

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Misura della massa del quark top usando
eventi $t\bar{t}$ selezionati da CMS a 13 TeV nel
regime *all-jet boosted*

Relatore:
Prof. Andrea Castro

Presentata da:
Noemi Cavalli

Anno Accademico 2016/2017

Sommario

Il presente elaborato riporta una misura della massa del quark *top* effettuata impiegando dati raccolti tramite il rivelatore CMS a LHC durante il Run del 2016, con energia del centro di massa per le collisioni *pp* pari a 13 TeV e luminosità integrata di 37.0 fb^{-1} .

Sono stati studiati i prodotti di decadimento di coppie $t\bar{t}$ secondo il canale totalmente adronico (*all jets*) in regime *boosted*. Tale canale risulta essere caratterizzato dalla presenza di due jet ampi, associati sia al decadimento dei bosoni W che al quark *b*. Per migliorare la purezza del campione gli eventi candidati sono stati selezionati richiedendo la presenza di quark *b* in ciascuno di essi.

Un'ulteriore richiesta di selezione è stata imposta tramite un'analisi multivariata per aumentare la discriminazione degli eventi di fondo.

La misura della massa del quark *top* è stata ottenuta effettuando un fit di massima verosimiglianza e considerando possibili cause di incertezza sistematica. Il valore finale risulta essere:

$$m_t = 171.9 \pm 0.3 \text{ (stat)} \pm 1.7 \text{ (sist)} \text{ GeV.}$$

La misura risulta pertanto consistente con la media mondiale di $173.1 \pm 0.6 \text{ GeV}$.

Indice

1	Il quark top nel Modello Standard	3
1.1	Il Modello Standard	3
1.1.1	Particelle fondamentali: fermioni	3
1.1.2	Particelle fondamentali: bosoni	6
1.1.3	Interazioni fondamentali	6
1.1.4	Limiti del Modello Standard	8
1.2	Il quark top	8
1.2.1	Scoperta del quark top	9
1.2.2	Produzione e decadimento di coppie $t\bar{t}$	9
1.2.3	Boosted jets	11
2	Il rivelatore CMS	13
2.1	CERN	13
2.2	LHC	14
2.3	CMS	15
2.3.1	Tracciatori	15
2.3.2	Calorimetri	17
2.3.3	Camere a muoni	18
3	Analisi dati e misura della massa del quark top	19
3.1	Campioni di dati	19
3.2	Selezione dati	20
3.2.1	Preselezione e studio dell'efficienza del <i>trigger</i>	21
3.2.2	N_{Jets} e $N_{B_{Jets}}$	23
3.3	Il fondo multijet	24
3.3.1	Studio della variabile m_{va}	24
3.3.2	variabile category	26
3.3.3	Modello per il fondo	27
3.4	Misura della massa	28
3.4.1	Template	28
3.4.2	Fit con funzione di massima verosimiglianza	29

3.4.3	Valore di massa	33
3.4.4	Errore statistico	35
3.4.5	Errore sistematico	36
3.4.6	Risultato finale	38
	Conclusioni	39
	Bibliografia	40

Capitolo 1

Il quark top nel Modello Standard

1.1 Il Modello Standard

I fenomeni fisici legati alle particelle richiedono la trattazione mediante una teoria che consenta di conciliare meccanica quantistica e relatività ristretta, in quanto la dinamica dei processi indagati risulta essere relativistica e su scale dimensionali inferiori al femtometro.

Il Modello Standard consente la descrizione di tali fenomeni utilizzando teorie di campo quantizzato (*Quantum Field Theories*, QFT), che prevedono l'esistenza di campi di materia e di campi associati a tre delle quattro interazioni fondamentali (elettromagnetica, debole e forte). Si hanno dunque due categorie di particelle: fermioni (particelle di materia) e bosoni (quanti scambiati tra particelle nelle interazioni).

Inoltre, attraverso la trattazione di Dirac, la teoria prevede l'esistenza di antiparticelle, che hanno le medesime proprietà cinematiche delle particelle cui sono associate, ma numeri quantici interni opposti, come ad esempio la carica elettrica e il numero barionico o leptonico.

Nello studio dei fenomeni subnucleari risultano rilevanti, oltre alle proprietà cinematiche e dinamiche quali quantità di moto, energia e massa, nuove caratteristiche delle particelle, identificate attraverso numeri quantici associati. Tra queste vi sono il sapore, che determina la differenza nominale tra le particelle fondamentali, il colore, collegato all'interazione forte, lo spin, momento angolare intrinseco, e l'isospin, che consente di categorizzare determinate particelle come diversi stati quantici di una stessa particella.

1.1.1 Particelle fondamentali: fermioni

I fermioni sono le particelle fondamentali (e dunque puntiformi) costituenti la materia; hanno tutte spin $1/2$, pertanto soddisfano il principio di esclusione di Pauli, e sono suddivise in due categorie:

- quark:

sono raggruppati in tre "famiglie" (riportate in Figura 1.1) ed esistono unicamente in stati legati quali barioni, costituiti da tre quark o tre antiquark, e mesoni, formati da coppie quark-antiquark. Essi prendono parte a processi deboli, forti ed elettromagnetici, pertanto possiedono carica elettrica, colore e sapore. I quark *up*, *charm*, *top* (*u*, *c*, *t*) hanno carica elettrica positiva pari a $+2/3$ la carica fondamentale ($e = 1.602 \times 10^{-19}$ C) mentre *down*, *strange*, *beauty* (*d*, *s*, *b*) possiedono carica negativa pari a $-1/3$ la carica fondamentale.

Le principali proprietà dei quark sono riportate in Tabella 1.1 [1]; si noti in particolare che i valori di massa sono riferiti alle masse costituenti, cioè sono i valori con cui le particelle prendono parte ai processi di formazione degli adroni (barioni e mesoni). I valori "nudi" non risultano misurabili in quanto i quark non esistono allo stato libero.



Figura 1.1: Quark del Modello Standard.

Quark	Carica elettrica	Spin	Massa
<i>u</i>	$+2/3$	$1/2$	$2.2^{+0.6}_{-0.4}$ MeV
<i>d</i>	$-1/3$	$1/2$	$4.7^{+0.5}_{-0.4}$ MeV
<i>c</i>	$+2/3$	$1/2$	$1.28^{+0.03}_{-0.03}$ GeV
<i>s</i>	$-1/3$	$1/2$	96^{+8}_{-4} MeV
<i>t</i>	$+2/3$	$1/2$	$173.1^{+0.6}_{-0.6}$ GeV
<i>b</i>	$-1/3$	$1/2$	$4.18^{+0.04}_{-0.03}$ GeV

Tabella 1.1: Valori caratteristici dei quark.

- leptoni:

come i quark, sono raggruppati in tre "famiglie" (riportate in Figura 1.2), tuttavia, a differenza dei primi, esistono unicamente come particelle libere. Elettrone, muone e leptone *tau* (e , μ , τ) hanno la stessa carica elettrica negativa (pari alla carica fondamentale), pertanto interagiscono elettromagneticamente. Non possiedono carica di colore e dunque non prendono parte ai processi forti; mentre sono coinvolti in quelli deboli.

I neutrini elettronici, muonici e tauonici (ν_e, ν_μ, ν_τ) prendono unicamente parte a processi deboli, poiché risultano privi di carica elettrica e di colore. A causa di ciò risultano difficili da rivelare sperimentalmente e dunque sono ancora oggi oggetto di studio e di ricerca.

In Tabella 1.2 sono riportati i valori delle principali proprietà dei leptoni [1].

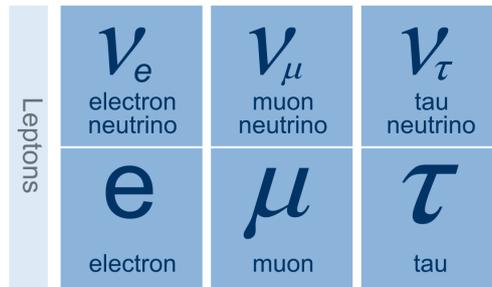


Figura 1.2: Leptoni del Modello Standard.

Leptone	Carica elettrica	Spin	Massa
ν_e	0	1/2	$< 2.05 \text{ eV}$
e	-1	1/2	$0.5109989461 \pm 0.0000000031 \text{ MeV}$
ν_μ	0	1/2	$< 0.19 \text{ MeV}$
μ	-1	1/2	$105.6583745 \pm 0.0000024 \text{ MeV}$
ν_τ	0	1/2	$< 18.2 \text{ MeV}$
τ	-1	1/2	$1776.86 \pm 0.12 \text{ MeV}$

Tabella 1.2: Valori caratteristici dei leptoni.

Poiché ad ogni fermione fondamentale risulta associato un antifermione con cariche interne opposte, e poiché ogni quark esiste in tre differenti varianti determinate dal colore,

le particelle fermioniche fondamentali previste dal Modello Standard risultano essere 48 in totale.

1.1.2 Particelle fondamentali: bosoni

I bosoni (riportati in Tabella 1.3 [1]) sono le particelle fondamentali di spin intero che vengono scambiate da quark, o più generalmente da adroni, e da leptoni durante i processi di interazione.

Ogni tipologia di interazione risulta mediata da un determinato bosone associato:

- il fotone γ :
mediatore per l'interazione elettromagnetica, osservato in processi quali effetto fotoelettrico, produzione di coppia elettrone-positrone ed effetto Compton. È elettricamente neutro pertanto non interagisce elettromagneticamente con se stesso.
- i bosoni W^+ , W^- , Z^0 :
vengono scambiati durante i processi deboli e sono rispettivamente dotati di carica elettrica positiva, negativa e nulla. Interagiscono quindi elettromagneticamente e debolmente e danno luogo a fenomeni di autointerazione.
Il fatto che abbiano massa non nulla appare in contrasto con le teorie di gauge di campo quantizzato, che prevedono come mediatori per le interazioni bosoni non massivi. Tale contrasto risulta superato attraverso il meccanismo di Higgs che conferisce massa ai mediatori deboli.
- i gluoni g :
mediatori per l'interazione forte che sono elettricamente neutri e privi di carica debole, quindi non prendono parte a fenomeni elettromagnetici o deboli. Sono invece dotati di carica di colore e dunque danno luogo a processi forti di autointerazione. Il Modello Standard prevede l'esistenza di otto distinte tipologie di gluoni.

Una menzione a parte spetta al bosone di Higgs H^0 , di spin nullo, rilevato sperimentalmente dalle collaborazioni ATLAS e CMS il quattro luglio 2012 [2][3]. Attraverso il meccanismo di Brout-Engler-Higgs, H^0 conferisce massa ai mediatori delle interazioni, in particolare ai bosoni W^+ , W^- , Z^0 , che secondo le teorie di gauge dovrebbero avere massa nulla, come risulta invece per il fotone e i gluoni.

1.1.3 Interazioni fondamentali

Le teorie di campo quantizzato per le interazioni forte, debole ed elettromagnetica prevedono l'invarianza della funzione Lagrangiana associata al sistema rispetto a trasformazioni di gauge locali.

Il meccanismo che determina l'interazione tra particelle comporta lo scambio di quanti

Bosone	Interazione	Carica elettrica	Spin	Massa
γ	Elettromagnetica	$< 10^{-35}$	1	$< 10^{-18}$ eV
W^+, W^-	Debole	± 1	1	80.385 ± 0.015 GeV
Z^0	Debole	0	1	91.1876 ± 0.0021 GeV
g	Forte	0	1	0
<i>gravitone</i>	Gravitazionale	0	2	$< 6 \times 10^{-32}$ eV
H^0	Meccanismo di Brout-Engler-Higgs	0	0	125.09 ± 0.24 GeV

Tabella 1.3: Bosoni del Modello Standard con valori caratteristici associati.

(bosoni di gauge dell'interazione), che vengono ceduti da una prima particella e assorbiti da una seconda.

La QED (*Quantum Electrodynamics*) coinvolge tutte le particelle dotate di carica elettrica e prevede l'esistenza di un unico campo (e quindi un unico mediatore) associato. Il fatto che il bosone mediatore (γ) abbia massa nulla comporta il lungo raggio dell'interazione elettromagnetica.

Procedendo per analogia con la QED è stata formulata la teoria per l'interazione forte, nominata QCD (*Quantum Chromodynamics*) [4]. Sebbene i gluoni abbiano massa nulla l'interazione forte risulta essere a corto raggio; ciò è conseguente al fenomeno del confinamento, che vincola i quark negli stati legati. Il potenziale forte risulta infatti espresso dalla seguente formula:

$$V_S = -\frac{4}{3} \frac{\alpha_S}{r} + Kr$$

dove α_S è la costante di accoppiamento forte, adimensionale e associata all'intensità di interazione, mentre la costante K risulta essere dell'ordine di 1 GeV/fm [5]. A grandi distanze prevale dunque il termine elastico del potenziale, che determina il confinamento dei quark all'interno delle strutture adroniche.

La QCD prevede l'esistenza di otto campi (e dunque otto gluoni) associati a tre distinte cariche, dette cariche di colore R,G,B (*red, green, blue*). Poiché i leptoni risultano privi di colore, non costituiscono stati legati, come invece accade per i quark.

La teoria di campo quantizzato per l'interazione debole, formulata a seguito dello studio dei decadimenti radioattivi dei nuclei atomici [4], prevede l'esistenza di tre distinti campi di interazione, mediati dai bosoni massivi W^+, W^-, Z^0 , pertanto risulta essere a corto raggio.

A differenza delle interazioni forte ed elettromagnetica, i processi deboli non conservano

il sapore delle particelle e coinvolgono anche i neutrini (che risultano dotati di "carica debole" senza avere colore o carica elettrica).

Le interazioni debole ed elettromagnetica sono state unificate tramite la teoria elettrodebole, sviluppata negli anni '60 da Weinberg, Salam e Glashow. Tale teoria prevede la conservazione dell'iper carica, definita attraverso la relazione di Gell-Mann-Nishijima:

$$Y_W = 2(Q - I_3) \quad (1.1)$$

dove Q indica la carica elettrica (in unità di carica elementare) e I_3 la terza componente di isospin.

1.1.4 Limiti del Modello Standard

Il Modello Standard, nonostante risulti efficace per la descrizione della molteplicità dei fenomeni fisici legati alla fisica delle particelle, non costituisce una teoria completa. Ad oggi, infatti, è stata formulata la teoria elettrodebole, che unifica interazione debole ed elettromagnetica, ed è stata costruita Teoria della Grande Unificazione (GUT) per unire alle due interazioni precedenti l'interazione forte. Tuttavia, la GUT non risulta verificabile direttamente utilizzando acceleratori di particelle [4], perché la scala di unificazione va molto oltre le energie accessibili.

Il Modello Standard non fornisce spiegazione al vasto range di valori assunti dalle masse delle particelle fondamentali, che va dai neutrini ($m < 2$ eV per ν_e) al quark top ($m \approx 173$ GeV). Anche il fatto che oggi la materia sia maggiormente presente nell'Universo rispetto all'antimateria risulta essere un'evidenza inspiegata; vi sono inoltre la materia oscura e l'energia oscura che, pur costituendo il 96% dell'Universo, non trovano riscontro nelle teorie formulate.

Sono ad oggi attivi numerosi esperimenti per colmare le mancanze del Modello Standard e fornire risposte alle questioni irrisolte.

1.2 Il quark top

Il quark *top* risulta essere la particella fondamentale più massiva presente nel Modello Standard e pertanto ha tempo di vita molto breve ($\approx 0.5 \times 10^{-24}$ s [7]). A causa di queste caratteristiche, non costituisce stati legati e decade per interazione debole e non forte.

Esso possiede carica elettrica pari a $Q = +2/3 e$, e numeri quantici di isospin $I = 1/2$, $I_3 = +1/2$, pertanto costituisce assieme al quark *beauty* un doppietto di isospin (terza generazione dei quark).

Lo studio del quark *top* risulta rilevante anche perché fornisce informazioni riguardo il meccanismo di Higgs; in particolare attraverso la misura della massa di t è possibile estrapolare il valore del potenziale di Higgs su grande scala [6].

1.2.1 Scoperta del quark top

A seguito dell'evidenza della violazione di simmetria CP da parte dei kaoni (mesoni composti dai quark e antiquark u, d, s), venne formulato da Kobayashi e Maskawa il miscelamento di sapore, che forniva spiegazione alla violazione di simmetria e che prevedeva l'esistenza di almeno tre generazioni di quark affiancate a tre generazioni di leptoni. Pertanto oltre ai quark u, d, s, c , allora già conosciuti, dovevano esistere due ulteriori particelle fondamentali, denominate *top* e *beauty*. La prima di queste ad essere rivelata sperimentalmente fu il quark *beauty*, scoperto attraverso la risonanza Υ al FNAL (*Fermi National Accelerator Laboratory*) nel 1977. Le proprietà del b vennero determinate dalla collaborazione PLUTO a DESY (*Deutsches Elektronen-Synchrotron*), studiando la distribuzione angolare di adroni costituiti da b , prodotti tramite collisioni elettrone-positrone e^+e^- . Il quark *beauty* risultò possedere carica elettrica $Q = -1/3$ e numero quantico associato alla terza componente di isospin $I_3 = -1/2$. Dunque b risultava fare parte di un doppietto di isospin, e ciò rafforzò la convinzione dell'esistenza di un ultimo quark con carica elettrica $Q = +2/3$ e numero di isospin $I_3 = +1/2$.

Inizialmente si riteneva che tale quark, il *top*, dovesse avere massa simile al *beauty*, poichè facenti parte del medesimo doppietto di isospin, pertanto ci si aspettava di rivelare processi $e^+e^- \rightarrow t\bar{t}$ ad energie dell'ordine di 10 GeV.

Tuttavia i risultati ottenuti dagli esperimenti condotti a PEP (SLAC, *Stanford Linear Accelerator Center*), PETRA (DESY), TRISTAN (*High Energy Accelerator Research Organization*), SLC (SLAC) e LEP 1 (CERN), rispettivamente con energie massime associate al centro di massa pari a 30, 45, 64, 91 e nuovamente 91 GeV, mostrarono la completa assenza di segnale associato al decadimento ricercato; pertanto il *top* doveva essere dotato di massa significativamente maggiore rispetto al *beauty*.

Ulteriori conferme dell'elevata massa di tale quark vennero fornite tramite gli esperimenti UA1, UA2 (CERN) e ARGUS (DESY); infine misure di precisione di parametri elettrodeboli al LEP (CERN) permisero di stimarne la massa come $m_t = 173_{-13}^{+12}$ GeV.

La scoperta del *top* avvenne nel 1995 ad opera delle collaborazioni CDF e D0 a Tevatron, l'acceleratore di fasci $p\bar{p}$ presente al FNAL. La rivelazione sperimentale venne effettuata impiegando energie di centro di massa pari a 1.8 TeV [8] e studiando i decadimenti $t \rightarrow Wb$, ammessi cinematicamente a tali energie.

A Tevatron vennero misurate: la sezione d'urto associata alla produzione di eventi $t\bar{t}$, la massa del *top* e altre proprietà, quali ad esempio la polarizzazione del bosone W nel decadimento $t \rightarrow Wb$.

1.2.2 Produzione e decadimento di coppie $t\bar{t}$

Il quark *top* può essere prodotto tramite interazione debole (produzione singola), oppure attraverso l'interazione forte (produzione della coppia $t\bar{t}$).

La produzione singola, osservata per la prima volta a Tevatron nel 2009[7], può avvenire

attraverso tre distinti canali chiamati:

- *t-channel*: $qb \rightarrow q't$;
- *s-channel*: $q\bar{q}' \rightarrow t\bar{t}$;
- *tW-channel*: $gb \rightarrow tW^-$ [9];

Questi processi risultano rilevanti nello studio della fisica oltre il Modello Standard in quanto, se il bosone di Higgs avesse carica elettrica non nulla, influenzerebbe differemente i canali *s* e *t*.

A LHC il processo dominante risulta essere la produzione della coppia $t\bar{t}$, che avviene mediante uno o più gluoni, come è mostrato nei diagrammi di Feynman riportati in Figura 1.3. La produzione da annichilazione di gluoni risulta essere più probabile ($\approx 90\%$) rispetto alla produzione da annichilazione di quark-antiquark ad energie del centro di massa pari a 13 TeV[8].

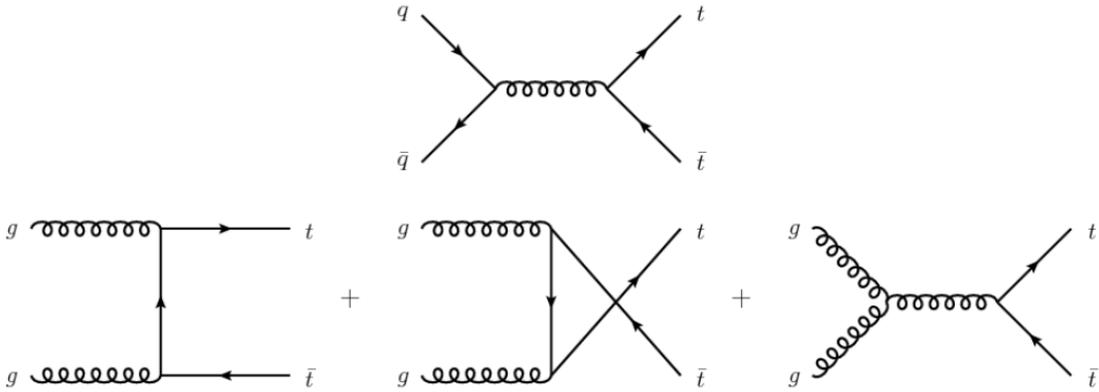


Figura 1.3: Diagrammi di Feynman relativi alla produzione di coppie $t\bar{t}$. Diagramma di annichilazione (sopra) e di fusione gluonica (sotto).

Il quark *top* decade in Wb con *Branching Ratio*:

$$BR(t \rightarrow Wb) = \Gamma(t \rightarrow Wb) / \sum_{q=d,s,b} \Gamma(t \rightarrow Wq) = 0.998 \quad (1.2)$$

Il quark *bottom* prodotto genera jet adronici, mentre il bosone W può decadere in $q\bar{q}'$, che a loro volta generano jet adronici, oppure in leptoni $\ell\nu$. Complessivamente possono avvenire i seguenti decadimenti della coppia $t\bar{t}$:

- canale totalmente adronico (canale *all-jet*): $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bq''\bar{q}'''\bar{b}$ con $BR = 4/9 \approx 45\%$;
- canale semi-leptonico: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell^+\nu_\ell b\bar{q}\bar{q}'\bar{b}$
oppure $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'b\ell^-\bar{\nu}_\ell\bar{b}$ con $BR = 8/27 \approx 30\%$;
- canale leptonico: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell^+\nu_\ell b\ell^-\bar{\nu}_\ell\bar{b}$ con $BR = 4/81 \approx 5\%$.

Mentre nella rivelazione delle prime due tipologie di decadimento il fondo risulta riducibile richiedendo la presenza di leptoni isolati (prodotti dal decadimento della W), nel regime completamente adronico ciò non è possibile, pertanto il fondo di eventi multijet QCD risulta essere significativamente intenso. In questo caso la distinzione tra fondo e rumore avviene richiedendo come selezione sugli eventi la presenza di due jet *b-tagged*, comprendenti mesoni composti dal quark *beauty*.

Attraverso la rivelazione dei mesoni B e dei bosoni W si risale al decadimento iniziale, ed è possibile misurare indirettamente la massa invariante le quark *top* che ha dato luogo al decadimento.

Nel presente elaborato si considera unicamente quest'ultima tipologia di decadimento di coppie $t\bar{t}$; in particolare si farà riferimento al regime *Boosted jets*.

1.2.3 Boosted jets

I quark *top* prodotti tramite collisioni ad alte energie, come avviene a LHC, possono acquisire valori di quantità di moto molto elevati. Ciò comporta che i prodotti di decadimento dei quark *top* con $p_t \geq 200$ GeV risultino spazialmente collimati, fino a costituire dei jet di particelle, come mostrato in Figura 1.4. Pertanto, al fine di ricostruire i singoli prodotti di decadimento della coppia $t\bar{t}$, è necessario l'impiego di specifici algoritmi. Tali algoritmi operano inizialmente la ricostruzione utilizzando un raggio ampio (ricostruzione dei *fat jets*), successivamente ne esaminano la struttura, ricercando al loro interno i singoli jet connessi al decadimento ricercato e distinguendo questi dai jet prodotti da altri eventi QCD.

La massa del quark *top* viene ricostruita tramite la tecnica nominata *Soft Drop de-clustering*[10], che rimuove dai jet in analisi la radiazione poco intensa e con ampia distribuzione angolare, proveniente da eventi di fondo.

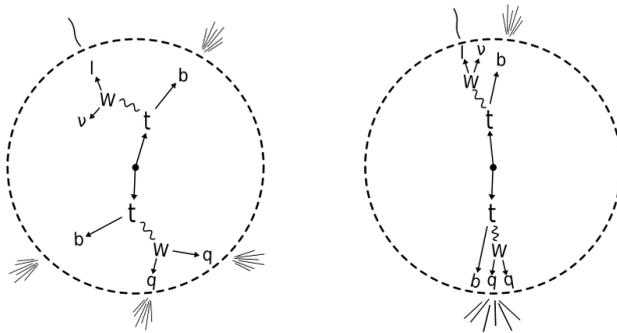


Figura 1.4: Rappresentazione del decadimento semi-leptonico di una coppia $t\bar{t}$ in regime standard (a sinistra) e in regime *boosted* (a destra).

Capitolo 2

Il rivelatore CMS

2.1 CERN

Il Centro Europeo per la Ricerca Nucleare (CERN), fondato nel 1952 con lo scopo di costituire un'organizzazione di ricerca europea, a oggi conta 22 paesi membri e continua nello studio dei fenomeni fisici coinvolgenti le particelle fondamentali costituenti l'Universo.

Il meccanismo alla base degli esperimenti presenti al CERN prevede dei fasci di particelle accelerate a elevate energie che vengono fatti collidere generando nuove particelle. L'impianto sperimentale risulta pertanto composto da acceleratori, che consentono alle particelle costituenti dei fasci di raggiungere velocità prossime a quella della luce, e da rivelatori, che permettono di determinare le proprietà fisiche delle particelle prodotte dalle collisioni e di studiare i fenomeni subnucleari che hanno luogo durante l'interazione tra fasci.

Principalmente vengono impiegati fasci costituiti da protoni, generati applicando un campo elettrico a un volume di gas idrogeno, separando i nuclei dagli elettroni. Tali protoni vengono accelerati fino a 50 MeV tramite Linac 2 (primo acceleratore) e immessi sequenzialmente in PSB (*Proton Synchrotron Booster*), PS (*Proton Synchrotron*), SPS (*Super Proton Synchrotron*) dove raggiungono progressivamente energie pari a 1.4 GeV, 25 GeV, 450 GeV, rispettivamente. Infine i fasci entrano in LHC (*Large Hadron Collider*) dove raggiungono 6.5 TeV e vengono fatti collidere in corrispondenza dei rivelatori ATLAS, CMS, ALICE e LHC-b [11].

Una rappresentazione schematica del complesso sperimentale è riportata in Figura 2.1.

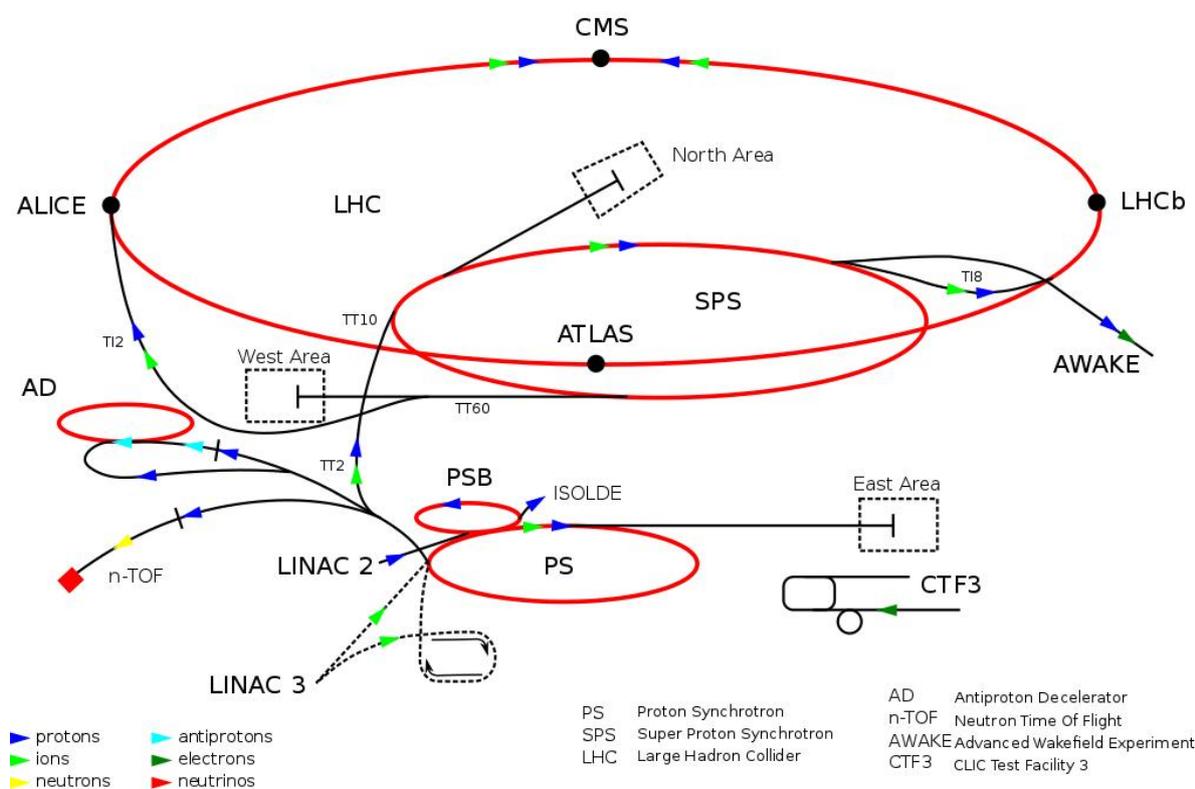


Figura 2.1: Complesso di acceleratori e rilevatori al CERN.

2.2 LHC

LHC è il più grande acceleratore di particelle esistente al mondo. Esso è costituito da un anello di circonferenza pari a 27 km posto a 100 m di profondità in prossimità del confine tra Svizzera e Francia.

Attraverso l'intenso campo magnetico (≈ 8 T) generato dai magneti superconduttori che lo compongono, due fasci di particelle paralleli e opposti (generalmente protoni, ma anche ioni pesanti quali ad esempio il piombo) vengono deviati all'interno dell'anello in due canali distinti (*beam pipes*) in cui è stato generato il vuoto, e accelerati fino al raggiungimento di energie di circa 7 TeV per fascio. Successivamente i fasci sono fatti collidere in corrispondenza delle posizioni degli apparati costituenti i quattro esperimenti: CMS, LHCb, ATLAS e ALICE. Vi sono diverse tipologie di magneti impiegate al fine di garantire la collisione frontale dei due fasci di particelle: dipoli magnetici per deviare i fasci, quadrupoli magnetici per concentrarli in modo tale da massimizzare la probabilità di collisione, e sestupoli/ottupoli/decapoli aggiunti per correggere piccole imperfezioni

del campo magnetico.

Per potere operare efficacemente occorre che i magneti superconduttori siano mantenuti a temperature prossime allo zero assoluto, pertanto l'acceleratore è connesso ad un sistema di distribuzione di elio liquido, che porta i magneti al raggiungimento di temperature pari a -271.3 °C. Tale sistema criogenico è il più grande esistente al mondo [11].

Tra i quattro esperimenti attivi a LHC, ATLAS e CMS hanno come scopo l'indagare la grande vastità di fenomeni fisici collegati alle dimensioni extra, alla materia oscura e, in generale, al Modello Standard. Sebbene abbiano scopi comuni, i due esperimenti impiegano tecniche e apparati differenti.

LHCb è finalizzato allo studio delle differenze tra materia e antimateria in fenomeni fisici coinvolgenti il quark *beauty* e alla violazione della simmetria CP.

Infine ALICE ha come scopo lo studio della fisica di interazione tra materia a forte densità energetica, più precisamente allo stato QGP (*Quark Gluon Plasma*) della materia.

2.3 CMS

Con i suoi 15 metri di altezza e 21 di lunghezza, CMS (*Compact Muon Solenoid*) [12] costituisce un rivelatore di avanguardia avente come scopo l'osservare ogni possibile fenomeno fisico generato dalla collisione tra particelle a elevatissima energia.

Per potere rivelare eventi simmetricamente da ogni direzione, CMS è costruito mediante componenti cilindriche sovrapposte, centrate attorno ai tubi a vuoto di LHC in cui scorrono i fasci di protoni (Figura 2.2).

Il solenoide, da cui prende il nome, è costituito da materiale superconduttore mantenuto a -268.5 °C in grado di generare un campo magnetico di 4 T necessario a deviare le particelle prodotte dalla collisione dei due fasci in modo tale da tracciarne il raggio di curvatura e quindi risalire alla quantità di moto. Misure precise sono possibili grazie all'impiego di tracciatori ad elevata precisione, calorimetri e rivelatori muonici.

2.3.1 Tracciatori

I tracciatori in silicio costituiscono la parte del rivelatore più vicina ai tubi a vuoto di LHC dove interagiscono i due fasci di protoni. Sono costruiti in modo tale da essere danneggiati il meno possibile dall'intensa quantità di radiazione cui sono sottoposti e allo stesso tempo rivelare con alta precisione il passaggio delle particelle generate dalle collisioni, senza alterarne significativamente la dinamica. Vi sono due tipologie di tracciatori in silicio:

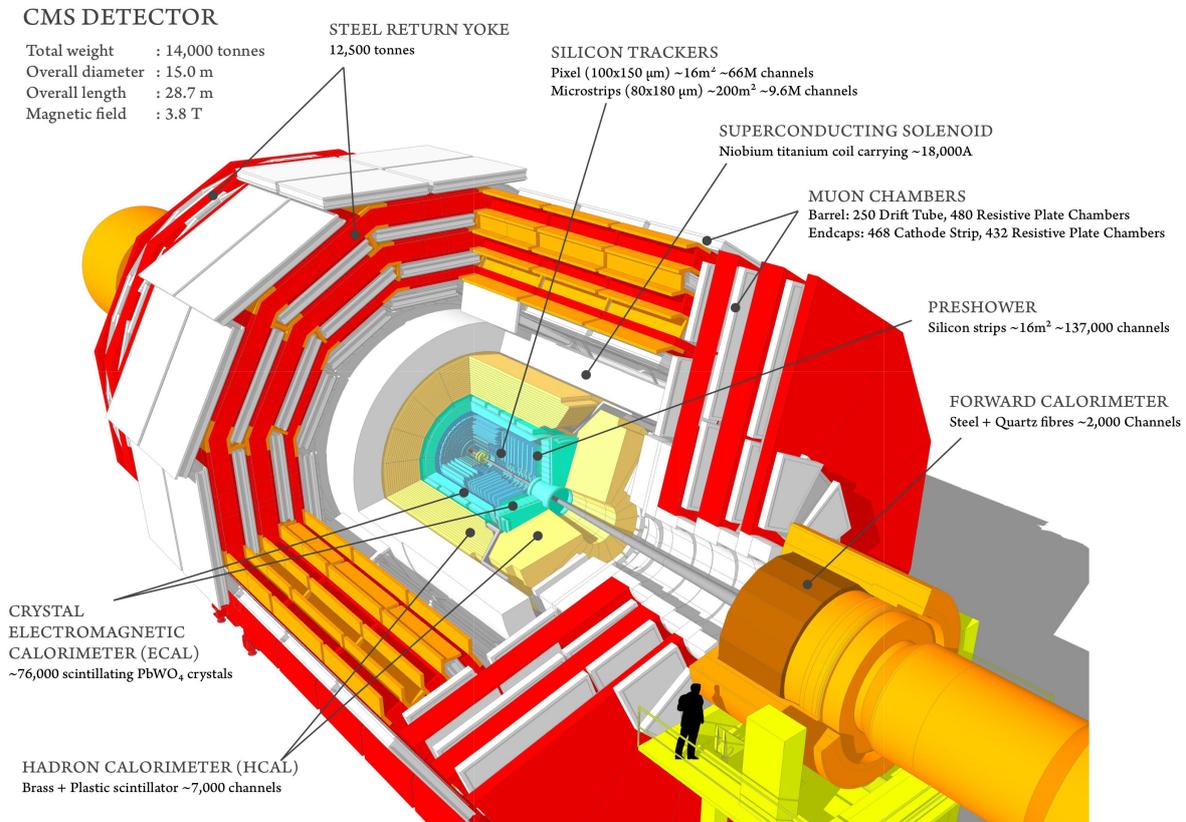


Figura 2.2: Schema del rivelatore CMS.

- Rivelatore a pixel (*Silicon pixel detector*):
 è l'impianto tracciante più vicino ai canali dei fasci. Costituito da strati cilindrici concentrici distanti dai *beam pipes* 4 cm, 7 cm, 11 cm e da due dischi posti alle estremità, riceve segnali da circa 10 milioni di particelle per centimetro quadro per secondo (a 8 cm). Risulta estremamente rilevante nel tracciamento delle particelle con vita media ridotta, che decadono durante l'attraversamento dei primi strati del rivelatore.
 Ogni strato è suddiviso in segmenti 100 μm per 150 μm , ciascuno corrispondente a un sensore in grado di rivelare il passaggio di una particella carica elettricamente. Il segnale elettrico, generato dai sensori quando avviene l'interazione, viene amplificato e successivamente registrato.
 Conoscendo le posizioni dei sensori che hanno rivelato il passaggio di una particella se ne può ricostruire la traiettoria.
- Rivelatore a strisce di silicio (*Silicon strip detector*):
 si estende in dieci strati fino ad una distanza di 130 cm dai canali dei fasci. Contiene

15200 moduli, ciascuno dei quali è costituito da un set di sensori a risposta rapida in grado di ricevere molteplici segnali in intervalli di tempo ridotti; tali segnali sono in seguito amplificati, processati tramite supporti elettronici abbinati ai sensori ed infine trasmessi mediante cavi in fibra ottica.

2.3.2 Calorimetri

Per potere effettuare misure di energia delle particelle, CMS impiega due distinti calorimetri, il primo associato alla rivelazione di elettroni e fotoni, il secondo alla rivelazione di adroni.

- **ECAL (*Electromagnetic Calorimeter*):**
situato all'esterno dei tracciatori, è costituito da un cilindro contenente 36 "supermoduli", ciascuno dei quali contiene 1700 cristalli a elevata densità in grado di rivelare il passaggio di elettroni e fotoni. Quando le particelle entrano in contatto con il cristallo questo genera scintillazioni di intensità proporzionale all'energia della particella incidente.
I segnali luminosi prodotti dai cristalli vengono ricevuti e amplificati da fotomoltiplicatori posti sul retro dei cristalli, costruiti in modo tale da operare in presenza di un intenso campo magnetico quale quello generato dal solenoide.
Alle estremità del cilindro sono posti altri 15000 cristalli insieme a rivelatori in grado di distinguere un fotone con elevata energia da coppie di fotoni vicini a basse energie (*Preshower detectors*).
- **HCAL (*Hadron Calorimeter*):**
consente la misurazione dell'energia di adroni e la rilevazione indiretta di particelle neutre (non interagenti con i rivelatori).
È costituito da un cilindro interno formato da strati "assorbitori" e da scintillatori distribuiti in modo tale da rivelare le particelle generate sia dalla collisione tra fasci sia dai decadimenti dei prodotti di collisione.
L'interazione fra un adrone e strati "assorbitori", in ottone o acciaio, genera nuove particelle prodotte dai decadimenti. Tali particelle secondarie possono a loro volta decadere interagendo con strati successivi del calorimetro. Si vengono dunque a creare "cascate adroniche" che interagiscono con gli scintillatori. I segnali ottenuti vengono in seguito amplificati da fotosensori (*Hybrid Photodiodes*).
Per potere arrestare tutti gli adroni emessi, HCAL è costituito da un ulteriore cilindro costituito da strati di ottone e acciaio. Componenti analoghe sono poste esternamente anche in corrispondenza delle direzioni dei fasci incidenti, sempre per prevenire la fuoriuscita di particelle a alta energia dal rivelatore.

2.3.3 Camere a muoni

Le camere a muoni, poste all'esterno del solenoide, costituiscono lo strato più esterno di CMS e hanno come finalità la rivelazione di leptoni (muoni) le cui traiettorie non risultano arrestarsi nei calorimetri.

Queste fungono inoltre da *trigger*, ovvero permettono di determinare se i segnali ottenuti siano associati al passaggio di un muone o meno. Vi sono in totale 1400 camere a muoni divise in tre tipologie:

- DTs (*Drift Tubes*):
ciascuna delle 250 camere DTs ha dimensioni pari a 2 m x 2.5 m ed è costituita da 12 strati di alluminio suddivisi in tre gruppi tra i quali uno ha come scopo la determinazione delle coordinate della traiettoria del muone lungo la direzione del fascio, mentre gli altri due lungo le direzioni perpendicolari a esso.
Ogni strato è costituito da tubi contenenti ciascuno un cavo elettrico immerso in un gas. Il passaggio di un muone nei tubi genera il rilascio di elettroni da parte del gas, i quali vanno a collidere con il cavo generando impulsi elettrici.
- CSCs (*Cathode Strip Chambers*):
sono composte da intrecci di cavi carichi positivamente (anodi) e strisce di rame cariche negativamente e immerse in un volume di gas (catodi).
Il passaggio di un muone genera il rilascio di elettroni da parte del gas. Si vengono così a creare due correnti opposte: la corrente di elettroni che fluisce verso l'anodo e la corrente di ioni positivi verso il catodo. I due distinti impulsi elettrici consentono la determinazione della posizione associata al passaggio del muone.
- RPCs (*Resistive Plate Chambers*):
sono costituite da due piani paralleli caricati elettricamente uno positivamente e uno negativamente. Nello spazio tra i piani è posto del gas che rilascia elettroni quando viene attraversato da particelle cariche. Tali elettroni oltrepassano i piani carichi e interagiscono con strisce metalliche poste all'esterno, consentendo la determinazione della quantità di moto della particella carica rivelata.
Una volta misurato la quantità di moto, il *trigger* determina se il segnale ricevuto sia associato al passaggio di un muone oppure a un'altra particella carica. Nel primo caso il segnale viene trasmesso, nel secondo caso scartato.

Capitolo 3

Analisi dati e misura della massa del quark top

3.1 Campioni di dati

L'analisi sperimentale è stata eseguita operando su campioni contenenti dati ottenuti attraverso simulazioni Monte Carlo (MC) e su dati reali ottenuti dal rivelatore CMS al CERN durante il Run del 2016 e immagazzinati in file ROOT sotto forma di *trees* contenenti i valori delle variabili ricostruite dal rivelatore. In questo elaborato sono state utilizzate due tipologie di campioni ottenuti tramite simulazioni MC, riferite a eventi $t\bar{t}$ (file TT) e generici eventi *multijets* prodotti da interazione forte (file QCD). Questi ultimi risultano divisi in sei distinti sottocampioni, ciascuno riferito a uno specifico valore della sezione d'urto e a un range della variabile H_T , avente come valore la somma scalare degli impulsi trasversi dei jet.

I campioni simulati di eventi $t\bar{t}$ utilizzati sono dodici, di cui sei riferiti al valore di massa del quark top pari a $m_t = 172.5$ GeV, simulati in condizioni diverse, e i rimanenti sei riferiti rispettivamente a masse pari a 166.5, 169.5, 171.5, 173.5, 175.5, 178.5 GeV. Tali campioni risultano ottenuti tramite il programma POWHEG [13] affiancato a PYTHIA [14], responsabile della simulazione della *parton shower*. L'unica eccezione è costituita da uno tra i sei campioni riferiti alla massa 172.5 MeV, che è stato prodotto anche utilizzando in combinazione gli algoritmi *parton shower* di HERWIG [15].

Gli eventi QCD di fondo risultano invece generati tramite MADGRAPH [16] (anche questo affiancato a PYTHIA).

Di seguito sono riportati in Tabella 3.1 il numero di eventi e la sezione d'urto associati ai file QCD e al file $t\bar{t}$ utilizzato come campione "standard" per il calcolo della massa del quark t .

Infine, il campione dati ottenuto da CMS è riferito al Run effettuato nel 2016, con energia del centro di massa pari a 13 TeV, relativo a collisioni pp e luminosità integrata di

37.0 fb^{-1} .

Campione dati simulati	$N_{evt} (\times 10^7)$	$\sigma(pb)$
$t\bar{t} m_t = 172.5$ GeV (campione "standard")	7.722934	832
eventi QCD <i>multijet</i> ($300 < H_T < 500$ GeV)	5.45379	367000
eventi QCD <i>multijet</i> ($500 < H_T < 700$ GeV)	6.227134	29400
eventi QCD <i>multijet</i> ($700 < H_T < 1000$ GeV)	1.562925	6524
eventi QCD <i>multijet</i> ($1000 < H_T < 1500$ GeV)	1.512729	1064
eventi QCD <i>multijet</i> ($1500 < H_T < 2000$ GeV)	1.18267	121.5
eventi QCD <i>multijet</i> ($H_T > 2000$ GeV)	0.6039005	25.42

Tabella 3.1: Numero di eventi e sezione d'urto associati ai campioni MC utilizzati.

3.2 Selezione dati

Per potere individuare gli eventi di interesse tra la moltitudine di eventi prodotti dalle collisioni pp a LHC, rilevati da CMS e archiviati in un campione dati, occorre selezionare determinati valori o range di valori per le variabili associate agli eventi prodotti dalle interazioni tra particelle.

Con lo scopo di risalire ai decadimenti completamente adronici delle coppie $t\bar{t} \rightarrow WbWB$, con la formazione di due *boosted jet*, per poi calcolare il valore della massa del quark *top*, sono state effettuate le seguenti selezioni sia sui dati che sui campioni MC:

- selezione di trigger;
- selezione sui leptoni elettricamente carichi rivelati (variabile $N_{Leptons}$);
- selezione sulla variabile quantità di moto trasversa dei jet, definita come:

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad (3.1)$$

dove p_x e p_y indicano le proiezioni della quantità di moto del jet sugli assi perpendicolari alla direzione dei fasci pp .

- selezione sulla variabile pseudorapidità (η), definita come:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \quad (3.2)$$

dove θ indica l'angolo tra la direzione del jet e il fascio di protoni;

- selezione sul numero di jet (variabile N_{Jets});
- selezione sul numero di jet *b-tagged*, dunque associati all'adronizzazione di quark *bottom* (variabile N_{BJets}) e sul numero di sotto-jet *b-tagged* (variabile *category*);
- selezione finalizzata alla distinzione tra eventi segnale ed eventi di fondo, basata su una analisi multivariata (variabile *mva*).

Tali selezioni sono state effettuate realizzando apposite macro per il *software* ROOT e accedendo ai file dati e MC, organizzati nelle strutture a *trees*.

3.2.1 Preselezione e studio dell'efficienza del *trigger*

Il trigger *HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20_v*, impiegato nell'analisi dei dati, seleziona eventi in modo tale da ridurre la frequenza a 1 kHz. Tale valore risulta essere il massimo consentito dall'apparato sperimentale per potere acquisire e immagazzinare dati di interesse [17], e viene ottenuto attraverso due distinti trigger: *Level 1* (L1) che riduce il *rate* di eventi prodotti dalle collisioni *pp* (\sim MHz) a 100 kHz, seguito dal *High Level Trigger* (HLT) che porta la frequenza di eventi a 1 kHz.

È inoltre richiesto che il jet più energetico (*Leading jet*) e il secondo jet abbiano quantità di moto trasversa rispettivamente maggiori di 280 GeV e 200 GeV.

Infine sono richiesti almeno un jet taggato *b* e un valore di *TrimMass* maggiore di 30 GeV, che facilita il rigetto di eventi QCD di fondo [18].

Prima di procedere alla selezione dei dati di interesse, è stata studiata l'efficienza del trigger:

$$\epsilon = \frac{n_{sel}}{n_{ref}} \quad (3.3)$$

definita come il rapporto tra il numero di eventi che hanno superato le selezioni intrinseche del trigger sopracitato (n_{sel}) e il numero di eventi che hanno invece superato unicamente le condizioni imposte da un trigger scelto come riferimento (n_{ref}). Quest'ultimo è nominato *HLT_AK8PFJet200_v*, e richiede unicamente che la quantità di moto trasversa del *leading jet* sia superiore a 200 GeV.

Per poter confrontare gli eventi che superano ciascuno dei due trigger, occorre che entrambi abbiano le medesime restrizioni minimali sulle variabili. Poiché il trigger di selezione richiede intrinsecamente la presenza di almeno un jet *b-tagged* e due jet generici, sono state inserite, insieme al passaggio delle condizioni del trigger di riferimento, le ulteriori richieste: $N_{Jets} > 1, N_{BJets} > 0$.

Per entrambi i trigger è stata richiesta inoltre l'assenza di segnale proveniente da leptoni carichi ($N_{Leptons} = 0$), poiché risultano di interesse unicamente gli eventi associati ad adroni.

L'efficienza del trigger in funzione della quantità di moto trasversa del primo e del secondo jet è riportata nelle Figure 3.1 e 3.2. Come si può osservare, per valori superiori

ai 400 GeV gli andamenti per dati ed eventi simulati risultano compatibili, pertanto è stato possibile trattare ugualmente i differenti campioni aggiungendo come richiesta di selezione sugli eventi $p_T > 400$ GeV per entrambi i jet.

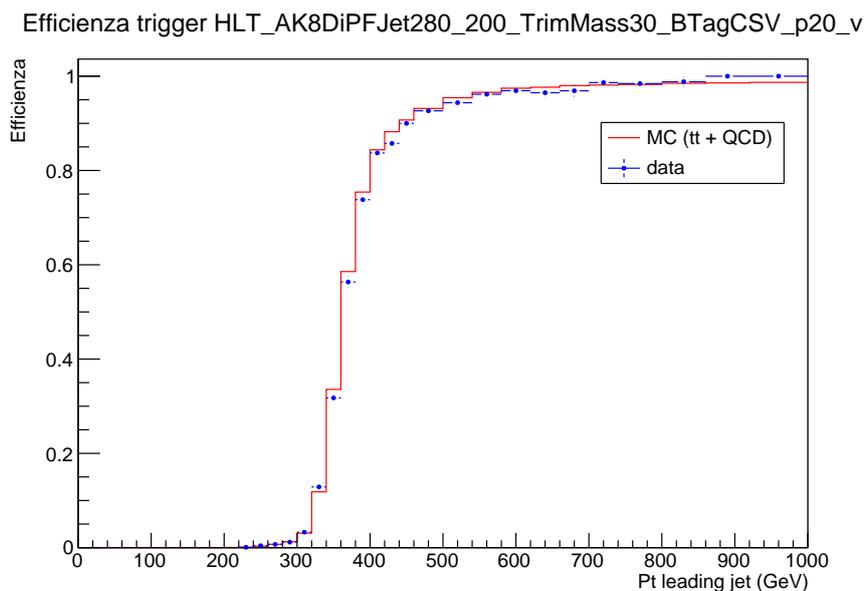


Figura 3.1: Efficienza del trigger $HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20_v$ in funzione del p_T del *leading jet* per dati ed eventi MC.

Efficienza trigger HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20_v

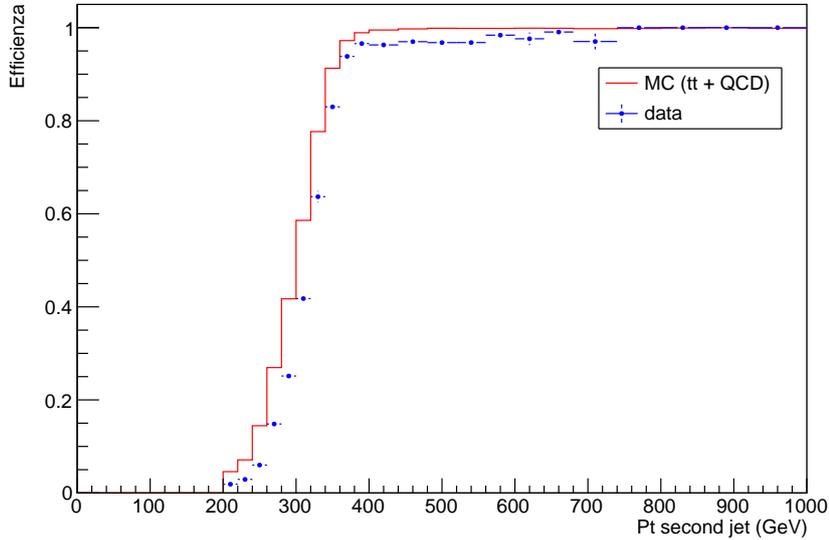


Figura 3.2: Efficienza del trigger $HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20_v$ in funzione del p_T del secondo jet per dati ed eventi MC.

3.2.2 N_{Jets} e $N_{B_{Jets}}$

Una volta determinate le richieste di selezione sulle variabili cinematiche, che nello specifico risultano essere:

- $p_T > 400$ GeV per *leading* e secondo jet,
- $|\eta| < 2.4$ per *leading* e secondo jet, corrispondente ad angoli compresi tra $\theta = +10^\circ$ e $\theta = -10^\circ$ tra la direzione dei jet e la direzione del fascio,

sono stati studiati gli eventi che hanno superato tali selezioni, assieme alla richiesta del trigger. Nelle Figure 3.3 e 3.4 è riportata la distribuzione degli eventi secondo le variabili associate al numero di jet e di jet con *b-tag*.

Si osserva che la maggior parte degli eventi risulta avere due jet, in accordo con il regime *boosted*. Inoltre si nota che negli eventi con due jet *b-tagged* il segnale associato alle coppie $t\bar{t}$ emerge distintamente rispetto agli eventi reali e QCD, pertanto, al fine di formulare il modello per il fondo, è stata aggiunta la richiesta di selezione sugli eventi $N_{B_{Jets}} > 1$.

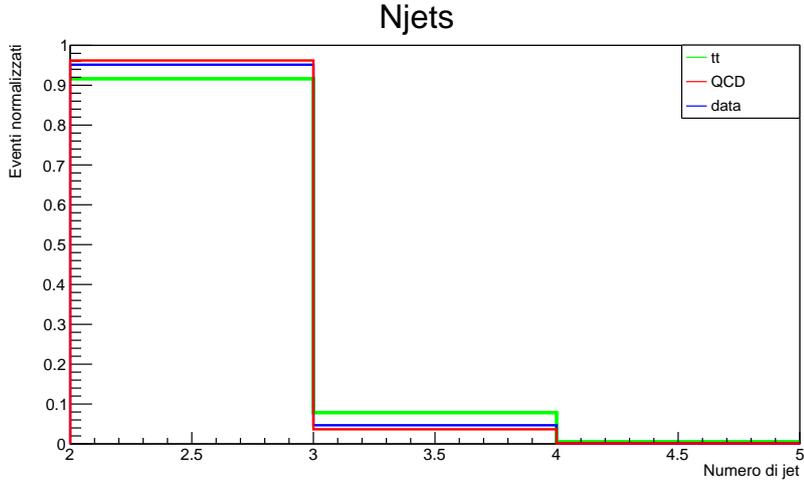


Figura 3.3: Distribuzione (normalizzata) della variabile N_{Jets} .

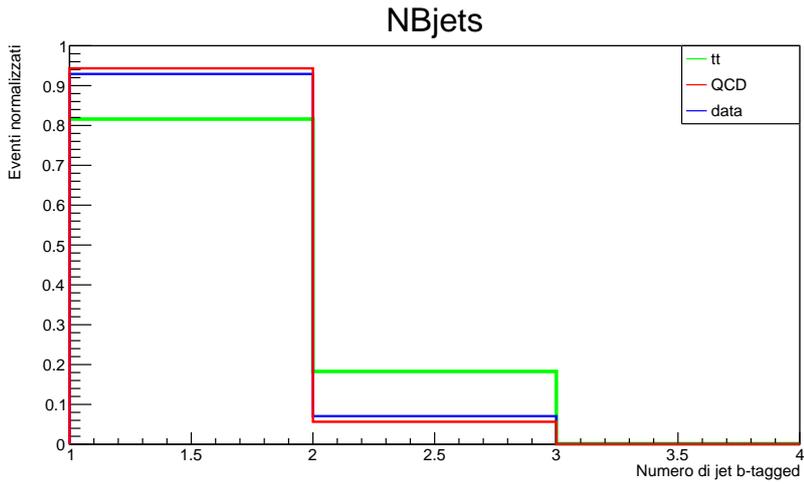


Figura 3.4: Distribuzione (normalizzata) della variabile N_{BJets} .

3.3 Il fondo multijet

3.3.1 Studio della variabile mva

Per potere giungere a una adeguata riduzione del fondo, in modo tale da avere predominanza di eventi associati ai decadimenti prodotti da coppie $t\bar{t}$, è stata utilizzata la variabile mva associata agli eventi. Tale variabile risulta ottenuta attraverso l'impiego di una rete neurale artificiale, e può assumere valori compresi tra 0 e 1, associati alla classificazione di un evento rispettivamente come fondo o segnale $t\bar{t}$. Tale rete neurale

è stata costruita per la misura della sezione d'urto di produzione $t\bar{t}$, ma risulta essere utile anche nel nostro caso. La rete neurale usa come input delle variabili τ_i chiamate *sub-jettiness* che sono associate a sottostrutture presenti entro i due *boosted jet*.

Gli andamenti in funzione di mva riportati in Figura 3.5, mostrano infatti il potere discriminante della rete neurale. Per valori ridotti si ha prevalenza di eventi associati al fondo QCD, che risulta riprodotto anche dai dati reali, mentre per valori elevati domina nettamente il campione MC $t\bar{t}$ e si osserva un allontanamento tra i dati e il fondo, in accordo con la presenza di segnale associato ai decadimenti del quark *top* nel campione di dati reali.

Utilizzando il campione MC $t\bar{t}$ come segnale e i campioni QCD come fondo, sono stati studiati il rapporto segnale rumore $\frac{S}{B}$ e la significanza $\frac{S}{\sqrt{B}}$ in modo tale da determinare la selezione sulla variabile mva da imporre agli eventi per ottenere una riduzione del fondo, e delle sue fluttuazioni, senza diminuire significativamente il segnale (e dunque aumentarne l'errore statistico associato).

Poiché la significanza risulta essere massima in corrispondenza del valore 0.91, come si osserva in Figura 3.6, è stata posta come ulteriore richiesta sugli eventi la condizione $mva > 0.91$, corrispondente a valori del rapporto segnale/rumore superiori a 1.4 (Figura 3.6).

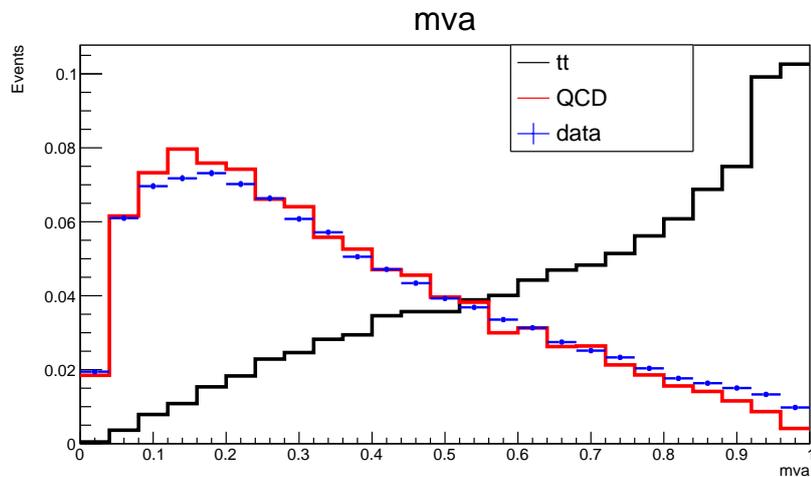


Figura 3.5: Distribuzione (normalizzata) della variabile mva .

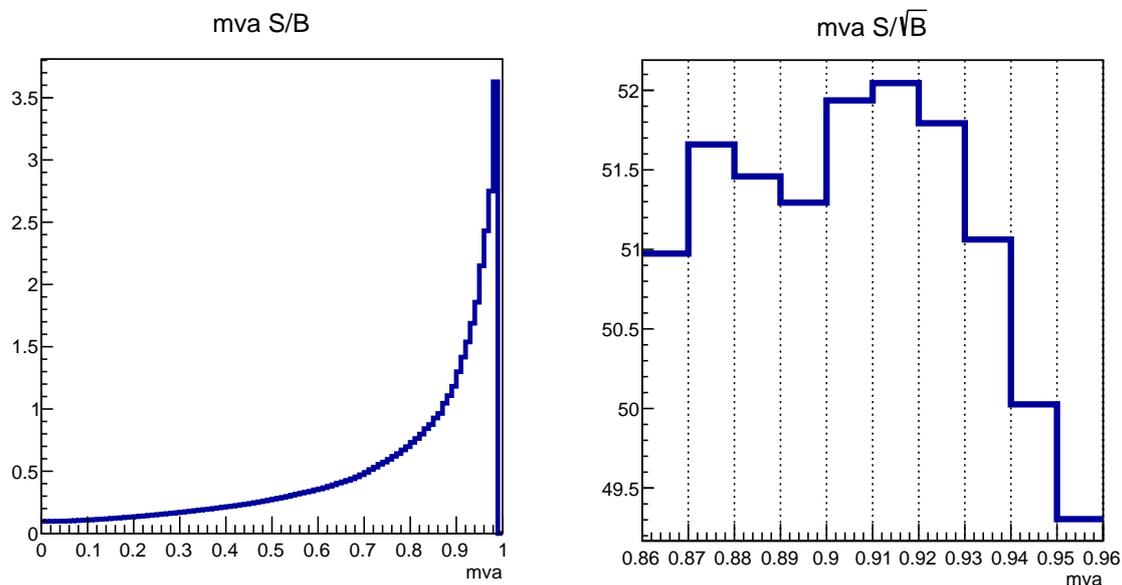


Figura 3.6: Grafico del rapporto segnale/rumore S/B (a sinistra) e particolare del picco di significanza S/\sqrt{B} (a destra).

3.3.2 variabile category

Un'ultima richiesta di selezione sui campioni dati viene effettuata sulla variabile *category*, la quale indica il numero di sotto-jet taggati b presenti in jet distinti. Pertanto, richiedendo la condizione $category = 2$ si otterrà una maggiore distinzione tra i dati e il fondo, come si può osservare in Figura 3.7

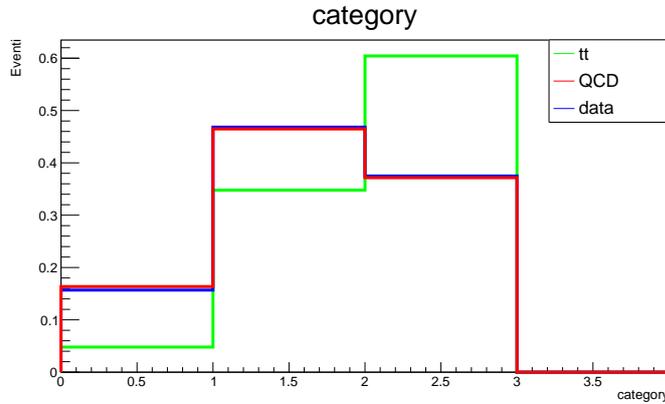


Figura 3.7: Grafico degli eventi (normalizzati) in funzione della variabile *category*.

3.3.3 Modello per il fondo

Il modello per il fondo può essere formulato a partire dal campione di dati reali oppure utilizzando i campioni MC QCD. Tuttavia, poiché questi ultimi risultano avere grandi incertezze associate alle sezioni d'urto degli eventi multijet, un modello costruito a partire dai campioni QCD risulterebbe essere meno preciso rispetto a uno formulato utilizzando il campione di dati reali. Pertanto, al fine di misurare la massa del quark *top*, è stato realizzato un modello per il fondo a partire da un campione complementare di dati, accertandosi che questo riproducesse l'andamento del fondo ottenuto mediante i campioni QCD.

Nello specifico, entrambi i modelli sono stati realizzati impiegando la variabile *Soft Drop Mass* riferita al jet più energetico e richiedendo l'assenza di eventi associati a jet taggati *b*. Come si osserva in Figura 3.8, gli andamenti risultano avere forma simile, tuttavia non sono perfettamente coincidenti a causa di contributi residui presenti nel campione dei dati associati a eventi di segnale.

Inoltre ci si aspetta che l'andamento del fondo QCD non muti considerevolmente richiedendo la presenza di uno o due jet *b-tagged*, come si osserva in Figura 3.9, 3.10.

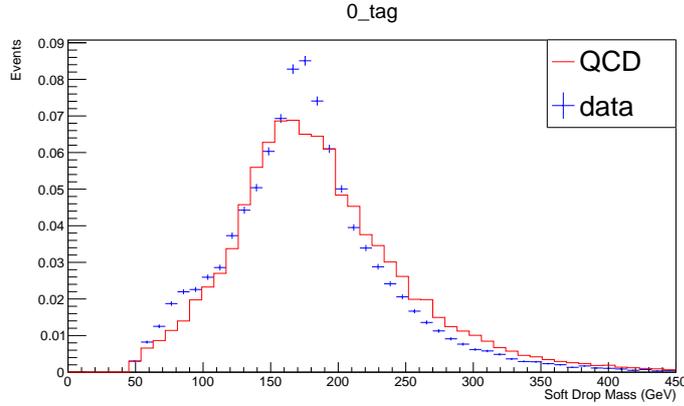


Figura 3.8: Distribuzione della variabile *Soft Drop Mass* per dati e eventi QCD in assenza di jet taggati b .

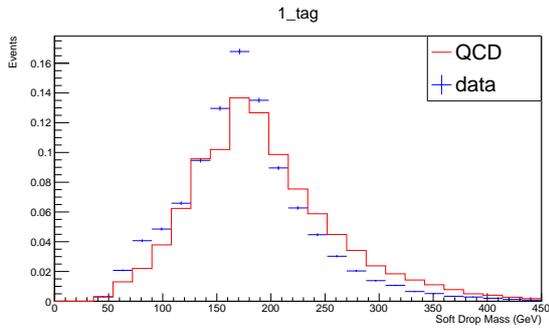


Figura 3.9: Distribuzione della variabile *Soft Drop Mass* per dati in assenza di jet taggati b e eventi QCD con la richiesta di un jet b -tagged.

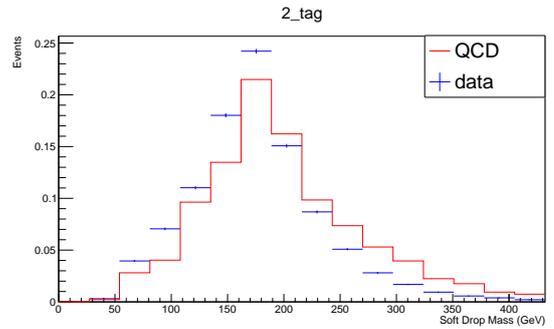


Figura 3.10: Distribuzione della variabile *Soft Drop Mass* per dati in assenza di jet taggati b e eventi QCD con la richiesta di due jet b -tagged.

3.4 Misura della massa

3.4.1 Template

Per procedere alla misura effettiva della massa del quark top , sono stati utilizzati, oltre al campione "standard" $t\bar{t}$ riferito al valore di massa pari a 172.5 GeV, sei ulteriori campioni MC relativi ad eventi $t\bar{t}$ riferiti a differenti valori per la massa del quark top .

Per ogni campione, attraverso una macro ROOT, sono state costruite delle distribuzioni normalizzate (*template*) per segnale e fondo, utilizzando la variabile *Soft Drop Mass* e normalizzando le distribuzioni, in modo tale da ottenere delle funzioni di densità di probabilità (*pdf*) associate al segnale proveniente da eventi $t\bar{t}$ e a eventi QCD di fondo.

Pertanto il contenuto di ogni bin (N_i) in tali distribuzioni risulta essere dato dalla

seguinte relazione:

$$N_i = \frac{n_i}{\sum_{i=1}^{N_{tot}} n_i} \quad (3.4)$$

dove n_i indica il numero di eventi attesi nel bin e N_{tot} il numero complessivo di bin. In particolare, per il fondo è stato utilizzato il campione di dati reali imponendo l'assenza di segnale b -tag sia per il *leading* che per il secondo jet.

3.4.2 Fit con funzione di massima verosimiglianza

Il metodo statistico della massima verosimiglianza consente di stimare i valori di uno o più parametri relativi ad una funzione di distribuzione di probabilità nota partendo da un campione finito di dati [19]. Una volta definita la funzione di massima verosimiglianza (*Likelihood function*) \mathcal{L} , si determina il valore del set di parametri (o del singolo parametro) ricercato che massimizza la verosimiglianza.

Nella trattazione di seguito riportata le *pdf* risultano essere le distribuzioni normalizzate di segnale e fondo per la variabile *Soft Drop Mass*, mentre la funzione di *Likelihood* risulta definita da:

$$\mathcal{L}(n_s, n_b) = \mathcal{L}_{norm} \cdot \mathcal{L}_{shape} \quad (3.5)$$

dove n_s e n_b indicano rispettivamente il numero di eventi attesi per segnale e fondo, e risultano essere i due parametri ricercati attraverso il fit della distribuzione effettuato tramite \mathcal{L} .

\mathcal{L}_{norm} è il fattore di normalizzazione Poissoniano, e risulta essere definito come segue:

$$\mathcal{L}_{norm} = e^{-\mu} \frac{\mu^n}{n!} \quad (3.6)$$

dove n indica il numero di eventi totali utilizzati, che risulta essere $n = 8658$, mentre μ è dato dalla somma tra eventi segnale ed eventi fondo attesi: $\mu = n_s + n_b$.

Il termine associato alla forma delle *pdf* di segnale e fondo impiegate, \mathcal{L}_{shape} , risulta essere definito come:

$$\mathcal{L}_{shape} = \prod_{i=1}^{N_{tot}} \frac{n_s P_s(i) + n_b P_b(i)}{\mu} \quad (3.7)$$

dove $P_s(i)$ e $P_b(i)$ indicano la densità di probabilità relativa allo i -esimo bin di ciascuna distribuzione di densità di probabilità considerata.

Definita la funzione di verosimiglianza, è stato effettuato un fit per le distribuzioni di dati e fondo utilizzando separatamente ciascuno dei sette distinti campioni MC relativi agli eventi $t\bar{t}$.

Tale fit non è stato ottenuto massimizzando la *Likelihood*, ma minimizzando $-\ln(\mathcal{L})$, mediante l'algoritmo MINUIT. Pertanto la funzione da minimizzare tramite il fit risulta essere:

$$-\ln(\mathcal{L}(n_s, n_b)) = -\mu + n \ln(\mu) + \ln(n!) - \sum_{i=1}^{N_{tot}} \ln \left(\frac{n_s P_s(i) + n_b P_b(i)}{\mu} \right) \quad (3.8)$$

Di seguito sono riportate nelle Figure 3.11-3.17 le distribuzioni restituite dai fit per i dati, gli eventi simulati $t\bar{t}$ e gli eventi di fondo per i diversi campioni MC $t\bar{t}$ impiegati; i parametri ottenuti sono riportati in Tabella 3.2.

In Figura 3.18 sono rappresentati i differenti valori della funzione di *Likelihood* ottenuti attraverso i distinti campioni MC $t\bar{t}$. Tali valori sono stati traslati rispetto all'asse y , ponendo il minimo a zero. Come si può osservare, il campione che minimizza $-\ln(\mathcal{L})$, ovvero massimizza la verosimiglianza, è quello relativo alla massa $m_t = 172.5$ GeV, come è ulteriormente confermato dai valori in Tabella 3.2.

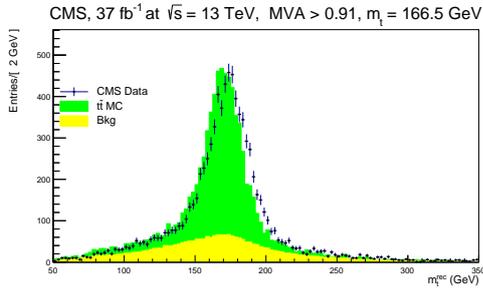


Figura 3.11: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 166.5 \text{ GeV}$. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

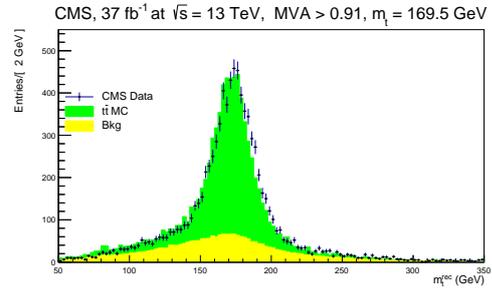


Figura 3.12: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 169.5 \text{ GeV}$. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

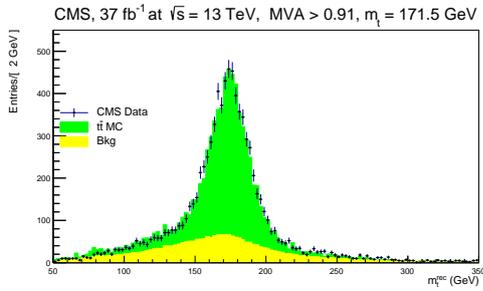


Figura 3.13: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 171.5 \text{ GeV}$. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

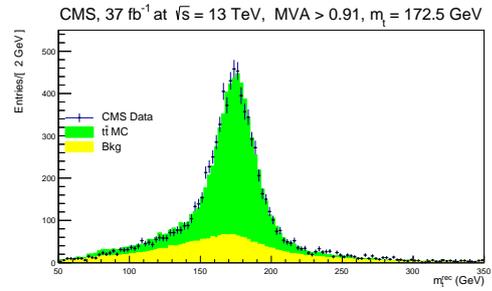


Figura 3.14: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 172.5 \text{ GeV}$. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

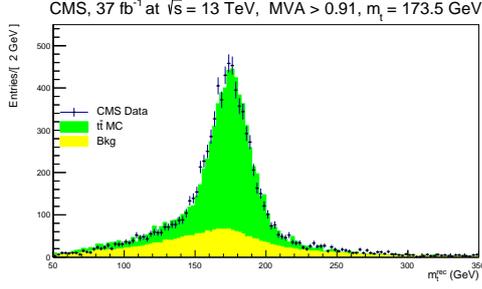


Figura 3.15: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 173.5$ GeV. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

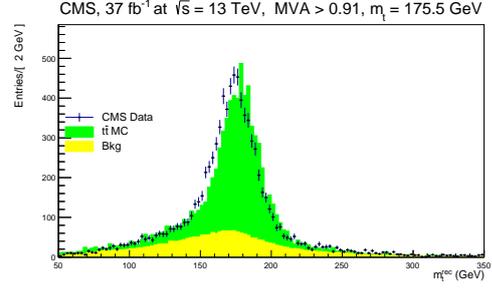


Figura 3.16: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 175.5$ GeV. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

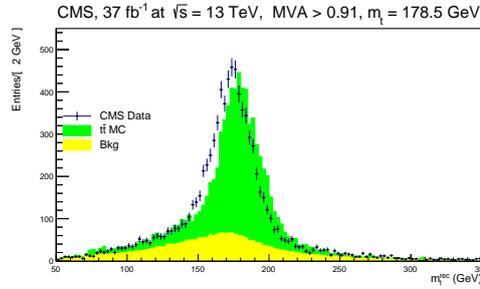


Figura 3.17: Distribuzione relativa al campione MC $t\bar{t}$ per $m_t = 178.5$ GeV. Il segnale e il fondo sono normalizzati ai valori restituiti dal fit.

<i>Template</i>	n_s	n_b	n	$-\ln(\mathcal{L})$
166.5 GeV	5487 ± 125	3171 ± 115	8658	34454.4
169.5 GeV	6067 ± 132	2591 ± 118	8658	34348.7
171.5 GeV	6038 ± 129	2620 ± 115	8658	34308.9
172.5 GeV	6062 ± 130	2596 ± 116	8658	34293.3
173.5 GeV	5862 ± 126	2796 ± 113	8658	34318.4
175.5 GeV	5451 ± 123	3207 ± 114	8658	34422.3
178.5 GeV	5355 ± 127	3303 ± 118	8658	34508.2

Tabella 3.2: Parametri del fit tramite la massima verosimiglianza ottenuti utilizzando i sette distinti campioni $t\bar{t}$.

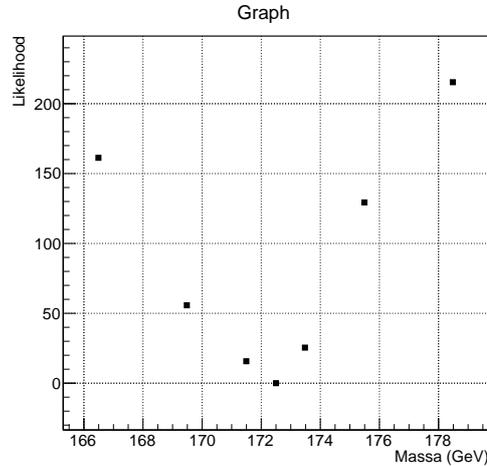


Figura 3.18: Valori della funzione $-\ln(\mathcal{L})$ ottenuti dai differenti campioni MC $t\bar{t}$.

3.4.3 Valore di massa

Per potere associare degli errori ai valori di *Likelihood* ottenuti mediante il fit, è stata utilizzata un'ulteriore macro ROOT in modo tale da fare fluttuare 100 volte secondo Poisson il contenuto di ogni bin dei diversi *template*, mantenendo tuttavia la normalizzazione.

Ogni valore della *Likelihood* ottenuto risulta dunque dalla media dei valori assunti in ognuna delle 100 fluttuazioni:

$$\mathcal{L} = \frac{\sum_{n=1}^{100} \mathcal{L}_n}{100} \quad (3.9)$$

L'errore associato a tali valori è stato calcolato tramite la relazione statistica:

$$\sigma_{\mathcal{L}} = \sqrt{\frac{\sum_{n=1}^{100} \mathcal{L}_n^2}{100} - \left(\frac{\sum_{n=1}^{100} \mathcal{L}_n}{100}\right)^2} \quad (3.10)$$

In Figura 3.19 è mostrato il grafico delle distribuzioni di dati, eventi $t\bar{t}$ e eventi di fondo, ottenuto conseguentemente alle fluttuazioni statistiche.

Infine, per ottenere il valore associato alla massa del quark *top*, è stato eseguito un fit tramite una funzione polinomiale di terzo grado sui sette valori della *Likelihood* ottenuti a seguito delle fluttuazioni dei bin, come è mostrato in Figura 3.20. I parametri di fit sono riportati in Tabella 3.3.

Il minimo della funzione di fit risulta essere in corrispondenza del valore di massa:

$$m_t = 171.9 \text{ GeV.}$$

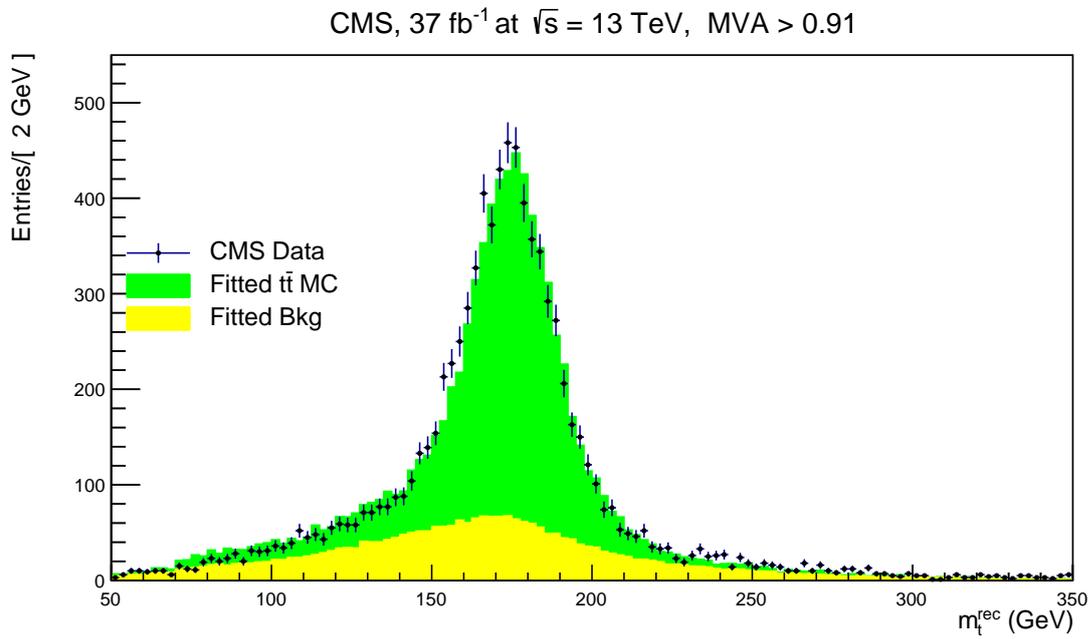


Figura 3.19: Distribuzioni di dati, segnale e fondo relativi al campione $t\bar{t}$ che massimizza la *Likelihood* ottenute a seguito del fit e delle fluttuazioni statistiche di ciascun bin.

p0	p1	p2	p3
380671	-5650.27	27.0952	-0.0413482

Tabella 3.3: Valori dei parametri della funzione polinomiale di terzo grado ottenuti tramite il fit.

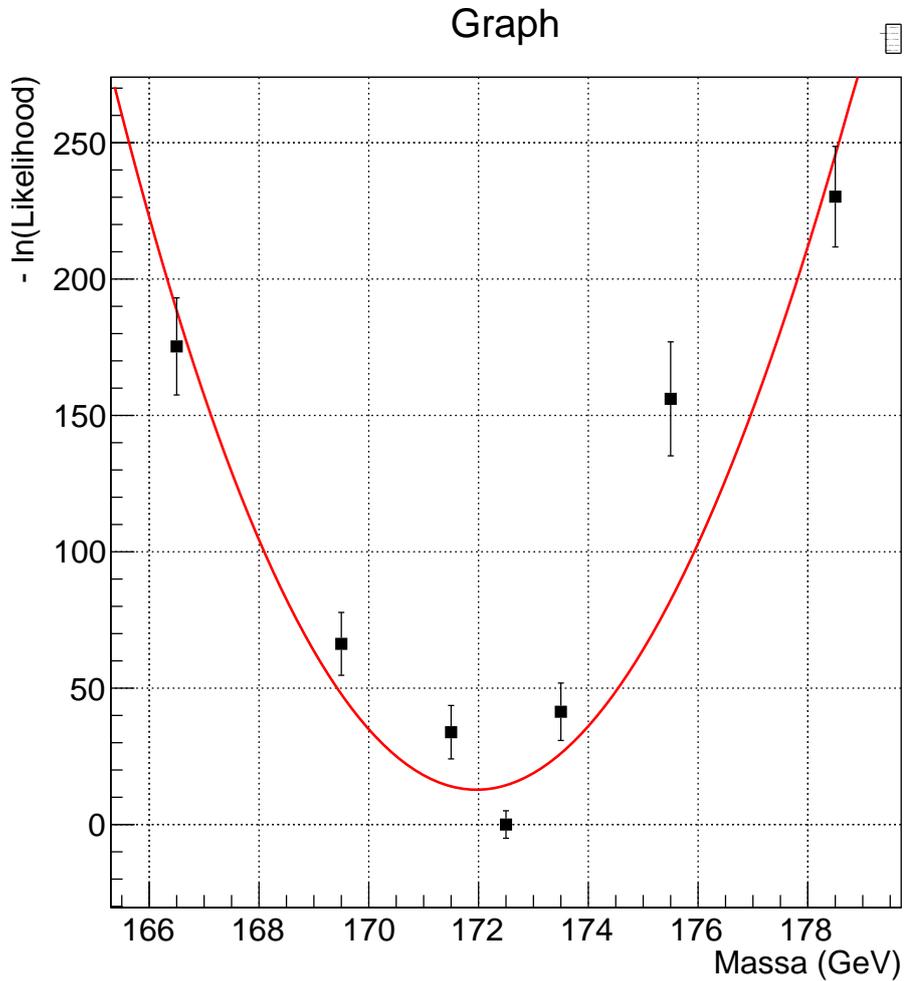


Figura 3.20: Fit tramite una funzione polinomiale di terzo grado sui valori di $-\ln(\text{Likelihood})$.

3.4.4 Errore statistico

Poiché una variazione sul valore della funzione di massima verosimiglianza pari ad una deviazione standard corrisponde a una variazione sul valore di $-\ln(\mathcal{L})$ pari a mezza unità, per determinare l'errore statistico da associare a m_t , si sono considerati i valori di massa corrispondenti ad una variazione del minimo della funzione polinomiale pari a $+0.5$. Tramite tali valori, è stato determinato l'errore statistico associato utilizzando la formula:

$$\delta m_{stat} = \frac{|m_{+0.5} - m'_{+0.5}|}{2} \quad (3.11)$$

Tale valore risulta essere:

$$\delta m_{stat} = 0.3 \text{ GeV.}$$

3.4.5 Errore sistematico

Per stimare l'errore sistematico da associare alla misura di massa del quark *top* sono stati utilizzati dei campioni MC $t\bar{t}$, riferiti alla massa di 172.5 GeV, contenenti eventi ottenuti tramite simulazioni di differenti fenomeni che possono avvenire durante il decadimento della coppia $t\bar{t}$ e di limiti strumentali di rivelazione che possono subentrare in regimi di elevata energia, quali quelli in cui si opera a LHC.

Nello specifico sono stati utilizzati:

- campioni associati a valori diversi per la emissione di radiazione (forte) finale e iniziale, rispettivamente nominati *fsr* (*final state radiation*) e *irs* (*initial state radiation*);
- campioni associati alle incertezze di rivelazione dipendenti dalla scala energetica (JES, *Jet Energy Scale*) e dalla risoluzione energetica del rivelatore (JER, *Jet Energy Range*), nominati *smeared* e *shifted*;
- un singolo campione ottenuto tramite l'algoritmo di *parton showering* HERWIG, da confrontare con il campione standard, che invece contiene eventi simulati di *parton showering* tramite PYTHIA.
- infine è stato effettuato il fit di verosimiglianza impiegando i campioni QCD per il modello del fondo. Il valore di massa ottenuto è stato confrontato con quello ricavato tramite la misura standard, il cui fondo risulta costruito a partire dai dati.

Non disponendo di campioni relativi a tutti i valori di massa del quark *top*, non si possono produrre degli insiemi di *template* modificati. Si può però fare una stima dell'effetto che le modifiche avrebbero sulla misura andando a guardare l'effetto rispetto al campione standard con $m_t = 172.5 \text{ GeV}$.

Per ciascun campione è stato ottenuto il valore medio della distribuzione della variabile *Soft Drop Mass* relativa al *leading jet*, utilizzando gli eventi che hanno superato tutte le condizioni di selezione richieste (trigger, assenza di leptoni, due jet taggati *b* distinti, $p_T > 400$, $|\eta| < 2.4$, $m_{va} > 0.91$). Tali valori risultano essere:

$$m_t^{fsr-up} = 159.5 \text{ GeV}$$

$$m_t^{fsr-down} = 155.6 \text{ GeV}$$

$$m_t^{isr-up} = 157.1 \text{ GeV}$$

$$m_t^{isr-down} = 156.8 \text{ GeV}$$

$$m_t^{smeared-up} = 157.0 \text{ GeV}$$

$$m_t^{smeared-down} = 157.0 \text{ GeV}$$

$$m_t^{shifted-up} = 156.9 \text{ GeV}$$

$$m_t^{shifted-down} = 157.2 \text{ GeV}$$

$$m_t^{PYTHIA} = 157.0 \text{ GeV}$$

$$m_t^{HERWIG} = 157.3 \text{ GeV}$$

L'errore relativo ricavato utilizzando le prime quattro coppie risulta:

•

$$\frac{|m_t^{fsr-up} - m_t^{fsr-down}|}{m_t^{fsr-up} + m_t^{fsr-down}} = 0.01$$

•

$$\frac{|m_t^{isr-up} - m_t^{isr-down}|}{m_t^{isr-up} + m_t^{isr-down}} = 0.001$$

•

$$\frac{|m_t^{smeared-up} - m_t^{smeared-down}|}{m_t^{smeared-up} + m_t^{smeared-down}} = 0$$

•

$$\frac{|m_t^{shifted-up} - m_t^{shifted-down}|}{m_t^{shifted-up} + m_t^{shifted-down}} = 0.001$$

L'errore relativo associato ai due diversi algoritmi *parton-shower* è stato invece ottenuto attraverso la seguente formula:

$$\frac{|m_t^{PYTHIA} - m_t^{HERWIG}|}{m_t^{PYTHIA}} = 0.002$$

Infine, tramite il fit di verosimiglianza sulle distribuzioni ottenute impiegando per il fondo i campioni QCD, è stato ottenuto:

$$m_t^{QCD} = 171.9 \text{ GeV}$$

Pertanto l'errore relativo associato all'impiego del campione dati o dei campioni QCD per il fondo risulta essere:

$$\frac{|m_t^{QCD} - m_t^{standard}|}{m_t^{standard}} = 0$$

I parametri di fit n_s e n_b riferiti al campione relativo alla massa di 172.5 GeV risultano tuttavia essere differenti rispetto a quelli ottenuti precedentemente:

$$n_s = 6361 \pm 120$$

$$n_b = 2297 \pm 101$$

Sommando gli errori ottenuti in quadratura, considerandoli dunque indipendenti, è stato ottenuto:

$$\frac{\delta m_{sist}}{m} = \sqrt{\sum_i \left(\frac{\delta m_i}{m}\right)^2} = 0.01$$

Come si può osservare, il contributo maggiore risulta essere dovuto al modello per la radiazione di stato finale.

3.4.6 Risultato finale

Il risultato finale per la massa del quark *top* risulta essere:

$$m_t = 171.9 \pm 0.3 (stat) \pm 1.7 (sist) \text{ GeV}, \quad (3.12)$$

dove la prima incertezza è statistica e la seconda sistematica.

Conclusioni

Lo studio delle caratteristiche del quark *top* risulta essere ancora oggi di notevole interesse nella fisica delle particelle, nonostante la prima rivelazione sperimentale di tale quark risalga al 1995. Infatti esso risulta essere l'unico tra i quark presenti nel Modello Standard a non adronizzare, ma a decadere unicamente per interazione debole nella coppia Wb . Tra le diverse proprietà della particella fondamentale, la massa costituisce una variabile di notevole interesse, poiché a partire da questa risulta possibile risalire ad ulteriori proprietà cinematiche del quark e studiare approfonditamente i processi cui prende parte. In questo elaborato è stata misurata la massa del quark *top* attraverso la ricostruzione dello stato finale in cui decade la coppia $t\bar{t}$ secondo il canale completamente adronico: $t\bar{t} \rightarrow WbWb$, che risulta avere un'elevata *Branching Ratio* ($\approx 45\%$) ma allo stesso tempo costituisce il canale in cui il segnale risulta meno distinguibile dai jet adronici causati da eventi QCD.

Partendo da campioni di dati reali (acquisiti tramite il rivelatore CMS a LHC) e campioni simulati secondo algoritmi Monte Carlo, sono stati selezionati gli eventi associati alla presenza di due jet ampi con $p_T > 400$ GeV ciascuno, che risultano pertanto essere in regime *boosted*, con l'ulteriore richiesta di essere entrambi taggati *b*.

Per potere distinguere i dati dal fondo, è stata imposta una selezione sulla variabile m_{vb} , ottenuta tramite una rete neurale artificiale e che utilizza in input variabili connesse alla sottostruttura dei due jet, richiedendo $m_{vb} > 0.91$. La distribuzione della massa invariante del jet più energetico in regime *boosted*, per eventi di segnale e di fondo, è stata usata come grandezza sperimentale associata alla massa del quark *top*. Utilizzando differenti campioni MC per eventi $t\bar{t}$, corrispondenti a diversi valori di massa per il quark *top*, si è ottenuta la distribuzione di massima verosimiglianza, e, successivamente si è risaliti al valore della massa del quark *top* che meglio riproduce la distribuzione di massa invariante osservata.

Utilizzando i campioni di dati acquisiti tramite CMS durante il Run 2016 ad energia del centro di massa per le collisioni pp pari a 13 TeV, e luminosità integrata di 37.0 fb^{-1} , è stato ottenuto il seguente valore:

$$m_t = 171.9 \pm 0.3 \text{ (stat)} \pm 1.7 \text{ (sist)} \text{ GeV}$$

che risulta essere in accordo, entro le incertezze, con la media mondiale di 173.1 ± 0.6 GeV.

Bibliografia

- [1] C. Patrignani et al. (Particle Data Group), *Chin. Phys. C*, 40 , 100001 (2016) and 2017 update.
- [2] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys.Lett. B* 716 (2012) 1
- [3] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *PHYS.IETT b* 716 (2012) 30
- [4] Sylvie Braibant, Giorgio Giacomelli, Maurizio Spurio. *Particelle ed interazioni fondamentali*. Springer Verlag; seconda edizione (2012).
- [5] Donald H. Perkins. *Introduction to high energy physics*. Cambridge University Press; quarta edizione (2000).
- [6] Marcel Vos (ATLAS and CMS collaborations), *Top-quark mass measurements at the LHC: alternative methods*, arXiv:1602.00428v2 [hep-ex], 4 Feb 2016
- [7] F. Maltoni (Univ. Catholique de Louvain) and A. Quadt (Univ. Gottingen). Updated September 2015 by T.M. Liss (The City College of New York) review del pdg.
- [8] Ulrich Husemann, *Top-Quark Physics: Status and Prospects*. arXiv:1704.01356v2 [hep-ex] 19 Apr 2017.
- [9] Enrico Bothmann, Frank Krauss, and Marek Schonherr, *Single top-quark production with SHERPA*, arXiv:1711.02568v1 [hep-ph] 7 Nov 2017.
- [10] Andrew J. Larkoski, Simone Marzani, Gregory Soyez and Jesse Thaler, *Soft Drop*, arXiv:1402.2657v2 [hep-ph], 13 May 2014
- [11] *Sito web ufficiale del CERN*.
<https://home.cern/about>

- [12] *Sito web ufficiale di CMS.*
<https://cms.cern>
- [13] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [14] T. Sjöstrand, S. Mrenna and P. Skands, *A Brief Introduction to PYTHIA 8.1*, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [15] Daniel Reichelt, Peter Richardson, Andrzej Siodmok, *Improving the Simulation of Quark and Gluon Jets with Herwig 7*, arXiv:1708.01491v1 [hep-ph].
- [16] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5: Going Beyond*, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522v1.
- [17] Andrea Perrotta, *Performance of the CMS High Level Trigger 2015* J. Phys.: Conf. Ser.664082044.
- [18] Dylan Rankin (Boston University), on behalf of the CMS Collaboration, *CMS triggers with boosted objects or jet substructure*, <https://indico.cern.ch/event/382815/contributions/910647/attachments/1139947/1632553/CMSboostedtriggers.pdf>
- [19] Glen Cowan, *Statistical data analysis*, Clarendon Press, Oxford, 1998.