# Università degli Studi di Padova



Dipartimento di Fisica e Astronomia "Galileo Galilei" Corso di Laurea in Fisica

Tesi di Laurea

# Ricostruzione della massa dei bosoni di Higgs nel processo $HH \rightarrow b\bar{b}b\bar{b}$ con l'esperimento CMS

**Relatore**: Dott. Tommaso Dorigo **Correlatore**: Dott. Martino Dall'Osso

> Laureanda: Valeria Fioroni Matricola: 1103232

Anno Accademico 2016-2017

#### Sommario

In questo lavoro di tesi è presentata l'analisi di eventi simulati per il processo HH  $\rightarrow$  bbbb. Nel corso dell'analisi degli eventi simulati per il Run 2 di LHC (collisioni protone-protone con energia nel centro di massa  $\sqrt{s} = 13$  TeV), è stato studiato l'impatto sulla ricostruzione della massa dei bosoni di Higgs della scelta del valore di soglia  $\Delta R^*_{cut}$  (con  $\Delta R^*$  valore medio della distanza angolare tra jets ricostruiti e b-quarks generati di un evento), in modo da ottimizzare la strategia di controllo della corretta selezione dei b-jets. Il valore ottimale ottenuto è  $\Delta R^*_{cut} = 0.4$ . La massa dei bosoni di Higgs è stata ricostruita dai b-jets selezionati ed è stato stimato il valore medio della distribuzione attraverso un fit gaussiano:  $m_{\rm H} = 116.3 \pm 0.4 \frac{\text{GeV}}{\text{c}^2}$ . Successivamente sono stati analizzati gli eventi simulati per HL-LHC, studiando l'andamento dell'efficienza di selezione degli eventi e della frazione di jets correttamente selezionati al variare del pileup (PU=0, PU=140, PU=200) e dei valori di taglio sulle variabili dei jets  $p_T$  e  $\eta$ .

# Indice

1	LHC	C e l'esperimento CMS	1							
	1.1	LHC: Large Hadron Collider	1							
		1.1.1 HL-LHC: High Luminosity LHC	2							
	1.2	CMS, il Compact Muon Solenoid	3							
		1.2.1 Il sistema di coordinate	3							
		1.2.2 Il tracker	4							
		1.2.3 Il calorimetro elettromagnetico ECAL	4							
		1.2.4 Il calorimetro adronico HCAL	4							
		1.2.5 Il rivelatore di muoni	5							
		1.2.6 Il trigger	6							
		1.2.7 CMS Phase 2 upgrade	6							
2	Proc	luzione di coppie di bosoni di Higgs	9							
	2.1	Produzione di coppie HH	9							
	2.2	Stato delle ricerche della produzione di coppie di bosoni di Higgs	10							
3	Iets		11							
	3.1	Adronizzazione	11							
	3.2	Ricostruzione dei jets	11							
	3.3	b-tagging dei jet	12							
4	Ana	lisi	13							
	4.1	Obiettivi e strategia di analisi	13							
	4.2	Simulazioni	13							
	4.3	3 Analisi degli eventi simulati a $\sqrt{s} = 13$ TeV e pileup 20								
		4.3.1 Selezione degli eventi	14							
		Selezione di 4 b-jets e accoppiamento in di-jets	15							
		4.3.2 Pairing dei b-jets selezionati con i quark b generati	17							
		4.3.3 Ricostruzione della massa dei bosoni di Higgs	19							
	4.4	Analisi degli eventi simulati per HL-LHC	22							
		4.4.1 Campione di eventi con pileup PU=0	22							
		4.4.2 Campione di eventi con pileup PU=140	23							
		4.4.3 Campione di eventi con pileup PU=200	23							
5	Con	clusioni	27							

# Capitolo 1 LHC e l'esperimento CMS

#### 1.1 LHC: Large Hadron Collider

Il *Large Hadron Collider* (LHC) [1] è un acceleratore di particelle costruito presso il CERN di Ginevra allo scopo di ottenere collisioni tra fasci di particelle ad alta energia. LHC è formato da una sequenza di magneti superconduttori, dipoli e quadrupoli, posti all'interno di un tunnel di circonferenza pari a 27 km (Fig. 1.1). L'apparato è stato progettato



Figura 1.1: Schema del sistema di acceleratori del CERN

1

per far collidere fasci di protoni a un'energia nel centro di massa  $\sqrt{s} = 14$  TeV, con una luminosità  $\mathcal{L}=10^{34}$  cm<sup>-2</sup>s<sup>-1</sup> e per far collidere ioni pesanti con un'energia  $\sqrt{s} = 5.5$  TeV per coppia di nucleoni. I fasci circolano uno in verso orario (*Beam 1*) e l'altro in verso antiorario (*Beam 2*) in due anelli distinti, e possono collidere in quattro punti di interazione. In corrispondenza dei punti di interazione sono posizionati quattro detectors: CMS, ATLAS, ALICE e LHCb.

Dall'inizio dell' attività di LHC ad oggi si sono susseguite due fasi denominate Run 1 e Run 2. Il Run 1 è iniziato nel 2009 ed è terminato nel 2012. In questa fase l'energia nel centro di massa delle collisioni p-p è stata incrementata fino a raggiungere il valore  $\sqrt{s} = 8$  TeV nel 2012. Al Run 1 segue un perido di shut-down negli anni 2013 e 2014 in cui viene effettuato un upgrade per permettere il raggiungimento della luminosità istantanea prevista dal progetto originale di LHC. Nel corso del Run 2, fase iniziata nel 2015 e che terminerà nel 2018, nelle collisioni p-p è stata raggiunta un'energia nel centro di massa  $\sqrt{s} = 13$  TeV.

#### 1.1.1 HL-LHC: High Luminosity LHC

A partire dal 2024 è previsto un miglioramento delle prestazioni di LHC grazie al progetto denominato High Luminosity LHC (HL-LHC). La sostituzione dei magneti superconduttori e di altri elementi della macchina permetteranno di raggiungere una luminosità istantanea pari a  $5 \cdot 10^{34}$  cm<sup>-2</sup>s<sup>-1</sup> e quindi una luminosità integrata di 250 fb<sup>-1</sup> per anno, con l'obiettivo di raggiungere i 3000 fb<sup>-1</sup> entro una dozzina di anni dall'upgrade. Questa luminosità integrata è dieci volte maggiore rispetto a quella prevista per i primi 12 anni di attività di LHC [2], come mostrato in Fig. 1.2.



Figura 1.2: Programma HL-LHC [2]. La linea viola indica l'energia delle collisioni mentre la verde la luminosità. In basso sono riportati i valori previsti per la luminosità integrata.

Il pileup, ovvero il numero di collisioni che avvengono durante un singolo bunch crossing, è proporzionale alla luminosità istantanea. Con l'aumento della luminosità istantanea previsto dal programma HL-LHC, ci si aspetta un pileup medio di 140, fino a un massimo di 200 [3]. L'upgrade dei quattro esperimenti principali di LHC è previsto a partire dal 2024 per far fronte a queste nuove condizioni di operatività.

#### 1.2 CMS, il Compact Muon Solenoid

Il detector CMS [4] (Fig. 1.3) è dotato di un magnete solenoidale superconduttore lungo 13 m e con un diametro interno di 5.9 m che genera un campo magnetico di 3.8 T. All'esterno del magnete è collocato un sistema di rivelatori di muoni lungo 21 m e con un diametro di 14.6 m. All'interno del criostato del magnete sono posizionati un dispositivo adibito al tracking di particelle cariche e un calorimetro diviso in due sezioni (il calorimetro elettromagnetico ECAL e il calorimetro adronico HCAL).



Figura 1.3: Schema del detector CMS [4]

#### 1.2.1 Il sistema di coordinate

Il sistema di coordinate di CMS ha l'origine nel punto di collisione dei fasci ed è definito in modo che l'asse x punti verso il centro dell'anello di LHC, l'asse y punti verticalmente verso l'alto (perpendicolarmente al piano di LHC) e l'asse z sia diretto lungo il fascio che circola in senso antiorario (Beam 2). L'angolo azimutale  $\phi$  è misurato dall'asse x positivo nel piano xy e l'angolo polare  $\theta$  è definito nel piano rz ma solitamente è espresso attraverso la pseudorapidità  $\eta$ = -ln(tan $\frac{\theta}{2}$ ). La coordinata radiale r indica la distanza dall'asse z. Il momento trasverso  $p_T$  è definito come la componente del momento delle particelle sul piano trasverso xy.

#### 1.2.2 Il tracker

Il tracker è un dispositivo che ha lo scopo di ricostruire le traiettorie delle particelle cariche coprendo angoli fino a  $|\eta| = 2.5$  [5]. In CMS tale rivelatore può essere diviso in due sub-detectors:

- un rivelatore a pixel in silicio, posto vicino al vertice di interazione. E' formato da 3 strati cilindrici nel Barrel, nella regione 4.4 cm < r < 10.2 cm, e da 2 dischi a ogni estremità.
- uno strip detector in silicio, posto nella regione più esterna del tracker. Nel Barrel sono presenti 4 strati cilindrici dell'Inner Barrel (TIB) e 6 strati dell'Outer Barrel (TOB), mentre gli Endcaps sono formati da 9 layers ciascuno.

#### 1.2.3 Il calorimetro elettromagnetico ECAL

Il calorimetro elettromagnetico ha lo scopo di misurare l'energia di elettroni e fotoni. In CMS il calorimetro elettromagnetico ECAL è un calorimetro omogeneo, costituito da cristalli di tungstanato di piombo PbWO<sub>4</sub>. ECAL è composto da tre parti principali:

- il Barrel EB (regione con  $|\eta| < 1.48$ )
- l'Endcap EE (regione con  $1.48 < |\eta| < 3.0)$
- il Pre-shower (posto di fronte agli endacaps, regione con  $1.653 < |\eta| < 2.6$ )

Nel Barrel il segnale viene letto attraverso fotodiodi a valanga (Avalanche Photodiodes APD) mentre nell'Endcap si utilizzano fototriodi a vuoto, più resistenti alla radiazione rispetto agli APD.

La risoluzione energetica del calorimetro ECAL è parametrizzata da tre diversi contibuti:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{1.1}$$

dove il primo termine è di natura statistica (a=1.8%) il secondo tiene conto del rumore elettronico e del pileup (b=4%) e il terzo termine è legato alla calibrazione (c=0.5%).

#### 1.2.4 Il calorimetro adronico HCAL

Il calorimetro adronico HCAL [6] è un calorimetro di tipo sampling, ovvero un calorimetro che alterna strati di materiale assorbitore (ottone) a strati di materiale attivo (scintillatore).

HCAL è composto da tre parti principali (Fig. 1.4):

- il Barrel HB (regione con  $|\eta| < 1.3$ )
- l'Endcap HE (regione con  $1.3 < |\eta| < 3.0$ )

• il Forward HF (regione con  $3.0 < |\eta| < 5.0$ )

e da una quarta porzione di calorimetro posta oltre il primo strato assorbitore di muoni, l'Outer Hadronic Calorimeter (HO).

Il Barrel e l'Endcap alternano strati di ottone (spesso 5 cm in HB e 8 cm in HE) a strati di scintillatore plastico spesso 4 cm, mentre nel Forward il materiale attivo è costituito da fibre di quarzo longitudinali spaziate di 5 mm. La luce di scintillazione viene raccolta utilizzando un wavelenght shifter (WLS) incorporato in una scanalatura nello scintillatore. Fuori dallo scintillatore la fibra di WLS è connessa ad una fibra ottica che trasporta il segnale al dispositivo di lettura.

La risoluzione energetica del calorimetro HCAL (espressa in GeV) è

$$\frac{\sigma_E}{E} \sim \frac{a}{\sqrt{E}} \oplus 5\% \tag{1.2}$$

dove a=65% nel Barrel, a=85% nell'Endcap e a=100% nel Forward.



Figura 1.4: Sezione del detector CMS. In blu sono indicate le tre componenti principali del calorimetro adronico HB, HE e HF

#### 1.2.5 Il rivelatore di muoni

Il rivelatore di muoni [7] è stato progettato per svolgere tre funzioni principali: identificazione dei muoni, misura del momento e triggering dei muoni. Il sistema è immerso in un campo magnetico di verso opposto rispetto a quello generato dal solenoide, in modo da permettere una misura del momento indipendente da quella del tracker. Il rivelatore di muoni è costituito da tre diverse tipologie di rivelatori a gas:

• Drift Tubes (DT): sono utilizzati nel Barrel ( $|\eta| < 1.2$ ) dove il campo magnetico residuo e il flusso di particelle sono bassi. Quattro strati di Drift Tube Chambers sono disposti su 5 ruote lungo l'asse *z*, ognuna divisa in 12  $\phi$ -sezioni, dove  $\phi$  è l'angolo azimutale. Ogni camera ha una risoluzione di circa 100  $\mu$ m in r $\phi$  e fino a 150  $\mu$ m in *z* e può misurare la direzione della particella con un accuratezza di ~ 1 mrad.

- Cathode Strip Chambers (CSC): sono utilizzate negli Endcaps ( $0.8 < |\eta| < 2.4$ ) dove il rate di particelle è maggiore e il campo magnetico non è uniforme. Le camere sono disposte su 4 dischi per ogni endcap, divisi in anelli concentrici. Ogni camera ha una risoluzione spaziale di circa 200 mm in r e 75x150 mm nella coorditata r $\phi$ .
- **Resistive Plate Chambers (RPC)**: 912 RPC sono posizionate sia nel Barrel che nell'Endcap. Le RPC hanno una risoluzione spaziale peggiore di quella di DT e CSC ma forniscono una risposta veloce e con una buona risoluzione temporale (1 ns), pertanto sono utili per il trigger.

#### 1.2.6 Il trigger

Alla luminosità prevista dal progetto di LHC il rate delle interazioni protone-protone è maggiore di 1 GHz. Solo una piccola frazione di queste collisioni contiene eventi di interesse per l'esperimento CMS e solo una piccola frazione di queste può essere immagazinata per l'analisi offline. Il compito del trigger è quello di selezionare gli eventi interessanti da archiviare offline.

Per selezionare gli eventi di interesse, CMS utilizza due livelli di trigger:

- L1: è il primo livello del trigger di CMS. Utilizza le informazioni provenienti dai calorimetri e dal rivelatore di muoni per decidere ento 4 μs se accettare o rigettare l'evento, abbassando il rate a ~100kHz [8].
- HLT: il High Level Trigger è il secondo livello. Per ogni evento vengono ricostruiti alcuni oggetti come ad esempio elettroni, muoni e jets e vengono applicati dei criteri di identificazione al fine di selezionare solo gli eventi di possibile interesse per l'analisi dei dati. Il rate di output viene abbassato a ~ 400 Hz [8].

#### 1.2.7 CMS Phase 2 upgrade

Durante lo shut-down di LHC che avverrà nel 2024, per far fronte all'elevata luminosità e al conseguente aumento del pileup previsti con HL-LHC (Sezione 1.1.1), è previsto un upgrade dell'esperimento CMS, denominato Phase 2 [9]. Il progetto prevede la completa sostituzione di molte componenti dell'esperimento tra cui il tracker, il calorimetro adronico nella regione forward e il sistema di trigger. Il nuovo tracker presenterà una maggiore granularità ed ermeticità. Verranno infatti installati moduli a pixel e strip di silicio in un maggior numero di strati lungo l'asse del fascio e in dodici dischi posti longitudinalmente alla linea di fascio come mostrato in Fig. 1.5. Ciò permetterà la tracciatura di particelle cariche in un range  $|\eta| < 4$  con conseguente aumento dell'accettanza del rivelatore. Le potenzialità di CMS nel contesto di HL-LHC sono state finora studiate effettuando proiezioni dei risultati ottenuti con analisi dei dati raccolti nel 2015 a  $\sqrt{s} = 13$  TeV [10]. Studi più accurati basati su simulazioni dell'esperimento nella sua configurazione finale sono attualmente in corso. Nella Sezione 4.4 verranno riportati i risultati relativi all'impatto del nuovo tracker sulla raccolta di segnale proveniente dalla produzione di coppie di bosoni di Higgs.



Figura 1.5: Schema del tracker previsto per il CMS Phase 2 upgrade; è mostrato un quarto di sezione longitudinale; sono evidenziate le diverse tipologie di rivelatore al silicio: strip (rosso), strip e pixel (blu), pixel (giallo e verde). [9].

## **Capitolo 2**

# Produzione di coppie di bosoni di Higgs

#### 2.1 Produzione di coppie HH

La scoperta di una particella con massa 125 GeV, con caratteristiche coincidenti, entro l'attuale precisione delle misure sperimentali, con il bosone di Higgs previsto dal modello standard (SM) [11], ha aperto nuovi scenari per ricerche con collisioni protone-protone a LHC. In particolare lo studio della produzione di coppie di bosoni di Higgs (di-Higgs) rappresenta il miglior canale per studiare il potenziale del campo di Higgs.

In un collider adronico come LHC, la produzione di coppie di bosoni di Higgs avviene principalmente, secondo il SM, attraverso la fusione gluone-gluone (Fig. 2.1).



Figura 2.1: Diagrammi di Feynman della produzione di coppie di bosoni di Higgs attraverso la fusione gluone-gluone

Questo processo ha una sezione d'urto prevista dal SM molto bassa rispetto al background pari a 33.53 + 4.3% - 6.0% fb in collisioni p-p a  $\sqrt{s} = 13$  TeV [12]. Tale valore è di diversi ordini inferiore rispetto ai processi di background e questo ne rende difficile la rivelazione. Pertanto, per osservare questo processo raro, è necessaria una grande luminosità integrata; alla luminosità integrata di 3000 fb<sup>-1</sup> (che si prevede di raggiungere con il programma HL-LHC (Sezione 1.1.1), ci si aspetta che la produzione di coppie HH potrà essere osservata con una significatività di 1.9 deviazioni standard.

Nella Figura 2.2 sono riportati i possibili stati finali della produzione di coppie HH e le relative probabilità di decadimento (branching ratio). In particolare il processo  $HH \rightarrow b\bar{b}b\bar{b}$ è quello con il maggior branching ratio, ma è affetto da un background irriducibile dovuto principalmente a processi di QCD.



Figura 2.2: Probabilità di decadimento di coppie di bosoni di Higgs.

# 2.2 Stato delle ricerche della produzione di coppie di bosoni di Higgs

I risultati della ricerca di coppie HH, utilizzando i dati raccolti da CMS nel Run II di LHC, non hanno individuato variazioni rispetto alle previsioni del SM e hanno permesso di impostare un limite superiore nella sezione d'urto del processo. Nella Tabella 2.1 sono riportati i valori  $\frac{\sigma}{\sigma_{SM}}$  ottenuti, dove  $\sigma$  è il valore ottenuto del limite superiore della sezione d'urto del processo e  $\sigma_{SM}$  è la sezione d'urto prevista dal Modello Standard [13] [14] [15] [16]. I risultati ottenuti dallo studio del processo HH  $\rightarrow$  bbbb verranno aggiornati con un'analisi estesa a tutti i dati raccolti nel 2016 (35.9 fb<sup>-1</sup>).

HH to	$L_{int}(\mathrm{fb}^{-1})$	Observed (expected) $\frac{\sigma}{\sigma_{SM}}$ 95% CL limits (fb)
bbbb	2.3	342 (308)
$b\bar{b}l\nu l\nu$	35.9	79 (89)
$b\bar{b} au au$	35.9	28 (25)
${ m b}ar{ m b}\gamma\gamma$	35.9	19 (17)

Tabella 2.1: Valori osservati (e valori attesi) del rapporto $\frac{\sigma}{\sigma_{SM}}$ 

## **Capitolo 3**

# Jets

#### 3.1 Adronizzazione

L'adronizzazione è un processo non-perturbativo attraverso il quale i gluoni, i quark e gli antiquark producono jet adronici. Analogamente a una particella carica accelerata che irradia un fotone, anche i quark, dotati di carica di colore, emettono un gluone, ma con una probabilità più alta dovuta alla maggiore costante di accoppiamento [17]. I gluoni a loro volta possono suddividersi in coppie di gluoni o produrre coppie quark-antiquark che irradiano altri gluoni e così via, generando una cascata. Con lo svilupparsi della cascata l'energia dei suoi componenti diminuisce. Ad un'energia dell'ordine del GeV, i partoni risultanti si ricombinano in adroni con carica di colore nulla.

#### 3.2 Ricostruzione dei jets

La ricostruzione di un jet consiste nell'identificare tutte le tracce relative a un singolo jet e il vertice da cui ha avuto inizio il processo di adronizzazione. Una prima ricostruzione dei jets è basata sulle informazioni provenienti dalle towers dei calorimetri elettromagnetico e adronico. I jets così ricostruiti sono chiamati *CaloJets* e vengono principalmente utilizzati dal trigger, grazie alla risposta veloce dei calorimetri. Una migliore ricostruzione dei jets è basata sulla tecnica del Particle Flow (PF) [18], che ricostruisce e identifica ogni particella combinando le informazioni dei vari sub-detectors di CMS. Un esempio di algoritmo basato sul PF è l'algoritmo di clustering  $anti - k_T$ , che unisce coppie di particelle in un jet in ordine crescete di  $p_T$  relativo. Ciò avviene finché la distanza

$$d_{ij} = min(\frac{1}{p_{T_i}^2}, \frac{1}{p_{T_j}^2})\frac{\Delta R_{ij}^2}{R^2}$$
(3.1)

 $con \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$  non è maggiore di un certo valore (Fig. 3.1). Il momento del jet ricostruito viene determinato attraverso la somma vettoriale dei momenti delle particelle ad esso associate, mentre l'energia viene ricostruita tenendo conto delle correzioni necessarie a causa delle limitazioni intrinseche del sistema (risposta non lineare dei calorimetri, segmentazione del detector, presenza di materiale prima dei calorimetri, rumore dovuto all'elettronica e al pileup).



Figura 3.1: Evento con 4 jets registrato da CMS [19]. I jets sono stati ricostruiti utilizzando l'algoritmo di clustering  $anti - k_T \operatorname{con} \Delta R < 0.4$  (Sezione 3.2).

#### 3.3 b-tagging dei jet

Per lo studio del processo di decadimento del bosone di Higgs in una coppia  $b\bar{b}$  è importante identificare i jets originati da quark b, detti *b-jets*. L'identificazione dei b-jets è basata sulla vita media relativamente lunga del quark b ( $c\tau \sim 450 \ \mu s$ ) che permette alle particelle (B-hadrons) contenenti al loro interno un quark b di avere un parametro d'impatto considerevole rispetto al punto di origine del quark (Fig. 3.2). L'algoritmo di b-tagging utilizzato in questo lavoro di tesi è il **cMVAv2** (combined MultiVariate Algorithm version 2) [20], il quale combina gli output di diversi algoritmi di identificazione di b-jets attraverso un Boosted Decision Tree (BDT). Gli algoritmi di identificazione in questione sono il CSVv2, una variante del CSVv2 che utilizza una diversa ricostruzione dei vertici, il Jet Probability (JP), il Jet B Probability (JBP), il Soft Electron (SE) e il Soft Muon (SM) [20].



Figura 3.2: Schema del decadimento di un B-hadron e relativo b-jet

## **Capitolo** 4

# Analisi

#### 4.1 Obiettivi e strategia di analisi

Questa tesi ha come obiettivo la ricostruzione della massa del bosone di Higgs nel processo  $HH \rightarrow b\bar{b}b\bar{b}$  da eventi simulati e lo studio dell'efficienza e della corretta selezione dei jets di eventi simulati ai pileup previsti per HL-LHC. L'analisi si articola nel modo seguente:

- Selezione degli eventi imponendo opportuni tagli sulle variabili cinematiche e sul numero di jets per evento
- Selezione di 4 b-jets
- Associazione dei jets ricostrutiti con i quark b generati
- Ricostruzione della massa dei bosoni di Higgs

La strategia di analisi presentata verrà applicata ad un campione di eventi simulati a un'energia nel centro di massa  $\sqrt{s} = 13$  TeV e in una condizione di pileup PU=20, valore tipico degli eventi registrati nel corso del Run 2 di LHC (Fig. 4.1).

Successivamente la selezione degli eventi e il pairing dei jets ricostruiti con i b-quarks generati verranno effettuati su 3 campioni di dati simulati in diverse condizioni di pileup in modo da valutare l'efficienza e la frazione di b-jets correttamente selezionati degli eventi che verranno registrati alla luminosità prevista per HL-LHC.

#### 4.2 Simulazioni

La produzione di coppie di bosoni di Higgs non risonanti è stata simulata utilizzando come generatore Monte Carlo il *MG5 aMC@NLO*. Il decadimento dei due Higgs ciascuno in una coppia  $b\bar{b}$  e l'adronizzazione sono stati simulati attraverso il programma *Pythia8* [21], mentre la risposta del detector CMS è stata modellizzata utilizzando *GEANT4* [22]. E' stato simulato un campione di eventi in una condizione di pileup 20 a un'energia nel centro di massa  $\sqrt{s} = 13$  TeV. Sono stati simulati tre campioni di eventi a un' energia nel centro di massa  $\sqrt{s} = 14$  TeV, in tre diverse condizioni di pileup: PU=0, PU=140 e PU= 200. Tali eventi sono stati simulati modellizzando la risposta del detector CMS come previsto dall'upgrade Phase 2 descritto nella sezione 1.2.7.



Figura 4.1: Pileup medio registrato da CMS nel 2016 [23]

#### Analisi degli eventi simulati a $\sqrt{s} = 13$ TeV e pileup 20 4.3

#### 4.3.1 Selezione degli eventi

Sono stati effettuati i tagli  $p_T$  >30 GeV e  $|\eta|$  <2.4 su ciascun jet di ogni evento e sono stati selezionati gli eventi con almeno 4 jets. In questo modo dei 299789 eventi iniziali sono stati selezionati 138327 eventi. L'efficienza di selezione, ovvero il rapporto tra il numero di eventi selezionati e il numero di eventi totale  $N_{tot}$  è

$$\varepsilon_1 = 0.4614 \pm 0.0009 = 46.14\% \pm 0.09\% \tag{4.1}$$

 $\operatorname{con} \sigma_{\varepsilon_1} = \sqrt{\frac{\varepsilon_1(1-\varepsilon_1)}{N_{tot}}}$  [24]. Le figure 4.2 e 4.3 presentano le variabili cinematiche e il cMVA dei jets dopo questa prima selezione.



Figura 4.2:  $p_T$  e  $\eta$  dei jets degli eventi selezionati



Figura 4.3: cMVA dei jets degli eventi selezionati

#### Selezione di 4 b-jets e accoppiamento in di-jets

E' stata effettuata un'ulteriore selezione richiedendo che gli eventi avessero almeno 4 jets "b-taggati", ovvero con cMVA> 0.4432. Sono stati selezionati 8812 eventi, con un'efficienza di selezione

$$\varepsilon_2 = 0.0293 \pm 0.0003 = 2.93\% \pm 0.03\% \tag{4.2}$$

 $con \sigma_{\varepsilon_2} = \sqrt{\frac{\varepsilon_2(1-\varepsilon_2)}{N_{tot}}} [24].$ 

I jets di ogni evento sono stati ordinati in ordine decrescente di cMVA e sono stati selezionati i 4 jets con cMVA maggiore, ovvero quelli con probalità maggiore di essere jets originati da quark b. Nelle figure 4.4 e 4.5 sono rappresentati il cMVA e le variabili  $p_T$  e  $\eta$ dei 4 jets selezionati per ogni evento.



Figura 4.4: cMVA dei primi 4 jets in cMVA degli eventi selezionati

I quattro jets selezionati sono stati accoppiati in 2 di-jets richiendendo che la differenza di massa dei due bosoni di Higgs ricostruiti fosse minima ( $\Delta M = M_{H_2} - M_{H_1}$  minimo). Per ogni coppia di bosoni di Higgs ricostruiti, è stato convenzionalmente chiamato H<sub>1</sub> il bosone con massa minore e H<sub>2</sub> quello con massa maggiore (Fig. 4.6).



(a)  $p_{\scriptscriptstyle T}$  dei primi 4 jets in cMVA degli eventi selezionati



(b) $\eta$ dei primi 4 jets in cMVA degli eventi selezionati

Figura 4.5:  $p_{\scriptscriptstyle T}$ e $\eta$ dei primi 4 jets in cMVA degli eventi selezionati



Figura 4.6: Massa invariante dei candidati bosone di Higgs,  $H_1$  (sinistra) e  $H_2$  (destra), ricostruiti dagli eventi selezionati

#### 4.3.2 Pairing dei b-jets selezionati con i quark b generati

Nelle figure 4.7 e 4.8 sono riportati il momento trasverso  $p_T$ , la pseudorapidità  $\eta$  e l'angolo azimutale  $\phi$  dei quattro quark b generati per ogni evento simulato. Tali quark b rappresentano la "verità Monte Carlo", cioè i quark, provenienti dal decadimento dei due bosoni di Higgs, da cui sono generati i quattro jets finali. Uno studio della distanza nel piano  $\phi$ - $\eta$  tra i jets ricostruiti e i quark b generati permette di valutare la correttezza della selezione finale dei jets.



Figura 4.7:  $p_T$  (sinista) e  $\eta$  (destra) dei b-quark degli eventi selezionati



Figura 4.8:  $\phi$  dei b-quark degli eventi selezionati

A ogni quark b è stato associato il jet più vicino in  $\eta \in \phi$  minimizzando la distanza angolare  $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2}$ , con  $\Delta \phi = \pi - ||\phi_1 - \phi_2| - \pi| \in \Delta \eta = |\eta_1 - \eta_2|$ . Per ogni evento è stata calcolata la media quadratica  $\Delta R^*$  attraverso la formula

$$\Delta \mathbf{R}^* = \frac{\sqrt{\sum\limits_{k=1}^4 \Delta \mathbf{R}_k^2}}{4} \tag{4.3}$$

La distribuzione di  $\Delta R^*$  (Fig. 4.9) presenta un picco a piccoli  $\Delta R^*$  (pairing ottimo) e una coda a grandi  $\Delta R^*$  (pairing pessimo).



Figura 4.9: distribuzione di  $\Delta R^*$ 

Per valutare la bontà del pairing è necessario fissare un valore di soglia  $\Delta R_{cut}^*$  al di sotto del quale i jets degli eventi vengano considerati correttamente accoppiati ai b-quarks. Per scegliere tale valore di soglia è stata eseguita la ricostruzione della massa dei bosoni di Higgs impostando  $\Delta R_{cut}^* = 0.3$  e  $\Delta R_{cut}^* = 0.4$ , in modo da valutare la presenza di qualche differenza nella distribuzione della massa e scegliere di conseguenza il taglio più opportuno. Nelle figure 4.10 e 4.11 sono riportati gli istogrammi della massa di H<sub>1</sub> e H<sub>2</sub> ricostruita per gli eventi con  $\Delta R^* < \Delta R_{cut}^*$  e per gli eventi  $\Delta R^* > \Delta R_{cut}^*$ . L'area degli istogrammi è stata normalizzata a 1 per permettere il confronto delle distribuzioni relative ai due diversi valori di  $\Delta R_{cut}^*$ .



Figura 4.10: Massa di H<sub>1</sub> (sinistra) e H<sub>2</sub> (destra) ricostruita per eventi con  $\Delta R^* < \Delta R^*_{cut}$ 



Figura 4.11: Massa di H<sub>1</sub> (sinistra) e H<sub>2</sub> (destra) ricostruita per eventi con  $\Delta R^* > \Delta R^*_{cut}$ 

La forma della distribuzione della massa degli Higgs per  $\Delta R_{cut}^* = 0.3$  e quella per  $\Delta R_{cut}^* = 0.4$ non presentano evidenti differenze. Pertanto, per preservare un maggior numero di eventi (Tabella 4.1), è stato impostato il valore di soglia a  $\Delta R_{cut}^* = 0.4$ .

$\Delta R^*_{cut}$	Eventi con $\Delta R^* < \Delta R^*_{cut}$
0.3	6932
0.4	7010

Tabella 4.1: Numero di eventi con  $\Delta R^* < \Delta R^*_{cut}$  per i due diversi valori di  $\Delta R^*_{cut}$ 

#### 4.3.3 Ricostruzione della massa dei bosoni di Higgs

Nelle figure 4.12 e 4.13 sono riportati gli istogrammi della massa di  $H_1$  e  $H_2$  ricostruita dai jets selezionati e accoppiati come descritto nella Sezione 4.3.1.



Figura 4.12: Massa di H<sub>1</sub>. In rosso è rappresentato l'istrogramma della massa di H<sub>1</sub> ricostruita per gli eventi selezionati con  $\Delta R^* > 0.4$ , in blu la massa di H<sub>1</sub> ricostruita per eventi selezionati con  $\Delta R^* < 0.4$  e in verde è rappresentata la massa di H<sub>1</sub> ricostruita utilizzando tutti gli eventi selezionati.



Figura 4.13: Massa di H<sub>2</sub>. In rosso è rappresentato l'istrogramma della massa di H<sub>2</sub> ricostruita per gli eventi selezionati con  $\Delta R^* > 0.4$ , in blu la massa di H<sub>2</sub> ricostruita per eventi selezionati con  $\Delta R^* < 0.4$  e in verde è rappresentata la massa di H<sub>2</sub> ricostruita utilizzando tutti gli eventi selezionati.

Per stimare il valore medio della distribuzione, è stato eseguito un fit gaussiano del picco della distribuzione della massa di tutti i bosoni di Higgs ricostruiti a partire dagli eventi selezionati (Fig. 4.14).



Figura 4.14: Fit gaussiano della distribuione della massa degli Higgs ricostruiti

La stima della massa del bosone di Higgs ottenuta dal fit è  $m_{\rm H} = 116.3 \pm 0.4 \frac{\text{GeV}}{c^2}$ . Il valore ricostruito è inferiore al valore nominale della massa del bosone di Higgs (125 GeV) a causa della mancanza di una precisa calibrazione energetica dei jets originati da b-quarks, le cui caratteristiche (in particolare la frequente presenza di neutrini al loro interno) ne rendono diversa la ricostruzione. Metodologie di regressione delle variabili di ogni jets sono normalmente implementate nelle analisi di CMS per ricalibrare i jets sulla base del confronto con i campioni MC.

#### 4.4 Analisi degli eventi simulati per HL-LHC

In questa sezione sono presentati i risultati dell'analisi sull'efficienza di selezione e la frazione di jets correttamente selezionati degli eventi simulati per HL-LHC. I campioni di eventi a diversi pileup sono stati simulati con collisioni p-p a un'energia nel centro di massa  $\sqrt{s} = 14$  TeV, modellizzando il rivelatore come previsto dall'upgrade Phase 2 di CMS (Sezione 1.2.7).

Lo studio è stato effettuato selezionando gli eventi come descritto in 4.3.1 per diversi valori di taglio delle variabili  $p_T$  e  $\eta$ . In particolare per il momento trasverso  $p_T$  è stato impostato il taglio a 30 GeV e a 50 GeV, mentre il taglio sulla variabile  $\eta$  è stato impostato a 2.4, 3.0 e 4.0.

Per ogni set di valori di taglio sono state eseguite le seguenti operazioni:

- è stato registrato il numero  $N_1$  di eventi selezionati imponendo il taglio in  $p_T$  e  $\eta$  e la presenza di almeno 4 jets. E' stata quindi calcolata l'efficienza di selezione  $\varepsilon_1 = \frac{N_1}{N_{tot}}$ , con  $N_{tot}$  numero totale degli eventi del campione prima della selezione. L'errore  $\sigma_{\varepsilon_1}$  è stato calcolato attraverso la formula  $\sigma_{\varepsilon_1} = \sqrt{\frac{\varepsilon_1(1-\varepsilon_1)}{N_{tot}}}$  [24].
- è stato registrato il numero  $N_2$  di eventi selezionati imponendo la presenza di almeno 4 jets b-taggati, ovvero con cMVA> 0.4432. E' stata quindi calcolata l'efficienza di selezione  $\varepsilon_2 = \frac{N_2}{N_{tot}}$ , con  $N_{tot}$  numero totale degli eventi del campione prima della selezione. L'errore  $\sigma_{\varepsilon_2}$  è stato calcolato attraverso la formula  $\sigma_{\varepsilon_2} = \sqrt{\frac{\varepsilon_2(1-\varepsilon_2)}{N_{tot}}}$  [24].
- per gli  $N_2$  eventi selezionati è stato eseguito il pairing dei primi 4 jets (in ordine di cMVA decrescente) con i b-quarks generati come descritto in 4.3.2. E' stato registrato il numero  $N_B$  di eventi con  $\Delta R^* < 0.4$  ed è stata calcolata la frazione p di eventi selezionati i cui jets sono associati correttamemte:  $p = \frac{N_B}{N_2}$ . L'errore  $\sigma_p$  è stato calcolato attraverso la formula  $\sigma_p = \sqrt{\frac{p(1-p)}{N_2}}$  [24].

#### 4.4.1 Campione di eventi con pileup PU=0

Nella Tabella 4.2 sono riportati i risultati della selezione degli  $N_{tot} = 291904$  eventi simulati a pileup 0 per i diversi tagli in  $p_T$  e  $\eta$ .

	$N_1$	$N_2$	$N_B$	$\varepsilon_1$ (%)	$arepsilon_2$ (%)	p (%)
$p_T > 30 \text{ GeV}  \eta  < 2.4$	190620	40697	33420	$65.30\pm0.08$	$13.94\pm0.06$	$82.1\pm0.1$
$p_{_T} > 30 \text{ GeV}  \eta  < 3.0$	208286	44202	36024	$71.35\pm0.08$	$15.14\pm0.06$	$81.4\pm0.1$
$p_T > 30 \text{ GeV}  \eta  < 4.0$	208320	44205	36025	$71.36\pm0.08$	$15.14\pm0.06$	$81.4\pm0.1$
$p_T > 50 \text{ GeV}  \eta  < 2.4$	107322	16931	13267	$36.76\pm0.08$	$5.80\pm0.04$	$78.3\pm0.3$
$p_{_T} > 50 \text{ GeV}  \eta  < 3.0$	120398	18113	14047	$41.24\pm0.09$	$6.20\pm0.04$	$77.5\pm0.3$
$p_{\scriptscriptstyle T} > 50 \ {\rm GeV} \  \eta  < 4.0$	120450	18115	14048	$41.26\pm0.08$	$6.20\pm0.04$	$77.5\pm0.3$

#### 4.4.2 Campione di eventi con pileup PU=140

Nella Tabella 4.3 sono riportati i risultati della selezione degli  $N_{tot} = 296662$  eventi simulati a pileup 140 per i diversi tagli in  $p_T$  e  $\eta$ .

	$N_1$	$N_2$	$N_B$	$\varepsilon_1$ (%)	$\varepsilon_2$ (%)	<i>p</i> (%)
$p_{_T} > 30 \text{ GeV}  \eta  < 2.4$	207994	50983	38345	$70.11\pm0.08$	$17.18\pm0.06$	$75.2\pm0.1$
$p_{_T} > 30 \text{ GeV}  \eta  < 3.0$	219992	54314	40423	$74.15\pm0.08$	$18.30\pm0.07$	$74.4\pm0.1$
$p_{_T} > 30 \text{ GeV}  \eta  < 4.0$	219992	54314	40423	$74.15\pm0.08$	$18.30\pm0.07$	$74.4\pm0.1$
$p_{\scriptscriptstyle T} > 50 \ {\rm GeV} \  \eta  < 2.4$	121755	22493	16261	$41.04\pm0.09$	$7.58\pm0.04$	$72.2\pm0.2$
$p_{\scriptscriptstyle T} > 50~{\rm GeV} ~ \eta  < 3.0$	131844	23739	16952	$44.44\pm0.09$	$8.00\pm0.04$	$71.4 \pm 0.2$
$p_{\scriptscriptstyle T} > 50 \ {\rm GeV} \  \eta  < 4.0$	131844	23739	16952	$44.44\pm0.09$	$8.00\pm0.04$	$71.4\pm0.2$

Tabella 4.3

#### 4.4.3 Campione di eventi con pileup PU=200

Nella Tabella 4.4 sono riportati i risultati della selezione degli  $N_{tot} = 297558$  eventi simulati a pileup 200 per i diversi tagli in  $p_T$  e  $\eta$ .

	$N_1$	$N_2$	$N_B$	$\varepsilon_1$ (%)	$arepsilon_2$ (%)	p (%)
$p_{_T} > 30 \text{ GeV}  \eta  < 2.4$	194754	50303	36240	$65.45\pm0.08$	$16.90\pm0.06$	$72.0\pm0.2$
$p_{\scriptscriptstyle T} > 30 \ {\rm GeV} \  \eta  < 3.0$	205520	53392	38019	$69.06\pm0.08$	$17.94\pm0.07$	$71.2\pm0.1$
$p_{\scriptscriptstyle T} > 30 \ {\rm GeV} \  \eta  < 4.0$	205520	53392	38019	$69.06\pm0.08$	$17.94\pm0.07$	$71.2\pm0.1$
$p_{\scriptscriptstyle T} > 50~{\rm GeV} ~ \eta  < 2.4$	113357	21888	15088	$38.09\pm0.08$	$7.35\pm0.04$	$68.9\pm0.3$
$p_{\scriptscriptstyle T} > 50~{\rm GeV} ~ \eta  < 3.0$	121772	22979	15662	$40.92\pm0.09$	$7.72\pm0.04$	$68.1\pm0.3$
$p_{\scriptscriptstyle T} > 50 \ {\rm GeV} \  \eta  < 4.0$	121772	22979	15662	$40.92\pm0.09$	$7.72\pm0.04$	$68.1\pm0.3$

Tal	bell	la ·	4.4
-----	------	------	-----

Dalle figure 4.15, 4.16 e 4.17 è possibile notare che i valori di  $\varepsilon_1$ ,  $\varepsilon_2$  e p per  $|\eta| < 3.0$  e  $|\eta| < 4.0$  sono gli stessi, indipendentemente dal  $p_T$ . Questo avviene perché l'algoritmo di b-tagging non è efficiente per alti valori della pseudorapidità.



Figura 4.15: Grafico dell' efficienza di selezione  $\varepsilon_1$ 



Figura 4.16: Grafico dell' efficienza di selezione  $\varepsilon_2$ 



Figura 4.17: Grafico della frazione di b-jets correttamente selezionati

## **Capitolo 5**

# Conclusioni

La massa dei bosoni di Higgs è stata ricostruita per gli eventi simulati a  $\sqrt{s} = 13$  TeV e pileup 20. Nel corso dell'analisi è stato studiato l'impatto sulla ricostruzione della massa della scelta del valore di soglia  $\Delta R^*_{cut}$  (con  $\Delta R^*$  valore medio della distanza angolare tra jets ricostruiti e b-quarks generati di un evento), ottenendo come valore ottimale  $\Delta R^*_{cut} = 0.4$ . Dal fit gaussiano del picco della distribuzione della massa dei bosoni di Higgs ricostruiti è stata ottenuta la stima  $m_H = 116.3 \pm 0.4 \frac{\text{GeV}}{\text{c}^2}$ .

Dallo studio effettuato sui campioni di eventi simulati per HL-LHC, risulta che sia le efficienze  $\varepsilon_1$  e  $\varepsilon_2$  che la frazione di b-jets correttamente selezionati diminuiscono sensibilmente impostando il taglio sul momento trasverso  $p_{\tau}$  a 50 GeV. Variando il valore di taglio della variabile  $\eta$  da 2.4 a 3.0 vengono selezionati più eventi ma la frazione di bjets correttamente selezionati diminuisce, mentre variandolo da 3.0 a 4.0 i valori di  $\varepsilon_1$ ,  $\varepsilon_2 \in p$  risultano stabili. Con l'aumentare del pileup aumenta il numero di jets per evento, pertanto ci si aspetta che l'efficienza di selezione cresca e che la frazione di b-jets correttamente selezionati diminuisca. Infatti, con più tracce e più jets per ogni evento, la probabilità di avere 4 jets che passino la selezione è maggiore ma, allo stesso tempo, aumenta anche la probabilità di selezionare jets che non coincidano con quelli originati dai b-quarks. Tali ipotesi sono confermate dall'analisi degli eventi simulati in una condizione di pileup pari a 140. Dall'analisi degli eventi simulati per il PU=200 risulta una diminuzione dell'efficienza di selezione; le ragioni di tale diminuzione sono tuttora oggetto di studio. L'andamento della frazione di jets correttamente selezionati, invece, è quello atteso: p diminuisce all'aumentare del pileup. Tale diminuzione è contenuta e ci si aspetta di poter migliorare la qualità della selezione con l'utilizzo di nuovi algoritmi per la ricostruzione dei jets e per il b-tagging.

# Bibliografia

- P. Lefèvre et al., "The Large Hadron Collider: Conceptual Design", CERN-AC-95-05-LHC
- [2] G. Apollinari et al., "High Luminosity Large Hadron Collider HL-LHC", ar-Xiv:1705.08830
- [3] The Atlas Collaboration, *Expected pileup values at the HL-LHC*, ATL-UPGRADE-PUB-2013-014, url: https://cds.cern.ch/record/1604492/
- [4] The CMS Collaboration, "The CMS experiment at the CERN LHC", 2008 JINST 3 S08004
- [5] The CMS Collaboration, The CMS tracker system project : Technical Design Report, CERN-LHCC-98-006; CMS-TDR-5
- [6] The CMS Collaboration, "The CMS hadron calorimeter project : Technical Design Report" CERN-LHCC-97-031; CMS-TDR-2
- [7] The CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV at the LHC", arXiv:1306.6905
- [8] The CMS Collaboration, "The CMS trigger system, arXiv:1609.02366
- [9] Contardo et al., Proposal for the Phase-II Upgrade of the CMS Detector, CERN-LHCC-2015-010, url: https://cds.cern.ch/record/2020886
- [10] The CMS Collaboration Projected performance of Higgs analyses at the HL-LHC for ECFA 2016, CMS-PAS-FTR-16-002, url: http://cds.cern.ch/record/2266165
- [11] The CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, arXiv:1207.7235 [hep-ex]
- [12] D. de Florian et al., Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, arXiv:1610.07922 [hep-ph]
- [13] The CMS Collaboration, Search for non-resonant pair production of Higgs bosons in the bbbb final state with 13 TeV CMS data, CMS-PAS-HIG-16-026, url: https://cds.cern.ch/record/2209572
- [14] The CMS Collaboration, Search for resonant and non-resonant Higgs boson pair production in the  $b\bar{b}l\nu l\nu$  final state at  $\sqrt{s} = 13$  TeV, CMS-PAS-HIG-17-006, url: https://cds.cern.ch/record/2257068

- [15] The CMS Collaboration, "Search for pair production of Higgs bosons in the two tau leptons and two bottom quarks final state using proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", CMS-PAS-HIG-17-002, url: https://cds.cern.ch/record/2256096
- [16] The CMS Collaboration, "Search for Higgs boson pair production in the final state containing two photons and two bottom quarks in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", CMS-PAS-HIG-17-008, url: http://cds.cern.ch/record/2273383
- [17] Alessandro Bettini, "Introduction to elementary particle physics", Cambridge University Press
- [18] The CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_T^{miss}$ "
- [19] Skovpen, Kirill, "Event displays highlighting the main properties of heavy flavour jets in the CMS Experiment", CMS-PHO-EVENTS-2017-006, url: https://cds.cern.ch/record/2280025
- [20] The CMS Collaboration, "Performance of b-Tagging Algorithms in Proton Collisions at 13 TeV using the 2016 Data", url: https://cds.cern.ch/record/2202967
- [21] S. M. T. Sjstrand and P. Skands, "A brief introduction to pythia 8.1", Computer Physics 495 Communications 2008 (2008), no. 178, 852.
- [22] S. Agostinelli et al., "Geant4a simulation toolkit", Nuclear Instruments and Methods in 506 Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 507 506 (2003), no. 3, 250 – 303, doi:10.1016/S0168-9002(03)01368-8.
- [23] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults Public CMS Luminosity Information
- [24] Robert D. Cousins, Kathryn E. Hymes, Jordan Tucker, "Frequentist Evaluation of Intervals Estimated for a Binomial Parameter and for the Ratio of Poisson Means", arXiv:0905.3831