

# Study of the Higgs boson production in association with a massive electroweak boson in final states with two b quarks and two leptons

---

Mesić, Benjamin

Doctoral thesis / Disertacija

2018

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:222888>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-05-08**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)





University of Zagreb

FACULTY OF SCIENCE

Benjamin Mesić

**STUDY OF THE HIGGS BOSON  
PRODUCTION IN ASSOCIATION WITH A  
MASSIVE ELECTROWEAK BOSON IN  
FINAL STATES WITH TWO B QUARKS  
AND TWO LEPTONS**

DOCTORAL DISSERTATION

Zagreb, 2018.





University of Zagreb

FACULTY OF SCIENCE

Benjamin Mesić

**STUDY OF THE HIGGS BOSON  
PRODUCTION IN ASSOCIATION WITH A  
MASSIVE ELECTROWEAK BOSON IN  
FINAL STATES WITH TWO B QUARKS  
AND TWO LEPTONS**

DOCTORAL DISSERTATION

Supervisor:  
Vuko Brigljević, PhD

Zagreb, 2018.







Sveučilište u Zagrebu

PRIRODOSLOVNO MATEMATIČKI FAKULTET

Benjamin Mesić

**PROUČAVANJE ZAJEDNIČKE TVORBE  
HIGGSOVA BOZONA I MASIVNOGA  
ELEKTROSLABOGA BOZONA U  
KONAČNIM STANJIMA S DVA B KVARKA  
I DVA LEPTONA**

DOKTORSKI RAD

Mentor:  
dr. sc. Vuko Brigljević

Zagreb, 2018.



## Supervisor's biography

### Personal

Name: Vuko Brigljević  
Date of birth: April 15, 1968  
Place of birth: Dubrovnik, Croatia  
Citizenship: Croatian and Swiss

### Education

1999 **PhD**, ETH Zurich  
Title : Measurements of Particle-Antiparticle oscillations in the  $B^0 - \bar{B}^0$  system  
Advisor : Prof. Hans Hofer  
1993 **Diploma in Physics**, ETH Zurich  
Title : Inklusive  $K_S^0$  production in Photoproduktion bei HERA  
Advisor : Prof. Ralph Eichler

### Positions

2016 - Head of the Laboratory for High Energy Physics at  
Ruđer Bošković Institute  
2014 - Senior Scientist, Ruđer Bošković Institute  
2008 - 2014 Senior Research Associate, Ruđer Bošković Institute  
2003 - 2008 Research Associate, Ruđer Bošković Institute  
2002 - 2003 CERN Research Fellow, working on the CMS Experiment  
1999 - 2002 Postdoctoral Research Associate at the Lawrence Livermore  
National Laboratory (Livermore, USA), working on the BABAR  
experiment at the Stanford Linear Accelerator Center  
1994 - 1999 Research and teaching assistant at the Institute for Particle  
Physics of ETH Zurich, working on the L3 experiment at CERN



## **Acknowledgements**

I would like to express my gratitude and appreciation to CMS collaboration and CMS group at Ruder Bošković Institute for their support in the development of the thesis.



# Abstract

In this thesis a study of the Higgs boson production in association with a W boson in the boosted topology is presented. The measurements are performed in two decay channels where the Higgs boson decays to two b quarks and W boson decays to electron or muon with their corresponding neutrinos,  $W(l\nu)H(bb)$ . Data samples analyzed in this work are collected by the CMS experiment at the CERN LHC in proton-proton collision and correspond to an integrated luminosity of  $35.9 \text{ fb}^{-1}$  at  $\sqrt{s}=13 \text{ TeV}$ .

Motivation for boosted topology comes from theoretical (BDRS) paper [1,2] in which authors suggest that substantial background reduction could be achieved with requirement that the Higgs boson candidate  $p_T$  is larger than 200 GeV. As a result, the signal event topology is characterized by the presence of a high  $p_T$  vector boson W recoiling from the Higgs boson candidate and two b quarks originating from the Higgs boson decay that are close to each other. Therefore, the key component is good reconstruction of the Higgs boson candidate, i.e. a jet with two genuine b quarks. For that purpose, AK08 jets have been used together with appropriate b-tagging algorithms for quantifying likelihood of having two b quarks inside. Definition of signal phase space is based on physical arguments: the W boson and the Higgs boson are expected to be back-to-back in the transverse plane, small jet multiplicity, and no b-hadron activity outside of the Higgs boson candidate.

Results are presented in terms of exclusion limits based on the  $CL_s$  prescription. The expected upper limits in the absence of a signal is 5.51 times the SM prediction, while the observed upper limit is 4.36. The expected pre-fit signal and background ratio is  $s/\sqrt{b} = 0.21$ , which shows that BDRS prediction ( $s/\sqrt{b} = 2.9$ ) was too optimistic. However, additional improvements of double b-tagger, better optimization of control region definitions, higher MC and data statistics, and perhaps usage of multivariate variables in final fit, could promote boosted analysis as a supplement to resolved analysis in high transverse momentum phase space, where two b-jets from the Higgs boson decay have significant overlap and represent difficult task for good reconstruction within resolved analysis.

Keywords: LHC, CMS, Standard Model, Higgs boson, WHbb, boosted





## Prošireni sažetak rada

### Uvod

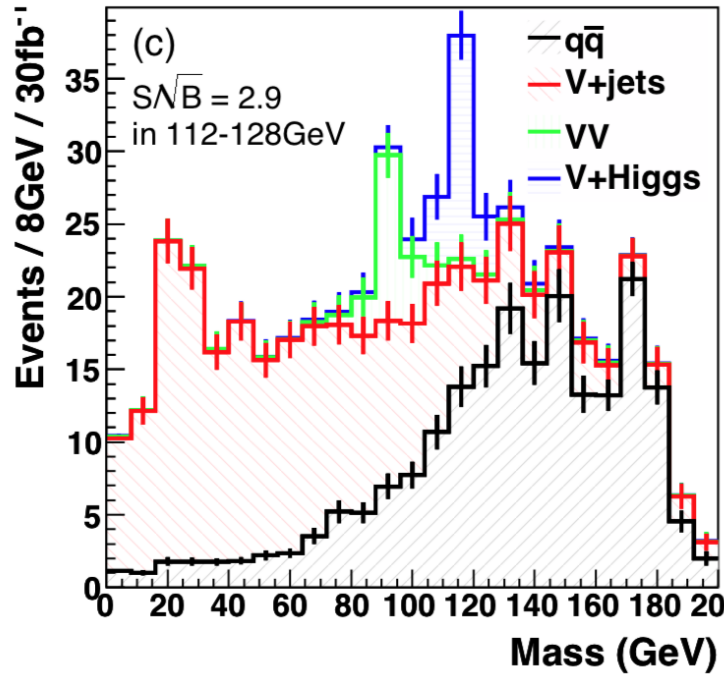
Ideja elementarnih čestica od kojih je načinjena sva materija pojavila se prije više od dvije i pol tisuće godina u okviru atomizma, filozofske teorije prema kojoj su svi objekti u svemiru sastavljeni od vrlo malih i nedjeljivih čestica, atoma. Utemeljiteljima antičkog atomizma smatraju se grčki filozofi Leukip i Demokrit. Iako se mnoštvo velikih znanstvenika zanimalo za filozofska pitanja od čega je materija sastavljena, uključujući i našeg Rudera Boškovića, tek je početkom 19. stoljeća John Dalton postavio formalne temelje teorije. Dmitrij Ivanovič Mendeljev se nadovezao na Daltonovu ideju te je 1869. godine izradio periodni sustav elemenata, sustavni poredak kemijskih elemenata po njihovim karakteristikama. Periodni sustav elemenata održao se kao fundamentalna teorija jedan kratak vremenski period do otkrića elektrona, protona i neutrona. Postuliranjem samo tri elementarne čestice, mogao se objasniti čitav periodni sustav elemenata. Međutim, sredinom 20. stoljeća otkriveno je veliko mnoštvo novih čestica te je time započela potraga za još elementarnijim česticama. Rezultat nekoliko sljedećih desetljeća eksperimentalnih i teorijskih istraživanja je Standardni Model, moderna teorija elementarnih čestica izgrađena na kvantnoj teoriji polja koja se održala sve do danas. Standardni model postulira 12 čestica materije i ukupno 4 čestice prijenosnika elektroslabe i jake nuklearne sile. Uz njih tu je i Higgsov bozon, posljednji dio slagalice eksperimentalno potvrđen 2012. godine. Važnost Higgsova bozona odnosno Higgsova polja može se promatrati iz nekoliko aspekata. Naime, on je zaslužan za generiranje mase ostalih elementarnih čestica kao i rješavanje mnogih tehničkih poteškoća koje se pojavljuju u Standardnom modelu te stoga ne čudi koliki je interes pobudio u znanstvenim krugovima posljednjih pedesetak godina.

Unatoč impresivnoj točnosti u predviđanju i objašnjenju fizikalnih pojava, Standardni model ostavlja i dalje mnoga otvorena pitanja te ukazuje na to da su potrebna njegova proširenja. Kvantna teorija gravitacije je sama po sebi neriješen izazov zbog kojeg se Standardni model i nakon nekoliko desetljeća pokušaja odupire ujedinjenju s gravitacijom. Jakost gravitacijske sile u usporedbi s ostalim silama je zanemariva za uvjete u kojima se odvijaju trenutna istraživanja. Međutim, za ekstremne uvjete definirane primjerice na Planckovoj skali, kvantna gravitacija je neizostavna komponenta neke veće teorije u kojoj su sve sile ujedinjene. Trenutno, najbolji kandidat je teorija struna. Sljedeće otvoreno pitanje je vezano za kozmološka mjerenja iz kojih slijedi kako samo 4% ukupne

energije svemira je objašnjeno Standardnim modelom. Postoje mnoga proširenja modela koja bi mogla objasniti kozmološka opažanja, ali do danas niti jedna čestica predviđena takvim proširenjima nije opažena. Nadalje, neutrimi su prema Standardnom modelu bez-masene čestice dok eksperimentalna mjerenja njihovih oscilacija govore da oni imaju masu. Maseni neutrimi se mogu dodati ručno u teoriju, ali to donosi sa sobom dodatne poteškoće na koje treba paziti. Zatim u trenutnom obliku modela, postoji 19 slobodnih parametara čije su vrijednosti određene iz različitih eksperimenata, ali sama priroda tolikog broja parametara nije poznata. Omjer materije i antimaterije u svemiru bi prema Standardnom modelu trebao biti jednak, što se iz eksperimenta, ali i svakodnevnog života može vidjeti da nije slučaj. Osim eksperimentalnih potvrda kako Standardni model nije kraj priče, postoji još mnoštvo tehničkih detalja koje je potrebno riješiti. Dakle, Standardni model je uspješna teorija elementarnih čestica koja se u pažljivoj sprezi eksperimenta i teorije gradila kroz mnoge generacije i čije su temelje postavili velikani svjetske znanosti. Ipak, i dalje ostaju mnoga pitanja čije je odgovore potrebno pričekati još neko vrijeme.

## Cilj rada

Nakon eksperimentalne potvrde Higgsova bozona, potrebno je bilo dokazati i njegova preostala svojstva predviđena teorijom. Prema Standardnom modelu, Higgsov bozon se najčešće raspada u dva b kvarka, međutim konkretan proces je teško opaziti na sudarivaču LHC zbog prevelike pozadine, odnosno, ostalih procesa koji ostavljaju gotovo identičan potpis u detektoru. Kako bi se smanjio doprinos pozadine, fizičari su se okrenuli nešto složenijim procesima. Primjer takvog procesa je zajednička tvorba Higgsova i W bozona prilikom kojeg se Higgsov bozon raspada u dva b kvarka, a W u dva leptona. Istraživanja su pokazala kako je u ovom slučaju pozadina značajno potisnuta, ali je ipak i dalje prisutna. Kao moguće rješenje, iz teorijskih razmatranja Butterwortha, Davisona, Rubina i Salama (BDRS članci) [1, 2] proizašla je ideja ultra-relativističke topologije koja se postiže zahtjevom na veliki iznos impulsa Higgsova bozona. Produkti raspada ultra-relativističke čestice su u sustava detektora jako rijetko udaljeni jedan od drugoga. Kako bi se označio kandidat Higgsova bozona potrebno je naći dva prostorno bliska potpisa dva b kvarka koja nisu karakteristična za pozadinu. Time je dobiveno novo oružje u razlikovanju Higgsova bozona, odnosno signala od pozadine. Distribucije signala i dominantnih doprinosa pozadinskih procesa prema BDRS članku, su prikazane na Slici 1. Istraživanje u ultra-relativističkom režimu ima i svoju cijenu. Iz teorijskih procjena slijedi kako se



Slika 1: Očekivane distribucije signala i pozadine u ultra-relativističkom režimu za kanal raspada  $W(l\nu)H(bb)$  prema BDRS.

samo 5% ukupnog signala nalazi u tom djelu faznog prostora. Ovaj rad je motiviran BDRS člancima te se u njemu eksperimentalno proučava ultra-relativistički raspad Higgsova bozona u dva  $b$  kvarka uz pridruženi  $W$  bozon koji se pritom raspada u dva leptona. Postoje ukupno dva  $WH(bb)$  kanala ovisno o konačnim stanjima raspada  $W$  bozona:  $H(bb)W(\mu\nu)$  i  $H(bb)W(e\nu)$ . U ovom radu su oba analizirana. Također, važno je spomenuti kako se  $WH(bb)$  proces može analizirati i bez zahtjeva na ultra-relativistički režim ("resolved" analiza). U tom slučaju, dva  $b$  kvarka ne moraju nužno biti jedan blizu drugoga.

### Eksperimentalni postav

Prema Standardnom modelu, elementarne čestice su naprosto pobuđenja kvantnih polja. Kako bi se stvorile nove čestice, potrebno je izazvati dovoljno jaku reakciju ili drugim riječima, potrebno je pobuditi polja što se postiže sudaranjem postojećih čestica. Povećanjem njihove kinetičke energije raste i vjerojatnost za nastanak masivnijih objekata. Postoje dva izvora visoko energetskih čestica, a to su kozmička zračenja i laboratoriji s ubrzivačima. Za razliku od kozmičkog zračenja, koje sadrži čestice različitih energija te one pritom dolaze iz različitih kuteva svemira, u laboratoriju su uvjeti dobro kontrolirani što je velika prednost. S druge pak strane, nedostatak laboratorija je ograničenje energije čes-

tica koje je uvjetovano samim ubrzivačem. Trenutno najveći ubrzivač čestica na svijetu LHC (Large Hadron Collider) nalazi se u CERN-u u Švicarskoj. Točnije, riječ je o nekoliko povezanih ubrzivača od kojih je posljednji u nizu LHC. LHC je enormno veliki stroj smješten u kružnom tunelu opsega 27 kilometara na dubini od prosječno 100 metara ispod površine zemlje. Konstruiran je za ubrzavanje protona na energije od 14 TeV (trenutno 13 TeV) u sustavu centra mase. Osim protona, mogu se sudarati i teški ioni. Iz analize podataka prikupljenih u sudarima, dolazi se do fizikalnih spoznaja o samim česticama. Postoje četiri mjesta na samom ubrzivaču gdje dolazi do sudara, a na njima je ukupno smješteno sedam detektora: CMS, ATLAS, ALICE, LHCb, LHCf i TOTEM. Svaki od detektora je jedinstven i prilagođen za posebne zadatke.

Za ovaj rad korišteni su podaci prikupljeni na detektoru CMS. Sa svojih 14 000 tona, 15 metara širine i 21 metar dužine predstavlja svojevrsnu kameru koja je u mogućnosti snimiti 40 milijuna fotografija sudara svake sekunde. Koristeći sve informacije iz detektora zajedno, može se vrlo precizno rekonstruirati i identificirati pojedine čestice, odnosno kvantitativno opisati njihova fizikalna svojstva. Konačna fotografija sudara dobiva se kombiniranjem informacija svih rekonstruiranih čestica. Za dobru rekonstrukciju čestica zaslužan je slojeviti dizajn detektora u kojem svaki sloj ima posebnu ulogu. Snažno magnetsko polje koristi se za zakretanje putanja nabijenih čestica. Iz smjera zakretanja putanje može se odrediti predznak naboja čestice jer se čestice različitih predznaka zakreću u suprotnom smjeru dok se iz samog iznosa zakrivljenosti putanje može izmjeriti količina gibanja čestice. Magnet je načinjen od supravodljivog materijala kroz koji prolazi struja od 18 500 A te u konačnici stvara magnetsko polje od 4 T, 100 000 puta jače od magnetskog polja Zemlje. Samo zakretanje putanje čestica nije dovoljno za njihovu identifikaciju. Potrebno je rekonstruirati putanje s velikom preciznošću za što je zaslužan sustav za detekciju tragova nabijenih čestica. On je načinjen od 75 milijuna poluvodičkih senzora poslaganih u koncentrične slojeve. Nabijene čestice prolaskom kroz poluvodički senzor izbijaju elektrone iz atoma silicija i time stvaraju elektron-šupljina parove. Električnim poljem pokupe se izbijeni elektroni te se dobije mali električni puls, odnosno mjesto gdje je nabijena čestica prošla. Kombinirajući električne signale iz više takvih slojeva moguće je rekonstruirati putanju pojedine čestice s velikom preciznošću. Iza sustava za detektiranje tragova čestica nalaze se dva sustava za mjerenje energija čestica. Elektromagnetski kalorimetar (ECAL) je bliže središtu detektora te se njime mjeri

energija elektrona i fotona čime se oni gotovo u potpunosti zaustavljaju i ne propagiraju kroz ostatak detektora. Hadronski kalorimetar (HCAL) nalazi se iza elektromagnetskog i služi za određivanje energije hadrona čiji se potpis u detektoru naziva hadronski pljusak. Preostale čestice koje se nisu zaustavile niti u jednom sloju detektora, a moguće ih je opaziti su mioni. Za njih postoji posebno izgrađen sustav mionskih komora. Detektor CMS je hermetički zatvoren, odnosno, pokriva sve moguće putanje čestica. Ako se pak izmjeri znatna neravnoteža u energiji i impulsu (MET), vjerojatno je riječ o čestici koja ne interagira s detektorom, poput neutrina. Kompleksan stroj poput detektora CMS je nužan kako bi se moglo pomicati granice razumijevanja i odgovarati na pitanja fizike elementarnih čestica. Glavna motivacija za izgradnju detektora je

- razumijevanje fizike na TeV skali
- otkriće i razumijevanje fizike Higgsova bozona
- potraga za fizikom izvan Standardnog modela
- razumijevanje fizike teških iona

## **Analiza podataka**

Prvi korak svake analize nekog fizikalnog procesa je razumijevanje pozadine koja bi se mogla pojaviti sa sličnim karakteristikama i tako značajno otežati samu analizu. U slučaju raspada  $WHbb$  s konačnim stanjima  $WH(bb)W(\mu\nu)$  i  $H(bb)W(e\nu)$ , postoji nekoliko takvih procesa unatoč ultra-relativističkoj topologiji, od kojih su najistaknutiji oni u kojima se pojavljuju top kvarkovi,  $W$  bozoni uz pridružene mlazove ili pak neki od parova ( $WZ$ ,  $WW$ ) bozona. Glavna značajka ultra relativističkog  $WHbb$  raspada po kojoj ga se može razlikovati od ostalih procesa su dva prostorno bliska  $b$ -hadrona s velikim ukupnim impulsom koji zajedno čine kandidata za Higgsov bozon ( $HC$ ) te rekonstruirani  $W$  bozon s približno jednakim impulsom suprotnog smjera. Također, u prosjeku se očekuje manji broj mlazova u događaju nego što je to u slučaju pozadine. S obzirom da su procesi pozadine učestaliji od signala i do nekoliko redova veličina, realno je za očekivati da će i nakon potpune selekcije ostati značajan dio pozadine. Temelj ove analize je dobra rekonstrukcija mlazova i njihovih svojstava: energije, mase i varijabli koje opisuju od čega su građeni. Mlazovi su rekonstruirani anti- $k_t$  algoritmom. Kada je riječ o

ultra-relativističkim topologijama, gdje se traže objekti s velikim impulsima i nekom specifičnom podstrukturom, obično se koriste barem dvostruko širi mlazovi ( $\Delta R = 0.8$ , AK08) od uobičajenih mlazova. U žargonu se često nazivaju i debeli mlazovi ("Fat jet"). Polazna točka za izgradnju Higgsova kandidata u ovoj analizi su AK08 mlazovi. Svakom AK08 mlazu je pridružena vrijednost koja opisuje kolika je vjerojatnost da se u njemu nalaze dva b-hadrona. Ona je rezultat posebnog algoritma nazvanog bb-označivač (bb-tagger) koji koristi informacije o svim nabijenim tragovima unutar samog mlaza i činjenicu da je vrijeme raspada b-hadrona značajno dulje od vremena raspada ostalih hadrona. Kandidat za Higgsov bozon je onaj AK08 mlaz koji ima najveću vrijednost bb-označivača od svih AK08 mlazova u događaju i  $p_T > 250$  GeV. Za rekonstrukciju W bozona koriste se MET i vodeći izolirani lepton u događaju. Osim AK08 mlazova, koriste se AK04 mlazovi i to za više funkcija: za ukupan broj dodatnih mlazova u događaju, potragu dodatne aktivnosti b-hadrona izvan i c-hadrona unutar Higgs bozon kandidata. Od 40 milijuna sudara snimljenih svake sekunde u detektoru CMS, samo su neki od njih zanimljivi za analizu. Zadatak odabira onih najzanimljivih pada na okidač ("trigger") koji koristi primitivne informacije iz detektora i u kratkom vremenskom periodu odlučuje hoće li se događaj prihvatiti ili odbaciti. U ovoj analizi korišten je okidač na izolirane leptone: mione ( $>24$  GeV) i elektrone ( $>27$  GeV). U 2016. godini, ukupno je prikupljeno  $35.9 \text{ fb}^{-1}$  na  $\sqrt{s} = 13$  TeV koji su ovdje i analizirani.

Za optimizaciju analize korišteni su simulirani uzorci koji su rezultat slijeda nekoliko različitih simulacija počevši od simulacije fizike sudara protona pa sve do simulacije odziva detektora. Kompletna selekcija za pronalaženje faznog prostora u kojem je signal najizraženiji izgrađena je na ukupno osam varijabli, koje su navedene u Tablici 1. HC bb označivač i HC  $\tau_2/\tau_1$  nameću zahtjev na postojanje dva b-hadrona unutar samog mlaza te time potiskuju pozadinu. Dodatnom selekcijom na omjer impulsa Higgs kandidata i W kandidata ostvaruje se u potpunosti ultra-relativistička topologija. S druge strane, broj dodatnih mlazova i veto na b-hadrone izvan mlaza Higgs kandidata iskorištavaju sama svojstva pozadine kako bi se ona dodatno potisnula. Veći broj dodatnih mlazova je karakteristika ponajviše procesa koji uključuju top kvarkove. Kako bi se precizno odredila normalizacija dominantnih pozadinskih procesa i također provjerilo slaganje simuliranih i izmjerenih distribucija najvažnijih opservali u analizi, potrebno je definirati nekoliko takozvanih kontrolnih uzoraka, područja koja su vrlo blizu faznom prostoru signalnog po-

dručja, ali su pritom potpuno ortogonalna. U ovoj analizi konstruirana su tri kontrolna uzorka gdje dominiraju sljedeći pozadinski procesi:  $t\bar{t}$ , W bozon uz teže mlazove i W bozon uz lakše mlazove. Prilikom prilagodbe iz koje se izvlače konačni rezultati jakosti signala, oblici pozadina se također prilagođavaju kako bi zajedno dali najbolje slaganje s izmjerenim podacima.

Tablica 1: Selekcija za signalno područje.

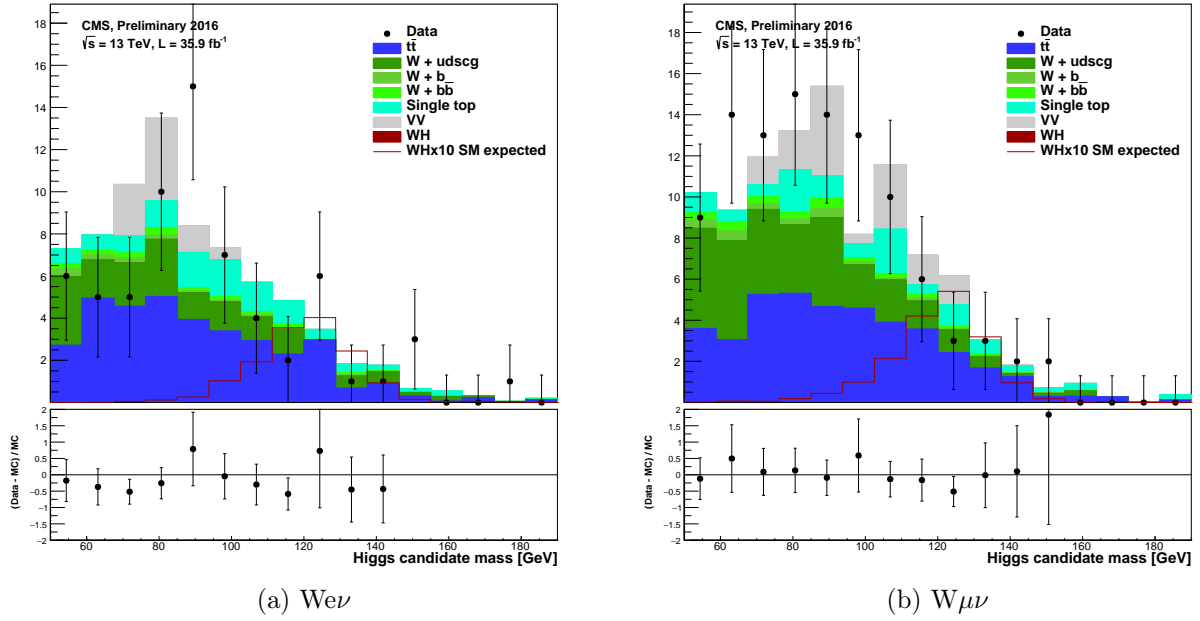
Varijabla	Selekcija
HC $p_T$	$> 250$ GeV
HC masa	$\in [50 - 190]$ GeV
HC bb označivač	$> 0.8$
V $p_T$ / HC $p_T$	$\in [0.8 - 1.2]$
HC $\tau_2/\tau_1$	$< 0.45$
N dod. mlazova	$< 2$
b tag veto	Da
c tag veto	Da

## Rezultati

Ključni rezultat ovog rada je izražen preko 95% C.L. ( $CL_S$ ) gornje granice. Za detaljnije objašnjenje statističke metode  $CL_S$  pogledati Odlomak 7.5. Distribucije mase Higgs bozon kandidata u signalnom uzorku nakon prilagodbe, prikazane su na Slici 2. Ukupan broj izmjerenih i očekivanih događaja prikazan je u Tablici 2. Dominantan izvor nesigurnosti u ovom mjerenju je statistička pogreška u simuliranim podacima koja proizlazi iz ograničenog broja simuliranih događaja. Rezultat standardne ("resolved") analize na podacima iz 2016. godine, gdje se Higgs bozon kandidat izgrađuje koristeći dva AK04 mlaza, je opažanje viška događaja za hipotezu postojanja samo pozadine. Signifikantnost opažanja je  $3.3\sigma$  [3]. Potencijalna mogućnost korištenja analize u ultra-relativističkom režimu je njena komplementarnost "resolved" analizi. Naime, ukupan broj očekivanih događaja u signalu koji prolazi ultra-relativističku selekciju je 3.71, a očekivani broj događaja koji prolazi dodatno i "resolved" selekciju je 2.61. Očekivana 95% C.L. gornja granica na omjer  $\sigma/\sigma_{SM}$  WHbb produkcije je 5.51 dok je opažena vrijednost 4.36, kao što



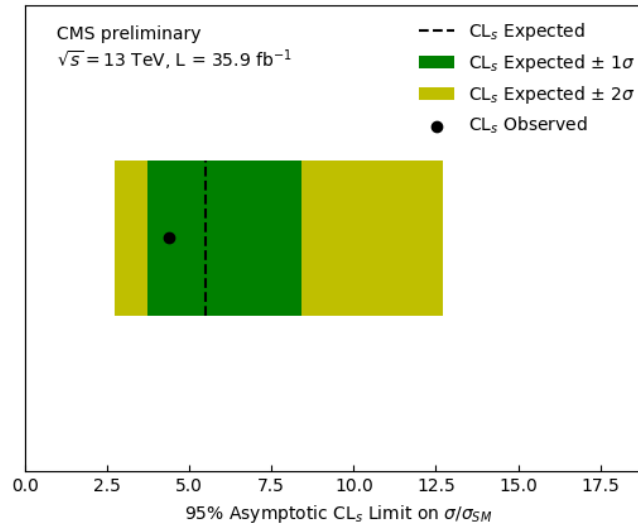
je prikazano na Slici 3.



Slika 2: Distribucija mase Higgs bozon kandidata u signalnom uzorku nakon prilagodbe.

Tablica 2: Ukupan broj izmjerenih i očekivanih događaja za pozadinu (B) i WH signal (S). Naveden je i očekivani omjer signala i pozadine (S/B).

Proces	Broj događaja	
	Pre-fit	Post-fit
$W + b\bar{b}$	30.61	4.69
$W + b$	14.20	3.81
$W + udscg$	14.13	45.83
$t\bar{t}$	101.25	75.23
Single-top-quark	32.22	22.63
VV	26.88	22.18
Ukupna pozadina	219.29	174.37
WH	3.71	0.0
Mjerenje	170	
S/B	0.016	0.0



Slika 3: Očekivana i izmjerena 95% C.L. gornja granica (u odsutnosti signala) na omjer  $\sigma/\sigma_{SM}$  WHbb produkcije na podacima prikupljenim u 2016. godini.

## Zaključak

U ovom su radu predstavljeni rezultati proučavanja zajedničke tvorbe Higgsova bozona i masivnog elektroslabog bozona u konačnim stanjima s dva b kvarka i dva leptona u ultra relativističkom režimu. Analizirani podatci su prikupljeni 2016. godine detektorom CMS u sudarima protona na energiji  $\sqrt{s} = 13 \text{ GeV}$ . Ukupna količina podataka odgovara integriranom luminozitetu od  $35.9 \text{ fb}^{-1}$ . Teorijska predviđanja Butterwortha, Davisona, Rubina i Salama (BDRS) koja su ujedno bila i motivacija za istraživanje predstavljeno u ovom radu, su pokazana kao preoptimistična. Očekivani omjer signala i pozadine je prema BDRS članku  $s/\sqrt{b} = 2.9$ , dok ova analiza pokazuje kako je taj broj red veličine manji, odnosno  $s/\sqrt{b} = 0.2$ . Očekivana 95% C.L. gornja granica na omjer  $\sigma/\sigma_{SM}$  WHbb produkcije je 5.51 dok je opažena vrijednost 4.36. Dominantan izvor nesigurnosti u ovom mjerenju je ograničena veličina simuliranog uzorka koja je ujedno i uobičajena za analize temeljene na ultra relativističkim topologijama. Uz dodatna poboljšanja identifikacije mlazova s dva b-hadrone te veću količinu simuliranih i prikupljenih podataka, ova analiza može biti dodatak standardnoj "resolved" analizi raspada Higgsova bozona u dva b kvarka.



# Contents

Supervisor's biography

Acknowledgements

Abstract

Prošireni sažetak rada

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Particle physics and the Standard Model . . . . .	1
1.1.1	Historical background . . . . .	3
1.1.2	Theoretical aspects . . . . .	4
1.1.3	Value of Standard Model parameters . . . . .	9
1.2	Physics beyond the Standard Model . . . . .	10
<b>2</b>	<b>The Higgs boson phenomenology at hadron colliders</b>	<b>13</b>
2.1	Phenomenology of proton-proton interactions . . . . .	13
2.2	The Higgs boson production in proton-proton collisions . . . . .	17
2.3	The Higgs boson decay . . . . .	19
2.4	Kinematic variables used in proton-proton collisions . . . . .	22

---

2.5	Boosted topologies . . . . .	22
<b>3</b>	<b>The Large Hadron Collider</b>	<b>25</b>
3.1	Introduction . . . . .	25
3.2	Collider design . . . . .	26
3.3	The Luminosity frontier . . . . .	34
3.4	Timeline of operations . . . . .	35
<b>4</b>	<b>The Compact Muon Solenoid Detector</b>	<b>37</b>
4.1	Introduction . . . . .	37
4.2	Detector design . . . . .	40
4.2.1	The Tracker . . . . .	42
4.2.2	The Electromagnetic Calorimeter (ECAL) . . . . .	46
4.2.3	The Hadronic Calorimeter (HCAL) . . . . .	49
4.2.4	The Solenoid . . . . .	51
4.2.5	The Muon System . . . . .	52
4.3	The Trigger . . . . .	54
4.3.1	Level-1 Trigger . . . . .	55
4.3.2	High level trigger (HLT) . . . . .	57
<b>5</b>	<b>Event simulation</b>	<b>59</b>
5.1	Introduction . . . . .	59
5.2	Matrix element (ME) . . . . .	61
5.3	Parton shower (PS) . . . . .	62

5.4	Underlying Event . . . . .	62
5.5	Merging Matrix Element and Parton Shower . . . . .	62
5.6	Hadronization . . . . .	63
5.7	Pile-Up Interactions . . . . .	64
5.8	Detector simulation . . . . .	64
<b>6</b>	<b>Event reconstruction</b>	<b>67</b>
6.1	The Particle Flow technique . . . . .	67
6.2	Tracks and vertices reconstruction . . . . .	68
6.3	Electrons . . . . .	70
6.4	Muons . . . . .	72
6.5	Lepton isolation . . . . .	74
6.6	Jets . . . . .	75
6.6.1	Jet energy corrections . . . . .	78
6.6.2	Jet energy resolution . . . . .	79
6.6.3	Jet identification . . . . .	79
6.6.4	Jet substructure . . . . .	80
6.6.5	Heavy flavour jet . . . . .	82
6.7	Missing transverse energy . . . . .	90
6.8	Vector boson reconstruction . . . . .	91
6.9	The Higgs boson reconstruction . . . . .	91

<b>7</b>	<b>Search for boosted <math>W(l\nu)H(b\bar{b})</math> production</b>	<b>95</b>
7.1	Data sets and triggers . . . . .	95
7.2	Monte Carlo samples . . . . .	96
7.3	Analysis Strategy . . . . .	98
7.3.1	Event reconstruction and selection . . . . .	99
7.4	Background and scale factor estimation . . . . .	107
7.5	Signal extraction and limit setting . . . . .	110
7.6	Systematic uncertainties . . . . .	111
7.7	Results . . . . .	114
7.8	Impacts of the systematic uncertainties . . . . .	117
7.9	Comparison with resolved analysis . . . . .	118
<b>8</b>	<b>Conclusion</b>	<b>121</b>
	<b>Bibliography</b>	<b>122</b>
<b>A</b>	<b>Conventions</b>	<b>137</b>
	<b>Curriculum vitae</b>	<b>139</b>

# List of Tables

1	Selekcija za signalno područje. . . . .	
2	Ukupan broj izmjerenih i očekivanih događaja za pozadinu (B) i WH signal (S). Naveden je i očekivani omjer signala i pozadine (S/B). . . . .	
1.1	Free parameters of the Standard Model with their measured values. . . . .	9
2.1	The SM Higgs boson ( $m_H=125$ GeV) branching ratios. . . . .	20
3.1	LHC technical parameters for proton-proton collisions. . . . .	30
4.1	A summary of the principal characteristics of the various tracker subsystems. . . . .	45
6.1	Preselection requirements used for electron identification. . . . .	72
6.2	Summary of the muon identification variables and the corresponding selection. . . . .	74
6.3	Working points and corresponding efficiency for b jets with $p_T > 20$ GeV in simulated $t\bar{t}$ events. The numbers in this table are for illustrative purposes since the efficiency is integrated over the $p_T$ and $\eta$ distributions of jets. . . . .	85
6.4	Working points and corresponding efficiency for c jets with $p_T > 20$ GeV in simulated $t\bar{t}$ events. The numbers in this table are for illustrative purposes since the efficiency is integrated over the $p_T$ and $\eta$ distributions of jets. . . . .	87



7.1	Transverse momentum thresholds applied in the lepton triggers at the HLT level. . . . .	96
7.2	Summary of the samples of simulated processes. . . . .	97
7.3	The number of additional jets definition. . . . .	103
7.4	b-tag veto definition. . . . .	104
7.5	c-tag veto definition. Events with at least one AK04 jet fulfilling these requirements are rejected. . . . .	105
7.6	Selection criteria that define the signal region. . . . .	106
7.7	The expected number of events in signal region after applying each requirement (cut-flow). All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . . . . .	107
7.8	Definition of signal and control regions. Flipped cuts are highlighted in red.	108
7.9	The total number of events in signal region for the expected prefit/postfit backgrounds (B), WH signal (S), and for data. Also shown is the signal-to-background ratio (S/B). . . . .	114
7.10	Overlap of boosted and resolved signal selection pre-fit yields. Normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . . . . .	119

# List of Figures

1	Očekivane distribucije signala i pozadine u ultra-relativističkom režimu za kanal raspada $W(l\nu)H(bb)$ prema BDRS. . . . .	
2	Distribucija mase Higgs bozon kandidata u signalnom uzorku nakon prilagodbe. . . . .	
3	Očekivana i izmjerena 95% C.L. gornja granica (u odsutnosti signala) na omjer $\sigma/\sigma_{SM}$ $WHbb$ produkcije na podatcima prikupljenim u 2016. godini.	
1.1	A diagram summarizing interactions between elementary particles according to the Standard Model. . . . .	3
2.1	Parametrization of the PDFs of the proton in the MSTW 2008 fit, at different scales. On the x-axis the longitudinal momentum fraction $x$ is shown, while on the y-axis the product $x \cdot f(x, Q^2)$ , where $f(x, Q^2)$ is the parton distribution function. Different partons are shown in different colors, including uncertainties. . . . .	15
2.2	Schematic representation of an event as produced by an event generator. The hard interaction (big red circle) is followed by the decay of both top quarks and the Higgs boson (small red circles). Additional hard QCD radiation is produced (red) and a secondary interaction takes place (purple blob) before the final-state partons hadronize (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow). . . . .	16

2.3	Leading-order Feynman diagrams contributing to the Higgs boson production. . . . .	18
2.4	The Higgs boson production cross section as a function of the LHC centre of mass energy and its mass at $\sqrt{s} = 14$ TeV. . . . .	19
2.5	The Higgs boson branching ratios as a function of the Higgs boson mass. . .	20
2.6	Leading-order Feynman diagrams representing the Higgs boson decay. . . .	21
2.7	Expected signal and background distributions for a 115 GeV SM Higgs in boosted topology ( $W(l\nu)H(bb)$ channel) according to calculation in BDRS paper. . . . .	23
2.8	Schematic decay of a particle $m$ with momentum $p$ into two particles separated by distance $\Delta R$ . . . . .	24
3.1	World map of countries involved in research at CERN. . . . .	26
3.2	Schematic description of the accelerator complex. . . . .	28
3.3	Schematic cross section of a LHC dipole magnet. . . . .	29
3.4	LHC schematic configuration showing clockwise beam colliding with counterclockwise beam with the primary goal of the LHC being at four intersecting points. . . . .	31
3.5	Schematic perspective view of the four main experiments installed at the LHC. From the top left, clockwise: CMS, LHCb, ATLAS and ALICE. . . .	32
3.6	The proposed LHC running schedule. By the end of its lifetime in 2035, it is expected to collect more than $3000 \text{ fb}^{-1}$ of data. . . . .	35
4.1	A view of the CMS detector with its subdetectors labeled. . . . .	38
4.2	Cumulative and peak luminosity versus day delivered to CMS by the LHC.	39

4.3	A sketch of the specific particle interactions in a transverse slice of the CMS detector, from the beam interaction region to the muon detector. The muon and the charged pion are positively charged, and the electron is negatively charged. . . . .	41
4.4	Material budget in units of radiation length as a function of pseudorapidity $\eta$ , for the different sub-detectors (left panel) and broken down into the functional contributions (right panel). . . . .	42
4.5	A schematic view of the pixel vertex detector. . . . .	43
4.6	Schematic cross section through the CMS tracker in the r-z plane. In this view, the tracker is symmetric about the horizontal line $r = 0$ , so only the top half is shown here. The centre of the tracker, corresponding to the approximate position of the pp collision point, is indicated by a star. Green dashed lines help the reader understand which modules belong to each of the named tracker subsystems. . . . .	45
4.7	Layout of the CMS electromagnetic calorimeter presenting the arrangement of crystal modules, supermodules, endcaps and the preshower in front. . . .	46
4.8	Geometric view of one quarter of the ECAL. . . . .	48
4.9	Geometric view of the CMS detector showing the HCAL subdetectors. . .	50
4.10	Geometric view of the CMS muon system, demonstrating the $\eta$ barrel (MB1–MB4, green), the four CSC stations in the endcap (ME1–ME4, blue), and the RPC coverage and overlap of all constituent subsystems. . . . .	52
4.11	Sketch of a cell showing drift lines and isochrones. The plates at the top and bottom of the cell are at ground potential. . . . .	53
4.12	Level-1 Trigger components . . . . .	56
5.1	Illustration of the individual steps of the event simulation procedure. . . .	60
6.1	Mean number of interactions per bunch crossing for the 2016 proton-proton (pp) collisions at 13 TeV. . . . .	70

6.2	Muon $p_T$ resolution as a function of the muon $p_T$ in the barrel (left) and in the endcap (right) region. . . . .	73
6.3	Jets reconstructed with different algorithms starting from the same set of simulated particles. . . . .	77
6.4	The PF algorithm attempts to fully reconstruct jet by combining information from all CMS subdetectors. . . . .	78
6.5	Schematic representation of the AK08 jet with corresponding $\tau$ axes. . . .	81
6.6	Illustration of a heavy-flavour jet with a secondary vertex (SV) from the decay of a heavy-flavour hadron resulting in charged-particle tracks that are displaced with respect to the primary interaction vertex (PV) and with a large impact parameter (IP) value. . . . .	82
6.7	Distribution of the CSVv2 discriminator values for jets of different flavours in simulated $t\bar{t}$ events. The distributions are normalized to unit area. . . .	83
6.8	Misidentification probability for c and light-flavour jets versus b jet identification efficiency for CSVv2 and also other b tagging algorithms applied to jets in $t\bar{t}$ events. . . . .	84
6.9	Distribution of the c-tagging discriminator values for jets of different flavours in $t\bar{t}$ events. The spikes originate from jets without a track passing the track selection criteria. . . . .	86
6.10	Correlation between CvsL and CvsB taggers for the various jet flavours. The L, M, and T working points discussed in the text are indicated by the dashed lines. . . . .	87
6.11	Distribution of the double-b tagger discriminator for $H \rightarrow b\bar{b}$ jets and for jets in an inclusive multijet sample containing zero, one, or two b quarks. .	88
6.12	Efficiency to correctly tag $H \rightarrow b\bar{b}$ jets and misidentification probability using jets in an inclusive multijet sample for four working points of the double-b tagger as a function of the jet $p_T$ . The AK08 jets are selected with $p_T > 300$ GeV and pruned jet mass between 50 and 200 GeV. . . . .	89

6.13	The Higgs boson candidate bb-tag value. Depending on the number of b quarks within HC, three categories are defined: 0b (red), 1b (blue), and 2b (purple). . . . .	92
6.14	The Higgs boson candidate $p_T$ distribution. Boosted topology starts at 250 GeV, when two b quarks are successfully caught within AK08 jet. . . . .	93
6.15	The Higgs boson candidate mass distribution. . . . .	93
7.1	Cumulative measured luminosity versus day delivered for the 2016 proton-proton (pp) collisions at 13 TeV. . . . .	96
7.2	Multiplicative weights to apply the differential NLO electroweak signal correction. . . . .	98
7.3	Feynman diagrams showing dominant background processes. . . . .	100
7.4	N-1 plot of HC double b-tag distribution. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	101
7.5	N-1 plot of HC $\tau_2/\tau_1$ distribution. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	102
7.6	N-1 plot of $V \text{ } p_T / \text{ HC } p_T$ distribution. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	102
7.7	N-1 plot of number of additional jets distribution. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	103
7.8	N-1 plot of b-tag veto distribution. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	104

7.9	N-1 plot of c-tag veto distribution. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	105
7.10	Correlation between CvsL and CvsB taggers for the various jet flavours. Combination of L and T working point has been used in this analysis. . . .	106
7.11	Pre-fit HC $p_T$ distribution in $t\bar{t}$ control region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	108
7.12	Pre-fit HC $p_T$ distribution in light flavor (LH) control region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	109
7.13	Pre-fit HC $p_T$ distribution in heavy flavor (HF) control region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	109
7.14	Post-fit HC $p_T$ distribution in $t\bar{t}$ control region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	115
7.15	Post-fit HC $p_T$ distribution in light flavor (LH) control region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	115
7.16	Post-fit HC $p_T$ distribution in heavy flavor (HF) control region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	116
7.17	Post-fit HC mass distribution in signal region. All samples are normalized to an integrated luminosity of $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown alone with a yield 10 times as large as its SM prediction. . . . .	116

7.18	Expected and observed 95% C.L. upper limits (in the absence of signal) on the ratio $\sigma/\sigma_{SM}$ of WHbb production for 13 TeV, 2016 data. The median expected limit, observed limit and the $1\sigma$ and $2\sigma$ bands are obtained with the full LHC CL <sub>S</sub> method. . . . .	117
7.19	List of nuisances with the largest effect on the signal strength uncertainty. The plot also shows the best fit value of $\mu$ ( $\hat{\mu}$ ) at the top and its uncertainty.	118
7.20	Results of resolved VH(bb) analysis on data recorded by the CMS experiment at the LHC in 2016. Figure a) The best fit value of the signal strength $\mu$ , at $m_H = 125.09$ GeV, is shown in black with a green uncertainty band. Figure b) Weighted dijet invariant mass distribution. Shown are data and the VH and VZ processes with all other background processes subtracted. .	119





# Chapter 1

## Introduction

In this chapter, the main aspects of the Standard Model, the theory that contains our current knowledge of elementary particles and their interactions, are described. The motivation for the Standard Model is given through a brief historical overview of important discoveries, while its implications and mathematical background are given in a separate subsection. The current knowledge of the Standard model parameters based on experimental data is summarized at the end of this chapter.

### 1.1 Particle physics and the Standard Model

Elementary particle physics addresses the question "What is the matter made of?" at the most fundamental level. The current theory of elementary particles and their interactions is called the Standard Model (SM) (see [4] for pedagogical overview). The Standard Model is formulated in the language of quantum field theory. In this framework, elementary particles are excitations of quantum fields that also govern their interaction. In order to create new particles, i.e. excite the fields, one needs to force the interaction between them. This can be achieved by collision of already existing particles which requires tremendous amount of energy. That is why elementary particle physics is also called high energy physics. The Standard Model describes three of the four known fundamental forces, namely the electromagnetic force, the weak force, and the strong force, where the weak and the electromagnetic force are unified in the electroweak interaction. The fourth force,

gravity, is not described by the model because it is not possible to extend the Standard Model to general relativity [4].

In every theory where space has three dimensions and quantum mechanics and special relativity are obeyed, like in the Standard Model, all particles must be either fermions or bosons. This is known as the spin-statistics theorem [5]. Fermions are particles that follow Fermi-Dirac statistics, i.e. obey the Pauli exclusion principle, while bosons are particles that follow Bose-Einstein statistics. It also follows from the spin-statistics theorem that particles with integer spin are bosons, while particles with half-integer spin are fermions. As a consequence of the Pauli exclusion principle, only one fermion can occupy a particular quantum state at any given time. Fermions are usually associated with ordinary matter, whereas bosons are generally force carrier particles. In particular, the Standard Model has two types of bosons: gauge and the Higgs boson. The gauge bosons are responsible for mediating the interaction between the elementary particles: the gluon ( $g$ ) for the strong force; the  $W$  and  $Z$  bosons ( $W^+$ ,  $W^-$ ,  $Z$ ) for the weak force, and the photon ( $\gamma$ ) for the electromagnetic force. The Higgs boson ( $H$ ) on the other side has a unique role in the Standard Model. It is responsible for the generation of particle masses. All gauge bosons are vector particles with spin 1, while the Higgs boson is the only scalar particle with the spin 0 in the Standard Model. Fermions are classified into two categories: quarks and leptons. There are six quarks: up ( $u$ ), down ( $d$ ), charm ( $c$ ), strange ( $s$ ), top ( $t$ ), bottom ( $b$ ) and six leptons: electron ( $e$ ), electron neutrino ( $\nu_e$ ), muon ( $\mu$ ), muon neutrino ( $\nu_\mu$ ), tau ( $\tau$ ), tau neutrino ( $\nu_\tau$ ). An important fermion characteristic is that each has its own corresponding antiparticle, i.e. a particle with the same quantum numbers but opposite physical charges. All fermions in the Standard Model are subject to the weak force, especially neutrinos who are only interacting through the weak force, which makes them very difficult to detect. While both the quarks and the leptons, excluding neutrinos, interact electromagnetically, only quarks are interacting via the strong force. One of the most interesting phenomenon emerging from the strong force is color confinement, which says that quarks and gluons cannot be isolated and therefore cannot be directly observed in normal conditions [6, 7]. Instead, they form hadrons. The two main types of hadrons are mesons (one quark, one antiquark) and baryons (three quarks). In Figure 1.1 list of all elementary particles, fermions (orange) and bosons (green), with their interactions is given.

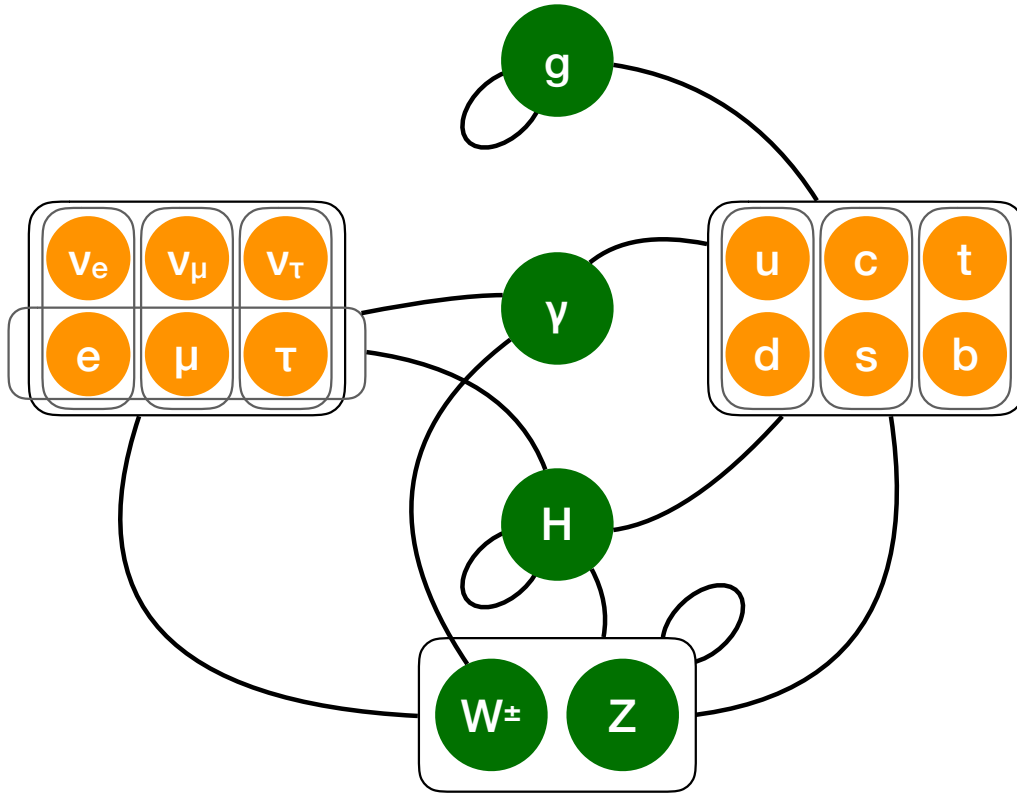


Figure 1.1: A diagram summarizing interactions between elementary particles according to the Standard Model.

### 1.1.1 Historical background

The term Standard Model was first proposed by Abraham Pais and Sam Treiman in 1975, with reference to the electroweak theory with four quarks [8]. Initial experimental discoveries that actually paved the way for the Standard model appeared in early 1930s. In a period of almost 50 years, there were both significant experimental and theoretical breakthroughs. Discovery of positron [9] and neutron [10] in 1932, muon in 1936 [11], neutrino in 1956 [12] and many other particles, e.g. pion [13] were experimental triumphs that shaped theoretical ideas. The first step towards the Standard Model was Sheldon Glashow's discovery in 1961 of a way to combine the electromagnetic and weak interactions [14]. In 1967, Steven Weinberg [15] and Abdus Salam [16] incorporated the Higgs mechanism [16–19] into Glashow's electroweak interaction, giving it its modern form. The neutral weak currents caused by  $Z$  boson exchange were discovered at CERN in 1973 [20–22], while the  $W^\pm$  bosons were discovered in 1983 [23, 24]. The theory of the

strong interaction acquired its modern form in early 1970s when asymptotic freedom was proposed [6, 7]. At the same time there were experimental evidences that confirmed that hadrons were composed of fractionally charged fermions, called quarks [25]. The final confirmation of the theory occurred in 2012 with the discovery of the Higgs boson [26, 27]. There is a vast amount of knowledge and anecdotes behind the Standard Model and it is simply impossible to mention everything in such a short introduction. Complete historic overview of elementary particles is given in [4]. For the last couple of decades, particle physics held dominant position in physics, which is proved by more than 30 Nobel prizes in the field.

### 1.1.2 Theoretical aspects

The development of the Standard Model was driven by theoretical and experimental physicists. From the theoretical point of view, quantum field theory provides the mathematical framework for the Standard Model in which a Lagrangian describes the dynamics and kinematics of the theory. Each kind of particle is described in terms of a field that is defined in space-time. The construction of the Standard Model proceeds first by postulating a set of symmetries of the system and then by writing down the most general renormalizable Lagrangian from its particle content that obeys these symmetries. The global Poincare symmetry is postulated, i.e. translational symmetry, rotational symmetry and Lorentz invariance. The Standard Model is a gauge theory, based on the symmetry group  $SU(3)_C \times SU(2)_L \times U(1)_Y$ . In other words local gauge symmetry is an internal symmetry that essentially defines it. The three factors of the gauge symmetry give rise to the three fundamental interactions. The fields belong to different representations of the various symmetry groups of the Standard Model [28]. Quantum Chromodynamics (QCD), the gauge field theory that describes the strong interactions of colored quarks and gluons, is the  $SU(3)_C$  component.

The Lagrangian of QCD is given by

$$\mathcal{L}_{QCD} = \sum_q \bar{\psi}_{q,a} (i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t_{ab}^C \mathcal{A}_\mu^C - m_q \delta_{ab}) \psi_{q,b} - \frac{1}{4} F_{\mu\nu}^A F^{A\mu\nu} \quad (1.1)$$

where repeated indices are summed over. The  $\gamma^\mu$  are the Dirac  $\gamma$ -matrices. The  $\psi_{q,a}$  are quark-field spinors for a quark of flavor  $q$  and mass  $m_q$ , with a color-index  $a$  that runs from  $a = 1$  to  $N_c = 3$ , i.e. quarks come in three “colors.” Quarks are said to be in the fundamental representation of the  $SU(3)_C$  color group<sup>1</sup>. The  $\mathcal{A}_\mu^C$  correspond to the gluon fields, with  $C$  running from 1 to 8, i.e. there are eight kinds of gluons. Gluons transform under the adjoint representation of the  $SU(3)_C$  color group. The  $t_{ab}^C$  correspond to eight  $3 \times 3$  matrices and are the generators of the  $SU(3)_C$  group. They encode the fact that a gluon’s interaction with a quark rotates the quark’s color in  $SU(3)_C$  space. The quantity  $g_s$  is the QCD coupling constant. Finally, the field tensor  $F_{\mu\nu}^A$  is given by

$$F_{\mu\nu}^A = \partial_\mu \mathcal{A}_\nu^A - \partial_\nu \mathcal{A}_\mu^A - g_s f_{ABC} \mathcal{A}_\mu^B \mathcal{A}_\nu^C \quad [t^A, t^B] = i f_{ABC} t^C \quad (1.2)$$

where the  $f_{ABC}$  are the structure constants of the  $SU(3)_C$  group. The free parameters of QCD are the coupling constant  $g_s$  or  $\alpha_s = g_s^2/4\pi$  and the quark masses  $m_q$ . There is freedom for an additional CP-violating term to be present in the QCD Lagrangian,  $\propto \theta_{QCD} F_{\mu\nu}^A \tilde{F}^{A\mu\nu}$  where  $\tilde{F}^{A\mu\nu}$  is the dual of the gluon field tensor. The coupling constant  $\alpha_s$  at a given momentum transfer scale  $Q^2$  can be parameterized as

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)/12\pi(11N_c - 2N_f)\ln(Q^2/\mu^2)} \quad (1.3)$$

where  $N_f = 3$  is the number of quark flavours and  $\mu^2$  is a renormalization energy scale [4]. For large  $Q^2$  values, i.e. at high energies, the coupling becomes small, which is known as asymptotic freedom. In this regime, quarks can be described as almost free particles, which is the basis of the perturbative approach to QCD calculation. For low  $Q^2$  values the coupling approaches unity. If a pair of quarks begin to separate, the exchanged gluons interact with each other and the strong coupling constant increases. This increasing force either binds the quarks together or it breaks when the energy density of the color field between the quarks is large enough to create an additional pair of quarks. The final state is made only by colorless bound states, known as hadrons. This effect is known as color confinement. It justifies the non-observation of free quarks and gluons. The energy scale  $\Lambda_{QCD}$ , at which the coupling constant becomes large and the perturbative approach

---

<sup>1</sup>The definition of group theory concepts can be found e.g. in [29]

breaks is given by

$$\Lambda_{QCD}^2 = \mu^2 \exp \left( \frac{-12\pi}{(11N_c - 2N_f)\alpha_s(\mu^2)} \right) \quad (1.4)$$

and consequently

$$\alpha_s(Q^2) = \frac{12\pi}{(11N_c - 2N_f) \ln(Q^2/\Lambda_{QCD}^2)} \quad (1.5)$$

The value of  $\Lambda_{QCD}$  is important when it comes to phenomenological description of hadron interactions. In typical hadron-hadron collisions, where  $Q^2 \gg \Lambda_{QCD}$  (hard scattering), the amplitude of the process is calculated with the perturbative QCD tools; the interaction products lose energy until the particle energies are of the order of  $Q^2 \approx \Lambda_{QCD}$  and the non-perturbative process dominates resulting in neutralization of colored quantum numbers and the production of non-colored states through the so-called hadronization process. In the formed hadrons, the quark color confinement happens again through the strong interaction. The value of  $\Lambda_{QCD}$  is not predicted from the theory and it is determined from experiments, which is approximately 100 MeV [28]. Free quarks have never been observed, which is understood to be a consequence of a long-distance confining property of the strong QCD force. As already mentioned, all quarks, with exception of top quark, hadronize, i.e. become part of a meson or baryon, on a timescale  $\approx 1/\Lambda_{QCD}$ ; the top quark instead decays before it has time to hadronize. This means that the question of what one means by the quark mass is a complex one, which requires that one adopts a specific prescription. A perturbatively defined prescription is the pole mass,  $m_q$ , which corresponds to the position of the divergence of the propagator. This is close to the physical picture of mass. However, when relating it to observable quantities, it suffers from substantial non-perturbative ambiguities [30].

The Standard Model of electroweak interactions (EW) is based on the gauge group  $SU(2)_L \times U(1)_Y$ , with gauge bosons  $W_\mu^i, i = 1, 2, 3$  and  $B_\mu$  corresponding to the  $SU(2)_L$  and  $U(1)_Y$  symmetries, respectively, and the associated gauge coupling constants  $g$  and  $g'$ . The left-handed fermion fields of the  $i^{th}$  fermion family transform as doublets  $\Psi_i = \begin{pmatrix} \nu_i \\ l_i^- \end{pmatrix}_L$  and  $\begin{pmatrix} u_i \\ d_i' \end{pmatrix}_L$  under  $SU(2)_L$ , where  $d_i' = \sum_j V_{ij} d_j$ , and  $V$  is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix.  $CP$  violation is incorporated into the EW model by a single observable complex phase in  $V_{ij}$ . Four independent parameters are required to fully define the CKM matrix. Many different parameterizations have been proposed. In this work the Wolfenstein parametrization is used [31]. The right-handed fields are  $SU(2)_L$  singlets,  $\psi_i$ .

A complex scalar Higgs doublet  $\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ , is added to the model for mass generation through spontaneous symmetry breaking with a potential given by

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2 \quad (1.6)$$

For  $\mu^2$  negative,  $\phi$  has a vacuum expectation value,  $v/\sqrt{2} = \frac{\mu}{\lambda}$ , where  $v \approx 246$  GeV, breaking part of the electroweak gauge symmetry, after which only one neutral Higgs scalar,  $H$ , remains in the physical particle spectrum. After the symmetry breaking the Lagrangian for the fermion fields,  $\psi_i$ , is

$$\begin{aligned} \mathcal{L}_F = & \sum_i \bar{\psi}_i (i\gamma^\mu \partial_\mu - m_i - \frac{m_i H}{v}) \psi_i \\ & - \frac{g}{2\sqrt{2}} \sum_i \bar{\Psi}_i \gamma^\mu (1 - \gamma^5) (T^+ W_\mu^+ + T^- W_\mu^-) \Psi_i \\ & - e \sum_i Q_i \bar{\psi}_i \gamma^\mu \psi_i A_\mu \\ & - \frac{g}{2\cos\theta_W} \sum_i \bar{\psi}_i \gamma^\mu (g_V^i - g_A^i \gamma^5) \psi_i Z_\mu \end{aligned} \quad (1.7)$$

Here  $\theta_W \equiv \arctan(g'/g)$  is the weak-mixing angle,  $e = g \sin \theta_W$  is the positron electric charge, and  $A_\mu \equiv B_\mu \cos \theta_W + W_{3\mu} \sin \theta_W$  is the photon field  $\gamma$ .  $W^\pm \equiv \frac{1}{\sqrt{2}}(W_1 \pm iW_2)$  and  $Z \equiv -B \sin \theta_W + W_3 \cos \theta_W$  are the charged and neutral weak boson fields, respectively. The Yukawa coupling of  $H$  to  $\psi_i$  in the first term in  $\mathcal{L}_F$ , which is flavor diagonal, can be written as  $gm_i/2M_W$ . The boson masses in the EW sector at lowest order in perturbation theory are given by

$$M_H = \lambda v, \quad M_W = \frac{1}{2}gv = \frac{ev}{2\sin\theta_W}, \quad M_Z = \frac{1}{2}\sqrt{g^2 + g'^2}v = \frac{ev}{2\sin\theta_W \cos\theta_W}, \quad M_\gamma = 0 \quad (1.8)$$

The second term in  $\mathcal{L}_F$  represents the charged-current weak interaction, where  $T^+$  and  $T^-$  are the weak isospin raising and lowering operators. The third term in  $\mathcal{L}_F$  describes electromagnetic interactions (QED) and the last describes the weak neutral-current interaction. The vector and axial-vector couplings are

$$g_V^i \equiv t_{3L}(i) - 2Q_i \sin^2 \theta_W, \quad g_A^i \equiv t_{3L}(i) \quad (1.9)$$



where  $t_{3L}(i)$  is the weak isospin of fermion  $i$  ( $+1/2$  for  $u_i$  and  $\nu_i$ ;  $-1/2$  for  $d_i$  and  $l_i$ ) and  $Q_i$  is the charge of  $\psi_i$  in units of  $e$ . Without the Higgs boson, the cross section for various weak processes violates unitarity, i.e. the probability of certain interactions is larger than unity. The Higgs boson, amongst other things, adds counter-acting interactions that prevent any violation of unitarity.

In particle physics one is interested in computing probabilities for different outcomes when incoming particles interact. To describe and calculate scattering cross-sections, the S-matrix formalism is used [32]. The S-matrix is closely related to the transition probability amplitude in quantum mechanics. In the Hamiltonian approach to QFT, the S-matrix can be calculated as a time-ordered exponential of the integrated Hamiltonian in the interaction picture. The number of times the interaction Hamiltonian acts is the order of the perturbation expansion. The leading-order (LO) term is the one with the largest order of magnitude, smaller ones are known as next-to leading order terms (NLO), next-to-next-to-leading order terms (NNLO), etc. The perturbative calculation of the S-matrix leads to famous Feynman diagrams [33], which are graphical representation of a perturbative contribution to the transition amplitude, invaluable tool for particle physics.

### 1.1.3 Value of Standard Model parameters

The Standard Model Lagrangian depends on 19 parameters, whose numerical values are established by experiment [28]. They are summarized in Table 1.1.

Table 1.1: Free parameters of the Standard Model with their measured values.

Parameter	Description	Measured value	Unit
$m_u$	Up quark mass	$2.2^{+0.6}_{-0.4}$	MeV
$m_d$	Down quark mass	$4.7^{+0.5}_{-0.4}$	MeV
$m_c$	Charm quark mass	1.28(3)	GeV
$m_s$	Strange quark mass	$96^{+8}_{-4}$	MeV
$m_t$	Top quark mass	173.1(6)	GeV
$m_b$	Bottom quark mass	$4.18^{+0.04}_{-0.03}$	GeV
$m_e$	Electron mass	510.9989461(31)	keV
$m_\mu$	Muon mass	105.6583745(24)	MeV
$m_\tau$	Tau mass	1776.82(16)	MeV
$\lambda$	Wolfenstein parameters	0.22496(48)	-
$A$		0.823(13)	-
$\rho$		0.141(19)	-
$\eta$		0.349(12)	-
$m_H$	Higgs boson mass	125.09(24)	GeV
$v$	Fermi coupling constant	$1.1663787(6) \cdot 10^{-5}$	$\text{GeV}^{-2}$
$\alpha$	Fine structure constant	$7.2973525664(17) \cdot 10^{-3}$	-
$\alpha_S$	Strong coupling constant	0.1182(12)	-
$\sin^2(\theta_W)$	Weak-mixing angle	0.23129(5)	-
$\theta_{QCD}$	QCD vacuum angle	$< 10^{-10}$	-

In the Standard Model, neutrinos are assumed to be massless. However, experiments showed [34–36] that neutrinos cannot be massless. The Standard Model can be extended to accommodate finite neutrino masses by introducing 3 neutrino masses and 4 neutrino mixing angles that relate the physical neutrino states to lepton flavour eigenstates.

## 1.2 Physics beyond the Standard Model

The Standard Model is an extremely successful theory, yet it leaves many theoretical and experimental questions unanswered. Various forms of "Beyond the Standard Model" new physics are proposed that would modify the Standard Model in such a way to be consistent with existing data but able to solve remaining issues:

- **Hierarchy problem** - Within the Standard Model, the mass of the Higgs boson gets some very large quantum corrections due to the presence of virtual particles (mostly virtual top quarks). These corrections are much larger than the actual mass of the Higgs boson. Which means that the bare mass parameter of the Higgs mass in the Standard Model must be fine tuned in such a way to almost completely cancel the quantum corrections.
- **Number of parameters** - The Standard Model depends on 19 numerical parameters. Their values are known from experiment, but the origin of the values is unknown.
- **Quantum triviality** - In a quantum field theory, charge screening can restrict the value of the observable charge of a classical theory. If the only resulting value of the renormalized charge is zero, the theory is said to be noninteracting. Thus, surprisingly, a classical theory that appears to describe interacting particles can, when realized as a quantum field theory, become a theory of noninteracting free particles. This phenomenon is referred to as quantum triviality. It may not be possible to create a consistent quantum field theory involving elementary scalar particles [37].
- **Strong CP problem** - According to QCD there could be a violation of CP symmetry in the strong interactions. However, no violation of the CP-symmetry is known to have occurred in experiments. As there is no known reason for it to be conserved in QCD specifically, this is a "fine tuning" problem known as the strong CP problem.
- **Self-consistency** - Self-consistency of the Standard Model has not been mathematically proven. A key question is the Yang–Mills existence and mass gap problem [38].
- **Neutrino mass** - Neutrinos are massless particles according to the Standard Model. However, neutrino oscillation experiments have shown that neutrinos do have mass

[34–36]. Mass terms for the neutrinos can be added by hand, but these lead to new theoretical problems. Nature of their mass is not understood, i.e. are they Dirac or Majorana particles.

- Standard model of cosmology - cosmological observations tell us that the Standard Model explains about 4 percent of the energy present in the universe. Of the missing 96 percent, about 24 should be dark matter, which would behave just like the other matter, but which only interacts weakly with the Standard Model fields. Yet, the Standard Model does not supply any fundamental particles that are good dark matter candidates.
- Gravity - For decades, the formulation of a quantum theory of gravity has been a major challenge in theoretical physics. A lot of effort has been spent in building theoretical models unifying gravity and quantum mechanics. It is however impossible to this day to know whether any of these attempts is a correct description of our physical world. While this fact is embarrassing, it has not been problematic for particle physics and the Standard Model, as gravity is much too feeble to play a significant role in elementary particle interactions. At extremely high energies (Planck scale:  $10^{19}$  GeV) quantum gravity must come into the picture, with the expected unification of all forces of nature. Currently, the best candidate for unified theory is String theory [39].
- Matter/Antimatter symmetry - The universe is exclusively made out of matter. However, matter and anti-matter should be created in the equal amounts in the Big Bang. The observed asymmetry can be explained only if the laws of physics do not treat matter and antimatter particles in the same way, resulting in a violation of the CP symmetry. Such a violation of the CP symmetry is indeed contained in the Standard Model through the complex phase of the CKM matrix. The size of this CP asymmetry present in the Standard Model is however considered to be much too small to explain the observed matter-antimatter asymmetry in the universe [40].



# Chapter 2

## The Higgs boson phenomenology at hadron colliders

This chapter is devoted to the introduction of concepts related to the hadron collider and the Higgs boson phenomenology. The procedure to connect the physical observables and partons is discussed. The production and decay mechanisms of the Higgs boson are described with a focus on definition of boosted topologies.

### 2.1 Phenomenology of proton-proton interactions

According to the parton model, all hadrons are made of valence and sea quarks. Valence quarks contribute to the hadron quantum numbers, e.g. the proton has two up and one down quark, while sea quarks, are virtual quark-antiquark ( $q\bar{q}$ ) pairs which form when a gluon splits. This process is also reversible and as a result there is a constant flux of gluon splits and creations, so-called "sea". In proton-proton collisions, the interaction generally involves a pair of partons: valence quarks, sea quarks or gluons. It is therefore likely that any of these partons (quarks  $q$  and  $q'$  or gluon  $g$ ) interact, making a  $qq$ ,  $q\bar{q}$ ,  $qq'$ ,  $q\bar{q}'$ ,  $gq$ ,  $g\bar{q}$  or  $gg$  interaction possible. Moreover, only a fraction  $x_1$  and  $x_2$  of the proton momenta ( $P_1$ ,  $P_2$ ) is carried by the colliding parton momenta  $p_1$  and  $p_2$

$$p_{1,2} = x_{1,2} \cdot P_{1,2} \tag{2.1}$$

If  $\sqrt{s}$  is the energy in the center of mass frame of the incident particles, the center of mass energy of the partonic interaction is given by  $\sqrt{s'} = \sqrt{s x_1 x_2}$  and varies event by event. While this represents an experimental difficulty, it allows to explore a wider range of energies than lepton colliders. That is why hadron colliders are better suited for discoveries and lepton colliders for precise measurements. As already mentioned in the previous chapter, gluons and quarks interact via strong force (QCD) and the corresponding coupling constant depends on the scale of the process. For high momentum transfers, it becomes sufficiently small to make perturbative expansion. In order to do calculations within QCD, one needs the factorization theorem [41]. The theorem is introduced to separate the cross section in two parts: the contribution from the hard process calculated using perturbative QCD and the contribution from the internal structure of the proton. Therefore, hard scattering between partons is independent from the proton internal structure. The production cross section of a final state  $X$  in a process  $(A + B \rightarrow X)$ , where  $a_i$  and  $b_j$  are the partons that constitute the hadrons  $A$  and  $B$  can be written as [42]

$$\sigma_{AB \rightarrow X} = \sum_{i,j} \int dx_{a_i} dx_{b_j} f_{A/a_i}(x_{a_i}, \mu_F^2) f_{B/b_j}(x_{b_j}, \mu_F^2) \hat{\sigma}_{a_i b_j \rightarrow X}(\mu_R^2, \mu_F^2) \quad (2.2)$$

where

- $x_1$  and  $x_2$  are the fractions of the initial momentum carried by the partons  $a_i, b_j$
- $f_{A/a_i}(x_{a_i}, \mu_F^2)$ ,  $f_{B/b_j}(x_{b_j}, \mu_F^2)$  are the parton distribution functions (PDFs) which describe probability densities for a parton  $i$  with momentum  $x_p$  to be found within the colliding hadrons (A,B). The PDFs are expressed in terms of momentum fraction  $x_{a_i}, x_{b_j}$  and factorization scale  $\mu_F$ . The factorization scale is introduced to take care of soft singularities. Depending on the emitted parton's transverse momentum, there are two cases. If it is higher than the scale  $\mu_F$ , it is absorbed into the PDF, else it is part of the short distance cross section. The PDFs cannot be determined by perturbative QCD. They are instead parametrized from experimental data collected on many experiments, the Tevatron at Fermilab [43], the hadron-electron colliders such as HERA at DESY [44], with additional updates from the study of tau decays [45] and quarkonium measurements [46]. Their evolution with the factorization scale is computed by the DGLAP equations [47–50], allowing to use the PDFs measured

at one scale to predict the results of experiments at other scales. The PDFs of the MSTW 2008 parametrization [51] for the quark and gluon constituents of a proton at two different values of momentum transfer are shown in Figure 2.1. The up and down components are dominated by valence quarks at high  $x$ , while the strange, charm, bottom, corresponding anti-quarks, and gluons are more prevalent at low  $x$ . The gluon fraction at low  $x$  increases by several orders of magnitude at large momentum transfers.

- The  $\hat{\sigma}_{a_i b_j \rightarrow X}$  term corresponds to the partonic cross section for the production of a  $X$  state from a pair of partons  $a_i, b_j$ . It is normally computed at a given order of perturbative QCD and at a renormalization scale  $\mu_R$ . The renormalization scale  $\mu_R$  is another arbitrary parameter introduced to treat a class of infinities where loop momenta can be very large.

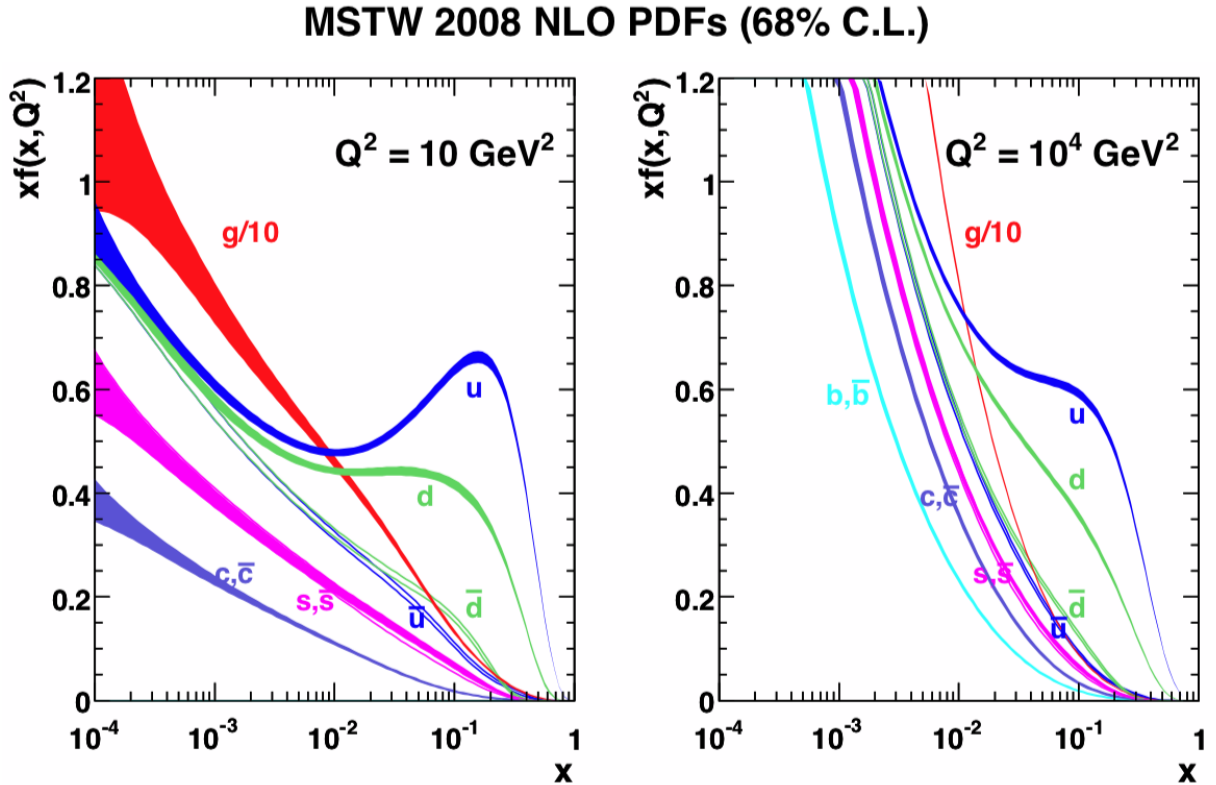


Figure 2.1: Parametrization of the PDFs of the proton in the MSTW 2008 fit, at different scales. On the x-axis the longitudinal momentum fraction  $x$  is shown, while on the y-axis the product  $x \cdot f(x, Q^2)$ , where  $f(x, Q^2)$  is the parton distribution function. Different partons are shown in different colors, including uncertainties.



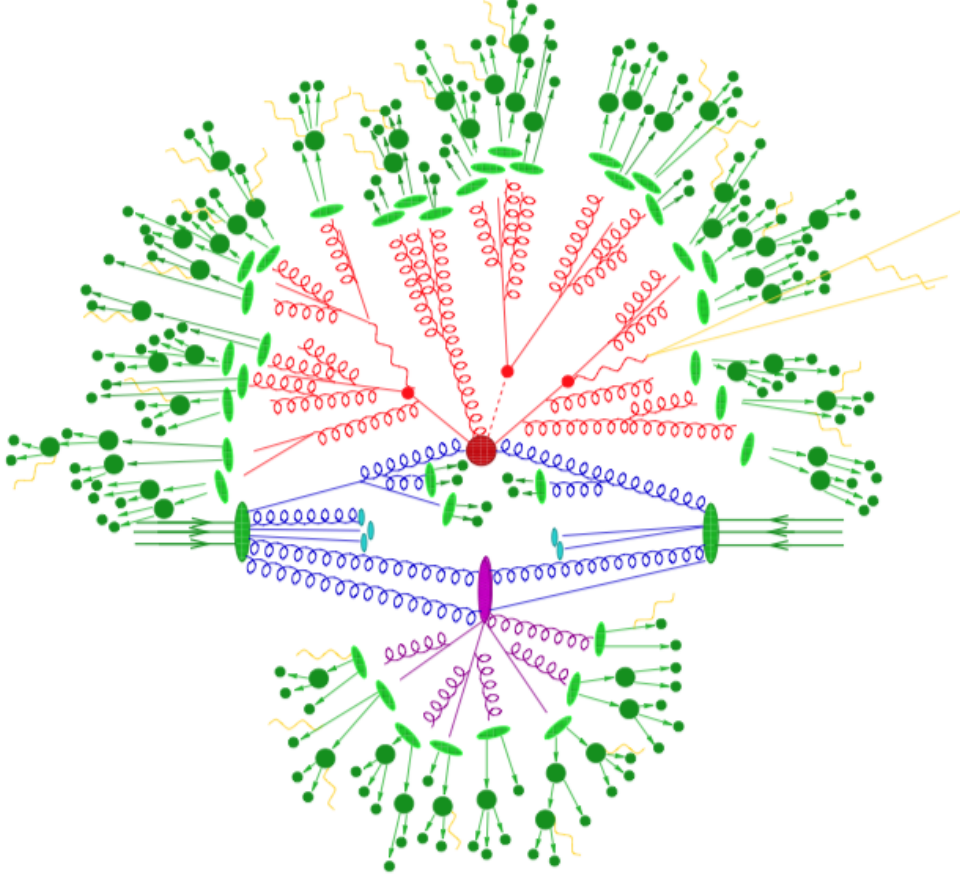


Figure 2.2: Schematic representation of an event as produced by an event generator. The hard interaction (big red circle) is followed by the decay of both top quarks and the Higgs boson (small red circles). Additional hard QCD radiation is produced (red) and a secondary interaction takes place (purple blob) before the final-state partons hadronize (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow).

After hard scattering, each final state parton can lose its energy through irradiation of gluons. Emitted gluons fragment into a parton showers of  $q\bar{q}$  pairs and additional gluons. The process continues until a low threshold energy is reached, and the hadronization process takes place. Hadronization process is described with more details in Section 5.6. Schematic description of the event generation is shown in Figure 2.2 [52]. The process cannot be treated in perturbative way because of low energy exchanged. Therefore, the dynamics of the process is described by fragmentation functions that represent the probability that a parton can fragment into a hadron in the final state. The cross section can

be modified in order to describe the  $A + B \rightarrow C + X$  process:

$$\hat{\sigma}_{a_i b_j \rightarrow C+X} = \int dz_C D_{c_k}(z_C, \mu_f^2) \hat{\sigma}_{a_i b_j \rightarrow c_k+X}(\mu_R^2, \mu_F^2) \quad (2.3)$$

where  $C$  is a hadron,  $D_{c_k}(z_C, \mu_f^2)$  is the fragmentation function that represents the probability that the produced parton  $c_k$  gives a final state particle  $C$  with momentum fraction  $z_C$  during the fragmentation process at some fragmentation scale  $\mu_f$ . The fragmentation scale is introduced with the same motivation as the factorization scale and under similar prescription, to absorb singularities due to final state collinear radiation. The fragmentation functions are not calculable, as the PDFs, but it is possible to calculate their dependence on the scale. Finally, the production cross section becomes

$$\sigma_{AB \rightarrow C+X} = \sum_{i,j,k} \int dx_{a_i} dx_{b_j} f_{A/a_i}(x_{a_i}, \mu_F^2) f_{B/b_j}(x_{b_j}, \mu_F^2) \hat{\sigma}_{a_i b_j \rightarrow C+X}(\mu_R^2, \mu_F^2, \mu_f^2) \quad (2.4)$$

Factorization and renormalization procedures have similar purpose but at opposite ends of the energy range of perturbative QCD. Both scale parameters  $\mu_F$  and  $\mu_R$  are not intrinsic to QCD. Their value should be carefully set to be of the order of the hard scale that characterizes the parton-parton interaction. The dependence on their value is weaker as more terms are included in the perturbative expansion. Since they are totally arbitrary, any physical observable must be independent from their particular choice. The simplifying assumption of a single scale  $\mu = \mu_F = \mu_R = \mu_f$  is often made, and the standard choice is  $\mu^2 = Q^2$ , the hard scattering scale.

## 2.2 The Higgs boson production in proton-proton collisions

The main production mechanism at high energy hadron colliders are gluon fusion, vector boson fusion, quark fusion and associated production with a vector boson, also known as the Higgs Strahlung. Representative diagrams at leading-order for these processes are shown in Figure 2.3.

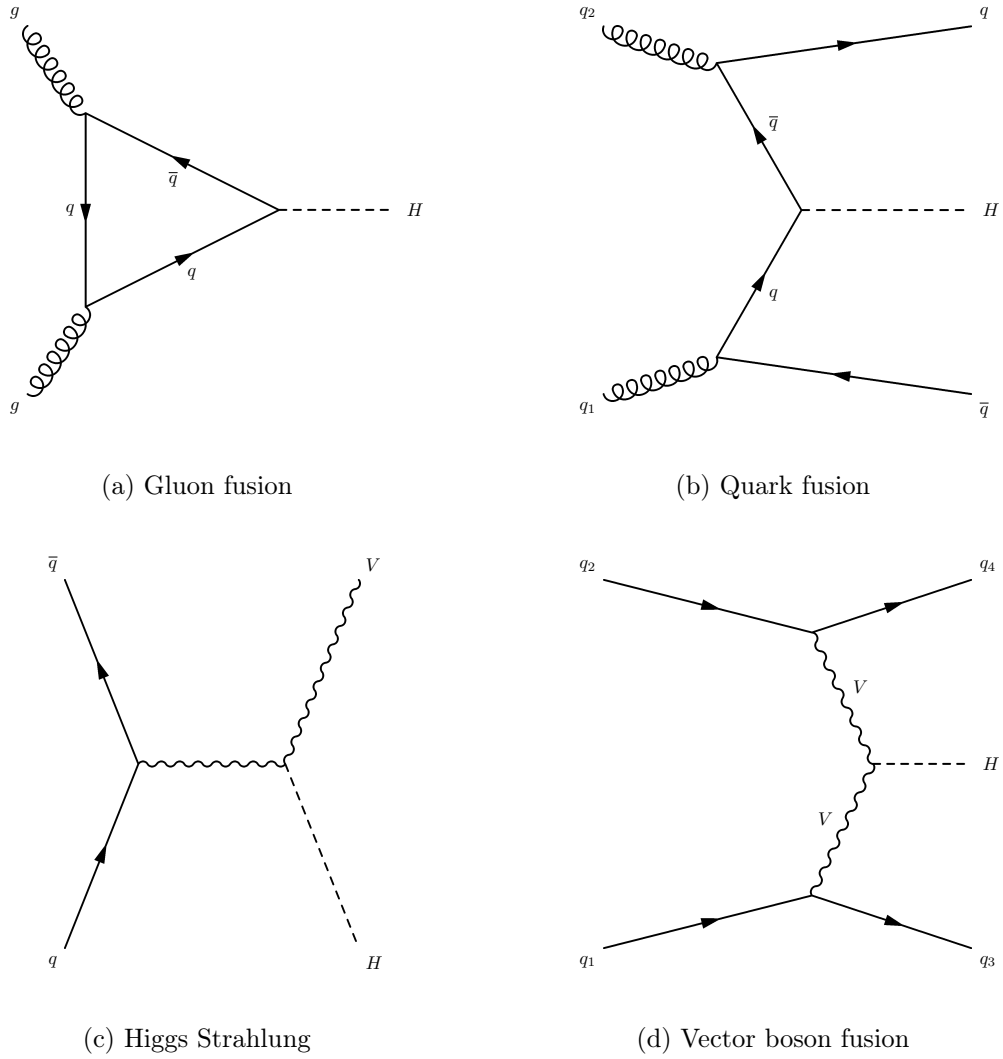


Figure 2.3: Leading-order Feynman diagrams contributing to the Higgs boson production.

The Higgs boson production mechanism with the largest cross section is the gluon fusion. The dominant contribution comes from the exchange of a virtual, top quark while contributions from lighter quarks propagating in the loop are suppressed due to their lower masses. The mechanism with the second-largest cross section is vector boson fusion (VBF), in which the scattering of two (anti-)quarks is mediated by exchange of vector bosons, from which the Higgs boson is radiated off. Because of the color-singlet nature of the vector boson exchange, gluon radiation from the central-rapidity regions is strongly suppressed, resulting in a large rapidity gap between the Z jets coming from the scattered quarks. These characteristic features of VBF processes can be exploited to distinguish them from overwhelming backgrounds and used as a clean environment not only for Higgs searches but also for the determination of the Higgs boson couplings. The next relevant

production mechanisms are associated production with vector bosons. As neither the Higgs boson nor the vector bosons are stable particles, their decay channels have to be considered. Associated production with vector bosons provide a relatively clean environment for studying the decay of the Higgs boson into bottom quarks. The Higgs boson production in association with  $t\bar{t}$  (Figure 2.3, (b) with  $q = t$ ) provides a direct probe of the top-Higgs coupling. Production in association with a single top quark can bring valuable information, in particular regarding the sign of the top Yukawa coupling. The cross sections for the production of the Higgs boson with their theoretical uncertainties are shown in Figure 2.4 [53].

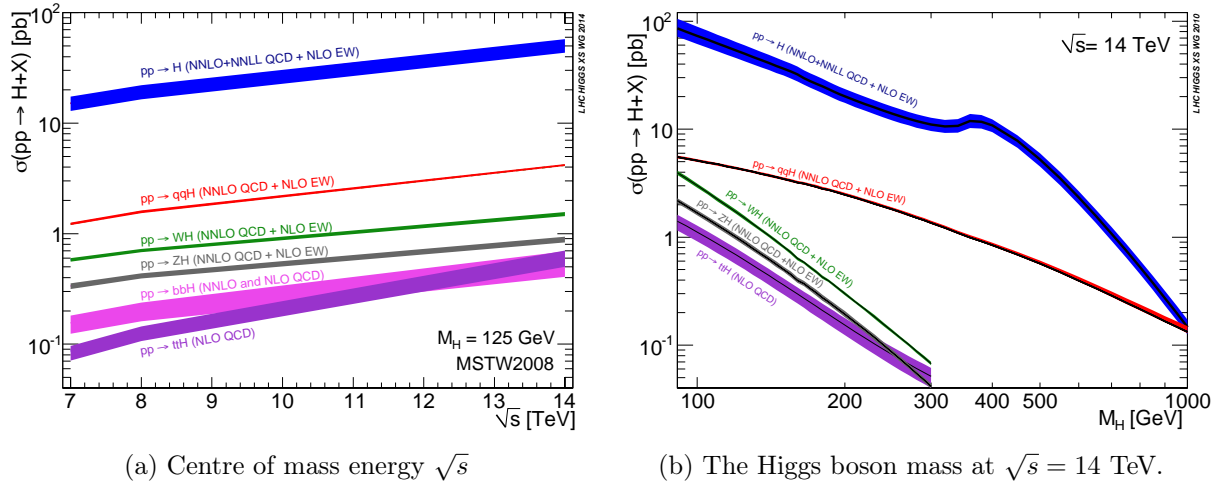


Figure 2.4: The Higgs boson production cross section as a function of the LHC centre of mass energy and its mass at  $\sqrt{s} = 14$  TeV.

## 2.3 The Higgs boson decay

The branching ratio (BR) is the fraction of time a particle decays to a particular final state. In the SM, the branching ratios for the Higgs boson decays in various final states depend on the Higgs boson mass, as shown in Figure 2.5 [53]. The dominant decay modes of the Higgs boson with mass of 125 GeV are  $H \rightarrow b\bar{b}$  and  $H \rightarrow WW^+$ , followed by  $H \rightarrow gg$ ,  $H \rightarrow \tau^+\tau^-$ ,  $H \rightarrow c\bar{c}$  and  $H \rightarrow ZZ^*$ . With much smaller rates follow the Higgs boson decays into  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow \gamma Z$  and  $H \rightarrow \mu^+\mu^-$ . Since the decays into gluons, diphotons and  $Z\gamma$  are loop induced, they provide indirect information on the Higgs boson couplings to  $WW$ ,  $ZZ^*$  and  $t\bar{t}$  in different combinations. The uncertainties in the BR

include the missing higher-order corrections in the theoretical calculations as well as the uncertainties in the SM input parameters, in particular for fermion masses and the QCD gauge coupling. Exact values of BRs are given in Table 2.1. Despite having almost the smallest BR, the Higgs boson has been observed firstly in diphoton and  $ZZ^*$  channel. The Higgs boson decay mode with the largest branching ratio is  $H \rightarrow b\bar{b}$ , however, this channel is challenging due to significant background contributions. Representative diagrams for these processes are shown in Figure 2.6.

Table 2.1: The SM Higgs boson ( $m_H=125$  GeV) branching ratios.

Decay channel	Branching ratio
$H \rightarrow \gamma\gamma$	$2.27 \times 10^{-3}$
$H \rightarrow ZZ^*$	$2.62 \times 10^{-2}$
$H \rightarrow W^-W^+$	$2.14 \times 10^{-1}$
$H \rightarrow \tau^-\tau^+$	$6.27 \times 10^{-2}$
$H \rightarrow b\bar{b}$	$5.84 \times 10^{-1}$
$H \rightarrow Z\gamma$	$1.53 \times 10^{-3}$
$H \rightarrow \mu^-\mu^+$	$2.18 \times 10^{-4}$

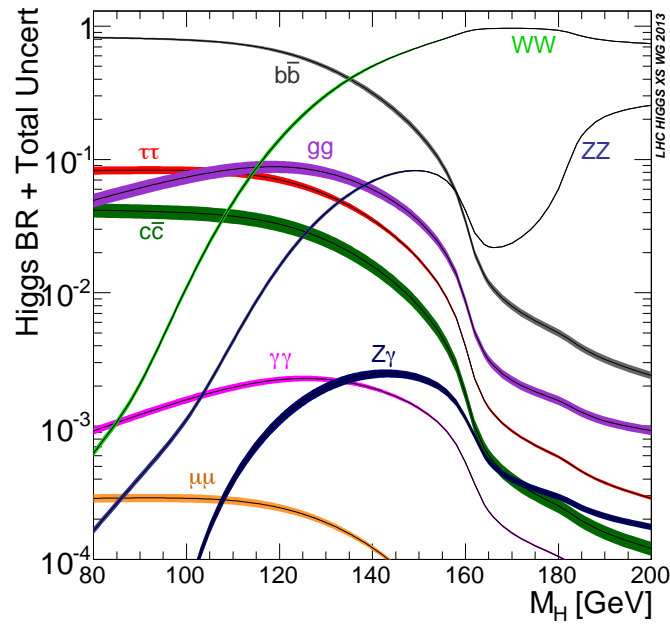
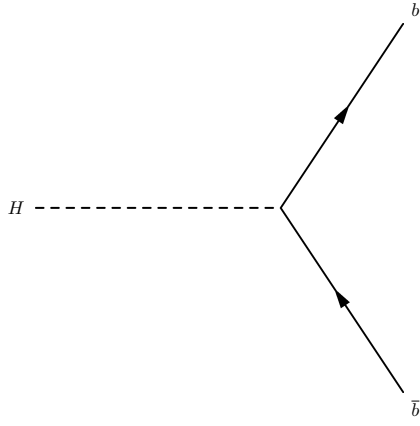
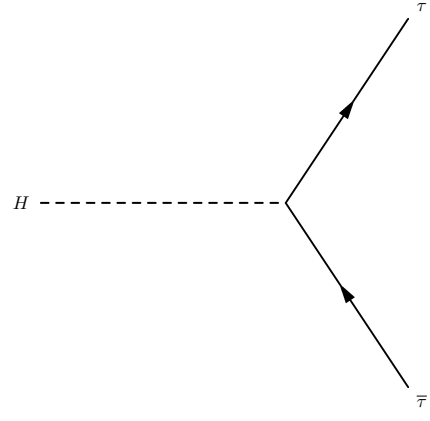


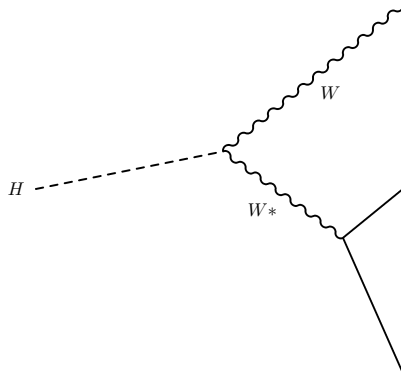
Figure 2.5: The Higgs boson branching ratios as a function of the Higgs boson mass.



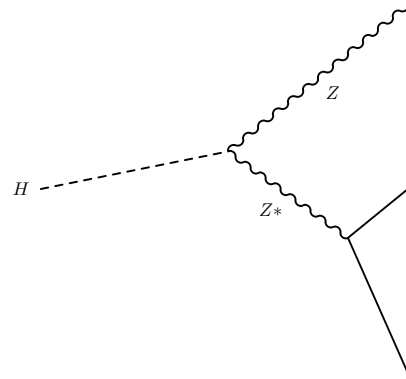
(a)  $H \rightarrow b\bar{b}$



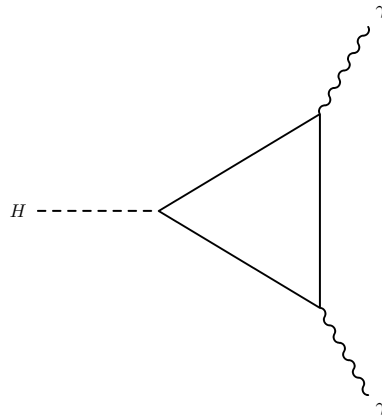
(b)  $H \rightarrow \tau\bar{\tau}$



(c)  $H \rightarrow WW^*$



(d)  $H \rightarrow ZZ^*$



(e)  $H \rightarrow \gamma\gamma$

Figure 2.6: Leading-order Feynman diagrams representing the Higgs boson decay.

## 2.4 Kinematic variables used in proton-proton collisions

The composite nature of the colliding protons (hadrons) means that parton-parton interactions rarely have their centre of mass frame coincident with the detector rest frame. Therefore, it is useful to analyze collisions in terms of variables that are invariant under Lorentz transformation along beam axis. The commonly used kinematic variables to describe final state particles are therefore rapidity  $y$ , transverse momentum  $p_T$  and azimuth angle  $\phi$ . The four-momentum of a particle of mass  $m$  is  $p = (E, p_x, p_y, p_z) = (m_T \cosh y, p_T \sin \varphi, p_T \cos \varphi, m_T \sinh y)$  where  $p_x$ ,  $p_y$  and  $p_z$  are the Cartesian coordinates of the momentum and the transverse mass is defined as  $m_T = \sqrt{p_T^2 + m^2}$ . The rapidity  $y$  is defined by the relation

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (2.5)$$

which is not Lorentz invariant since it transforms as  $y = y + c$  for boosts along beam axis, where  $c$  is the relative velocity between two frames. The rapidity difference  $\Delta y$ , however, is invariant and is used for measuring the angular distance between two objects. In the ultra-relativistic regime, the rapidity can be approximated by the pseudorapidity  $\eta$ , defined as

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right) \approx y \quad (2.6)$$

where  $\theta$  is the angle between the particle and the beam direction ( $\cos \theta = p_z/|p|$ ) and is easily measurable in the detector. The transverse energy  $E_T = E \sin \theta$  is also a useful variable. Finally, the angular distance  $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \varphi)^2}$  is Lorentz invariant variable to measure distance between two objects in azimuth  $\Delta \varphi$  and pseudorapidity  $\Delta \eta$  plane.

## 2.5 Boosted topologies

Butterworth, Davison, Rubin, and Salam (BDRS papers) [1, 2] proposed novel approach where they claim that going into the boosted regime should become an advantage in terms of reducing uninteresting background processes. Their research was focused on the Higgs

boson production in association with a massive electroweak boson in final states with two  $b$  quarks and two leptons. As shown in Figure 2.7 [1], signal (blue color) should be visible due to significant reduction of background processes.

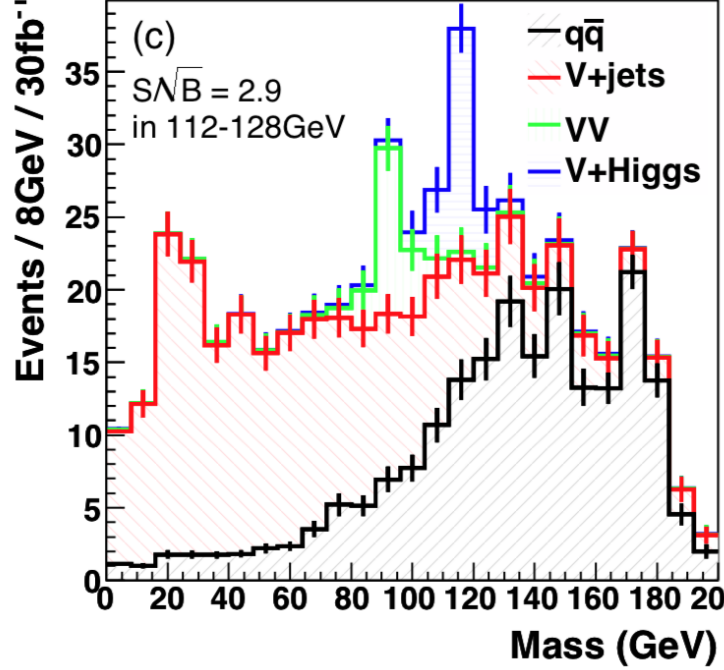


Figure 2.7: Expected signal and background distributions for a 115 GeV SM Higgs in boosted topology ( $W(l\nu)H(bb)$  channel) according to calculation in BDRS paper.

The idea of the boosted topology is the following. When a particle of mass  $m$  and momentum  $p_T$  decays into two particles of momenta  $p_1$  and  $p_2$  as shown in Figure 2.8, angular separation between new particles will vary significantly with  $p_T$  and decay orientation, according to

$$\Delta R \simeq \frac{1}{\sqrt{z(1-z)}} \frac{m}{p_T} \quad (2.7)$$

where  $z$ ,  $1 - z$  are the momentum fractions of a new particles. With increasing value of momentum  $p_T$ , distance between the particles decreases. If background processes have the same particles in final state as the signal process, requirement on high value of  $p_T$  gives implicit constraint on the event topology that is characteristic for signal only and it results in additional background suppression. On the other side, this represents a big challenge from detector point of view, it is more difficult to successfully identify closer particles. BDRS papers were the motivation for the work presented.



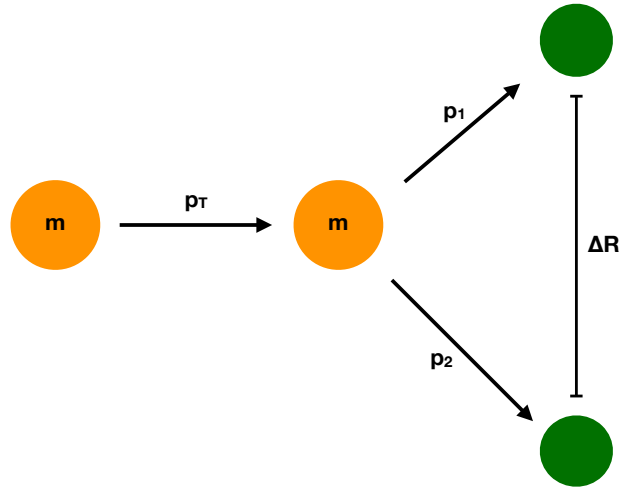


Figure 2.8: Schematic decay of a particle  $m$  with momentum  $p$  into two particles separated by distance  $\Delta R$ .

# Chapter 3

## The Large Hadron Collider

In this chapter an overview of the Large Hadron Collider (LHC) is given, with a focus on its design and technical details including a brief description of all detectors attached to the collider.

### 3.1 Introduction

The Large Hadron Collider (LHC) [54–56] is the most powerful particle collider and the largest single machine beneath the France–Switzerland border near Geneva. It represents the culmination of years of dedicated effort by roughly 10000 scientists and engineers from over 100 countries. The world map of countries involved in research at CERN is shown in Figure 3.1 [57]. The story of the LHC began in the late eighties when it was first proposed. It took several years of serious planning before installation of LHC, which started in 2001 in the pre-existing underground tunnel which housed the Large Electron Positron Collider (LEP) [58]. On 10. 9. 2008, the LHC was officially inaugurated and the new era of high energy physics has started. One of the key motivations for building such an enormous machine was to investigate the possible existence of the Higgs boson, a key part of the Standard Model of physics, which was predicted by theory but had not been observed before. It was estimated that, if the Standard Model were correct, the LHC would be able to produce (at  $\sqrt{s} = 14$  TeV and  $L \approx 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ ) a couple of the Higgs bosons ( $m_H=125$  GeV) every minute. Additionally, the LHC allowed the search for

supersymmetric particles and other hypothetical particles. Nowadays, after 10 years of work one can easily claim that the LHC fulfilled expectations. Its first run discoveries included the long sought Higgs boson [26, 27], several composite particles [59, 60], and the creation of a quark–gluon plasma [61–63]. Making discoveries and being the largest machine comes with a huge price. The total cost of the project is approximately 4.6 billion Swiss francs. Unfortunately, the LHC had to pay the price in terms of human life. On 25 October 2005, José Pereira Lages, a technician, was killed during the LHC construction when a switchgear that was being transported fell on top of him. The LHC gained a substantial amount of attention from outside the scientific community and its progress is followed by most popular science media. The LHC has also influenced works of fiction including TV shows, popular science books, novels, movies and is considered as one of the biggest achievements of humankind.

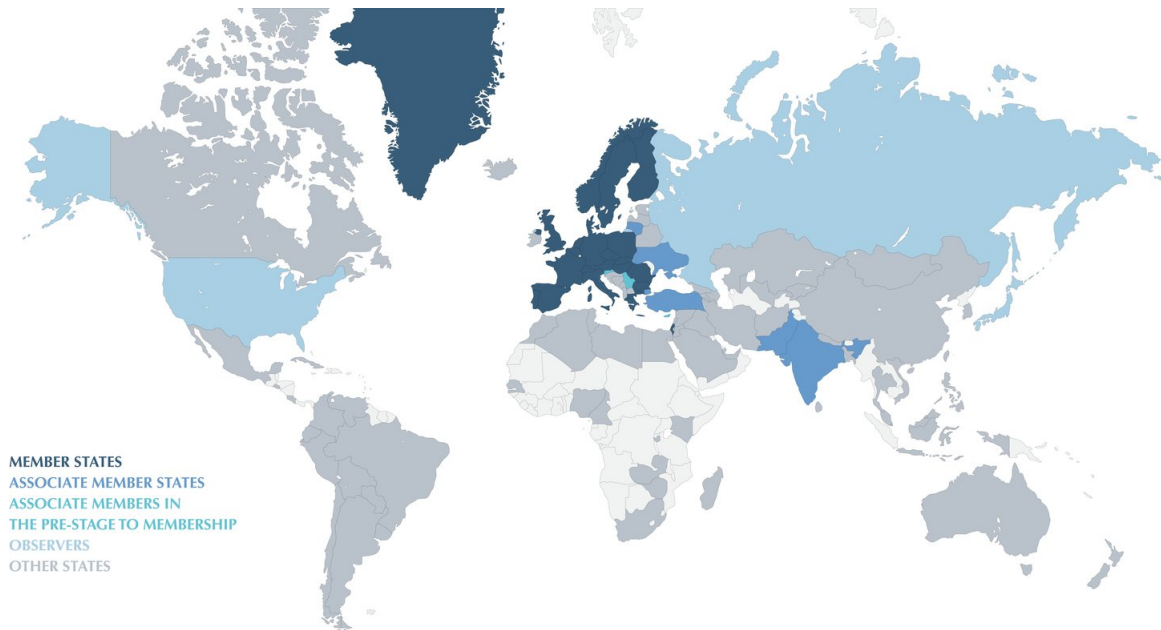


Figure 3.1: World map of countries involved in research at CERN.

## 3.2 Collider design

The transition from the Large Electron Positron Collider to the Large Hadron Collider was motivated by several reasons. Installing the new machine in a pre-existing tunnel substantially reduced the cost and time of construction, colliding hadrons instead of leptons allowed to reach a higher energy in the center of mass frame. Every charged particle,

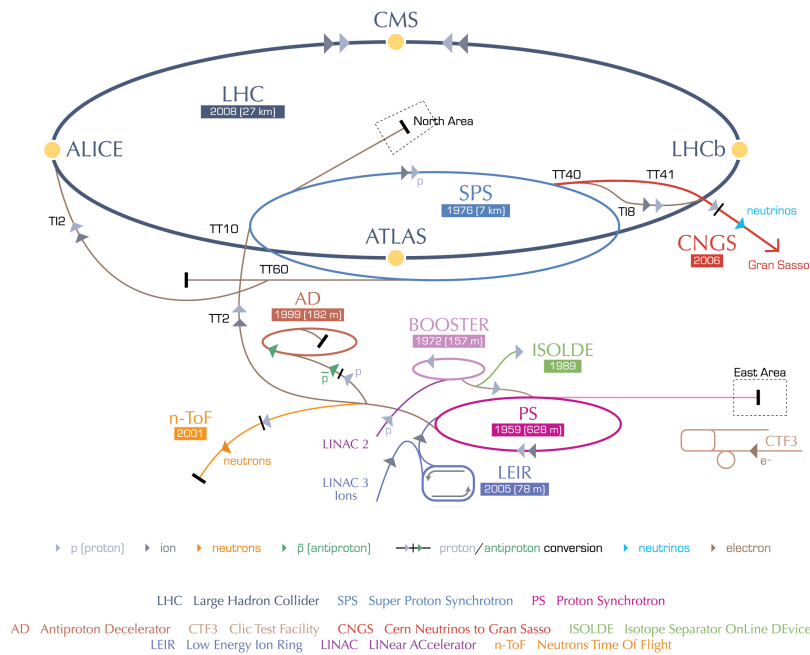
while being accelerated, loses its energy through the process called synchrotron radiation which is proportional to the fourth power of the ratio  $E/m$  between its energy and its mass. Comparing leptons accelerated by LEP to hadrons accelerated by the LHC, the synchrotron radiation energy loss at the same energy  $E$  is reduced by 12 orders of magnitude, which explains why it is possible to accelerate hadrons to significantly higher energies. Secondly, the composite structure of hadrons compared to the elementary structure of leptons allows the LHC to produce hard collisions at much wider energy spectrum. Having access to hard collisions with higher energies and wider spectrum, promotes LHC as a perfect candidate for discovering new physics. Another design decision was the choice to collide particles of the same charge instead of colliding particles with their antiparticles which was the case in other accelerators, e.g. antiprotons with protons at the Tevatron. Achieving large bunch populations is straightforward with proton beams because it is easier to produce protons. Additionally, with the primary goal of the LHC being to study electroweak symmetry breaking, the benefits of using antiprotons are small as the Higgs boson is produced primarily through gluon fusion in high-energy proton-(anti)proton and the gluon PDFs of protons and antiprotons are the same. Therefore, since it is more cost effective and easier to provide higher luminosities with proton-proton collisions and they yield the same physics potential, the LHC was designed to collide two 7 TeV counter-circulating beams of protons resulting in a center of mass energy of 14 TeV, or two beams of heavy ions, in particular lead nuclei, at an energy of 2.76 TeV/nucleon in the center of mass frame [64]. Due to issues with commissioning the superconducting magnets for the LHC and, in particular, the protection system for those magnets, the LHC presently operates at energies of up to 6.5 TeV per beam giving in total 13 TeV. In Figure 3.2 a schematic description of the accelerator complex installed at CERN is shown [65]. The acceleration procedure that brings protons to the LHC collision energies is separated into stages that combine various accelerators. Each of these accelerators provide beams for other experiments in addition to providing the beams for the LHC. From the historical point of view, they were all representing the final stage of acceleration in the experiments for which they were built [56]. The LHC acceleration chain consists:

- **The Duoplasmatron** is a proton or hadron source. To obtain protons, hydrogen gas is injected into the metal cylinder, surrounded by an electrical field which that breaks down the gas into protons and electrons. The protons are then accelerated

to 90 keV. This process yields roughly  $10^{14}$  protons.

- **The LINAC** is a linear accelerator 36 m long that generates a pulsed beam with an energy of 50 MeV using Radio Frequency Quadrupoles (RFQ) and focusing quadrupole magnets.
- **The Proton Synchrotron Booster (PSB)** [56] consists of four superimposed synchrotron rings of circumference equal to 157 m. The radio frequency techniques used for accelerating and bunching the hadron beams in the LINAC results in very long bunches, while the LHC is designed for 25 nanoseconds bunch spacing. In order to achieve the required bunch length, the bunches are accelerated and simultaneously split in the boosters. The PSB splits the input bunches into groups of four bunches and accelerates them to 1.4 GeV.

### CERN's accelerator complex



European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

© CERN 2008

Figure 3.2: Schematic description of the accelerator complex.

- **The Proton Synchrotron (PS)** is a synchrotron ring 628 m long, where the bunches are twice split in two, accelerated to 26 GeV, interspersed by 25 ns, and then passed to the following accelerator in the chain.

- **The Super Proton Synchrotron (SPS)** [56] is a synchrotron ring with a circumference of approximately 7 km in which the energy of the protons is brought to 450 GeV. It provides primarily high-intensity proton beams for the LHC but also for a number of active fixed-target experiments, e.g. COMPASS, NA61/SHINE and NA62.
- **The Large Hadron Collider (LHC)** is the largest and final synchrotron ring 27 km long, which is located at a depth varying between 50 m to 175 m. It is fed by transmission lines, in order to form two beams that run in opposite directions in separate parallel pipes, where each of them is accelerated to an energy of 6.5 TeV.

The transition from the LEP to the LHC collider brought numerous advantages but also design challenges. In order to accommodate the increase of the beam momentum, it was required to have stronger magnetic field, i.e. 8.33 T at the design energy of 7 TeV per beam, which was achieved by using superconducting magnets. Secondly, the same magnetic field cannot be used to bend counter circulating proton beams in the same direction.

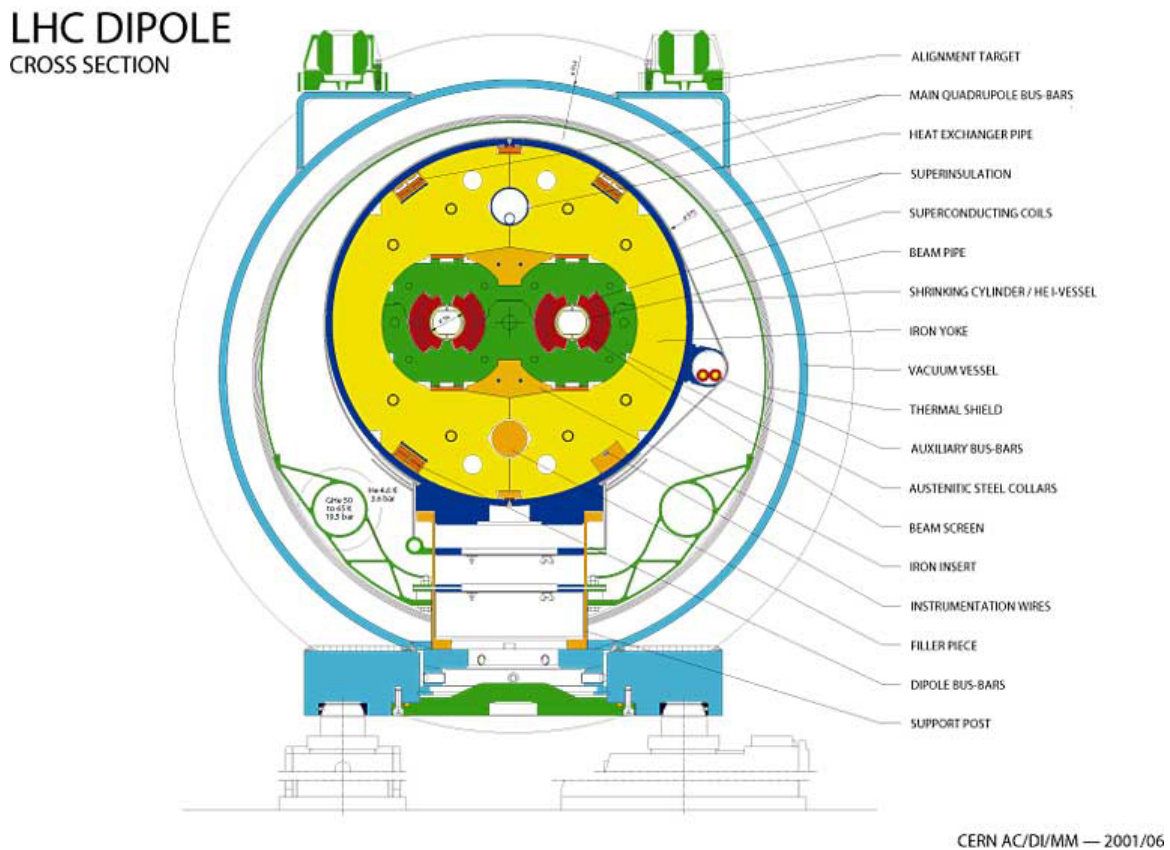


Figure 3.3: Schematic cross section of a LHC dipole magnet.

Therefore, the accelerator uses pairs of magnets with opposing field directions, mounted around the beam pipes within the same container in order to exploit the same structure and the same cryogenic system. A schematic cross section of a beam pipe in a LHC dipole magnet is shown in Figure 3.3 [66]. To minimize collisions of protons with the residual gas, high vacuum inside of beam pipes (in the range  $10^{-10}$ - $10^{-9}$  torr) is maintained. The LHC reaches the nominal energy of 6.5 TeV in a time interval of about 20 minutes after SPS provides the beam. If all the requirements for the collisions are satisfied, the stable beams condition is declared and data taking begins. Once the stable beam condition is over, the beam is extracted from the machine, dumped, and the process starts again. On average, full cycle before beam is dumped lasts approximately 15 hours. To achieve and maintain stable condition, a high level of coordination is required. The LHC ring is divided into eight main sectors called octants which are further separated by so-called arcs, as shown in Figure 3.4 [54]. Octants contain 14.3 m long magnetic dipoles responsible for bending the beam. In total, there are 1232 dipole magnets along the entire machine. In the central part of each octant there is a rectilinear structure the Long Straight Section (LSS), which is roughly 528 m long. At each LSS, beams are either injected, collimated, accelerated, intersected or extracted. In particular, there are 2 points where beams are injected.

Table 3.1: LHC technical parameters for proton-proton collisions.

Parameter	Value (Nominal)	Value (2016)
Maximum dipole magnetic field	8.33 T	8.33 T
Dipole operating temperature	1.9 K	1.9 K
Beam energy at injection	450 GeV	450 GeV
Beam energy at collision	7 TeV	6.5 TeV
Maximum instantaneous luminosity	$10^{34} \text{ cm}^{-2}\text{s}^{-1}$	$1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
Number of bunches per proton beam	2808	2220
Maximum number of protons per bunch	$1.69 \times 10^{11}$	$1.69 \times 10^{11}$
Bunch collision frequency	40 MHz	40 MHz
Energy loss per turn at 14 TeV	7 keV	7 keV

For collimation, the straight sections contain 386 focusing and defocusing quadrupoles, which keep the beam stable along the orbit; 360 sextupoles and 336 octupoles are used to make additional minor corrections to the beam direction. There are 2 lattice modules, called Dispersion Suppressor Left or Right (DSL and DSR), located at either end of the arcs, for the reduction of the horizontal dispersion. To accelerate the beams, the Radio Frequency (RF) acceleration system is used. It is housed in the cavern previously used for the ALEPH [67] experiment at LEP. It consists of 16 superconducting RF resonant cavities, 8 for each beam, with a frequency of 400 MHz for the carrier wave. Within each cavity a specific RF acceleration system establishes the value of the potential (2 MV) and the corresponding electric field (5.5 MV/m). The system increases the proton energy by 0.5 MeV with each revolution of the beam. After the nominal energy is reached, the RF cavity only supplies the beam with the energy lost during each revolution by synchrotron radiation, approximately 7 keV. The maximum energy allowed on the LHC beams is not determined by the electric field generated by the RF cavity but by the magnetic field necessary to maintain the protons in orbit.

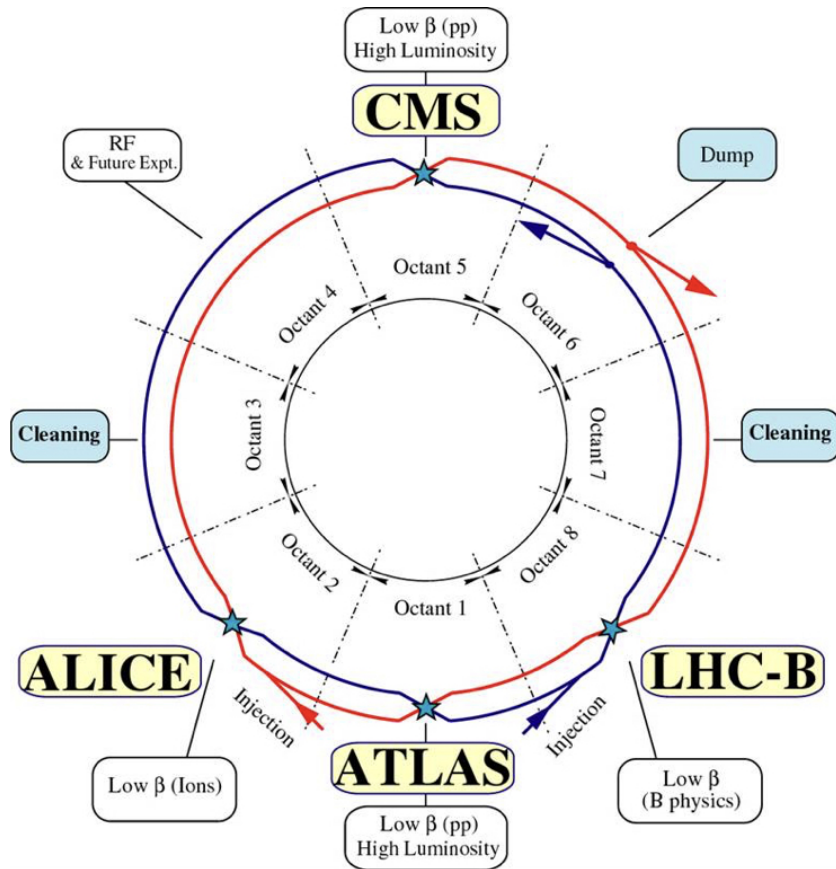


Figure 3.4: LHC schematic configuration showing clockwise beam colliding with counter-clockwise beam with the primary goal of the LHC being at four intersecting points.



Finally, to remove the beams without damaging the machine, they are dumped into 7m long segmented carbon cylinder of 700 mm diameter where 362 MJ of beam energy is dissipated in the  $90 \mu\text{s}$ . The LHC has four crossing points, around which are positioned seven detectors. Each of them is designed to address different questions. Four main experiments are:

- **Compact Muon Solenoid (CMS)** [68] is one of two general-purpose detectors, designed to observe any new physics phenomena that the LHC might reveal.
- **A Toroidal LHC ApparatuS (ATLAS)** [69] is other general-purpose detector. It has the same objectives as CMS but is designed independently and using different technological solutions.

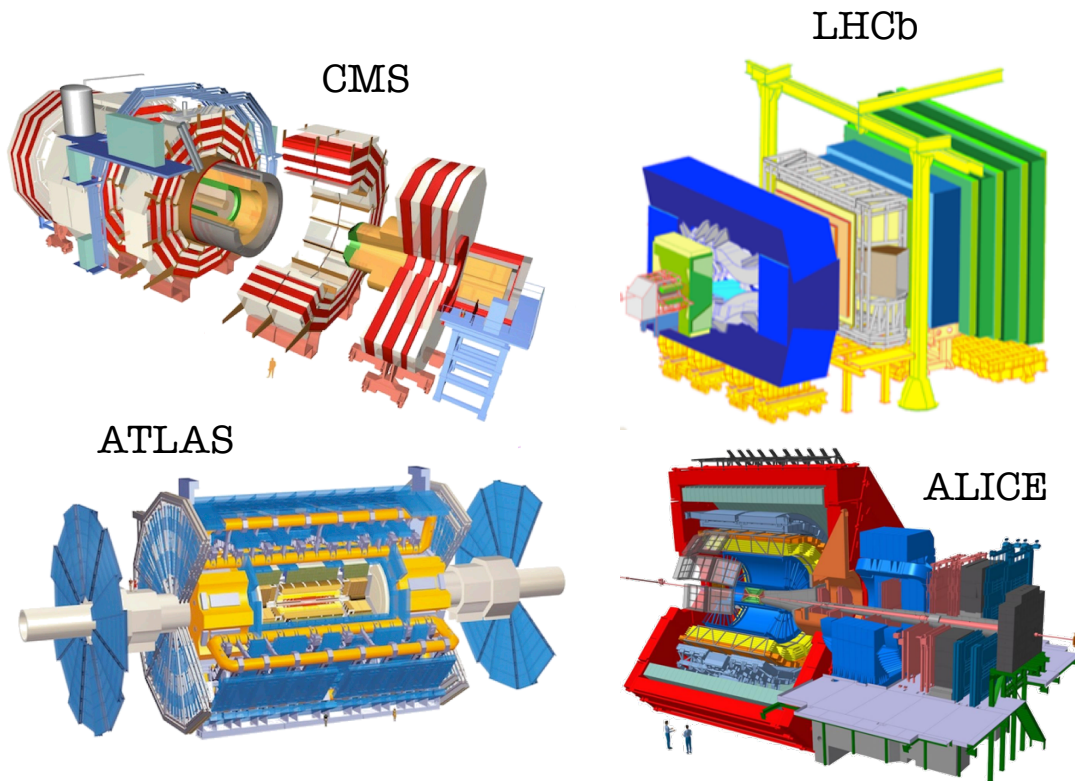


Figure 3.5: Schematic perspective view of the four main experiments installed at the LHC. From the top left, clockwise: CMS, LHCb, ATLAS and ALICE.

- **A Large Ion Collider Experiment (ALICE)** [70] is an experiment primarily designed to study collisions between heavy ions, e.g. lead ions in which a new state of matter, quark-gluon plasma can be produced. According to the Big Bang theory,

in the very early stage of the Universe formation, quark-gluon plasma existed for a short period of time.

- **LHC beauty (LHCb)** [71] is an experiment designed to study matter and antimatter asymmetry through the analysis of rare decays of hadrons containing b quarks, collectively known as B mesons. In comparison with the other three main detectors, LHCb is the only single arm forward detector.

In addition to the four main detectors, there are three much smaller detectors with very specific roles.

- **TOTAL Elastic and diffractive cross section Measurement (TOTEM)** [72] is dedicated for monitoring the LHC luminosity, precise measurement of the proton-proton interaction cross section and also to the in-depth study of the proton structure. It is made up of detectors positioned on either side of the CMS interaction point. Although the two experiments are scientifically independent, TOTEM complements the results from the CMS detector and from other LHC experiments as well.
- **LHC forward (LHCf)** [73] is made up of two detectors which sit along the LHC beamline, at 140 m on either side of the ATLAS collision point. It uses particles scattered at low angles by collisions in the LHC as a source to simulate cosmic rays in laboratory conditions. The LHCf is intended to measure the energy and numbers of neutral pions produced in nuclear collisions. This will hopefully help explain the origin of high energy cosmic rays. The results will provide input for other measurements from the Pierre Auger Observatory in Argentina and the Telescope Array Project in Utah.
- **Monopole and Exotics Detector at the LHC (MoEDAL)** [74] is dedicated to a direct search for magnetic monopoles and other exotic highly ionizing stable massive particles and pseudo-stable massive particles. To detect these particles, the project uses nuclear track detectors, which suffer characteristic damage due to highly ionizing particles, and are therefore perfectly suited for the purpose of detection. It is located in the same cavern as LHCb.

### 3.3 The Luminosity frontier

Finally yet importantly, luminosity, i.e. measure of the number of collisions that can be produced in a detector. For a process having a cross section  $\sigma$  that occurred  $N$  times in unit of time, the total luminosity  $\mathcal{L}$  is defined by the relation

$$N = \sigma \mathcal{L} \quad (3.1)$$

In a particle collider, the luminosity can be expressed as a function of the characteristic parameters of the beam [54]

$$\mathcal{L} = \frac{N_b^2 n_b f \gamma_r}{4\pi \epsilon_b \beta^*} \mathcal{F}(\vartheta) \quad (3.2)$$

The numerator of this equation expresses the rate at which protons enter an interaction region:  $N_b$  is the number of protons in a bunch,  $n_b$  is the number of colliding bunches at the interaction point,  $f$  the frequency at which they orbit the LHC and  $\gamma_r$  is the relativistic gamma factor. The denominator is the transverse geometrical cross section of the luminous region at the interaction point:  $\epsilon_n$  is the normalized emittance, i.e. the measure for the average spread of particle coordinates in position and momentum phase space,  $\beta^*$  is the beta function at the collision point. Finally, there is one more relativistic correction factor,  $\mathcal{F}(\vartheta)$ , defined as

$$\mathcal{F}(\vartheta) = \sqrt{1 + \left( \frac{\vartheta \sigma_z}{2\sigma^*} \right)^2} \quad (3.3)$$

It determines the reduction in the luminosity in the case that the bunches are not colliding head on. Here,  $\vartheta$  is the crossing angle itself,  $\sigma_z$  is the lab frame bunch length and  $\sigma^*$  is the bunch transverse width in the lab frame [64]. As the accelerator energy is limited by the technological constraints of the superconducting magnets and by the size of the ring, a method for increasing the number of rare beyond Standard Model physics events and with low cross section is to build an accelerator with high instantaneous luminosity rate. At the LHC, during the 13 TeV run in 2016, was achieved a maximum instantaneous luminosity of  $15.3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  [75].

### 3.4 Timeline of operations

On 10 September 2008, the first beam was successfully circulated through the LHC. The protons were fired around the tunnel in stages. It took less than one hour to guide the stream of particles around its inaugural circuit. Unfortunately, one week from the start a magnet quench occurred. This was a major incident that postponed the LHC operations until November 2009 [76] when the LHC became the world's highest-energy particle accelerator achieving 1.18 TeV per beam, beating the Tevatron's previous record of 0.98 TeV per beam held for eight years. On 30 March 2010, the two beams collided for the first time at center of mass 7 TeV (3.5 TeV per beam) marking the start of the LHC research program. One year later, on 21 April 2011, the LHC became the world's highest luminosity hadron accelerator achieving a peak luminosity of  $4.67 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ . Finally, on 4 July 2012 the CMS and ATLAS experiments discovered a new particle which was later confirmed as the Higgs boson. For the next two years, the LHC was in the first long shutdown because of preparation for a higher energy and luminosity. On 20 May 2015, protons collided in the LHC at the record-breaking collision energy of 13 TeV. Since then, the machine has been working smoothly with constant increase of the luminosity.

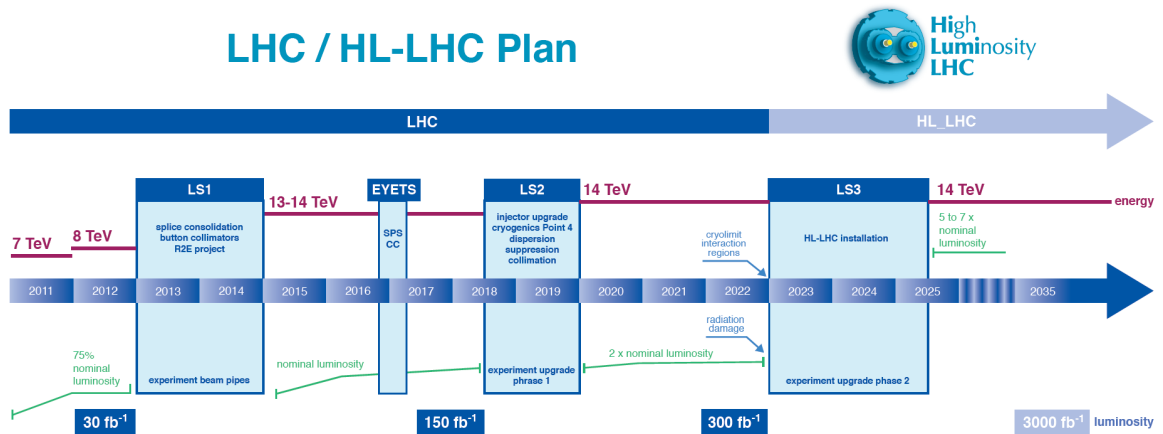


Figure 3.6: The proposed LHC running schedule. By the end of its lifetime in 2035, it is expected to collect more than  $3000 \text{ fb}^{-1}$  of data.

As any other particle physics experiment, the LHC is entering the period when it begins to suffer from diminishing major discoveries. Expected design characteristics have been reached and almost entire phase space for research with current setup has been exploited.

Next logical step is accelerator upgrade, typically in collision energy, luminosity or improved detectors which is also planned for the LHC. In addition to a possible increase of the collision energy to 14 TeV in 2018, a luminosity upgrade (the High Luminosity LHC) will boost the accelerator's potential for new discoveries, starting in 2026. The upgrade aims at increasing the luminosity of the machine by a factor of 10, providing a better chance to see rare processes and improving statistically marginal measurements. The LHC schedule is shown in Figure 3.6 [77].

# Chapter 4

## The Compact Muon Solenoid Detector

In this chapter an overview of the CMS detector is given.

### 4.1 Introduction

The Compact Muon Solenoid (CMS) [68] is a general purpose detector designed with focus on discovering the Higgs boson but also exploring new physics along with a wide array of possible Standard Model physics measurements. The detector is 21.6 m long, 15 m in diameter and weighs about 14000 tonnes. It is located in an underground cavern near the village of Cessy in France, just across the border from Geneva. The CMS experiment is one of the largest international scientific collaborations in history, involving 4300 particle physicists, engineers, technicians, students and support staff from 182 institutes in 42 countries (February 2014) [78] which published more than 700 articles [79]. The detector acts as a giant camera, taking 3D images of particle collisions provided by the LHC every 25 ns, which results in 40 million shots each second. By reconstructing nearly all particles from each collision and then using those pieces of information together, the detector can recreate an image of the collision for further analysis.

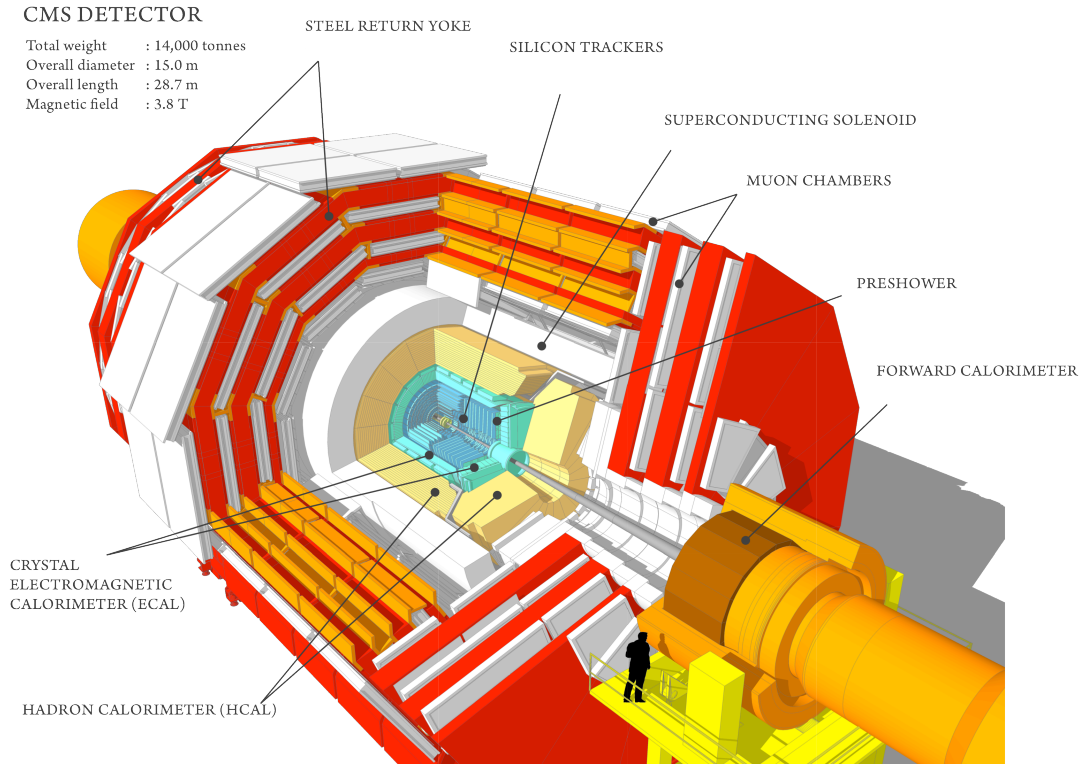
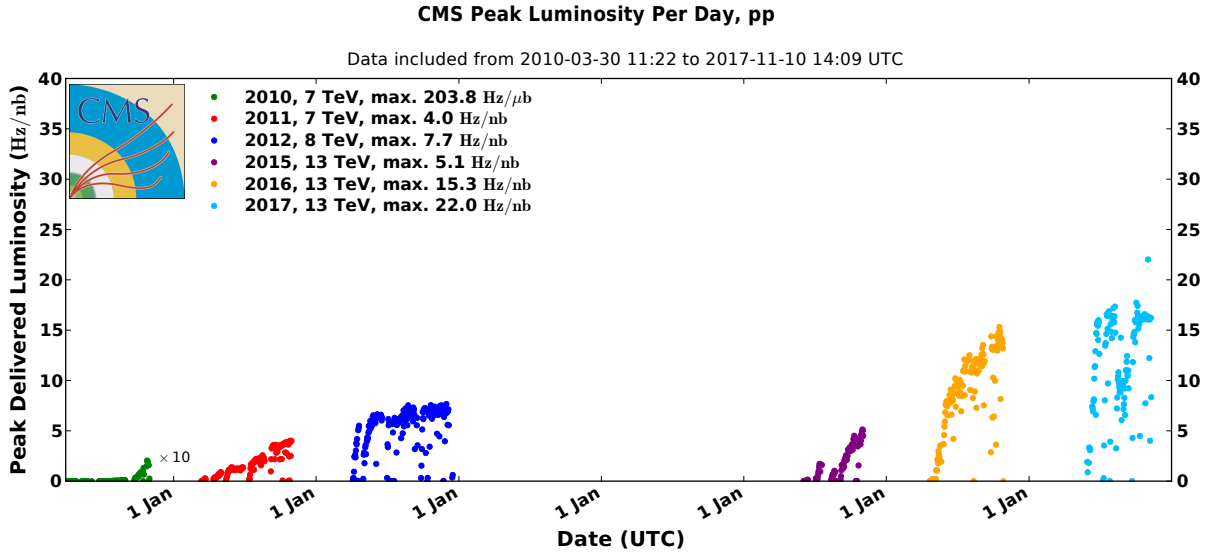


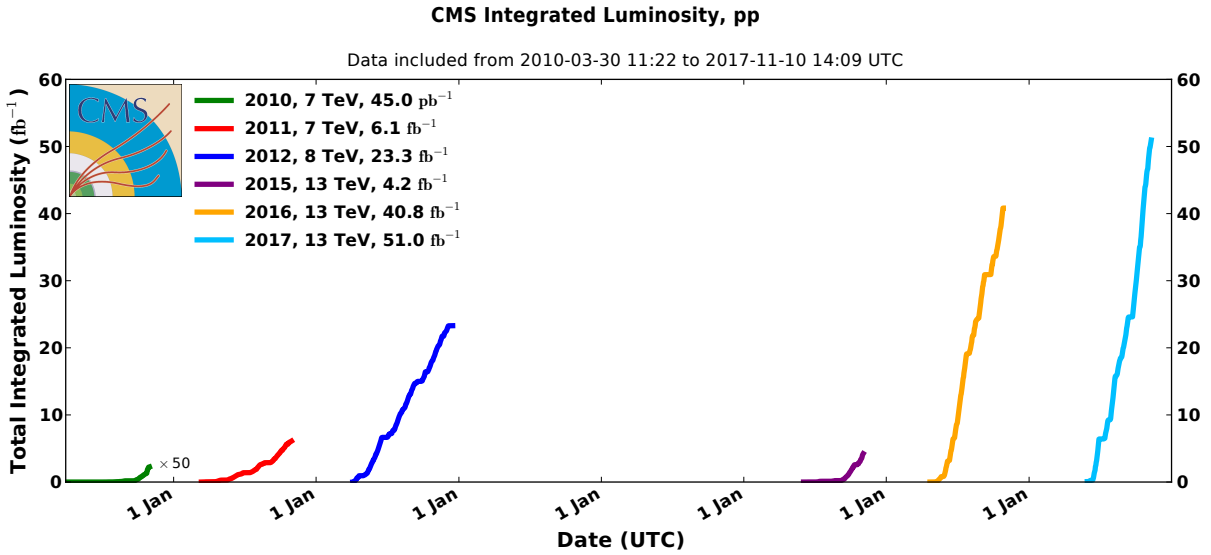
Figure 4.1: A view of the CMS detector with its subdetectors labeled.

The energy of particles is measured by the calorimeters. They have to be large enough to absorb the full particle energy if possible. An important part of the detector is a magnet. The trajectories of charged particles are bent by its field. The direction of curvature gives the information of particle's charge while the curvature radius is used to calculate the transverse momentum. For particles with high kinetic energy and therefore small curvature, a sufficiently long trajectory must be measured to accurately determine the radius. A view of the detector with its subdetectors is shown in Figure (4.1). The detector is designed to enclose the interaction region so that the total energy and momentum balance of each event is precisely measured and reconstructed. Combining the information from the different layers of the detector, it is possible to determine the type of particles. Each sub-detector is connected to an electronic readout system via thousands of cables. When an set of electronic impulses is registered, it is forwarded to few thousand computers that work together to collect them all efficiently. Out of all recorded events, only small part of them is selected. It is a role of so-called trigger system to select the most interesting events and this decision needs to be taken very fast. Therefore, data rate is reduced

by selecting only these events that fulfill certain criteria from 40 million events to few hundred events per second. Despite this significant data reduction, CMS still records approximately 600 MB/s. To quantify the amount of collected data, the cumulative or integrated luminosity is usually used. It is shown in Figure (4.2) for each year since its beginning in 2010 [75].



(a) Peak luminosity delivered to CMS.



(b) Cumulative luminosity delivered to CMS.

Figure 4.2: Cumulative and peak luminosity versus day delivered to CMS by the LHC.



## 4.2 Detector design

The CMS detector is shaped like a cylindrical onion, with several concentric layers of components. The central part (barrel) contains several layers of detectors with cylindrical symmetry and coaxial with respect to the direction of the beams. A set of detector disks (endcaps) close the detector at both ends to ensure its tightness. CMS uses a right-handed coordinate system with the origin in the nominal interaction point. z-axis is pointing along the beam line. x-axis is pointing towards the center of the LHC ring and y-axis points upwards. From the inner region to the outer one, the components are:

- **The Silicon tracker** [80–82] is placed in the region  $r < 1.2$  m and  $|\eta| < 2.5$ . It consists of a silicon pixel vertex detector and a surrounding silicon microstrip detector, with a total active area of about 215 m<sup>2</sup>. It is used to reconstruct charged particle tracks and decay vertices.
- **The Electromagnetic Calorimeter (ECAL)** [83, 84] is placed in the region  $1.2$  m  $< r < 1.8$  m and  $|\eta| < 3$ . It consists of scintillating crystals of lead tungstate (PbWO<sub>4</sub>) and it is used to measure the energy of photons and electrons.
- **The Hadron Calorimeter (HCAL)** [85] is placed in the region  $1.8$  m  $< r < 2.9$  m and  $|\eta| < 5$ . It consists of brass layers alternated with plastic scintillators and it is used to measure the direction and the energy released by hadrons produced in the interactions.
- **The Superconducting Solenoidal Magnet** [86] is placed in the region  $2.9$  m  $< r < 3.8$  m and  $|\eta| < 1.5$ . It generates an internal uniform magnetic field of 3.8 T along the direction of the beams, necessary to deflect the charged particles in order to allow a measurement of their momentum through the curvature observed in the tracking system. The magnetic field is closed with an iron yoke 21.6 m long with a diameter of 14 m, in which a residual magnetic field of 1.8 T is present, in the opposite direction with respect to the 3.8 T field in the interior of the solenoid.
- **The Muon system** [87] is placed in the region  $4$  m  $< r < 7.4$  m and  $|\eta| < 2.4$ . It consists of Drift Tubes (DT) in the barrel region and Cathode Strip Chambers (CSC) in the endcaps. A complementary system of Resistive Plate Chambers (RPC) is used

both in the barrel and in the endcaps. This composite tracking system for muons is used to reconstruct muons tracks that pass through it. The muons chambers are housed inside the iron (high magnetic permeability) structure of the return yoke that encloses the magnetic field.

Charged particles such as electrons, pions and muons leave traces through ionisation. Electrons are very light and therefore lose their energy quickly, while pions penetrate further through the layers of the detector. Photons themselves leave no trace, but in the calorimeters, each photon converts into electron-positron pair, the energies of which are then measured. Muons are the only particles that reach (and are detected by) the outermost layers of the detector as can be seen in Figure (4.3) [88]. A detailed discussion of each sub-detector is given in the following sections.

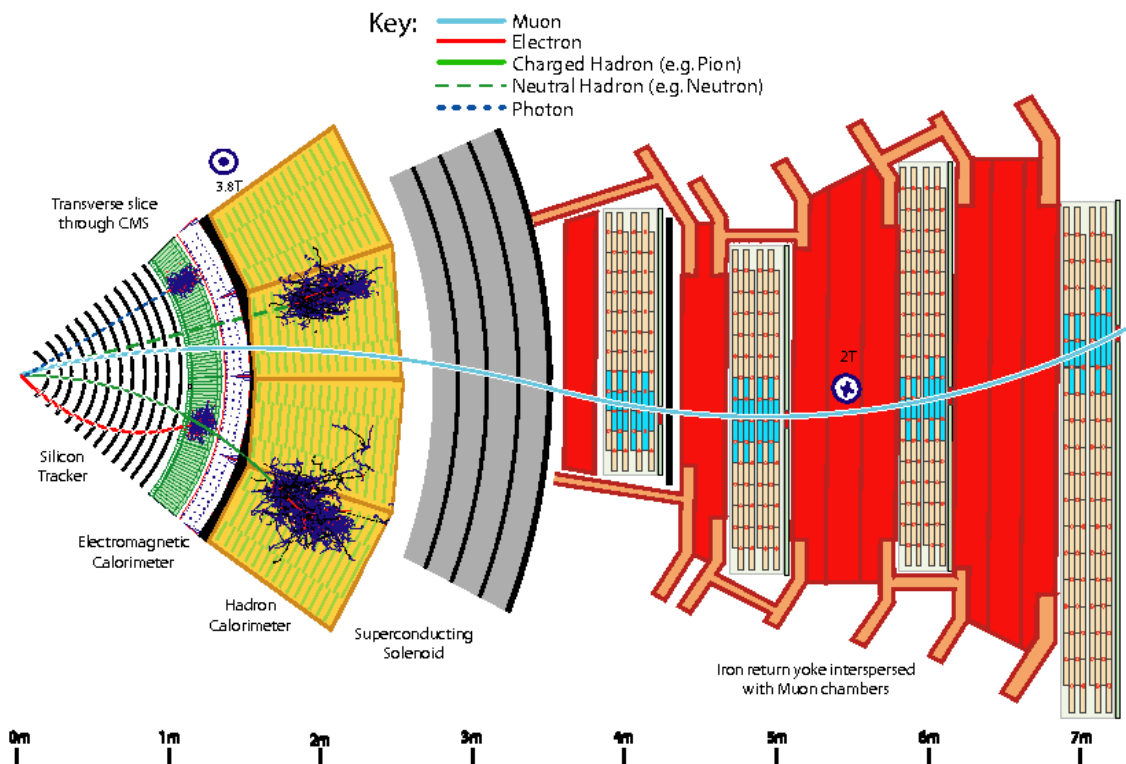


Figure 4.3: A sketch of the specific particle interactions in a transverse slice of the CMS detector, from the beam interaction region to the muon detector. The muon and the charged pion are positively charged, and the electron is negatively charged.

### 4.2.1 The Tracker

The Silicon Tracker is the closest detector to the collision point. Its goal is to reconstruct the trajectory of charged particles originating from the interaction point and also to identify the position of secondary vertices, if they exist, in the events containing composite particles with short lifetime (in particular hadrons containing the quark  $b$  that decay after a few hundred  $\mu\text{m}$ ). In order to reconstruct the trajectories with the highest possible resolution it is necessary to have low occupancy detector with large redundancy of measured points per track. The low occupancy is obtained by building a detector with high granularity. The redundancy of hits per track is necessary to reduce the ambiguity in assigning hits to the track itself, which was achieved by building multiple detection layers. However, the amount of material composing the detector has been limited to avoid deterioration of the measurement itself. An excessive amount of material would compromise the accuracy of the track measurement because of the particle multiple scattering. Secondly, traversing the detector can induce pair-production of electrons from real photons, causing the photon to appear as though it is an electron since the pair-produced electrons will leave tracks. To mitigate these effects, the detection layers in the trackers have been limited in number and thickness.

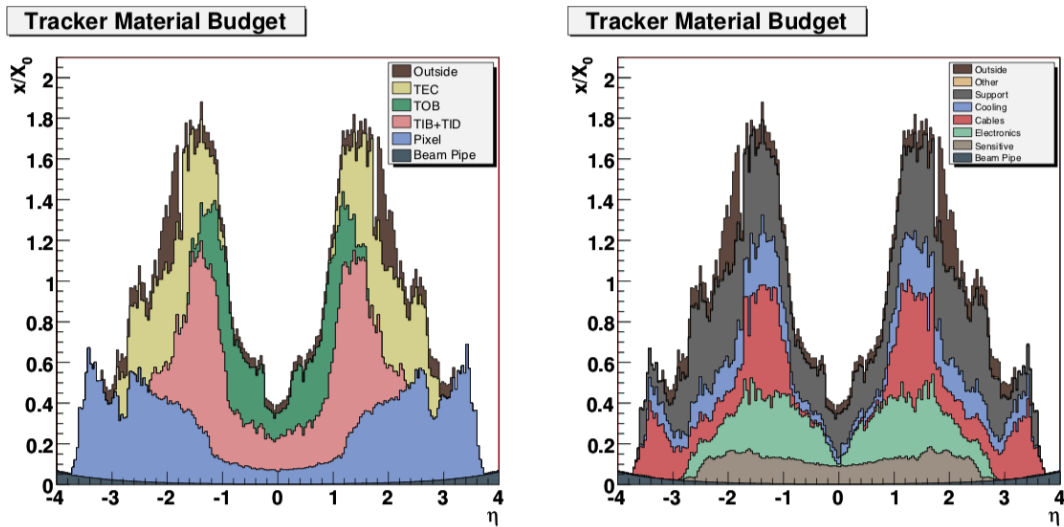


Figure 4.4: Material budget in units of radiation length as a function of pseudorapidity  $\eta$ , for the different sub-detectors (left panel) and broken down into the functional contributions (right panel).

In Figure (4.4) the material budget of the the tracker is reported in units of radiation length  $t/X_0$  as a function of  $\eta$  obtained from the full simulation of the tracker. The material budget is higher in the region  $1 < \eta < 2$  due to the presence of cables and services. The tracker comprises a large silicon strip tracker with a small silicon pixel tracker inside it. Both operate using the same principles of finding the location of a charged particle in a collection of pixels or on a doped strip of silicon using the ionization charge left by the incident particle. For that purpose, the sensors are constructed as reversed-biased p-n diodes, which give a detectable current when the bias voltage across the diode is lowered by the ionization charge flow. The use of semi-conductor sensors is critical to the operation of the tracker due to its small thickness, which results in very short charge collection times, allowing for fast readout. Additionally, the use of semi-conductor, i.e. silicon, allows for high detector granularity, granting excellent position resolution for reconstructed tracks, for a low cost in material.

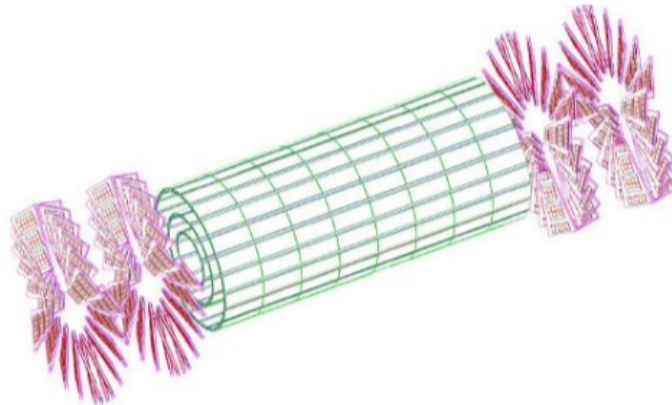


Figure 4.5: A schematic view of the pixel vertex detector.

The tracker is a cylindrical volume 5.8 m in length and 2.5 m in diameter, with its axis closely aligned to the LHC beam line. A schematic drawing of the tracker is shown in Figure (4.6). In the central pseudorapidity region, the pixel tracker consists of three co-axial barrel layers at radii between 4.4 cm and 10.2 cm and the strip tracker consists of ten co-axial barrel layers extending outwards to a radius of 110 cm. On both sides of the barrel are endcaps, each consisting of two disks in the pixel tracker, and three small plus nine large disks in the strip tracker. The endcaps extend the acceptance of the tracker up to a pseudorapidity of  $|\eta| < 2.5$ .

**The pixel detector** consists of cylindrical barrel layers at radii of 4.4, 7.3 and 10.2 cm,

and two pairs of endcap disks at  $z = \pm 34.5$  and  $\pm 46.5$  cm. It provides three-dimensional position measurements of the hits arising from the interaction of charged particles with its sensors. The hit position resolution is approximately  $10\ \mu\text{m}$  in the transverse coordinate and  $20 - 40\ \mu\text{m}$  in the longitudinal coordinate, while the third coordinate is given by the sensor plane position. In total, its 1440 modules cover an area of about  $1\ \text{m}^2$  and have astonishing 66 million pixels.

**The strip tracker** has 15 148 silicon modules, which in total cover an active area of about  $198\ \text{m}^2$  and have 9.3 million strips. It is composed of four subsystems. The Tracker Inner Barrel (TIB) and Disks (TID) cover  $r < 55$  cm and  $|z| < 118$  cm, and are composed of four barrel layers, supplemented by three disks at each end. These provide position measurements in  $r$ - $\phi$  with a resolution of approximately  $13$ - $38\ \mu\text{m}$ . The Tracker Outer Barrel (TOB) covers  $r > 55$  cm and  $|z| < 118$  cm and consists of six barrel layers providing position measurements in  $r$ - $\phi$  with a resolution of approximately  $18$ - $47\ \mu\text{m}$ . The Tracker EndCaps (TEC) cover the region  $124 < |z| < 282$  cm. Each TEC is composed of nine disks, each containing up to seven concentric rings of silicon strip modules, yielding a range of resolutions similar to that of the TOB.

The modules of the pixel detector use silicon of  $285\ \mu\text{m}$  thickness, and achieve resolutions that are roughly the same in  $r$ - $\phi$  as in  $z$ , because of the chosen pixel cell size of  $100 \times 150\ \mu\text{m}^2$ . The modules in the TIB, TID and inner four TEC rings use silicon that is  $320\ \mu\text{m}$  thick, while those in the TOB and the outer three TEC rings use silicon of  $500\ \mu\text{m}$  thickness. In the barrel, the silicon strips usually run parallel to the beam axis and have a pitch (i.e., the distance between neighbouring strips) that varies from  $80\ \mu\text{m}$  in the inner TIB layers to  $183\ \mu\text{m}$  in the inner TOB layers. The endcap disks use wedge-shaped sensors with radial strips, whose pitch varies from  $81\ \mu\text{m}$  at small radii to  $205\ \mu\text{m}$  at large radii. The modules in the innermost two layers of both the TIB and the TOB, as well as the modules in rings 1 and 2 of the TID, and 1, 2 and 5 of the TEC, carry a second strip detector module, which is mounted back-to-back to the first and rotated in the plane of the module by a ‘stereo’ angle of  $100\ \text{mrad}$ . The hits from these two modules, known as ‘ $r$ - $\phi$ ’ and ‘stereo hits’, can be combined into matched hits that provide a measurement of the second coordinate ( $z$  in the barrel and  $r$  on the disks). The achieved single-point resolution of this measurement is an order of magnitude worse than in  $r$ - $\phi$ . The principal characteristics of the tracker are summarized in a Table (4.1) and Figure (4.6).

Table 4.1: A summary of the principal characteristics of the various tracker subsystems.

Tracker subsystem	Layers	Pitch	Location
Pixel tracker barrel	3 cylindrical	$100 \times 150 \mu m^2$	$4.4 < r < 10.2 \text{ cm}$
Strip tracker inner barrel (TIB)	4 cylindrical	$80 - 120 \mu m$	$20 < r < 55 \text{ cm}$
Strip tracker outer barrel (TOB)	6 cylindrical	$122 - 183 \mu m$	$55 < r < 116 \text{ cm}$
Pixel tracker endcap	2 disks	$100 \times 150 \mu m^2$	$34.5 <  z  < 46.5 \text{ cm}$
Strip tracker inner disks (TID)	3 disks	$100 - 141 \mu m$	$58 <  z  < 124 \text{ cm}$
Strip tracker endcap (TEC)	9 disks	$97 - 184 \mu m$	$124 <  z  < 282 \text{ cm}$

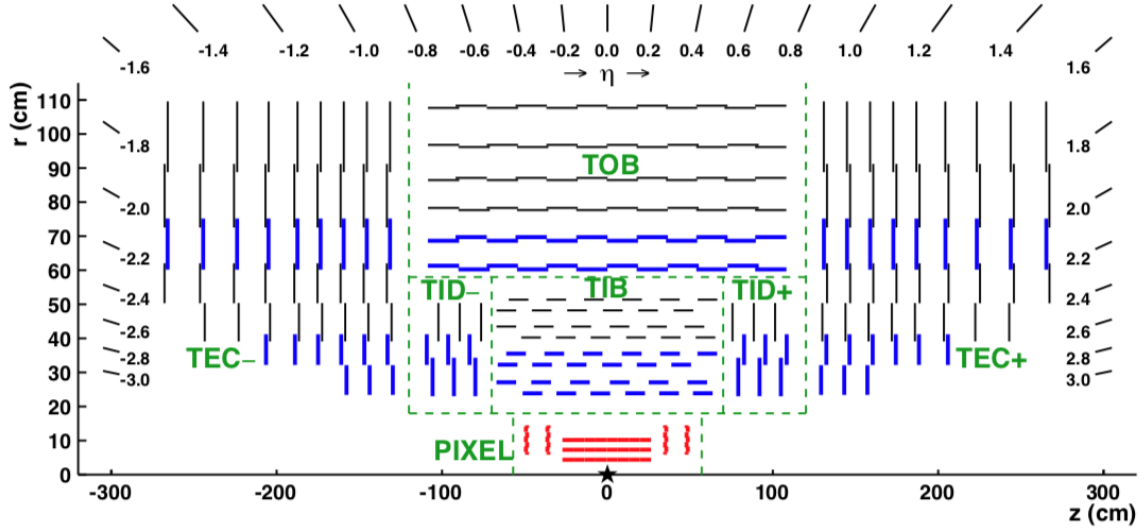


Figure 4.6: Schematic cross section through the CMS tracker in the  $r$ - $z$  plane. In this view, the tracker is symmetric about the horizontal line  $r = 0$ , so only the top half is shown here. The centre of the tracker, corresponding to the approximate position of the  $pp$  collision point, is indicated by a star. Green dashed lines help the reader understand which modules belong to each of the named tracker subsystems.

Since the tracker is within the magnetic field of the solenoid, the charged particles bend with radius of curvature  $R$  in the transverse plane of the detector according to:  $R = p_T/qB$  where  $p_T = p_x^2 + p_y^2$ . Since the tracking detector has excellent position resolution, the

momentum measurements derived from it have very high resolution [64]

$$\frac{\sigma_{pT}}{p_T} = (15p_T \oplus 0.5) \% (\text{TeV}), |\eta| < 1.6$$

$$\frac{\sigma_{pT}}{p_T} = (60p_T \oplus 0.5) \% (\text{TeV}), |\eta| = 2.5$$

## 4.2.2 The Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter (ECAL, Electromagnetic CALorimeter) is the second closest detector to the collision point. It is designed to measure accurately the energy of electrons and photons. Using information from the tracker and ECAL, electrons and photons can be distinguished due to the fact that electrons leave the trace in the tracker while photons do not. The ECAL consists of 76 832 ( $\text{PbWO}_4$ ) crystals. The characteristics of the lead tungstate crystals make them an appropriate choice for operation at LHC. The high density ( $\rho = 8.3 \text{ g/cm}^3$ ), short radiation length ( $X_0 = 0.89 \text{ cm}$ ) and small Moliere radius (2.2 cm) allow the construction of a compact and high granularity calorimeter.

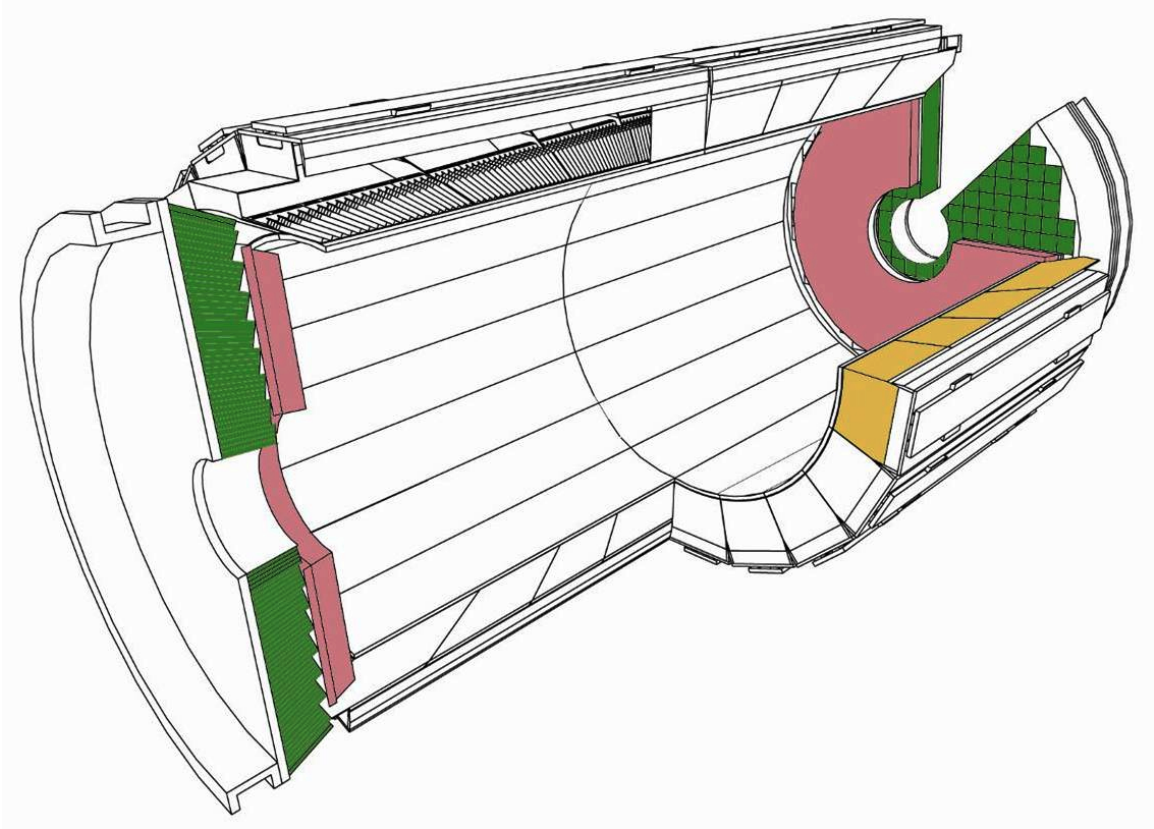


Figure 4.7: Layout of the CMS electromagnetic calorimeter presenting the arrangement of crystal modules, supermodules, endcaps and the preshower in front.

Another advantage of this material is the radiation hardness and the fast scintillation decay time ( $\tau = 10$  ns), which permits to collect about 80% of the produced light within the 25 ns interval between two consecutive bunch crossings. The main drawbacks of this material are the low light yield ( $\sim 10$  photoelectrons/MeV) and the strong dependence on the operating temperature, that makes it necessary to keep the crystals at a stabilized temperature (18°C). Low light yield is solved through the use of silicon avalanche photodiodes (Avalanche Photodiodes, APD) in the barrel part and single stage photomultipliers (Vacuum Photo-Triode, VPT) in the endcaps, both resistant to the radiation and to the strong magnetic field of CMS. They are located on the ‘back’ of each crystal, collecting all of the light emitted in the electromagnetic shower. The layout of the CMS ECAL is given in Figure (4.7). The ECAL is divided into the ECAL barrel (EB) in the range  $|\eta| < 1.4442$ , and the endcaps (EE) and endcap preshower detector (ES) both in the range  $1.566 < |\eta| < 3.0$ . Out of 76 832 crystals, EB contains 61 200 crystals the while two endcaps (EE) contain 7324 crystals.

**The ECAL barrel (EB)** has an inner radius of 129 cm, a length of 630 cm and it extends in the region  $|\eta| < 1.479$ . It consists of 36 supermodules, each one with a length equal to the half of the barrel length. Each supermodule consists of a  $20 \times 85$  crystal matrix in the  $\phi - \eta$  plane. Supermodules are divided into 4 modules along the  $\eta$  direction and each module is in turn divided into sub-modules. Each submodule, consists of a  $5 \times 2$  crystal arrays mounted on a glass fiber structure. EB crystals in the barrel region are rectangular prisms, with a length of 23 cm, a frontal area equal to  $22 \times 22$  mm<sup>2</sup> and a rear area equal to  $26 \times 26$  mm<sup>2</sup> which corresponds to  $25.8 X_0$  in longitudinal shower depth and 1 Moliere radius in the transverse shower size. The granularity of a single crystal is equal to  $\Delta\eta \times \Delta\varphi = 0.0175 \times 0.0175$  (i.e. about 1°). Crystals are grouped into  $5 \times 5$  matrices called trigger towers that provide some useful information for the trigger. The crystal axes are tilted by 3° with respect to the direction pointing to the interaction point, both in the  $\eta$  and  $\varphi$ . This is done to avoid the gaps between adjacent crystals to be aligned with the direction pointing to the interaction point.

**The ECAL endcap (EE)** covers the region  $1.479 < |\eta| < 3$  and is formed by two semicircular halves of aluminum called Dees. On each dee, a total of 3662 crystals with trapezoidal shape with a length of 22 cm, frontal area equal to  $28.6 \times 28.6$  mm<sup>2</sup> and a rear area equal to  $30 \times 30$  mm<sup>2</sup> are placed. EE crystals are arranged in 18 supercrystals with



5×5 elementary unit. Unlike the EB crystals which are arranged in a  $\eta - \varphi$  symmetry, the EE crystals are arranged according to a  $x - y$  geometry. To ensure good tightness, there is an overlap of one half crystal between the barrel and the endcaps. Moreover, in order to avoid the presence of gaps between crystals, their axes are pointing 130 cm beyond the interaction point.

**The ECAL endcap preshower (ES)** is installed on the inner side of the endcaps in order to separate the showers produced by a primary  $\gamma$  from those produced by forward emitted  $\pi_0$ . This detector, which covers the region  $1.653 < |\eta| < 2.6$ , is a sampling calorimeter consisting of two lead disks ( $2 X_0$  and  $1 X_0$  thick respectively) that initiate the electromagnetic shower from incoming photons or electrons, with silicon strip sensors after each disk, which measure the deposited energy as well as the shower transverse profile. The silicon detector is  $320 \mu\text{m}$  thick, has area of  $63 \times 63 \text{ mm}^2$  and it is divided into 32 strips with a 1.9 mm pitch. Schematic 2D representation of one quarter of the ECAL is shown in Figure (4.8).

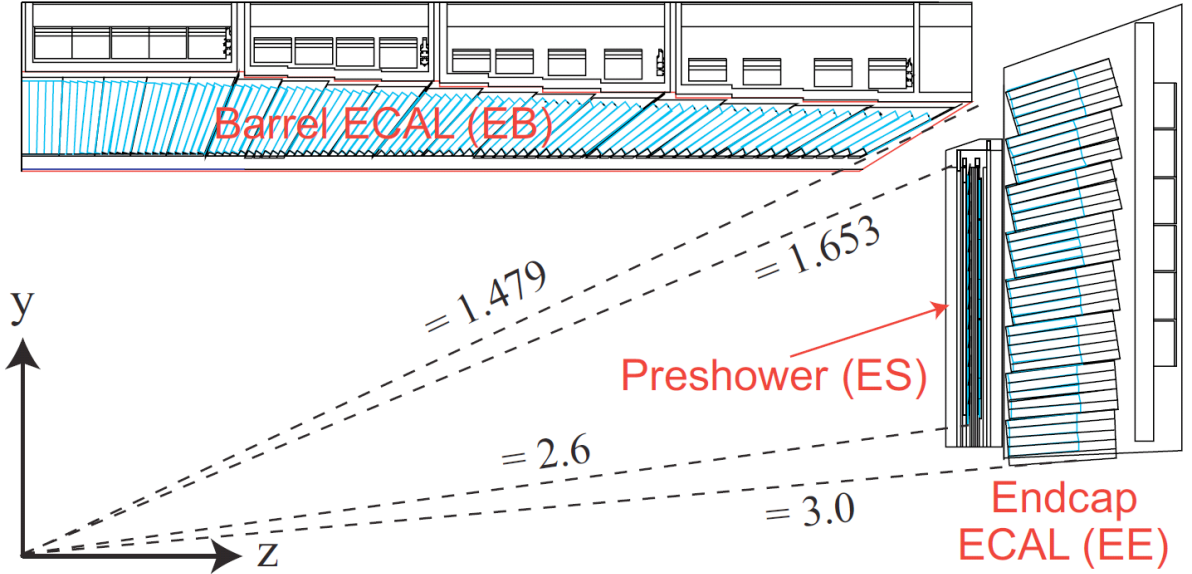


Figure 4.8: Geometric view of one quarter of the ECAL.

The energy resolution of ECAL for electrons in beam tests has been measured [89] and can be expressed by the sum in quadrature of three terms

$$\frac{\sigma_E}{E} = \frac{2.8 \% \text{ GeV}^{1/2}}{\sqrt{E}} \oplus \frac{12 \% \text{ GeV}}{E} \oplus 0.3\% \quad (4.1)$$

where  $E$  is the energy measured in GeV. The first term dominates at low energies. It represents the statistical error coming from the stochastic nature of electromagnetic shower evolution. The second term represents the error in the energy measurement coming from electronic noise and a small contribution generated by energy depositions from additional soft interactions. It varies with pseudorapidity. The last, constant, term dominates at high energies. It takes into account detector non-uniformities and calibration uncertainties, e.g. the stability of the operating conditions such as temperature and voltage, the presence of dead material in front of crystals, the electromagnetic shower leakage into the HCAL, the intercalibration errors and the radiation damage of the crystals. The radiation damage manifests itself as a change in transparency. This is monitored and corrected using a laser calibration system [83].

### 4.2.3 The Hadronic Calorimeter (HCAL)

The hadronic calorimeter (HCAL, Hadronic CALorimeter) measures strongly interacting particles by forcing them to interact with a dense material interleaved or embedded with a scintillator. Together with the ECAL, it makes a complete calorimetric system. Furthermore, thanks to its tightness and coverage  $|\eta| < 5$ , it can provide a measurement of the features of non-interacting particles, such as neutrinos, by measuring the missing energy in the transverse plan,  $E_T^{\text{miss}}$ . Additionally, the large angular coverage of the HCAL allows for a complete description of pileup energy depositions. As shown in (4.9) [90], the HCAL is divided into four subdetectors: Barrel Hadronic Calorimeter, located in the barrel region inside the magnet; Endcap Hadronic Calorimeter, situated in the endcap regions inside the magnet; Outer Hadronic Calorimeter, placed along the inner wall of the magnetic field return yoke, just outside of the magnet; Forward Hadronic Calorimeter, consisting of two units placed in the very forward region outside the magnetic coil. The choice of the material is driven by two factors: in order to maximize particle containment for a precise missing transverse energy measurement, the amount of absorber material inside the magnetic coil was maximized. Since HCAL is mostly placed inside the magnetic coil, a non-magnetic material like brass was chosen.

**The Barrel Hadronic Calorimeter (HB)** is a sampling calorimeter consisting of alternating plates of brass absorber and scintillator tiles embedded with wavelength shifting

(WLS) fibers. The WLS fibers are spliced to clear fibers which direct the light to an optical decoder unit. The optical decoder unit arranges the fibers into readout towers and directs the light from each tower to separate channels of a hybrid photo diode. It is 9 m long, extends into region  $178 < r < 288$  cm, and therefore covers  $|\eta| < 1.4$ . HB consists of 36 azimuthal wedges assembled into two half-barrels. Each half-barrel is divided along  $\eta$  into 16 sectors and the detector contains a total of 2304 calorimetric towers with granularity of  $0.087 \times 0.087$  in  $\eta - \varphi$ . Each tower consists of 15 brass layers, with a thickness of 50 mm each, arranged in direction parallel to the beams. Interleaved with the absorber layers, there are 17 layers of plastic scintillators with a thickness of 3.7 mm, except for the innermost one that is 9 mm thick.

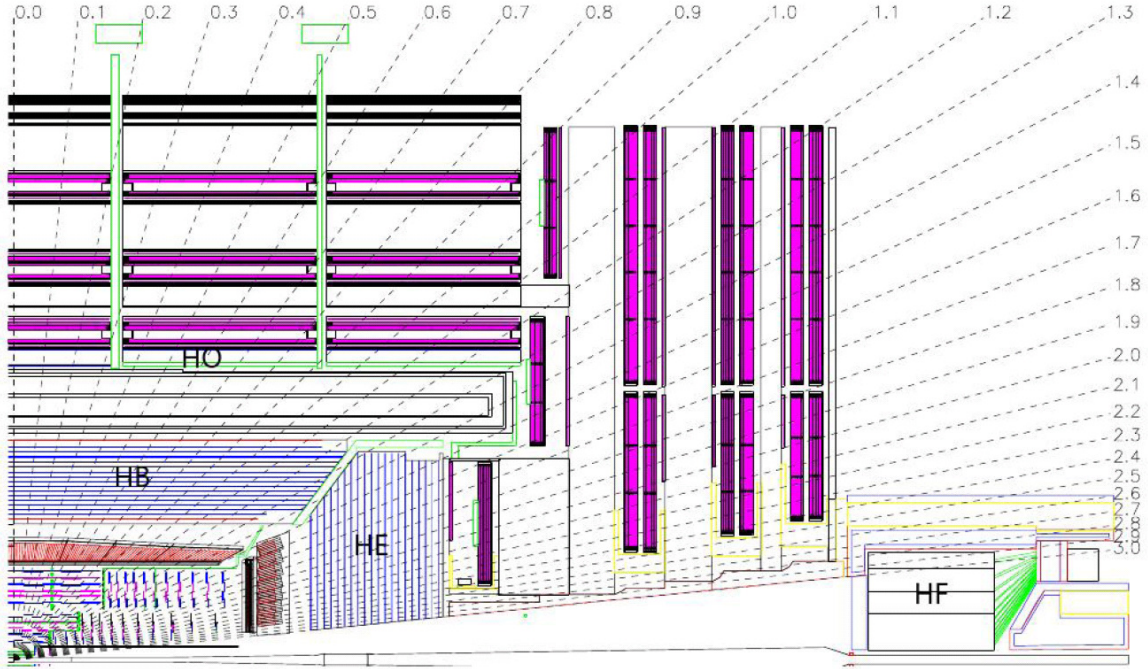


Figure 4.9: Geometric view of the CMS detector showing the HCAL subdetectors.

**The Endcap Hadronic Calorimeter (HE)** shares the same working principles and almost the same design as the HB. It extends the pseudorapidity region to  $1.3 < |\eta| < 3$ , slightly overlapping with the barrel. The empty region between the two detectors is used to pass cables and optical fibers and does not point toward the center of interaction in order to preserve the calorimeter tightness.

**The Outer Hadronic Calorimeter (HO)** is located outside of the solenoid and acts as “tail catcher” effectively increasing the thickness of the calorimeter in the central pseudo

rapidity region. It consists of several layers of plastic scintillators, which increase the effective size of the calorimeter to more than 10 nuclear interaction length. The scintillators have a thickness of 10 mm and they have the same HB granularity in order to have a 1-1 correspondence between HB calorimetric towers and HO segments.

**The Forward Hadronic Calorimeter (HF)** is a Cerenkov light detector made up of quartz fibers embedded within a 1.65 m long steel absorber. It is placed in the very forward region,  $|\eta| \sim 5$ , at a distance of 11.2 m from the nominal point of collision of beams. This detector is optimized for the identification of those processes which produce very forward jets. The choice of these materials was due to the high radiation dose of the forward region, which does not allow the use of plastic scintillators. Each unit is composed of 18 sections, each covering an angle of  $20^\circ$  in  $\varphi$  and containing 24 calorimetric towers with fibers 1.65 m long interleaved to fibers 1.43 m long. There are 13 towers in  $\eta$  with segmentation variable from 0.1 to 0.3 depending on the distance from the beam while the  $\varphi$  segmentation ranges between  $10^\circ$  and  $20^\circ$ . Overall, the apparatus consists of 900 towers, sampled with 1800 readout channels.

The energy resolution in the different geometrical regions of HCAL as function of energy is

$$\text{HB, HE:} \quad \frac{\sigma_E}{E} = \frac{90 \% \text{ GeV}^{1/2}}{\sqrt{E}} \oplus 4.5\% \quad (4.2)$$

$$\text{HF:} \quad \frac{\sigma_E}{E} = \frac{172 \% \text{ GeV}^{1/2}}{\sqrt{E}} \oplus 9\% \quad (4.3)$$

#### 4.2.4 The Solenoid

The CMS magnet is the biggest superconducting solenoid ever built in the world. The solenoid achieves a magnetic field of 3.8 T in a barrel of 6 m in diameter and 12.5 m in length. The energy stored in the magnet is about 2.6 GJ at full current, 19 500 A. The superconductor is made of four Niobium-Titanium layers, cooled to  $-268.5^\circ\text{C}$ . In case of a quench, when the magnet loses its superconducting property, the energy is dumped to resistors within 200 ms. The magnet return yoke of the barrel has 12-fold rotational symmetry and is composed of three sections along the z-axis; each one is split into 4 layers. Most of the iron volume is saturated or nearly saturated, and the field in the yoke is about half (1.8 T) of the field in the central volume. The magnet had to be built offsite

and transported to CMS by road, which meant it physically could not be more than 7 metres in diameter, or else it would not fit through the streets on its way to Cessy.

### 4.2.5 The Muon System

The CMS Muon System is dedicated to accurately track muons, as minimum ionizing particles that do not deposit much of their energy in the calorimeters. It is placed outside the magnetic coil, and is embedded in the return yoke, to fully exploit the 1.8 T return flux. In combination with tracker measurements, the muon system precisely traces a muon's path giving enhanced resolution, especially at large transverse momentum. As shown in Figure (4.10) [91], the system consists of three independent subsystems: the drift tubes, the resistive plate chambers and cathode strip chambers.

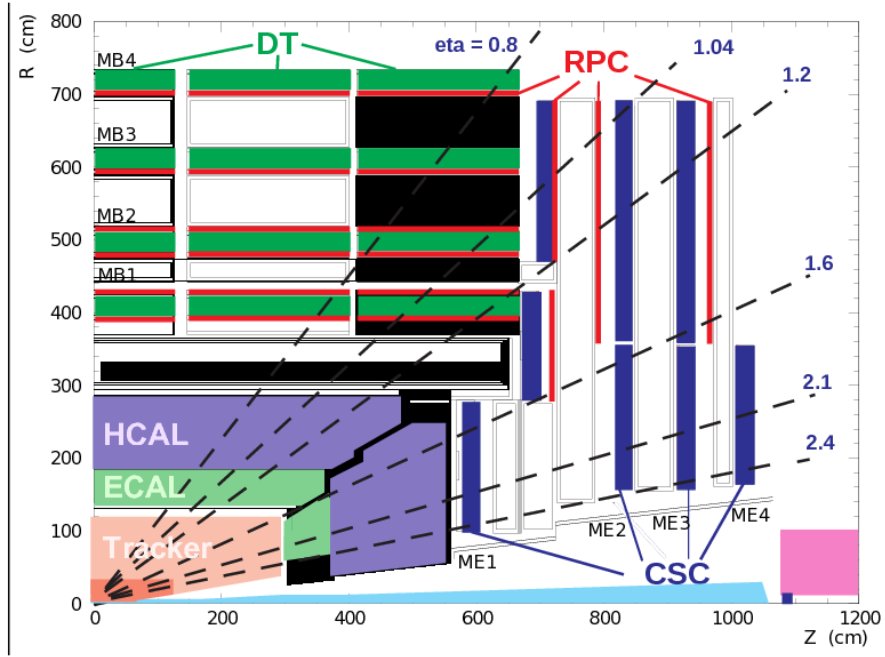


Figure 4.10: Geometric view of the CMS muon system, demonstrating the  $\eta$  barrel (MB1–MB4, green), the four CSC stations in the endcap (ME1–ME4, blue), and the RPC coverage and overlap of all constituent subsystems.

**The Drift Tubes (DT)** are placed in the barrel region with pseudorapidity coverage  $|\eta| < 1.2$ , where the particle flux is low enough and the magnetic field is sufficiently weak and homogeneous. The DT system layout follows the yoke segmentation and consists of 5 iron wheels composed of 12 azimuthal sectors, covering an angular region of approxima-

tively  $30^\circ$  each. Each wheel contains four concentric rings of chambers, called stations. Every station is formed by 12 DT chambers. The basic detector unit of the DT system is a rectangular drift tube cell with a transverse size of  $13 \times 42 \text{ mm}^2$  and whose length varies from 2 to 4 m. As the particles traverse the drift tube, they ionize a gas mixture of Ar (85%) and  $\text{CO}_2$  (15%). The ionized atoms are collected by the anode, creating a detectable electronic pulse. Distance of the charged particle from the anode wire can be deduced from the arrival time and shape of the electronic pulse using the value of the electric field generated by the anode of the square drift tube in the return field, which allows to determine the drift velocity of ionization electrons. To measure the charge particle's track over the full chamber, there are multiple drift tubes measuring this distance organized into two layers in the chamber. The spatial resolution of a DT chamber is  $100 \mu\text{m}$  in the  $r-\varphi$  plane, and  $150 \mu\text{m}$  in the  $z$  direction, with a drift time of 386 ns and timing resolution of 3.8 ns, using 3 consecutive staggered drift cells, as estimated using test beam data. The high timing resolution is important for triggering and allows for accurate assignment of identified muons to correct bunch crossings. A schematic layout of a single DT chamber is shown in Figure (4.11) [92].

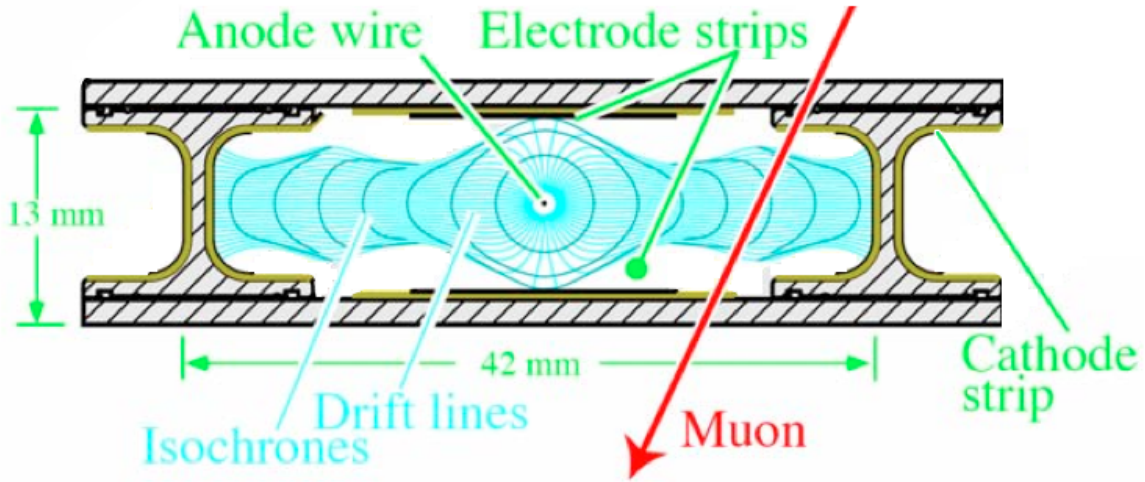


Figure 4.11: Sketch of a cell showing drift lines and isochrones. The plates at the top and bottom of the cell are at ground potential.

**The Resistive Plates Chambers (RPC)** are fast gaseous detectors that consist of two parallel plates, a positively-charged anode and a negatively-charged cathode, both made of a very high resistivity plastic material and separated by a gas (mostly  $\text{C}_2\text{H}_2\text{F}_4$ ) volume. When a muon passes through the chamber, electrons are knocked out of gas

atoms. These electrons in turn hit other atoms causing an avalanche of electrons. The electrodes are transparent to the signal (the electrons), which are instead picked up by external aluminum strips after a small but precise time delay. The pattern of hit strips gives a quick measure of the muon momentum, which is then used by the trigger to make immediate decisions about whether the data are worth keeping. RPCs combine a good spatial resolution with a time resolution of just one nanosecond [93].

**The Cathode Strip Chambers (CSC)** are located in the endcap disks, at  $0.8 < |\eta| < 2.4$ . The huge particle flux at large  $\eta$  and high magnetic field in this region do not allow usage of DT. Therefore, a solution based on CSC detectors has been adopted. CSCs are multiwire proportional chambers filled with a gas mixture of Ar (40%), CO<sub>2</sub> (50%), and CF<sub>4</sub> (10%) in which the cathode plane is clustered into strips oriented radially and transversely with respect to the anode wires. This allows the simultaneous measurement of two coordinates. These detectors can be placed in regions with higher flows of charged particles and less homogeneous magnetic fields since the drift path of the charge carriers is shorter in comparison to drift tubes. However, due to the spatial size of the strips, the timing resolution from this measurement is very poor, greater than 25 ns, and cannot be used for precision timing. To overcome this, the timing information from the anode wires is used, yielding timing resolution of 15 ns for a single plane. The CSC is composed of 4 superimposed disks (called stations), mounted on the iron disks of the return yoke and orthogonal to the direction of the beams. Each station is formed by two rings (three for the innermost station), divided into 18 or 36 CSCs with trapezoidal shape. Each CSC used in the system consists of six layers of sensitive wires. The strip spatial resolution is about  $80 - 85 \mu\text{m}$  while  $r$  can be determined with a precision of 0.5 cm.

### 4.3 The Trigger

The LHC produces events at a rate of 40 MHz at the interaction point in CMS. The average size of an event at design luminosity, including all the detector information, is roughly 1 MB. In order to save all the events, one would need to deal with 40 TB of data per second which represents a big challenge in terms of data storage and financial cost. However, most of the events produced are not interesting from the physics point of view, since they have low transverse momentum interactions. Therefore, a trigger

system is required to select only interesting events, at the design rate of events around 500 Hz which also substantially reduces the amount of data stored. CMS decided to go with a two-level trigger system, the Level-1 Trigger (L1) [94] and the High Level Trigger (HLT) [95]. The L1 trigger needs to reduce the data flow from 40 MHz to 100 kHz. Based on the coarse information from calorimetry and muon system, it has to decide whether the event is interesting or not. The HLT is responsible to reduce the L1 rate down to desired 500 Hz using information from all the subdetectors.

#### 4.3.1 Level-1 Trigger

The L1 is composed of dedicated processors which are designed to perform online physics calculations and reduce the data rate by two orders of magnitude. It relies only on the information of the calorimeters and muon system identifying muons, electromagnetic objects, missing transverse energy and jets. Due to the high input data rate, the information from the tracker is not used since it cannot be read out and processed fast enough. The trigger achieves its goal by using embedded reconstruction and identification algorithms to loosely select events with possibly interesting physics. The time available to make a decision is about 3  $\mu$ s. It consists of three main subsystems

- L1 Calorimeter Trigger
- L1 Muon Trigger
- L1 Global Trigger

The organization of CMS Level-1 Trigger is schematically illustrated in Figure (4.12) [96]. **Calorimeter Trigger** is relatively simple. The calorimeter information is first processed by the regional calorimeter trigger, which looks for clusters of signals collected both from ECAL and HCAL. It finds out electron/photon,  $\tau$  and jet candidates along with their transverse energy and sends this information to the Global Calorimeter Trigger. The Global Calorimeter Trigger sorts the candidates according to their transverse energy and sends the first four objects to the L1 Global Trigger.

**The L1 Muon Trigger** is a complex system: a pieces of information from RPC, CSC and DT specific triggers are combined in the so called L1 Global Muon Trigger. The RPC



trigger electronics builds Track Segments, gives an estimate of the  $p_T$  and sends these segments to the Global Muon Trigger. It also provides the CSC logic unit with information to solve hit position ambiguities in case two or more muon tracks cross the same CSC chamber. The CSC trigger builds track segments made out of the cathode strips only, where a  $p_T$  value and a quality flag is assigned to each segment. The best three segments in each sector of nine CSC chambers are passed to the CSC Track Finder, which uses the full CSC information to build tracks, assigns them a  $p_T$  and a quality flag, and sends them to the Global Muon Trigger. DT are equipped with Track Identifier electronics, which is able to find groups of aligned hits in the four chambers of a superlayer. Then the DT Track Finder builds tracks and sends them to the Global Muon Trigger. The Global Muon Trigger sorts the RPC, CSC and DT muon tracks and tries to combine them. The final set of muons is sorted according to their quality and the best four tracks are passed to the L1 Global Trigger.

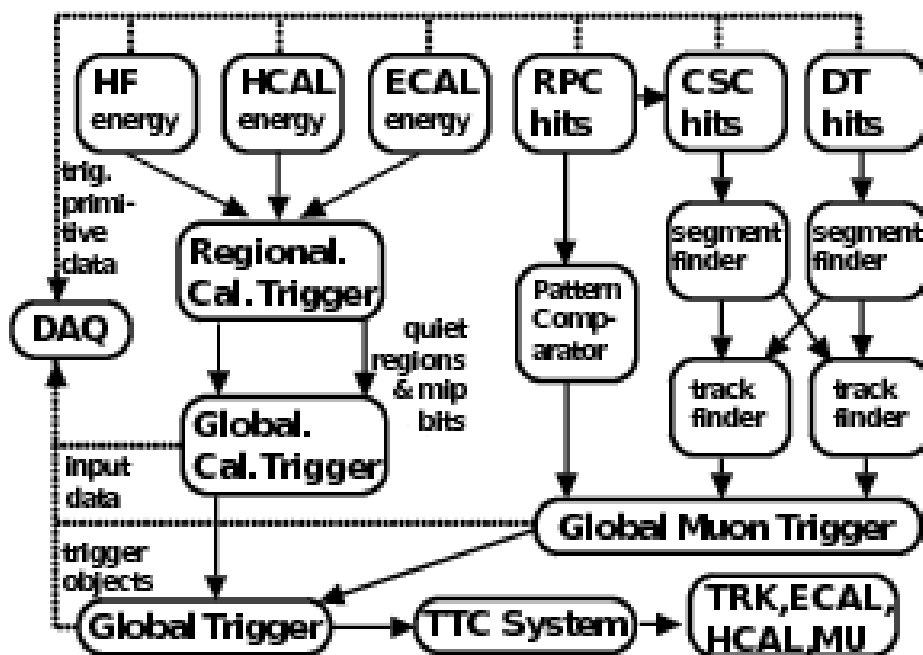


Figure 4.12: Level-1 Trigger components

The **L1 Global Trigger** is on the top of the chain and is responsible for making a decision whether to keep the event or not. If the event is accepted, the decision is sent to the Timing Trigger and Control System, that commands the readout of the remaining

subsystems. In order to make the decision, the L1 Global Trigger sorts the ranked objects produced by calorimetry and muon system and checks if at least one of the criteria in the Level-1 Trigger table is passed. L1 trigger table is composed of many different selection criteria based on e.g. muon  $p_T$ , number of muon candidates, amount of energy stored in calorimeter, etc.

#### 4.3.2 High level trigger (HLT)

The final level of event selection before writing to disk for offline analysis is the HLT. Its code runs on a computing farm, where it performs event reconstruction using the information from all subdetectors, including tracker which was not used at L1 trigger. The time available to make a decision is about 300 ms. The reliability of HLT algorithms is of capital importance, because events not selected by the HLT are lost. As in the case of L1 trigger, HLT also has a set of different selection criteria (decision paths) of which at least one needs to be satisfied. To reduce the overall CPU usage, the HLT retains a pipelined structure and stops executing a decision path when a step in that path fails. The L1 trigger completely determines which HLT decision paths will be run. If there are only L1 muon candidates present in the event and no interesting L1 electromagnetic candidates then only muon reconstruction algorithms will be run.



# Chapter 5

## Event simulation

The procedure for event simulation is described in this chapter.

### 5.1 Introduction

In order to describe interactions between elementary particles, the Standard Model is sufficient. However, at the LHC, the high energy colored partons are produced which results in considerably more complex final states and analytical description of those interactions is impossible. Solution for this problem are numerical techniques. They are used for simulating collision processes and also for implementation of many non-deterministic processes in detector physics such as shower stochastics, energy loss, and detector noise. The most widely used and flexible of these techniques are known as Monte Carlo (MC) algorithms, due to their reliance on random numbers. In the high energy physics community they are also known as event generators. Different event generators, such as PYTHIA8 [97], POWHEG [98] and MADGRAPH5 [99] implement computations at different levels of precision and with different techniques. Commonly, the highest precision calculations, that take into account several orders in perturbation theory, are only available for a limited number of processes, thus making it hard to derive predictions on inclusive quantities. On the other hand these quantities can often be described with reasonable precision with programs that implement lower order calculations. A schematic representation of the different components (and calculation steps) that are implemented in event generators is

shown in Figure (5.1), [100].

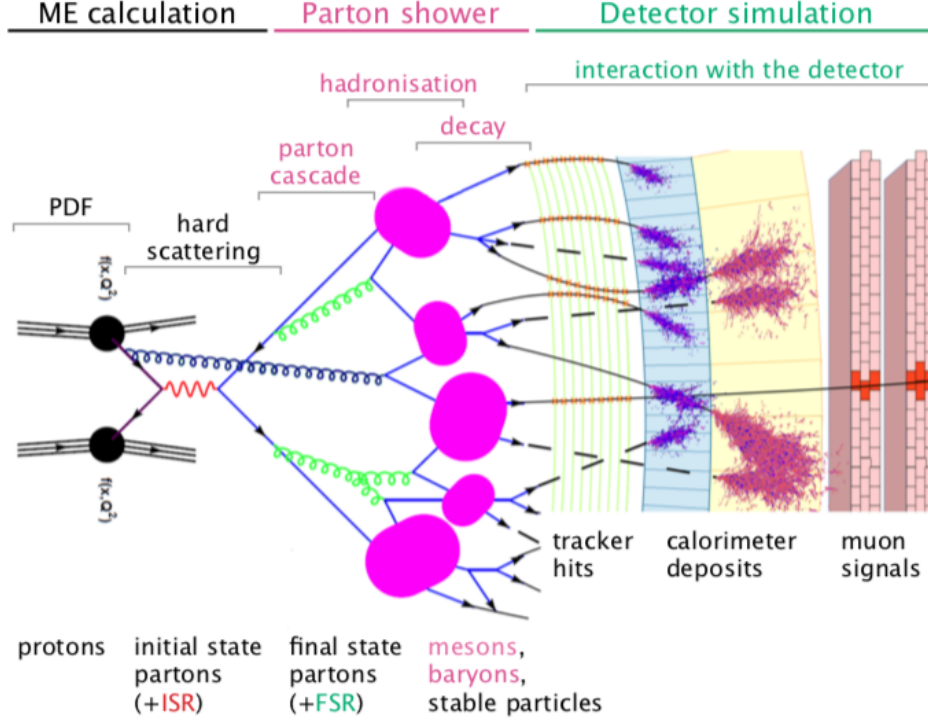


Figure 5.1: Illustration of the individual steps of the event simulation procedure.

The event simulation is the result of the following chain of calculations:

- Initial step is to compute the scattering matrix elements associated to the Feynman diagrams for the process of interest considering the incoming partons, extracted from the colliding hadrons as free particles.
- Hard scattering generation is actually the first step of event production in which two colliding hadrons with momenta are given. One parton of each hadron is selected to enter the scattering process. Using differential cross sections from the previous step, final state partons and leptons are produced.
- Partons which are involved in the hard event may emit additional gluons or photons due to their color and electric charge. If at the time of the hard scattering, parton involved in interaction emits gluon or photon, total momentum of all particles coming from this scattering will be different from zero because of momentum conservation. This effect is known as Initial State Radiation (ISR). Also, the resulting particle can

produce further radiation, Final State Radiation (FSR). They are simulated with the Parton Shower (PS) algorithm.

- Apart from partons which are involved in the hard interaction, other parton pairs could also interact during a hadron-hadron collision. These Multiple Parton Interactions (MPI) contribute to the underlying structure of the event (underlying event). Such interactions need to be simulated, in order to produce realistic events, and ISR and FSR need to be simulated for these collisions as well.
- All the final states after hard scattering are not observed free but are subjected to the hadronization process which is simulated with some empirical models resulting in colorless hadrons.
- In the end, particles with very short lifetime such as  $\tau$  and B-hadrons are allowed to decay by the generator itself. If they live enough to reach the detector, they are left undecayed.

## 5.2 Matrix element (ME)

Computation of the matrix elements translates to the order-by-order calculation of Feynman diagrams in perturbative QCD. Leading order (LO) or tree-level matrix element calculations still play an important role in the simulation of events produced at hadron colliders. They can be performed up to several (on the order of eight) partons in the final state. Next-to-leading order (NLO) is more precise with all the virtual loop corrections included up to given order in coupling constant. However, loop calculations are complex and they are available for a limited number of processes. The ME comes with few complications. Process with a high number of partons with low transverse momentum makes the calculation complicated due to soft divergences. Secondly, the method is not able to describe situations in which the emitted parton is collinear with respect to the radiating parton, which is known as collinear divergences. The loop corrections that should cancel these divergences are omitted in finite order calculations. Thus, the phase-space should be carefully chosen to avoid the problematic regions. In other words, the matrix element cross section calculations are performed away from soft and collinear divergences. To pro-

duce realistic events, phase-space regions omitted from the matrix element calculations have to be recovered, with care to avoid divergences.

### 5.3 Parton shower (PS)

Parton Shower (PS) algorithms simulate the evolution of a final state of colored partons into jets, together with the approximation of the effects of ISR and FSR of gluons, quark pairs and photons. This process is done in few steps. First, the parton is taken from the matrix element generator. The probability of the splitting processes  $q \rightarrow qg$ ,  $g \rightarrow q\bar{q}$  and  $g \rightarrow gg$  for some energy of the emitted parton are evaluated and stochastically applied. Once the parton splitting has occurred in the simulation, the same algorithm is applied recursively to all daughter partons until the energies of all partons are individually less than  $\Lambda_{QCD}$ . The introduction of a threshold value and the application of an angular sorting procedure in the emission of partons allows to eliminate soft and collinear divergences typical of the ME method. After parton cascade evolution, non perturbative effects take place and the hadronization is applied.

### 5.4 Underlying Event

The remaining constituents of collided hadrons which were not involved in the hard interaction are taking part in soft interactions among each other. They are simulated based on the multiple parton-parton interaction approach. Each interaction is represented by an independent perturbative calculation with non-perturbative form factors. The total number of interaction is extracted from a Poisson distribution based on the parton density and the parton-parton cross section.

### 5.5 Merging Matrix Element and Parton Shower

Matrix Element and Parton Shower algorithm are fundamental building blocks for simulation of collisions at LHC. As already mentioned, ME is exact to a given order in per-

turbation theory. It describes well separated parton configurations but it cannot describe the internal structure of a jet due to divergences in the soft and collinear regions. Since hadrons are observed in the experiments, fragmentation models need to be applied to partons. On the other side, PS is universal. Given the basic hard process, it will produce realistic parton configurations. PS is derived in the collinear limit and handles divergences by requiring conservation of total probability, which is particularly useful for describing jet evolution. Additionally, it can be used to evolve partons down to a common scale which removes the need of tuning fragmentation models at different scales. Since they are derived in the collinear approximation, they may fail in efficiently filling the phase-space for well separated parton configurations. The ME and PS are complementary procedures and their combined application in the intermediate cases allows to exploit the characteristics of the two algorithms in their respective limits of validity. Several prescriptions exist to perform ME-PS matching avoiding double-counting or holes in the phase-space, such as MLM [101] or CKKW [102].

## 5.6 Hadronization

Hadronization is the step in which PS partons are turned into hadrons. The process is non-perturbative and it is described by several phenomenological models. One of them is The Lund string model [103] and it is used by the PYTHIA generator. In a nutshell, the model treats all but the highest-energy gluons as field lines, which are attracted to each other due to the gluon self-interaction, and therefore form a narrow tube (or string) of strong color field. In order to describe the energetic flow in a Lorentz-covariant way, the string is considered massless. With increasing distance between quarks, their potential energy is growing on the account of kinetic energy. In the situation where  $g \rightarrow q\bar{q}$  splittings are not allowed, this process would continue until the endpoint quarks have lost all their momentum and they would reverse direction due to acceleration by the now shrinking string. However, in this model quark-antiquark fluctuations inside the string field can make the transition to become real particles by absorbing energy from the string, therefore screening the original endpoint charges from each other and breaking the string into two separate color-singlet pieces.



## 5.7 Pile-Up Interactions

Event simulation has to take into account not only proton-proton collision that caused the hard process but also other proton-proton collisions occurring in the same or nearby bunch crossing with small momentum transfer. These additional collisions are known as pile-up. There are two different types of pile-up contributions: in-time pile-up and out-of-time pile-up. The in-time pile-up is result of additional proton-proton collisions in the same bunch crossing. Out-of-time pile-up, on the other side, is caused by proton-proton collision from previous or following bunch crossings due to finite decay time of detector signals and the fact that some detectors integrate over more than one bunch crossing. The out-of-time pile-up is described by separately simulating minimum-bias collisions including all of the steps described in the previous sections. For a proper description of out-of-time pile-up, the pulse shapes of subdetectors are considered. For each event, a random instantaneous luminosity is drawn from the luminosity profile. Together with the total inelastic proton-proton cross section, an expected number of interactions for this luminosity is determined while the number of interactions used for pile-up is chosen randomly from a Poisson distribution with a mean value at the expected number of interactions.

## 5.8 Detector simulation

The simulation of detector response is the final step in event generation where simulated particles are passed through a simulated version of the CMS detector. The entire detector simulation is based on the software package GEANT4 [104] which has a variety of models and parameterizations used to describe the interaction of particles with matter. It takes into account the geometry of the detector, materials comprising sensitive detector elements, readout electronics together with electronic responses, uninstrumented structural material and a detailed model of the magnetic field provided by the solenoid, which is derived from measurements taken from the real detector, allowing for accurate simulation of the trajectories of particles. GEANT can statistically model how different particles interact, both electromagnetically and strongly, with the detector materials, predicting a variety of distributions such as charge depositions in doped silicon and raw light yields in scintillating material. After the raw information that various detectors collect has been

generated, a detailed electronics simulation of every subdetector is applied. This allows for a fine grained and tuneable estimation of additional detector effects, such as analog noise, and yields a realistic representation of the detector output given the estimation of the input from GEANT. The simulated detector response is stored in the same raw data format as used by the CMS detector so that exactly the same reconstruction algorithms are used in simulation as in real data. In addition to the detector response, the simulation and generator information is saved and can be used to calculate measurement extrapolations, efficiencies and calibrations.



# Chapter 6

## Event reconstruction

Object reconstruction and identification is based on standard algorithms which are designed for all physics analyses in CMS. In this section, the techniques used for the reconstruction and identification of the physics objects of interest are described.

### 6.1 The Particle Flow technique

The raw signals collected by the various subsystems are not suited for physics analysis. In order to obtain meaningful objects, the Particle Flow (PF) event reconstruction technique [105] is used. PF aims to reconstruct and identify observable particles, i.e. electrons, muons, photons and hadrons by using information from all subdetectors. PF objects are then used to build jets, measure missing transverse energy  $E_T^{\text{miss}}$  or to create other more complex features, e.g. charge isolation, b-tag. The initial step of PF is the reconstruction of the elementary objects, such as charged-particle tracks, muon tracks and calorimeter clusters. The tracking system has a vital role in the PF and is achieved by means of an iterative tracking strategy based on the Kalman Filter algorithm. The calorimeter clustering algorithm is performed separately in each subdetector and is divided into three steps. In the first step, cluster seeds are identified as local calorimeter cells with an energy deposit above a given threshold. Next, topological clusters are grown from the seeds by merging cells with at least one side in common with a cell already in the cluster and with an energy above a given threshold. Since topological clusters usually give rise to as many

PF clusters as there are seeds, the energy of each cell is shared among the particle flow clusters according to the cell-cluster distance. All the elements from initial step of PF are then combined based on topological compatibility using a link algorithm. For instance, a calorimeter PF cluster is linked to a charged-particle track if the extrapolated position from the track to the calorimeter is compatible with the cluster boundaries. In general, PF candidates are identified in the following order:

- **Muon:** A global muon is declared as PF muon if its combined  $p_T$  measurement is compatible within 3 standard deviations with the one provided by the tracker alone.
- **Electron:** PF electron is identified as charged-particle track linked with one or more ECAL clusters.
- **Charged hadrons:** The remaining tracks are identified as PF charged hadrons. They can be linked to ECAL and HCAL clusters.
- **Photons and neutral hadrons:** clusters which are not linked with the tracks are identified as PF photons (ECAL clusters) and PF neutral hadrons (HCAL clusters).
- **Jets:** After the identification of all PF candidates in the event, PF jets are clustered.
- **Missing transverse momentum:** The last step is the reconstruction of the missing transverse momentum.

After each PF candidate has been identified, corresponding elements used for particular reconstruction are removed from the list. The reconstruction of each PF object is explained in more details in the following subsections.

## 6.2 Tracks and vertices reconstruction

Tracks in CMS are reconstructed through an iterative procedure based on the Kalman filter. After each step, hits associated to reconstructed tracks are removed from the hit collections which substantially reduces complexity and simplifies the following iterations. The reasoning for iterative tracking is that initial iterations search for tracks that are easiest to find (e.g. high  $p_T$ ) and the following ones are dedicated to find more complicated

tracks (e.g. low  $p_T$ , displaced tracks). The first iteration, where the majority of tracks are reconstructed, is designed to identify prompt tracks with  $p_T > 0.8$  GeV that have three hits in the three layers of the pixel detector. A second iteration is used to recover prompt tracks that have only two pixel hits. Further iterations aim at finding low- $p_T$  prompt tracks and tracks that originate outside the collision point, i.e. tracks produced by a secondary vertex. There are six iterations and the main differences between them lies in the configuration of the seed generation and the final track selection. Starting from the reconstructed hits, the track reconstruction is decomposed in four logical parts:

- **seeding** : Initial track candidates are obtained from any combination of two hits from different pixel detector layers that are compatible with the beam spot. The first estimate of the transverse momentum of the seed is required to be above a certain threshold.
- **pattern recognition** : This step is based on a combinatorial Kalman filter method. The filter proceeds iteratively from the seed layer, starting from a coarse estimate of the track parameters provided by the seed, and including the information of the successive detection layers one by one. In each layer, i.e. with every additional measurement, the track parameters are known with better precision, up to the last point, where they include the full tracker information.
- **track fitting** : In this step a fit of the trajectory is performed, using its associated hits and providing an estimate of the track parameters ( $p_T$ ,  $\eta$ ,  $\varphi$ ,  $d_{xy}$ , and  $d_z$ .)
- **selection** : Tracks are eventually selected based on quality requirements.

More details on the CMS tracks reconstruction can be found in the corresponding publications [106,107]. Once when reconstruction of tracks is finished, the following step is to determine location of vertices, i.e. points where tracks intersect [108,109]. Vertices are the origins of the observed particles and accordingly hint at the occurrence of a physics interaction. There are three types of vertices: the primary-interaction vertex, additional pile-up-interaction vertices, and secondary vertices. The primary-interaction vertex represents the proton-proton collision with the largest momentum transfer. This vertex is assumed to include the most interesting hard process.

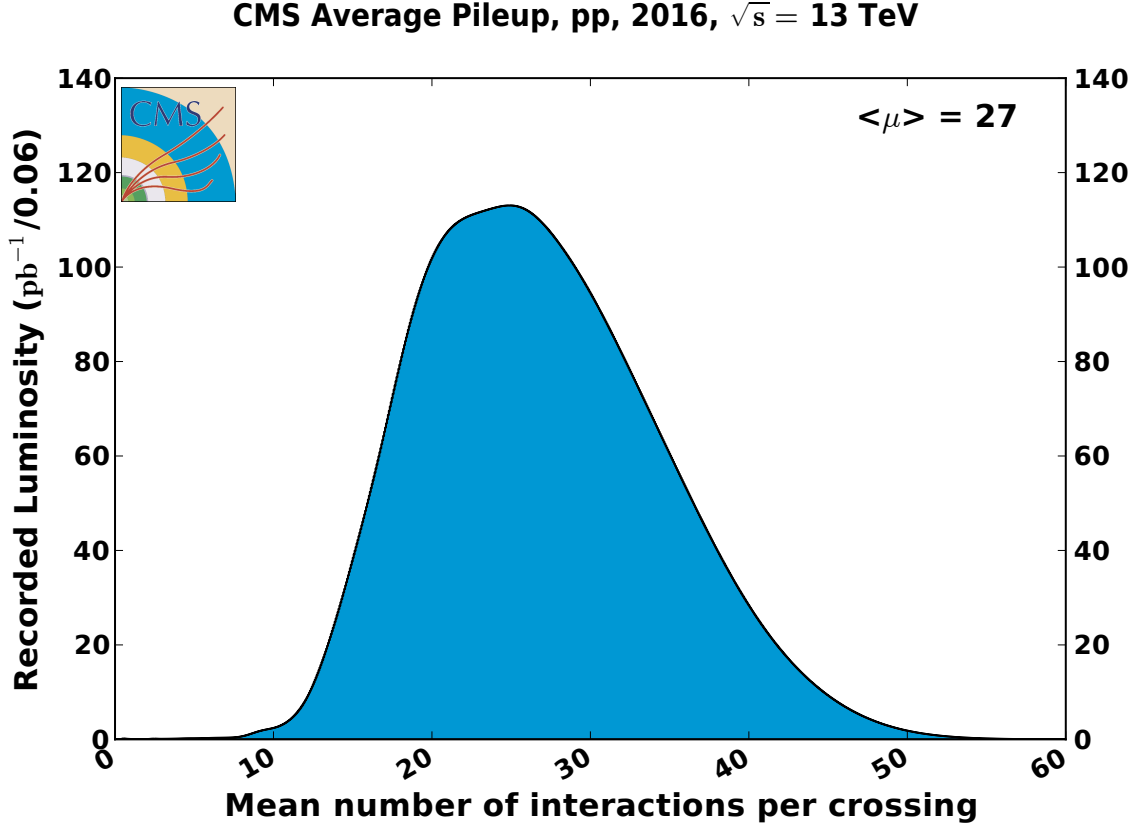


Figure 6.1: Mean number of interactions per bunch crossing for the 2016 proton-proton (pp) collisions at 13 TeV.

The pile-up vertices are caused by additional proton-proton collisions in the collision event and introduce noise in the form of additional particles. The mean number of pile-up vertices per bunch crossing for the 2016 proton-proton collisions is shown in Figure (6.1), [82]. Secondary vertices originate from the delayed decays of hadrons. Therefore, these vertices are displaced with respect to the primary-collision vertex. Appearance of secondary vertices is an essential element in the identification of B-hadron decays, also referred to as b-tagging, which is crucial for this analysis.

### 6.3 Electrons

The electron reconstruction is based on the information from tracking system and ECAL. The process begins by measuring the energy deposits in ECAL, which form a supercluster (SC). A SC is a collection of one or more ECAL clusters associated using an algorithm that

takes into account the characteristic shape of the energy deposited by electrons emitting bremsstrahlung radiation in the tracker material. Initial step in a reconstruction of SC is the identification of the crystal with the highest energy deposit, which becomes the seed for SC. Its shape is specified by a narrow width profile in the  $\eta$  coordinate spread over the  $\phi$  direction. Superclusters are matched to tracks reconstructed in the tracker with the GSF algorithm [110] in order to obtain electron candidates. To identify electrons and distinguish them from other particles, several variables are used:

- $\Delta\eta_{\text{trk,SC}}$  and  $\Delta\phi_{\text{trk,SC}}$  measure the spatial matching between the track and the supercluster in the  $\eta$  and  $\phi$  coordinates.
- $\sigma_{i\eta,i\eta}$  quantifies calorimeter shower shape by measuring the width of the ECAL supercluster along the  $\eta$  direction computed for all the crystals in the  $5\times 5$  block of crystals centered on the supercluster seed which is the highest energy crystal of the SC.
- $\mathbf{H}/\mathbf{E}$  measures the ratio between the energy deposited in the HCAL tower behind the SC seed and SC energy.
- $\mathbf{I}_{\text{ECalPF,SC}}$  and  $\mathbf{I}_{\text{HCaIPF,SC}}$  are ECAL and HCAL PF Cluster Isolation variables, which are explained in the following chapter.
- **dR03TkSumPt** represents the sum of tracks transverse momenta for all tracks that are within  $dR < 0.3$  of the electron candidate.
- $\mathbf{d}_{xy}(\mathbf{PV})$  and  $\mathbf{d}_z(\mathbf{PV})$  are the transverse and longitudinal impact parameters with respect to the primary vertex.

Electrons are first preselected by requiring  $p_T > 7$  GeV,  $|\eta| < 2.4$ ,  $d_{xy} < 0.05$  cm,  $d_z < 0.2$  cm and a very loose relative isolation cut of 0.4. In order to reproduce the selection applied by the most common electron triggers, additional requirements given in Table (6.1) are used. In the end, a tighter identification is applied using a multivariate approach [111].



Table 6.1: Preselection requirements used for electron identification.

Observable	Selection	
	$ \eta_{SC}  \leq 1.4442$	$ \eta_{SC}  \geq 1.5660$
$p_T$	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$
$\sigma_{i\eta, i\eta}$	$< 0.012$	$< 0.033$
H/E	$< 0.09$	$< 0.09$
dR03TkSumPt/ $p_T$	$< 0.18$	$< 0.18$
$I_{\text{ECalPF,SC}}/p_T$	$< 0.4$	$< 0.45$
$I_{\text{HCalPF,SC}}/p_T$	$< 0.25$	$< 0.28$
$\Delta\eta_{\text{trk,SC}}$	$< 0.0095$	-

Multivariate ID discriminator is provided with two different working points based on the expected selection efficiency. The working point with 80 % selection efficiency is used in this analysis to suppress the fake background in that final state, with an additional  $p_T$  threshold of 30 GeV.

## 6.4 Muons

Muons which are detected in the CMS detector can have two different origins. They are either coming from cosmic radiation, which penetrated through Earth's surface, or they are produced in collisions. Cosmic muons are useful for the calibrations of subdetectors, e.g. for the tracker. For physics analysis, muons produced at the collision point are used. They can go through the entire detector without significant energy loss. Muon tracks are reconstructed by the tracker (tracker tracks) and independently by the muon system (standalone muon tracks). Based on these objects, two reconstruction approaches can be used: in the first method (outside-in), for each standalone muon track a tracker track is searched for by extrapolating the two tracks onto a common surface. In the case of positive match, they are merged to a global muon. The second approach (inside-out) consists in considering all tracker tracks with  $p_T > 0.5 \text{ GeV}$  as potential muon candidates. They

are extrapolated to the muon chamber system taking into account the magnetic field, the expected energy losses, and the multiple scattering in the detector material. If at least one muon segment matches the extrapolated tracks, the corresponding tracker track is identified as a tracker muon.

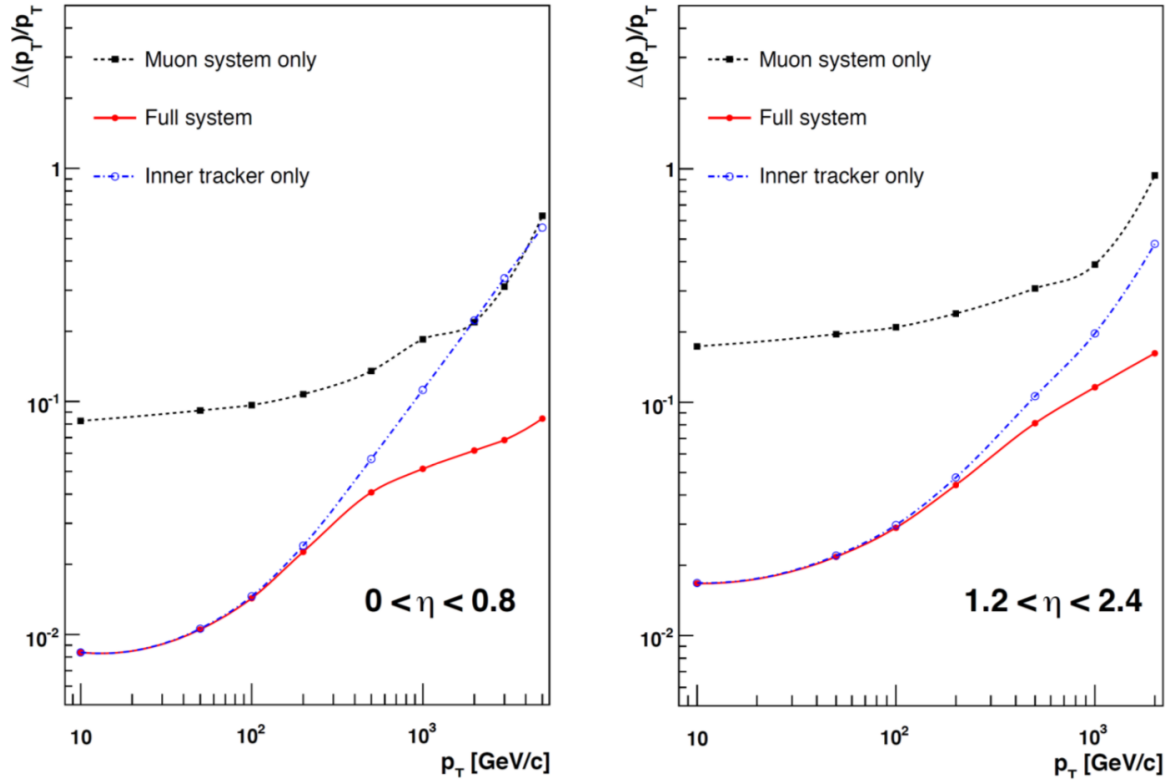


Figure 6.2: Muon  $p_T$  resolution as a function of the muon  $p_T$  in the barrel (left) and in the endcap (right) region.

Measuring tracks with two independent systems improves significantly the muon  $p_T$  resolution, especially in the region with  $p_T > 200$  GeV, as shown in Figure (6.2) [92]. In general, muon ID depends on the type of the physics analysis. Different muon definitions can be made by changing requirement on muon observables. The most widely used definition in physics analyses at CMS, which is the one used in this analysis, is the so-called Tight muon selection. The requirements defining the Tight Muon identification are summarized in Table (6.2).

Table 6.2: Summary of the muon identification variables and the corresponding selection.

Observable	Selection
Is Global muon	True
Is PF muon	True
Track layers with valid hits	$> 5$
Number of hits in Global muon track fit	$> 0$
Number of valid pixel hits	$> 0$
Number of matched muon stations	$> 1$
$\chi^2/d.o.f.$	$< 10$
$d_{xy}(PV)$	$< 0.2 \text{ cm}$
$d_z(PV)$	$< 0.5 \text{ cm}$
$p_T$	$> 25 \text{ GeV}$
$I_{PF}$	$< 0.06$

## 6.5 Lepton isolation

Lepton isolation is used for differentiating prompt muons and electrons produced in the decay of massive particles such as Z or W bosons and the ones produced in jets through the decay of hadrons. It is quantified by the total  $p_T$  of the particles emitted around the direction of the lepton. The particle-flow based isolation relative to the lepton  $p_T$  is defined as [88]

$$I_{PF} \equiv \frac{1}{p_T^\ell} \left( \sum p_T^{\text{charged}} + \max \left[ 0, \sum p_T^{\text{neutral}} + \sum p_T^\gamma - p_T^{\text{PU}}(\ell) \right] \right) \quad (6.1)$$

There are three independent sum terms over: charged hadrons ( $p_T^{\text{charged}}$ ), photons ( $p_T^\gamma$ ), and neutral hadrons ( $p_T^{\text{neutral}}$ ). In all three sums only particles with distance  $\Delta R$  to the lepton smaller than 0.3 (0.4) are considered. The  $\sum p_T^{\text{charged}}$  is the scalar sum of the transverse momenta of charged hadrons originating from the chosen primary vertex of the event. The  $\sum p_T^{\text{neutral}}$  and  $\sum p_T^\gamma$  are the scalar sums of the transverse momenta for neutral hadrons and photons, respectively. Since the isolation variable is particularly sensitive to

energy deposits from pileup interactions, a  $p_T^{\text{PU}}(\ell)$  contribution is subtracted, using two different techniques. For muons, it is defined as  $p_T^{\text{PU}}(\mu) \equiv 0.5 \times \sum_i p_T^{\text{PU},i}$ , where  $i$  runs over the momenta of the charged hadron PF candidates not originating from the primary vertex, and the factor of 0.5 corrects for the fraction of charged and neutral particles in the cone. For electrons, it is defined as  $p_T^{\text{PU}}(e) \equiv \rho \times A_{\text{eff}}$ , where the effective area  $A_{\text{eff}}$  is the geometric area of the isolation cone scaled by a factor that accounts for the residual dependence of the average pileup deposition on the  $\eta$  of the electron, and  $\rho$  is the median of the  $p_T$  density distribution of neutral particles within the area of any jet in the event. Both muon and electron candidates in this analysis are required to have a relative isolation smaller than 0.06.

## 6.6 Jets

Jets, as the experimental signature of quarks and gluons produced in high energy processes, are very important for the measurements presented in this thesis. They are used for testing predictions of high-energy QCD processes, and also for identifying the hard partonic structure of decays of massive particles. Jet algorithms cluster partons, or particles, or sets of reconstructed objects through sequential iterative clustering procedures that use only the four-momenta of input objects. There are several different algorithms available, characterized by different features. From a "theoretical standpoint", a good jet clustering algorithm should contain following features [112]:

- **Infrared safety** - the solution of the algorithm should be insensitive to soft radiation in the event.
- **Collinear safety** - collinear singularities should not exist in the perturbative calculations and the algorithm should be insensitive to collinear radiation in the event.
- **Invariance under boosts** - there should be no dependence on boosts in the longitudinal direction, i.e. direction along the beam of colliding particles.
- **Order independence** - the algorithm should find the same jets at parton, particle and detector level.

Some experimental principles should be followed as well: its performance should be as independent as possible of the detector, the algorithm should not amplify the inevitable effects of resolution smearing, and should not be strongly affected by pile-up and high beam luminosities. Additionally, the algorithm should be easy to implement, efficient in jet identification and in the usage of computing resources. There are two main classes of jet clustering algorithms. The first one consists in the recombination, where jets are reconstructed associating together particles whose trajectories lie within a cone of radius  $\Delta R$  in the  $\eta - \phi$  plane. The second class of algorithms uses the sequential recombination scheme, that iteratively recombine the closest pair of particles according to some distance measure.

### The $k_t$ and anti- $k_t$ Algorithms

The  $k_t$  and anti- $k_t$  are IRC-safe sequential recombination scheme algorithms. For their description, it is necessary to introduce two definitions of distances:  $d_{ij}$ , the distance between two objects (e.g. particles)  $i$  and  $j$ , and  $d_{iB}$ , the distance between the object  $i$  and the beam. They are defined as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}$$

$$d_{iB} = k_{ti}^{2p}$$

where  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  and  $k_{ti}$ ,  $y_i$  and  $\phi_i$  are the transverse momentum, rapidity and azimuth of particle  $i$ . In the two expressions,  $R$  is the distance parameter and  $p$  is a parameter that sets the energy relative power versus the geometrical scale  $\Delta_{ij}$ . The sequential clustering algorithms work by first finding the minimum of the entire set  $(d_{ij}, d_{iB})$ . If  $d_{ij}$  is the minimum then particles  $i$  and  $j$  are combined into one particle  $(ij)$  using summation of four-vectors after which  $i$  and  $j$  are removed from the list of particles. If  $d_{iB}$  is the minimum,  $i$  is labelled a final jet and removed from the list of particles. This process is repeated until either all particles are part of a jet with the distance between the jet axes  $\Delta_{ij}$  greater than  $R$ , which is inclusive clustering. Or until a desired amount of jets have been found, this is exclusive clustering. By changing the value of the  $p$  parameter, different jet clustering algorithms are obtained:

- 1 for the  $k_t$  algorithm
- -1 for the anti- $k_t$  algorithm [113]
- 0 for Cambridge/Aachen algorithm (CA) [114]

The difference between the three algorithms is the momentum weighting. For the  $k_t$  algorithm, the weighting proportional to  $k_t^2$  implies that jets are reconstructed starting from particles with low transverse momenta. Additionally, this algorithm produces jets with irregular border and therefore complicates the correction of effects like pile-up. In the case of CA algorithm, particles are merged based on the distance  $\Delta_{ij}$  only. As with  $k_t$ , CA leads to jets with irregular borders.

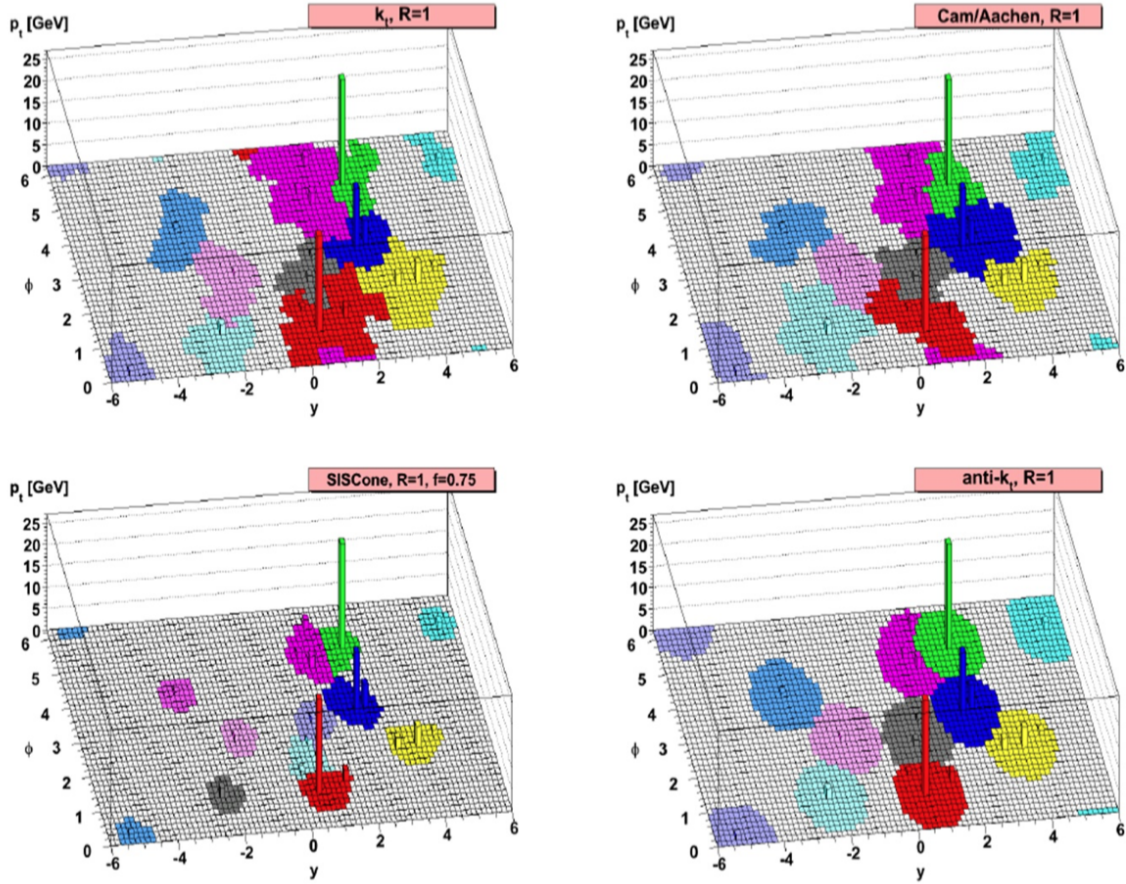


Figure 6.3: Jets reconstructed with different algorithms starting from the same set of simulated particles.

Finally, the anti- $k_t$  algorithm uses a weighting proportional to  $1/k_t^2$ . In this case the jets are created around particles with the highest transverse momenta and have a circular

shape. Jets reconstructed with different algorithms starting from the same set of simulated particles are shown in Figure 6.3 [115]. In the measurements presented in this thesis, the anti- $k_t$  algorithm is used for two types of jets ( $R = 0.4$  (AK04) and  $R = 0.8$  (AK08)) and they are built using Particle Flow objects, as shown in Figure 6.4. Compared to other jets, for instance jets built using calorimetric clusters, they have significantly better momentum and spatial resolution mostly due to inclusion of tracker information.

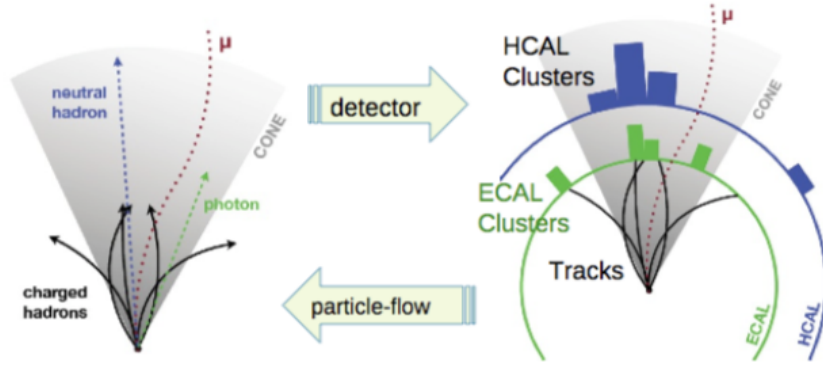


Figure 6.4: The PF algorithm attempts to fully reconstruct jet by combining information from all CMS subdetectors.

### 6.6.1 Jet energy corrections

Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum, defined as the momentum of corresponding particle level jet, over the whole  $p_T$  spectrum and detector acceptance. Jet energy corrections are derived from simulation studies so that the average measured response of jets becomes identical to that of particle level jets. A mismatch is mainly caused by the nonuniform and nonlinear response of the CMS calorimeters, electronics noise, and pile-up. For that purpose, CMS has developed a procedure to calculate and apply the jet energy corrections (JEC) [116]. The corrections are applied as a multiplicative factor  $\mathcal{C}$  to each component of the raw jet four-momentum

$$p_\mu^{cor} = \mathcal{C} \cdot p_\mu^{raw}$$

where  $p_{\mu}^{cor}$  is the corrected jet four-momentum. The correction factor is made of the offset correction  $C_{offset}$ , the MC calibration factor  $C_{MC}$ , and the residual calibrations  $C_{rel}$  and  $C_{abs}$  for the relative and absolute energy scales, respectively. The  $C_{offset}$  removes the extra energy due to noise and pile-up,  $C_{MC}$  removes the nonuniformity in  $\eta$  and the nonlinearity in  $p_T$ , and the residual corrections account for the small differences between data and simulation. The various components are applied in exact sequence as described by the equation below

$$\mathcal{C} = C_{offset}(p_T^{raw}) \cdot C_{MC}(p_{T,1}, \eta) \cdot C_{rel}(\eta) \cdot C_{abs}(p_{T,2})$$

where  $p_{T,1}$  is the jet momentum after applying the offset correction and  $p_{T,2}$  is the jet momentum after applying all previous corrections.  $C_{offset}$  and calibration factors  $C_{rel}$  and  $C_{abs}$  are applied to both data and simulation while  $C_{MC}$  is applied only to data.

### 6.6.2 Jet energy resolution

Studies have shown that the jet energy resolution (JER) in simulation is better than in data. Therefore the simulated jets need to undergo a smearing procedure in order to have a better agreement with the data. Reconstructed jets in simulated events are corrected in a two step procedure. Firstly, the reconstructed jet  $p_T$  is scaled for the observed  $p_T$  difference between reconstructed and particle level jets, also called generated jets. This step works only for jets that are matched to generated ( $\Delta R$  and  $\Delta p_T$ ) jets. In the second step, a gaussian smearing of the  $p_T$  distribution of the reconstructed jet is applied in order to get the desired resolution. If matching requirement to generated jets failed, only the second step to reconstructed jet is applied.

### 6.6.3 Jet identification

To avoid usage of fake jets that are originating from noisy calorimetric cells or electronics noise, some basic quality criteria for jets are introduced. They are collectively called jet identification (or Jet ID) and allow the rejection of badly reconstructed jets while maintaining a high fraction (about 99%) of real jets. Criteria are based on the following



variables: Jet  $p_T$ , number of constituents, charged hadron fraction, neutral hadron fraction, charged multiplicity, charged electromagnetic fraction, and neutral electromagnetic fraction. Additionally, to reduce the incidence of jets originating from pile-up, a pile-up jet identification algorithm is used. It makes use of both vertex information, exploiting the charged component of the jet, and jet shape information to identify jets belonging to pile-up.

### 6.6.4 Jet substructure

As a consequence of the high centre-of-mass energy of the LHC, particles can be produced with a large Lorentz boost. If those particles are decaying to a pair of quarks or gluons, the resulting jets are often overlapping. In order to capture all the decay products, the jets have to be reconstructed with a larger distance parameter. In this analysis  $R = 0.8$  (AK08) is used. Because of its size, AK08 jet is also known as "fat jet". Jet substructure techniques can then be applied to resolve the subjets corresponding to the decay products in the AK08 jet. The subjet axes are obtained by reclustering the jet constituents using the  $k_t$  algorithm and undoing the last step of the clustering procedure.

#### Subjettiness

A variable able to discriminate jet substructure is the N-subjettiness,  $\tau_N$  [117], which is a jet shape variable, computed under the assumption that the jet has N subjets, and it is defined as the  $p_T$  weighted distance between each jet constituent and its nearest subjet axis ( $\Delta R$ ):

$$\tau_N = \frac{1}{d_0} \sum_k p_T^k \min(\Delta R_{1,k}, \dots, \Delta R_{N,k})$$

where  $k$  runs over all jet constituents. The normalization factor is  $d_0 = \sum_k p_T^k R_0$  and  $R_0$  is the original jet distance parameter, i.e.  $R_0 = 0.8$ . The  $\tau_N$  variable has a small value if the jet is consistent with having N subjets. The subjet axes are determined by the exclusive- $k_T$  clustering algorithm, forcing it to return exactly N jets [118, 119]. As a result, the  $\tau_N$  axes, also called  $\tau$  axes, are obtained. These are then used to estimate

the directions of the partons giving rise to the subjets, as schematically illustrated in Figure 6.5.

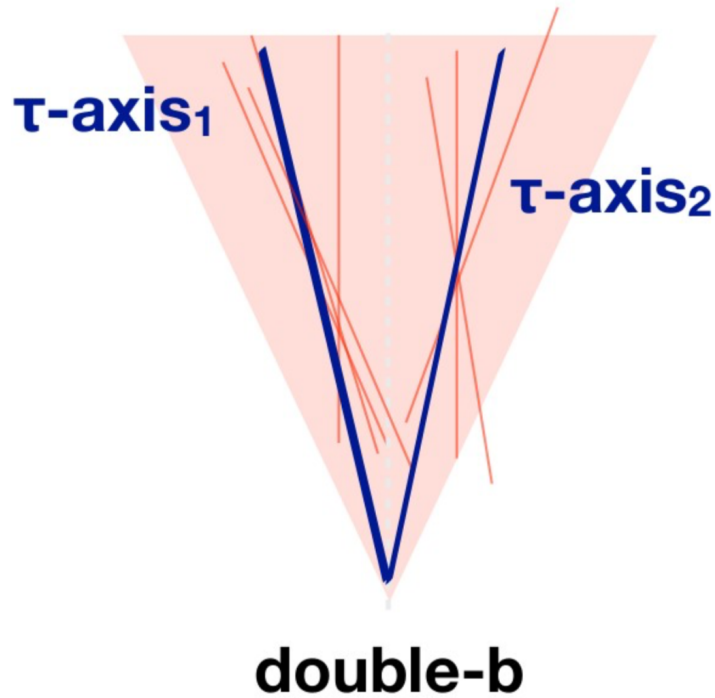


Figure 6.5: Schematic representation of the AK08 jet with corresponding  $\tau$  axes.

### Pruned mass

Pruning [120] is a technique designed for removing soft and wide-angle radiation. It will also remove the uncorrelated contributions from underlying event and pile-up that can make significant contributions to the jet mass. Pruning is based on the re-clustering of the fat-jet constituents. Instead of discarding soft subjets, pruning removes contamination by vetoing soft and large-angle recombinations during reclustering. The requirements for vetoing a recombination of two constituents  $j_1$  and  $j_2$  with  $p_{T,j_1} > p_{T,j_2}$  to a resulting jet  $j$  are

$$\frac{p_{T,j_2}}{p_{T,j}} < z_{\text{cut}}$$

$$\Delta R_{j_1,j_2} > D_{\text{cut}}$$

where  $z_{\text{cut}}$  represents a lower threshold for the transverse momentum of the softer constituent with respect to the combined jet. The parameter  $D_{\text{cut}}$  determines the minimum angular distance for a recombination to be pruned.

### 6.6.5 Heavy flavour jet

Algorithms for heavy-flavour jet identification exploit the long lifetime of heavy-flavour hadrons present in jets. The lifetime of hadrons containing  $b$  quarks is approximately 1.5 ps, while the lifetime of hadrons containing  $c$  quarks is less than 1 ps. As a consequence, displaced tracks appear within a jet. Their typical displacement is of the order of few millimeters for  $b$  hadron momentum of  $\mathcal{O}(10 \text{ GeV})$ . Using displaced tracks, a secondary vertex (SV) can be reconstructed, as shown in Figure 6.6. The displacement of tracks with respect to the primary vertex is characterized by the distance between the primary vertex and the tracks at their point of closest approach, also known as impact parameter. Compared to lighter quarks and massless gluons,  $b$  and  $c$  quark have harder fragmentation, which means that the decay products of the heavy-flavour hadron have, on average, a larger  $p_T$  relative to the jet axis than the other jet constituents. Additionally, in approximately 20% (10%) of the cases, a muon or electron is present in the decay of a heavy  $b$  ( $c$ ) hadron.

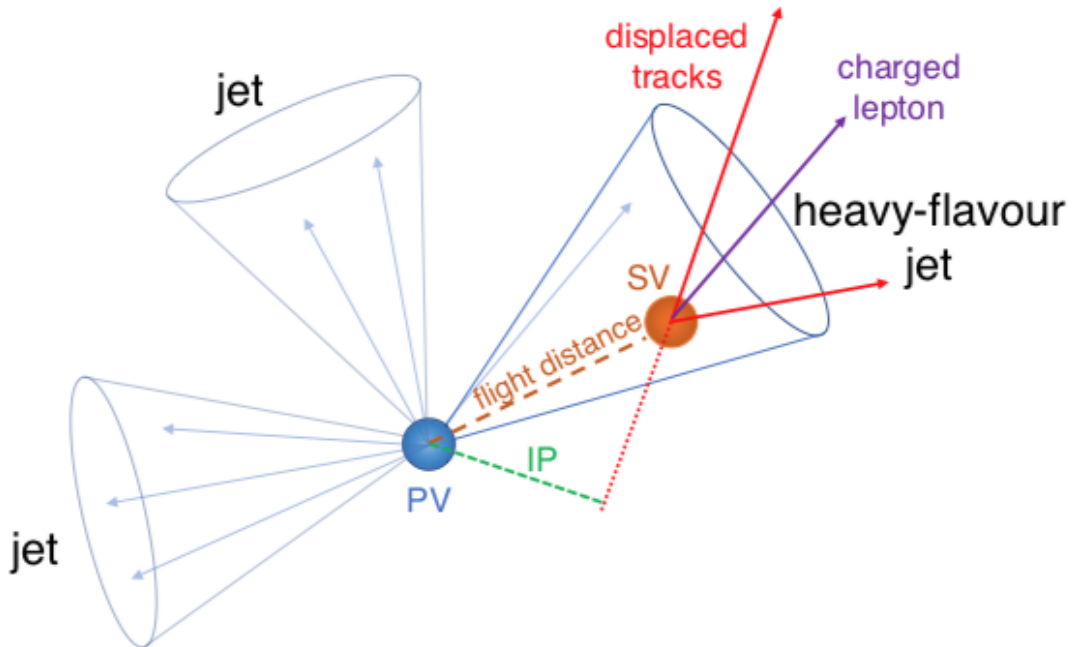


Figure 6.6: Illustration of a heavy-flavour jet with a secondary vertex (SV) from the decay of a heavy-flavour hadron resulting in charged-particle tracks that are displaced with respect to the primary interaction vertex (PV) and with a large impact parameter (IP) value.

Powerful heavy-flavour tagging algorithms can be built using a variety of observables based on tracks, vertices and identified leptons. Several algorithms have been developed by CMS [121], each one based on different input information. A common feature of all the algorithms is that each one yields a single discriminator value for every jet, which measures the likelihood that the jet has been produced by the hadronization of a heavy quark.

### b-tagging

For b-tagging (AK04 jets), in this analysis CSVv2 algorithm has been used, which is based on the Combined secondary vertex (CSV) algorithm [122] and combines the information of displaced tracks with the information on secondary vertices using a multivariate technique, i.e. neural network. Specifically a feed-forward multilayer perceptron with one hidden layer is used. To reconstruct secondary vertices, inclusive vertex reconstruction is exploited (IVR) [123] which makes use of all tracks in the event, with appropriate selections. More details of the algorithm are given in [121].

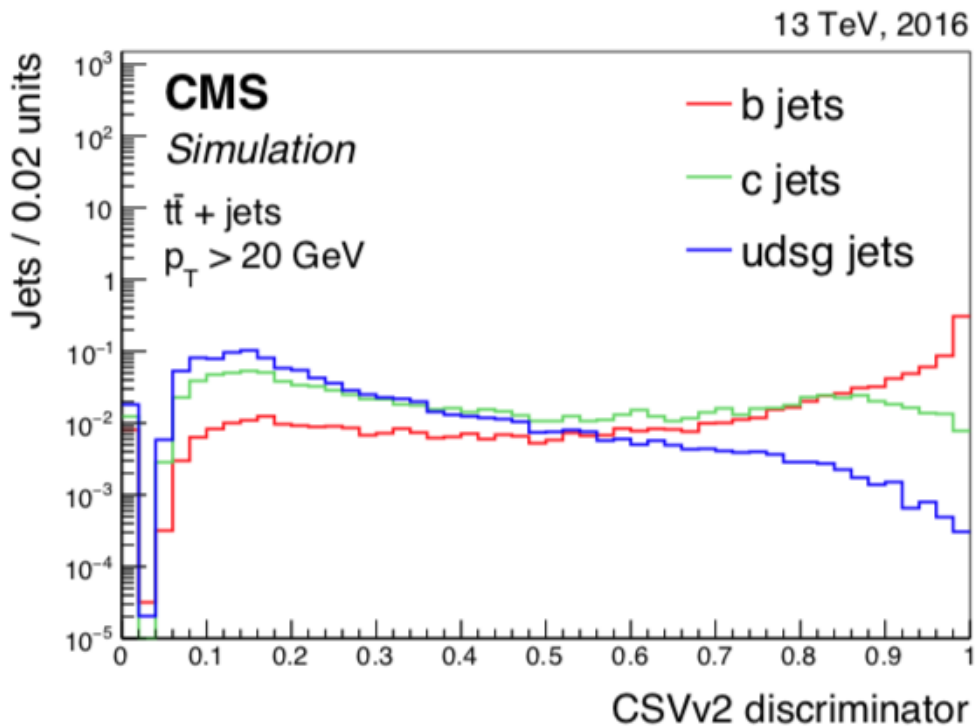


Figure 6.7: Distribution of the CSVv2 discriminator values for jets of different flavours in simulated  $t\bar{t}$  events. The distributions are normalized to unit area.

Figure 6.7 shows the distribution of the discriminator values for the various jet flavours in simulated  $t\bar{t}$  events. The tagging efficiency of the CSVv2 is determined using simulated pp collision events. The efficiency (misidentification probability) to correctly (wrongly) tag a jet with flavour  $f$  is defined as the number of jets of flavour  $f$  passing the tagging requirement divided by the total number of jets of flavour  $f$ . Figure 6.8 shows the b jet identification efficiency versus the misidentification probability for either c or light-flavour jets in simulated  $t\bar{t}$  events requiring jets with  $p_T > 20$  GeV and  $|\eta| < 2.4$ . Three standard working points are defined using jets with  $p_T > 30$  GeV in simulated multijet events with  $80 < p_T < 120$  GeV. These working points, "loose" (L), "medium" (M), and "tight" (T), correspond to thresholds on the discriminator after which the misidentification probability is around 10%, 1%, and 0.1%, respectively, for light-flavour jets. The efficiency for correctly identifying b jets in simulated  $t\bar{t}$  events for each of the three working points of the various taggers is summarized in Table 6.3.

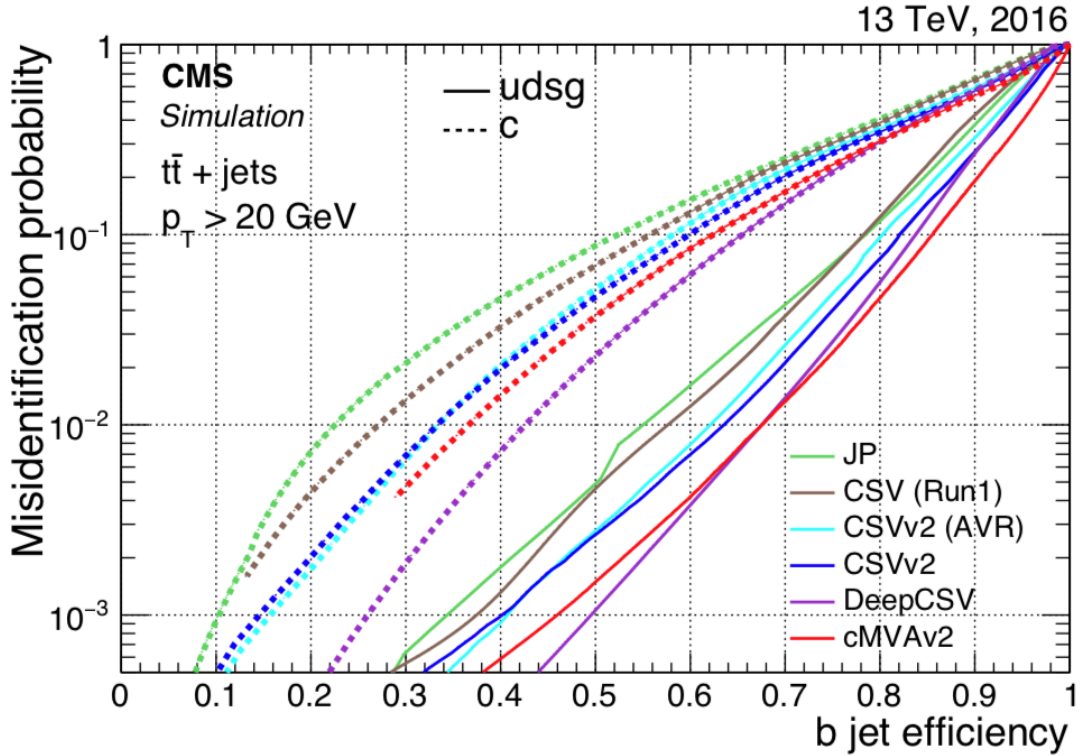


Figure 6.8: Misidentification probability for c and light-flavour jets versus b jet identification efficiency for CSVv2 and also other b tagging algorithms applied to jets in  $t\bar{t}$  events.

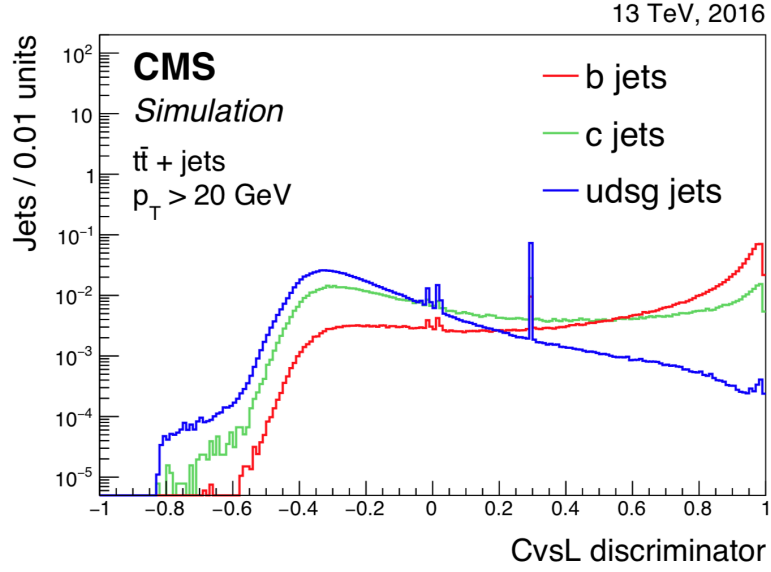
Table 6.3: Working points and corresponding efficiency for b jets with  $p_T > 20$  GeV in simulated  $t\bar{t}$  events. The numbers in this table are for illustrative purposes since the efficiency is integrated over the  $p_T$  and  $\eta$  distributions of jets.

Working point	$\epsilon_b(\%)$	$\epsilon_c(\%)$	$\epsilon_{uds g}(\%)$
CSVv2 (L)	81	37	8.9
CSVv2 (M)	63	12	0.9
CSVv2 (T)	41	2.2	0.1

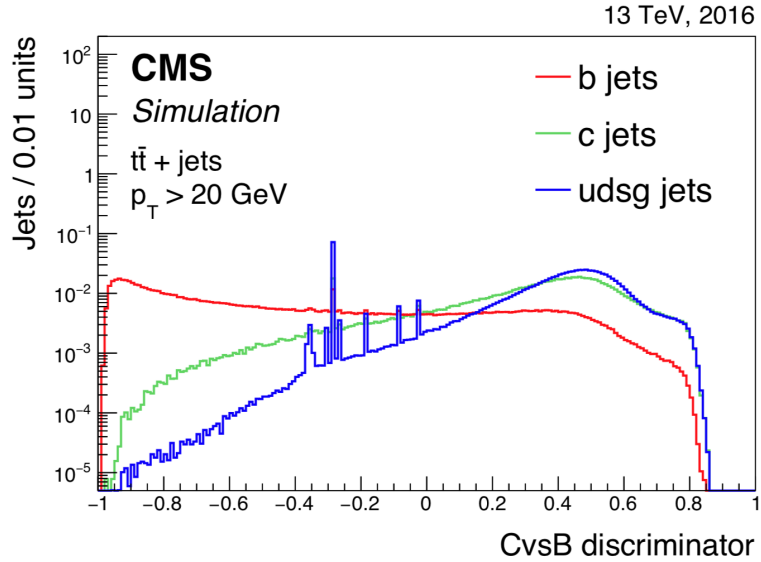
### c-tagging

Since the lifetime of c hadrons is shorter than that of b hadrons, but long enough to result in measureable displaced decay vertices, the distributions of the tagging variables for c jets lie in between the distributions for b and light-flavour jets, as can be seen from Figure 6.7. The secondary vertex multiplicity is for instance, lower and the smaller c quark mass results in a smaller track  $p_T$  relative to the jet axis. Therefore, it is particularly challenging to efficiently identify jets originating from c quarks. The c jet identification algorithm uses properties and variables similar to the ones used in the CSVv2 algorithm: displaced tracks, secondary vertices, and soft leptons inside the jets. Two algorithms were made for c-tagging on AK04 jets: one for discriminating c jets from light-flavour jets (CvsL) and another one for discriminating c jets from b jets (CvsB). The training of the two discriminators was performed using a multivariate technique gradient boosting classifier as implementation of the boosted decision trees. Figure 6.9 shows the output discriminator distributions for CvsL and CvsB. The performance of the c tagger is evaluated using jets with  $p_T > 20$  GeV and  $|\eta| < 2.4$  in a sample of simulated  $t\bar{t}$  events. Figure 6.10 shows the correlation between the CvsL and CvsB discriminators for various jet flavours. Discriminator values close to one for both discriminators correspond to signal-like c jets. Therefore, the c jets populate the upper right corner of this figure, whereas b jets and light-flavour jets populate the region near the bottom right and the upper left corners, respectively. In the upper left corner there is a relatively large fraction of c jets because of the similarity of c jets and light-flavour jets at CvsL discriminator values, as can be seen in Figure 6.10. In order to discriminate c jets from other jet flavours and to evaluate

the performance of the  $c$  tagger, thresholds are applied on both CvsL and CvsB to select the upper right corner of this phase space. Three working points have been defined corresponding to the efficiency for correctly identifying  $c$  jets. These are indicated by the dashed lines. The loose working point has a high efficiency for  $c$  jets and rejects primarily  $b$  jets, whereas the tight working point rejects primarily light-flavour jets. Table 6.4 summarizes the efficiencies for the three working points.



(a) CvsL



(b) CvsB

Figure 6.9: Distribution of the  $c$ -tagging discriminator values for jets of different flavours in  $t\bar{t}$  events. The spikes originate from jets without a track passing the track selection criteria.

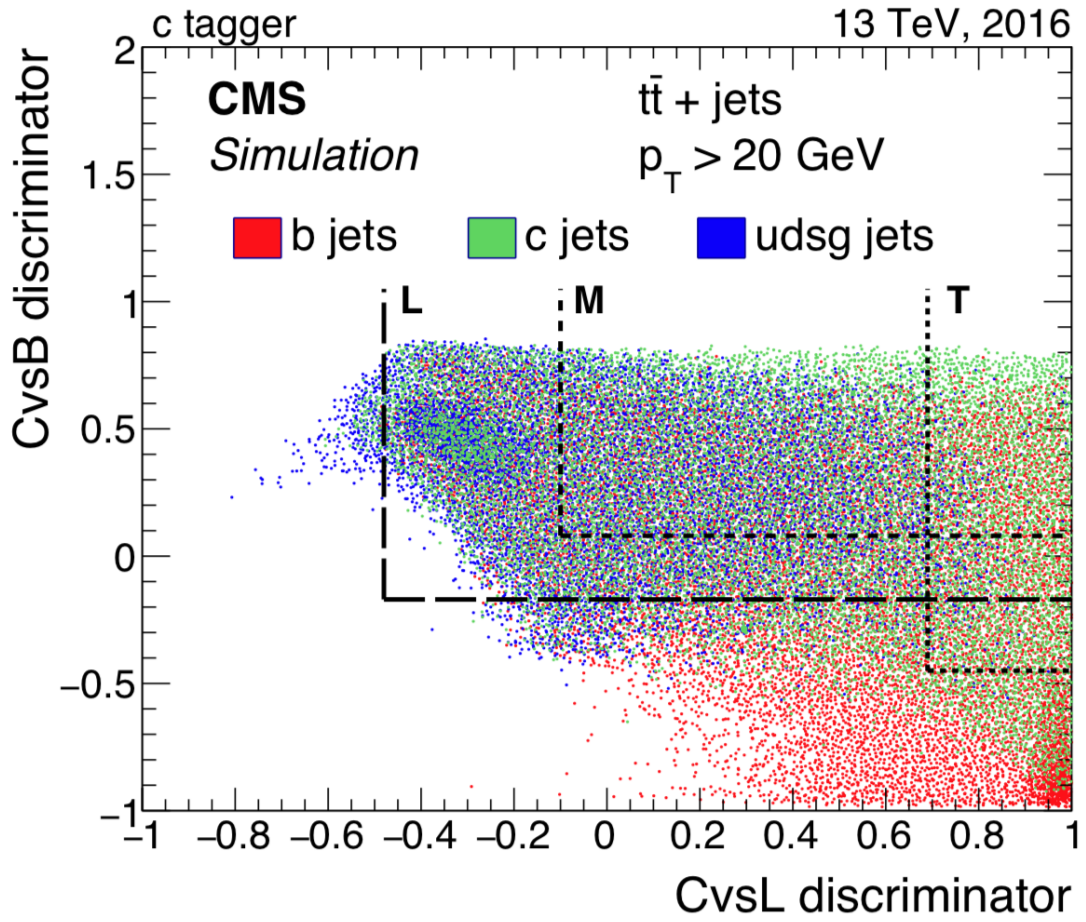


Figure 6.10: Correlation between CvsL and CvsB taggers for the various jet flavours. The L, M, and T working points discussed in the text are indicated by the dashed lines.

Table 6.4: Working points and corresponding efficiency for c jets with  $p_T > 20$  GeV in simulated  $t\bar{t}$  events. The numbers in this table are for illustrative purposes since the efficiency is integrated over the  $p_T$  and  $\eta$  distributions of jets.

Working point	$\epsilon_c(\%)$	$\epsilon_b(\%)$	$\epsilon_{udsg}(\%)$
c tagger (L)	88	36	91
c tagger (M)	40	17	19
c tagger (T)	19	20	1.2



## bb-tagging

A double b-tagging algorithm was developed for tagging AK08 jets with two b hadrons. It exploits not only the presence of two b hadrons inside the AK08 jet but also the correlation between the directions of their momenta. Any dependence of the algorithm performance on the mass or  $p_T$  of the  $b\bar{b}$  pair is avoided. This strategy allows the usage of the tagger in physics analyses with a large range of jet  $p_T$ . The dependence on the jet mass is avoided as this variable is often used to define a signal region. In addition, this strategy also permits the use of the double-b tagger for the identification of boosted  $Z \rightarrow b\bar{b}$  jets or any other boosted  $b\bar{b}$  resonance where the kinematics of the decay products are similar. The algorithm is based on similar variables as CSVv2. It relies on reconstructed tracks and secondary vertices obtained using the IVF algorithm. Tracks with  $p_T > 1$  GeV are associated with jets in a cone of  $\Delta R < 0.8$  around the jet axis. Additionally, each track is then associated with the closest  $\tau$  axis, where the distance of a track to the  $\tau$  axis is defined as the distance at their point of closest approach.

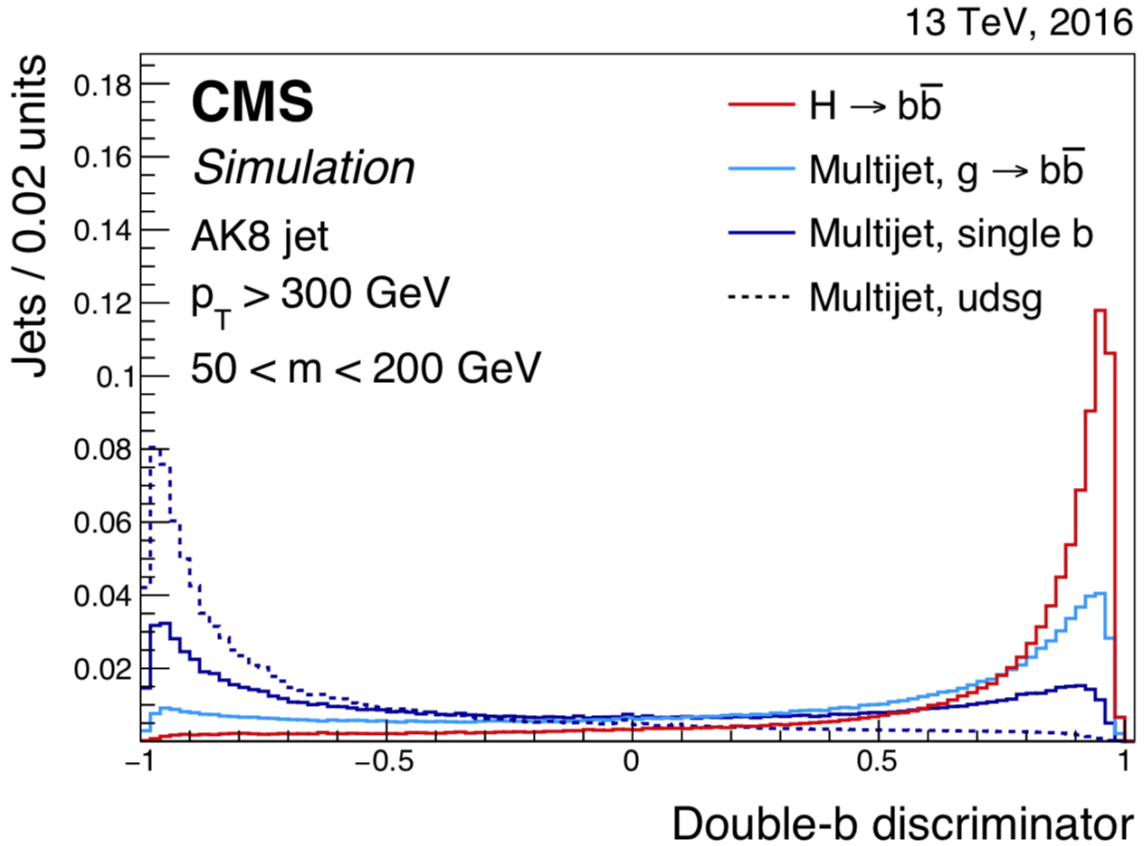
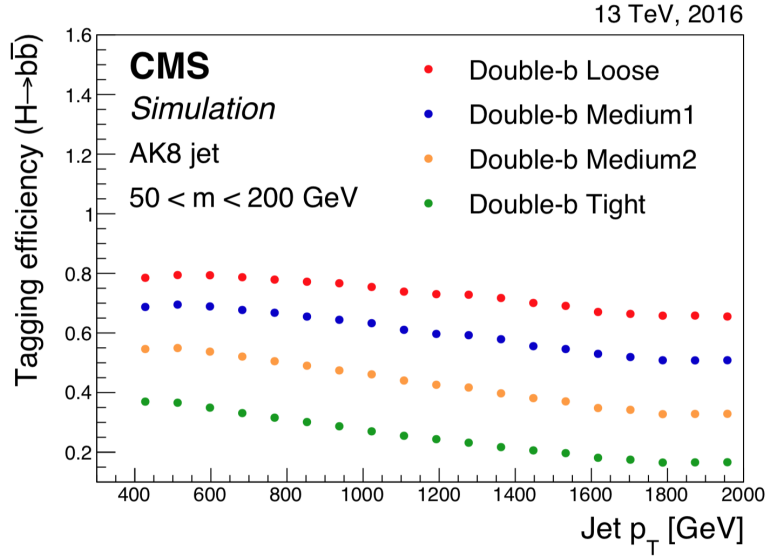
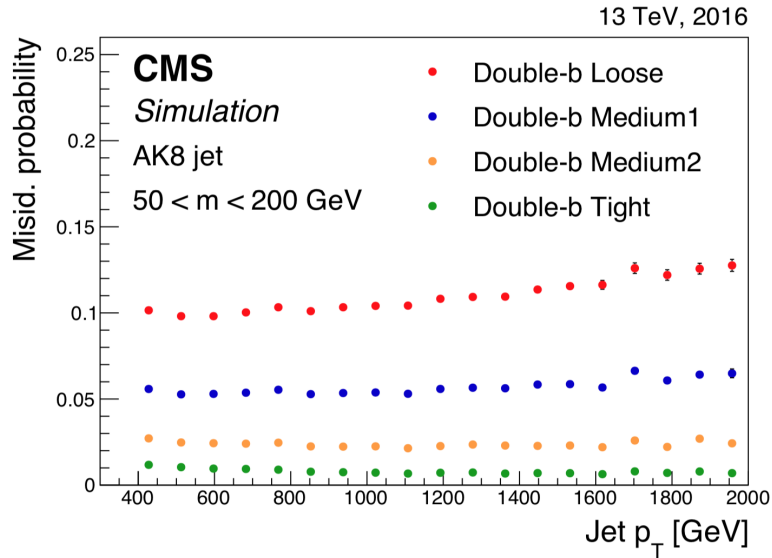


Figure 6.11: Distribution of the double-b tagger discriminator for  $H \rightarrow b\bar{b}$  jets and for jets in an inclusive multijet sample containing zero, one, or two b quarks.

The same selection requirements applied to tracks as in the CSVv2 algorithm are also applied here. The only difference is that instead of the jet axis, the  $\tau$  axis is used. The reconstructed secondary vertices are associated first with jets in a cone  $\Delta R < 0.7$  and then to the closest  $\tau$  axis within that jet. For each  $\tau$  axis, the track four-momenta of the constituent tracks from all the secondary vertices associated with a given  $\tau$  axis are added to compute the secondary vertex mass and  $p_T$  for that  $\tau$  axis.



(a) Tagging efficiency



(b) Tagging misidentification

Figure 6.12: Efficiency to correctly tag  $H \rightarrow b\bar{b}$  jets and misidentification probability using jets in an inclusive multijet sample for four working points of the double-b tagger as a function of the jet  $p_T$ . The AK08 jets are selected with  $p_T > 300$  GeV and pruned jet mass between 50 and 200 GeV.

The most discriminating variables are the impact parameter significance for the most displaced tracks, the 2D impact parameter significance for the first track above the (5.2 GeV) b-hadron mass threshold, and the secondary vertex energy ratio for the secondary vertex with the smallest 3D flight distance uncertainty ( $SV_0$ ). The secondary vertex energy ratio is defined as the total energy of all secondary vertices associated with a given  $\tau$  axis divided by the total energy of all the tracks associated with the AK08 jet that are consistent with the primary vertex, for each of the two  $\tau$  axes. The variables are combined using a BDT multivariate technique. The distribution of the double-b discriminator values is shown in Figure 6.11. Four working points are defined corresponding to about 75%, 65%, 45% and 25% efficiency for a jet  $p_T$  of around 1 TeV. The efficiencies and misidentification probabilities as functions of the jet  $p_T$  for these four working points are shown in Figure 6.12. The decreasing efficiency at high jet  $p_T$  is due to the stronger collimation of particles, which results in a lower track reconstruction efficiency and therefore in a lower tagging efficiency for high jet  $p_T$ .

## 6.7 Missing transverse energy

Only weakly interacting neutral particles, e.g. neutrinos, do not interact with the detector material and therefore leave no trace. The direction and the energy of these particles can be measured only by an indirect approach. The incoming protons which participate in a collision have only longitudinal component of momentum (along the beam axis). Due to momentum conservation, the transverse components of all particles produced in a collision should add up to zero. The missing transverse momentum is the momentum imbalance in the transverse plane of all detected particles in the event defined as

$$\vec{p}_T^{\text{miss}} = - \sum_{\text{PF obj}} \vec{p}_T^{\text{PF obj}} \quad (6.2)$$

where the sum extends over all PF objects. Its magnitude is the missing transverse energy  $E_T^{\text{miss}}$ . A non vanishing value of missing energy is a potential signature of the presence of particles in the event that have not interacted with the detector. However, it can be also caused by detector inefficiencies or resolution.

The use of missing transverse energy is important for this analysis in the reconstruction

of the  $W \rightarrow l\nu$  decays. The estimation of the  $E_T^{miss}$  in simulated events is improved by correcting it for the difference between raw (i.e. uncorrected) and calibrated jets, including the scale and resolution corrections, with  $p_T > 15$  GeV,  $|\eta| < 4.7$ , and passing a set of filters meant to remove electron and muon candidates. Events are rejected if they fail to pass set of recommended filters designed to suppress events affected by know issues such as instrumental noise.

## 6.8 Vector boson reconstruction

The reconstruction of W boson candidates begins with the identification and selection of charged leptons and MET described in the previous sections. MET selection is implicitly required in the  $p_T(V)$  selection criteria. Given the unique signature of a boosted vector boson recoiling from two jets, the dominant background is from real W decays. Therefore, a minimal selection is sufficient to identify highly pure samples of W+jets events. Decays  $W \rightarrow l\nu$  are identified primarily by the topology of a single isolated lepton and additional missing transverse energy. The transverse momentum  $p_T$  of the W candidate is computed as

$$p_T(W) = \sqrt{(E_{T,x}^{miss} + p_{l,x})^2 + (E_{T,y}^{miss} + p_{l,y})^2}$$

It is observed that in the boosted regime, where the QCD background is much reduced, simply requiring  $p_T(W) > 100$  GeV is sufficient to select a relatively clean sample of real W decays.

## 6.9 The Higgs boson reconstruction

The Higgs boson candidate (HC) definition was optimized studying simulated samples. In each event, the AK08 jet with the highest bb-tag value ( $|\eta| < 2.0$ ) is considered as HC without any additional requirements. Distribution of HC bb-tag value is shown in Figure 6.13 for the signal sample. For the purpose of evaluating the performance of the

HC definition, HC is split into three categories based on the number of true b-hadrons contained within jet: 0, 1, and 2. The tagger and therefore this HC definition, has good performance in discriminating 2b and 0b case, while additional cuts are required for differentiating 2b and 1b cases. They are discussed in the next chapter.

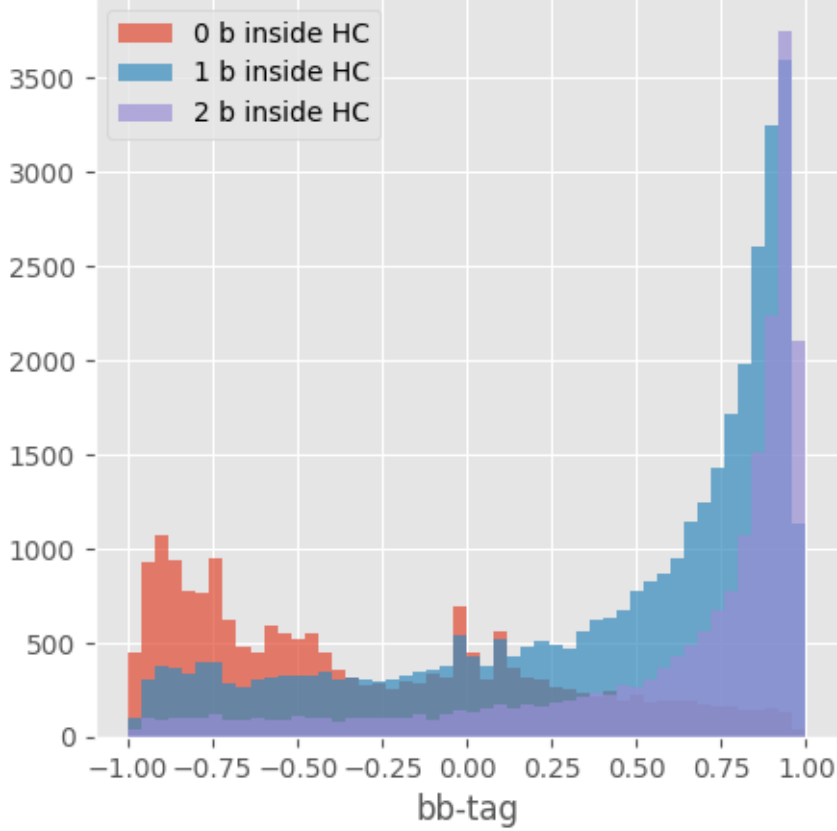


Figure 6.13: The Higgs boson candidate bb-tag value. Depending on the number of b quarks within HC, three categories are defined: 0b (red), 1b (blue), and 2b (purple).

According to the BDRS paper [1, 2], the boosted HC should have  $p_T > 200$  GeV. In this work the threshold for the HC  $p_T$  is set to 250 GeV. This value is chosen to suppress cases where the 2 b-hadrons are not contained in AK08 jet defined as HC, as can be seen in Figure 6.14. After selection on  $p_T$ , there is still a substantial fraction of HC candidates which contain only one b-hadron. They are suppressed efficiently with requirements on the value of the bb-tag and subjetiness. The Higgs boson candidate mass distribution is given in Figure 6.15.

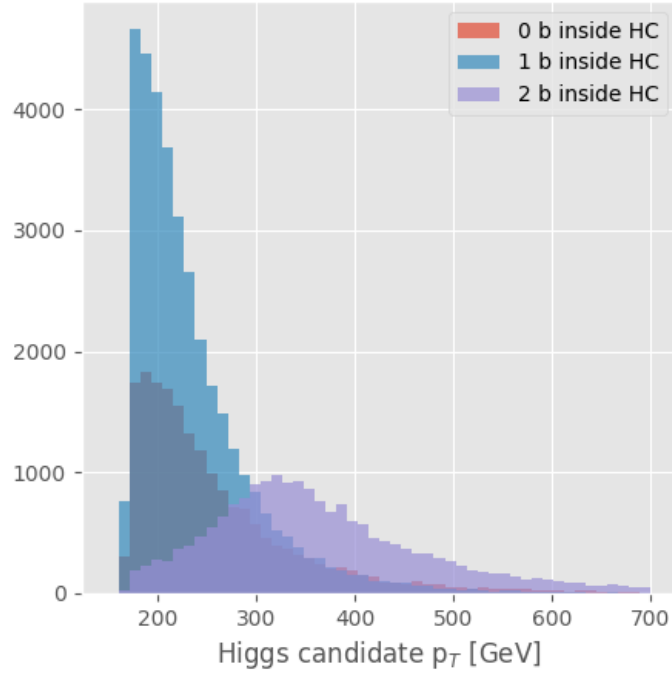


Figure 6.14: The Higgs boson candidate  $p_T$  distribution. Boosted topology starts at 250 GeV, when two b quarks are successfully caught within AK08 jet.

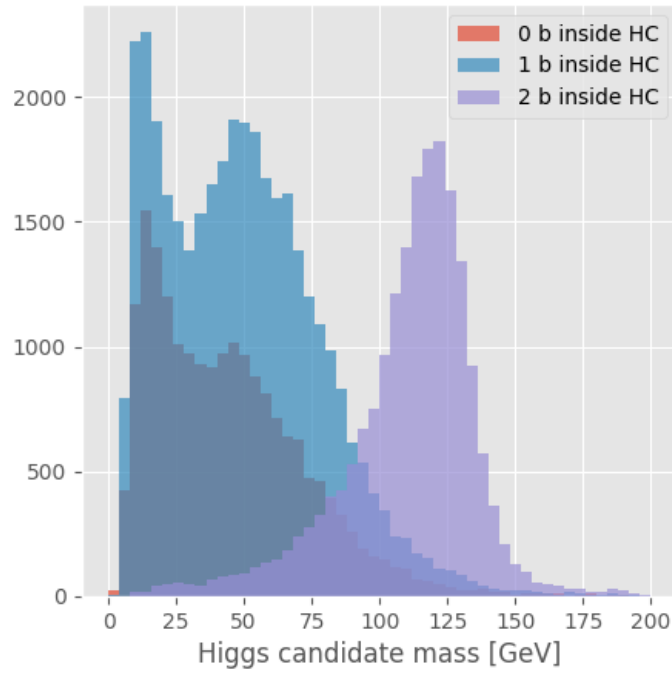


Figure 6.15: The Higgs boson candidate mass distribution.



# Chapter 7

## Search for boosted $W(l\nu)H(b\bar{b})$ production

This analysis is focused on a search for  $pp \rightarrow W(l\nu)H(bb)$  in boosted topology. At the beginning of this chapter the simulated samples used for signal and background processes, and the triggers used to collect the data are described. The extensive part of this chapter is dedicated to the event selection, the estimation of the background, and related systematic uncertainties. In the end, the fit procedure for the signal extraction is explained and the results are given.

### 7.1 Data sets and triggers

This measurement is based on data taken by CMS experiment in proton proton collisions at a center-of-mass energy  $\sqrt{s} = 13$  TeV in 2016, corresponding to an integrated luminosity  $L = 35.9 \text{ fb}^{-1}$ . This luminosity is smaller than the total delivered luminosity to CMS by the LHC for two reasons. On one side, a small fraction of the delivered luminosity was not recorded by CMS due to a small data taking inefficiency, as shown in Figure 7.1. On the other side, due to various issues with the detector, some of the recorded data have not passed some quality requirements which make them suitable for analysis. The overall uncertainty on the integrated luminosity is estimated to be 2.5% [124].



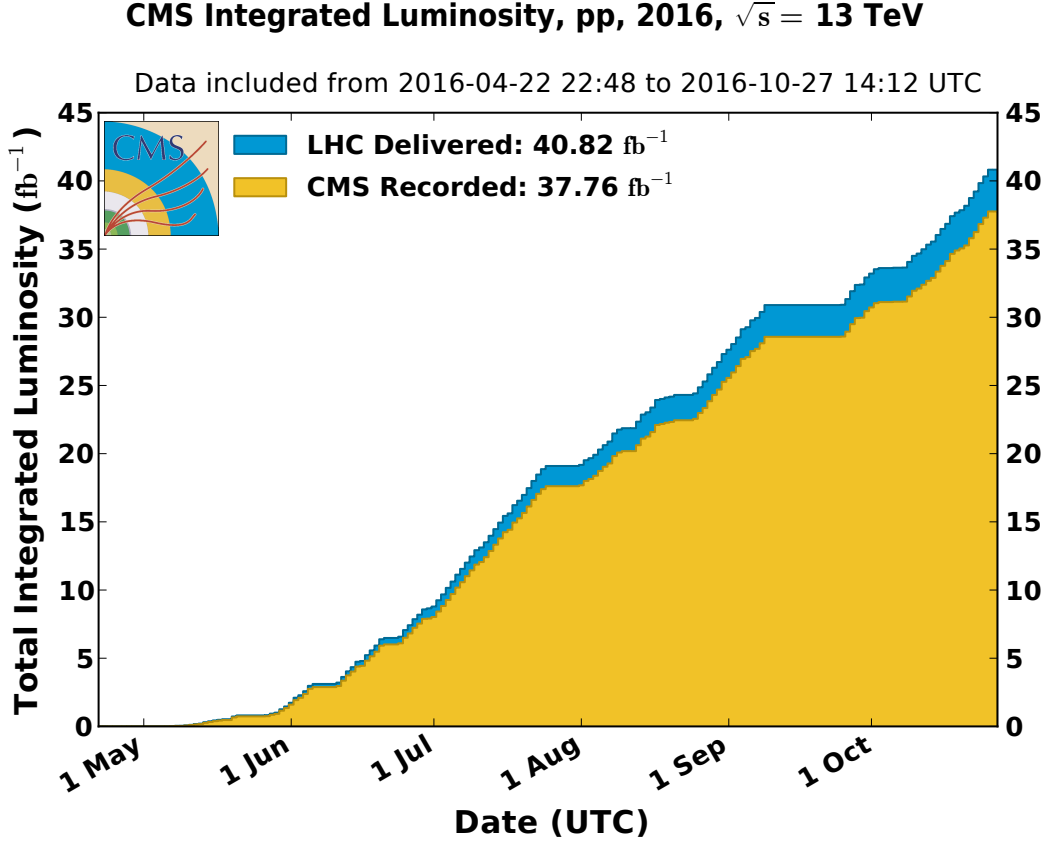


Figure 7.1: Cumulative measured luminosity versus day delivered for the 2016 proton-proton (pp) collisions at 13 TeV.

Events used in this measurement are required to pass single-electron or single-muon trigger. A brief overview of the HLT  $p_T$  criteria on the leptons is given in Table 7.1.

Table 7.1: Transverse momentum thresholds applied in the lepton triggers at the HLT level.

Trigger path	Threshold
Single Electron	$p_T > 27$ GeV
Single Muon	$p_T > 24$ GeV

## 7.2 Monte Carlo samples

Samples of simulated signal and background events are produced using the Monte Carlo (MC) event generators listed in Table 7.2. The CMS detector response is modeled with

Geant4 [104]. The signal samples have Higgs bosons with  $m_H = 125$  GeV produced in association with a W boson. The WH processes are generated at next-to-leading order (NLO) using the POWHEGv2 event generator [125,126]. The MADGRAPH5\_aMC@NLO [127] generator is used at NLO with the FxFx merging scheme [128] for the W+jets background samples. The  $t\bar{t}$  production process as well as the single top quark sample for the tW are produced with POWHEG v2. The single top quark samples for the t-channel are instead produced with POWHEG v1. For parton showering and hadronization the POWHEG and MadGraph5\_aMC@nlo samples are interfaced with PYTHIA 8.212 [97]. The PYTHIA8 parameters for the underlying event description correspond to the CUETP8M1 tune derived in [129].

Table 7.2: Summary of the samples of simulated processes.

Process	Event generator configuration	$XS \times BR[\text{pb}]$
$W(l^+\nu)H(bb)$	POWHEG+PYTHIA8	$2.52 * 0.108 \times 0.5824$
$W(l^-\nu)H(bb)$	POWHEG+PYTHIA8	$1.659 * 0.108 \times 0.5824$
$t\bar{t}$	POWHEG+PYTHIA8	831.76
Single top tW (t)	POWHEG+PYTHIA8	35.6
Single top tW ( $\bar{t}$ )	POWHEG+PYTHIA8	35.6
Single top t-channel (t)	POWHEG+PYTHIA8	$136.02*0.325$
Single top t-channel ( $\bar{t}$ )	POWHEG+PYTHIA8	$80.95*0.325$
Wjets ( $100 < W_{pT} < 250$ )	MadGraph5_aMC@nlo+PYTHIA8	628.3
Wjets ( $250 < W_{pT} < 400$ )	MadGraph5_aMC@nlo+PYTHIA8	22.37
Wjets ( $400 < W_{pT} < 600$ )	MadGraph5_aMC@nlo+PYTHIA8	2.67
Wjets ( $600 < W_{pT}$ )	MadGraph5_aMC@nlo+PYTHIA8	0.408
WZ	PYTHIA8	47.13
WW	PYTHIA8	118.7
ZZ	PYTHIA8	16.523

### EWK signal WH corrections

The production cross sections for the signal samples are rescaled to next-to-next-to-leading order (NNLO) and they are applied as a function of the vector boson transverse momentum. Figure 7.2 shows the shape of these corrections. The total cross section  $\sigma_{WH}$  is given by [130].

$$\sigma_{WH} = \sigma_{NNLOQCD}^{WH,DY}(1 + \delta_{EW}) + \sigma_{t-loop} + \sigma_{\gamma} \quad (7.1)$$

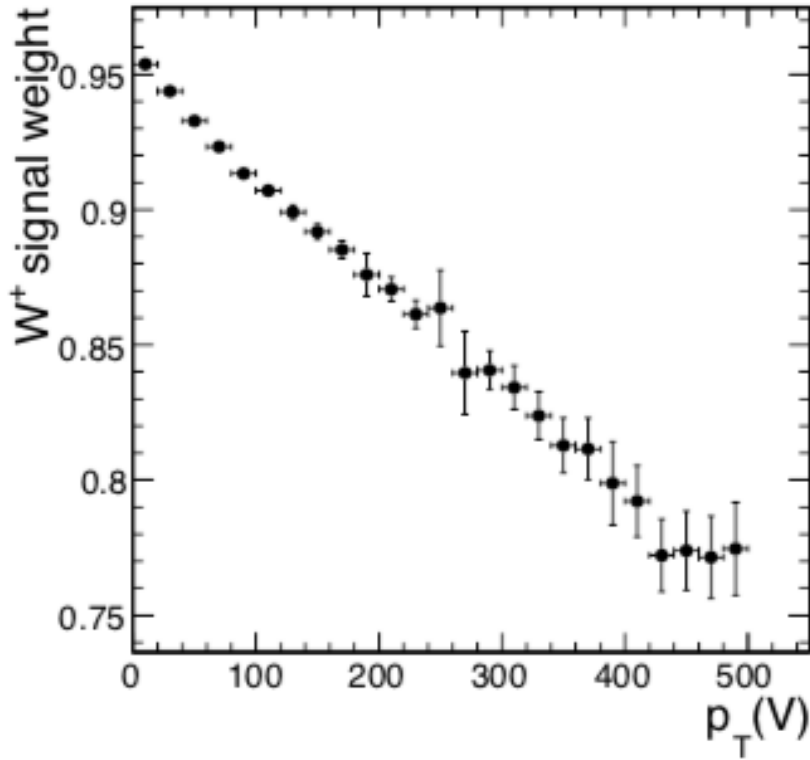


Figure 7.2: Multiplicative weights to apply the differential NLO electroweak signal correction.

## 7.3 Analysis Strategy

Simulated samples of signal and background events are used to optimize the search. The first step in the analysis is to select a signal region enriched in WH events. The next step is to define several control regions, each enriched in events from individual background processes. The purpose of control regions is to test the accuracy of the simulated samples and to extract the normalization of backgrounds. Finally, simultaneous likelihood fit to

the shape and normalization of specific distributions for the signal and control regions is used to extract the strength of a possible Higgs boson signal. In the signal region, the HC mass distribution is fitted, while the HC transverse momentum distribution is used in the control region.

### 7.3.1 Event reconstruction and selection

Signal events feature presence of a vector boson recoiling from the HC with an expected invariant mass in the range  $90 < M_H < 150$  GeV and  $p_T > 250$  GeV. HC should contain both b quarks coming from the Higgs decay. Therefore, no additional b jet activity outside the HC should be present. Isolated leptons other than those arising from the decay of a W boson are not expected and additional jet activity is reduced. The measurement is limited to the phase space in which both the W and Higgs bosons are central. In the absence of significant additional (ISR or FSR) jet activity, the W and Higgs bosons are expected to be back to back in the transverse plane. The dominant backgrounds arise from three general sources. It is important, to mention that QCD background is not among them because its contribution is negligible in this part of phase space.

- **W + jets** - production of W boson in association with one or more jets. This background looks very much like signal topologically, as shown in Figure 7.3, but has a generally softer  $p_T$  spectrum, a sharply falling HC mass distribution. The contributions from light jets should be much reduced after the application of selection on HC related variables. Contribution of W + bb nearly degenerate with signal.
- **Top quarks** - production of pair of top quarks, as well as single top quarks represent a particularly challenging background. These backgrounds include one or two real W decays, at least two b jets (for  $t\bar{t}$  production). The main way of reducing the  $t\bar{t}$  background relies on topological differences: the jet multiplicity extends well beyond two (typically from the hadronic decay of the second W), the azimuthal opening angle between the vector boson and HC is more broadly distributed than in signal events. Single top events are more difficult to reject relative to signal, but the cross section is such that it typically represents only 10-20% of the total background in WH. Diagram of  $t\bar{t}$  process is illustrated in Figure 7.3.

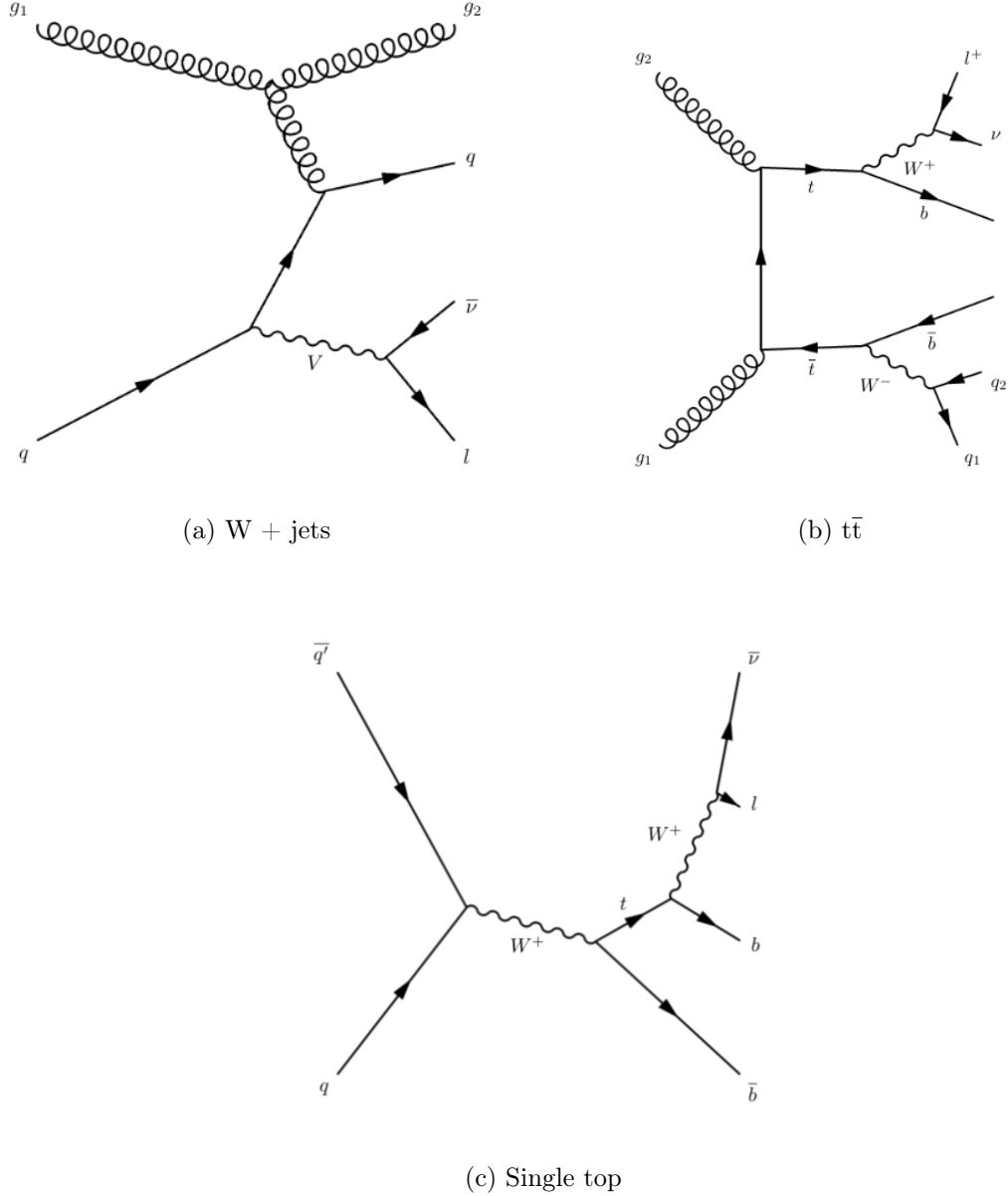


Figure 7.3: Feynman diagrams showing dominant background processes.

- **Dibosons (WW, WZ, ZZ)** - production of vector-boson pairs is another important background. The dominant contribution arises when one boson decays leptonically and the second boson decays to jets. The WZ can lead directly to a  $V + b\bar{b}$  combination with two real  $b$  jets coming from the hadronic  $Z \rightarrow b\bar{b}$  decay. This background is almost irreducible with respect to signal events. The peak position of the HC mass is shifted towards the  $Z$  boson mass. Good mass resolution is the key handle to separate the signal from this background.

Each background process has its own unique footprint which can be exploited while cre-

ating dedicated variables for differentiating signal and background events. The full list of variables used for selection is described below. For the evaluation of each variable, so-called "N-1 plots" have been used. In such a plot when showing one of the used variables, all selection requirements are applied but the one on the shown variable.

- **bb-tag** - The presence of two b quarks within the HC is the most important signal characteristic and can be enhanced by a tight selection on the bb-tag value, which is required to be  $> 0.8$ . This requirement mostly suppresses contributions from  $W + \text{light jets}$  and  $t\bar{t}$ .

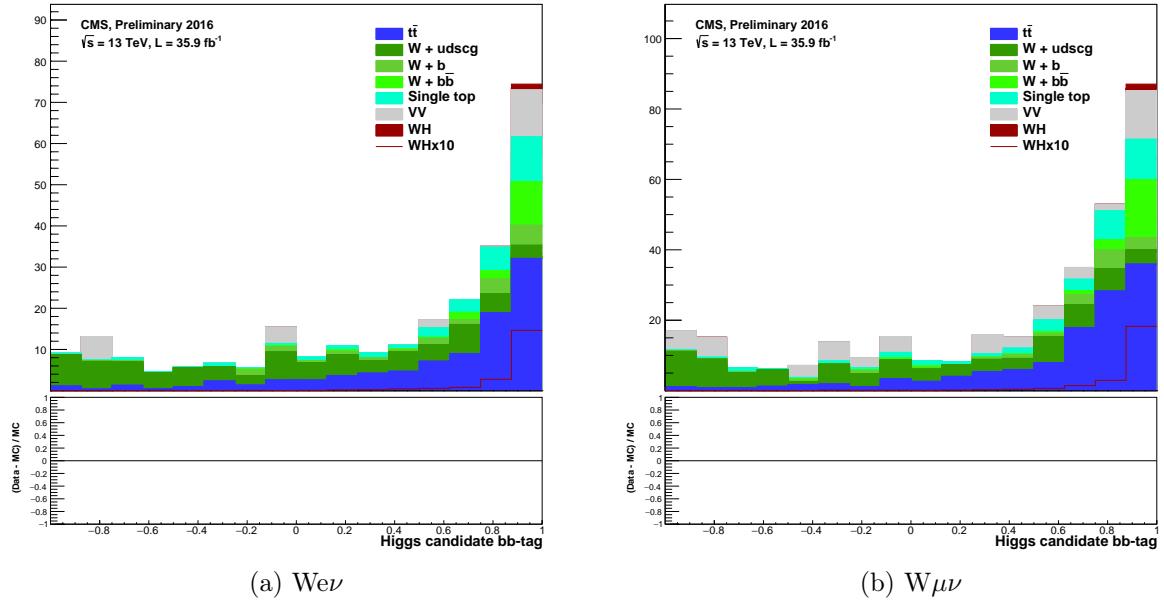


Figure 7.4: N-1 plot of HC double b-tag distribution. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown along with a yield 10 times as large as its SM prediction.

- $\tau_2/\tau_1$  - b quarks inside of HC should on average have equal momentum, resulting in two symmetric subjets for which it is easier to identify two distinct  $\tau$  axes. This doesn't hold for the background. For instance, there is no physical reason why two b quarks coming from  $t\bar{t}$  decay should carry the same momentum. Lower values of  $\tau_2/\tau_1$  are more likely for HC with two real subjets. In order to pass the selection, HC  $\tau_2/\tau_1$  must be lower than 0.45.

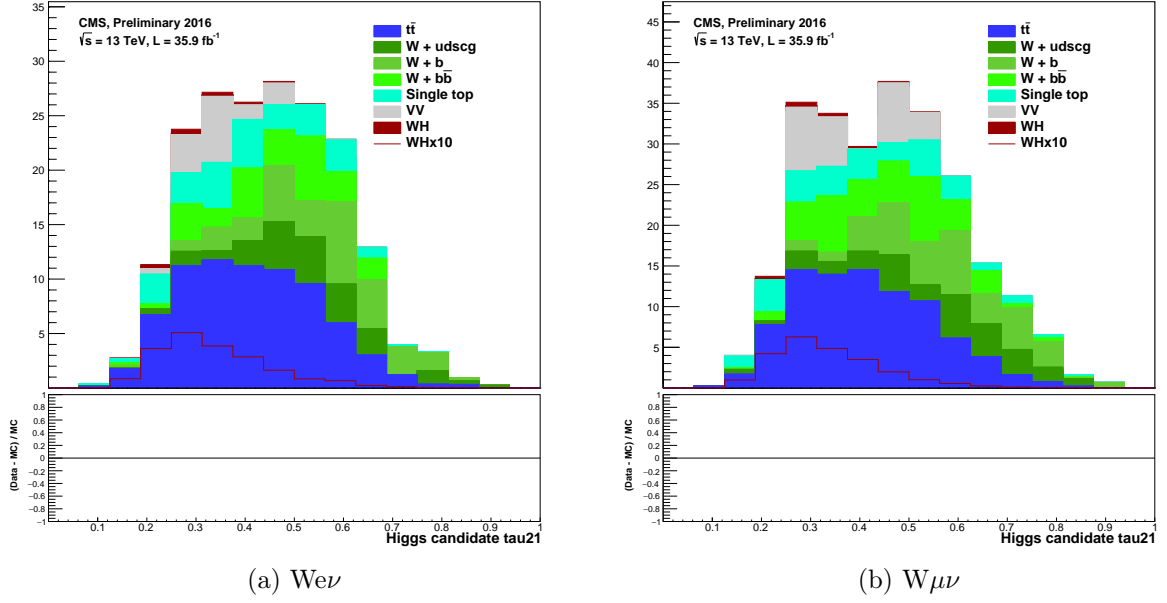


Figure 7.5: N-1 plot of HC  $\tau_2/\tau_1$  distribution. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction.

- $V \text{ p}_T / \text{HC p}_T$  - In the case of signal, the W boson and HC are expected to be balanced in the transverse plane. Therefore, the ratio of their momenta is required to be in the range,  $0.8 < V \text{ p}_T / \text{HC p}_T < 1.2$ , as can be seen in Figure 7.6.

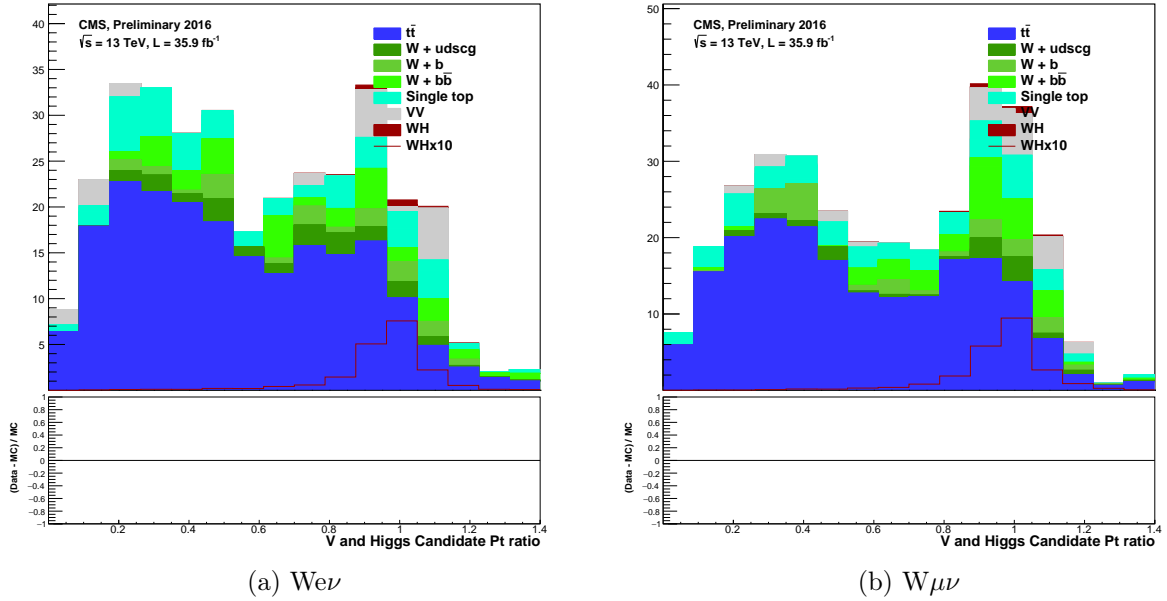


Figure 7.6: N-1 plot of  $V \text{ p}_T / \text{HC p}_T$  distribution. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction.

- **The number of additional jets** - Additional jet multiplicity is characteristic for  $t\bar{t}$  background. Events are rejected if they contain more than one additional jet. Only AK04 jets fulfilling the requirements listed in Table 7.3 are considered.

Table 7.3: The number of additional jets definition.

Observable	Selection
$\Delta R(\text{jet}, \text{HC})$	$> 0.8$
Jet $p_T$	$> 20 \text{ GeV}$
Jet $\eta$	$< 2.4$
Jet ID	True

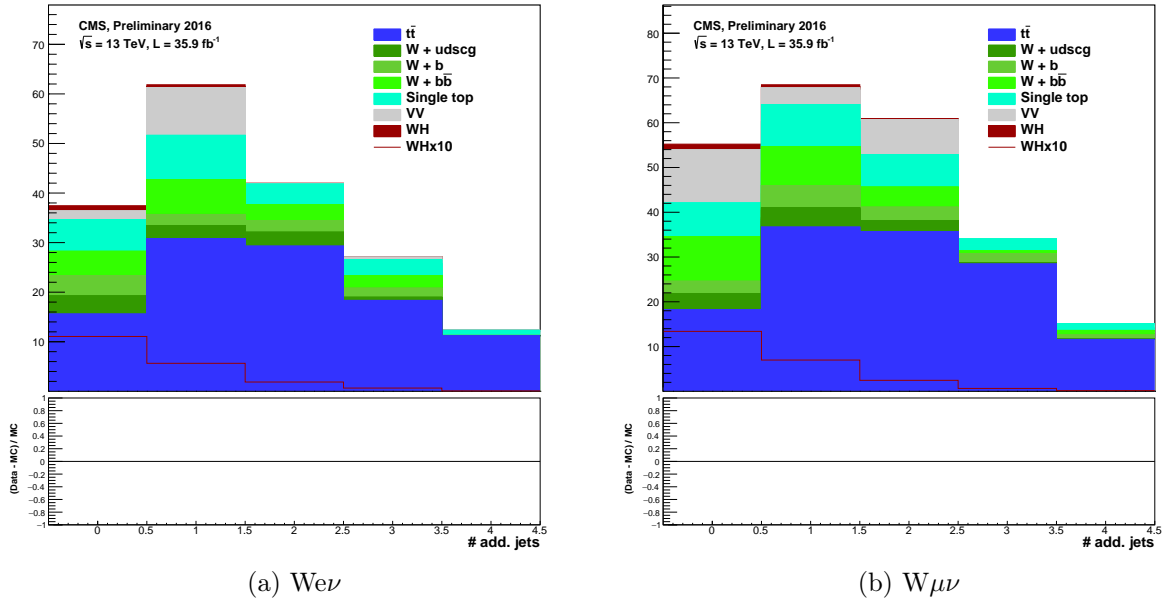


Figure 7.7: N-1 plot of number of additional jets distribution. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown along with a yield 10 times as large as its SM prediction.

- **b-tag veto** - No additional b-jet activity outside of the HC is expected for the signal, while this is not the case for background and in particular for  $t\bar{t}$ . Events are rejected if there is any AK04 jet passing the selection given in Table 7.4.



Table 7.4: b-tag veto definition.

Observable	Selection
$\Delta R(\text{jet}, \text{HC})$	$> 0.8$
Jet $p_T$	$> 20 \text{ GeV}$
Jet $\eta$	$< 2.4$
Jet b-tag	$> 0.5426$
Jet ID	True

b-tag veto provides a significant reduction of top background with almost no reduction of signal, as can be seen in Figure 7.8. In left(right) bin are the events that fail(pass) the b-tag veto.

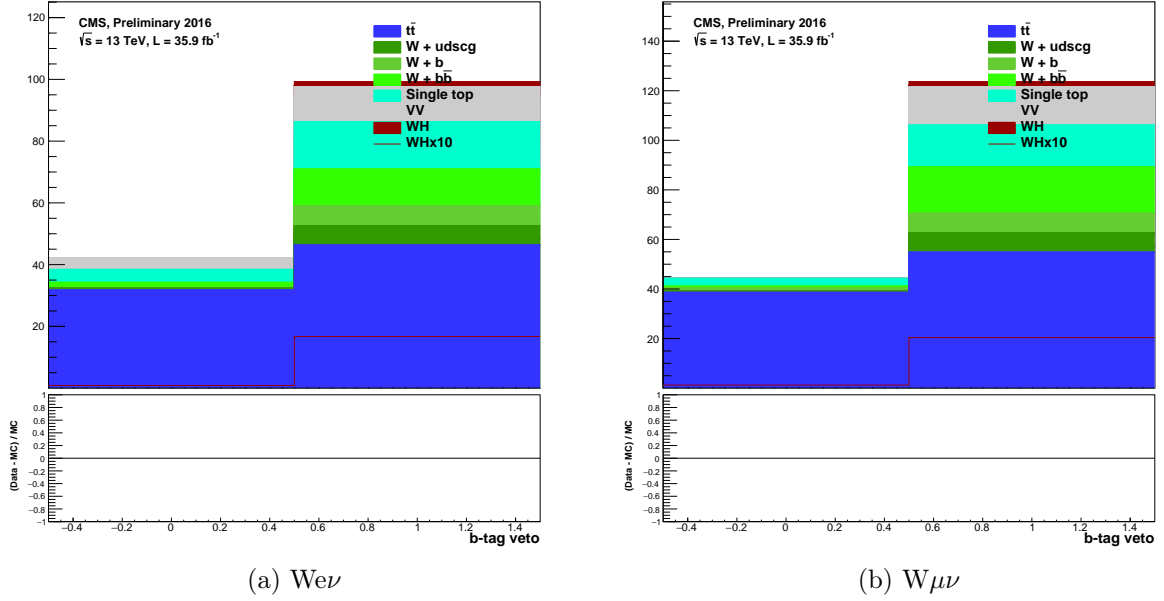


Figure 7.8: N-1 plot of b-tag veto distribution. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction.

- **c-tag veto** - The idea of the c-tag veto is to suppress events with c-jet activity inside of HC. Event is rejected if there is any AK04 jet in the HC passing selection listed in Table 7.5.

Table 7.5: c-tag veto definition. Events with at least one AK04 jet fulfilling these requirements are rejected.

Observable	Selection
$\Delta R(\text{jet}, \text{HC})$	$< 0.8$
Jet $p_T$	$> 20 \text{ GeV}$
Jet $\eta$	$< 2.4$
Jet c-tagVsB	$> -0.17$
Jet c-tagVsL	$> -0.48$
Jet ID	True

c-tag veto reduces background dramatically but this comes with a cost of reducing signal by 50%. Figure 7.9 shows c-tag veto distribution. Left(right) bin are the events that fail(pass) the c-tag veto. This particular combination of c-tagger cut values is chosen because they are officially validated by CMS and are at the same time very efficient in background reduction, as seen in Figure 7.10.

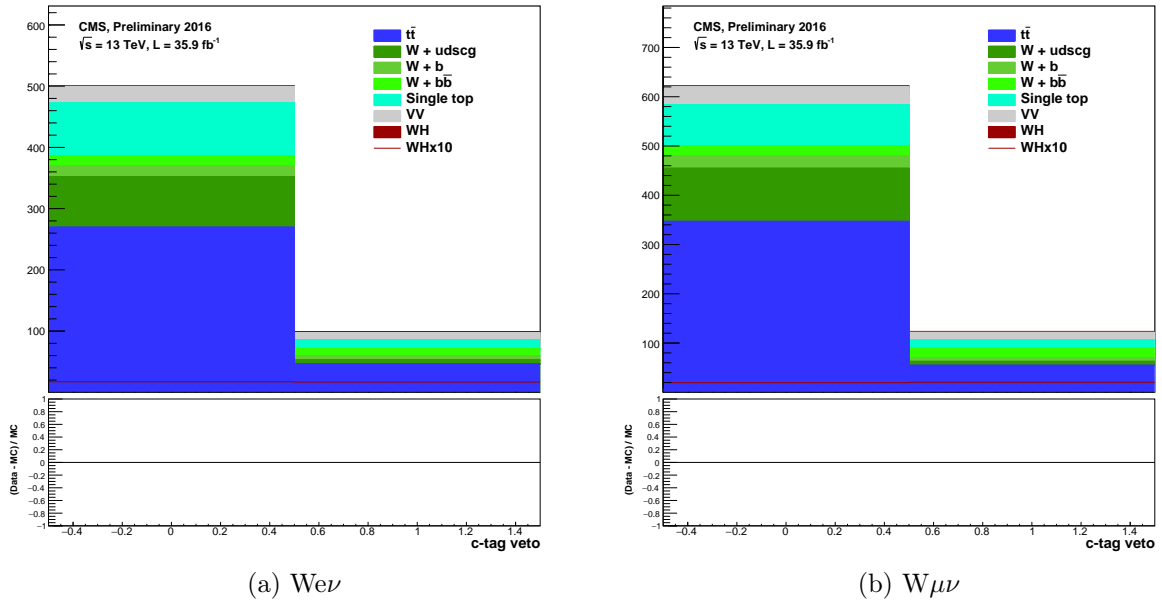


Figure 7.9: N-1 plot of c-tag veto distribution. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown along with a yield 10 times as large as its SM prediction.

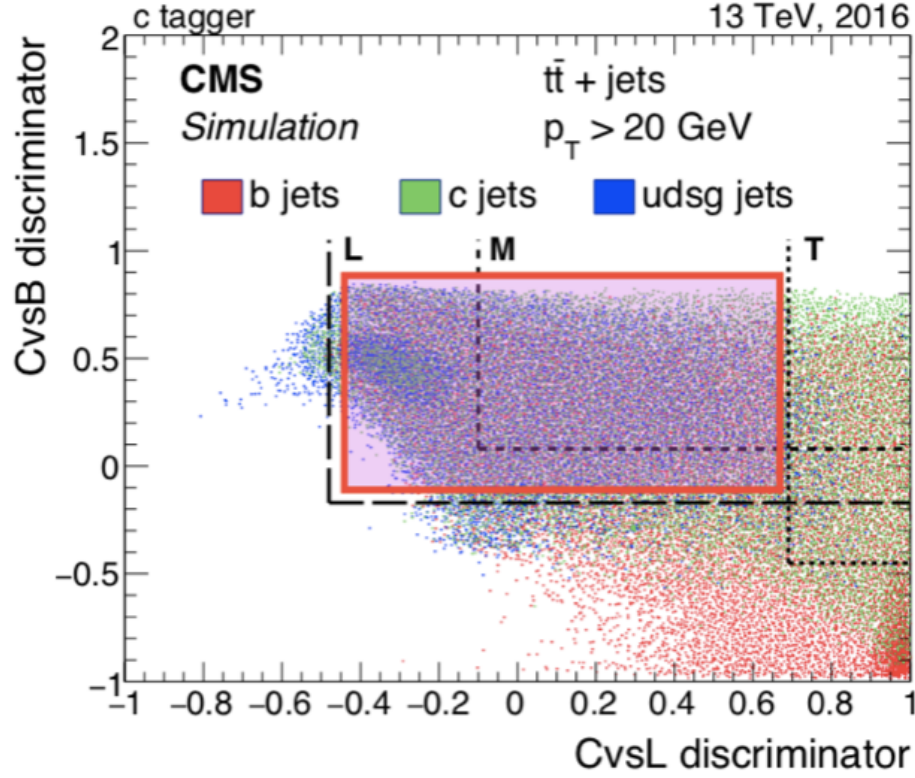


Figure 7.10: Correlation between CvsL and CvsB taggers for the various jet flavours. Combination of L and T working point has been used in this analysis.

The final selection which defines the signal region is given in Table 7.6.

Table 7.6: Selection criteria that define the signal region.

Observable	Selection
HC $p_T$	$> 250$
HC mass	$\in [50 - 190]$
HC $b\bar{b}$ tag	$> 0.8$
$V p_T / HC p_T$	$\in [0.8 - 1.2]$
HC $\tau_2/\tau_1$	$< 0.45$
N add. jets	$< 2$
b tag veto	True
c tag veto	True

The expected number of events in signal region after applying each requirement is given in Table 7.7.

Table 7.7: The expected number of events in signal region after applying each requirement (cut-flow). All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ .

Observable	Signal	$t\bar{t}$	W + udscg	W + b	W + bb	Single top	VV
HC bb tag	16.53	39133.18	1662.41	881.39	817.91	2837.75	406.17
V $p_T$ / HC $p_T$	13.05	26610.61	736.71	273.38	383.40	1778.04	243.17
HC $\tau_2/\tau_1$	9.60	4453.49	320.00	105.35	128.58	539.60	123.00
N add. jets	7.81	1193.36	224.69	68.58	73.61	252.66	101.76
b tag veto	7.39	717.91	203.81	57.82	67.73	201.60	89.64
c tag veto	3.71	101.25	14.13	14.20	30.61	32.22	26.88

## 7.4 Background and scale factor estimation

For the purpose of determining the normalization of the main background processes, and to validate how well the simulated samples model the distributions of variables most relevant to the analysis, several control regions are carefully chosen in data. They are supposed to be as close as possible to the signal region phase space while at the same time being completely orthogonal. Table 7.8 list the selection criteria used to define them. Separate control regions are specified for  $t\bar{t}$ , for the production of W boson in association with mainly heavy-flavor (HF) or light-flavor (LF) jets. While the  $t\bar{t}$  control region is pure in its targeted background process, other control regions contain slightly larger contribution of other processes. As a result, distinct HC  $p_T$  distributions are used to extract the normalization scale factors of the various simulated background samples when fit to data in combination with the HC mass distributions in the signal region to search for a possible WH signal. In this signal-extraction fit, the shape and normalization of these distributions are allowed to vary, for each background component, within the systematic and statistical uncertainties described in Section 7.6.

Table 7.8: Definition of signal and control regions. Flipped cuts are highlighted in red.

Observable	Signal	$t\bar{t}$	W + LF	W + HF
HC $p_T$	$> 250$	$> 250$	$> 250$	$> 250$
HC mass	$\in [50 - 190]$	$\in [50 - 190]$	$\in [50 - 190]$	$\in [50 - 190]$
HC $b\bar{b}$ tag	$> 0.8$	$> 0.8$	$< 0.8$	$> 0.8$
V $p_T$ / HC $p_T$	$\in [0.8 - 1.2]$	$\in [0.8 - 1.2]$	$\in [0.8 - 1.2]$	$\in [0.8 - 1.2]$
HC $\tau_2/\tau_1$	$< 0.45$	$< 0.45$	$> 0.45$	$> 0.45$
N add. jets	$< 2$	$> 1$	$< 2$	$< 2$
b tag veto	True	True	True	True
c tag veto	True	True	True	True

Pre-fit HC  $p_T$  distributions for all control regions are shown in Figures 7.11, 7.12, and 7.13.

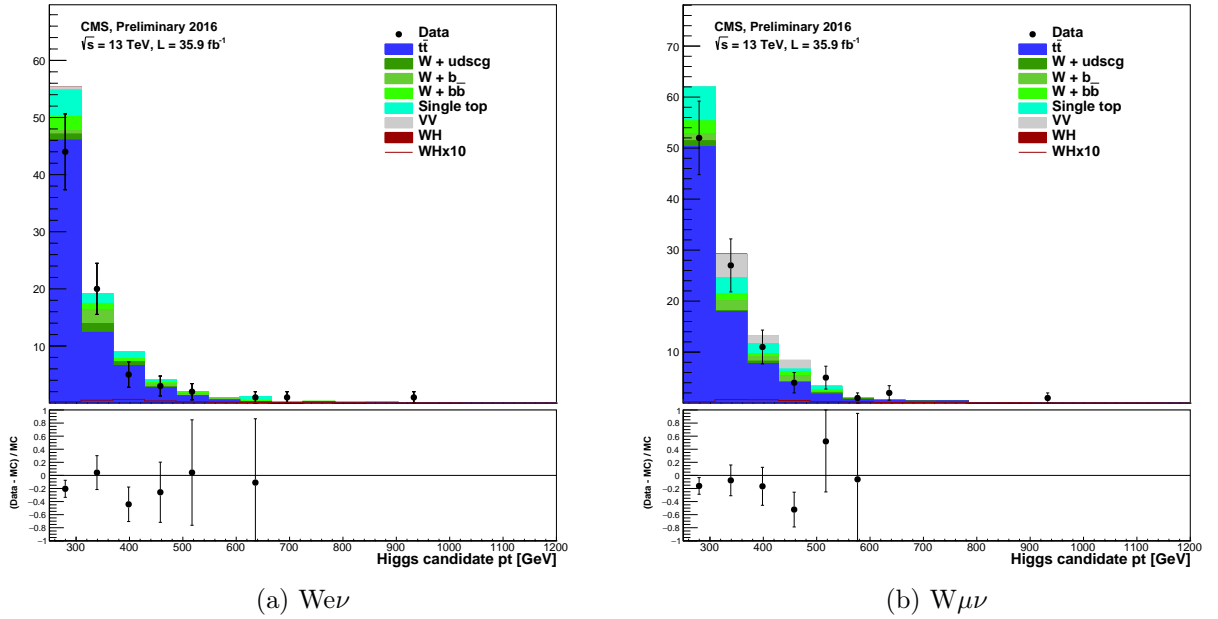


Figure 7.11: Pre-fit HC  $p_T$  distribution in  $t\bar{t}$  control region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown along with a yield 10 times as large as its SM prediction.

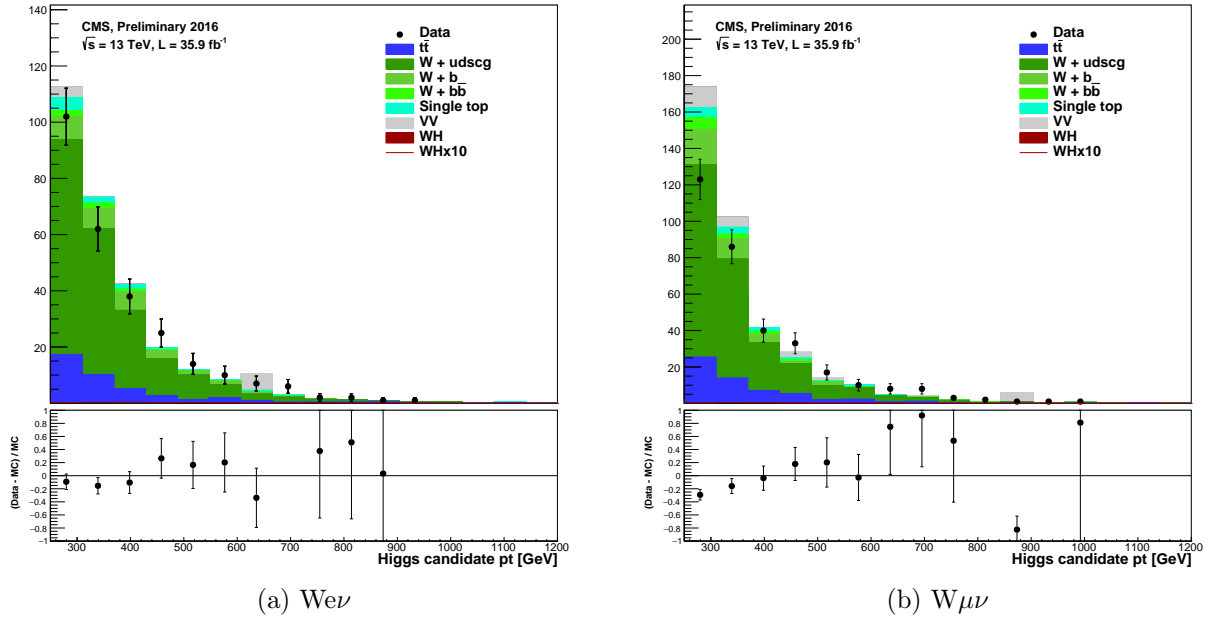


Figure 7.12: Pre-fit HC  $p_T$  distribution in light flavor (LH) control region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction.

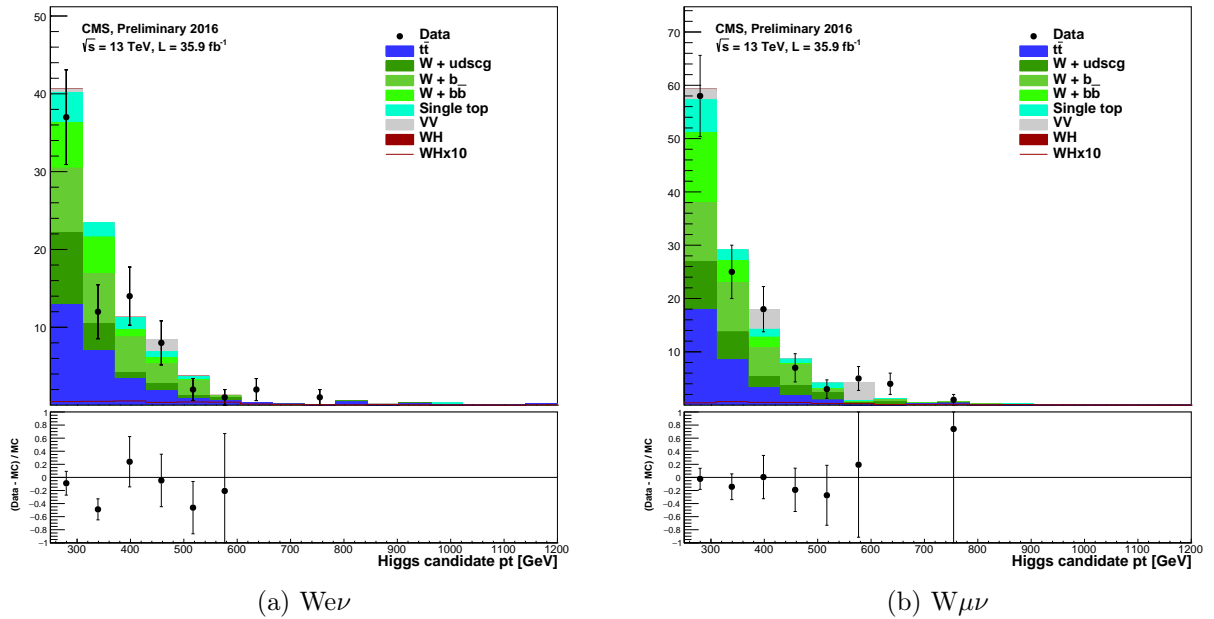


Figure 7.13: Pre-fit HC  $p_T$  distribution in heavy flavor (HF) control region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The signal is also shown alone with a yield 10 times as large as its SM prediction.

## 7.5 Signal extraction and limit setting

The procedure for deriving exclusion limits is based on the CL<sub>s</sub> prescription [131], which is used with the profile likelihood test statistic

$$\bar{q}_\mu = -2 \ln \frac{\mathcal{L}(\text{data} | \mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data} | \hat{\mu}, \hat{\theta})} \quad (7.2)$$

The parameter  $\mu = \sigma/\sigma_{SM}$  is the signal strength modifier, while  $\theta$  in general represents the full set of nuisance parameters. The maximum likelihood estimates (best-fit-values) of  $\mu$  and  $\theta$  are denoted by  $\hat{\mu}$  and  $\hat{\theta}$  respectively.  $\hat{\theta}_\mu$  is the conditional maximum likelihood estimate of all nuisance parameters with  $\mu$  fixed. In this analysis the range of  $\mu$  is restricted to the physically meaningful regime, i.e.  $0 \leq \hat{\mu} \leq \mu$ . The likelihood  $\mathcal{L}$  is given by the product of the individual bin likelihoods

$$\mathcal{L}(\text{data} | \mu, \theta) = \prod_i \text{Poisson}(N_i | \mu \cdot s_i(\theta) + b_i(\theta)) \cdot p(\tilde{\theta} | \theta) \quad (7.3)$$

where  $\text{Poisson}(N_i | \mu \cdot s_i(\theta) + b_i(\theta))$  stands for the Poisson probabilities to observe  $N_i$  events in the  $i$ -th bin given the expected event rate  $\mu \cdot s_i(\theta) + b_i(\theta)$ :

$$\text{Poisson}(N_i | \mu \cdot s_i(\theta) + b_i(\theta)) = \frac{(\mu s_i + b_i)^{N_i}}{N_i!} e^{-(\mu s_i + b_i)} \quad (7.4)$$

Information on the systematic uncertainties is contained in the probability density functions  $p(\tilde{\theta} | \theta)$  and treated in two different ways. If the systematic uncertainty does not change the shape of the fitted distribution, a log-normal probability density function is used:

$$\rho(\theta) = \frac{1}{\sqrt{2\pi} \ln(\kappa)} \exp\left(-\frac{\ln(\theta/\tilde{\theta})^2}{2(\ln(\kappa))^2}\right) \frac{1}{\theta} \quad (7.5)$$

where the parameter  $\kappa$  defines the width of the log-normal distribution and  $\tilde{\theta}$  is the default value of the nuisance parameter. If the systematic uncertainty changes the shape of the fitted distribution, different procedure is followed. Two additional distributions with different shapes for each process affected by some uncertainty are produced, by shifting that parameter up and down by one standard deviation. In other words, this uncertainty defines the variations of the nominal shape as a function of the value of the nuisance

parameter.

Using the frequentist approach, Monte-Carlo pseudo-experiments are generated to construct the corresponding pdfs. The  $\text{CL}_S$  value is calculated as the ratio of two probabilities

$$\text{CL}_S(\mu) = \frac{\text{P}(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \mu, \hat{\theta}_\mu^{\text{obs}})}{\text{P}(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | 0, \hat{\theta}_0^{\text{obs}})} \quad (7.6)$$

If for  $\mu = 1$ ,  $\text{CL}_S = 0.05$ , the SM Higgs boson with a nominal production rate is said to be excluded at 95% Confidence Level (C.L.). In this analysis, the result is expressed as 95% C.L. upper limits on  $\mu$ . To quote the 95% Confidence Level upper limit on  $\mu$ , to be further denoted as  $\mu^{95\% \text{CL}}$ , we adjust  $\mu$  until we reach  $\text{CL}_S = 0.05$ .

To express the sensitivity of an experiment to exclude certain hypothesis, expected limits are determined using the following procedure [132]. The expected limit is expressed as median limit and as  $\pm 1\sigma$  and  $\pm 2\sigma$  bands for the background-only hypothesis. These are obtained by generating a large set of background-only pseudo-data and calculate  $\text{CL}_S$  and  $\mu^{95\% \text{CL}}$  for each of them, as if they were real data. Then, one builds a cumulative probability distribution of results by starting integration from the side corresponding to low event yields. The point at which the cumulative distribution crosses the quantile of 50% is the median expected value. The  $\pm 1\sigma$  (68%) band is defined by the crossing of the 16% and 84% quantiles, while the crossings at 2.5% and 97.5% define the  $\pm 2\sigma$  (95%) band.

## 7.6 Systematic uncertainties

The primary physics result in this thesis is the upper limit on  $\mu$  for WH production, assuming a Higgs boson mass of 125 GeV. It is affected by systematic uncertainties on the expected signal and background yields and their shapes. The following experimental systematic sources have been taken into account.

- **Luminosity** - The uncertainty in the integrated luminosity collected during the 2016 data taking period is found to be 2.5%. No shape uncertainties are considered.
- **Lepton efficiency** - Lepton trigger, reconstruction, and identification efficiencies are evaluated from data using the standard tag-and-probe technique with Z bosons.



The systematic uncertainty is computed from the statistical uncertainties in the bin-by-bin efficiencies. The total uncertainty is a constant 3% per charged lepton. No shape uncertainties are considered.

- **Pile-up multiplicity** - The uncertainty on the average number of pile-up interactions measured in data is used to produce modified signal and background shapes, which are used in the final fit.
- **Monte Carlo statistics** - The finite size of the signal and background MC samples is included in the normalization uncertainties. The shapes are allowed to vary within the bin-by-bin statistical uncertainties from the MC samples.
- **Jet energy scale** - The jet energy scale uncertainty is estimated separately for each jet energy correction and added in quadrature to get the final uncertainty. It is varied within one standard deviation based on  $p_T$  and  $\eta$ . Systematic uncertainties are estimated by producing new shapes for all signal and background processes by varying the jet energy scale within these uncertainties.
- **Jet energy resolution** - The jet energy resolution in data is worse than in the simulation. In order to take into account the differences between data and MC, MC jet energies are smeared. The uncertainty on the smearing factors is used to produce modified signal and background shapes. These modified shapes are then used in the final fit both for AK04 and AK08 jets.
- **b-tagging** - Jet (AK04) b-tagging efficiencies are determined from control samples in data and MC in order to account for the difference in efficiencies between data and MC. Simulated events with selected jets are reweighted by the data/MC efficiency ratios, referred to as scale factors. These scale factors depend on  $p_T$ ,  $\eta$ , and flavour of the selected jets [121]. To estimate the effect of the uncertainties on the efficiency determination, each scale factor is shifted up and down by the corresponding uncertainty, and event weights are recalculated. The procedure is done separately for b and c jet efficiencies and for light flavour mistag rate.
- **c-tagging** - Jet (AK04) c-tagging efficiencies and corresponding systematic uncertainties have been computed following a similar procedure as in the case of b-tagging. Data/MC efficiency ratios are applied as weight factors for each event. To estimate

the effect of the uncertainties, each scale factor is shifted up and down and event weights are recalculated.

- **bb-tagging** - The bb-tag scale factors are determined for the selected AK08 jets and applied as weight factors with  $p_T$  dependence for each event. For estimation of the uncertainties, each event weight is shifted up and down following the procedure provided by CMS.
- $\tau_2/\tau_1$  - The systematic uncertainty on  $\tau_2/\tau_1$  takes into account uncertainties on the  $\tau_2/\tau_1$  selection efficiency and extrapolation uncertainties on the  $\tau_2/\tau_1$  selection due to propagation to higher momenta. The effects of the corresponding uncertainties are assessed by varying the corresponding event weights by one standard deviation and producing the modified signal and background shapes.

Theoretical uncertainties generally arise from missing higher-order QCD corrections and PDF uncertainties.

- **QCD scale uncertainties** - The uncertainties from normalization ( $\mu_R$ ) and factorization ( $\mu_F$ ) scales are estimated by varying both scales independently in the range  $(0.5\mu_0, 2\mu_0)$  around their nominal values  $\mu_0$  for all processes and taking them as uncorrelated sources of shape uncertainties.
- **PDF uncertainties** - The imperfect knowledge of the proton quark content is encoded in a set of NNPDF parton distribution functions [133]. The utilization of different PDF sets can affect the shape of the signal and background contributions. The PDF uncertainty effect on the distribution shapes is evaluated by using the PDF replicas associated to the NNPDF set. The effect on the signal strength is found to be at most 1%.
- **Theoretical  $p_T$  spectrum** - potential differences in the  $p_T$  spectrum of W and the Higgs boson between data and Monte Carlo generators could introduce systematic effects in the signal acceptance and efficiency estimates. Electroweak corrections have been applied to the signal MC samples. The estimated uncertainty from higher order EWK corrections is 2% for WH [134, 135].

## 7.7 Results

Expected and observed 95% C.L. upper limits on  $\mu$  (in the absence of signal) are calculated for  $35.9 \text{ fb}^{-1}$  of 13 TeV (2016) data and are shown in Figure 7.18. For the  $1\sigma$  and  $2\sigma$  bands, the  $\text{CL}_s$  frequentist calculation currently recommended by the LHC Higgs Combination Group (summarized in section 7.5) is employed. The post-fit distributions of control and signal regions are shown in Figures 7.14, 7.15, 7.16, 7.17. They consider the adjustments of all nuisance parameters in the final maximum likelihood fit to extract the signal strength. Yields before and after the fit are shown in Table 7.9.

Table 7.9: The total number of events in signal region for the expected prefit/postfit backgrounds (B), WH signal (S), and for data. Also shown is the signal-to-background ratio (S/B).

Process	Number of events	
	Pre-fit	Post-fit
W + bb	30.61	4.69
W + b	14.20	3.81
W + udscg	14.13	45.83
$t\bar{t}$	101.25	75.23
Single-top-quark	32.22	22.63
VV	26.88	22.18
Total backgrounds	219.29	174.37
WH	3.71	0.0
Data	170	
S/B	0.016	0.0

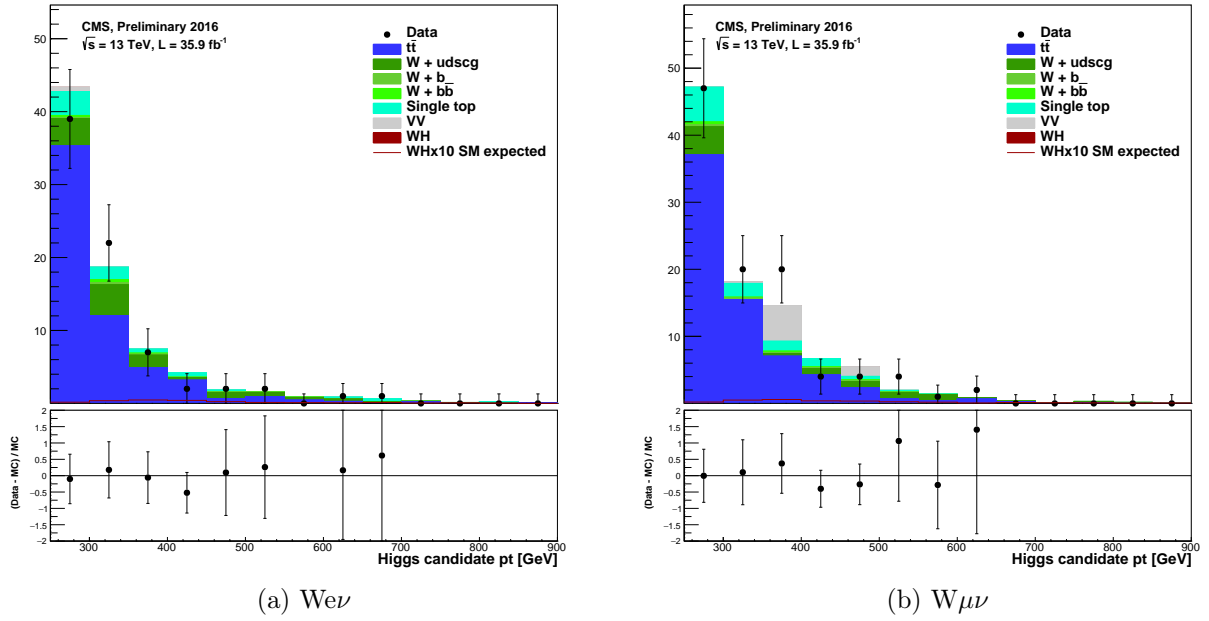


Figure 7.14: Post-fit HC  $p_T$  distribution in  $t\bar{t}$  control region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown along with a yield 10 times as large as its SM prediction.

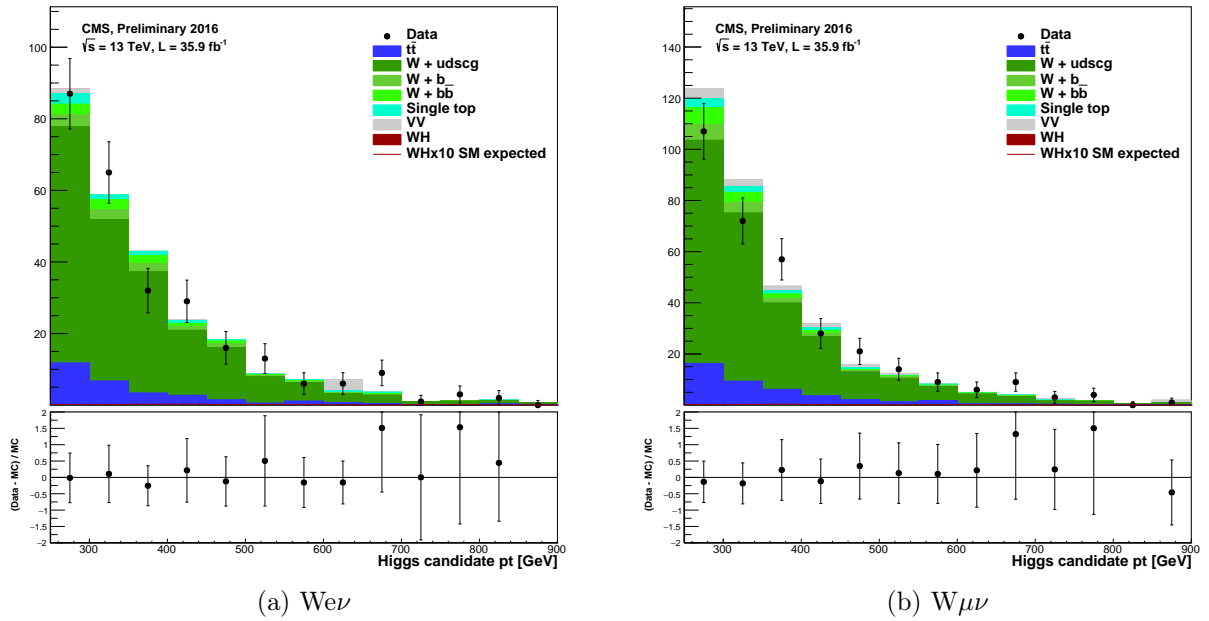


Figure 7.15: Post-fit HC  $p_T$  distribution in light flavor (LH) control region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown along with a yield 10 times as large as its SM prediction.

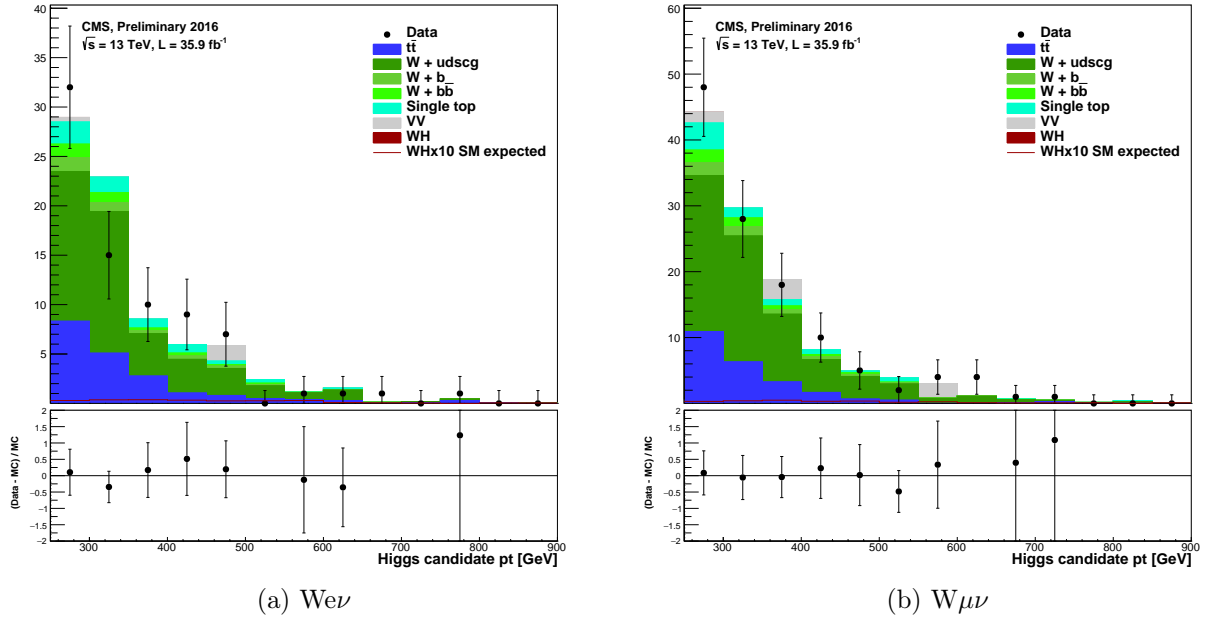


Figure 7.16: Post-fit HC  $p_T$  distribution in heavy flavor (HF) control region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown along with a yield 10 times as large as its SM prediction.

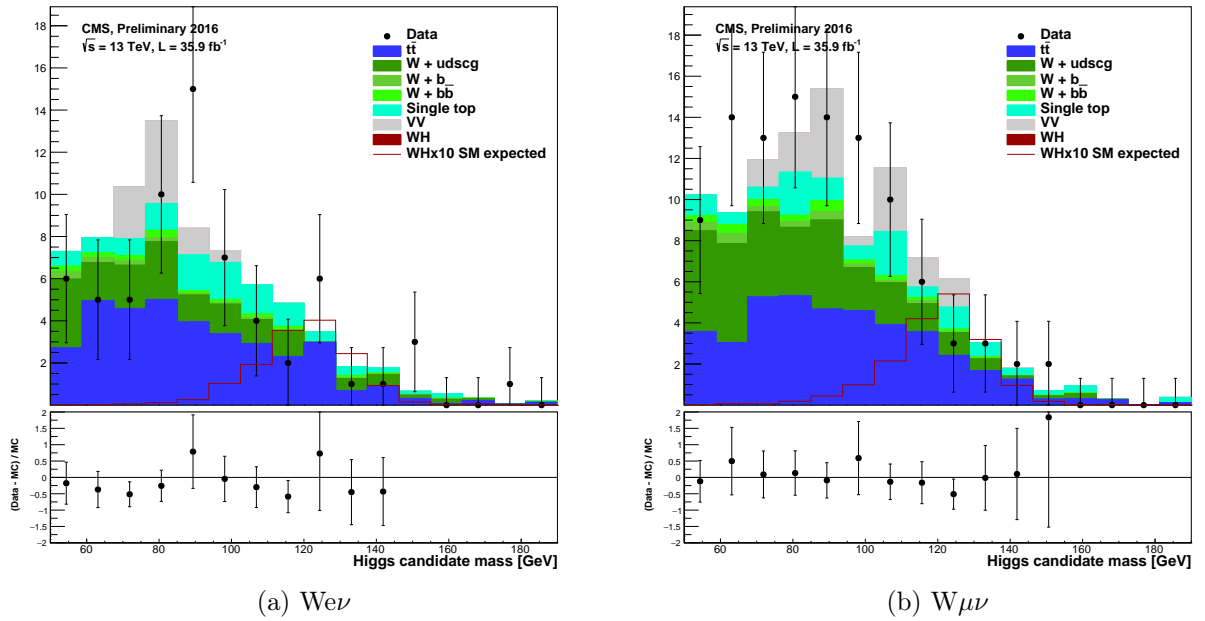


Figure 7.17: Post-fit HC mass distribution in signal region. All samples are normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The expected pre-fit signal is also shown along with a yield 10 times as large as its SM prediction.

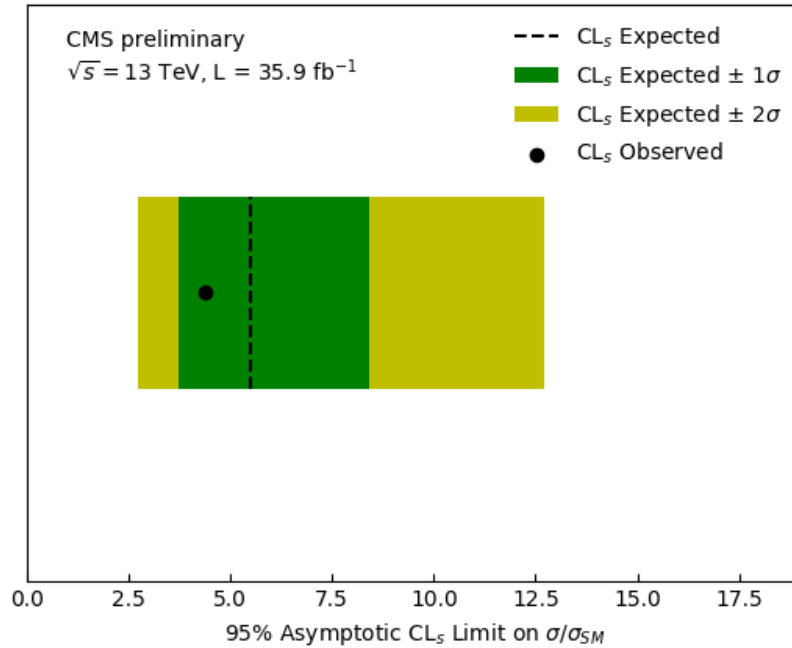


Figure 7.18: Expected and observed 95% C.L. upper limits (in the absence of signal) on the ratio  $\sigma/\sigma_{SM}$  of WHbb production for 13 TeV, 2016 data. The median expected limit, observed limit and the  $1\sigma$  and  $2\sigma$  bands are obtained with the full LHC CL<sub>S</sub> method.

## 7.8 Impacts of the systematic uncertainties

The impact of a nuisance parameter  $\theta$  on a signal strength  $\mu$  is defined as the shift  $\Delta\mu$  that is induced as  $\theta$  is fixed and brought to its  $+1\sigma$  or  $-1\sigma$  post-fit values, with all other parameters profiled as normal. This is effectively a measure of the correlation between the nuisance parameter and the signal strength, and is useful for determining which nuisance parameters have the largest impact on the signal strength uncertainty. The list of the parameters with the largest effect is shown in Figure 7.19. The left panel in the plot shows the value of  $(\theta - \theta_0)/\Delta_\theta$  where  $\theta$  and  $\theta_0$  are the post and pre-fit values of the nuisance parameter and  $\Delta_\theta$  is the post-fit uncertainty. The asymmetric error bars show the pre-fit uncertainty divided by the post-fit uncertainty meaning that parameters with error bars smaller than  $\pm 1$  are constrained in the fit. The signal strength uncertainty is dominated by MC statistics as shown in Figure 7.19.

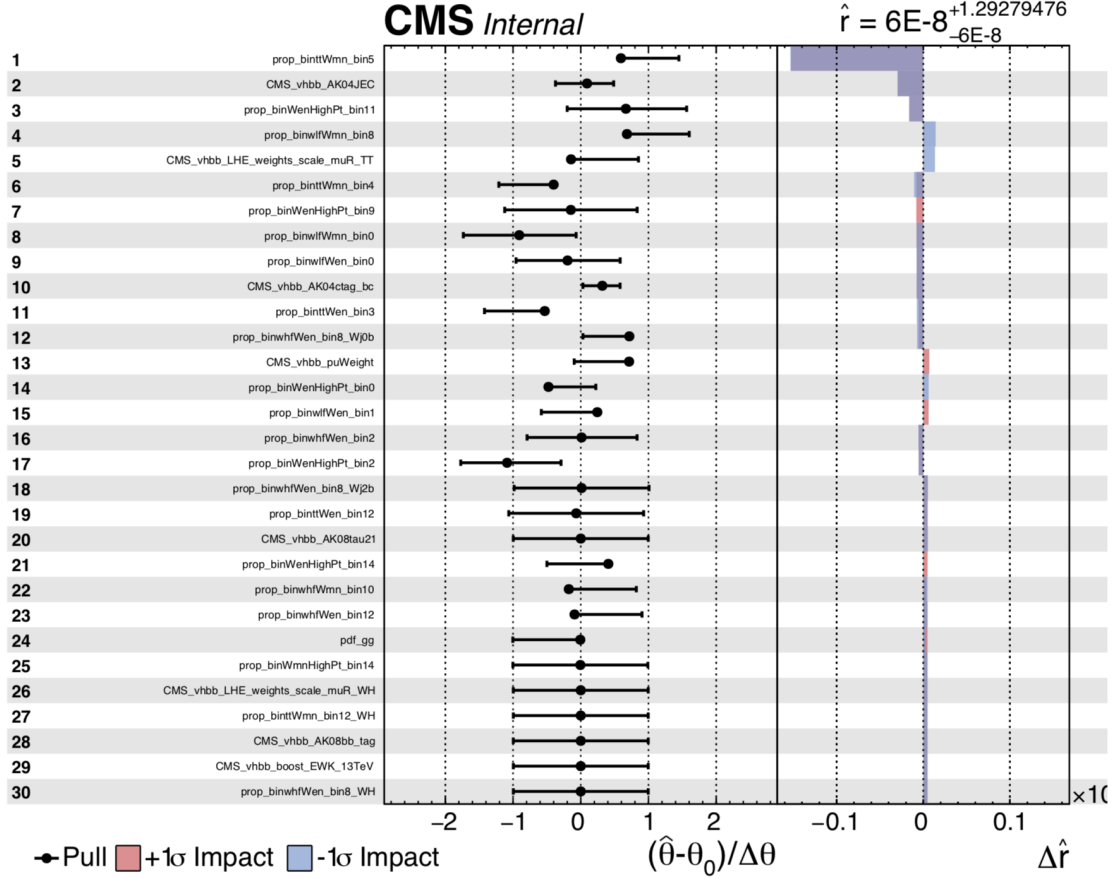


Figure 7.19: List of nuisances with the largest effect on the signal strength uncertainty. The plot also shows the best fit value of  $\mu$  ( $\hat{r}$ ) at the top and its uncertainty.

## 7.9 Comparison with resolved analysis

In the resolved  $VH(bb)$  analysis, one searches for a Higgs boson in association with a  $W$  or  $Z$  boson ( $VH$ ), in final states including 0, 1, or 2 charged leptons, and two identified bottom quark jets (AK04). The resolved  $VH(bb)$  analysis on data recorded by the CMS experiment at the LHC in 2016 observed an excess of events in data compared to the expectation in the absence of a signal [3]. The significance of this excess is 3.3 standard deviations, where the expectation from SM Higgs boson production is 2.8. The signal strength corresponding to this excess, relative to that of the SM Higgs boson production, is  $1.2 \pm 0.4$ , as shown in Figure 7.20. Combining 2017 data with previous searches by the CMS experiment for  $H(bb)$  in other production processes yields an observed significance of 5.6 standard deviations and a signal strength of  $1.04 \pm 0.2$ . [136].

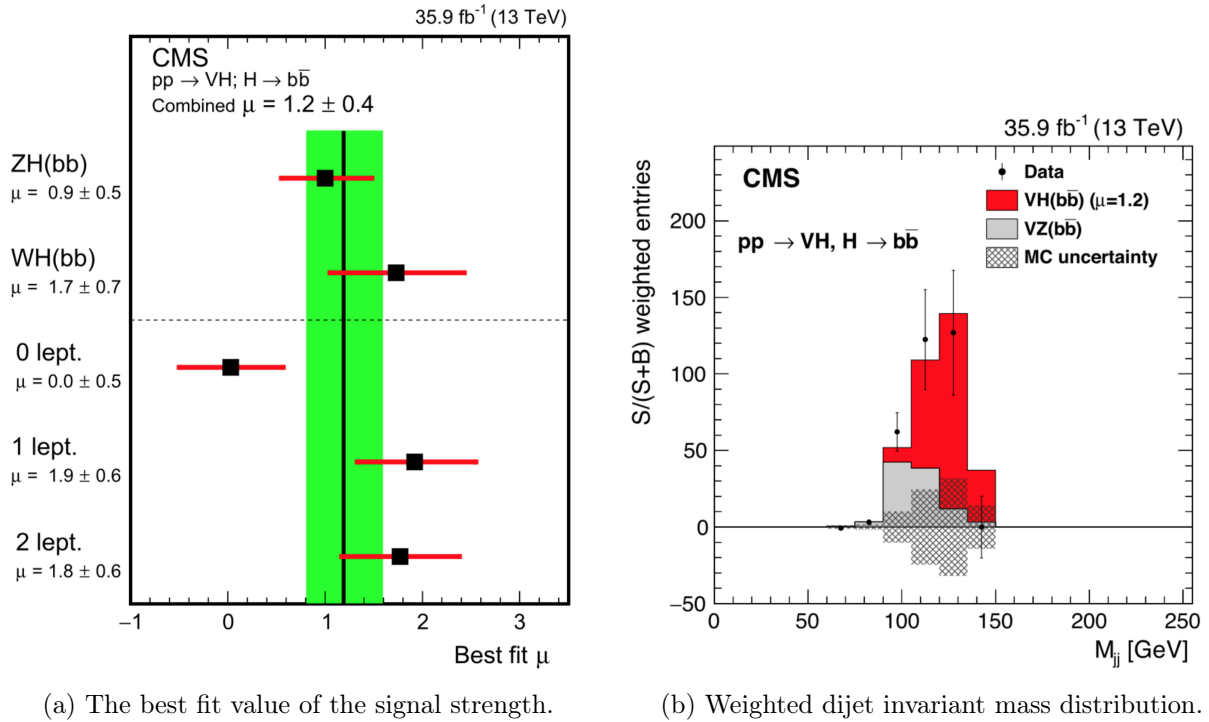


Figure 7.20: Results of resolved VH(bb) analysis on data recorded by the CMS experiment at the LHC in 2016. Figure a) The best fit value of the signal strength  $\mu$ , at  $m_H = 125.09$  GeV, is shown in black with a green uncertainty band. Figure b) Weighted dijet invariant mass distribution. Shown are data and the VH and VZ processes with all other background processes subtracted.

In addition to an improved S/B ratio according to the BDRS paper, the boosted analysis could be used to improve the sensitivity of the resolved analysis by increasing the signal sample. In the boosted regime, where the two b jets are merged, AK08 jets should provide better Higgs candidate identification and mass reconstruction compared to AK04 jets. The expected additional signal yield from the inclusion of the boosted topology for the WH signal is given in Table 7.10. Out of 3.7 expected signal events in the boosted selection, 2.6 also pass the resolved selection, which means that there is approximately only one additional expected event from the boosted topology.

Table 7.10: Overlap of boosted and resolved signal selection pre-fit yields. Normalized to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ .

The expected number of events	
Boosted	3.71
Boosted + Resolved	2.61





# Chapter 8

## Conclusion

The study of the Higgs boson production in association with a W boson in the boosted topology is presented. Two decay channels have been analyzed where the Higgs boson decays to two b quarks and the W boson decays to an electron or a muon with their corresponding neutrinos,  $W(l\nu)H(bb)$ . Measurements have been performed using data from proton-proton collisions at a centre-of-mass energy of 13 TeV collected by the CMS experiment at the LHC in 2016 and corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . Despite the fact that the Higgs boson dominantly decays to a pair of b quarks, this channel is particularly challenging due to large backgrounds, even after requiring an associated W boson. The authors of the BDRS paper [1, 2] suggested to address this problem by studying the boosted Higgs phase space. With the requirement that the Higgs boson  $p_T$  is larger than 200 GeV, and using merged jets topologies, the expected background was found to be much reduced. This motivated the work presented in this thesis. Therefore, the signal event topology is characterized by the presence of a high  $p_T$  vector boson W recoiling from the Higgs boson candidate. As a consequence of boosted regime, two b quarks from the Higgs decay should be close to each other. The key component of this analysis is good reconstruction of the Higgs boson candidate, i.e. a jet with two genuine b quarks. For that purpose, AK08 jets have been used together with appropriate b-tagging algorithms for quantifying the likelihood of having two b quarks inside the jet. The signal selection and event reconstruction are based on physical arguments: the W boson and the Higgs boson are expected to be back-to-back in the transverse plane, events should not contain many additional jets, and there should in particular not be any b-hadron activity

outside of the Higgs boson candidate.

To help control the normalization of the main background processes, and to check how well the simulated samples model the distributions of variables most relevant to the analysis, several control regions are selected in data:  $t\bar{t}$ , W boson in association with mainly heavy-flavor (HF) or light-flavor (LF) jets. A simultaneous signal-extraction fit is done on the signal and control region distributions, where the shape and normalization of these distributions are allowed to vary, for each background component, within the systematic and statistical uncertainties.

The expected upper limits in the absence of a signal is 5.51 times the SM prediction, while the observed upper limit is 4.36. The results presented in this thesis, with the expected pre-fit signal and background ratio of  $s/\sqrt{b} = 0.2$ , shows that BDRS prediction ( $s/\sqrt{b} = 2.9$ ) was too optimistic. However, additional improvements of double b-tagger, better optimization of control region definitions, higher MC and data statistics, and perhaps usage of multivariate variables in the final fit, could promote the boosted analysis as a supplement to the resolved analysis in the high transverse momentum phase space, where two b-jets from the Higgs boson decay have significant overlap and represent a difficult task for good reconstruction within the resolved analysis.

# Bibliography

- [1] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, “Jet substructure as a new higgs-search channel at the large hadron collider,” *Physical Review Letters*, vol. 100, no. 24, 2008.
- [2] J. M. Butterworth, I. Ochoa, and T. Scanlon, “Boosted higgs  $\rightarrow b\bar{b}$  in vector-boson associated production at 14TeV,” *The European Physical Journal C*, vol. 75, no. 8, 2015.
- [3] CMS Collaboration, “Evidence for the Higgs boson decay to a bottom quark-antiquark pair,” *Physics Letters B*, vol. 780, pp. 501–532, may 2018.
- [4] D. Griffiths, *Introduction to Elementary Particles*. Weinheim: WILEY - VCH, 2 ed., 2008.
- [5] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That*. New York: Addison-Wesley, 2 ed., 1989.
- [6] H. D. Politzer, “Reliable perturbative results for strong interactions?,” *Physical Review Letters*, vol. 30, no. 26, pp. 1346–1349, 1973.
- [7] D. J. Gross and F. Wilczek, “Ultraviolet behavior of non-abelian gauge theories,” *Physical Review Letters*, vol. 30, no. 26, pp. 1343–1346, 1973.
- [8] T. Y. Cao, *Conceptual developments of 20th century field theories*. Cambridge University Press, 1998.
- [9] C. D. Anderson, “The positive electron,” *Physical Review*, vol. 43, pp. 491–494, mar 1933.

- [10] J. Chadwick, "Possible existence of a neutron," *Nature*, vol. 129, pp. 312–312, feb 1932.
- [11] J. Street and E. Stevenson, "New evidence for the existence of a particle of mass intermediate between the proton and electron," *Physical Review*, vol. 52, pp. 1003–1004, nov 1937.
- [12] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, "Detection of the free neutrino: a confirmation," *Science*, vol. 124, pp. 103–104, jul 1956.
- [13] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini, and C. F. Powell, "Processes involving charged mesons," *Nature*, vol. 159, pp. 694–697, may 1947.
- [14] S. L. Glashow, "Partial-symmetries of weak interactions," *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961.
- [15] S. Weinberg, "A model of leptons," *Physical Review Letters*, vol. 19, no. 21, pp. 1264–1266, 1967.
- [16] A. Salam, "Elementary particle physics: Relativistic groups and analyticity," *Proceedings of the Eighth Nobel Symposium*, p. 367, 1968.
- [17] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Physical Review Letters*, vol. 13, no. 9, pp. 321–323, 1964.
- [18] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Physical Review Letters*, vol. 13, no. 16, pp. 508–509, 1964.
- [19] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Physical Review Letters*, vol. 13, no. 20, pp. 585–587, 1964.
- [20] F. J. Hasert *et al.*, "Search for elastic muon-neutrino electron scattering," *Physics Letters B*, vol. 46, no. 1, pp. 121–124, 1973.
- [21] F. Hasert *et al.*, "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment," *Physics Letters B*, vol. 46, no. 1, pp. 138–140, 1973.

- [22] F. Hasert *et al.*, “Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment,” *Nuclear Physics B*, vol. 73, no. 1, pp. 1–22, 1974.
- [23] UA1 Collaboration, “Experimental observation of isolated large transverse energy electrons with associated missing energy at,” *Physics Letters B*, vol. 122, no. 1, pp. 103–116, 1983.
- [24] UA2 Collaboration, “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN p collider,” *Physics Letters B*, vol. 122, no. 5-6, pp. 476–485, 1983.
- [25] J. I. Friedman and H. W. Kendall, “Deep inelastic electron scattering,” *Annual Review of Nuclear Science*, vol. 22, pp. 203–254, dec 1972.
- [26] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012.
- [27] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012.
- [28] C. Patrignani *et al. Particle Data Group*, *Chin. Phys. C*, vol. 40, (2016) and 2017 update.
- [29] H. Georgi, *Lie Algebras In Particle Physics: from Isospin To Unified Theories (Frontiers in Physics)*. Westview Press, 1999.
- [30] M. Beneke, “Renormalons,” *Physics Reports*, vol. 317, no. 1-2, pp. 1–142, 1999.
- [31] L.-L. Chau and W.-Y. Keung, “Comments on the parametrization of the kobayashi-maskawa matrix,” *Physical Review Letters*, vol. 53, no. 19, pp. 1802–1805, 1984.
- [32] J. A. Wheeler, “On the mathematical description of light nuclei by the method of resonating group structure,” *Physical Review*, vol. 52, no. 11, pp. 1107–1122, 1937.
- [33] R. P. Feynman, “The theory of positrons,” *Physical Review*, vol. 76, no. 6, pp. 749–759, 1949.

- 
- [34] Super-Kamiokande Collaboration, “Evidence for oscillation of atmospheric neutrinos,” *Physical Review Letters*, vol. 81, no. 8, pp. 1562–1567, 1998.
- [35] SNO Collaboration, “Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory,” *Physical Review Letters*, vol. 89, no. 1, 2002.
- [36] KamLAND Collaboration, “First results from KamLAND: Evidence for reactor antineutrino disappearance,” *Physical Review Letters*, vol. 90, no. 2, 2003.
- [37] D. J. Callaway, “Triviality pursuit: Can elementary scalar particles exist?,” *Physics Reports*, vol. 167, no. 5, pp. 241–320, 1988.
- [38] “Quantum Yang–Mills Theory.” <http://www.claymath.org/sites/default/files/yangmills.pdf>. Accessed: 2018-10-22.
- [39] S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*. Pearson, 2003.
- [40] L. Canetti, M. Drewes, and M. Shaposhnikov, “Matter and antimatter in the universe,” *New Journal of Physics*, vol. 14, p. 095012, sep 2012.
- [41] R. Ellis, H. Georgi, M. Machacek, H. Politzer, and G. G. Ross, “Factorization and the parton model in QCD,” *Physics Letters B*, vol. 78, no. 2-3, pp. 281–284, 1978.
- [42] J. M. Butterworth, G. Dissertori, and G. P. Salam, “Hard processes in proton-proton collisions at the large hadron collider,” *Annual Review of Nuclear and Particle Science*, vol. 62, no. 1, pp. 387–405, 2012.
- [43] “The Tevatron: 28 years of discovery and innovation.” <http://www.fnal.gov/pub/tevatron/>. Accessed: 2018-07-19.
- [44] F.-P. Schilling, “Inclusive diffraction at HERA,” *Nuclear Physics B - Proceedings Supplements*, vol. 117, pp. 403–407, 2003.
- [45] D. Boito, O. Catà, M. Golterman, M. Jamin, K. Maltman, J. Osborne, and S. Peris, “New determination of  $\alpha_s$  from hadronic tau decays,” *Physical Review D*, vol. 84, no. 11, 2011.

- [46] K.-T. Chao, H.-W. Huang, and Y.-Q. Liu, “Gluonic and leptonic decays of heavy quarkonia and the determination of  $\alpha_s(m_c)$  and  $\alpha_s(m_b)$ ,” *Physical Review D*, vol. 53, no. 1, pp. 221–230, 1996.
- [47] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language,” *Nuclear Physics B*, vol. 126, no. 2, pp. 298–318, 1977.
- [48] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics,” *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].
- [49] V. Gribov and L. Lipatov, “Deep inelastic electron scattering in perturbation theory,” *Physics Letters B*, vol. 37, no. 1, pp. 78–80, 1971.
- [50] L. N. Lipatov, “The parton model and perturbation theory,” *Sov. J. Nucl. Phys.*, vol. 20, pp. 94–102, 1975. [Yad. Fiz.20,181(1974)].
- [51] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC,” *The European Physical Journal C*, vol. 63, no. 2, pp. 189–285, 2009.
- [52] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with SHERPA 1.1,” *Journal of High Energy Physics*, vol. 2009, pp. 007–007, feb 2009.
- [53] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. 2017. doi:10.23731/CYRM-2017-002. arXiv:<http://arxiv.org/abs/1610.07922>.
- [54] O. S. Brüning *et al.*, *LHC Design Report, Vol. 1: The LHC Main Ring*. Geneva: CERN, 2004.
- [55] O. S. Brüning *et al.*, *LHC Design Report, Vol. 2: The LHC Infrastructure and General Services*. Geneva: CERN, 2004.
- [56] M. Benedikt *et al.*, *LHC Design Report, Vol. 3: The LHC Injector Chain*. Geneva: CERN, 2004.
- [57] CERN, “CERN member states map.” <https://home.cern/about/member-states>.



- 
- [58] S. Myers, “The lep collider, from design to approval and commissioning,” 1991.
  - [59] LHCb Collaboration, “Observation of Two New  $\Xi_b^-$  Baryon Resonances,” *Physical Review Letters*, vol. 114, no. 6, 2015.
  - [60] LHCb Collaboration, “Observation of  $J/\psi p$  Resonances Consistent with Pentaquark States in  $\Lambda_b^0 \rightarrow J/\psi K^- p$  Decays,” *Physical Review Letters*, vol. 115, no. 7, 2015.
  - [61] CMS Collaboration, “Evidence for collective multiparticle correlations in p-PbCollisions,” *Physical Review Letters*, vol. 115, no. 1, 2015.
  - [62] ALICE Collaboration, “Long-range angular correlations on the near and away side in p-pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV,” *Physics Letters B*, vol. 719, no. 1-3, pp. 29–41, 2013.
  - [63] ATLAS Collaboration, “Observation of associated near-side and away-side long-range correlations in  $\sqrt{s_{NN}}=5.02$  TeV Proton-lead collisions with the ATLAS detector,” *Physical Review Letters*, vol. 110, no. 18, 2013.
  - [64] L. Gray, *A Measurement of the  $Z\gamma$  Cross Section and Limits on Anomalous Triple Gauge Couplings at  $\sqrt{s} = 7$  TeV Using CMS*. PhD thesis, University of Wisconsin, Madison, 2012.
  - [65] C. Lefèvre, “The CERN accelerator complex Complexe des accélérateurs du CERN.” <https://cdsweb.cern.ch/record/1260465?ln=en>.
  - [66] V. Duvivier, “Cross section of LHC dipole. Dipole LHC: Coupe transversale.” <http://cds.cern.ch/record/843195>.
  - [67] ALEPH Collaboration, “ALEPH: A detector for electron-positron annihilations at LEP,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 294, no. 1-2, pp. 121–178, 1990.
  - [68] CMS Collaboration, “The CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, 2008.

- [69] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider,” *Journal of Instrumentation*, vol. 3, 2008. doi:10.1088/1748-0221/3/08/S08003.
- [70] ALICE Collaboration, “The ALICE experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, 2008. doi:10.1088/1748-0221/3/08/S08002.
- [71] LHCb Collaboration, “The LHCb Detector at the LHC,” *Journal of Instrumentation*, vol. 3, 2008. doi:10.1088/1748-0221/3/08/S08005.
- [72] TOTEM Collaboration, “The TOTEM experiment at the CERN Large Hadron Collider,” *Journal of Instrumentation*, vol. 3, 2008. doi:10.1088/1748-0221/3/08/S08007.
- [73] LHCf Collaboration, “The LHCf detector at the CERN Large Hadron Collider,” *Journal of Instrumentation*, vol. 3, 2008. doi:10.1088/1748-0221/3/08/S08006.
- [74] MoEDAL Collaboration, “The Physics Programme Of The MoEDAL Experiment At The LHC,” *Int.J.Mod.Phys. A*, vol. 29, 2014. doi:10.1142/S0217751X14300506.
- [75] CMS Collaboration, “CMS Luminosity - Public Results.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [76] CERN, “Interim summary report on the analysis of the 19 september 2008 incident at the lh.” [https://edms.cern.ch/ui/file/973073/1/Report\\_on\\_080919\\_incident\\_at\\_LHC\\_\\_2\\_.pdf](https://edms.cern.ch/ui/file/973073/1/Report_on_080919_incident_at_LHC__2_.pdf).
- [77] CERN, “The new HL-LHC timeline.” [https://cds.cern.ch/record/1975962/files/new\\_timeplan\\_24Sept2014\\_1\\_image.png](https://cds.cern.ch/record/1975962/files/new_timeplan_24Sept2014_1_image.png).
- [78] “About CERN.” <https://home.cern/about/experiments/cms>. Accessed: 2018-07-24.
- [79] “CMS publishes 700th research paper.” <https://cms.cern/news/cms-publishes-700th-research-paper>. Accessed: 2018-07-24.
- [80] CMS Collaboration, V. Karimäki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi, K. Freudenreich, G. Hall, R. Horisberger, M. Huhtinen and A. Cattai,

- “The CMS tracker system project : Technical Design Report,” *CMS-TDR-5*, p. 600, 1997. <http://cds.cern.ch/record/368412?ln=it>.
- [81] CMS Collaboration, “The CMS tracker : addendum to the Technical Design Report,” *CMS-TDR-5-add-1*, p. 82, 2000. <http://cds.cern.ch/record/490194?ln=it>.
- [82] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *Journal of Instrumentation*, vol. 9, 2014. doi:10.1088/1748-0221/9/10/P10009.
- [83] CMS Collaboration, “The CMS electromagnetic calorimeter project: Technical Design Report,” *CMS-TDR-4*, 1997. <http://cds.cern.ch/record/349375?ln=it>.
- [84] CMS Collaboration, P. Bloch, R. Brown, P. Lecoq and H. Rykaczewski, “Changes to CMS ECAL electronics: addendum to the Technical Design Report,” *CMS-TDR-4-add-1*, 2002. <http://cds.cern.ch/record/581342?ln=it>.
- [85] CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [86] CMS Collaboration, *The CMS magnet project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [87] CMS Collaboration, *The CMS muon project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.
- [88] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” *Journal of Instrumentation*, vol. 12, 2017. doi:10.1088/1748-0221/12/10/P10003.
- [89] CMS Collaboration, “Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at  $\sqrt{s}=7$  TeV,” *Journal of Instrumentation*, vol. 8, pp. P09009–P09009, sep 2013.
- [90] CMS Collaboration. <http://xdaq.web.cern.ch/xdaq/setup/images/HCAL.png>. Accessed: 2018-07-25.

- [91] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7\text{TeV}$ ,” *Journal of Instrumentation*, vol. 7, no. 10, pp. P10002–P10002, 2012.
- [92] CMS Collaboration, “The CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, no. 08, pp. S08004–S08004, 2008.
- [93] CMS Collaboration, “Resistive Plate Chambers.” <http://cms.web.cern.ch/news/resistive-plate-chambers>. Accessed: 2018-11-20.
- [94] CMS Collaboration, *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS, CERN, 2000.
- [95] S. Cittolin, A. Racz, and P. Spiccas, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS, Geneva: CERN, 2002.
- [96] CMS Collaboration, “The CMS trigger system,” *Journal of Instrumentation*, vol. 12, no. 01, pp. P01020–P01020, 2017.
- [97] T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to PYTHIA 8.1,” *Computer Physics Communications*, vol. 178, pp. 852–867, jun 2008.
- [98] C. Oleari, “The POWHEG BOX,” *Nuclear Physics B - Proceedings Supplements*, vol. 205-206, pp. 36–41, aug 2010.
- [99] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, “MadGraph 5: going beyond,” *Journal of High Energy Physics*, vol. 2011, jun 2011.
- [100] S. D. Williamson, “Search for Higgs-Boson Production in Association with a Top-Quark Pair in the Boosted Regime with the CMS Experiment,” 2016.
- [101] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, “Matching matrix elements and shower evolution for top-pair production in hadronic collisions,” *Journal of High Energy Physics*, vol. 2007, pp. 013–013, jan 2007.
- [102] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, “QCD matrix elements + parton showers,” *Journal of High Energy Physics*, vol. 2001, pp. 063–063, nov 2001.

- 
- [103] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, “Parton fragmentation and string dynamics,” *Physics Reports*, vol. 97, pp. 31–145, jul 1983.
- [104] S. Agostinelli *et al.*, “Geant4, a simulation toolkit,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, pp. 250–303, jul 2003.
- [105] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET,” Tech. Rep. CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009.
- [106] W. Adam, B. Mangano, T. Speer, and T. Todorov, “Track reconstruction in the CMS tracker,” 2005.
- [107] F.-P. Schilling, “Track Reconstruction and Alignment with the CMS Silicon Tracker,” Tech. Rep. CMS-CR-2006-061, CERN, Geneva, Sep 2006. Manuscript not submitted to the proceedings ICHEP06.
- [108] W. Waltenberger, “Adaptive Vertex Reconstruction,” Tech. Rep. CMS-NOTE-2008-033, CERN, Geneva, Jul 2008.
- [109] T. Speer, K. Prokofiev, R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Vertex Fitting in the CMS Tracker,” Tech. Rep. CMS-NOTE-2006-032, CERN, Geneva, Feb 2006.
- [110] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, “Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC,” 2005.
- [111] CMS Collaboration, “Electron and photon performance in CMS with the full 2016 data sample.” [https://cds.cern.ch/record/2255497/files/DP2017\\_004.pdf](https://cds.cern.ch/record/2255497/files/DP2017_004.pdf). Accessed: 2018-11-21.
- [112] G. C. Blazey, J. R. Dittmann, S. D. Ellis, V. D. Elvira, K. C. Frame, S. Grinstein, R. Hirosky, R. Piegaiia, H. Schellman, R. Snihur, V. Sorin, and D. Zeppenfeld, “Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop,” Tech. Rep. hep-ex/0005012, May 2000.

- [113] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, pp. 063–063, apr 2008.
- [114] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, “Better jet clustering algorithms,” *Journal of High Energy Physics*, vol. 1997, pp. 001–001, aug 1997.
- [115] R. Atkin, “Review of jet reconstruction algorithms,” *Journal of Physics: Conference Series*, vol. 645, p. 012008, oct 2015.
- [116] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS,” *Journal of Instrumentation*, vol. 6, pp. P11002–P11002, nov 2011.
- [117] J. Thaler and K. V. Tilburg, “Identifying boosted objects with n-subjettiness,” *Journal of High Energy Physics*, vol. 2011, mar 2011.
- [118] S. Catani, Y. Dokshitzer, M. Seymour, and B. Webber, “Longitudinally-invariant  $k_{\perp}$ -clustering algorithms for hadron-hadron collisions,” *Nuclear Physics B*, vol. 406, pp. 187–224, sep 1993.
- [119] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Physical Review D*, vol. 48, pp. 3160–3166, oct 1993.
- [120] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Techniques for improved heavy particle searches with jet substructure,” *Physical Review D*, vol. 80, sep 2009.
- [121] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV,” *Journal of Instrumentation*, vol. 13, pp. P05011–P05011, may 2018.
- [122] CMS Collaboration, “Identification of b-quark jets with the CMS experiment,” *Journal of Instrumentation*, vol. 8, pp. P04013–P04013, apr 2013.
- [123] T. Müller, C. Piasecki, G. Quast, and C. Weiser, “Inclusive Secondary Vertex Reconstruction in Jets,” Tech. Rep. CMS-NOTE-2006-027, CERN, Geneva, Jan 2006.
- [124] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period,” Tech. Rep. CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

- 
- [125] P. Nason, “A new method for combining NLO QCD with shower monte carlo algorithms,” *Journal of High Energy Physics*, vol. 2004, pp. 040–040, nov 2004.
- [126] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX,” *Journal of High Energy Physics*, vol. 2010, jun 2010.
- [127] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *Journal of High Energy Physics*, vol. 2014, jul 2014.
- [128] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO,” *Journal of High Energy Physics*, vol. 2012, dec 2012.
- [129] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements,” *The European Physical Journal C*, vol. 76, mar 2016.
- [130] CERN, “Cern yellow reports: Monographs, vol 2 (2017): Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector,” 2017.
- [131] A. L. Read, “Presentation of search results: theCLstechnique,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 28, pp. 2693–2704, sep 2002.
- [132] ATLAS, CMS, LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in summer 2011,” 2011.
- [133] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. D. Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, “Parton distributions for the LHC run II,” *Journal of High Energy Physics*, vol. 2015, apr 2015.
- [134] M. Ciccolini, A. Denner, and S. Dittmaier, “Strong and electroweak corrections to the production of a higgs boson and 2 jets via weak interactions at the large hadron collider,” *Physical Review Letters*, vol. 99, oct 2007.
- [135] M. Ciccolini, D. A., and D. S., “Electroweak and QCD corrections to higgs production via vector-boson fusion at the CERN LHC,” *Physical Review D*, vol. 77, jan 2008.

- [136] CMS Collaboration, “Observation of Higgs Boson Decay to Bottom Quarks,” *Physical Review Letters*, vol. 121, sep 2018.





# Appendix A

## Conventions

- The conventional units have been used  $\hbar = c = 1$ .
- The electric charge is measured in units of electron charge  $e = 1.602 \cdot 10^{-19}$  C.
- The antielectron ( $e^+$ ) is traditionally called positron.
- The radiation length  $X_0$  is defined as the path that a highly energetic particle must traverse in a material to reduce its energy to a fraction equal to  $1/e$  of the initial value.



# Curriculum vitae

## Personal

Name: Benjamin Mesić  
Date of birth: September 1, 1990  
Place of birth: Osijek, Croatia  
Citizenship: Croatian  
Email: benjamin.mesic@cern.ch

## Education

2014 -        **Doctoral studies**, University of Zagreb, Faculty of Science  
Title : Study of the Higgs boson production in association with a massive electroweak boson in final states with two b quarks and two leptons  
Advisor : Prof. Vuko Brigljević  
Thesis defense is scheduled for December, 19th 2018  
2009 - 2014   **MSc**, University of Zagreb, Faculty of Science  
Title : Birkhoff's Theorem  
Advisor : Prof. Ivica Smolić

## Positions

2014 - 2018   Research assistant at Ruđer Bošković Institute and member of the CMS collaboration at CERN

## Publications

1. **“Studies of  $B_{s2}^*(5840)^0$  and  $B_{s1}(5830)^0$  mesons including the observation of the  $B_{s2}^*(5840)^0 \rightarrow B^0 K_S^0$  decay in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 11, 939 (2018)
2. **“Performance of reconstruction and identification of  $\tau$  leptons decaying to hadrons and  $\nu_\tau$  in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JINST **13**, no. 10, P10005 (2018)
3. **“Search for physics beyond the standard model in high-mass diphoton events from proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **98**, no. 9, 092001 (2018)
4. **“Charged-particle nuclear modification factors in XeXe collisions at  $\sqrt{s_{NN}} = 5.44$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1810**, 138 (2018)
5. **“Observation of Higgs boson decay to bottom quarks”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, no. 12, 121801 (2018)
6. **“Measurement of jet substructure observables in  $t\bar{t}$  events from proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **98**, 092014 (2018)
7. **“Search for a charged Higgs boson decaying to charm and bottom quarks in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 115 (2018)
8. **“Evidence for the associated production of a single top quark and a photon in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, 221802 (2018)
9. **“Search for long-lived particles with displaced vertices in multijet events in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **98**, no. 9, 092011 (2018)
10. **“Searches for pair production of charginos and top squarks in final states with two oppositely charged leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 079 (2018)

- 
11. **“Measurements of the differential jet cross section as a function of the jet mass in dijet events from proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 113 (2018)
  12. **“Precision measurement of the structure of the CMS inner tracking system using nuclear interactions”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JINST **13**, no. 10, P10034 (2018)
  13. **“Search for supersymmetry in events with a  $\tau$  lepton pair and missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 151 (2018)
  14. **“Measurement of charged particle spectra in minimum-bias events from proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 9, 697 (2018)
  15. **“Search for the decay of a Higgs boson in the  $\ell\ell\gamma$  channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 152 (2018)
  16. **“Search for dark matter produced in association with a Higgs boson decaying to  $\gamma\gamma$  or  $\tau^+\tau^-$  at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1809**, 046 (2018)
  17. **“Observation of the  $Z \rightarrow \psi\ell^+\ell^-$  decay in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, no. 14, 141801 (2018)
  18. **“Search for resonant pair production of Higgs bosons decaying to bottom quark-antiquark pairs in proton-proton collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 152 (2018)
  19. **“Search for a singly produced third-generation scalar leptoquark decaying to a  $\tau$  lepton and a bottom quark in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1807**, 115 (2018)
  20. **“Search for pair-produced resonances each decaying into at least four quarks in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, no. 14, 141802 (2018)

21. **“Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 9, 701 (2018)
22. **“Search for narrow and broad dijet resonances in proton-proton collisions at  $\sqrt{s} = 13$  TeV and constraints on dark matter mediators and other new particles”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 130 (2018)
23. **“Search for Higgs boson pair production in the  $\gamma\gamma b\bar{b}$  final state in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **788**, 7 (2019)
24. **“Search for beyond the standard model Higgs bosons decaying into a  $b\bar{b}$  pair in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 113 (2018)
25. **“Observation of the  $\chi_{b1}(3P)$  and  $\chi_{b2}(3P)$  and measurement of their masses”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, 092002 (2018)
26. **“Constraints on models of scalar and vector leptoquarks decaying to a quark and a neutrino at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **98**, no. 3, 032005 (2018)
27. **“Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two b quarks and two  $\tau$  leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **785**, 462 (2018)
28. **“Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1810**, 117 (2018)
29. **“Search for black holes and sphalerons in high-multiplicity final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 042 (2018)
30. **“Search for top squarks decaying via four-body or chargino-mediated modes in single-lepton final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1809**, 065 (2018)

- 
31. **“Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state of two muons and two  $\tau$  leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1811**, 018 (2018)
32. **“Search for vector-like T and B quark pairs in final states with leptons at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 177 (2018)
33. **“Measurement of the groomed jet mass in PbPb and pp collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1810**, 161 (2018)
34. **“Constraining gluon distributions in nuclei using dijets in proton-proton and proton-lead collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, no. 6, 062002 (2018)
35. **“Measurement of the top quark mass with lepton+jets final states using p p collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 11, 891 (2018)
36. **“Elliptic flow of charm and strange hadrons in high-multiplicity pPb collisions at  $\sqrt{s_{\text{NN}}} = 8.16$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **121**, no. 8, 082301 (2018)
37. **“Search for disappearing tracks as a signature of new long-lived particles in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 016 (2018)
38. **“Measurement of differential cross sections for Z boson production in association with jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 11, 965 (2018)
39. **“Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JINST **13**, no. 06, P06015 (2018)
40. **“Observation of  $t\bar{t}H$  production”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 23, 231801 (2018)



41. **“Search for a new scalar resonance decaying to a pair of Z bosons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 127 (2018)
42. **“Search for high-mass resonances in final states with a lepton and missing transverse momentum at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 128 (2018)
43. **“Search for a heavy right-handed W boson and a heavy neutrino in events with two same-flavor leptons and two jets at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1805**, no. 05, 148 (2018)
44. **“Search for a heavy resonance decaying into a Z boson and a Z or W boson in 2l2q final states at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1809**, 101 (2018)
45. **“Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 11, 112003 (2018)
46. **“Search for new physics in dijet angular distributions using proton–proton collisions at  $\sqrt{s} = 13$  TeV and constraints on dark matter and other models”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 9, 789 (2018)
47. **“Search for  $t\bar{t}H$  production in the all-jet final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 101 (2018)
48. **“Search for additional neutral MSSM Higgs bosons in the  $\tau\tau$  final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1809**, 007 (2018)
49. **“Search for high-mass resonances in dilepton final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 120 (2018)
50. **“Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying  $\tau$  leptons at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 066 (2018)

- 
51. **“Observation of proton-tagged, central (semi)exclusive production of high-mass lepton pairs in pp collisions at 13 TeV with the CMS-TOTEM precision proton spectrometer”**  
A. M. Sirunyan *et al.* [CMS and TOTEM Collaborations], JHEP **1807**, 153 (2018)
  52. **“Search for a heavy resonance decaying into a Z boson and a vector boson in the  $\nu\bar{\nu}q\bar{q}$  final state”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1807**, 075 (2018)
  53. **“Measurements of differential cross sections of top quark pair production as a function of kinematic event variables in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 002 (2018)
  54. **“Search for third-generation scalar leptoquarks decaying to a top quark and a  $\tau$  lepton at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, 707 (2018)
  55. **“Jet properties in PbPb and pp collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1805**, 006 (2018)
  56. **“Search for a heavy resonance decaying to a pair of vector bosons in the lepton plus merged jet final state at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1805**, 088 (2018)
  57. **“Search for narrow resonances in the b-tagged dijet mass spectrum in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 20, 201801 (2018)
  58. **“Measurement of the  $\Lambda_b$  polarization and angular parameters in  $\Lambda_b \rightarrow J/\psi \Lambda$  decays from pp collisions at  $\sqrt{s} = 7$  and 8 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 7, 072010 (2018)
  59. **“Search for heavy neutral leptons in events with three charged leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 22, 221801 (2018)
  60. **“Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1807**, 161 (2018)

61. **“Search for natural and split supersymmetry in proton-proton collisions at  $\sqrt{s} = 13$  TeV in final states with jets and missing transverse momentum”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1805**, 025 (2018)
62. **“Search for single production of vector-like quarks decaying to a b quark and a Higgs boson”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 031 (2018)
63. **“Search for lepton-flavor violating decays of heavy resonances and quantum black holes to  $e\mu$  final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1804**, 073 (2018)
64. **“Comparing transverse momentum balance of b jet pairs in pp and PbPb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 181 (2018)
65. **“Search for dark matter in events with energetic, hadronically decaying top quarks and missing transverse momentum at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 027 (2018)
66. **“Combined search for electroweak production of charginos and neutralinos in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 160 (2018)
67. **“Measurement of the  $Z\gamma^* \rightarrow \tau\tau$  cross section in pp collisions at  $\sqrt{s} = 13$  TeV and validation of  $\tau$  lepton analysis techniques”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 9, 708 (2018)
68. **“Search for new physics in events with two soft oppositely charged leptons and missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **782**, 440 (2018)
69. **“Search for decays of stopped exotic long-lived particles produced in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1805**, 127 (2018)
70. **“Electroweak production of two jets in association with a Z boson in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 7, 589 (2018)

- 
71. **“Measurement of prompt and nonprompt charmonium suppression in PbPb collisions at 5.02 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 6, 509 (2018)
  72. **“Search for  $R$ -parity violating supersymmetry in pp collisions at  $\sqrt{s} = 13$  TeV using b jets in a final state with a single lepton, many jets, and high sum of large-radius jet masses”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **783**, 114 (2018)
  73. **“Search for Physics Beyond the Standard Model in Events with High-Momentum Higgs Bosons and Missing Transverse Momentum in Proton-Proton Collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 24, 241801 (2018)
  74. **“Bose-Einstein correlations in  $pp$ ,  $p\text{Pb}$ , and PbPb collisions at  $\sqrt{s_{NN}} = 0.9 - 7$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. C **97**, no. 6, 064912 (2018)
  75. **“Search for lepton flavour violating decays of the Higgs boson to  $\mu\tau$  and  $e\tau$  in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 001 (2018)
  76. **“Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JINST **13**, no. 05, P05011 (2018)
  77. **“Search for the  $X(5568)$  state decaying into  $B_s^0\pi^\pm$  in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 20, 202005 (2018)
  78. **“Azimuthal correlations for inclusive 2-jet, 3-jet, and 4-jet events in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 7, 566 (2018)
  79. **“Search for  $Z\gamma$  resonances using leptonic and hadronic final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1809**, 148 (2018)
  80. **“Measurement of the associated production of a single top quark and a Z boson in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **779**, 358 (2018)

81. **“Search for the flavor-changing neutral current interactions of the top quark and the Higgs boson which decays into a pair of b quarks at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1806**, 102 (2018)
82. **“Search for new physics in final states with an energetic jet or a hadronically decaying  $W$  or  $Z$  boson and transverse momentum imbalance at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 9, 092005 (2018)
83. **“Constraints on the double-parton scattering cross section from same-sign  $W$  boson pair production in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1802**, 032 (2018)
84. **“Search for pair production of excited top quarks in the lepton + jets final state”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **778**, 349 (2018)
85. **“Study of jet quenching with isolated-photon+jet correlations in PbPb and pp collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **785**, 14 (2018)
86. **“Search for new long-lived particles at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **780**, 432 (2018)
87. **“Search for gauge-mediated supersymmetry in events with at least one photon and missing transverse momentum in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **780**, 118 (2018)
88. **“Search for excited quarks of light and heavy flavor in  $\gamma$ + jet final states in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **781**, 390 (2018)
89. **“Search for  $ZZ$  resonances in the  $2\ell 2\nu$  final state in proton-proton collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 003 (2018)
90. **“Measurement of the underlying event activity in inclusive  $Z$  boson production in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1807**, 032 (2018)

- 
91. **“Measurement of the inclusive  $t\bar{t}$  cross section in pp collisions at  $\sqrt{s} = 5.02$  TeV using final states with at least one charged lepton”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 115 (2018)
  92. **“Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1808**, 011 (2018)
  93. **“Measurement of associated Z + charm production in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 4, 287 (2018)
  94. **“Search for top squarks and dark matter particles in opposite-charge dilepton final states at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 3, 032009 (2018)
  95. **“Search for new physics in events with a leptonically decaying Z boson and a large transverse momentum imbalance in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 4, 291 (2018)
  96. **“Search for supersymmetry in proton-proton collisions at 13 TeV using identified top quarks”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 1, 012007 (2018)
  97. **“Measurement of quarkonium production cross sections in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **780**, 251 (2018)
  98. **“Search for standard model production of four top quarks with same-sign and multilepton final states in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 2, 140 (2018)
  99. **“Pseudorapidity distributions of charged hadrons in proton-lead collisions at  $\sqrt{s_{NN}} = 5.02$  and 8.16 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1801**, 045 (2018)
  100. **“Search for supersymmetry in events with at least three electrons or muons, jets, and missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1802**, 067 (2018)

101. **“Measurement of b hadron lifetimes in pp collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 6, 457 (2018),  
Erratum: [Eur. Phys. J. C **78**, no. 7, 561 (2018)]
102. **“Measurement of differential cross sections in the kinematic angular variable  $\phi^*$  for inclusive Z boson production in pp collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 172 (2018)
103. **“Pseudorapidity and transverse momentum dependence of flow harmonics in pPb and PbPb collisions”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. C **98**, no. 4, 044902 (2018)
104. **“Search for a massive resonance decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **781**, 244 (2018)
105. **“Measurement of angular parameters from the decay  $B^0 \rightarrow K^{*0}\mu^+\mu^-$  in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **781**, 517 (2018)
106. **“Study of dijet events with a large rapidity gap between the two leading jets in pp collisions at  $\sqrt{s} = 7$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, no. 3, 242 (2018)
107. **“Search for pair production of vector-like quarks in the  $bW\bar{b}W$  channel from proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **779**, 82 (2018)
108. **“Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1801**, 097 (2018)
109. **“Search for supersymmetry in events with one lepton and multiple jets exploiting the angular correlation between the lepton and the missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **780**, 384 (2018)
110. **“Observation of Correlated Azimuthal Anisotropy Fourier Harmonics in  $pp$  and  $p + Pb$  Collisions at the LHC”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 9, 092301 (2018)

- 
111. **“Search for new phenomena in final states with two opposite-charge, same-flavor leptons, jets, and missing transverse momentum in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 076 (2018)
112. **“Measurements of the  $pp \rightarrow ZZ$  production cross section and the  $Z \rightarrow 4\ell$  branching fraction, and constraints on anomalous triple gauge couplings at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **78**, 165 (2018), Erratum: [Eur. Phys. J. C **78**, no. 6, 515 (2018)]
113. **“Evidence for the Higgs boson decay to a bottom quark–antiquark pair”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **780**, 501 (2018)
114. **“Observation of top quark production in proton-nucleus collisions”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 24, 242001 (2017)
115. **“Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 8, 081801 (2018)
116. **“Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 7, 071802 (2018)
117. **“Search for electroweak production of charginos and neutralinos in multilepton final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 166 (2018)
118. **“Combination of inclusive and differential  $t\bar{t}$  charge asymmetry measurements using ATLAS and CMS data at  $\sqrt{s} = 7$  and 8 TeV”**  
M. Aaboud *et al.* [ATLAS and CMS Collaborations], JHEP **1804**, 033 (2018)
119. **“Search for Higgsino pair production in  $pp$  collisions at  $\sqrt{s} = 13$  TeV in final states with large missing transverse momentum and two Higgs bosons decaying via  $H \rightarrow b\bar{b}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 3, 032007 (2018)



120. **“Search for supersymmetry with Higgs boson to diphoton decays using the razor variables at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **779**, 166 (2018)
121. **“Measurement of the Splitting Function in  $pp$  and Pb-Pb Collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 14, 142302 (2018)
122. **“Search for heavy resonances decaying to a top quark and a bottom quark in the lepton+jets final state in proton–proton collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **777**, 39 (2018)
123. **“Search for Evidence of the Type-III Seesaw Mechanism in Multilepton Final States in Proton-Proton Collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 22, 221802 (2017)
124. **“Measurement of normalized differential  $t\bar{t}$  cross sections in the dilepton channel from  $pp$  collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1804**, 060 (2018)
125. **“Principal-component analysis of two-particle azimuthal correlations in PbPb and  $p$ Pb collisions at CMS”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. C **96**, no. 6, 064902 (2017)
126. **“Search for massive resonances decaying into  $WW$ ,  $WZ$ ,  $ZZ$ ,  $qW$ , and  $qZ$  with dijet final states at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **97**, no. 7, 072006 (2018)
127. **“Nuclear modification factor of  $D^0$  mesons in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **782**, 474 (2018)
128. **“Search for resonant and nonresonant Higgs boson pair production in the  $b\bar{b}l\nu l\nu$  final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1801**, 054 (2018)
129. **“Measurement of prompt  $D^0$  meson azimuthal anisotropy in Pb-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 20, 202301 (2018)

- 
130. **“Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **774**, 682 (2017)
131. **“Constraints on the chiral magnetic effect using charge-dependent azimuthal correlations in  $p$ Pb and PbPb collisions at the CERN Large Hadron Collider”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. C **97**, no. 4, 044912 (2018)
132. **“Search for single production of a vector-like T quark decaying to a Z boson and a top quark in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **781**, 574 (2018)
133. **“Observation of the Higgs boson decay to a pair of  $\tau$  leptons with the CMS detector”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **779**, 283 (2018)
134. **“Search for a light pseudoscalar Higgs boson produced in association with bottom quarks in pp collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1711**, 010 (2017)
135. **“Search for the pair production of third-generation squarks with two-body decays to a bottom or charm quark and a neutralino in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **778**, 263 (2018)
136. **“Search for supersymmetry in events with at least one photon, missing transverse momentum, and large transverse event activity in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1712**, 142 (2017)
137. **“Measurement of the differential cross sections for the associated production of a W boson and jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **96**, no. 7, 072005 (2017)
138. **“Search for natural supersymmetry in events with top quark pairs and photons in pp collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1803**, 167 (2018)
139. **“Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 005 (2017)

140. **“Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at  $\sqrt{s}=13\text{TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **778**, 101 (2018)
141. **“Search for heavy resonances that decay into a vector boson and a Higgs boson in hadronic final states at  $\sqrt{s}=13\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 9, 636 (2017)
142. **“Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **775**, 1 (2017)
143. **“Search for Higgs boson pair production in the  $bb\tau\tau$  final state in proton-proton collisions at  $\sqrt{s}=8\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **96**, no. 7, 072004 (2017)
144. **“Measurement of charged pion, kaon, and proton production in proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **96**, no. 11, 112003 (2017)
145. **“Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at  $\sqrt{s}=13\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1711**, 047 (2017)
146. **“Search for electroweak production of charginos and neutralinos in WH events in proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1711**, 029 (2017)
147. **“Measurement of the semileptonic  $t\bar{t} + \gamma$  production cross section in pp collisions at  $\sqrt{s}=8\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 006 (2017)
148. **“Suppression of Excited  $\Upsilon$  States Relative to the Ground State in Pb-Pb Collisions at  $\sqrt{s_{\text{NN}}}=5.02\text{TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **120**, no. 14, 142301 (2018)
149. **“Measurements of jet charge with dijet events in pp collisions at  $\sqrt{s}=8\text{ TeV}$ ”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 131 (2017)
150. **“Particle-flow reconstruction and global event description with the CMS detector”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JINST **12**, no. 10, P10003 (2017)

- 
151. **“Search for top squark pair production in pp collisions at  $\sqrt{s} = 13$  TeV using single lepton events”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 019 (2017)
152. **“Searches for  $W'$  bosons decaying to a top quark and a bottom quark in proton-proton collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1708**, 029 (2017)
153. **“Search for new physics in the monophoton final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 073 (2017)
154. **“Search for pair production of vector-like T and B quarks in single-lepton final states using boosted jet substructure in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1711**, 085 (2017)
155. **“Search for dark matter produced in association with heavy-flavor quark pairs in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 12, 845 (2017)
156. **“Search for top quark partners with charge 5/3 in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1708**, 073 (2017)
157. **“Search for Low Mass Vector Resonances Decaying to Quark-Antiquark Pairs in Proton-Proton Collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 11, 111802 (2017)
158. **“Measurements of  $t\bar{t}$  cross sections in association with  $b$  jets and inclusive jets and their ratio using dilepton final states in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **776**, 355 (2018)
159. **“Combination of searches for heavy resonances decaying to WW, WZ, ZZ, WH, and ZH boson pairs in proton–proton collisions at  $\sqrt{s} = 8$  and 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **774**, 533 (2017)
160. **“Measurement of the  $B^\pm$  Meson Nuclear Modification Factor in Pb-Pb Collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**

- A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 15, 152301 (2017)
161. **“Search for Supersymmetry in  $pp$  Collisions at  $\sqrt{s} = 13$  TeV in the Single-Lepton Final State Using the Sum of Masses of Large-Radius Jets”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 15, 151802 (2017)
  162. **“Search for new phenomena with the  $M_{T2}$  variable in the all-hadronic final state produced in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 10, 710 (2017)
  163. **“Search for Charged Higgs Bosons Produced via Vector Boson Fusion and Decaying into a Pair of  $W$  and  $Z$  Bosons Using  $pp$  Collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 14, 141802 (2017)
  164. **“Measurement of the triple-differential dijet cross section in proton-proton collisions at  $\sqrt{s} = 8$  TeV and constraints on parton distribution functions”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 11, 746 (2017)
  165. **“Search for black holes in high-multiplicity final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **774**, 279 (2017)
  166. **“Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **96**, no. 3, 032003 (2017)
  167. **“Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton–proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 9, 578 (2017)
  168. **“Measurement of the top quark mass in the dileptonic  $t\bar{t}$  decay channel using the mass observables  $M_{b\ell}$ ,  $M_{T2}$ , and  $M_{b\ell\nu}$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. D **96**, no. 3, 032002 (2017)
  169. **“Search for  $t\bar{t}$  resonances in highly boosted lepton+jets and fully hadronic final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1707**, 001 (2017)

- 
170. **“Measurements of the  $pp \rightarrow W\gamma\gamma$  and  $pp \rightarrow Z\gamma\gamma$  cross sections and limits on anomalous quartic gauge couplings at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 072 (2017)
171. **“Search for new physics with dijet angular distributions in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1707**, 013 (2017)
172. **“Search for a heavy resonance decaying to a top quark and a vector-like top quark at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1709**, 053 (2017)
173. **“Measurement of the jet mass in highly boosted  $t\bar{t}$  events from pp collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 7, 467 (2017)
174. **“Search for anomalous couplings in boosted  $WW/WZ \rightarrow \ell\nu q\bar{q}$  production in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **772**, 21 (2017)
175. **“Search for associated production of dark matter with a Higgs boson decaying to  $b\bar{b}$  or  $\gamma\gamma$  at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1710**, 180 (2017)
176. **“Search for third-generation scalar leptoquarks and heavy right-handed neutrinos in final states with two tau leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1707**, 121 (2017)
177. **“Measurement of the top quark mass using single top quark events in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 5, 354 (2017)
178. **“Search for dark matter produced with an energetic jet or a hadronically decaying W or Z boson at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1707**, 014 (2017)
179. **“Measurement of double-differential cross sections for top quark pair production in pp collisions at  $\sqrt{s} = 8$  TeV and impact on parton distribution functions”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 7, 459 (2017)

180. **“Search for standard model production of four top quarks in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **772**, 336 (2017)
181. **“Measurement of prompt and nonprompt  $J/\psi$  production in pp and pPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 4, 269 (2017)
182. **“Search for associated production of a Z boson with a single top quark and for tZ flavour-changing interactions in pp collisions at  $\sqrt{s} = 8$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1707**, 003 (2017)
183. **“Study of Jet Quenching with  $Z + \text{jet}$  Correlations in Pb-Pb and  $pp$  Collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **119**, no. 8, 082301 (2017)
184. **“Azimuthal anisotropy of charged particles with transverse momentum up to 100 GeV/  $c$  in PbPb collisions at  $\sqrt{s_{NN}}=5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **776**, 195 (2018)
185. **“Measurement of the inclusive energy spectrum in the very forward direction in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1708**, 046 (2017)
186. **“Search for single production of vector-like quarks decaying into a b quark and a W boson in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **772**, 634 (2017)
187. **“Search for single production of vector-like quarks decaying to a Z boson and a top or a bottom quark in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1705**, 029 (2017)
188. **“Measurement of the  $t\bar{t}$  production cross section using events with one lepton and at least one jet in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1709**, 051 (2017)
189. **“Search for dark matter and unparticles in events with a Z boson and missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1703**, 061 (2017), Erratum: [JHEP **1709**, 106 (2017)]

- 
190. **“Search for light bosons in decays of the 125 GeV Higgs boson in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
V. Khachatryan *et al.* [CMS Collaboration], JHEP **1710**, 076 (2017)
  191. **“Mechanical stability of the CMS strip tracker measured with a laser alignment system”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JINST **12**, no. 04, P04023 (2017)
  192. **“Search for high-mass  $Z\gamma$  resonances in proton-proton collisions at  $\sqrt{s} = 8$  and 13 TeV using jet substructure techniques”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **772**, 363 (2017)
  193. **“Search for massive resonances decaying into WW, WZ or ZZ bosons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1703**, 162 (2017)
  194. **“Measurements of the charm jet cross section and nuclear modification factor in pPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **772**, 306 (2017)
  195. **“Search for electroweak production of a vector-like quark decaying to a top quark and a Higgs boson using boosted topologies in fully hadronic final states”**  
A. M. Sirunyan *et al.* [CMS Collaboration], JHEP **1704**, 136 (2017)
  196. **“Searches for pair production of third-generation squarks in  $\sqrt{s} = 13$  TeV pp collisions”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Eur. Phys. J. C **77**, no. 5, 327 (2017)
  197. **“Search for CP violation in  $t\bar{t}$  production and decay in proton-proton collisions at  $\sqrt{s} = 8$  TeV”**  
V. Khachatryan *et al.* [CMS Collaboration], JHEP **1703**, 101 (2017)
  198. **“Search for dijet resonances in proton–proton collisions at  $\sqrt{s} = 13$  TeV and constraints on dark matter and other models”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **769**, 520 (2017), Erratum: [Phys. Lett. B **772**, 882 (2017)]
  199. **“Charged-particle nuclear modification factors in PbPb and pPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
V. Khachatryan *et al.* [CMS Collaboration], JHEP **1704**, 039 (2017)



200. **“Suppression of  $\Upsilon(1S)$ ,  $\Upsilon(2S)$  and  $\Upsilon(3S)$  production in PbPb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV”**  
V. Khachatryan *et al.* [CMS Collaboration], Phys. Lett. B **770**, 357 (2017)
201. **“Relative Modification of Prompt  $\psi(2S)$  and  $J/\psi$  Yields from pp to PbPb Collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **118**, no. 16, 162301 (2017)
202. **“Cross section measurement of  $t$ -channel single top quark production in pp collisions at  $\sqrt{s} = 13$  TeV”**  
A. M. Sirunyan *et al.* [CMS Collaboration], Phys. Lett. B **772**, 752 (2017)
203. **“Observation of charge-dependent azimuthal correlations in  $p$ -Pb collisions and its implication for the search for the chiral magnetic effect”**  
V. Khachatryan *et al.* [CMS Collaboration], Phys. Rev. Lett. **118**, no. 12, 122301 (2017)