muon.
- Il parametro d'impatto, calcolato con 2 metodi differenti, in modo da ridurre al minimo il rischio di confondere il muone come proveniente da una collisione secondaria, da adrone di un jet, o da un raggio cosmico.
- Il Mini-Isolamento.
- Le variabili cinematiche del muone, componenti del quadrimomento: impulso trasverso, energia, pseudorapidità, ϕ .
- Le variabili cinematiche del jet.
- Le variabili cinematiche del top.

L'unica differenza tra le due configurazioni sta nell'uso della variabile JetDCSV: in un caso è usata per eseguire la preselezione e nell'altro è usata come variabile aggiuntiva nell'allenamento. Le due configurazioni sono indicate con bt e nbt, rispettivamente. La configurazione bt è stata usata solo sul campione di eventi resolved, mentre la configurazione nbt su entrambi i campioni merged e resolved. Questa scelta proviene naturalmente dal fatto che selezionare gli eventi del campione merged sulla base del b-tagging sarebbe del tutto superfluo, poiché i jet del campione merged sono necessariamente jet b.

3.3.4 Efficienze

Completato l'allenamento otteniamo un set di dati in cui a ogni evento è stato assegnato un valore tra 0 e 1. Di seguito sono mostrate le distribuzioni del top-tagging per diversi allenamenti, in rosso il segnale e in blu il fondo. Per la precisione in

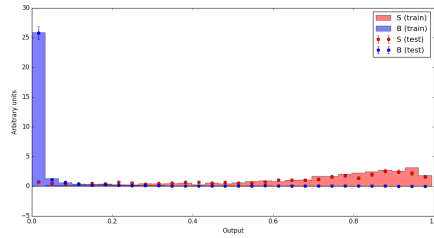


Figura 3.5: BDT per la configurazione bt resolved

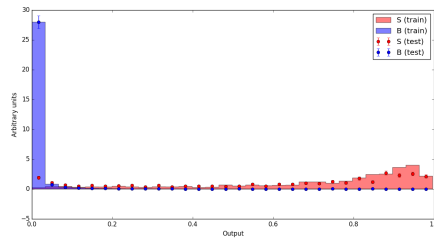


Figura 3.6: BDT per la configurazione nbt resolved

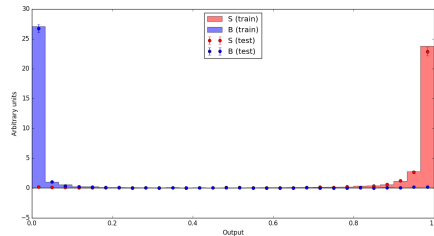


Figura 3.7: BDT per la configurazione nbt merged

figura 3.5 la distribuzione per l'allenamento sul campione bt resolved, in figura 3.6 la distribuzione per il campione nbt resolved, e in figura 3.7 la distribuzione per il campione nbt merged. Come è evidente dalla figura l'algoritmo di machine learning si dimostra particolarmente efficace nel classificare gli eventi nella configurazione merged, ovvero la configurazione in cui il muone proveniente dal quark top è contenuto nel jet b del quark stesso. Per la configurazione resolved, invece, diventa più difficile classificare il segnale.

Definiamo efficienza di segnale il rapporto tra il numero di quark top veri al di sopra del taglio e il numero di quark top totali:

$$\epsilon_{signal} = \frac{n_{above}}{n_{tot}}$$

e analogamente l'efficienza di background come il rapporto tra in numero di quark top falsi al di sopra del taglio e il numero di quark top totali

$$\epsilon_{background} = \frac{n_{below}}{n_{tot}}$$

l'obiettivo è avere l'efficienza di segnale alta e l'efficienza di background bassa.

Nelle tabelle 3.2, 3.3 e 3.4 sono indicate le efficienze per gli allenamenti con taglio al top-tagging > 0.5 e per il taglio standard. Si evince che in tutte le configurazioni e per tutte le fasce di impulso, l'efficienza di segnale di XGBoost è uguale o di poco superiore all'efficienza di segnale dell'analisi standard, mentre l'efficienza di background è considerevolmente più bassa per l'analisi con machine learning.

top p_t	std signal	std background	ML signal	ML background
0-500	74.516	3.811	74.93	1.835
500-1000	72.551	4.187	72.409	1.128
1000-2000	71.856	4.19	71.979	1.346
>2000	71.856	4.173	72.000	1.232

Tabella 3.2: efficienze per la configurazione nbt resolved

top p_t	std signal	std background	ML signal	ML background
0-500	74.516	3.811	74.926	2.612
500-1000	72.551	4.187	73.036	3.168
1000-2000	71.856	4.19	71.835	2.983
>2000	71.856	4.173	71.739	2.517

Tabella 3.3: efficienze per la configurazione bt resolved

top p_t	std signal	std background	ML signal	ML background
0-500	60.112	2.912	61.333	0.776
500-1000	69.076	8.752	69.549	0.38
1000-2000	68.706	2.978	71.195	0.712
>2000	69.072	6.148	69.363	1.258

Tabella 3.4: efficienze per la configurazione nbt merged

Volendo modificare la soglia del top-tagging, uno strumento utile per facilitare la scelta è la *receiver operating curve* (ROC). In una ROC sull'asse delle x è

graficata l'efficienza di background e sull'asse delle y l'efficienza di segnale. Immaginiamo di tracciare una linea verticale vicina all'asse delle y, questa indica un taglio molto vicino a 1, l'intersezione della retta con la roc individua le efficienze.

Se la roc fosse una retta con coefficiente angolare 1, qualunque taglio sarebbe inutile, perché le efficienze di segnale e fondo sono sempre uguali. Se la roc fosse un gradino, qualunque taglio sarebbe perfetto.

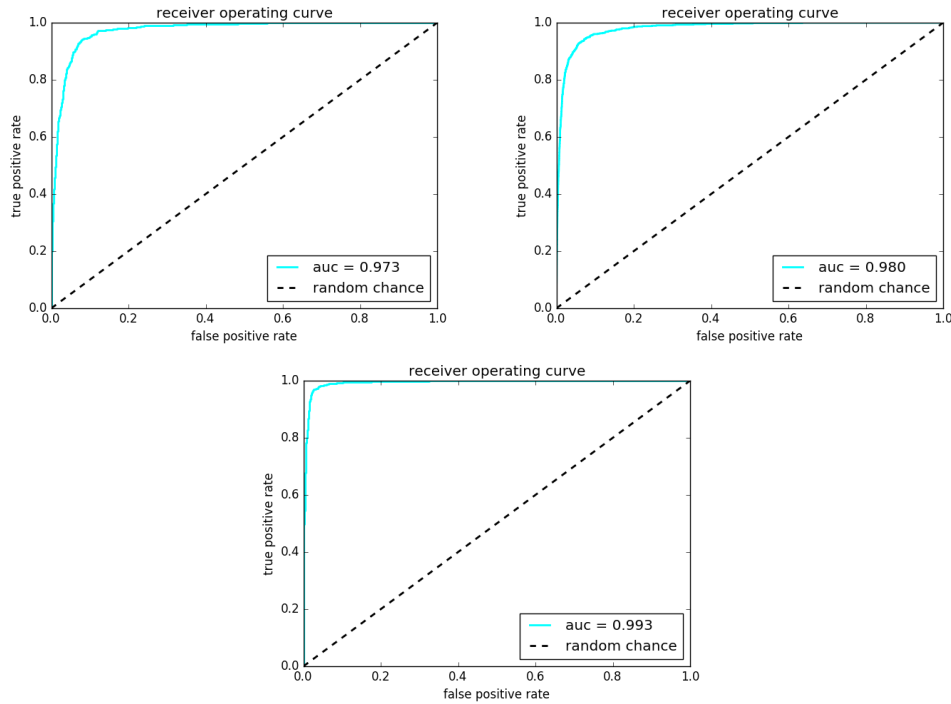


Figura 3.8: ROC per la configurazione bt resolved (in alto a sinistra), nbt resolved (in alto a destra), nbt merged (in basso)

Nella figura 3.8, sono raffigurate le ROC per le configurazioni bt resolved, nbt resolved e nbt merged. In seguito, per la ricostruzione di W' , abbiamo deciso di classificare come quark top di segnale quelli con top-tagging > 0.7 .

3.4 Ricostruzione W'

3.4.1 Ricostruzione della massa invariante

Con le informazioni di XGBoost è possibile ricostruire il bosone W' . Una volta individuato il quark top con top-tagging più elevato in un evento, ricostruiamo la massa invariante di W' sommando il quadrimpulso del quark top a quello del jet b più energetico, a meno che questo non sia il jet b proveniente dal decadimento del quark top stesso; ci aspettiamo che questo avvenga in un numero esiguo di casi, infatti quando W' decade, la sua energia si divide quasi equamente tra il quark top e il quark bottom, l'energia del quark top poi si divide ulteriormente tra il jet b, il muone e il neutrino; è naturale che l'energia del primo jet sia maggiore di quella del secondo.

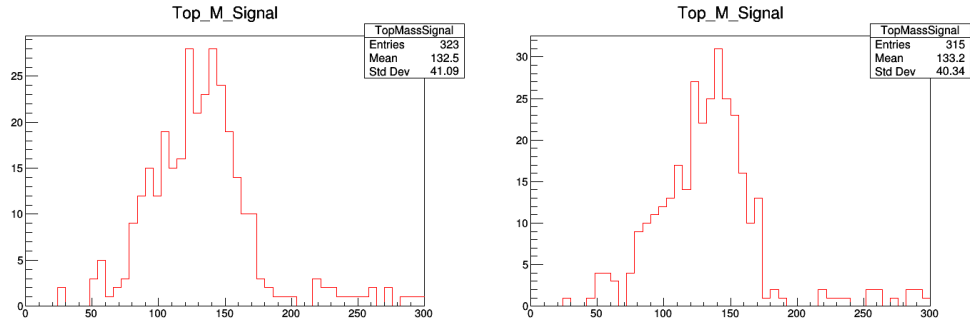


Figura 3.9: Distribuzione della massa dei quark top di segnale per la configurazione bt resolved (a sinistra) e nbt resolved (a destra).

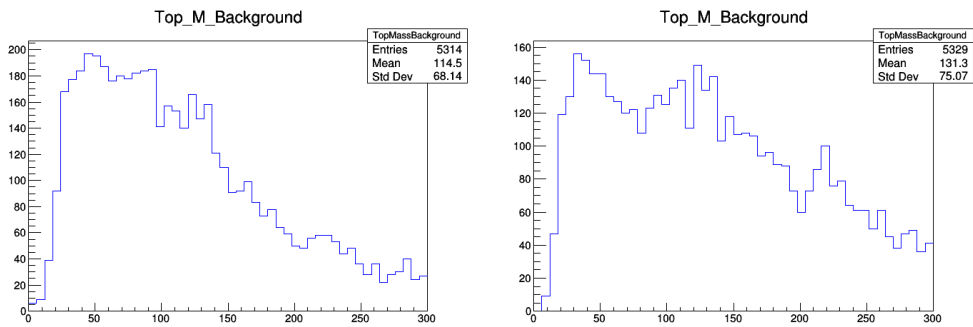


Figura 3.10: Distribuzione della massa dei quark top di fondo per la configurazione bt resolved (a sinistra) e nbt resolved (a destra).

Per le configurazioni bt resolved e nbt resolved l'algoritmo si dimostra efficace nel distinguere i quark top di segnale (figura 3.9) da i quark top di fondo

(figura 3.10). Tuttavia gli eventi contenenti W' nelle configurazioni resolved sono esegui, poichè, considerato l'alto impulso dei quark top, raramente i suoi prodotti di decadimento sono risolti. Pertanto non si è in grado di ricostruire la massa di W' a partire da questi campioni.

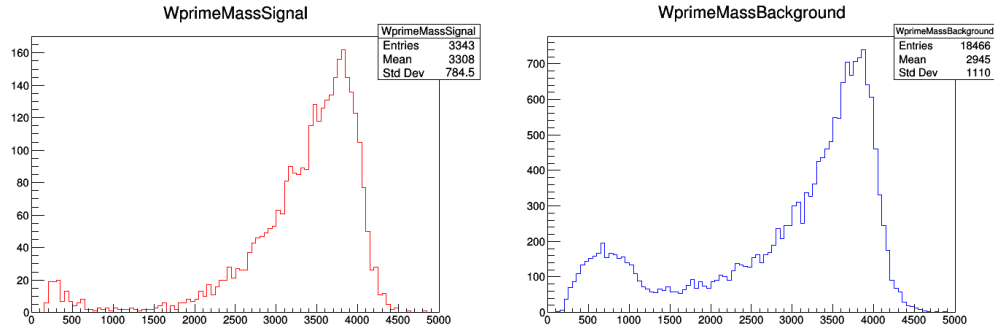


Figura 3.11: Massa di W' di segnale (a sinistra) e di fondo (a destra) nella configurazione nbt merged

In figura 3.11 è rappresentata la distribuzione per la massa di W' nella configurazione nbt merged. Per il segnale è ben riconoscibile un picco attorno ai 4 TeV. La larghezza del picco sulla destra è dovuta a gli effetti di risoluzione del rivelatore, mentre la lunga coda a sinistra è dovuta al fatto che fallisce la ricostruzione della componente z del neutrino tramite l'energia mancante. Si nota anche un basso picco in prossimità dello zero, che è dovuto alla porzione di quark top classificata incorrettamente. Anche per il fondo è presente un picco a 4 TeV, Questo è dovuto in parte al fatto che l'algoritmo è programmato per riconoscere solo i muoni, e pertanto classifica i quark top provenienti da W' , ma che decadono in elettroni e tauoni, come fondo; e in parte ai quark top veri classificati incorrettamente.

3.4.2 Confronto con l'analisi standard

La ricostruzione con il metodo standard si esegue richiedendo che il muone proveniente dal decadimento del quark top sia un muone tight con Mini-Isolamento < 0.1 e che il jet b proveniente dal quark top abbia JetDCSV = 1. Anche in questo caso la ricostruzione fallisce per le configurazioni resolved, mentre otteniamo un picco distinto a 4 TeV nella configurazione merged (figura 3.12). Possiamo apprezzare il vantaggio dell'uso del machine learning confrontando questa distribuzione con quella di XGBoost.

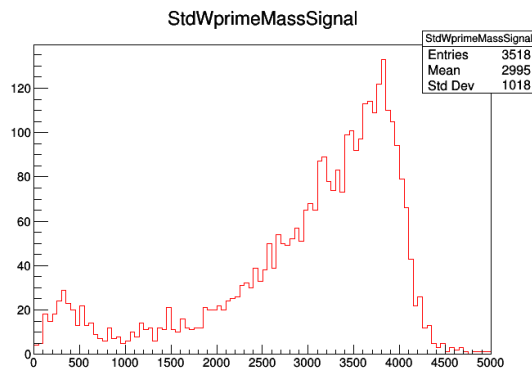


Figura 3.12: Massa di W' ricostruita con il metodo standard per la configurazione nbt merged

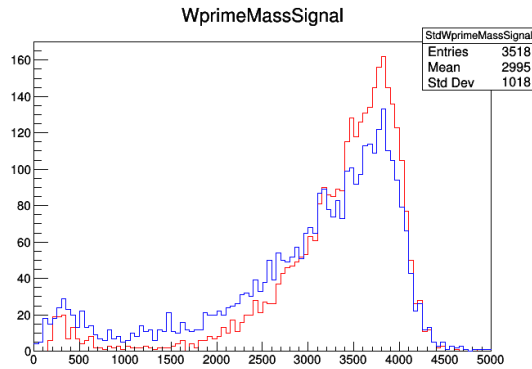


Figura 3.13: Confronto tra la distribuzione ottenuta con XGBoost (rosso) e la distribuzione ottenuta col metodo standard (blu) per la configurazione nbt merged.

Come già evidenziato dalla tabella 3.4, la figura 3.13 mostra che l'algoritmo di machine learning presenta un picco più pronunciato, ideale per le analisi di fisica, ed è più efficace nella reiezione del fondo.

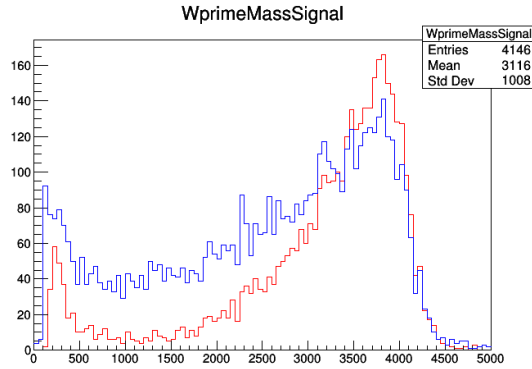


Figura 3.14: Confronto tra la distribuzione ottenuta con XGBoost (rosso) e la distribuzione ottenuta col metodo standard (blu) per la configurazione nbt merged, senza restrizioni sull'impulso del quark top.

Questa differenza è ancora più lampante se non ci limitiamo a osservare solo i quark top con $1000 < p_t < 2000$, ma utilizziamo quark t dell'intero campione. La figura 3.14 mostra le distribuzioni di massa di W' senza restrizioni sull'impulso del quark top.

Conclusioni

In questo lavoro di tesi è stata presentata la ricerca di un nuovo bosone massivo, carico chiamato W' sfruttando un algoritmo di machine learning supervisionato denominato XGBoost. Lo stato finale considerato è caratterizzato da un jet b altamente energetico e da i prodotti di decadimento semileptonico di un quark top.

Il Modello Standard è attualmente la teoria che meglio descrive le particelle fondamentali e le interazioni tra di esse. Nel corso degli ultimi decenni ha ricevuto numerose conferme sperimentali tra cui ricordiamo la scoperta del bosone di Higgs nel 2012 da parte degli esperimenti ATLAS e CMS al Large Hadron Collider. Tuttavia il Modello Standard presenta anche numerose mancanze, la più importante delle quali è l'assenza della descrizione della forza gravitazionale, per superare questo e molti altri problemi sono stati sviluppati numerosi modelli "Beyond Standard Model". Alcuni di questi modelli prevedono l'esistenza di un nuovo bosone W' in grado di accoppiarsi ai quark. L'esistenza del bosone W' con accoppiamenti preferenziali alla terza generazione potrebbe spiegare recenti osservazioni di un eccesso di decadimenti semileptonici per il mesone B , in disaccordo con i valori previsti dal MS di almeno 4σ da parte degli esperimenti BaBar, Belle e LHCb.

Il bosone W' è ricercato in numerosi canali sia adronici che leptonici ad LHC, al fine di ricostruirlo con efficienze superiori a quelle ottenute con tecniche di analisi standard, si è fatto ricorso all'algoritmo di machine learning XGBoost. Sono stati considerati decadimenti del bosone $W' \rightarrow tb$, dove il quark top decade in $t \rightarrow \mu\nu b$. Sono state considerate 2 configurazioni per lo stato finale di muone, neutrino e jet b chiamate merged e resolved. Su queste sono stati eseguiti tre allenamenti aventi diversi parametri. In tutti e tre i casi la condizione di early stopping si è rivelata necessaria per evitare l'overtraining. Gli allenamenti sono stati ripetuti dividendo i campioni in quattro parti a seconda dell'impulso del quark top, per ottimizzare la classificazione. Sono state mostrate le distribuzioni della variabile risultante "top-tagging" per le configurazioni resolved e merged, per quest'ultima in particolare l'algoritmo svolge un eccellente lavoro nel classificare correttamente i quark top di segnale da quelli di fondo. Si è mostrato che XG-

Boost è in grado di raggiungere efficienze di segnale più alte e allo stesso tempo ridurre considerevolmente le efficienze di fondo rispetto all'analisi standard.

Utilizzando i quark top classificati da XGBoost è stata ricostruita la massa invariante del bosone W' . Per la configurazione merged risulta un picco in prossimità dei 4 TeV. Il confronto con la massa ricostruita in modo standard evidenzia i miglioramenti ottenuti dall'algoritmo di machine learning nella selezione del fondo.

Bibliografia

- [1] *The CMS electromagnetic calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.
- [2] *The CMS hadron calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.
- [3] *The CMS tracker: addendum to the Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 2000.
- [4] M. Aaboud, et al. (The ATLAS Collaboration) *Search for $W' \rightarrow tb$ decays in the hadronic final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.* <https://doi.org/10.1016/j.physletb.2018.03.036>.
- [5] M. Abdullah, J. Calle, B. Dutta, A. Florez, and D. Restrepo. *Probing a simplified, W' model of $R(D^{(*)})$ using b -tags, τ leptons and missing energy.* <https://arxiv.org/pdf/1805.01869v2>.
- [6] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees.* Wadsworth, 1984.
- [7] M. Cacciari, G. P. Salam, and G. Soyez. *The anti-kt jet clustering algorithm.* JHEP, 04:063, 2008.
- [8] T. Chen, C. Guestrin. *XGBoost: A Scalable Tree Boosting System.* <https://arxiv.org/pdf/1603.02754>.
- [9] F. Englert, and R. Brout. *Broken Symmetry and The Mass of Gauge Vector Mesons.* Physical Review Letters, Volume 13, Numero 9, 1964.
- [10] L. Evans, P. Bryant. *LHC Machine.* <https://doi.org/10.1088/1748-0221/3/08/S08001>.
- [11] J. G. Layter. *The CMS muon project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

- [12] R. Kalman. *A new approach to linear filtering and prediction problems*. Transaction of the ASME-Journal of Basic Engineering, 82 (series D): 35-45, 1960.
- [13] P. Langacker. *Introduction to the Standard Model and Electroweak Physics*. <https://arxiv.org/pdf/0901.0241>.
- [14] A. M. Sirunyan et al. (The CMS collaboration). *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*. 2018, JINST 13 P05011.
- [15] A. M. Sirunyan et al. (The CMS collaboration). *The CMS experiment at the CERN LHC*. <https://doi.org/10.1088/1748-0221/3/08/S08004>.
- [16] A. M. Sirunyan et al. (The CMS collaboration). *Search for heavy resonances decaying to a top quark and a bottom quark in the lepton+jets final state in proton-proton collisions at 13 TeV*. doi:10.1016/j.physletb.2017.12.006.
- [17] A. M. Sirunyan et al. (The CMS collaboration). *Searches for W' bosons decaying to a top quark and a bottom quark in the lepton+jets final state in proton-proton collisions at 13 TeV*. [https://doi.org/10.1007/JHEP08\(2017\)029](https://doi.org/10.1007/JHEP08(2017)029).
- [18] M. Tanabashi et al. (Particle Data Group). *Review of Particle Physics*. Phys. Rev. D 98, 030001 (2018).
- [19] https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMuonAnalysis#Muon_identification.