Departamento de Investigación Básica
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Departamento de Física Teórica
Facultad de Ciencias, Universidad Autónoma de Madrid

Measurement of the tZq production cross section in pp collisions at $\sqrt{s} = 13$ TeV using CMS data

---

Thesis submitted by

María del Mar Barrio Luna

for the degree of Doctor of Physics (Ph.D.)

Supervised by

Dr. Mara Senghi Soares

Madrid, July 2021

Departamento de Investigación Básica
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Departamento de Física Teórica
Facultad de Ciencias, Universidad Autónoma de Madrid

Medida de la sección eficaz de producción de tZq en colisiones pp a $\sqrt{s} = 13$ TeV en CMS

Tesis presentada por

María del Mar Barrio Luna

para optar al grado de Doctora en Física

Supervisada por

Dra. Mara Senghi Soares

Madrid, julio 2021

I don't know what's coming next. But I do know it's gonna be just like this. Hard, painful.

Buffy, the Vampire Slayer (2003)

A mis padres, mi hermano, Mara y Caridad.

Abstract

Measurement of the tZq production cross section in pp collisions at $\sqrt{s} = 13$ TeV using CMS data

by María del Mar Barrio Luna

This document presents a measurement of the production cross section of a single top quark in association with a Z boson (and an additional quark), a rare standard model process which is also an irreducible background to many important searches at the Large Hadron Collider (LHC).

The process is studied using events with three leptons (electrons or muons) in the final state for an integrated luminosity of $35.9\,\text{fb}^{-1}$ recorded by the Compact Muon Solenoid (CMS) detector at the LHC using proton-proton collisions at a centre-of-mass energy of 13 TeV. In these accelerator conditions, the reference next-to-leading-order (NLO) cross section for $\text{tZq} \rightarrow \text{Wb}\ell^+\ell^-\text{q}$ (considering only leptonic decays of the Z boson to electrons, muons or tau leptons) is $94.2^{+1.9}_{-1.8}\,(\text{scale}) \pm 2.5\,(\text{PDF})\,\text{fb}$, which includes lepton pairs from off-shell Z bosons with invariant mass $m_{\ell^+\ell^-} > 30\,\text{GeV}$.

Previous analyses had been conducted at the LHC at 8 TeV but it was not until 2017, when the analysis described in this document was performed, that the first evidence for this process was presented, also announced by the ATLAS collaboration in the same year. The evidence of this process along with a measurement of its cross section compatible with the prediction, represents an important test of the standard model. The final state of the process is identical to flavour-changing neutral current tZq production. Flavour-changing neutral currents (FCNC) are a phenomenon which is highly suppressed in the standard model, predicted to occur at rates that are not accessible at the current accelerator conditions. In fact, analyses looking for FCNC tZq signatures are conducted in parallel to tZq standard model searches by the experimental community. The cross section measurement presented is sensitive

to the contribution from FCNC tZq production, were there incompatibilities with the theoretical standard model prediction.

For this analysis, a multivariate classification approach is used to achieve a powerful discrimination between signal-like events and other standard model processes (background). The cross section is extracted from a maximum likelihood fit performed simultaneously on three statistically independent regions for the four different leptonic channels (three electrons, three muons, two electrons and one muon, and one electron and two muons). The first region is defined to be populated mostly by signal events while the other two control regions are defined so that they contain mostly events from the main background processes. The measurement yields a cross section value $\sigma(\mathrm{pp} \to \mathrm{tZq} \to \mathrm{Wb}\ell^+\ell^-\mathrm{q}) = 123^{+33}_{-31}\,(\mathrm{stat})\,^{+29}_{-23}\,(\mathrm{syst})\,\mathrm{fb}$, calculated so that it contains the contribution from tau leptons too. The observed (expected) significance is reported to be 3.7 (3.1) standard deviations.

Resumen

Medida de la sección eficaz de producción de tZq en colisiones pp a $\sqrt{s} = 13$ TeV en CMS

por María del Mar Barrio Luna

Este documento presenta la primera medida del experimento CMS de la sección eficaz de producción de un quark top asociado con un bosón Z y un quark adicional en colisiones protón-protón a una energía en el sistema centro de masas de 13 TeV. La tasa de producción de este proceso en el contexto del modelo estándar es muy pequeña por lo que no había sido observado experimentalmente hasta este momento. Este proceso constituye también un fondo irreducible en muchos análisis y búsquedas de nueva física que se están llevando a cabo en el Gran Colisionador de Hadrones (*Large Hadron Collider* o LHC) del CERN (Ginebra, Suiza).

El proceso se ha estudiado analizando una selección de eventos con tres leptones en el estado final (electrones o muones), de la muestra total de datos recogidos por el detector CMS (*Compact Muon Solenoid*) en el LHC, correspondiente a una luminosidad integrada de $35.9\,\text{fb}^{-1}$.

La sección eficaz de producción de tZq en colisiones protón-protón, a una energía de 13 TeV, calculada en segundo orden en teoría de perturbaciones en Cromodinámica Cuántica (*next-to-leading order*, NLO), multiplicada por la fracción de desintegración del bosón Z en electrones, muones o taus, es $94.2^{+1.9}_{-1.8}\,(\text{escala}) \pm 2.5\,(\text{PDF})\,\text{fb}$. El cálculo incluye la contribución de pares de leptones procedentes de bosones Z fuera de la capa de masas, con una masa invariante de los dos leptones superior a 30 GeV.

Previamente a este análisis se habían llevado a cabo búsquedas de este proceso en el LHC con los datos tomados a una energía de las colisiones de 8 TeV, pero no fue hasta el año 2017 (año en el que se realizó el estudio que se recoge en esta tesis),

con el análisis de una muestra estadística más abundante, tomada a una energía de las colisiones protón-protón significativamente superior, cuando se tuvo por primera vez evidencia experimental de este proceso. En ese mismo año, ATLAS presentaba también, de forma independiente, evidencia del mismo. Estos resultados, y una medida de la sección eficaz compatible con la predicción teórica, confirmaban de nuevo la solidez del modelo estándar. La importancia de este análisis reside también en el hecho de que el estado final es idéntico al que tendría la producción de tZq mediante corrientes neutras con cambio de sabor o, en inglés, *flavour-changing neutral currents* (FCNC). La predicción del modelo estándar dicta que los procesos mediados por este tipo de corrientes ocurrirían a tasas de producción tan bajas que no han podido ser aún observadas en las condiciones actuales del LHC. Análisis similares son llevados a cabo en paralelo para la búsqueda de eventos tZq mediados por corrientes FCNC. La medida de la sección eficaz de tZq es por lo tanto sensible a posibles contribuciones de FCNC-tZq, que podrían inferirse en el caso de que aparecieran incompatibilidades entre la predicción del modelo estándar y los resultados experimentales.

Con el objetivo de optimizar la discriminación entre eventos de señal y de otros procesos del modelo estándar (fondos o *backgrounds*) se han utilizado métodos estadísticos multivariable. La medida de la sección eficaz se ha obtenido mediante un ajuste realizado de manera simultánea sobre tres regiones de control independientes estadísticamente, diseñadas para contener en su mayoría eventos de señal y de los principales fondos, respectivamente. La sección eficaz de producción obtenida del proceso tZq es $\sigma(\mathrm{pp} \to \mathrm{tZq} \to \mathrm{Wb}\ell^+\ell^-\mathrm{q}) = 123^{+33}_{-31}\,(\text{estadístico})\,^{+29}_{-23}\,(\text{sistemático})\,\text{fb}$. La señal medida corresponde a una significancia estadística observada de 3.7 desviaciones estándar, con una significancia esperada de 3.1 desviaciones estándar. La medida obtenida de la tasa de producción de este suceso es compatible con la predicción del modelo estándar.

Agradecimientos

No ha resultado un camino fácil, sin estas personas no habría sido posible. Gracias por hacerlo más fácil.

En primer lugar (no podía ser de otra forma) me gustaría dar las gracias a mi directora, Mara, por las incontables horas que ha dedicado en ayudarme no sólo a nivel académico, sino también a nivel personal. Gracias, Mara, por tu infinita paciencia, empatía y consejos. Gracias porque, sin ti, creo que en muchas ocasiones me habría rendido. Desde el primer momento has sido para mí un referente, un ejemplo a seguir. Gracias por preocuparte en mi formación, por las largas horas mirando código juntas y por "obligarme" a hacer un análisis de principio a fin, desde la creación de mis primeras ntuplas hasta la obtención del resultado final. Siento no haber sabido estar a la altura en la última etapa de la tesis.

En segundo lugar, y aunque quizás ellos no se lo esperen, me gustaría dar las gracias a otras dos personas clave durante estos años: Javier Brochero y Óscar González López. Gracias, Óscar, por enseñarme a cambiar mis "me quiero morir" por "quiero matar a alguien" y así dejar de volcar mi frustración sobre mí misma. Una de las mejores experiencias de la tesis ha sido compartir despacho contigo y con el *entrañable* Alberto. Por haber soportado mis llantos, mis tomaduras de pelo, por no haberme lanzado por la ventana en mil ocasiones cuando habría sido la mejor de las opciones, por haber compartido ventilador, por haber guardado mi caja del tesoro llena de botes de champú a medio terminar y por dejarte cuidar por Wilson la planta de plástico. Si vuelvo a ir por el CERN, guárdame una tarde para que te invite a uno de esos flanes que venden en Cornavin.

Havié, a ti tengo que agradecerte el gran apoyo moral brindado en la etapa de escritura. Por haber venido hasta mi casa para "tirarme de las orejas" cuando yo no sabía de dónde sacar fuerzas para continuar. Gracias por ayudarme a avanzar cada vez que me quedaba bloqueada. Me va a dar un poco de pena no seguir compartiendo despacho contigo y con *el irlandés*, pero espero seguir compartiendo cervezas contigo y con Dermot...¡El overlíiiiiii!

El doctor Alberto Escalante del Valle se merece también su propio párrafo. Gracias, *Albertismo*, por acordarte de mí cada poco y enviarme apoyo cuando más lo necesitaba. Aunque la mayor parte de la tesis la hayamos pasado a lo *tú en Londres (Ginebra) y yo en California (Madrid)*, para mí eres mi hermano mayor del doctorado y no se me ocurre un mejor compañero de despacho (junto con Óscar).

Os echo de menos siempre.

Gracias a Eduísmo, Manueliño, Javismo y Brunete, mis primeros compañeros de despacho. Gracias por no estar ahí mi primera semana de contrato, cuando me di de bruces con una beca-cueva en la que sólo había una planta muerta, una suerte de *cadáver de zambomba*. No sabéis cómo os agradezco la ilusión con la que me hacíais llegar cada lunes al centro. También a Mariano, Manu, Jose, Jorge, Bárbara, Miguel Ángel y Sara. A mis compañeras de despacho: Diana, Chiara y Ana, por aguantarme y por el esfuerzo que le ponen a la decoración del despacho en navidad. A quienes vinieron después a animar el cotarro: Mabel, Carmen, Martín, Edgar, Molero, Iker. Y, por supuesto, a mis *hijismos CMSismos*: Sergio, Irene y Adrián. Espero que os vaya genial con vuestras respectivas tesis y que no os desesperéis demasiado.

Gracias a CRAB, a ROOT y a las miles de millones de versiones de CMSSW por haber convertido algunos de mis días en una auténtica pesadilla.

Al resto de personas del departamento: María Cepeda, Begoña, Isabel, Marcos, Juan, Nica, Pablo, Chema, Juan Pablo, Cruz, Nacho, Miguel Cárdenas, Ricardo e Ignacio. A Antonio, Juanjo y Calonge por *salvarme* cada vez que la liaba. A Cris, por el tiempo dedicado durante el análisis en el grupo de DTs.

A mis compañeras de *tortura*: Andrea, Jose, Víctor y Julia. Por estar ahí siempre, aún en la distancia. Por el apoyo, por hacerme sentir acompañada, por escucharme y animarme. Sois de mis personas favoritas y creo que lo sabéis. No voy a extenderme demasiado: sobran las palabras, pero me muero por daros un abrazo fuerte. Gracias por formar parte de mi vida.

A Nuria, por existir (junto a mí). A Mimar, mi mar: madre, hermana, amiga, confidente. Gracias por enseñarme que nacer mujer te hace más fuerte, por escucharme, rebatirme, ayudarme a tirar *p'alante*. Por enseñarme a hacer, y hacerme, revolución.

A toda mi familia, por su constante apoyo y ayuda incondicional. Por los cuidados. A mi hermano, Jorge, porque pensar en ti me da fuerzas para cualquier cosa desde el 93. A mi madre, Mercedes, por estar ahí siempre, por animarme, cuidarme, escucharme; y ti, papá, por ayudarme siempre en cualquier aspecto. A veces lo pienso y nunca sé qué sería de mi sin vuestra ayuda y sacrificio. Sabéis que os quiero más de lo que puedo expresar con un par de frases. A mis tíos y primos: Paco, Maribel, Dani, Carmela, Juanjo, Paco y María.

A mis amigas: mis *atunes*, *girasolas*, mis *chicas de los ochenta*, mi gente de Barcelona, a las que llegaron a mi vida a través del baile para quedarse, a *las que llevan rulos*, a Anita, Andrea, Ceci, Roger, mis compañeras de la carrera (Mo, Marta, Emi, Juan, Josete, el Martin y Miguel) y a todas aquellas que seguramente me esté olvidando porque, en fin, soy un poco desastre y vosotras sois muchas.

Gracias.

Si aún alguien sigue leyendo, este párrafo lo escribo en parte gracias a Caridad, quien llegó a mi vida en un momento en el que estaba al borde del colapso. Gracias, Caridad, por ayudarme a salir del pozo negro en el que anduve metida y del que no sabía salir (igual ni siquiera te mostré cuán negro y profundo era). Fuiste como un hada madrina y no miento cuando digo que ahora mismo se me ha saltado alguna que otra lágrima y ando medio desquiciada porque no tengo tu paquete de pañuelos de papel al lado. Una de las copias de esta tesis es para ti, porque no sé si habría podido acabarla si en parte no hubiera contado contigo.

También me gustaría dar las gracias a las personas que me han ayudado a llegar aquí: a la gente de LHCb (Lluís Garrido, Carla Marín y Vicente Rives), a Juan Terrón, Claudia Glasman y la gente de DESY. A Anette Knebe por sacarme siempre del fango con las gestiones de la universidad.

Thanks to everyone that helped me at CERN. To my flatmate, Claudia. To Nicolas Tonon and Jéremy Andrea for their dedication, hard work and patience while working with me in the tZq analysis. To Sara Fiorendi, Benjamin Radburn-Smith and Martijn Mulders for their help. I want to thank you specially, Emma, for hugging me and helping me relax in the middle of the panic attack I had when presenting my first poster. Thanks, Fabio, for the good times at Evora and Madrid, I am glad you entered our lives to stay longer.

Contents

List of Figures

List of Tables

List of Abbreviations

2HDM	**2H**iggs **D**oublet **M**odel
ALICE	**A** **L**arge **I**on **C**ollider **E**xperiment
AOD	**A**nalysis **O**bject **D**ata
ATLAS	**A** **T**oroidal **L**HC **A**pparatu**S**
AVR	**A**daptive **V**ertex **R**econstruction
BDT	**B**oosted **D**ecision **T**ree
BSM	**B**eyond **S**tandard **M**odel
CB	**C**ut **B**ased
CERN	**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire
CKM	**C**abbibo-**K**obayashi-**M**askawa
CMB	**C**osmic **M**icrowave **B**ackground
CMS	**C**ompact **M**uon **S**olenoid
CMSSW	**CMS** **S**oft**W**are
CSC	**C**athode **S**trip **C**hamber
CSV	**C**ombined **S**econdary **V**ertex
DT	**D**rift **T**ube
DY	**D**rell **Y**an
EB	**E**lectromagnetic calorimeter **B**arrel
ECAL	**E**lectromagnetic **CAL**orimeter
EE	**E**lectromagnetic calorimeter **E**ndcap
EFT	**E**ffective **F**ield **T**heory
EW	**E**lectro**W**eak
EWSB	**E**lectro**W**eak **S**ymmetry **B**reaking
FCNC	**F**lavour-**C**hanging **N**eutral **C**urrent
FS	**F**lavour **S**cheme
GCT	**G**lobal **C**alorimeter **T**rigger
GIM	**G**lashow-**I**liopoulos-**M**aiani
GSF	**G**aussian **S**um **F**ilter
GT	**G**lobal **T**rigger
HB	**H**adronic calorimeter **B**arrel
HCAL	**H**adronic **CAL**orimeter
HE	**H**adronic calorimeter **E**ndcap
HF	**H**adronic **F**orward calorimeter
HLT	**H**igh **L**evel **T**rigger
HO	**H**adronic **O**uter calorimeter
IP	**I**mpact **P**arameter
IVF	**I**nclusive **V**ertex **F**inder
JEC	**J**et **E**nergy **C**orrections

JER	**J**et **E**nergy **R**esolution
JES	**J**et **E**nergy **S**cale
JP	**J**et **P**robability
L1	**L**evel **1** trigger
LEP	**L**arge **E**lectron **P**ositron collider
LHC	**L**arge **H**adron **C**ollider
LO	**L**eading **O**rder
MC	**M**onte**C**arlo
MEM	**M**atrix **E**lement **M**ethod
MET	**M**issing **E**nergy (in the) **T**ransverse plane
MSSM	**M**inimal **S**upersymmetric **S**tandard **M**odel
MVA	**M**ulti**V**ariate **A**nalysis
NLO	**N**ext to **L**eading **O**rder
NNLO	**N**ext to **N**ext to **L**eading **O**rder
NPL	**N**on**P**rompt **L**epton
OSSF	**O**pposite-**S**ign **S**ame-**F**lavour
PDF	**P**arton **D**istribution **F**unction
PDG	**P**article **D**ata **G**roup
PF	**P**article **F**low
PS	**P**arton **S**hower & **P**roton **S**ynchrotron
PSB	**P**roton **S**ynchrotron **B**ooster
PU	**P**ile **U**p
PV	**P**rimary **V**ertex
QCD	**Q**uantum **C**hromo**D**ynamics
QED	**Q**uantum **E**lectro**D**ynamics
QFT	**Q**uantum **F**ield **T**heory
RCT	**R**egional **C**alorimeter **T**rigger
RPC	**R**esistive **P**late **C**hamber
SC	**S**uper**C**luster
SF	**S**cale **F**actor
SL	**S**uper**L**ayer
SM	**S**tandard **M**odel
SPS	**S**uper **P**roton **S**ynchrotron
SSB	**S**pontaneous **S**ymmetry **B**reaking
SV	**S**econdary **V**ertex
TE	**T**racker **E**ndcap
TGC	**T**riple **G**auge **C**oupling
TIB	**T**racker **I**nner **B**arrel
TID	**T**racker **I**nner **D**isk
TOB	**T**racker **O**uter **B**arrel
WP	**W**orking **P**oint

Introduction

The standard model of particle physics, or simply *the standard model* (SM) is the most rigorous theory of particle physics up to date, incredibly precise and accurate in its predictions. The model was developed during the past mid-century, its current formulation being finalized in the mid-1970s upon experimental confirmation of the existence of quarks. Within its framework, all visible matter in the universe is described by spin-$\frac{1}{2}$ fermions grouped into *quarks* and *leptons*. In its current formulation, the standard model contains three fermion generations, with increasing masses. Matter particles interact via the exchange of gauge *bosons*, a different type of particles directly connected with the fundamental interactions of nature. A very basic scheme of the particle content of the standard model is presented in figure 1, and a brief introduction to the theory can be found in chapter 1.

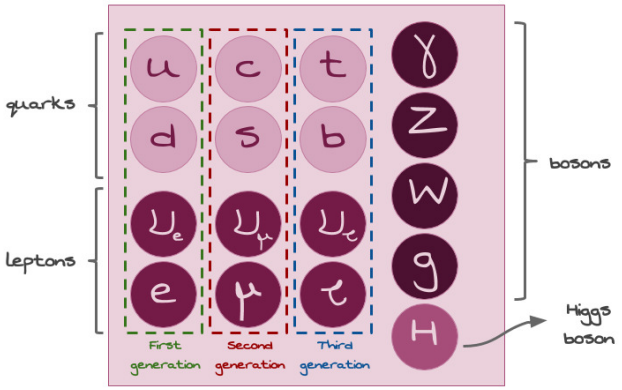


FIGURE 1: Standard model particles.

Originally, before the formulation of the standard model as we know it today, the quark model developed by Gell-Man and Zweig in 1964 to explain the experimental observations in accelerator and cosmic ray experiments contained only three quarks. These were the up and down quarks (which formed protons and neutrons) and the strange quark, which were enough to accurately describe the different properties of the observed hadrons. However, the discovery of the J/ψ meson in 1974, interpreted as a bound state of a new quark and its antiparticle led to the inclusion of the charm in the quark model, even though it had been theoretically predicted to explain the decay properties of charged and neutral K mesons. Up to this point, only two

generations of quarks and leptons were needed. Later on, in the mid-seventies, the first particles from the third generation of leptons and quarks were discovered. The tau lepton was first observed in 1975 [3] and the bottom quark, predicted to have a mass of about 5 GeV and electric charge of $\frac{-1}{3}$ was inferred from the discovery of the Υ-meson at Fermilab, which resulted to be a $b\bar{b}$ bound state. The remaining particles of the third generations, the top quark and the τ neutrino were not discovered until 1995 and 2000, respectively.

The search for the top quark started in the late seventies. Physicists had known that the top must exist since 1977, when its partner, the bottom quark, was discovered. However, top quark search resulted a long and arduous process as it turned out to be much more massive than originally expected: about 200 times larger than the mass of the proton and 40 times higher than the mass of the next-lightest quark. This search culminated in 1995 with the observation of the production of top quark-antiquark pairs via strong interactions by the CDF [4] and DØ[5, 6] collaborations at the Tevatron collider at Fermilab using proton-antiproton collision data.

Another milestone in top quark research was the observation of single top quark production via electroweak interactions in 2009 by the CDF [7] and DØ[8] collaborations at the Tevatron collider. The identification of top quarks in the electroweak single top channel is much more difficult than in the QCD $t\bar{t}$ channel, due to a less distinctive signature and significantly larger backgrounds.

With the Tevatron having made the first precious thousands top quarks, the subsequent efforts in the study of this quark and its properties were carried mostly at the Large Hadron Collider (LHC), the largest and most powerful of modern high energy physics experiments up to date. It was built by the *European Organization for Nuclear Research* (CERN) between 1998 and 2008 in a 27 km-tunnel in the frontier between France and Switzerland. The LHC collides protons at an energy only achieved in the first 10^{-12} s of the universe, close to the speed of light. The construction of the LHC made it possible to confirm the existence of the Higgs boson in 2012, the last missing piece of the standard model to be discovered [9, 10]. The results of the collisions taking place at the LHC are recorded by several detectors placed along the LHC ring (ALICE, ATLAS, CMS and LHCb, among others). The LHC is itself a genuine top quark factory, millions of them being produced each year of running. The high production rate opens the way to precision studies of top quark properties, unattainable at previous machines. From the theoretical point of view, a very intense effort for improving theory calculations has been made, achieving a level of precision such that it can be used to spot new physics in subtle deviations from the calculations. In particular, the analysis presented in this thesis has been performed using data collected by the *Compact Muon Solenoid* (CMS) experiment, located in an underground cavern at Cessy (France). A short introduction to CMS and the LHC can be found in chapter 4.

This work presents a measurement of the production cross section of a single top in association with a Z boson and an additional quark (tZq), a suppressed

electroweak process within the standard model context. This process is sensitive to the tZ coupling, and an enlargement in the coupling strength could increase the tZq cross section beyond the standard model prediction. Therefore, the choice of this study does not only provide a test for the current model, but it could also yield hints of new physics.

The analysis is performed considering only leptonic decays of both the Z boson and the top quark, as this channel leaves a cleaner signature in the detector. The measurement has been performed using a likelihood fit in three statistically independent regions simultaneously. The first region, named *signal region*, is designed so as to be populated mostly by signal events, while the other two (*control*) regions are populated mostly by events from the main background processes. By doing this, background contributions are better constrained in the analysis. The LHC provided a first opportunity to study this rare standard model process, as the previous generation of accelerators could not access this process due to low beam colliding energies. Previous studies were conducted during the first data taking period at the LHC (*Run I*) at 8 TeV, when CMS reported a signal with a significance of 2.4 standard deviations [11].

A brief introduction to the basic theory of the standard model is given in chapter 1. A deeper insight to top quark physics and, particularly, tZq production, are given in chapters 2 and 3, respectively. Chapter 4 provides the reader with a description of the Large Hadron Collider and the CMS detector. A detailed explanation of the different subdetectors that compose the CMS apparatus is also presented in this chapter.

Chapter 5 describes how the different physical objects are reconstructed by the CMS experiment. The signature studied in this analysis contains electrons, muons, missing transverse energy and jets coming from the hadronization of b quarks, and all these objects are reconstructed using the particle flow (PF) algorithm, which is also described in this chapter.

The three subsequent chapters discuss the main analysis. Chapter 6 describes the data and simulation datasets, along with a description of the object and event selection, and the three different kinematic regions considered in the analysis. This chapter also describes the data-driven estimation of one of the main sources of background in the analysis. Chapter 7 describes the multivariate analysis tools used to optimize signal-to-background separation, and a description of the shape analysis is presented.

Postfit results are presented in chapter 8. Here, the measurement of the tZq production cross section is presented, along with some other postfit results, and a discussion on how the different sources of systematics affect this result. Finally, conclusions are presented in chapter 9.

Chapter 1

The standard model

An overview of the standard model of particle physics is presented in this chapter. First, its particle content is reviewed, along with its description as a quantum field theory; a brief overview of quantum chromodynamics and the electroweak theory is also given. Some important aspects of electroweak symmetry breaking and flavour physics are also described. At the end of the chapter we review the limitations of the model, and the need to search for signatures of new physics.

The standard model of particle physics, based on quantum field theory, is the framework that currently provides the best description of elementary particles and three of the four fundamental interactions (electromagnetism, weak interactions, strong force and gravity). Developed in the early 1970s, it has successfully explained almost all experimental results in particle physics and precisely predicted a wide variety of phenomena. The validity of the SM is constantly being tested with high precision at high energy physics experiments by comparing theoretical predictions to experimental results.

A lot of bibliography can be found elsewhere regarding the building blocks, mathematical description and phenomenology of the SM. Detailed introductions to its theoretical formulation may be found, for instance, in [12] or [13], should the reader have special interest therein.

This chapter provides a brief introduction to the current model, addressing those topics which are most related to the analysis. Section 1.1 describes the building blocks of the SM, section 1.2 describes the formalism of the model as a quantum field theory providing a brief introduction to the theory of strong and electroweak interactions as well as the mechanism of spontaneous electroweak symmetry breaking. This section ends with a quick look at flavour physics and how flavour changing neutral interactions might take place within the context of the SM. The final section (1.3) of this chapter profiles some of the known limitations of the model, according to the experimental results.

1.1 Particles in the standard model

Elementary particles are objects for which experimental results have not yet revealed any sign of internal structure. They are grouped according to their spin in two different categories: *fermions* or spin-$\frac{1}{2}$ matter particles, and the force carriers, referred to as *bosons*, that are the spin-1 vector bosons and the spin-0 (scalar) Higgs boson.

Fermions: matter particles

Fermions are further categorized into *leptons* and *quarks*, where only the latter take part on strong interactions. Quarks and leptons are paired in *isospin* partners[1] forming three different *generations*, with increasing masses. The reason behind the number of quark and lepton generations remains still unknown. Among the leptons, there are three types of charged leptons, namely electrons (e), muons (μ) and taus (τ), and their corresponding neutrinos (ν_e, ν_μ and ν_τ, respectively). Quarks are classified depending on their isospin into up-type ($Q = +2/3$) and down-type quarks ($Q = -1/3$). Isospin (I_3) and electric charge (Q) are related by

$$Q = I_3 + \frac{1}{2}Y \tag{1.1}$$

where Y is the *hypercharge*.

Up-type quarks include the up (u), charm (c) and top (t) quarks, whereas down (d), strange (s) and bottom (b) are down-type quarks. Being the most massive, fermions of second and third generations decay into first generation fermions. These lightest quarks form ordinary matter (proton and neutron compositions are uud and udd, respectively) . An overview of the different fermion generation particles, along with some of their main properties is provided in table 1.1.

Bosons: force carriers or mediators

The second category of SM particles, the bosons, are connected to the fundamental interaction fields. The SM theory is invariant under local transformations of the gauge group $SU(3)_C \times SU(2)_L \times U(1)_\gamma$. The three groups correspond roughly to the three interactions described by the model: the $SU(3)_C$ gauge field corresponds to the *strong* interaction, and acts only on particles carrying *colour charge* (quarks) by the exchange of eight kind of massless gluons (g). The $SU(2) \times U(1)$ group corresponds to the electroweak interaction. Before electroweak symmetry breaking,

[1] Quarks can be grouped into doublets with opposite isospin values, $\pm\frac{1}{2}$. Quarks with positive isospin will transform weakly to their negative isospin partner, and vice versa.

	generation	name	charge	mass	interactions
Quarks	1^{st} gen.	u	$+\frac{2}{3}$	2.2 MeV	Weak, EM, Strong
		d	$-\frac{1}{3}$	4.7 MeV	
	2^{nd} gen.	c	$+\frac{2}{3}$	1.28 GeV	
		s	$-\frac{1}{3}$	0.95 GeV	
	3^{rd} gen.	t	$+\frac{2}{3}$	173 GeV	
		b	$-\frac{1}{3}$	4.18 GeV	
Leptons	1^{st} gen.	e	-1	0.511 MeV	Weak, EM (except neutrinos)
		ν_e	0	< 2 eV	
	2^{nd} gen.	μ	-1	106 MeV	
		ν_μ	0	< 0.19 MeV	
	2^{nd} gen.	τ	-1	1.78 GeV	
		ν_τ	0	< 18.2 MeV	

TABLE 1.1: Classification of the three generations of spin-$\frac{1}{2}$ SM fermions (quarks and leptons). The values of the different masses are taken from the latest review by the Particle Data Group [2].

Name	Mass	Interaction	Gauge group
Photon	0	Electromagnetic	U(1)
Z boson	91.19 GeV	Weak interaction	SU(2)
W boson	80.38 GeV		
Gluon (g)	0	Strong interaction	SU(3)
Higgs boson	125.2 GeV	Yukawa interaction	SU(2) $\otimes$ U(1)

TABLE 1.2: The gauge bosons of the SM and their associated interactions. Masses are taken from the latest Particle Data Group review to date [2].

$SU(2)$ is mediated by three weak isospin, massless bosons and $U(1)$ by a weak hypercharge massless boson. After electroweak symmetry breaking, these gauge bosons are recombined into the massive carriers of the weak force, the charged $W^{\pm}$ and the neutral Z bosons, along with the massless photon γ, associated to the electromagnetic interaction. With the exception of the Higgs boson, all other bosons have spin 1. The Higgs boson is the only scalar (spin 0) fundamental particle in the SM, introduced in the theory as a solution that would break the electroweak symmetry in a spontaneous way. An overview of this mechanism is given in section 1.2.2. The SM bosons, along with some of their properties, are reviewed in table 1.2.

1.2 The standard model as a quantum field theory

In the framework of a quantum field theory, particles are described as excitation modes of quantized fields that are operators acting on the quantum mechanical Hilbert space. Within the SM, each type of particle is described by a specific type of field:

- spin-0 particles, described by scalar fields $\phi(x)$
- spin-1 particles, described by vector fields $A_\mu(x)$
- spin-$\frac{1}{2}$ particles, described by spinor fields $\psi(x)$

The dynamics of the fields is described by the corresponding Lagrangian density $\mathcal{L}(\phi_i, \partial_\mu \phi_i)$, which is a function of the fields ϕ_i and their space-time derivatives $\partial_\mu \phi_i$. Interactions in the SM are connected to local gauge transformations under which the Lagrangian remains invariant.

The model itself is a renormalizable quantum field theory based on a $SU(3)_C \times SU(2)_L \times U(1)_\gamma$ local gauge symmetry. The SM Lagrangian, $\mathcal{L}_{\mathcal{SM}}$ is invariant under this symmetry group, and the fundamental interactions are contained in two main pieces: the strong sector $\mathcal{L}_{\mathcal{QCD}}$ and the electroweak sector, $\mathcal{L}_{\mathcal{EW}}$. Thus, within the context of the SM, three of the four fundamental interactions (strong, weak and electromagnetism) are described with two gauge theories:

- The theory of electroweak interactions, that unifies the electromagnetic (QED) and weak interactions.
- Quantum chromodynamics (QCD), or the theory of strong interactions.

Local gauge invariance under U(1), SU(3) and $\mathrm{SU(2)}_L \otimes \mathrm{U(1)}_Y$ leads to the $\mathcal{L}_{\mathcal{QED}}$, $\mathcal{L}_{\mathcal{QCD}}$ and $\mathcal{L}_{\mathcal{EW}}$ Lagrangians, respectively. A brief introduction to these gauge theories will be given in the following sections.

1.2.1 The strong interaction

Quantum chromodynamics is the theory that accounts for strong interactions, described by an $SU(3)_C$ gauge theory. Eight types of gluons that represent the massless spin-1 gauge bosons of the group are responsible for mediating strong interactions. The conserved charge of the group is called *colour*, and it can take on three values, which are equal in strength, namely: red (R), green (G) and blue (B).

The only fermions sensitive to the strong interaction are quarks, which were proposed by Gell-Man and Zweig in 1964 and were first observed in deep inelastic scattering experiments at the Stanford Linear Accelerator Center (SLAC) four years after their prediction. As gluons carry themselves colour charge, they can either interact with quarks or with each other.

The running coupling constant α_S, related to the intensity of the strong interaction, depends on the energy scale of the interaction, and is given by

$$\alpha_s(Q^2) \approx \frac{1}{ln(Q^2/\Delta)} \tag{1.2}$$

where Q is the momentum transfer involved in the process and Δ is the non-perturbative scale of QCD.

This effect leads to two important physical implications of the strong interaction:

- *Quark confinement within hadrons* As the coupling constant depends on the energy scale of the strong interaction, its strength increases with distance from the charge. This implies that when a quark-antiquark pair begins to separate, the colour field generated by the exchanged gluons will increase its intensity to a point where the creation of a new quark-antiquark pair becomes more energetically preferable rather than increasing further the interaction strength. This explains why quarks cannot be found isolated but rather forming colour neutral states called *hadrons*, which are either formed by quark-antiquark pairs (called *mesons*, which are $R\bar{R}$, $B\bar{B}$ and $G\bar{G}$ colour states) or by groups of three RGB ($\bar{R}\bar{G}\bar{B}$) quarks (named *baryons*). Due to their short lifetime, top quarks, object of study of this thesis, do not form hadrons (more details on the next chapter).
- *Asymptotic freedom* The interaction strength decreases with increasing energy. This causes the quarks inside hadrons to behave more or less as free particles, when probed at large enough energies, such as those reached at the LHC.

1.2.2 The Electroweak Theory

The electromagnetic interaction, which acts on all charged particles and is mediated by photons, was the first interaction to be described via a quantum field theory known as *quantum electrodynamics* (QED), developed by Tomonaga, Dyson, Schwinger and Feynman. QED is subjected to a local invariance under the gauge group $U(1)$ and successfully pictures the interaction of electrically charged fermions with photons.

In 1934, Fermi proposed his theory to describe β decays ($n \rightarrow pe^{-}\bar{\nu}$), which required the introduction of a new *weak* interaction. This interaction is mediated

by the $W^{\pm}$ and Z^0 bosons, and acts between quarks and leptons, allowing for flavour-changing transitions of fermions.

Weak interactions, however, were not fully depicted until 1960, when Glashow [14], Weinberg [15], Ward and Salam developed the theory of electroweak interactions, in which electromagnetic and weak forces are unified.

A special feature of the weak interactions is that parity is not conserved, a phenomenon that was observed by experimentalist Chien-Shiung Wu in 1957 [16], during her study of nuclear β ${}^{60}_{27}\mathrm{Co} \rightarrow {}^{60}_{28}\mathrm{Ni} + e\overline{\nu}_e\gamma\gamma$ decays by analyzing the direction of the escaping electron with respect to the polarization of the cobalt probe through an external magnetic field. Parity violation had been previously predicted by Yang and Lee [17]. In the electroweak theory, fermion fields are decomposed into chiral eigenstates to account for parity violation, chirality being a Lorentz invariant quantity corresponding to the eigenvalues of the operator $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$, ± 1. Eigenvectors associated to the eigenvalues -1 and +1 are said to have *left-handed* and *right-handed* chirality, respectively. Thus, any Dirac fermion field can be decomposed using the projections

$$\psi_L = P_L\psi = \frac{1}{2}\left(1 - \gamma_5\right)\psi \qquad \psi_R = P_R\psi = \frac{1}{2}\left(1 + \gamma_5\right)\psi \tag{1.3}$$

where ψ_L and ψ_R are referred to as *left-handed* and *right-handed* fermion states, respectively. Unlike QED and QCD, weak interactions act differently on particles with opposite chiralities. Left-handed fermions transform as $SU(2)_L \otimes U(1)_Y$ doublets, whereas right-handed fermions transform as $U(1)_Y$ singlets,

$$\psi_L = \left\{ \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \begin{pmatrix} e_L \\ \nu_{e,L} \end{pmatrix}, \begin{pmatrix} \mu_L \\ \nu_{\mu,L} \end{pmatrix}, \begin{pmatrix} \tau_L \\ \nu_{\tau,L} \end{pmatrix} \right\} \tag{1.4}$$

$$\psi_R = \left\{u_R, d_R, c_R, s_R, b_R, t_R, e_R, \mu_R, \tau_R, \nu^{?}_{e,R}, \nu^{?}_{\mu,R}, \nu^{?}_{\tau,R}\right\} \tag{1.5}$$

Within the context of the SM, right-handed neutrinos do not participate in any of the described interactions. However, the observation of neutrino oscillations suggests that such particles may exist, even though they do not take part within the SM framework.

The addition of a fermionic mass term in the Lagrangian describing electroweak interactions would mix left-handed and right-handed chiralities, which is not possible since the two type of fermions have different transformation properties. Similarly, a mass term for the gauge fields would violate local gauge invariance and is also forbidden. A mechanism for generating non-zero masses while preserving the consistency of the theory at high energies therefore needs to be introduced: the Higgs mechanism for spontaneous electroweak symmetry breaking.

Spontaneous electroweak symmetry breaking and the generation of particle masses

All fields produced by imposing gauge invariance are strictly massless. A mass term for a boson field is not invariant under $SU(2) \otimes U(1)$. However, the vector bosons $W^{\pm}$ and Z have a consistently non-zero mass, which gives weak interactions their short-range characteristics. The W bosons were discovered in 1983 at CERN by the UA1 [18] and UA2 [19] collaborations, with an estimated mass of $M_W = 80.379 \pm 0.012\text{GeV}$. The Z boson was discovered a few months after the W bosons in the UA1 [20] and UA2 [21] experiments and its mass is currently estimated at $M_Z = 91.1876 \pm 0.0021\text{GeV}$ (these values are taken from the latest PDG review [2]). A simple solution to this puzzling situation is to introduce a scalar field to the theory, spontaneously break the original symmetry and generate masses for the different particles.

Spontaneous electroweak symmetry breaking is governed by the Higgs mechanism. In this mechanism, an additional complex scalar field (the Higgs field), doublet of $SU(2)_L \otimes U(1)_Y$, is introduced:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.6}$$

where the superscripts indicate the electric charge of the field. ϕ^+ and ϕ^0 form an isospin doublet with quantum numbers $I_3 = \pm 1/2$ and $Y = 1/2$, respectively; this choice is imposed in order to keep the photon massless. ϕ^+ annihilates positively charged particles, creating antiparticles with negative charge; ϕ^0 annihilates neutral particles to create neutral antiparticles. It can also be written in terms of four real scalar fields (ϕ_1,ϕ_2,ϕ_3 and ϕ_4).

Renormalizability and invariance under $SU(2)_L \otimes U(1)_Y$ require the Higgs potential $V(\phi)$ to take on the form

$$V(\phi) = -\mu^2 \Phi^\dagger \Phi + \lambda(\phi^\dagger \Phi)^2 \tag{1.7}$$

where μ and λ are, respectively, the mass and self-interaction coupling constants. In order to preserve vacuum stability, it is required that $\lambda > 0$. If $\mu^2 < 0$, the minimum of the potential $V(\phi)$ is found at

$$\Phi^\dagger \Phi = -\frac{\mu^2}{\sqrt{2\lambda}} \equiv \frac{v^2}{2} \tag{1.8}$$

and the scalar field has a non-vanishing *vacuum expectation value* (VEV), $\langle \Phi \rangle = v/\sqrt{2} \neq 0$.

As the previous condition is fulfilled by an infinity of possible ground states, the choice of a particular one spontaneously breaks the symmetry. This choice is

arbitrary, and :

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.9}$$

prevents electromagnetism and electric charge conservation from being broken by the scalar VEV, and $Q(\Phi) = 0$. The introduction of the Higgs potential allows for the direct generation of mass terms for the electroweak gauge bosons ($W^{\pm}$, Z^0), except for the photon, which remains massless. Three of the real scalar fields in the complex doublet in (1.6) are absorbed to generate mass to the heavy gauge bosons, while the fourth one emerges as a new massive scalar boson, the Higgs boson.

Once the Higgs field acquires a vacuum expectation value, the gauge bosons gain mass through interactions with the Higgs field. As there is nothing preventing the Higgs doublet to couple to fermion fields, quarks and charged leptons also gain mass when interacting with the Higgs boson through the *Yukawa couplings*, λ_f. These couplings (and as a consequence, particle masses) are not predicted by the theory and must be measured experimentally. The mass of the fermions is obtained from the following expression

$$m_f = \lambda_f \frac{v}{\sqrt{2}} \tag{1.10}$$

The value of the Higgs boson mass is related to the parameters of the Higgs potential

$$m_H = v \cdot \sqrt{2\lambda} \tag{1.11}$$

where the vacuum expectation value v of the Higgs field is related to the mass of the vector bosons by

$$m_W = \frac{g}{2} \cdot v \qquad m_Z = \frac{\sqrt{g^2 + g'^2}}{2} \cdot v \tag{1.12}$$

inferred from their respective experimental measurements. In the last expressions, g and g' are, respectively, the coupling constants of the $SU(2)_L$ and $U(1)_Y$ groups of the electroweak theory. The Higgs self coupling λ can be only determined from direct Higgs boson detection.

This solution was confirmed by the experimentalists at the LHC in 2012, when a massive scalar boson compatible with the SM Higgs was detected both by the ATLAS [10] and CMS [9] collaborations. This result also closed the search for the SM particles, since all of them had already been detected experimentally, the Higgs boson remaining the only missing piece until then.

Flavour in the standard model

Flavour physics covers the study of the different types of quarks (or *flavours*), their spectrum and the transmutations among them. Within the SM, interactions

mediated by the charged $W^{\pm}$ bosons are the only source of flavour- and generation-changing interactions. The flavour quantum number is nonetheless conserved in strong and electromagnetic interactions. Neutral current weak interactions (mediated by the neutral Z boson) are also flavour-conserving. Therefore, flavour-changing neutral processes do not occur in the SM at tree level, but they can be induced by loop processes.

At the heart of flavour physics lies the *Cabibbo-Kobayashi-Maskawa* matrix [22], [23], also known as CKM matrix, or V_{CKM}. The CKM matrix describes quark flavour transitions within the SM and has the form

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.13}$$

where each $V_{q_1 q_2}$ element is proportional to the coupling strength between each two pair of quarks (or flavours) q_1 and q_2. These elements must satisfy the unitarity condition of the CKM matrix:

$$V_{CKM} V_{CKM}^{\dagger} = V_{CKM}^{\dagger} V_{CKM} = 1 \tag{1.14}$$

so that, for instance, $\sum_q |V_{tq}|^2 = 1$. The quark mixing parameters are experimentally well tested and constrained from global fits to many measurements in the flavour sector of the SM but only under the assumption of three quark generations. The most precise experimental values of the elements of the CKM matrix are [2]:

$$V_{CKM} = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix} \tag{1.15}$$

The SM Lagrangian of electroweak interactions between fermions and the gauge bosons can be expressed via three terms:

$$\mathcal{L}^{matter}_{SU(2)\times U(1)} = -e J^{\mu}_{em} A_{\mu} - \frac{g}{2}\left(J^{\mu}_{CC} W^{+}_{\mu} + J^{\mu\dagger}_{CC} W^{-}_{\mu}\right) - \frac{g}{2\cos\theta_W} J^{\mu}_{NC} Z_{\mu} \tag{1.16}$$

where each of them contemplates the contribution from the electromagnetic current J^{μ}_{em}, the weak charged current J^{μ}_{CC} and the weak neutral current J^{μ}_{NC}, respectively.

The second term in the previous expression describes flavour transitions of quarks and leptons via *charged currents*:

$$J^{\mu}_{CC} = \bar{\mathcal{U}}_L \gamma^{\mu} \mathcal{D}_L = (\bar{u}, \bar{c}, \bar{t})_L \, \gamma^{\mu} V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \tag{1.17}$$

which couples up-type antiquarks to down-type quarks. Processes mediated by charged currents therefore violate flavour conservation.

The last term describes quark and lepton interactions via *neutral currents*, which are of the form:

$$\begin{aligned} J^{\mu}_{NC} = & \quad \bar{\mathcal{U}}_L \gamma^{\mu} \left(1 - \frac{4}{3}\sin^2\theta_W\right) \mathcal{U}_L \\ & - \quad \bar{\mathcal{U}}_R \gamma^{\mu} \frac{4}{3}\sin^2\theta_W \mathcal{U}_R \\ & - \quad \bar{\mathcal{D}}_L \gamma^{\mu} \left(1 - \frac{2}{3}\sin^2\theta_W\right) \mathcal{D}_L \\ & + \quad \bar{\mathcal{D}}_R \gamma^{\mu} \frac{2}{3}\sin^2\theta_W \mathcal{D}_R \end{aligned} \tag{1.18}$$

These neutral currents couple quarks with same-flavoured quarks only: flavour and charge are conserved. Flavour-violating charge-conserving transitions do not happen at tree level in the SM. However, they can take place at loop level via flavour-changing vertices, as seen in figure 1.1.

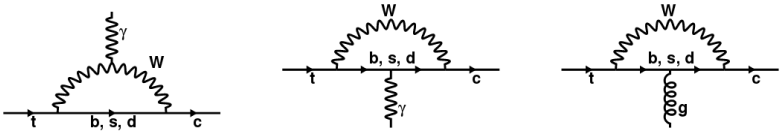


FIGURE 1.1: Examples of flavour-changing neutral current SM loop diagrams for the process $t \to c\gamma$ and $t \to cg$.

These contributions are proportional to the splitting between the quark masses and, in the case of the top quark, $B(t \to Bq)$ where $B = \gamma, Z, g, H$ and $q = u, c$ are predicted to be of order 10^{-12} or smaller. An increase in the value of the $t \to c$ or $t \to u$ rates could be an indication of the existence of new flavour-violating neutral current interactions that are absent in the SM. tZq production, subject of the current analysis, is specially sensitive to these kind of processes. FCNC interactions of the top quark are further described in section 3.2.3.

1.3 Limitations of the standard model

Even though the SM represents both a very elegant and successful theoretical framework, that has been finely tested throughout the last decades in the different particle physics experiments, it leaves several fundamental questions unsolved, which will be briefly outlined in the current section.

- **Gravity** As already mentioned, the SM does not describe gravity within its framework. The hypothetical mediator of gravity, the graviton, would have to

be a spin-2 boson, leading to non-renormalizable divergences. However, at the energy scales reached by particle colliders, gravity is negligible compared to the strong and weak interactions, and its effects would pass unnoticed. General relativity provides with a fine description of gravitation at astrophysical scales and, although some attempts have been made to unify gravity with the rest of the SM interactions at quantum level, no success in this topic has taken place yet.

- **Dark matter and dark energy** Different astrophysical and cosmological observations show effects that the SM cannot seem to explain. The observation of the rotation speeds of galaxies, mapping of matter distributions, and acoustic oscillations in the cosmic microwave background (CMB) suggest that there exists an unidentified type of matter, the so-called *dark matter*. According to the observations, dark matter should be made of stable, neutral particles. The only SM particles satisfying this description would be the neutrinos, but this has not been proven yet. Some theories beyond the SM (BSM) also provide candidate particles for dark matter. Another important building block of the universe and a crucial ingredient to describe its expansion, *dark energy*, is also missing in the description provided by the SM.

- **Nature of EW symmetry breaking and hierarchy problem** Another open question is the nature of electroweak symmetry breaking, why is there such difference between the intensity of the different fundamental forces (in particular, why the weak force is orders of magnitude weaker than gravity) and why the Higgs boson has its particular mass value. Even though some models, like supersymmetry, propose solutions to the problem of the Higgs mass, the SM does not provide a suitable answer.

- **Matter-antimatter asymmetry** The SM does not provide an answer to the asymmetry between matter and antimatter observed in the universe. CP-symmetry states that the laws of physics should be the same if a particle is interchanged with its antiparticle (C symmetry) while its spatial coordinates are inverted (P symmetry). Initially, equal amounts of matter and antimatter should have been produced if CP-symmetry was preserved. As it is not the case, physical laws must have acted in different ways for matter and antimatter, violating charge-parity conservation. Successful as it may seem, the SM provides no source for CP-violation which is strong enough to explain the observed asymmetry.

- **Neutrino mass** According to the Higgs mechanism in the SM, particles in the vacuum acquire mass as they interact with the Higgs boson. Photons are massless because they do not interact with the Higgs boson. All particles change *handedness* when they interact with the Higgs boson: left-handed particles become right-handed, and vice versa. Experiments have shown that neutrinos are always left-handed. Since right-handed neutrinos do not exist in the SM, the theory predicts that neutrinos can never acquire mass. However, neutrinos have shown to have non-vanishing mass values, but the SM does not explain the mechanism responsible for generating neutrino masses.

All these open questions for which the SM does not provide a plausible answer lead to the development of several theories trying to give an explanation to the phenomena described above. However, the description of such models is beyond the scope of this thesis and will not be reviewed here. The study of some rare processes is sensitive to signatures of new physics and future experimental results could shed some light on how the current theory could be extended to cover the missing parts the current model cannot account for.

Chapter 2

The top quark

The top quark is the heaviest particle in the SM and could play a relevant role in complementary models that go beyond its current formalism. A good understanding of top quark phenomenology is of key importance in the development of all these models. The current analysis lays in the frontier between the SM and new physics, as it serves as a direct, stringent SM test, being at the same time sensitive to BSM effects. A brief overview of the top quark discovery is given in the introduction of this thesis. This chapter offers an introduction to the main properties of the top quark and its role in particle physics. Its different production and decay modes will also be reviewed. Further details and introductory aspects related to the theoretical basis of the analysis will be reviewed in the next chapter.

2.1 Top quark properties

Within the context of the SM, the top quark has the same quantum numbers and interactions as all other up type quarks. The left-handed top quark is the weak isospin partner from the doublet formed along with the bottom quark, with weak isospin $T_3 = +\frac{1}{2}$ and electric charge $Q_{em}^{top} = +\frac{2}{3}$. The right-handed top forms an $SU(2)_L$ singlet.

Two empirical facts distinguish the top from all other quarks and dictate its phenomenology: its much larger mass (it is more than 40 times heavier than the second heaviest quark, the bottom quark) and its very small mixing with quarks of the first and second generations. Quark mixing is encoded in the matrix elements of the CKM matrix (see 1.2.2). The matrix element V_{tb} is close to unity, whereas the elements V_{ts} and V_{td} are significantly smaller.

The value of the top quark mass is currently estimated at

$$m_t = 172.9 \pm 0.4 \text{ GeV}$$

from direct measurements [2]. Latest CMS measurements using 13 TeV data [24] yield a value of

$$m_t^{CMS} = 172.25 \pm 0.08(\text{stat+JSF}) \pm 0.62(\text{syst}) \text{ GeV}$$

where JSF is the uncertainty related to an overall jet scale factor, stat and syst stand for the statistical and systematic unceratainties, respectively. Figure 2.1 shows a summary of the most recent experimental top quark mass measurements. The latest CMS combination using Run I LHC data is highlighted in red.

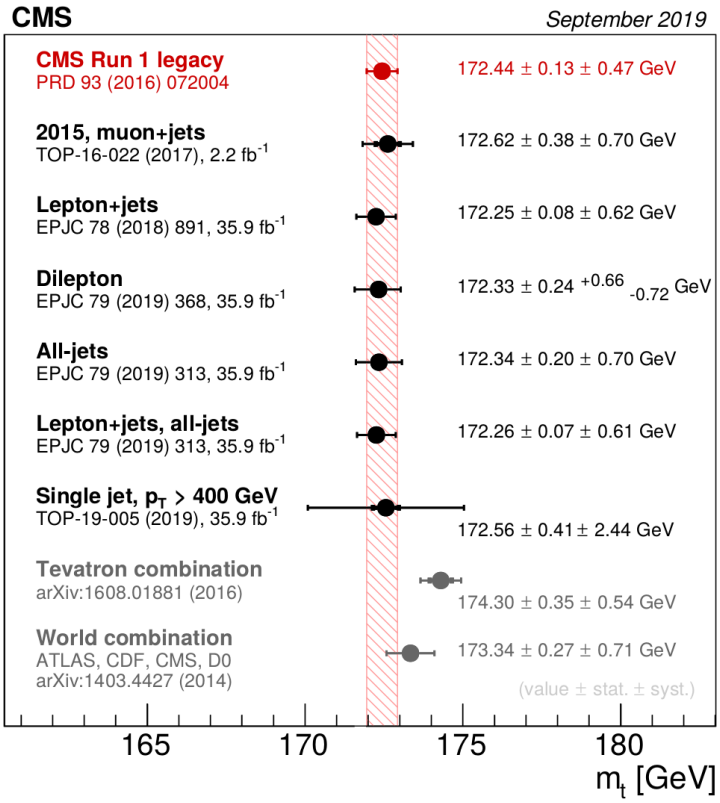


FIGURE 2.1: Summary of CMS top quark measurements in Run II. The Run I legacy, Tevatron and world combination measurements are also shown.

Neglecting the mass of the b quark and higher order terms, the total width of the top quark (Γ_t) is given by [25]

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \tag{2.1}$$

where G_F is the Fermi constant, α_s is the strong coupling and m_t and m_W are the masses of the top quark and the W boson, respectively. Due to its mass value, the top quark has a large decay width of $\Gamma = 1.41^{+0.19}_{-0.15}$ GeV.

Being heavier than a W boson, and because of the large $|V_{tb}|^2$ value, the top quark is the only quark that decays semi-weakly and almost exclusively into the two body system formed by a b quark and a W boson. Due to its large mass, the top quark has a very short lifetime of about about 0.5×10^{-24} s [2]. This value is significantly smaller than the typical time for the formation of QCD bound-state hadrons ($\tau_{QCD} \approx 1/\Lambda_{QCD} \approx 3 \times 10^{-24}$ s), and as such, the top is the only quark that decays before hadronization (no hadrons containing a top quark are expected to exist in the SM), offering unique possibilities to study bare-quark properties. These properties are crucial in calculations in the SM and beyond, as top quark properties are consequently not hidden by hadronization effects, providing a clean source of fundamental information.

2.1.1 Top quark decay

Top quarks are produced either in top-antitop pairs, or individually (the corresponding details can be found in the following section 2.2). Events from top quark pair production and single top quark production are classified by the decay products of the W boson that arises from the top quark two-body decay, the weak boson being able to decay into nine different modes.

On one hand, the three possible leptonic decays of the W boson are

$$W \to e\nu_e \quad W \to \mu\nu_\mu \quad W \to \tau\nu_\tau \tag{2.2}$$

all of them being almost equally favoured, as $BR(e^+\nu_e) = BR(\mu^+\nu_\mu) = BR(\tau^+\nu_\tau) = \frac{1}{9}$.

However, in about a 68% of the cases, a W boson decays hadronically to quark-antiquark pairs in the three possible color combinations ($R\bar{R}$, $G\bar{G}$ and $B\bar{B}$) as in

$$W \to c\bar{d}, \boldsymbol{c\bar{s}}, c\bar{b} \quad \text{or} \quad W \to \boldsymbol{u\bar{d}}, u\bar{s}, u\bar{b} \tag{2.3}$$

As the couplings are proportional to the $|V_{q\bar{q}'}|^2$ CKM matrix elements, the branching ratios of W hadronic decays are dominated by the CKM-favoured $\boldsymbol{u\bar{d}}$ and $\boldsymbol{c\bar{s}}$ final states. The different W boson decay modes, along with their corresponding branching fractions, are listed in table 2.1.

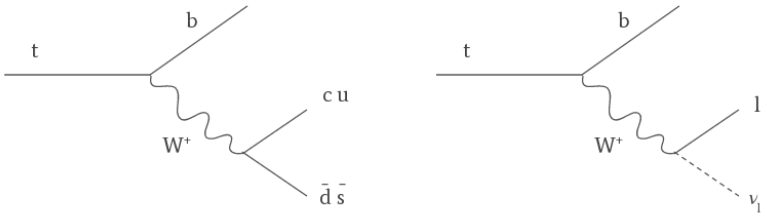


FIGURE 2.2: Top quark decay modes: hadronic (left) and leptonic (right). Only the CKM-favoured $u\bar{d}$ and $c\bar{s}$ final states are shown in the hadronic diagram.

W decay mode	Branching ratio (%)
$e\nu_e$	10.71 ± 0.16
$\mu\nu_\mu$	10.63 ± 0.15
$\tau\nu_\tau$	11.38 ± 0.21
hadrons	67.41 ± 0.27

TABLE 2.1: W boson decay modes and their corresponding branching fractions. The different top quark decay rates are proportional to these values, taken from [2].

2.2 Top quark production modes

Two classes of top quark production exist. The first proceeds by the strong QCD force to produce a top-antitop quark pair, the second one being mediated by electroweak interactions leads to the production of a single top quark (or antiquark).

2.2.1 Pair production

The dominant mechanism for top quark production at hadron colliders is pair production through strong interactions. The production of $t\bar{t}$ pairs proceeds either through the annihilation of a quark and an antiquark ($q\bar{q} \rightarrow t\bar{t}$), or through interaction between gluons (gg$\rightarrow t\bar{t}$) in the colliding beam particles. Figure 2.3 shows the Feynman diagrams for top pair production at leading order.

At the $p\bar{p}$ Tevatron collider, the quark annihilation production mode was dominant (accounting for about an 85% of the cross section) whereas gluon fusion vastly domains top quark pair production at the LHC (contributing approximately 80% - 90% to the total $t\bar{t}$ production cross section), leading to a large cross section at the LHC compared to the Tevatron because of the increasing gluon PDF towards smaller momentum fractions.

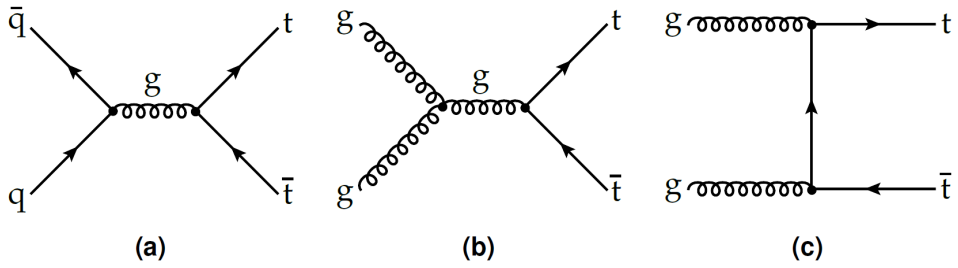


FIGURE 2.3: Feynman diagrams of $t\bar{t}$ production at LO: (a) quark annihilation; (b) s-channel and (c) t-channel gluon fusion.

Shall the reader have special interest, complementary bibliography regarding top pair production can be found elsewhere, but no further discussion will be provided throughout this document, as it plays no special role in the current analysis.

2.2.2 Single top production

Besides pair production via strong interactions, top quarks can also be produced individually in high energy collisions through electroweak processes involving the Wtb vertex. Single top quark production is of significant phenomenological relevance, and provides information that complements the one obtained from top quark pair production. In particular, it represents an optimal tool to study charged-current interactions of the top quark and is also sensitive to new physics effects and anomalous couplings.

In the SM, three types of electroweak single top quark production modes exist, that can be distinguished according to the virtuality ($Q^2 = -p_\mu p^\mu$) of the involved W boson:

- *t-channel production* if the W boson is *space-like* or has a virtuality $Q^2 > 0$. Of key importance in the current analysis, this process will be reviewed in more detail further in the next chapter.

- *s-channel production.* It is the single top quark production channel with smaller cross section, as the *time-like* ($Q^2 < 0$) mediating W boson should have a large virtuality in order to produce the heavier top quark. In various BSM scenarios however the cross section of this process is expected to increase due to new heavy particles such as W' or charged Higgs bosons which may even be produced on their mass shell, $p_\mu p^\mu - m^2 = 0$, and hence occur as a resonance.

- *tW associated production* represents the third production mode, in which the top quark is produced along with a W boson which can be produced on-shell

($Q^2 = -m_W^2$). The initial b quark is a sea quark inside the proton. The main partonic processes for t and $\bar{t}$ productions are $gb \to tW^-$ and $g\bar{b} \to \bar{t}W^+$, since other CKM-suppressed contributions from gs and gd initial states are negligibly small.

Examples of the tree-level Feynman diagrams contributing to the three different channels are shown in figure 2.4.

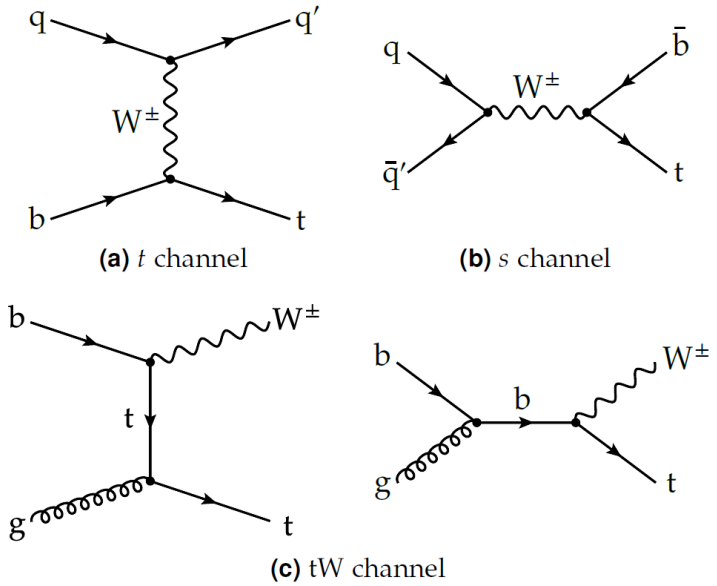


FIGURE 2.4: Examples of feynman diagrams of single top quark production at LO.

The dominant contribution to the single top production cross section at the LHC is predicted to come from the t-channel process, followed by the tW associated production. The contribution from s-channel production at the LHC is relatively small compared to the dominant process. This can be seen in figure 2.5, where the predicted and measured cross sections for these processes are presented, among other SM processes.

A peculiarity of the t- and s- channels, specific to pp collisions, is the difference between production cross sections of single t and $\bar{t}$ that results from the different parton distribution functions (PDF) of incident up and down quarks involved in the hard scattering.

Due to the valence content (uud) of the colliding protons at the LHC, the rate of top quark production at the LHC is roughly twice the rate of anti-top quark production, because it is initiated by a quark-antiquark collision.

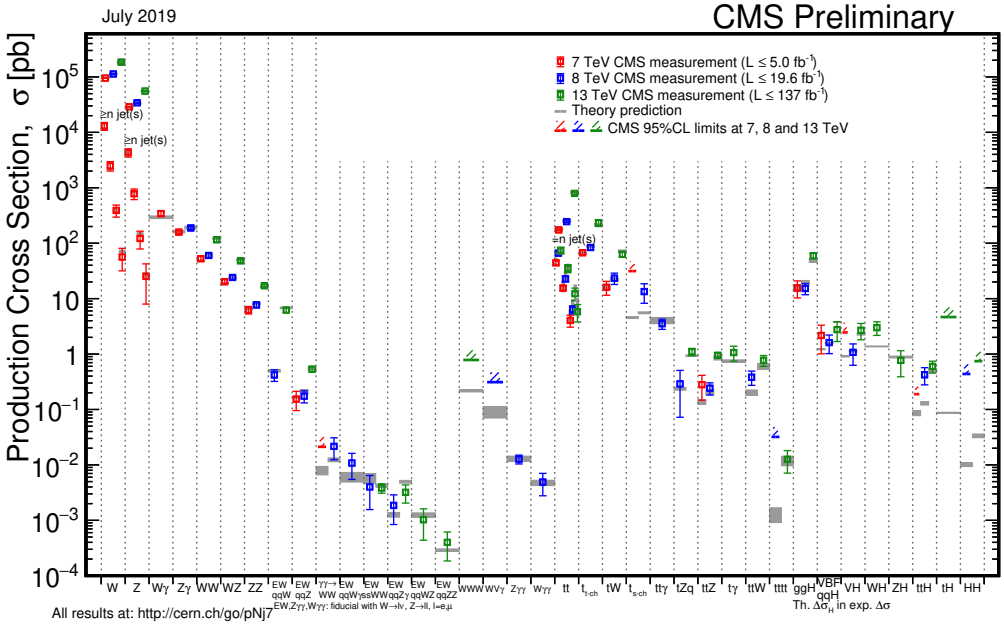


FIGURE 2.5: Summary of the cross section measurements of different SM processes provided by the CMS collaboration up to July of 2019.

2.3 LHC top quark studies in the standard model and beyond

Owing to its unique experimental properties, the top quark plays a special role among the SM fermions and offers a valuable testing ground for many SM predictions. Since its discovery in 1995, a lot of effort from both the experimental and theory communities has been put in studying its properties and processes it is involved in, testing the SM to levels of precision which are challenging the theory. The amount of top quarks produced at the LHC allowed to improve and go beyond the pioneering studies carried out at the Tevatron. In the past years, major interest has been set in studying the role of the top quark in potential TeV scale new physics, which could in fact be within reach of the LHC.

One interesting feature of the top quark is that its Yukawa coupling to the Higgs boson y_t is of order unity, substantially greater than all other Yukawa couplings, which are of order $\mathcal{O}(10^{-2})$. A precise measurement of y_t is of great importance in the understanding of EWSB mechanism and could also shed light on new physics beyond the SM. The top quark Yukawa coupling [26] is at the moment one of the most promissing SM quantities accessible at the LHC that can shed a light on the scale of new physics [1]. Events containing top quarks are also backgrounds to many

[1]The structure of the Higgs effective potential has a strong dependence on the top Yukawa coupling through the term $-6y_t^4$ in the renormalization group equation (RGE) of the quartic Higgs self coupling λ, the minus sign of the term being responsible for an extra minimum at high field values. Small changes in $\delta y_t \approx 10^{-6}$ have a significant impact on the structure of the potential

new physics processes, so the best understanding of top properties is required for experimental studies of processes beyond the SM.

In the following we present some interesting experimental top-related studies carried out at the LHC.

2.3.1 LHC studies of the top quark in the standard model

Given its role on the EWSB, an accurate knowledge of the top quark mass gives a valuable input for precision SM calculations. The top quark and the Higgs boson enter one-loop corrections in the calculation of $W^{\pm}$ and Z boson masses. Thus, a precise measurement of the top quark mass, along with the electroweak boson masses, allowed for an accurate prediction of the Higgs mass, as seen in figure 2.6 playing a key role in the discovery of the Higgs boson.

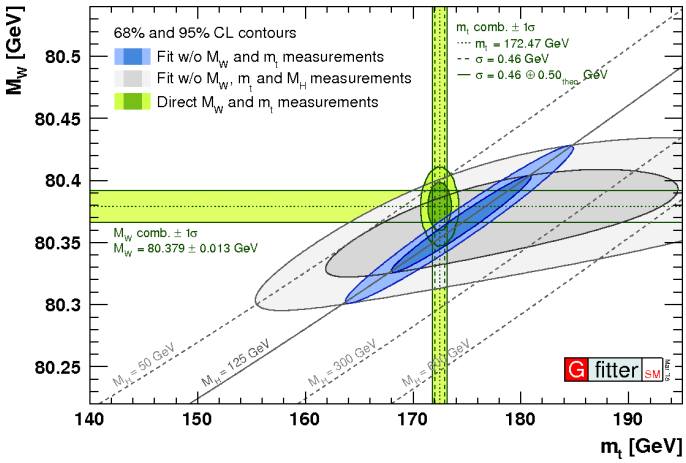


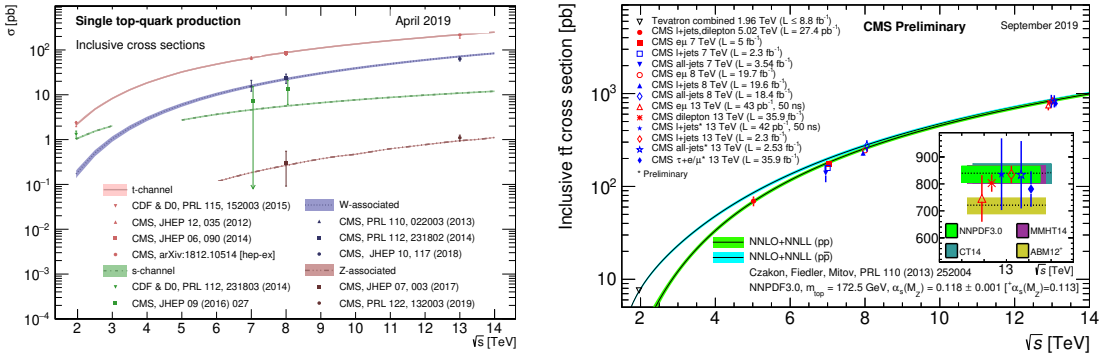
FIGURE 2.6: The 68% and 95% CL contours for the indirect determination of m_W and m_{top} from global SM fits to electroweak precision data [1].

Measurements of the single top quark cross sections allow to extract a limit on the CKM matrix element V_{tb} and study the Wtb vertex directly. The latest CMS measurement of this kind yields $|V_{tb}| = 0.998 \pm 0.038(\text{exp}) \pm 0.016(\text{theo})$ with the 95% confidence level limit being $|V_{tb}| > 0.92$ [28]. Some recent CMS single top quark cross section measurements in the different production channels are reported in figure 2.7a.

Measurements of the top pair cross section $\sigma_{t\bar{t}}$ provide important tests of the SM. The calculations depend on fundamental parameters such as the top quark mass m_t, the strong coupling constant α_S, and the parton distribution functions of the proton, so some $\sigma_{t\bar{t}}$ measurements have been used to determine these parameters

[27]. The stability of the EW vacuum could also be strongly affected by new physics, and a precise measurement of top properties serves as a bound on different BSM scenarios.

too [29]. A summary of the most recent CMS measurements of $\sigma_{t\bar{t}}$ is presented in figure 2.7b.



(A) (B)

FIGURE 2.7: Summary of single top 2.7a and top pair 2.7b production cross sections CMS measurements. Theoretical single top calculationsare courtesy of N. Kidonakis. Top pair cross section values are compared with the theory calculation at NNLO+NNLL accuracy. Tevatron measurements are also shown in both cases.

The invariance of the SM under CPT (charge, parity and time reversal) transformations predicts that particle and antiparticle masses should be equal. However, in some extensions of the SM, CPT-violating effects are present. A direct measurement of a mass difference between particle and anti-particle (Δm_t) would indicate a violation of the CPT symmetry, otherwise serving as a probe of the consistency of the SM in this aspect. The latest result regarding this topic was published by CMS using 8 TeV data, and yields a value of $\Delta m_t = -0.15 \pm 0.19(\text{stat}) \pm 0.09(\text{syst})$ GeV, consistent with the SM expectation [30].

Apart from top pair and single top production, it is possible within the context of the SM to produce four top quarks ($t\bar{t}t\bar{t}$), the representative leading order diagrams shown in figure 2.8. Being a rare SM process, its cross section can be used to constrain the magnitude and CP properties of y_t [31]: a value of the top Yukawa coupling larger than expected in the SM can lead to a significant increase in $t\bar{t}t\bar{t}$ production (see right diagram of Fig. 2.8).

Search for new physics using top quarks

Top quarks are present in many models beyond the SM. Some extensions of the theory predict new heavy particles that may decay into top quark pairs, resulting as resonance of $t\bar{t}$ pairs. Specific analyses are developed to look for these new particles through the top quark pair invariant mass distribution, which could either be scalars, vector or axial-vector particles (such as an hypothetical Z' boson [32]) or even spin-2 particles, their hypothetical mass ranging from 1 TeV up to several TeV.

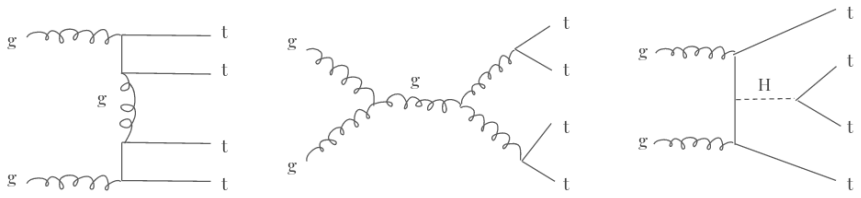


FIGURE 2.8: Representative Feynman diagrams for $t\bar{t}t\bar{t}$ production at LO in the SM.

Other models have an extended Higgs sector that includes additional charged and neutral Higgs bosons which may couple predominantly to top quarks given its mass. As an example, the Minimal Supersymmetric Standard Model (MSSM) and 2-Higgs doublet models (2HDM) include in their particle content charged Higgs bosons, which could be produced in association with top quarks via the channel $bg \to tH^-$ or $\bar{b}g \to tH^+$ [33, 34]. Compositeness models (top-color assisted technicolor, top-seesaw, etc) are based on effective operators containing the top, and it also plays a special role in models with extra warped dimensions. Moreover, $t\bar{t}t\bar{t}$ production can be significantly enhanced by BSM particles and interactions, having special sensitivity to new physics effects. As such, it has been used to constrain 2DHM and simplified dark matter models [35].

Another example where searches in the top sector would reveal new physics is through observations of flavour-changing neutral currents (FCNC) decays of the top quark. Within the context of the SM these decays are predicted to be extremely rare, whereas in many BSM models, such as the 2HDM and the MSSM, FCNC can be highly enhanced. The large amount of top quarks produced currently at the LHC allows to search for specific top rare decays. Even though the SM states that tops decay to a W boson and a b quark with a branching fraction of about 100%, some extensions predict that it may also decay to a Z boson and a quark, $t \to Zq$, where q can be either a u or a c quark. These processes are predicted to have very small branching fractions, of the order of 10^{-14}, any enhancement of the predicted cross sections therefore possibly presenting an indicative of new physics. Experimental signatures of FCNC processes in the top sector can be sought for either in single top or top quark pairs production. The first searches for FCNC decays were made at the Tevatron, but even now at the LHC no evidence for these processes has been found.

The latest analysis considering single top quark FCNC production in association with the Higgs boson ($pp \to tH$), top quark pair production with FCNC decay of the top quark ($t \to qH$) [36] and tZ-FCNC production [37] observed no significant deviation from the predicted background. Both studies set upper limits on the corresponding FCNC branching fractions, which are shown in table 2.2.

The current analysis, devoted to the measurement of the tZq production cross

Process	Observed	Expected
$\mathcal{B}(t \to uZ)$	0.022	0.027
$\mathcal{B}(t \to cZ)$	0.049	0.118
$\mathcal{B}(t \to uH)$	0.47	0.34
$\mathcal{B}(t \to cH)$	0.47	0.44

TABLE 2.2: Observed and expected upper limits on the tZ- and tH-FCNC branching ratios. The values are taken from [37] and [36].

section, is sensitive to BSM processes such as FCNC decays involving the direct coupling of the top quark to a Z boson and an up or charm quark, either at the production or decay. Deviations from the expected SM tZq cross section could be an indication of BSM-FCNC signatures. The next chapter offers an overview of the theory related to the analysis presented in this thesis.

Chapter 3

Single top rare processes: tZq

The main theoretical aspects relative to the search of the associated production of a single top quark and a Z boson, the topic of this thesis, are described in this chapter. Starting from the basics of single top production and, in concrete, the characteristics of the t-channel production mode, we close the chapter with the description of the SM tZq production, remarking its importance in the search of other interesting SM rare processes, as well as its relation with the search for new physics, and in particular with FCNC searches in the top sector.

Increasing luminosities and centre-of-mass energy at the LHC have motivated the search for rare SM processes which were not accesible with the previous conditions. One of such processes in the top quark sector is the associated production of a single top quark and a Z boson along with an additional quark (tZq). This particular process has a rather small cross section compared to other SM processes (see figure 2.5).

In this chapter we will present a brief introduction to tZq-SM production, a single top t-channel process in which a Z boson is radiated off one of the quark lines.

3.1 Single top t-channel production

Single top production via the t-channel has the highest cross section in proton-proton collisions (compared to s-channel at tW production). In this channel a top quark is produced through the exchange of a W boson between a light-flavoured quark and a bottom quark, changing the flavour of the bottom quark to a top quark. The W boson involved in this process has a virtuality $Q^2 > 0$ and is said to be *space-like*. A characteristic feature of this mode is the production of an additional spectator quark (q') which recoils against the W boson and tends to be scattered fairly forward in the CMS detector ($|\eta| > 3$), making it easier to

identify t-channel single top events in the analysis. In this production mode, top quarks appear roughly twice as often as antitops, given the valence composition of the proton (two up quarks, and a down quark). t-channel single top and antitop production diagrams at tree level are shown in figure 3.1.

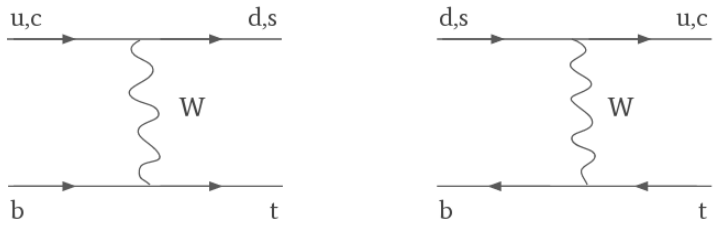


FIGURE 3.1: t-channel diagrams for single top (left) and single antitop (right) SM production.

The description of this process and its cross section calculations can be performed in two different scenarios or *schemes*. Processes involving b quarks can be described in QCD either in the *4-* or *5- flavour schemes* (FS), depending on whether the b quark is considered part of the proton or not. In the 4-flavour scheme, the b quarks are generated in hard scattering from gluon splitting and appear only in the final state. 5-flavour calculations include the b quark in the initial state, coming from a non-vanishing bottom PDF in the colliding proton. This makes the 5-flavour scheme a 2→2 process and the 4-flavour one a 2→3 process. Figure 3.2 shows the tree-level diagrams for t-channel single top production in both schemes. Predictions

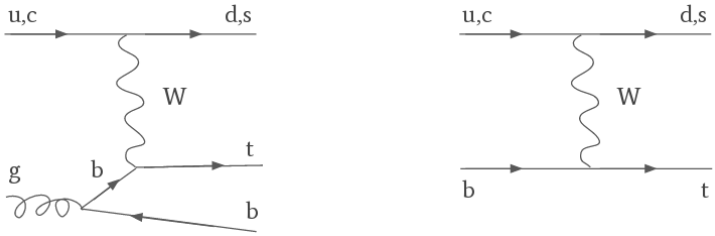


FIGURE 3.2: Leading order t-channel Feynman diagrams in the 4-flavour (left) and 5-flavour (right) schemes.

in these two schemes show substantial agreement, even though calculations in the 5-flavour scheme are considerably simpler, leading to more stable predictions and an easier computation at higher orders. However, its implementation in simulation depends on the gluon splitting model in the parton shower, whereas this is straightforward in the 4-flavour scheme.

In the SM, the flavour quantum number of fermions can be changed by charged currents, i.e., through the weak interactions mediated by the exchange of a charged $W^{\pm}$ boson. FCNC processes such as $t \to qX^0$ where X^0 is a charge-neutral boson (photon, Z, gluon or H) are absent at tree level. FCNC can happen in higher-order

loop diagrams with the help of a virtual W-boson; however this kind of processes are highly suppressed through the GIM mechanism. The SM estimation of the branching ratio for the FCNC transition $t \to qZ$ yields $\mathcal{B}(t \to qZ) \sim 10^{-14}$. Several BSM models suggest that the FCNC branching ratios can be significantly enhanced, by orders of magnitude.

3.2 tZq production in the standard model

Within the context of the SM, the production of a top quark in association with a Z boson and an additional quark proceeds via the single top t-channel mechanism, in which the Z boson is radiated off one of the quark lines (see figure 3.3). This predominantly occurs through the leading order electroweak processes

$$u + b \to d + t + Z \qquad \bar{d} + b \to \bar{u} + t + Z \tag{3.1}$$

for top quark production, and

$$d + \bar{b} \to u + \bar{t} + Z \qquad \bar{u} + \bar{b} \to \bar{d} + \bar{t} + Z \tag{3.2}$$

for antitop production, at a lower rate as explained in section 2.2.2. Processes initiated by charm and strange quarks do also take place, with smaller contributions. The main leading order diagrams that contribute to this final state can be seen in figure 3.3, where the contribution from non-resonant lepton pairs is also considered (bottom right diagram). Further information regarding the non-resonant contribution is given in chapter 6.

The Z boson may also be radiated from the W boson (bottom left diagram in figure 3.3), and as a consequence the process is sensitive to the *triple gauge coupling* (TGC) WWZ. In fact, an enhancement in the production rate could be an indication of a contribution from anomalous couplings in the WWZ vertex. Cubic (and quartic) self-interactions of the electroweak gauge bosons are present in the SM due to the non-abelian nature of the electroweak interaction, and are completely fixed by the gauge couplings. This is not the case in some BSM scenarios, and processes that are sensitive to gauge boson self-interactions are important tools to search for nonstandard effects.

There is also additional interest in searching for rare single top processes, such as tZq or tHq. Electroweak production of a single top quark in association with a Z or Higgs boson provides a natural opportunity to constrain possible deviations of the neutral couplings of the top quark with respect to the SM predictions. Asking for just one top quark (or antiquark) in the final state (instead of top pairs as in $t\bar{t}V$ or $t\bar{t}H$) implies that no QCD interactions are present at LO, making the process "purely" electroweak. As a consequence, QCD corrections are typically smaller and under control, and dimension-6 corrections of QCD interactions enter only at NLO,

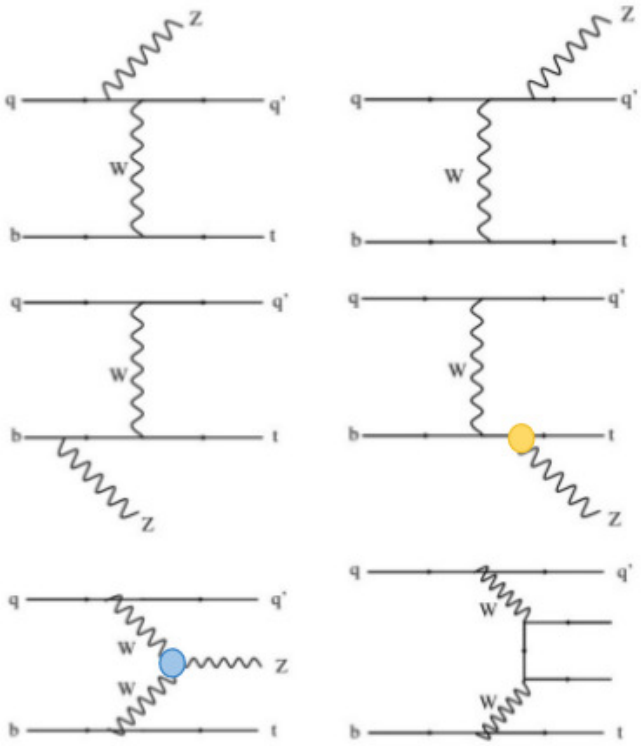


FIGURE 3.3: Leading-order $t\ell^+\ell^-q$ production diagrams. The lower right-hand diagram represents the non-resonant contribution to the process.

	$Br(t \to q\gamma)$ [38]	$Br(t \to qZ)$ [37]	$Br(t \to qg)$ [39]	$Br(t \to qH)$ [40]
$q = u$	$1.3 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	$4.0 \cdot 10^{-5}$	$5.5 \cdot 10^{-3}$
$q = c$	$1.7 \cdot 10^{-3}$	$4.9 \cdot 10^{-4}$	$20 \cdot 10^{-5}$	$4.0 \cdot 10^{-3}$

TABLE 3.1: Current limits for FCNC top processes.

do not spoiling the sensitivity of electroweak couplings. These measurements allow to put constrains on top-quark, triple gauge, and gauge-Higgs interactions.

3.2.1 tZq final states

Depending on the decay of the gauge bosons (the W boson coming from the top decay, and the associated Z boson) we may find different tZq decay channels:

- **Trilepton channel**: both bosons decay leptonically ($W \to \ell\nu_l$ and $Z \to \ell\ell$).
- **Dilepton channel**: the Z boson decays to a lepton pair, and the W boson decays hadronically to a pair of quarks ($W \to q\bar{q}'$ and $Z \to \ell\ell$).
- **Single lepton channel**: the Z boson decays to a quark pair, and the W boson decays leptonically ($W \to \ell\nu_l$ and $Z \to q\bar{q}$).
- **All-hadronic channel**: both the Z boson and the W boson decay to quarks, and no leptons are present in the final state ($W \to q\bar{q}'$ and $Z \to q\bar{q}$).

where ℓ denotes either electrons, muons or tau leptons. The trilepton channel is of special experimental interest, as it is the one with cleaner signal, i.e., with the lowest level of background, even though it is the channel with the smallest cross section. The next-to-leading-order cross section for the process $tZ(\ell^+\ell^-)q$ where the Z boson decays to leptons calculated for proton-proton collisions at a center-of-mass energy at 13 TeV in the 5F scheme is given by

$$\sigma(tZq \to Wb\ell^+\ell^-q) = 94.2^{+1.9}_{-1.8}(scale) \pm 2.5(PDF) \quad \text{fb} \tag{3.3}$$

This value was obtained using the Monte Carlo generator `MADGRAPH5@aMCatNLO` 2.2.2. The calculation, which includes lepton pairs from off-shell Z bosons with invariant mass $m_{ll} > 30GeV$, uses the NNPDF 3.0 set of parton distribution functions. The *scale* and *PDF* uncertainties are estimated, respectively, by changing the QCD renormalization and factorization scales by factors of $\frac{1}{2}$ and 2, and by using the 68% confidence level uncertainty on the PDF set. This cross section calculation will serve as reference in the rest of the analysis.

tZq production in the SM is a single-top process and, as such, a *jet* arising from the hadronization of the recoiling quark is found in the forward region of the

detector, a signature that allows to distinguish tZq-SM from $t\bar{t}-$FCNC production, which has exactly the same signal with no recoiling jet in the final state (see figure 3.6). Further details on the topology and properties of the trilepton tZq channel are presented in chapter 6. Figure 3.4 shows a schematic view of the final state of tZq production in the trilepton channel.

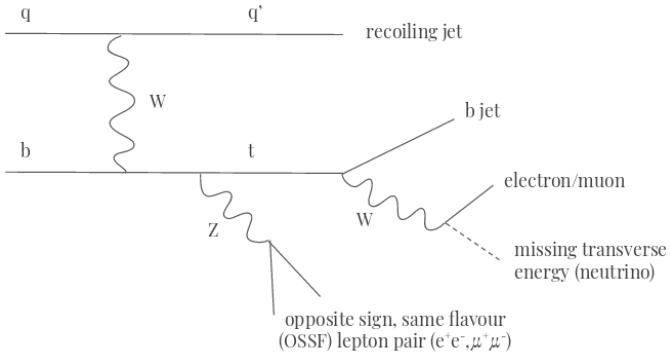


FIGURE 3.4: tZq final state in the trilepton channel.

3.2.2 tZq as background for other standard model rare processes

SM tZq production is a background to other SM procees, such as $t\bar{t}H$ and tHq production. As stated in chapter 2, these searches are specially interesting as they allow to measure, for instance, the top quark Yukawa coupling to the Higgs boson y_t, playing a key role in the study of the EWSB mechanism.

tZq associated production also represents a major background to other interesting SM processes, such as the production of top-quark pairs in association with massive vector bosons ($t\bar{t}V$ where V can be either a Z boson or the charged $W^{\pm}$). Precise measurements of the production cross section of $t\bar{t}W$ and $t\bar{t}Z$ are of particular interest because these topologies might have significant contributions from new physics beyond the SM and at the same time represent important background sources in some searches for new physics and other SM processes (such as $t\bar{t}H$ production).

Some other processes are also affected by the contribution from tZq production, although to a lesser extent. Those include tHq, $t\bar{t}H$ and four-tops production.

3.2.3 Top flavour-changing neutral interactions beyond the standard model

This search is also appealing in the context of the SM *effective field theory* (EFT). In this context, the SM is considered an effective low-energy theory applicable up to energies not exceeding a certain scale Λ, and deviations from its predictions can be described by the effect of higher-dimensional operators of the SM fields in the Lagrangian. Experimental results can help determining the way in which these deviations are affected by the different operators. In the absence of convincing evidence for new resonances, EFT provides an appealing model-independent approach to treat possible discrepancies with the current theory.

The extension of the Lagrangian includes the effect of higher dimensional operators ($d > 4$) that are suppressed by powers of Λ (in general, operators are suppressed by a factor Λ^{d-4} in which d is the dimension of the operator), resulting in the addition of interactions to the set of SM vertices, as in

$$\mathcal{L}_{SM} = \mathcal{L}_{SM}^{(4)} + \frac{1}{\Lambda}\sum_k c_k^5 Q_k^5 + \frac{1}{\Lambda^2}\sum_k c_k^6 Q_k^6 + \mathcal{O}\left(\frac{1}{\Lambda^3}\right) \tag{3.4}$$

where $\mathcal{L}_{SM}^{(4)}$ is the usual renormalizable part of the SM Lagrangian, containing only dimension two and four operators, Q_k^d stand for the different dimension-d operators, and c_k^d are the corresponding dimensionless coupling constants (also known as *Wilson coefficients*).

There is only one dimension-5 operator that generates a Majorana mass term for the neutrinos and mixing between the different flavour eigenstates, required to be non vanishing by neutrino phenomenology. However, there exist 59 different dimension-6 operators that include four-fermion operators as well as vector-scalar-fermion ones. Certain dimension-6 operators parametrize the top quark couplings to other SM particles, contributing at $\mathcal{O}(\Lambda^{-2})$, some of which violate lepton and baryon number or give corrections to flavour-changing processes suppressed by the GIM mechanism.

Flavour-changing neutral transitions as the tZ-FCNC and $t\bar{t}$-FCNC production modes are shown in figures 3.5 and 3.6, respectively, and can be incorporated to the SM Lagrangian using effective operators of dimension 4 and 5, its contributions quantified by a set of anomalous couplings [41]. A list of operators that can be constrained by different measurements related to top quark processes is shown in Figures 3.7 and 3.5. Here one can see how the measurement of the tZq (tZj in the table, j standing for *jet*) might help constraining most of these operators.

Many models for new physics predict new contributions to top flavour-changing neutral current interactions that are orders of magnitude in excess of the SM expectations. In particular, 2HDM leads to potentially measurable FCNCs, even though

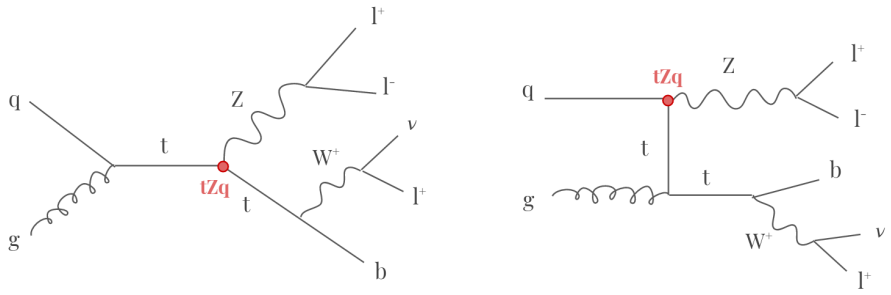


FIGURE 3.5: Single top FCNC production diagrams. The flavour violating vertices (tZq) are highlighted in red.

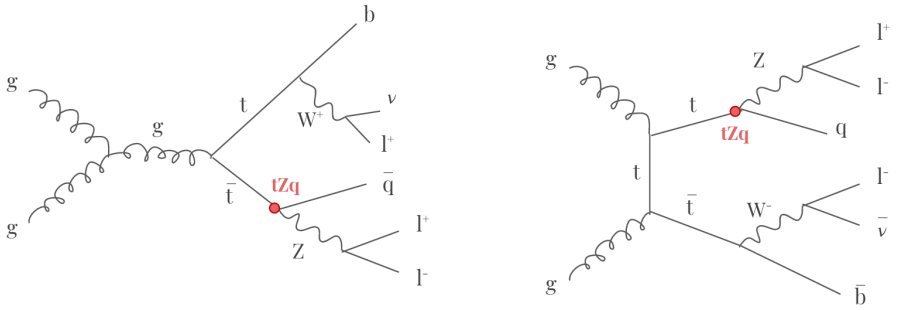


FIGURE 3.6: Top quark pair FCNC. The flavour violating vertices (tZq) are highlighted in red.

Process	O_{tG}	O_{tB}	O_{tW}	$O^{(3)}_{\varphi Q}$	$O^{(1)}_{\varphi Q}$	$O_{\varphi t}$	$O_{t\varphi}$	O_{bW}	$O_{\varphi tb}$	O_{4f}	O_G	$O_{\varphi G}$
$t \to bW \to bl^+\nu$	N		L	L				L^2	L^2	$1L^2$		
$pp \to tj$	N		L	L				L^2	L^2	1L		
$pp \to tW$	L		L	L				L^2	L^2	1N	N	
$pp \to t\bar{t}$	L									2L-4N	L	
$pp \to t\bar{t}j$	L									2L-4N	L	
$pp \to t\bar{t}\gamma$	L	L	L							2L-4N	L	
$pp \to t\bar{t}Z$	L	L	L	L	L	L				2L-4N	L	
$pp \to t\bar{t}W$	L								L	1L-2L		
$pp \to t\gamma j$	N	L	L	L				L^2	L^2	1L		
$pp \to tZj$	N	L	L	L	L	L		L^2	L^2	1L		
$pp \to t\bar{t}t\bar{t}$	L									2L-4L	L	
$pp \to t\bar{t}H$	L						L			2L-4L	L	L
$pp \to tHj$	N		L	L			L	L^2	L^2	1L		N
$gg \to H$	L						L				N	L
$gg \to Hj$	L						L				L	L
$gg \to HH$	L						L				N	L
$gg \to HZ$	L			L	L	L	L				N	L

FIGURE 3.7: Schematic representation of relation between EFT dimension-6 operators and processes, focusing on single top production and associated channel, entering different top quark processes, either at LO, LO^2 or at NLO in QCD.

the tree-level flavour conservation is still guaranteed in this model. However, it predicts measurable top FCNCs due to loop processes that involve the additional

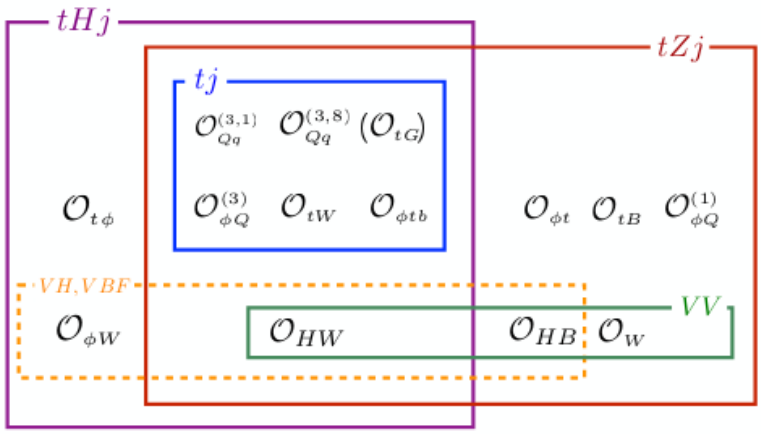


FIGURE 3.8: Dimension-6 EFT operators to which different analyses have sensitivity (tHq, tZq, VV).

charged Higgs bosons. The rate of the flavour-violating processes depends on the mass of the charged Higgs bosons. In the MSSM, top flavour-changing neutral interactions arise at one loop in the presence of flavour-violating mixing in the soft SUSY-breaking mass matrices. In models of warped extra-dimensions, top FCNCs arise when SM fermions propagate in the extra dimension with profiles governed by the Yukawa coupling. These non-trivial profiles lead to flavour-violating coupling between the SM fermions and the Kaluza-Klein excitations (KK) of the SM gauge bosons, and these couplings are predicted to be the largest for the top quark. R-parity violating supersymmetric models [42], top-color assisted technicolor models [43] and singlet quark models [44] do also predict enhancements of the FCNC rate, orders of magnitude above the prediction of the standard production.

Process	SM	2HDM(FV)	2HDM(FC)	MSSM	RPV	RS
$t \to Zu$	7×10^{-17}	–	–	$\leq 10^{-7}$	$\leq 10^{-6}$	–
$t \to Zc$	1×10^{-14}	$\leq 10^{-6}$	$\leq 10^{-10}$	$\leq 10^{-7}$	$\leq 10^{-6}$	$\leq 10^{-5}$
$t \to gu$	4×10^{-14}	–	–	$\leq 10^{-7}$	$\leq 10^{-6}$	–
$t \to gc$	5×10^{-12}	$\leq 10^{-4}$	$\leq 10^{-8}$	$\leq 10^{-7}$	$\leq 10^{-6}$	$\leq 10^{-10}$
$t \to \gamma u$	4×10^{-16}	–	–	$\leq 10^{-8}$	$\leq 10^{-9}$	–
$t \to \gamma c$	5×10^{-14}	$\leq 10^{-7}$	$\leq 10^{-9}$	$\leq 10^{-8}$	$\leq 10^{-9}$	$\leq 10^{-9}$
$t \to hu$	2×10^{-17}	6×10^{-6}	–	$\leq 10^{-5}$	$\leq 10^{-9}$	–
$t \to hc$	3×10^{-15}	2×10^{-3}	$\leq 10^{-5}$	$\leq 10^{-5}$	$\leq 10^{-9}$	$\leq 10^{-4}$

FIGURE 3.9: SM and new physics model predictions for branching ratios of top FCNC decays. The SM predictions are taken from [45], on 2HDM with flavour-violating Yukawa couplings [45], [46] (2HDM (FV) column), the 2HDM flavour-conserving (FC) case from [47], the MSSM with 1 TeV squarks and gluinos from [47], the MSSM for the R-parity violating case from [42], [48], and warped extra dimensions (RD) from [49], [50].

SM tZq production, object of the current analysis, is an irreducible background in these searches and thus its study, as it had already been stated, is of unprecedented importance in new physics searches as any substantial disagreement with the expectations could be a probe of flavour-violating anomalous contributions.

Chapter 4

Experimental setup: the CMS detector at the LHC

The measurement in this thesis is based on proton-proton collision data recorded with the CMS detector during 2016 at a centre-of-mass energy of 13 TeV. The aim of this chapter is to provide the reader with a brief description of CERN's Large Hadron Collider and CMS experiment, including the description of its various subdetectors (the tracking detector, the electromagnetic and hadronic calorimeters and the muon spectrometers) and the trigger system. A brief overview of the software and data acquisition process in CMS is also given.

4.1 The LHC at CERN

The *Large Hadron Collider* (LHC) [51] is currently the most powerful particle collider in the world. Put into service in 2008, it is a 27 km ring located 100 m under the surface of the French-Swiss border at CERN, in the vicinity of Geneva. It was constructed between 1998 and 2008 in the existing tunnel of its predecessor, the Large Electron Positron Collider (LEP).

It was designed to study proton-proton collisions at a centre-of-mass energy of 14 TeV with a luminosity up to 10^{34} cm^{-2} s^{-1}, with the goal of searching for undiscovered physics predicted by the SM and possible BSM phenomena. One of the main motivations of building an accelerator with such characteristics was the search of the Higgs boson, the last missing particle of the SM, whose discovery was presented in summer of in 2012. With all the SM particles in the table, the LHC presents itself as a tool to investigate deeper into the SM, looking for rare processes predicted by the model, and new physics signals.

The LHC hosts two beam pipes in which protons are accelerated in opposite directions, circulating in proton bunches. The beams are focused and collide at four different interaction points (IP) in which the four main detectors are located:

ALICE (A Large Ion Collider Experiment) [52], ATLAS (A Toroidal LHC ApparatuS) [53], CMS (Compact Muon Solenoid) [54], and LHCb (LHC beauty) [55]. ALICE is an experiment that involves the study of heavy ion collisions and the QCD phase diagram. The LHCb detector is a single-arm spectrometer designed to cover the b-quark physics sector, intimately related to CP-symmetry violation. Finally, ATLAS and CMS are the two *multipurpose* experiments of the LHC, which cover a wide physics programme, ranging from precision SM measurements to searches for yet unseen SM events or signs of new physics. They have similar goals but different technical solutions and design, and are situated at opposite locations in the ring.

In addition to the four main detectors, three smaller experiments (LHCf [56], TOTEM [57], and MoEDAL [58]) were installed at the LHC using certain fractions of the scattered particles from the IPs of the ATLAS, CMS, and LHCb detectors, respectively. LHCf is the smallest detector at the LHC, and is intended to measure the properties of forward-moving particles produced when protons crash together. Its goal is to test the capability of cosmic ray measuring devices. TOTEM is a long, thin detector connected to the LHC beam pipe designed to achieve precise measurements of protons as they emerge from collisions at small angles, inaccessible to other detectors. Finally, the goal of the MoEDAL experiment is to search directly for magnetic monopoles, hypothetical particles with magnetic charge.

The bending of the beams is achieved by a total of 1232 superconducting dipole magnets that keep the protons orbiting during acceleration. The beams are focused around the nominal orbit using a set of 392 quadrupole magnets, and additional multipole (sextu-, octu-, and decapole) magnets are used for further non-linear corrections. Special quadrupole triplets at each side of the four interaction points focus the beams for collision. The superconducting magnets are cooled down to 1.9K using liquid helium.

Acceleration process

Protons in the beams are accelerated in different consecutive steps, to produce the proton bunches that will finally collide at the four interaction points. Hydrogen atoms are first stripped of their electrons through an electric field, the remaining protons accelerated up to 50 MeV in the linear accelerator *LINAC 2*, and straightaway injected into the *Proton Synchrotron Booster* (PSB), where their energy is ramped up to 1.4 GeV before entering the *Proton Synchrotron* (PS). In this synchrotron, the 25 ns bunch spacing takes place and protons are further accelerated to 28 GeV right before entering the 7km-diameter *Super Proton Synchrotron* (SPS), where they are pushed up to the LHC injection energy of 450 GeV. Once in the LHC, protons are accelerated to their maximal energy. A basic sketch of CERN's accelerator complex is displayed in figure 4.1.

Luminosity calculation and beam parameters

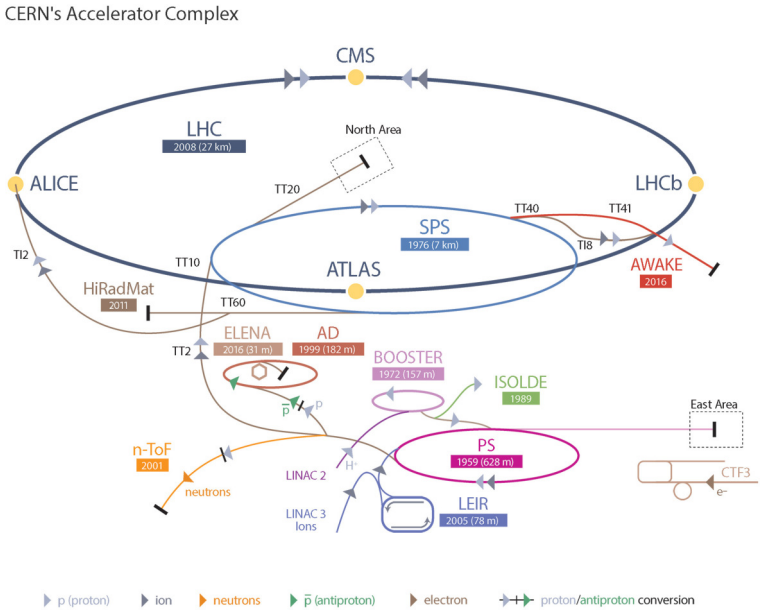


FIGURE 4.1: CERN accelerator complex.

The expected event rate for a given physics process i is connected to its production cross section σ_i in pp collisions though the *instantaneous luminosity* of the machine ($\mathscr{L}$):

$$\frac{dN_i}{dt} = \mathscr{L}(t) \times \sigma_i \tag{4.1}$$

the total number of events being $N_i = L \times \sigma_i$, where $L = \int \mathscr{L}(t)dt$ is the integrated luminosity over the entire data taking period.

This instantaneous luminosity at the interaction points, or the number of protons crossing a unit surface (cm^2) by unit time (s), depends only on beam parameters, which can be adjusted in order to achieve higher production rates at a given centre-of-mass energy in the LHC apparatus. For a Gaussian beam distribution, it can be written as follows

$$\mathscr{L} = \frac{N_p^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} \cdot F \tag{4.2}$$

In this equation,

- N_p is the proton population per beam, with a value of $1.18 \cdot 10^{11}$ in 2016.

- n_b is the number of colliding bunches per beam, 2808 for 25 ns spacing .
- f_{rev} is the bunch revolution frequency, 11245 Hz.
- $\gamma_r = \frac{E}{m}$ is the relativistic Lorentz factor, where E is the beam energy and m the particles' mass. For $E = 6.5$ TeV and $m = 0.938$ GeV, the relativistic gamma factor takes a value of 6930.
- ϵ_n is the normalized transverse beam emmitance, a parameter that gives an idea of the spatial and momentum dispersion of the beam. During 2016, it took values ranging between 3.5 μm-2.6 μm.
- β^* is the value of the betatron function $\beta(z)$ at the IP, which is related to the transverse size of the particle beam along the beam trajectory and, at position z,is given by

$$\beta(z) = \beta^* + \frac{z^2}{\beta^*} \tag{4.3}$$

  Its typical value in CMS is around 0.4-0.55 μm.
- Finally, F is a geometrical factor that accounts for the reduction in luminosity due to the crossing angle at the collision point.

Further description of the beam parameters can be found in [51].

During the data taking period comprising the whole 2016, the CMS detector was able to collect a total integrated luminosity of 37.76 fb^{-1} out of the 40.82 fb^{-1} delivered by the LHC. Prior to physics analysis, all data needs to pass a stringent certification procedure that ensures that only those with the highest possible quality are used. The total amount of luminosity certified for analysis is then reduced to 35.9 fb^{-1} in 2016.

In addition to proton-proton collisions, the LHC has a dedicated heavy ion physics programme. During these data taking periods, heavy ions (Pb-Pb and p-Pb) are accelerated at the LHC ring at a nucleon-nucleon centre-of-mass energy of 5.1 TeV (2.76 TeV during Run 1).

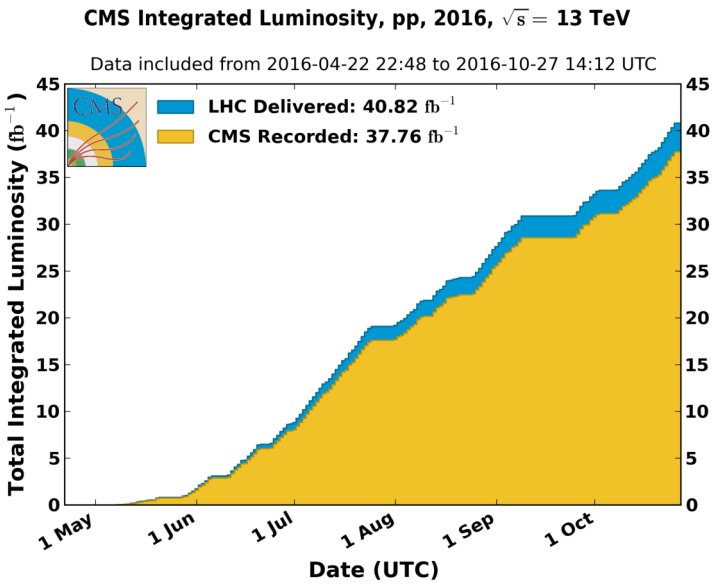


FIGURE 4.2: Cumulative day-by-day integrated luminosity delivered in by the LHC (blue) and recorded by CMS (yellow) in 2016.

4.2 The CMS detector

The Compact Muon Solenoid (CMS) [54] experiment is one of the two general-purpose detectors collecting data from LHC collisions. It is the heaviest detector at the LHC, weighting approximately 14000 tons, and is designed to collect data from proton-proton and heavy ion collisions at high luminosities. It is located at interaction point 5 (P5) of the LHC ring, in Cessy (France).

The detector is shaped cylindrically by layers in the barrel region and endcap disks in the forward regions around the beam pipe. It has an overall length of 21.6 m and a diameter of 14.6 m. Its compactness compared to ATLAS, the other multi-purpose experiment, is mainly due to its superconducting solenoid, which produces a magnetic field of 3.8 T.

One of the principal goals of the LHC when it was designed was the discovery of the Higgs boson, whose search had to be done over a wide mass range, as its value was very loosely constrained. As a consequence, the CMS detector was designed to measure as precisely as possible a huge variety of possible decay channels and some limitations or requirements had to be taken into account during its design. Some of these constraints were:

1. A high performance tracker system to distinguish particles originating from the b quark and τ lepton decays, and separate tracks from the main interaction from information associated to other vertices.

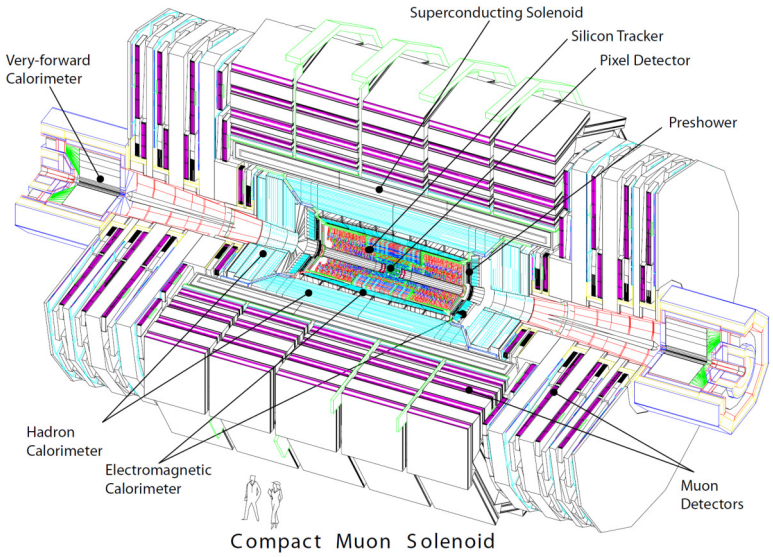


FIGURE 4.3: Layout of the CMS detector.

2. An optimal muon detection and triggering system, able to provide the best possible muon identification and resolution at a wide range of energy and angular values, and good dimuon mass resolution.

3. Good electromagnetic calorimetry, to obtain good lepton identification and isolation, and good diphoton and dielectron mass resolution.

4. Some BSM theories predict the existence of weakly interacting particles which, in case of being produced, will be associated to a significant amount of transverse missing energy. Thus, an accurate reconstruction of all particles in the event is crucial. Also, a good hermeticity is needed to provide good missing transverse energy and jet energy resolution.

The high instantaneous luminosity in collisions represents also a constraint in the detector design. With LHC collisions produced every 25 ns, a very fast trigger system and detector response are required. In addition, the large flux of particles in the forward regions had to be taken into account at the design of the subdetectors in those regions to make them resistant to radiation damage.

4.2.1 The CMS coordinate system

The origin of the CMS coordinate system is the interaction point at the centre of the detector. The x-axis points radially inward to the middle of the LHC ring, the y-axis pointing upwards and the z-axis directed along the beamline in anticlockwise direction. The (x, y) plane is usually referred to as the transverse plane, and

quantities in this plane are often denoted with a subscript T (i.e. the projection of particle momenta in the transverse plane or transverse momenta, p_T).

A cylindrical coordinate system (z, ϕ, θ) is often adopted. The azimuthal angle ϕ is measured from the x-axis on the transverse plane taking values between $[-\pi, +\pi]$, and the polar coordinate $\theta \in [0, \pi]$ is measured with respect to the z-axis. The *rapidity* y is defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{4.4}$$

for a particle with energy E and momentum p_z along the beam axis. Differences in rapidity are invariant under Lorentz boosts, making this variable useful in hadronic collision studies. A convenient reformulation of the polar angle, the *pseudorapidity* (η), is often used in hadronic collisions:

$$\eta = \arctan \frac{p_z}{|p|} = -\ln \left[\tan \frac{\theta}{2} \right] \tag{4.5}$$

The minus sign is added in the pseudorapidity definition to make η positive when points to the same direction as the magnetic field. The values for η are zero in the transverse plane and $\pm\infty$ along the beam axis. The description of the angular apertures of the different subdetector parts in CMS will be given in terms of the pseudorapidity. The adopted coordinate system is shown in figure 4.4

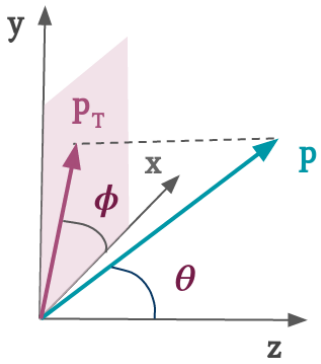


FIGURE 4.4: CMS coordinate system.

4.3 Solenoid Magnet

The central feature of the CMS detector is a superconducting solenoid of 12.5 m length and 6 m internal radius, the largest ever built, which provides a uniform magnetic field of 3.8 T in the beam direction for the set of inner detectors. The

intense magnetic field enables to measure the momentum of charged particles by analyzing their curved trajectories in the inner tracking system. In addition, muon momenta can be measured a second time in the muon system, placed right outside the magnet volume, using the outer return flux of the magnetic field.

The magnet itself consists of four layers of NbTi superconducting cables and is placed within a vacuum tank cooled down to 4.7 K. A large 10k-tons iron yoke, comprising 5 wheels and 2 endcaps, guides the return flux of the magnetic field, resulting in a homogeneous magnetic field within its full volume.

4.4 Inner tracker

The CMS inner tracking system [59] is located surrounding the interaction point along the beam line.

With a length of 5.8 m and a diameter of 2.5 m, it constitutes the first detector layer traversed by particles, subsequently dealing with a very intense flux of particles. The CMS tracking system is therefore designed to sustain high radiation levels, achieve an optimal reconstruction of charged particles trajectories bent by the magnetic field and measure their momenta, charge and point of origin after track reconstruction. In order to successfully address CMS's physics programme, it is required to achieve a reconstruction efficiency $> 95\%$ for isolated tracks and $> 90\%$ for tracks within jets in a pseudorapidity coverage of $|\eta| < 2.5$, along with a lepton resolution of $\Delta p_T/p_T = 10\%$.

The tracker can as well identify secondary vertices (SV) along the tracks, a feature used to identify jets arising from the hadronization process of b quarks, often referred to as *b jets* (see section 5.4).

The whole tracking system is divided into an inner pixel detector and an outer silicon strip tracker. Both tracker modules are based on doped silicon semiconductor diodes with embedded readout chips covering an angular acceptance of $|\eta| <2.5$. Its technology provides modules with a sufficiently high granularity and a fast response time that are also able to operate in high radiation environments. Particles are detected via the ionization of the silicon cells and the trajectories are then built associating different hits, their momentum determined from the curvature of their associated trajectories.

The pixel system is responsible for an optimized impact parameter resolution in both $r-\theta$ and z directions, that allows for a three-dimensional vertex reconstruction, therefore playing a key role in recovering secondary vertices from b and τ decays, even when their track multiplicities are low. It is also designed to separate tracks coming from the main interaction or *primary vertex* (PV) from the tracks associated

to additional vertices, referred to as *pileup* vertices (PU). It is segmented in 66 million pixels distributed in three barrel layers and two forward disks in each of the endcaps, that have a spatial resolution of 10 μm in the transverse plane, and about 20 μm along the z-axis. At the beginning of 2017, the pixel detector has been upgraded to improve the efficiency and resolution on tracks and allow recovery of degradation in outer tracker layers. Its geometry has thus been modified and a fourth additional layer in the barrel and an additional third disk in the endcaps have been added [60].

The silicon strip tracker that surrounds the pixel detector is divided into four different parts, as can be seen in figure 4.5: the Tracker Inner Barrel (TIB), the Tracker Outer Barrel (TOB), two Tracker Inner Disks (TID) and the Tracker Endcap (TE).

An overview of the tracker system of CMS can be seen in figure 4.5.

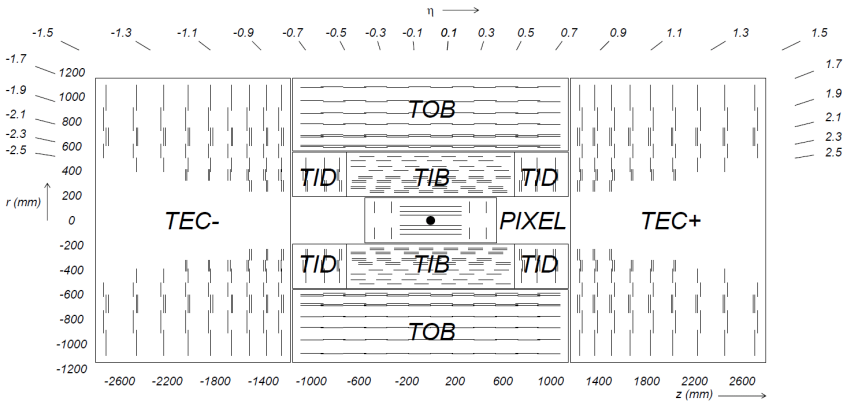


FIGURE 4.5: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

4.5 Electromagnetic Calorimeter

The CMS Electromagnetic Calorimeter (ECAL) [61] encloses the tracking system and is made up of 75848 lead tungstate ($PbWO_4$) scintillator crystals to detect electromagnetic showers originating from charged or neutral particles, such as electrons and photons, and measure their energy. It was specifically design to provide a fast response, have fine granularity to provide good energy resolution and minimize the probability of misidentifying jets as either electrons or photons, and to be radiation resistant in order to survive the LHC environment.

The ECAL covers two different regions in $|\eta|$: the barrel pseudorapidity coverage goes up to $|\eta| < 1.479$, and the endcap calorimeters cover the outermost regions

$1.479 < |\eta| < 3.0$. Both regions are made of lead tungstate $PbWO_4$ crystals. The compactness and fine granularity of the electronic calorimeter result from the properties of this material, which has short radiation length ($X_0 = 0.89$ cm) and *Moliere* radius (spatial extension of an electromagnetic shower inside a crystal) ($R_M = 2.2$ cm), as well as a high density (8.28 g/cm^3) value, making it optimal for electron and photon reconstruction. This material also benefits from a short scintillator decay time, which ensures that 80% of the light is emitted within the 25 ns interval between consecutive bunch crossings.

The ECAL is arranged as in figure 4.6. The barrel region (EB) is made of a total of 61200 crystals and is structured in 36 identical "supermodules", covering a volume of 8.14 m^3 and with a total weight of 67.4t. Each of the EB supermodules covers half the barrel length ($0 < |\eta| < 1.479$) and has a total of 1700 crystals. These are 23 cm long (corresponding to 25.8 X_0) and have a cross section of 2.2×2.2 cm^2 in the transverse plane.

A total of 7324 similar shaped crystals are placed in each of the endcaps (EE). These crystals, grouped in 156 mechanical units of 5×5 crystals (supercrystals or SCs), have a length of 22 cm (24.7 X_0) and a front cross section of 2.82×2.82 cm^2. The endcaps crystal volume is 2.9 m^3 and the weight is 24.0t.

Since the hadronic activity is particularly high in the forward region, the ECAL is completed by a preshower sampling calorimeter (ES), placed in front of the two endcaps, covering a pseudorapidity region of $1.652 < |\eta| < 2.6$. It consists on alternate layers of lead and silicon strips. The principal aim of the preshower detector is to identify neutral pions and improve the determination and position resolution of the showers from electrons and photons.

The relative resolution of the ECAL depends on the energy deposit and is approximately given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E/\mathrm{GeV}}}\right)^2 + \left(\frac{0.12}{E/\mathrm{GeV}}\right)^2 + (0.30)^2 \tag{4.6}$$

In this expression, the first term corresponds to the stochastic term which takes into account the fluctuation in the number of produced photo-electrons, the second describes the noise from the electronics and pileup events, and the final term is a constant term which accounts for the non-uniformity of the crystal response and instrumental errors.

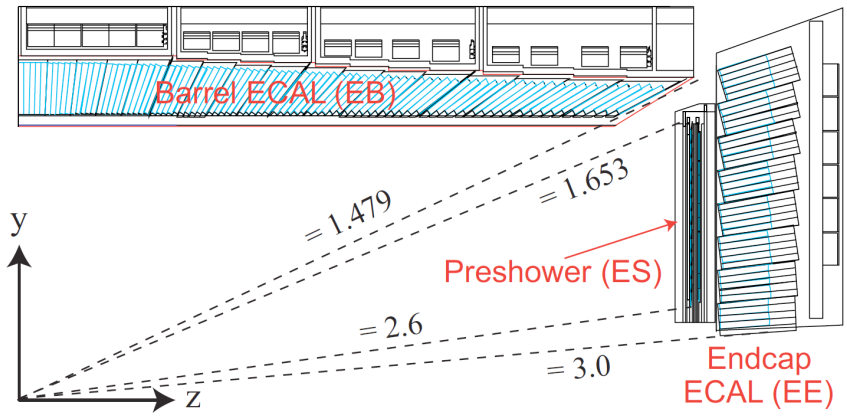


FIGURE 4.6: Longitudinal view of the CMS electromagnetic calorimeter.

4.6 Hadronic Calorimeter

The hadronic calorimeter (HCAL) [62] is the last CMS subdetector system located within the solenoid coil and aims for the measurement of the energy deposited by neutral and charged hadrons, as well as to estimate the missing transverse energy. Surrounding the ECAL, its design is influenced by the choice of magnet parameters and maximizes the amount of material within the magnet coil.

Hermeticity is essential to accurately measure the missing energy of neutrinos or other exotic particles. The central barrel and endcap HCAL subdetectors are fully immersed in the magnetic field, and cover up to $|\eta| = 3.0$. Since the magnet restricts the size of the barrel subdetector to a radius $\rho = 2.95$ m, an outer hadron calorimeter or tail catcher is placed outside the solenoid complementing the barrel calorimeter. In the forward direction, a Cherenkov-based, radiation-hard technology detector placed at 11.2 m from the interaction point extend the pseudorapidity coverage down to $|\eta| = 5.2$.

The HCAL is divided into four subsystem of sampling calorimeters:

- **Hadronic barrel** (HB): covers the pseudorapidity range $|\eta| < 1.3$ and is segmented into 36 wedges of 26t each, disposed in two half-barrel sections (HB+ and HB-). Each wedge is segmented in to four ϕ sectors. Each wedge is formed by flat brass absorber plates interleaved with the scintillator plastic material.

- **Hadronic endcap** (HE): covers a substantial portion of pseudorapidity ($1.3 < |\eta| < 3$) and is designed to have high radiation tolerance (this region contains about 34% of the particles produced in the final state). 79mm-thick brass

plates are intercalated with 9mm gaps to accommodate the scintillator material and achieve a granularity identical to the one in the barrel up to $|\eta| < 1.6$, which increases at larger pseudorapidities.

- **Hadronic forward** (HF): an additional calorimeter system, located 11.2 m away at both sides from the interaction point and covering a pseudorapidity range of $3 < |\eta| < 5.2$. The HF incorporates steel as absorber and quartz fibers as scintillator material. This design provides good radiation resistance, due to the very high hadronic activity in this region. The HF is of particular importance in the search for single top t-channel events, which feature a characteristic forward recoiling jet, that usually falls within its acceptance.

- **Hadronic outer** (HO): placed outside of the solenoid magnet, and utilizing it as an absorber, aims to provide adequate sampling for $|\eta| < 1.3$. It consists of one or two layers of scintillators, depending on the pseudorapidity, which match the granularity of the HB. Located within the muon system of CMS, its design and geometry are strongly influenced by that of the muon detector system. It is segmented in five consecutive rings, each covering 2.5 m along the beam line.

An overview of the CMS hadronic calorimeter and its different parts can be seen in figure 4.7.

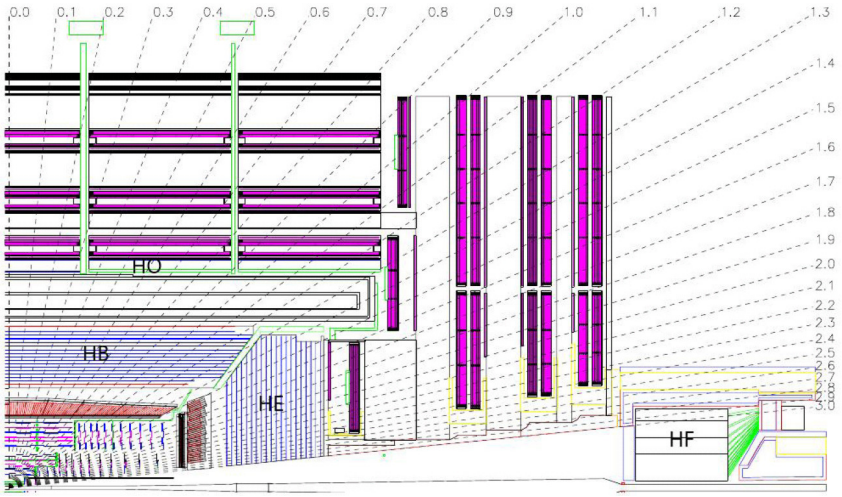


FIGURE 4.7: Longitudinal view of one quarter of the detector in the transverse plane, showing the positions of the HCAL parts: hadron barrel (HB), hadron outer (HO), hadron endcap (HE) and hadron forward (HF).

4.7 Muon System

The outermost part of the CMS detector consists of a set of detectors that aim to identify and improve the momentum resolution of muons traversing all the detector material [63].

Muons behave as minimally ionizing particles, and are the only ones (along with undetected neutrinos) that can pass through all the detector material, its detection therefore being relatively straightforward. Muons are a key ingredient in some SM and new physics processes.

As implied by the experiment's name, muon detection is of key importance within CMS's physics programme. The main tasks of the muon detector system are muon triggering, identification and momentum measurement. With this aim, three different detector technologies are used: the drift tube (DT) chambers, located along the barrel, the Cathode Strip Chambers (CSC) and the Resistive Plate Chambers (RPC). These are all sets of gaseous detectors, and the choice of the different technologies used respond to the different radiation environment to be covered in each case.

- **Drift Tube Chambers (DT)**

  The basic element of the drift tube is the drift cell, which is filled by a gas mixture of 85% Ar and 15% CO_2 and contains a 50μm thick and 2.4m long gold-plated steel wire spanned through the centre.

  Placed in the barrel region forming five consecutive rings or wheels along the beam axis [YB-2, YB-1, YB0, YB+1, YB+2], up to $|\eta|$ <1.2, the DT detectors receive a relatively low muon flux. Each of these wheels, with a longitude of 2.7 m, is itself segmented in 12 sectors along the ϕ coordinate in the transverse plane. These ϕ-sectors, each covering $\sim 30°$, are organized in four muon barrel (MB) DT stations located disposed at different radial distances from the beam axis: MB1, MB2, MB3 and MB4, this last being the outermost at about 7m from the beam axis. An outlay of the muon DT chamber distribution is displayed in figure 4.8.

  The three innermost DT stations (MB1, MB2 and MB3) have three *superlayers* (SL), the outer ones having their wires parallel to the beam axis to measure trajectories along the (r, θ)-coordinates, the innermost superlayer having them placed orthogonal to measure trajectories in the (r, z)-coordinates. This central superlayer is not present in the outer DT stations (MB4).

- **Cathode Strip Chambers (CSC)**

  Located in the endcaps, for 0.9< $|\eta|$ < 2.4, the Cathode Strip Chambers (CSC) are designed to resist the higher radiation levels and magnetic field present in this region. These are trapezoidal-shaped chambers installed in four disks per side, each disk containing 18 or 36 chambers disposed such

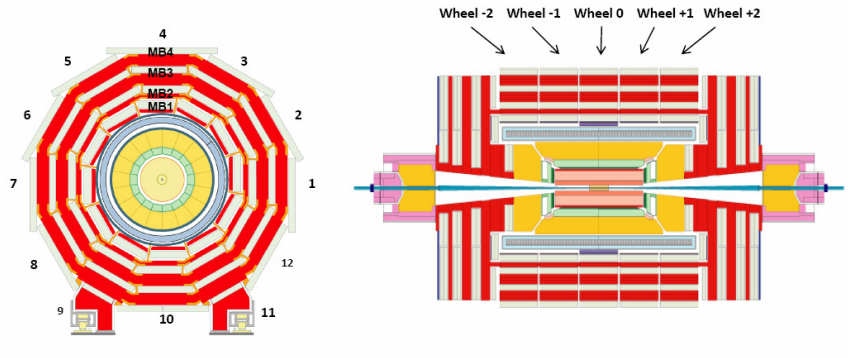


FIGURE 4.8: Transverse (left) and axial (right) view of the CMS Drift Tube Chamber distribution. The barrel is divided in five wheels [YB-2, YB-1, YB0, YB+1, YB+2] segmented in 12 sectors, each of them composed by 4 barrel stations (MB1, MB2, MB3, MB4).

that a full azimuthal coverage is achieved amounting to a total of 468 CSCs. All chambers, except those in the innermost station, are overlapped in the azimuthal angle ϕ to avoid gaps in the muon acceptance.

The CSCs are six-layered gaseous chambers filled with a mixture of Ar, CO_2 and CF_4 in which closely spaced anode wires are stretched between two radially oriented cathodes. The measurements from the cathodes are used to estimate the (r, ϕ) position, while the anodes are optimized to measure the η-coordinate and the bunch crossing the detected muon comes from. Thus, each CSC measures the space coordinates (r, ϕ, z) in each of its 6 layers.

- **Resistive Plate Chambers (RPC)**

  The set of muon detectors is completed by the Resistive Plate Chambers (RPC), a set of gaseous parallel-plate detectors that provide a fast and independent trigger over a large portion of the rapidity of the muon system ($|\eta| < 1.6$).

  Six of them are placed in the CMS barrel, two in each of the two first stations, to guarantee the good functioning of the trigger for low-p_T tracks that may not reach the outer chambers, and one in each of the last two stations. In the endcap region, there is a plane of RPCs in each of the first 3 stations intended for using the coincidences between stations to reduce background, to improve the time resolution for bunch crossing identification, and to achieve a good p_T resolution at trigger level.

  Despite having a coarser spatial resolution, the RPC system provides a much shorter time resolution (of about 1 ns) than the DTs and CSCs, and is therefore used to associate each muon signal to its corresponding bunch crossing.

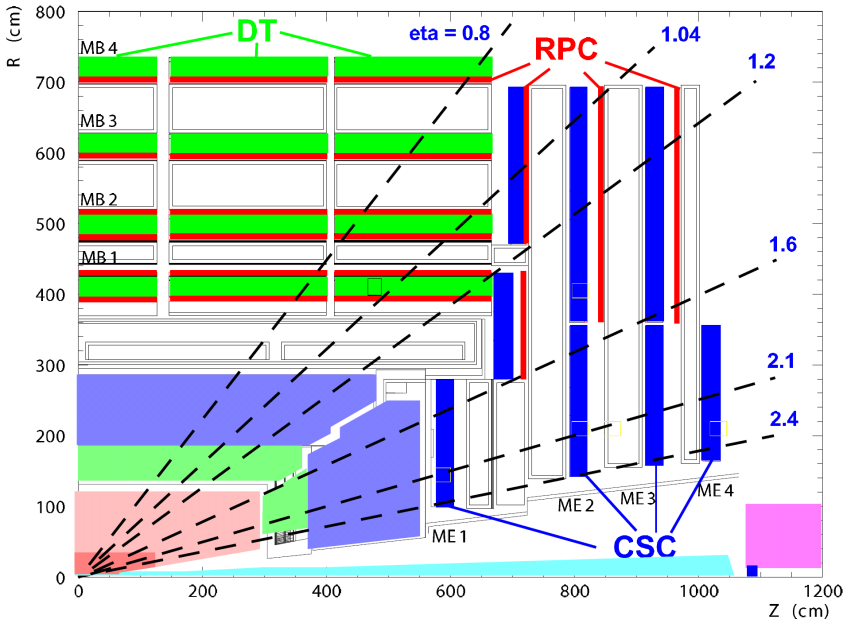


FIGURE 4.9: A longitudinal view of the muon system indicating the location of the three detector types contributing to the muon spectrometer.

4.8 Trigger system and data acquisition in CMS

At the nominal LHC design luminosity of 10^{34} cm^{2}s^{-1}, CMS will get an input rate of 10^9 interactions per second, which needs to be reduced by a factor 10^7 in order to make it possible to technologically store all the data produced everyday. The task of reducing this rate is accomplished by the CMS trigger system, which represents the first step of the physics event selection.

To select events of potential physics interest, the CMS trigger system [64] utilizes two levels: the first level (L1) trigger and the high level trigger (HLT). The first level is implemented in custom hardware and selects events containing candidate objects, which will be passed to the high level trigger. The aim of the L1 trigger is to reduce the data rate by an average factor of 400, from the input rate of 40MHz to 100 kHz.

The HLT is implemented in software and consists on a multi-stage iterative algorithm conducted on a farm of computers. It uses full detector information to reproduce the L1 trigger decision and then to iteratively improve on this decision by the staged introduction of calorimetry and tracking information. If, at any stage, the event is rejected, the process is halted and the processing resources freed for handling the next event. A time cut-off is applied to prevent resources becoming locked. The aim of the HLT is to further reduce the data rate so that it can then be stored at approximately 100 Hz.

These two systems will be described in the following sections.

4.8.1 Level 1 Trigger

The L1 trigger involves the calorimetry and muon systems as well as some correlation of information from them, and is hardware-implemented. Its main purpose is to reduce the number of events to be processed by the high level trigger, tentatively deciding whether or not an event should be accepted with information from the calorimeter and muon detectors only. It is therefore organized into three major subsystems

- L1 calorimeter trigger
- L1 muon trigger
- L1 global trigger

The L1 muon trigger is itself further organized into subsystems representing the different muon detector systems (drift tube trigger, cathode strip chambers trigger and resistive plate chamber trigger) in addition to a global muon trigger that combines information from each of the muon sub-systems. The calorimeter and muon triggers do not perform any selection themselves, but rather they identify trigger objects (electrons, photons, jets, muons) and send the best candidates along with the corresponding kinematical information to the L1 Global Trigger, responsible for combining the output of L1 Calorimeter Trigger and L1 Muon Trigger and for making the decision to either retain the event or discard it. Thus, the only selection is done at the L1 global trigger. A schematic view of the L1 trigger system can be seen in figure 4.10

The L1 trigger rate is limited by the speed of the detector electronics readout and the rate at which the data can be harvested by the data acquisition (DAQ) system.

The L1 global trigger (GT) is the final step of the CMS level 1 system and implements a menu of triggers, a set of selection requirements applied to the final list of objects (i.e., electrons/photons, muons, jets, or τ leptons), required by the HLT algorithms to meet the physics data-taking objectives. This menu includes trigger criteria ranging from simple single-object selections with E_T above a preset threshold to selections requiring coincidences of several objects with topological conditions among them. A maximum of 128 separate selections can be implemented in a menu.

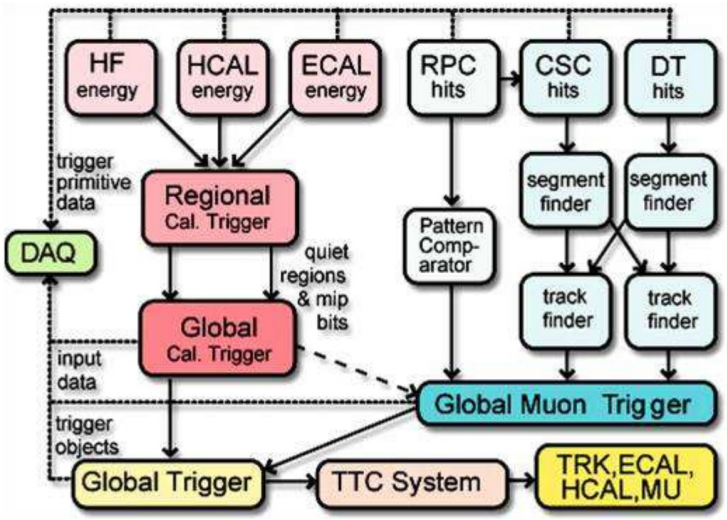


FIGURE 4.10: Overview of the level 1 trigger system. The L1 global trigger combines information from the L1 calorimeter trigger and the L1 muon global trigger.

The L1 calorimeter trigger

The calorimeter trigger uses raw data signals to identify isolated and non-isolated electrons and photons, along with central, forward and tau jet candidates. The main difficulty in performing this task is coping with the raw data rate in the calorimeters. This is achieved in two stages: first, 5×5 ECAL crystal groups are combined into *trigger towers* mapped to HCAL towers, and these trigger towers are then summed in a second step into regions whose size is closer to the natural scale of the trigger objects and are classified according to geometric and quality criteria.

The L1 calorimeter trigger comprises two stages: a regional calorimeter trigger (RCT) and a global calorimeter trigger (GCT). The regional trigger identifies all candidates along with their transverse energy, and sends this information to the global calorimeter trigger, which sorts the candidates according to their transverse energy and quality, and sends the first four objects to the L1 Global Trigger for further processing.

The L1 muon trigger

The primary task of the muon trigger is to identify muon candidates and reconstruct their momentum and transverse energy, along with information on the associated bunch crossing. All three detector sub-systems (drift tubes, cathode strip chambers and resistive plate chambers) contribute to the muon trigger, which benefits both from the good spatial resolution from the drift tubes and the cathode strip chambers, along with the excellent time resolution of the resistive plate chambers.

The redundancy of the muon system allows for a high performance trigger with an outstanding background rejection.

In the RPC system, the trigger electronics build track segments which are sent along with their reconstructed p_T to the Global Muon Trigger. It also provides the CSC trigger system with information to solve hit position ambiguities in case several muon tracks traverse a same chamber.

The CSC trigger builds local charged tracks (LCT) using cathode strip information only, and assigns them an estimate of their p_T and a quality flag. The best charged tracks are then passed to the CSC Track Finder, which builds tracks using the full CSC information. These tracks are then sent to the Global Muon Trigger.

The DT trigger is equipped with a Track Identifier system, which makes up segments from sequences of hits within a same superlayer. This segments are then sent to the DT Track Correlator which combines information from different superlayers. The best combined segments are sent to the DT Track Finder, which builds track candidates and passes the information to the Global Muon Trigger.

At the end, the Global Muon Trigger system is responsible for combining the different sub-detector information to build candidates. After sorting them according to their momentum and quality, the best four tracks are passed to the L1 Global Trigger.

The L1 global trigger

Based on trigger objects delivered by the global calorimeter and muon triggers, such as isolated or non-isolated electrons and photons, muons, central and forward jets and global event quantities (such as the missing energy in the transverse plane), the L1 Global Trigger makes the decision to either accept or reject an event before delivering the information to the high level trigger.

These trigger objects are synchronized to each other and to the LHC orbit clock and then sent to the global trigger logic (GTL) module, where the trigger algorithm calculations are performed. This algorithm imposes certain conditions on the received objects, such as the total E_T or p_T being above a given threshold, angular positions being within a selected window, or imposing certain objects to be separated in an event. These conditions are combined by simple combinatorial logic to form up to 128 different algorithms, which form a given L1 menu. These algorithms are finally combined into a final "OR" by the final decision logic module, and can be prescaled or blocked. These prescales allow to adjust the trigger rate for a certain algorithms. The set of algorithms in the menu along with the prescaled values completely define the L1 trigger selection. On a final step, the information is forwarded to the high level trigger for further scrutiny.

4.8.2 High Level Trigger

The high level trigger [65] aims to maximize the efficiency while keeping CPU time and rate low, it must be flexible to adapt to changes in data-taking conditions, have a robust performance with respect to changes in alignment and calibration conditions and be stable with respect to pileup. The HLT algorithms use the same software framework and most of the reconstruction code used for offline reconstruction and analyses.

The HLT is designed to reduce the L1 trigger input rate down to 1 kHz, compatible with the data acquisition capabilities, which is the amount that will be written to mass storage. The HLT itself is fully software implemented and runs on a farm of commercial processors.

As the HLT receives a lower event rate consequently having more time to process them (about 300 ms), it can go into finer details and use more refined algorithms closer to the offline reconstruction profiting from information from all subdetectors.

The reliability of the HLT is of key importance, as any event discarded will be subsequently lost. The still short time available for event processing imposes constraints on its design. Computationally sophisticated algorithms must only be run on good candidates for interesting events, and with this view, the HLT is organized in different sub-levels, each of which reduces the number of events to be processed in the following step. These different steps correspond roughly to what would have been distinct trigger systems; however, the CMS HLT architecture does not include a sharp division between these trigger steps, other than the order in which they are applied.

- **Level 2**: first selection step of the HLT process, receiving the maximum rate of events. It receives L1 candidates as input and uses complete information from the muon system and the calorimeters only, reducing the number of events to be processed by the following sub-levels.
- **Level 2.5**: intermediate step that adds pixel information and provides fast confirmation of the L2 candidates before delivering the information corresponding to the accepted events to the L3 step.
- **Level 3**: receives a reduced rate of events, and makes use of the full information provided by the tracking detectors to either discard or accept an event.

Computationally expensive and time consuming tasks, such as track reconstruction (due to the high number of channels, complex pattern recognition and higher combinatorics), are executed in the level 3 only in the region of interest. However, track reconstruction is performed only in a limited set of hits and is stopped once

the required object resolution is reached, since the ultimate precision is not required at this stage, but rather in the offline reconstruction to achieve the highest possible accuracy in the different physics analyses.

Data processing is structured around the concept of *HLT path*, a sequence of algorithms processed in a predefined order of increasing complexity that both reconstructs physics objects and applies selections on these objects. The reconstruction at this point is only partial, involving the reconstruction of particles only in a limited region of the detector, to minimize the CPU time required by the HLT. Full event reconstruction of events delivered by the high level trigger takes place offline and will be depicted in the following chapter. The size of the events reconstructed by the HLT is significantly smaller (typically about 1.5 kB per event) that those reconstructed offline (around 0.5 MB).

Accepted events are sent to the storage manager, where data is first stored locally on disk and eventually sent to the CMS Tier-0 computing centrefor offline processing and permanent storage. Even though most data is analyzed as soon as possible, lower-priority *parked* data is not analyzed until the end of the running period.

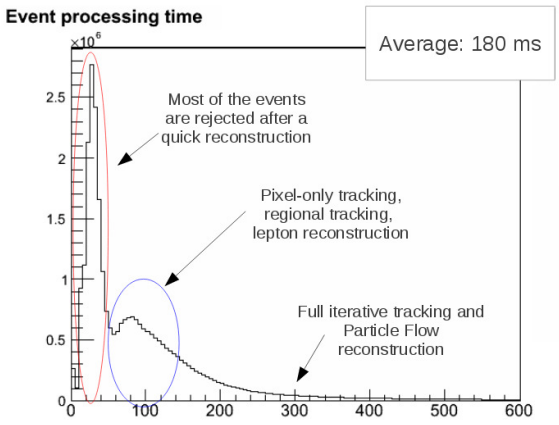


FIGURE 4.11: Processing time for events in CMS. The data shown correspond to a short run taken in November 2012.

4.9 Computing at CMS

The CMS software and computing system covers a wide variety of of activities, such as data storage, distribution and processing, event reconstruction, generation of Monte Carlo samples.

4.9.1 CMS data hierarchy

CMS data is arranged in a hierarchical manner, consisting on a number of event data formats with varying degrees of detail, size, and refinement. Starting from the raw data produced from the online system, successive degrees of processing refine this data, apply calibrations and create higher level physics objects. Each physics event is written in each of the different available CMS data formats.

4.9.2 Tier system

The computing centres available to CMS around the world are distributed and configured in a tiered architecture that functions as a single coherent system. Each Tier is made up of several computer centres and provides a specific set of services; they process, store and analyze all the data from the LHC.

Tier 0

The first tier in the CMS model [66] is the CERN Data Centre, which is located in Geneva, Switzerland. All data from the LHC passes through the central CERN hub, but CERN provides less than 20% of the total compute capacity. The Tier-0 does not provide analysis resources, but instead performs several functions (accept online raw data system and distribute it into primary datasets based on trigger information, perform a prompt calibration and object reconstruction from the raw data, and distribute these raw and reconstructed datasets to Tier-1's).

Tier 1

These are thirteen large computer centres with sufficient storage capacity and with round-the-clock support for the Grid. They are responsible for the safe-keeping of a proportional share of raw and reconstructed data, large-scale reprocessing and safe-keeping of corresponding output, distribution of data to Tier 2s and safe-keeping of a share of simulated data produced at these Tier 2s.

Tier 2

The Tier-2's are typically universities and other scientific institutes with substantial CPU resources, which can provide sufficient but limited disk space for data storage as well as adequate computing power for specific analysis tasks, calibration studies and Monte Carlo production. The activities related to the Tier-2's are organized by the responsible of each centre in collaboration with physics groups. As in 2018, there are 15 Tier-1's and up to 153 Tier-2 centres.

4.9.3 CMS software framework (CMSSW)

The CMS experiment uses its own software framework, called CMSSW, for data acquisition, triggering, reconstruction and analysis of CMS data. CMSSW is based on C++ and python languages. The CMS Event Data Model (EDM) is based on the concept of *events*. An event is a C++ class that contains all the raw and reconstructed information of the physics event from the collision. Additionally, every event has relevant information related to the configuration of the software used and the CMS calibration conditions. Events can be accessed and are written to ROOT files, which are are organized in *trees* in which information from all events is collected. ROOT [67] is an analysis package written in a C++ object-oriented structure. All the histograms and files with data objects are produced with user-compiled codes built with ROOT functions.

CMS consolidates its code regularly into *releases*, which are constantly revised and developed in order to aggregate specific features. To guarantee stable releases while supporting fast development, CMS implemented an advanced central release validation process.

Chapter 5

Event reconstruction

The reconstruction of basic analysis objects, such as leptons and jets, within an event is described in this chapter. A key ingredient in the event reconstruction of CMS is the particle flow (PF) algorithm. It creates particle candidates by combining various subdetector information for a global event interpretation which improves the identification, spatial resolution, and accuracy in the energy measurement of the different objects under study. Jet reconstruction from PF particles is also described. The focus of this chapter is set on the reconstruction of muons, electrons, jets, missing transverse energy, as well as heavy-flavour jet identification in CMS.

In order to perform cutting-edge physics analyses, particle detectors like CMS aim at achieving a meticulous knowledge of all objects of interest in the event. One of the main challenges is the reconstruction of the hadronic jets produced in the collisions. To achieve high level of precision, information coming from the calorimeters only does not suffice and it has to be combined with that from the tracker, which offers a better resolution for the measurement of charged particles' momenta, improving, as a direct consequence, jet reconstruction.

The main requirements of the reconstruction process involve achieving the highest possible efficiency and having a good resolution, while keeping a low fake rate (comprising *misreconstructed* objects). The algorithms must also be robust against detector problems and have almost no dependency on data-taking conditions, such as noise contamination, dead regions of the detector or possible increases in pileup conditions. Special interest is also set in minimizing CPU time consumption per event processed and keeping the memory usage to its possible minimum.

With this view, CMS has developed a set of dedicated algorithms which are periodically improved in order to satisfy these necessities. The building bricks in particle reconstruction are

- Charged particle tracks using information from the tracker system, used both in muon and electron reconstruction.

- Muon tracks reconstructed in the muon system.
- Energy superclusters made from clusters of energy deposits in the calorimeters.

These initial objects are then used to construct particle candidates, and can be matched and combined to form the different particle objects. Tracks and clusters are only used once in the event, in order to avoid double counting of particles. This information is then combined and sorted resulting in a list of physical objects for each event. This is done by the PF algorithm, described in section 5.2. Reconstruction and identification of electrons and muons using information from the tracker, calorimeters and muon chambers is described in 5.1. At the end of this chapter there is a short introduction to jet reconstruction (section 5.3), b jet identification along with a brief introduction to the different available algorithms (section 5.4) and the reconstruction of the missing transverse energy (section 5.5) using the PF algorithm.

5.1 Lepton reconstruction and identification

In this first section, the basics for lepton reconstruction are provided. Reconstruction of τ-leptons is not described as they are not directly used in the analysis, but information regarding this topic can be found elsewhere [68, 69].

5.1.1 Muon reconstruction

CMS aims at achieving an optimal muon reconstruction at a wide range of energy values: from low energetic muons as in B physics studies, to relatively energetic muons from electroweak processes up to very energetic ones, which could be a signature of new physics phenomena. Muons of interest for this thesis usually lay on an intermediate energy range between the two ends of the spectra.

Having the highest identification efficiency thanks to the three muon spectrometer subsystems (DT, CSC and RPC), muons are the first particles reconstructed by the particle flow algorithm. Muon tracks are reconstructed in a multi-faceted way, both in the tracker system and the muon chambers.

The reconstruction in the muon spectrometer starts with the reconstruction of hit positions in the DT, CSC and RPC subsystems. Hits within each DT and CSC chamber are then matched to form *segments*. The segments are collected and matched to generate *seeds* that are used as a starting point for the actual track fit of DT, CSC and RPC hits. The result is a reconstructed track in the muon

spectrometer, and is called *standalone muon track*. Standalone muon tracks are then matched with those reconstructed in the silicon tracker system (called *tracker tracks*) to generate *global muon tracks*, featuring the full CMS resolution. The final collection of high-level muon physics objects is comprised of three different muon types:

- *Standalone muons*: reconstructed from tracks in the muon spectrometers, which are extrapolated to the point of closest approach to the beam line. Their momentum resolution is improved by the application of a beam-spot constraint.
- *Tracker muons*: objects reconstructed by an algorithm that starts from reconstructed tracks in the tracker and looks for compatible segments in the muon calorimeters.
- *Global muons*: assembled by matching tracks in the muon system and the inner tracker.

which will be part of the PF objects used for analysis.

Two reconstruction approaches are used within CMS:

- **Global muon reconstruction** (*outside-in*): for each standalone track, a matching tracker track is searched for. Track parameters are extrapolated to a common surface where they are compared. A global muon track is then built combining hits from the tracker track and the standalone muon track, and fitted using a Kalman filter. This approach is particularly interesting for energetic muons ($p_T > 200$ GeV), for which the fit can improve momentum resolution compared to the tracker-only fit.
- **Tracker muon reconstruction** (*inside-out*): each tracker track with $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV is considered a possible muon track candidate and is extrapolated to the muon system, taking into account the effect of the magnetic fields, energy losses during the time of flight and multiple Coulomb scattering in the detector material. If at least one muon segment matches the extrapolated track, the corresponding track qualifies as tracker muon. Tracker muon reconstruction is more efficient for low-momentum muons ($p_T < 5$ GeV).

A more detailed description of the reconstruction process is presented in the following sections. The description of how the matching is done in order to reconstruct global muon objects will also be presented.

Muon reconstruction in the muon system

The muon reconstruction starts at the level of the individual chamber. The results are track segments in the DTs and in the CSCs, and three-dimensional points in the RPCs. This process is performed in four different stages, separately in each subsystem:

1. Seed generation or *seeding*: a search for initial track segments or *seeds* is performed on the first step.
2. Trajectory building: starting from seeds, track candidates are grown layer by layer.
3. Trajectory cleaning: duplicate tracks and bad fits are removed at this step.
4. Trajectory smoothing: a refit is performed on the remaining tracks to optimize track parameters.

Trajectory seeds are the starting point in track reconstruction, and they consist on position and direction vectors and an estimate of the muon transverse momentum. Two types of trajectory seeds are considered: *hit-based* and *state-based* seeds which take, respectively, hit doublets or triplets compatible with the beam spot, or a trajectory state on a detector to define the initial position and direction of the seed.

Trajectories are then built starting from the position specified by the corresponding seed, and the search for compatible measurements in the subsequent detector layers proceeds in the direction defined by this seed. A pre-filter is often applied in order to reduce possible bias from the seed. The full knowledge of the track parameters at each detector layer is used to find compatible hits in the next layer, and the propagated trajectory state is then updated accordingly using the information from the last added hits. The process is iterated until the outermost compatible layer of muon detector is reached. A suitable propagator, accounting for material effects such as multiple scattering processes and energy losses during ionization or bremsstrahlung, and fast enough so as to reduce processing time during reconstruction, is needed at this step.

Many tracks are produced during the building process. Ambiguities in track finding arise when the same track is reconstructed twice starting from different seeds or when a given seed develops into more than one final track candidate. These ambiguities must be resolved in order to avoid that the same charged particle is counted twice, or even multiple times. A dedicated cleaning algorithm determines the fraction of shared hits between two track candidates. If half of the hits are shared among mutually exclusive tracks, the algorithm removes the one with the smallest number of hits from the candidates collection. Alternatively, if both tracks have

the same number of hits, the one with highest χ^2 value is discarded. The procedure is repeated iteratively until all the pairs of track candidates in the collection share less than a 50% of their hits.

At the smoothing stage, a backward Kalman filter is applied on all remaining tracks in the outside-in direction down to the innermost detector layer. At each hit the updated parameters of this second filter are combined with the predicted parameters of the track building pre-filter. At this point the trajectory is finally built and the track parameters defined.

Muon reconstruction in the silicon tracker

The reconstruction process in the tracker system is an iterative process, carried out in separate steps:

- Seeding: starting from a few hits, initial track candidates are provided, along with an initial estimation of the trajectory parameters.
- Track finding (pattern recognition): extrapolates the seed trajectories along the expected flight path of a charged particle, searching for additional hits that can be assigned to the track candidate.
- Track fitting (final track fit): provides the best possible estimate of the trajectory parameters, using a Kalman filter and trajectory smoother.

Seed finding (or seeding) comprises the search for the starting points of tracks in the innermost detector parts. Seeds are made out of hit pairs or triplets, which must be pointing towards the interaction point and pass a configurable minimum p_T threshold, in order to limit the available hit combinations. Seeds contain enough information to define the starting trajectory parameters and associated uncertainties of potential tracks. Due to the uniformity of the magnetic field in the region, charged particles in the tracker follow helical paths, and therefore five parameters are needed to define a trajectory.

Once candidate seeds are found, track building is performed outwards (starting from the innermost layers to the outermost ones). Track building is based on the combinatorial track finder algorithm, an adaptation of the combinatorial Kalman filter, to allow pattern recognition and track fitting to occur within the same framework. Parting from the seed, the initial estimated trajectory is extrapolated to find additional hits in subsequent layers which are compatible with a particle track hypothesis, layer by layer. After adding a hit to the trajectory candidate, the algorithm updates the associated track parameters and their uncertainties. If multiple compatible hits are found on a given layer, the trajectory is cloned for each of them. However, if no hits are found within a layer, a ghost hit is created. After each

iteration, hits associated with tracks are removed, thereby reducing the combinatorial complexity and simplifying subsequent iterations. This process is performed iteratively until the outermost layer is reached or a stopping condition is satisfied.

Track fitting is performed using a Kalman filter. The fitting procedure is done iteratively in two separate steps: a first filter is applied over the full list of hits from the inside outwards updating the track trajectory estimate sequentially with each hit. Once the outermost layer is reached, a second filter is initialized with the result of the first filter (*smoothing*), which runs backward towards the beam-line. The resulting track parameters are obtained from the average of the track parameters of the two filters. To obtain the best precision, this filtering and smoothing process uses a *Runge-Kutta propagator* to extrapolate the trajectory from one hit to the next, which takes into account any effects the traversed material and the magnetic field might have on the trajectory.

In a final step, fitted track candidates are assigned different quality flags and are required to pass a quality selection to reduce the amount of fake tracks before they are considered in physics analyses. The criteria reflect the seed requirements and depend additionally on the total number of fitted hits, the χ^2/ndof of the fit, the amount of ghost hits, and the amount of shared hits with other tracks, amongst others. The result is a collection of muon tracker tracks.

Global muon reconstruction: track matching

Global muon reconstruction is performed associating muon tracks in the silicon tracker detector to standalone trajectories in the muon chambers, looking for compatibilities in momentum and position. Due to the large multiplicity of tracks in the silicon detector, a two-step process is used to minimize the number of possible tracker tracks associated to a single standalone trajectory.

In the first of these stages, a rectangular region in the $\eta - \phi$ space, referred to as the *region of interest* (*ROI*), is defined, which contains a subset of all possible track candidates. The selection of the region of interest has a strong impact on the reconstruction efficiency, fake rate, and CPU reconstruction time. This region is defined by a set of parameters related to the trajectory of the standalone muon (including direction and minimum p_T). The spread in $\Delta\eta$ and $\Delta\phi$ is extracted from the error estimates of the standalone muon direction.

With the tracking region of interest defined around the standalone muon, the matching algorithm iterates over all reconstructed tracker tracks and chooses a subset of tracks that are within this region. The collection of regional tracker tracks are then compared to the standalone muon track using more stringent matching criteria.

After the selection of a subset of tracker tracks that match the standalone muon track, the next step in making a global muon track is to fit a track using the hits from the tracker track and the standalone muon track. The global refit algorithm attempts to perform a track fit for each tracker track - standalone muon pair. If, at the end, there is more than one possible global muon track, the global muon track with the best χ^2 is chosen. Thus, for each standalone muon there is a maximum of one global muon that will be reconstructed.

Further details on muon reconstruction in CMS may be found in [70, 71].

Muon isolation and identification

To distinguish isolated leptons from non-isolated ones, a cone is constructed around the lepton direction. An isolation variable is constructed from the scalar sum of the transverse energy of all reconstructed particles contained within the cone, excluding the contribution from the lepton candidate itself.

Each bunch crossing, several proton-proton interactions take place. This makes the assignment of the vertex of the hard-scattering process non-trivial. For each reconstructed collision vertex, the quantity

$$\sum_{tracks} p_T^2$$

(for all tracks associated to the vertex) is computed, and that for which this value is largest, is referred to as the *primary vertex* . Particles from vertices other than the primary vertex may lie within the cone and contribute to the isolation value of the lepton. In the case of charged particles, only those associated to the primary vertex are considered in the isolation calculation. Since neutral particles do not leave tracks in the silicon tracker, they cannot be assigned to vertices. Therefore a correction is applied to account for this effect, by subtracting the energy deposited in the isolation cone by charged particles not associated with the primary vertex, multiplied by a factor 0.5[1]. This is commonly known as *Delta-beta* or $\Delta\beta$ correction.

PF-based muon isolation is $\Delta\beta$ corrected for pileup mitigation, and is defined as

$$I_{\text{rel}}^{\mu,\text{PF}} = \frac{I_{\text{charged hadrons}} + \max(0, I_\gamma + I_{\text{neutral hadrons}} - \beta \cdot I_{\text{pileup}})}{p_T^\mu} \tag{5.1}$$

where $\beta = 0.5$, and $I_{\text{charged hadrons}}$, I_γ and $I_{\text{neutral hadrons}}$ denote the summed transverse energies of charged hadrons, photons and neutral hadrons, respectively, within a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ around the muon candidate. The term I_{pileup} is used to correct the amount of considered neutral energy. It denotes the

[1]This value is motivated by assuming a 2:1 ratio of charged:neutral production rates, as π^0, π^+ and pi^- production rates are assumed to be equal.

summed transverse energies of charged-particle tracks within the $\Delta R < 0.4$ cone that are associated to pileup vertices. $\beta \cdot I_{\text{pileup}}$ can be interpreted as an approximation to the amount of neutral energy coming from pileup interactions that contributes to the neutral component, I_γ and $I_{\text{neutral hadrons}}$. Different isolation working points can be used, depending on the requirements of the analysis. One may either want to get as many candidates as possible (thus loosening the isolation constrain) or have a more refined definition of the muon objects, requiring a tighter isolation even though that means loosing statistics:

- Loose: $I_{\text{rel}}^{\mu,\text{PF}} < 0.25$ with $\epsilon \simeq 0.98$
- Medium: $I_{\text{rel}}^{\mu,\text{PF}} < 0.20$
- Tight: $I_{\text{rel}}^{\mu,\text{PF}} < 0.15$ with $\epsilon \simeq 0.95$

Reconstructed muons are then classified in different categories (or identification *working points*) in the PF algorithm in terms of certain quality criteria, in order to address different requirements (to either achieve a higher efficiency or to reduce considerably the number of fake muons).

- **Loose muon**: particle identified as a muon by the Particle-Flow event reconstruction, and that is also reconstructed either as a global-muon or as an arbitrated tracker-muon. This identification criteria is designed to be highly efficient, for prompt[2] muons, as well as from muons from heavy and light quark decays.
- **Medium tagged muon**: loose muon with additional track-quality and muon-quality requirements. The fraction of valid tracker hits is also required to be above a certain threshold. This identification criteria is designed to be highly efficient for prompt muons and for muons from heavy quark decays.
- **Tight muon**: tight muons correspond to those having tracks reconstructed both in the muon system and the tracker (meaning only *global* muons are considered, see section 5.1.1), which reduces the contamination from muons produced in hadron decays and from *punch-through*[3] particles. To suppress punch-through and accidental track-to-segment matches, muon segments must be formed from at least two matched stations. Only quality muon tracks for which the global track fit has a goodness-of-fit $\chi^2/\text{ndof} < 10$ are selected. In addition, selected tracks are required to include at least one valid hit in the muon chambers and at least one hit in the inner pixel detector. Muon candidates are excluded if a minimum of five hits in the tracker is not reached. This particular cut guarantees that a good p_T measurement is achieved. Residual

[2]Prompt leptons are leptons originating from the main collision taking place in the event.
[3]Hadron shower remnants that reach the muon system, that may cause charged hadrons to be misreconstructed as muons.

contamination from cosmic muons and muons arising from pileup interactions is rejected by requiring a minimal distance of the muon track to the primary vertex of $|d_0| < 2$ mm and $d_z < 5$ mm. Muons passing these criteria are the ones that will be used in the analysis (see section 6.5.1 in the next chapter).

More categories exist within the CMS PF framework (as those devoted to high-p_T muon analyses) but they will not be used in the current analysis.

5.1.2 Electron reconstruction

Right after muons, electrons are the next reconstructed particles in the PF algorithm. Electron reconstruction in CMS uses combined information from the pixel detector, the silicon strip tracker and the electromagnetic calorimeter. As with muons, electron reconstruction depends on the p_T of the electron themselves. Electrons of interest for the current analysis lie on an intermediate energy range. They are reconstructed by the association of a track reconstructed in the inner detector to an energy cluster in the ECAL, from which the properties of the electron candidate are extracted. Devoted algorithms have been developed and optimized during the LHC data taking periods.

Electron reconstruction itself represents a challenging task, as electrons traversing the tracker suffer from considerable energy loss due to bremsstrahlung of photons as the electron travels through the detector material. The fraction of the initial energy that reaches the calorimeter has therefore a significant spread in the azimuthal ϕ direction (figure 5.1). Calorimeter clustering takes into account this energy spread and collects bremsstrahlung radiation. The electron track reconstruction relies on a dedicated *Gaussian Sum Filter* (GSF) [72] using a specific energy loss modeling.

For a complete description of electron reconstruction in CMS, see [73].

Electron clustering in the calorimeter

As mentioned before, electrons radiate photons when traversing the tracker material, and the energy reaching the calorimeter is spread in the azimuthal plane. This effect can be quite large in some cases. In fact, about a 35% of the electrons radiate more than 70% of their initial energy before reaching the ECAL, and in a 10% of the cases more than 95% of this energy is radiated. This effect can be seen in figure 5.2.

In order to achieve an accurate measurement of the electron's energy, it is essential to collect the energy of these radiated photons. With this purpose, different

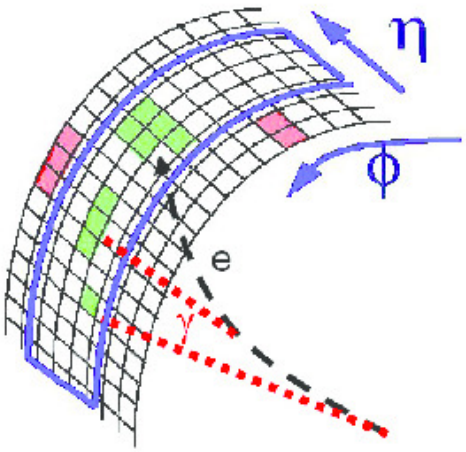


FIGURE 5.1: Typical energy deposit topology of an electron radiating bremsstrahlung photons. The squares represent the ECAL crystals while the green ones are those associated with a significant energy deposit. The electron and the bremsstrahlung photons produce each a cluster of crystals. The reconstruction procedure aims at gathering those clusters to form a supercluster.

algorithms have been developed within the CMS experiment. A set of dynamic clustering algorithms are used to merge clusters belonging to the same electromagnetic shower into so-called *superclusters* (SC).

In Run II, an alternative approach is used that is part of the PF reconstruction algorithm. In this approach, called 'mustache' clustering, clusters are reconstructed by grouping together all crystals contiguous to a seed crystal if their energy deposit is two standard deviations above the electronic noise. The requirement of a crystal to be taken as a seed is that its energy must be above these thresholds; $E^{seed} >$ 230 MeV in the barrel and $E^{seed} > 600$ MeV or $E_T^{seed} > 150$ MeV in the endcap regions [74]. This approach provides significant improvements to energy resolution with respect to previous algorithms.

Due to the increase in luminosity during the second run period, a devoted algorithm (the *Multifit* algorithm) has been developed to get rid of pileup contributions. Corrections accounting for crystal transparency losses during time are also taken into account. For endcap clusters the preshower energy (E_{ES}) is also used.

Electron tracking

Electron tracking proceeds in a similar way to muon tracking, the process comprising the general steps described in 5.1.1: seeding, track finding, fitting and selection. Electron candidates can be reconstructed following the standard reconstruction used

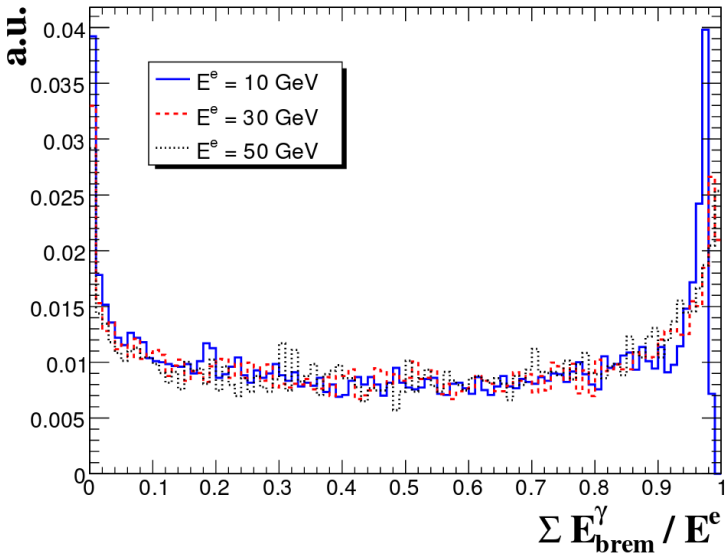


FIGURE 5.2: Distribution of the fraction, $\sum E^{\gamma}_{brem}/E^{e}$, of the generated energy (E^{e}) radiated as bremsstrahlung photons ($\sum E^{\gamma}_{brem}$) for electrons of 10, 30 and 50 GeV. The emission of photons is integrated up to the inner radius of the electromagnetic calorimeter.

for muon tracks. However, due to bremsstrahlung radiation, the energy loss distribution is highly non-Gaussian and, as a consequence, the standard Kalman filter, appropriate when all variable uncertainties are Gaussian, does not suit the purpose and specialized tracking is required.

As the standard tracking efficiency and resolution are not particularly good for electrons, trajectory seeds for the GSF tracking are constructed in two different ways, namely the *ECAL driven* or *tracker driven* seeding procedures.

The ECAL driven method starts by searching for energy clusters in the ECAL and looking for compatible tracker seeds in the pixel tracker. To recover the energy radiated by bremsstrahlung photons, ECAL superclusters are formed by merging energy clusters of similar η at close azimuthal (ϕ) positions, as the bremsstrahlung photons are expected to strike the ECAL at the same η value as the electron but spread in ϕ. ECAL cluster energy and position are used to infer the position of the hits expected in the innermost tracker layers assuming that the cluster is produced either by an electron or a positron. Regions in the tracker compatible with a given ECAL supercluster might have several tracks, from which only one could be that of the electron. To ensure only the more relevant tracks are reconstructed, a devoted seeding procedure is applied. If the event contains several matching seeds for a given supercluster, the one with smallest matching distance is chosen as best candidate. This method is particularly efficient for electrons with $p^{e}_{T} > 10$ GeV.

The tracker driven seeding method takes the track collection and attempts to identify a subset of these tracks compatible with being originated by electrons. A

boosted decision tree performs a preselection of the tracker clusters, in order to reduce the fake rate due to light hadrons. Electrons that do not radiate large amounts of energy will leave tracks that can be extrapolated to ECAL energy clusters. Those that suffer large bremsstrahlung losses will leave a fitted track with few associated hits and a poor χ^2 value. This strategy is more efficient for low p_T electrons and electrons within jets (non isolated).

The seed collections obtained by using these two methods are merged, and used to initiate electron track finding. This procedure is similar to that used in standard tracking, except that the χ^2 threshold, used by the Kalman filter to decide whether a hit is compatible with a trajectory, is weakened in order to accommodate tracks that deviate from their expected trajectory because of bremsstrahlung.

To obtain the best electron track parameter estimates, the final track fit is performed using a Gaussian Sum Filter. The energy loss experienced by an electron traversing material follows a distribution described by the Bethe-Heitler formula, which is non-Gaussian. The GSF technique solves this by approximating the Bethe-Heitler energy loss distribution as the sum of multiple weighted Gaussian functions to model the energy loss. Their widths, means and relative amplitudes are selected so as to optimize this approximation.

This algorithm provides significant improvements to both the electron's momentum and angular resolution compared to the standard algorithm. This effect can be seen in figure 5.3.

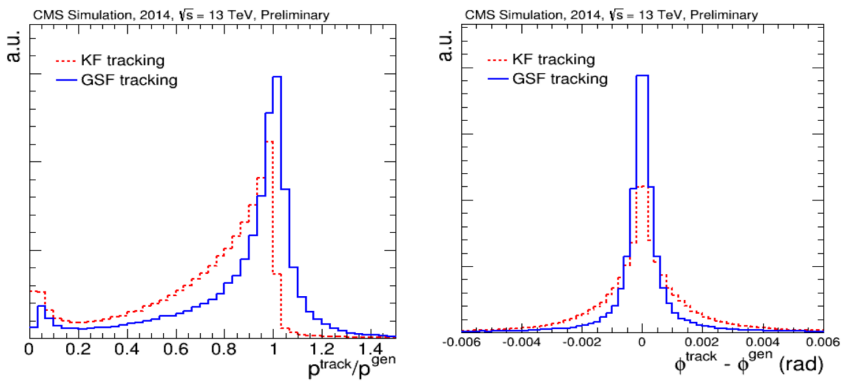


FIGURE 5.3: Comparison of (a) momentum and (b) angular resolutions of the different track reconstruction algorithms in CMS.

Electron identification and matching

A link between a charged-particle track and a calorimeter cluster proceeds as follows. The track is extrapolated from its last measured hit in the tracker to the two layers of the PS and the ECAL, at a depth corresponding to the expected maximum

of a typical electron shower profile. Then, the track is linked to a cluster if its extrapolated position relies within the cluster boundaries. A link distance in the (η, ϕ) plane is then defined between the track and the cluster.

To account for bremsstrahlung losses, tangents to the track are extrapolated to the electromagnetic calorimeter. If one of these tangents relies within a cluster, the latter is linked to the original track as a potential bremsstrahlung photon candidate.

Objects reconstructed as electrons will be the input of the PF algorithm.

The cut-based (CB) identification criteria corresponds to a set of requirements regarding shower-shape related variables, energy related variables, impact parameter cuts and isolation criteria. Different working points are defined depending on the identification efficiency needed for each analysis: *veto* (with an average efficiency of ∼95%), *loose* (∼90%), *medium* (∼80%) and *tight* (∼70%). The current analysis uses electrons passing the tight criteria. This selection includes a set of variables related to electron signal properties that increase the signal to background ratio, providing clean electron signals. These are described in the following.

- $\sigma^{\mathbf{5\times5}}_{\mathbf{i\eta i\eta}}$: selection variables based on the shower shape exploit the fact that electromagnetic showers are more concentric and dense than hadronic ones. $\sigma_{\eta\eta}$, defined as the energy weighted sum over all the crystals in a 5×5 array around the seed crystal[4] of the difference between the particular crystal η and the seed η squared

$$\sigma_{\eta\eta} = \sum_{i \in 5\times5} (\eta_i - \eta_{seed})^2 \frac{E_i}{E_{5\times5}} \tag{5.2}$$

  where i runs over all crystals in the 5×5 array, is independent on the particle p_T. In expression 5.2, $E_{5\times5}$ represents the total energy in the 5×5 array. However, this variable does not perform well near the cracks in the detector, so a redefinition of $\sigma_{\eta\eta}$ is needed. With this view, $\sigma_{i\eta i\eta}$ is defined to be almost identical to its predecessor, but stable in η applying corrected weights.

$$\sigma^{5\times5}_{i\eta i\eta} = \frac{\sum_{i \in 5\times5} w_i \, (\eta_i - \eta_{seed})^2 \cdot \Delta\eta_i^2}{\sum_{i \in 5\times5} w_i} \tag{5.3}$$

  In this definition, the distance of crystal i from the seed crystal $(\eta_i - \eta_{seed})$ is multiplied by the crystal width in η $(\Delta\eta_i)$ and the weight for a crystal i with energy E_i is defined to be

$$w_i = \max(0.47 + \log(E_i/E_{5\times5})) \tag{5.4}$$

  This redefinition is more robust to effects like noise and improves the electron identification performance, as seen in Fig. 5.4.

[4]The *seed* crystal is a local energy maximum above a certain threshold. When an electron is reconstructed, the associated track usually points towards the cluster seed.

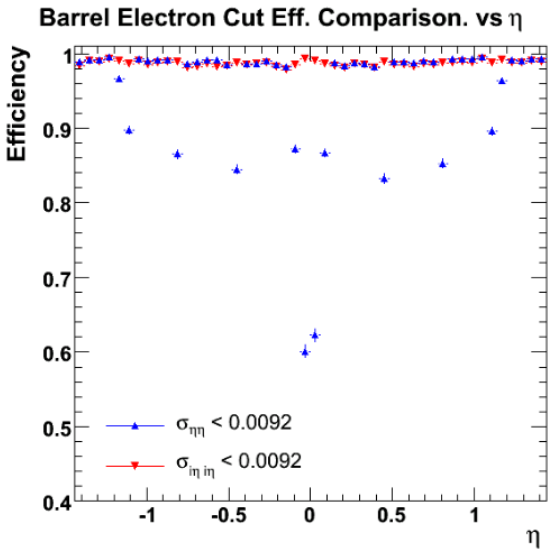


FIGURE 5.4: Comparison between $\sigma_{\eta\eta}$ and the improved $\sigma_{i\eta i\eta}$ which behaves well at the cracks.

Clusters initiated by electrons have their associated track pointing directly towards the weighted cluster center, and a precise matching between seeds and clusters is achieved by applying cuts on related variables:

- $\mathbf{\Delta\phi_{in}}$: defined as the absolute difference between the ECAL supercluster weighted position in ϕ, ϕ_{SC}, and the extrapolated position in the ECAL using the track parameters at the interaction vertex, ϕ_{in}^{extr} (see figure 5.5).

$$|\Delta\phi_{in}| = |\phi_{SC} - \phi_{in}^{extr}| \tag{5.5}$$

  This variable is sensitive to Bremsstrahlung.

- $\mathbf{\Delta\eta_{in}}$: similar to $\Delta\phi_{in}$ but in the η direction

$$|\Delta\eta_{in}| = |\eta_{SC} - \eta_{in}^{extr}| \tag{5.6}$$

These variables are already taken into account for electron preselection during reconstruction. However, accidental track-supercluster matching can be reduced by applying tighter cuts on both.

- **H/E, the relative hadronic over electromagnetic energy fraction**. The energy deposited by an electron is almost fully contained in the electromagnetic calorimeter. In contrast, hadrons will tend to leave energy in the hadronic calorimeter.

- **Expected missing inner hits**: photon conversions in the tracker material are a considerable source of non prompt electrons. Electrons from photons

converting further in the tracker than the first sensitive layer result in tracks without hits in the inner layers.

- **The ratio between the energy in the electromagnetic calorimeter and the track momentum**. For electrons, the ratio E_{SC}/p_{track} is close to unity, being approximately uniformly distributed in other cases. As this variable is too dependent on the electron energy, this variable is redefined as $|1/E_{SC} - 1/p_{track}|$ in order to make it less sensitive to high energy electrons.

- **Isolation**. The most powerful handle for electron identification is isolation. Hadrons that are misidentified as electrons are usually accompanied by other particles nearby, in contrast to prompt electrons that are well isolated.

  The way in which isolation is defined is similar to that of muons. A cone is defined around the electron, and a correction for pileup contribution has to be applied. In this case, the correction is referred to as *effective area correction*, which takes into account the energy density in the event (ρ) as well as the area the electron takes up in the detector or the *effective area* (A_{eff}). The term corresponding to the estimated neutral and photonic energy from pileup in the cone around the lepton is substituted by $\rho \cdot A_{eff}$ in contrast with the $\Delta\beta$ correction term for muons, $\beta \cdot I_{\text{pileup}}$.

  The relative isolation for electrons is defined in a cone of radius $\Delta R = 0.3$ as:

$$I^{\text{e}}_{\text{rel}} = \frac{I_{\text{charged hadrons}} + \max(0, I_{\gamma} + I_{\text{neutral hadrons}} - \rho \cdot A_{eff}(\eta_{SC}))}{p^{\text{e}}_{T}} \tag{5.7}$$

  The median of the transverse energy density ρ is calculated in $\delta\eta \times \delta\phi$ from charged-particle tracks associated to pileup vertices. Effective areas are estimated from simulation and denote the expected amount of neutral energy from pileup interactions per ρ within the isolation cone as a function of η_{SC}.

  A_{eff} is defined as the ratio of the slope, obtained from the linear fit to $Iso(N_{vtx})$ distribution, to the slope of the linear fit to the $\rho_{event}(N_{vtx})$ distribution, where N_{vtx} is the number of the primary vertices in the event. These corrections stabilize the efficiency with respect to the changing pileup conditions.

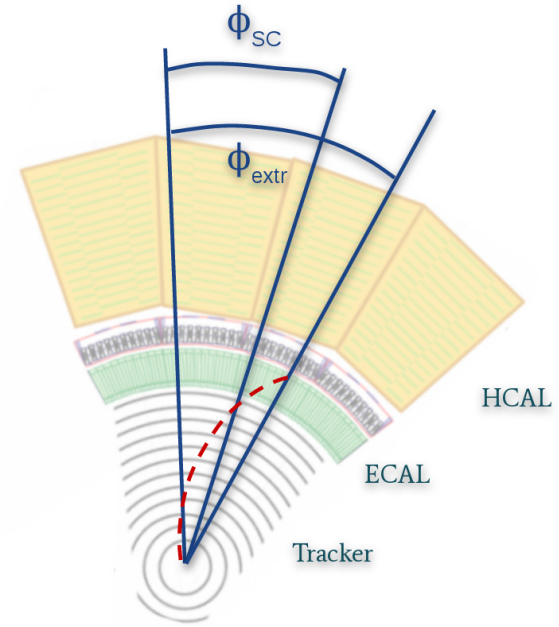


FIGURE 5.5: Definition of $\mathbf{\Delta\phi_{in}}$, the absolute difference between the ECAL supercluster weighted position in ϕ, ϕ_{SC}, and the extrapolated position in the ECAL using the track parameters at the interaction vertex, ϕ^{extr}.

5.2 The Particle Flow algorithm

The Particle Flow (PF) [74] event-reconstruction is an algorithm nowadays used in most CMS analyses. It aims at identifying and reconstructing individually all stable particles arising from LHC collisions (*i.e.* electrons, muons, photons, charged hadrons and neutral hadrons) by combining information from the different CMS subdetectors in an optimal way to determine their direction, energy and type. The accuracy achieved in the measurement and identification of the different particles leads to an improved performance of the reconstruction of jets (section 5.3) and missing transverse energy (section 5.5). The algorithm also serves to other purposes such as identifying b-quark initiated jets (section 5.4), quantifying lepton isolation with respect to other particles in the event or tau reconstruction and identification [75].

The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of the momentum of the track originated at the main interaction vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons attached to the track. The energy of muons is obtained from the corresponding track momentum. The energy of charged hadrons is determined from a combination of the track momentum and the corresponding ECAL and HCAL

energy, corrected for zero-suppression effects, and calibrated for the nonlinear response of the calorimeters. Finally, the energy of neutral hadrons is obtained from the corresponding calibrated ECAL and HCAL energy. The missing transverse energy is defined as the magnitude of the transverse momentum imbalance, which is the negative of the vectorial sum of the transverse momenta of all the particles reconstructed with the PF algorithm. Tracks belonging to the primary or secondary vertices of the most energetic pp interaction are retained, while particles identified as coming from pileup interactions are removed from the event.

5.2.1 The fundamental ingredients of the PF algorithm

The PF algorithm is based on the concept of global event reconstruction as it performs a correlation of the basic *PF elements* (tracks and clusters) obtained from all subdetector systems. The reconstruction of the particles proceeds with a link algorithm that connects these PF elements to form *PF blocks*. Particles are then reconstructed from these blocks, and sequentially removed from the algorithm.

Iterative tracking

The tracker not only provides a vastly superior momentum resolution than that of the calorimeters for charged particles, but also a precise measurement of these particles direction at the production vertex, before any deviation induced by the magnetic field in their way towards the calorimeters. It therefore plays a crucial role in particle reconstruction.

The tracking efficiency must remain as close to 100% as possible to avoid missing any charged hadron signatures, which would then be exclusively detected by the calorimeters, inducing an overall decay in the event reconstruction efficiency. At the same time, the tracking fake rate must remain as small as possible.

With these purposes, an iterative-tracking algorithm [76] was adopted. In the first iteration, tracks are seeded and reconstructed with very tight criteria, leading to a moderate tracking efficiency while keeping a negligible fake rate. In the following steps, track seeding criteria are progressively loosen to increase tracking efficiency, and hits unambiguously assigned to tracks from the first iteration are removed in order to keep the fake rate low. The last iterations have relaxed constraints on the origin vertex, allowing for the reconstruction of secondary charged particles, such as those originating from photon conversions.

Calorimeter clustering

The clustering is performed in each subdetector separately, in order to optimize the detection efficiency: ECAL barrel and endcaps, HCAL barrel and endcaps, and the two preshower layers. In the HF no clustering is performed: the electromagnetic or hadronic components of each cell give rise to an *HF-EM cluster* and an *HF-HAD cluster*. The algorithm aims to detect and measure the energy and direction of stable neutral particles (i.e. photons, neutral hadrons) and separate them from charged-hadron energy deposits, as well as to reconstruct and identify electrons (along with their associated Bremsstrahlung photons) and other charged hadrons for which the tracking parameters were not accurately measured.

The clustering is done in three separate steps. First, local energy maxima above a given threshold in the calorimeters are used to form *cluster seeds*. Secondly, *topological clusters* are formed by the aggregation of cells with at least one side in common with the existing cluster seeds, if the energy of these cells is above a certain minimum. Neighbours from a seed can not become a seed and the number of neighbours is a parameter of the algorithm, which can be either 4 or 8 (see figure 5.6). A topological cluster usually gives rise to as many *PF clusters* as the number of seeds it contains.

It is possible to share the energy of one crystal among two or more clusters.

Then, each seed in the topological cluster gives rise to a *PF cluster*. Calorimeter granularity is exploited by sharing the energy of each cell among all PF clusters, depending on the distance to each of the seeds, and an iterative computation of the cluster energy and position is carried out.

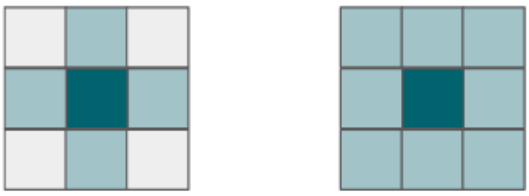


FIGURE 5.6: Seed with 4 and 8 neighbours (in light green) that must have a lower energy than the seed (in dark green).

The topological clustering starts with the maximum energy value in the calorimeter, then does a systematic walk around the detector face, checking if the cells pass a certain energy threshold. Once the program comes across a cell under the threshold, it changes direction until it discovers another region below the threshold. This process continues until it maps out a wide area in which to make a Particle Flow cluster. It repeats this process if there is another seed, or local maximum above a certain threshold, in a region outside of the first topological cluster. The program repeats this until all seeds are accounted for (Note: most seeds will be within the

same topological cluster). Once the topological clusters are mapped out, the program starts developing the PF clusters. Each seed gets a PF cluster that is the same size and shape as the topological cluster it belongs to, in terms of cells. The clustering program puts a bias on cell energies closer to the seed than those further away. Therefore, although each PF cluster will cluster the same cells, each seed's energy distribution will be different.

Link algorithm

A given particle is expected to give rise to several PF elements, namely charged-particle and muon tracks, and clusters in the calorimeters, which need to be linked by a devoted algorithm in order to reconstruct each single particle, while getting rid of any possible double counting from different detectors.

The link algorithm defines a distance between each pair of linked elements in the event to quantify the quality of the link, producing what is referred to as a *PF blocks*. These blocks usually contain only up to three elements, thanks to the good granularity of CMS detectors, the simplicity of the blocks making the algorithm performance independent of the complexity of the event topology.

Further information on the linking between tracks in the silicon detector and calorimeter clusters or signals in the muon chambers is given in sections 5.1.2 and 5.1.1, respectively.

5.2.2 Description of the Particle Flow algorithm

For each block, the algorithm proceeds as follows. On a first step, each global muon gives rise to a *PF muon* if the combined momentum (from the muon chambers and the tracker) is within three standard deviations from that computed solely with the tracker. The corresponding tracks are then removed from the processing.

Secondly, electrons are reconstructed and identified as explained in 5.1.2. Each pre-identified electron track is refitted with a dedicated filter that accounts for all the energy losses the electron suffers while traversing the tracker material (see section 5.1.2). The final identification is performed using a combination of tracker and calorimeter variables, giving rise to a *PF electron* and its corresponding track and calorimeter clusters (including those assigned to the bremsstrahlung photons) are removed from the algorithm.

In the final steps, more stringent criteria are applied to the remaining tracks, to establish links to the rest of the calorimeter clusters. If a track is linked to several HCAL or ECAL clusters, only the link to the closest cluster is kept for comparison.

After this cleaning in the event, each of the remaining tracks in the block is associated to a *PF charged hadron*, its momentum and energy taken from the track momentum. Only if the calibrated calorimetric energy is compatible with this measurement within the uncertainties, its momentum is redefined by a fit of the measurements from the tracker and the calorimeters.

On the other hand, if the calibrated energy of the closest ECAL or HCAL cluster is significantly larger than the associated charged-particle momentum (and the relative difference in energy exceeds the calorimeter expected energy resolution), it will give rise to a *PF photon* or a *PF neutral hadron*. If the excess is larger to the total ECAL cluster energy, a photon is created and the remaining energy excess is associated to a neutral hadron [5]. The remaining ECAL and HCAL clusters give rise to PF photons and neutral hadrons, respectively.

5.3 Jet reconstruction

Hadronic jets are amongst the most striking phenomena in high energy physics. Signatures involving jets almost always have the largest cross sections, but are the most difficult to interpret and to distinguish from backgrounds. Insight into jet properties is of significant importance to understand strong interactions and to look for signs of new physics phenomena.

Due to colour charge antiscreening, quarks and gluons produced in high energy collisions cannot appear as free particles. When trying to separate quarks, the generation a quark-antiquark pair from the vacuum becomes energetically preferable, forming a new bound state with the two initial quarks in a process referred to as *hadronization*. This process is responsible for the production of high multiplicity hadronic jets. Combined information from the tracking and calorimetry systems is used to reconstruct these objects and measure their properties, as well as to provide information on the flavour of the primary quarks from which they originated.

Jets are reconstructed at particle level by clustering particles identified by the particle flow algorithm, this approach allowing for a high reconstruction efficiency even at low momenta. Particles are clustered using the iterative anti-k_T algorithm [77] with a distance parameter $R = 0.4$, over the PF particles that are not identified as isolated leptons. Parameter R controls the cone size of the resulting jets. Preselected isolated muons or electrons are excluded from the jet clustering to prevent double counting of their momenta.

[5] The precedence given in the ECAL to photons over neutral hadrons is justified by the observation that more that 25% of jet energy is generally carried by photons, while only a 3% of this ECAL energy is related to neutral hadrons.

The anti-k_T algorithm is both infrared and collinear safe (see figures 5.7 and 5.8, respectively). Infrared and collinear (IRC) safety is the property that if one modifies an event by a collinear splitting or the addition of a soft emission, the set of hard jets that are found in the event should remain unchanged.

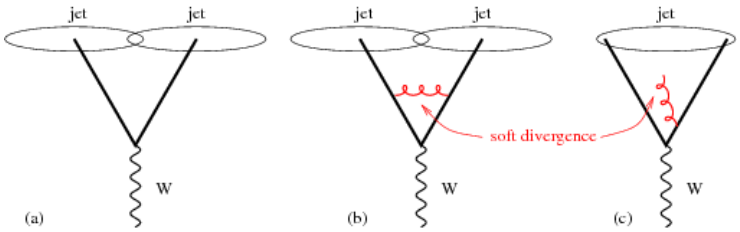


FIGURE 5.7: Configurations illustrating IR unsafety in events with a W and two hard partons. The addition of a soft gluon converts the event from having two jets to just one jet.

Hard partons undergo many collinear splittings during fragmentation. In addition, soft particles are often emitted in QCD events, both through perturbative and non-perturbative effects. These effects occur randomly and are hard to predict.

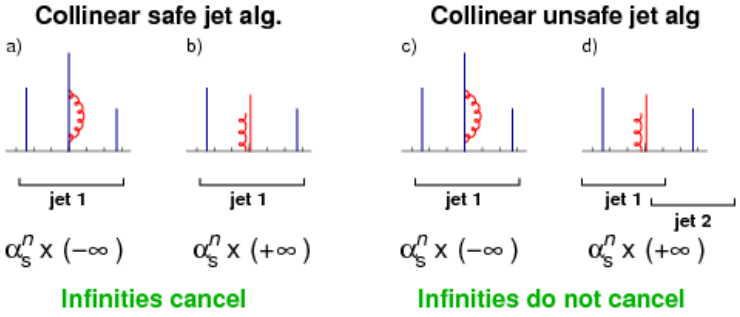


FIGURE 5.8: Illustration of collinear safety (left) and collinear unsafety in an iterative cone type algorithm (right) together with its implication for perturbative calculations (taken from the appendix of [77]). Partons are vertical lines, their height is proportional to their transverse momentum, and the horizontal axis indicates rapidity.

The iterative clustering is performed on particle flow entities, initially referred to as *protojets* or *pseudojets*. For each input object i and each pair i, j of them, the following parameters are calculated:

$$d_i = \frac{1}{p_{T,i}^2} \qquad d_{ij} = \min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right) \cdot \frac{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2}{R^2} \tag{5.8}$$

If $d_{ij} < d_i$, the two objects are merged into a single jet k, and their 4-momenta are summed $p_k = p_i + p_j$. Otherwise, if $d_i < d_{ij}$, object i is removed from the

algorithm and is considered a jet. Clusters are promoted to final jets when their p_T^{-2} is smaller than any remaining distance. The iterative process continues until no objects are left.

In this algorithm, soft particles will tend to cluster with hard ones long before they gather among themselves. If a hard particle has no hard neighbours within twice the cone radius, then it will simply accumulate all its surrounding soft particles forming a perfectly conical jet of radius R. Therefore, the presence of soft particles in the event will not modify the shape of the jets, which will nonetheless remain flexible with respect to hard radiation.

5.4 B tagging

Jets coming from the hadronization of b quarks, the so called *b jets*, are present in many physics processes of interest, such as top quark or Higgs boson decays. An accurate identification of b jets is therefore crucial to study and characterize top events. Different algorithms have thus been developed in CMS to identify (*tag*) this kind of jets, within the pseudorapidity acceptance of the tracker.

These b tagging algorithms [78] benefit from the characteristics of b-flavoured hadrons, such as their long lifetime, high mass and large momentum fraction, as well as from the presence of soft leptons from semileptonic b decays.

A common feature of most of them is the identification of a displaced secondary vertex, reconstructed from displaced tracks within a jet. After hadronization and due to their relatively long lifetimes, B mesons travel measurable distances away from the primary vertex before decaying. This is possible thanks to the excellent precision of the tracker system.

The output of these algorithms is a discriminator value on which cuts might be applied. The different taggers used during Run 2 are:

- **Jet Probability (JP) tagger**: is mostly used for performance measurements. Based on tracker information, it calculates the likelihood of the jet to originate from the primary vertex using the associated tracks. For each track, the probability to originate from the primary vertex is obtained, and the information from all tracks is then combined to extract the jet probability.
- **Combined Secondary Vertex (CSV and CSVv2)**: this algorithm exploits information from displaced tracks and secondary vertex information. This algorithm was optimized in Run 2, so from then on the CSVv2 algorithm is used. It is based on the CSV algorithm, significantly improved with respect to it by using a multivariate analysis (neural network) instead of a

WP name	WP value	Selection efficiency (%)	Mis-identification (%)
Loose	0.5426	82	11.5
Medium	0.8484	67	1.4
Tight	0.9535	47	0.15

TABLE 5.1: B-tagging working points and their selection and mistagging efficiencies for PF jets using the CSVv2 algorithm.

Likelihood Ratio and by including additional variables and an improved track selection.

Two algorithms for reconstructing secondary vertices are used: the *adaptive vertex reconstruction* (AVR) and the *inclusive vertex finder* (IVF) algorithms. The AVR algorithm was used in b jet identification during Run 1. Certain selection criteria are applied in order to remove vertices that are less likely to originate from a b hadron decay. The reconstructed secondary vertices are removed if they share more than a 65% of the tracks with the primary vertex, if they do not have at least two tracks, or if the distance between the primary vertex and the secondary vertex is smaller than 0.1mm or exceeds 2.5cm in the transverse plane. Secondary vertices are only considered if their mass is below 6.5GeV and not compatible with kaon decays.

In contrast to the AVR algorithm, the IVF is not seeded from tracks associated to PF jets, but instead collects information from the set of all tracks reconstructed in the event. Among them, displaced tracks are set as seeds if their distance parameter is of at least 50 μm with a significance as large as 1.2. Nearby tracks are then clustered using these seeds following some selection criteria. Multiple vertices might share a certain amount of tracks and, at this stage, a vertex is removed if it shares at least 70% of its tracks and its distance significance with respect to another vertex is less than 2, and the vertices are then refitted. The efficiency to reconstruct a secondary vertex for b jets using the IVF algorithm is about 10% higher compared to the AVR. However, the probability to find secondary vertices in light flavour jets with this algorithm increases by an 8%.

The CSVv2 algorithm has the best performance among the CMS heavy flavour taggers.

- **Combined Multivariate Analysis (cMVAv2) algorithm**: developed in Run 2, it combines the information from six different b tagging discriminators in a Boosted Decision Tree. In addition to the JP and two CSVv2 algorithms, it uses information from the Soft Electron (SE) and Soft Muon (SM) discriminators. These SE and SM algorithms look for the presence of soft leptons inside the jet cone

As can be seen in figure 5.9, the improvement of the CSVv2(AVR) algorithm with respect to the CSV is of order 2-4% in b jet identification efficiency at the same misidentification probability. The use of IVF vertices with respect to AVR vertices

in CSVv2 improves the efficiency by a 1-2% at the same misidentification rate. The cMVAv2 algorithm, however, outperforms the rest of the b jet identification algorithms.

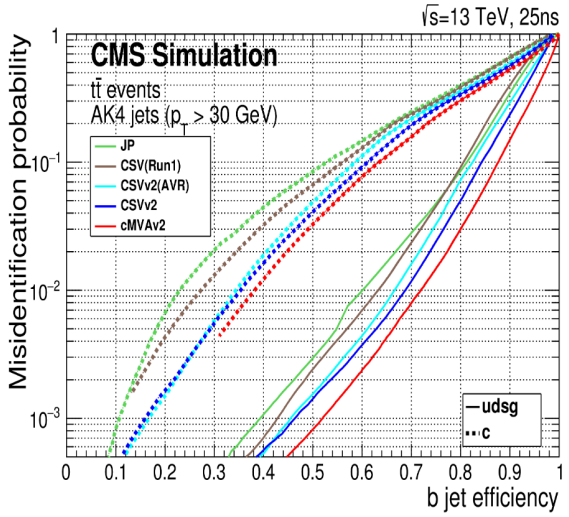


FIGURE 5.9: Performance of the b jet identification efficiency algorithms demonstrating the probability for non-b jets to be misidentified as b jet as a function of the efficiency to correctly identify b jets.

5.5 MET reconstruction

Particles such as neutrinos escape the detector without being detected due to their very low interaction cross section with matter, and cannot be reconstructed directly by the particle flow algorithm, but their presence can be inferred by conservation laws in the transverse plane ($\vec{p}_T^{\,total} = 0$, before and after the collision). Events with neutrinos in their final state are characterized by large missing transverse energy ($\not{E}_T$) values. Missing transverse energy $\not{E}_T$ and momentum $\vec{\not{p}}_T$ are calculated from the momentum imbalance of the summed PF candidate momenta in the transverse plane as

$$\not{E}_T = |\vec{\not{p}}_T| \qquad \vec{\not{p}}_T = -\sum_i^{PFcand} \vec{p}_{T,i} \tag{5.9}$$

Even though large $\not{E}_T$ values are often associated to the presence of undetected neutrinos in the event, smaller amounts of missing transverse energy are often due to resolution or instrumental effects, which need to be taken into account.

An optimal reconstruction of the missing transverse energy and momentum depends on an accurate identification and reconstruction of all particle flow objects in the event. Any errors or inefficiencies in the measurement of the activity in the transverse plane would translate in a bad reconstruction of the missing E_T (often referred to as *MET*) and a subsequent misidentification of the kind of events we would be looking at. In fact, the $\not\!\!E_T$ is one of the most important observables to be taken into account in many searches for new physics phenomena, such as supersymmetry, extra dimensions or dark matter searches. In addition, it plays a crucial role in SM measurements of processes involving top quarks and W bosons, subsequently being of special interest in the present analysis.

The measurement of the missing energy in the transverse plane is also sensitive to *in-time* and *out-of-time* pileup, meaning additional proton-proton collisions occurring either in the same bunch crossing as the collision of interest, or in the previous or following ones, respectively. A detailed understanding of all the sources of possible uncertainties in the $\not\!\!E_T$ measurement is therefore of crucial importance.

The reader may find more detailed information in [79].

Chapter 6

Measurement of the tZq production cross section

The following three chapters provide full description of the analysis. The current one is devoted to the explanation of the building elements of the analysis: the event topology and main backgrounds, the data and simulation samples used, along with the event and object selection and the corrections applied to the simulated samples. The strategy to separate signal events from the main backgrounds is also portrayed here. The next two chapters complete the description of the analysis. In chapter 7, the tools for the statistical analysis are described and the results will be presented in chapter 8.

6.1 Overview of the analysis

The measurement of the tZq production cross section in the final state with three high-p_T leptons, from now on referred to as *trilepton* channel, is based on a binned maximum likelihood fit performed simultaneously on three statistically independent regions: the signal region and two control regions. The latter are defined to be as close as possible to the signal region so that extrapolations from one region to another are minimized, in order to reduce background-related uncertainties. In the first place, events entering the analysis need to satisfy certain criteria. In particular, they are required to pass the *trilepton baseline selection* (will be described further in section 6.8) which requires the presence of exactly three high-p_T leptons in the event. This requirement reduces considerably the contribution from most of the background sources. After the initial trilepton selection, the sample separation is based on the jet and b jet multiplicities of the events. The two control regions are defined so that they are enriched in events from the main background sources in the analysis, and the fit performed simultaneously in the three regions allows to better constrain the contributions from these background processes.

In order to achieve an optimal signal-to-background separation, multivariate

techniques are used in the analysis. Two *boosted decision trees* (BDTs, theoretical description will be given in section A.2 while those in the analysis are described in section 7.1.1 in the next chapter) are trained in two of the three regions (the signal region and one of the two control regions). The output discriminant[1] distribution of both BDTs, along with the distribution of the W boson transverse mass obtained in the third region, are used as templates for the simultaneous fit.

The fit is performed using the Combine statistical tool developed by the Higgs Combination Group [80]. Based on the RooStats framework, this tool allows to run different statistical methods in a user-friendly way. The parts of the study concerning the statistical analysis will be discussed in more detail in the next chapter.

The analysis is performed on events containing only selected electrons or muons in the final state, and the resulting cross section is extrapolated to include the contribution from tau leptons to the trilepton tZq production[2]. Figure 6.1 provides a simplified description of the entire analysis.

[1]The *discriminant* obtained after a multivariate analysis is a single variable that offers an optimized discrimination between events from two different hypotheses (i.e. signal vs background). Ranging between two given values, events from one hypothesis would be next to the maximum value of the discriminator, and the opposite tail of the distribution would be populated by events from the other hypothesis.

[2]The contribution from taus decaying to electrons or muons that pass the selection criteria is included in the analysis.

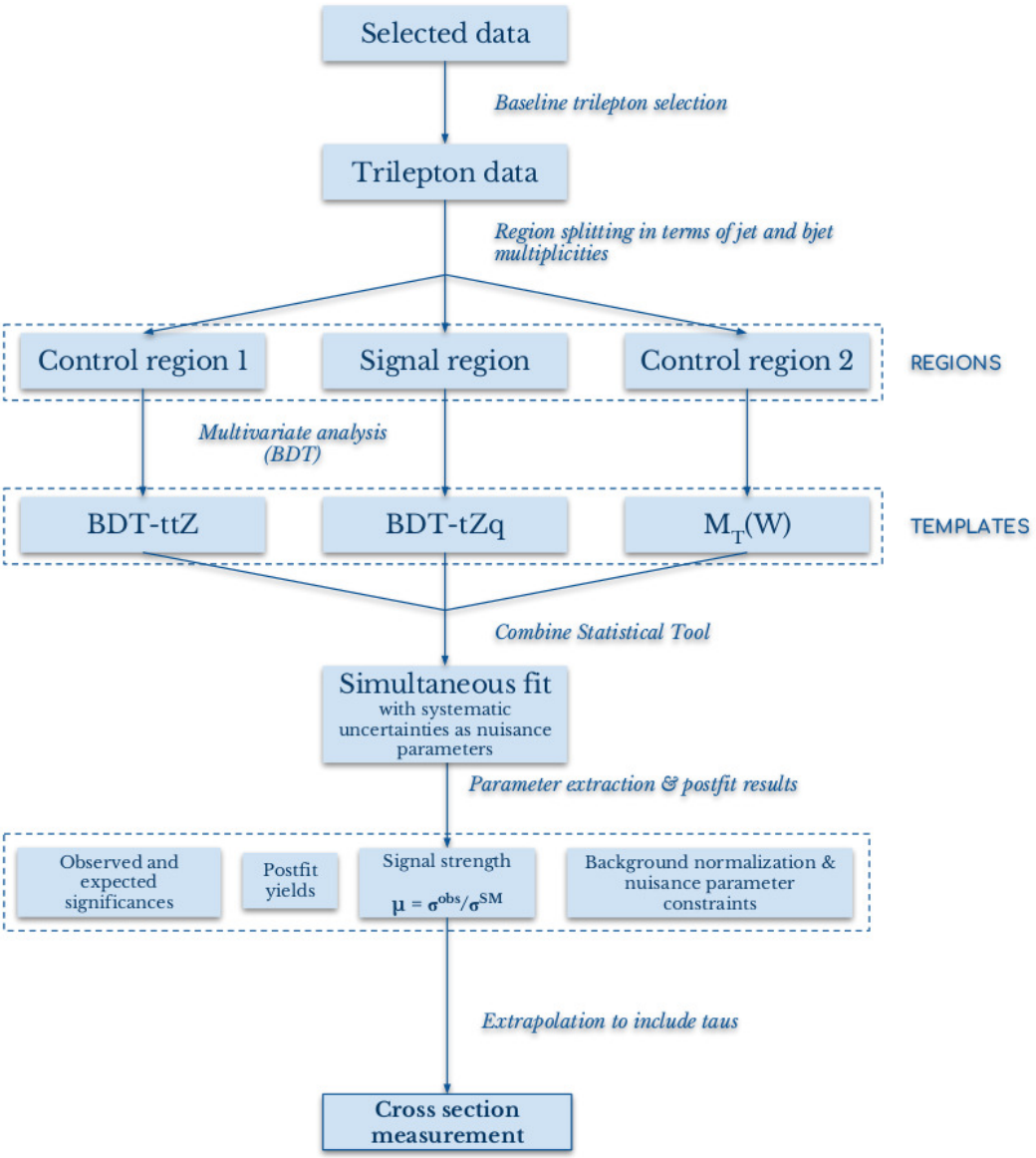


FIGURE 6.1: Schematical view of the analysis. Further details are contained within the next two chapters.

6.2 The tZq trilepton channel

As discussed in section 3.2.1, the production of a single top along with a Z boson can yield different final states, depending on the decay of the two gauge bosons involved in the process (the W from the top decay and the additional Z boson).

The focus of the current analysis is set on events with tree isolated high-p_T leptons in the final state, where the W boson decays leptonically and the two remaining leptons come from the leptonic decay of a Z boson (or when a non-resonant lepton pair is produced) as

$$
\begin{array}{l}
q + b \rightarrow t + Z + q' \\
\qquad\quad\; \hookrightarrow \ell^+ + \ell^- \\
\quad\; \hookrightarrow W + b \rightarrow \nu_\ell + \ell^+ + b
\end{array}
\tag{6.1}
$$

The cross section measurement could be conducted in any of the different decay channels. Dilepton tZq production (in which the W boson decays hadronically, $W \rightarrow qq'$) has a cross section twice as large as trilepton tZq:

$$
\underbrace{\sigma^{SM}(t(qqb)\ell\ell q)}_{\text{tZq}-\text{dilepton}} = \underbrace{\sigma^{SM}(t(\ell\nu_\ell b)\ell\ell q)}_{\text{tZq}-\text{trilepton}} \times \underbrace{\frac{BR(t \rightarrow qqb)}{BR(t \rightarrow \ell\nu_\ell b)}}_{\simeq 1.97}
\tag{6.2}
$$

since $BR(t \rightarrow qqb) = (66.5 \pm 1.4)\%$ and $BR(t \rightarrow \ell\nu_\ell b) = (33.8 \pm 1.0)\%$ (values taken from [81], [2]).

The dilepton channel is dominated by the contribution from $t\bar{t}$ dilepton production and processes with one Z boson produced in association with jets (*Z+jets*), whose separation from the dilepton tZq signal presents a difficult task. In addition, these two processes have large production rates in contrast with the dilepton tZq process, as shown in Fig. 2.5.

However, the case in which both heavy particles decay into charged leptons yields a trilepton topology for which the SM backgrounds are much reduced in comparison with the dilepton case. SM processes with very similar event topology as the tZq final state are considered backgrounds, which must be precisely determined in order to subtract their contributions. Two different kind of background sources can be present: the so-called *reducible* (those in which certain particle combination *mimics* the signal final state) and *irreducible* (those in which the final state is identical to the signal one) backgrounds.

A previous analysis was conducted by CMS at 8 TeV [11], yielding a measured cross section $\sigma(\text{tZq} \rightarrow \ell\nu\text{b}\ell^+\ell^-\text{q}) = 10^{+8}_{-7}$ fb with an observed statistical significance

of 2.4σ. In this same analysis, exclusion limits are set on FCNC branching ratios at 95% CL.

6.2.1 Trilepton event topology: tZq and main backgrounds

The signature of tZq production in the trilepton decay mode consists of a single top produced in the t-channel decaying leptonically (giving rise to a b quark jet, a lepton and its corresponding neutrino), a pair of opposite-sign same-flavour (OSSF) leptons compatible with a Z boson decay (but which might also come from non-resonant contributions), and an additional recoiling jet.A schematic view of the final state is shown in figure 3.4.

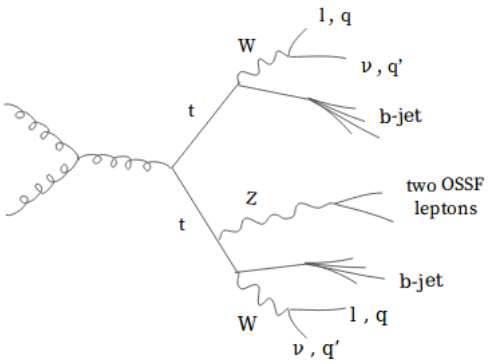
There are several other SM processes containing the same lepton topology as the tZq case in the final state, which have higher production rates than the process under study. The two most important in our analysis are diboson WZ production in association with jets (WZ+jets) and $t\bar{t}Z$ production, which contain two opposite sign, same flavour leptons compatible with a Z boson decay and an additional high-p_T lepton. A schematic view of the topology of the two main background processes in our analysis is shown in figure 6.2. The main difference between these background sources and tZq production comes from the jet and b jet multiplicities in the final state.

In $t\bar{t}Z$ production, the final state of events in which one of the top quark decays leptonically and the other one decays hadronically contains at least three jets. Two of these jets correspond to the two b quarks arising from the decays of the top and the antitop. In contrast, no b jets are expected in WZ production besides a small contribution from b quarks coming from gluon splitting. This feature will be exploited to achieve a better discrimination among the three processes, as suggested in [82].

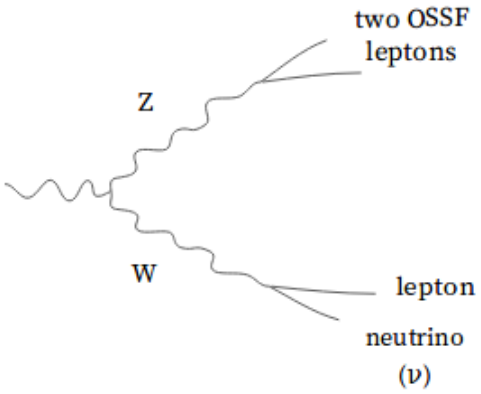
Even though the expected b jet multiplicity of these processes is different to the one in tZq, due to the limited efficiency of b jet identification, background events will inevitably fall in the signal region. For instance, if one of the two b jets expected in $t\bar{t}Z$ decays is identified as arising from a light quark (*light jet*), only one b jet will be selected and the corresponding event will pass the signal selection. On the other hand, if light jets in WZ+jets production are mistagged as b jets or if there are b jets coming from gluon splitting, there would be one identified b jet in the final state, and the event will look like signal.

Other SM processes involving top quarks such as $t\bar{t}H$ and $t\bar{t}W$ can have the same lepton composition as tZq, but they can be easily removed in the selection process by applying a requirement on the invariant mass of the OSSF lepton pair.

Another very important source of background in our search is is the so-called



(A) ttZ



(B) WZ+jets

FIGURE 6.2: Topology of the two main background processes of tZq trilepton production.

non-prompt lepton (NPL) background, sometimes also called *fake lepton background*. These are processes in which either another physical object (i.e. a misidentified jet) or real leptons coming from heavy flavour decays or photon conversion fake prompt leptons in the final state. Events with one fake lepton come mainly from Drell-Yan (DY), WW, $t\bar{t}$ and tW production. The dominating source of NPL background events is DY (Z+jets), followed by $t\bar{t}$ production. In this analysis, it was found that the contribution from background events with more than one fake lepton is negligible. tZq events contain a Z boson and Z+jets events can pass the signal selection if one of the jets is misidentified as a lepton. Fake lepton rates are not well modelled in simulation, so the contribution from NPL background is estimated using data-driven techniques.

6.3 Data samples and trigger strategy

The current analysis uses data collected by CMS from proton-proton collisions produced in the LHC during the 2016 data taking period, at a center-of-mass energy

of $\sqrt{s}$ = 13TeV, corresponding to an integrated luminosity of 36fb^{-1}. Raw data coming from the detector is processed at the CPU farms at CERN in order to create *datasets* containing the reconstructed physical objects. Calibration effects are already taken into account when delivering the different datasets available for the analyzers.

Experimental data in the analysis is taken from datasets containing at least one or two high-p_T leptons (electrons or muons) at the trigger level, and the online selection is performed on events triggered by the presence of either one, two or three leptons above certain p_T and isolation thresholds. The lowest thresholds are listed in table 6.1.

Trigger paths	**Lowest p_T thresholds**
Three lepton triggers	16, 12 and 8 GeV for electrons 12, 10 and 5 GeV for muons
Double lepton triggers	23 and 12 GeV for electrons 17 and 8 GeV for muons
Single lepton triggers	32 GeV for electrons 24 GeV for muons

TABLE 6.1: Lowest p_T thresholds in the trigger paths used in the analysis.

The different trigger paths used in the analysis are categorized as

- Paths triggered by at least one muon and one electron (ME)
- Paths triggered by at least two electrons (EE)
- Paths triggered by at least two muons (MM)
- Paths triggered by a single electron (E)
- Paths triggered by a single muon (M)

As the same event might be present in different datasets used, which are analyzed simultaneously, the trigger logic is designed specifically to avoid possible double counting of events on data. It consists in vetoing a given event in a dataset if it has already been selected in another one. If an event passes, for instance, any of the paths triggered by the presence of a muon and an electron (ME) it is correspondingly assigned the flag *ME*. This event is kept if it is in the MuonEG sample, but discarded in any other dataset. Events labeled *MM* (passing paths triggered by the presence of two muons) are kept if they are in the DoubleMuon sample, and discarded in any other sample. The same logic applies to the other samples. All events coming

from the datasets used in the analysis fall within one of these different paths, so no events are lost. Table 6.2 shows a schematic view of the logic used on data. In simulated events there is no such double-counting problem, therefore a simplified logic (consisting on a simple `OR` of all the different trigger paths used) is followed in this case.

The choice of this strategy guarantees not only that all events of interest are not thrown away at any point, getting the maximum number of events in the signal region, but it also enables to achieve full trigger efficiency (see section 6.6.1). The choice of the offline selection p_T cuts is such that these values are above the p_T thresholds of the trigger paths used, and the selected logic prevents from possible double counting of events.

Sample	Trigger logic
MuonEG	ME
DoubleMuon	MM & !ME
DoubleEG	EE & !MM & !ME
SingleMuon	M & !EE & !MM & !ME
SingleElectron	E & !M & !EE & !MM & !ME

TABLE 6.2: Trigger logic used on data samples.

6.4 Signal and background simulation

Various samples of simulated events for signal and background processes are generated. Simulated events are used extensively in this measurement to evaluate the detector resolution, the efficiencies and acceptance, and to estimate the contributions from background processes that have topologies similar to the trilepton tZq final state.

The default tZq sample is generated at NLO precision in QCD in the 4F scheme using the MADGRAPH5_aMC@NLO 2.2.2 generator [83]. The two main background processes, WZ+jets, $t\bar{t}Z$ and $t\bar{t}W$ are generated using the same event generator as the signal sample, with up to one additional hadronic jet at NLO QCD precision. $t\bar{t}H$ production is simulated with the same generator. Other minor processes in our analysis, such as ZZ and tWZ production, are generated using POWHEG v2.0 [84]-[89] at NLO and MADGRAPH5_aMC@NLO at LO precision, respectively. MADGRAPH5_aMC@NLO and POWHEG define the scheme or prescription for the matching of the fixed-order matrix elements to parton showers. The PDF set NNPDF 3.0 is used in all generators.

All samples are interfaced to the general purpose MC generator Pythia version 8.205 [90] with the CUETP8M1 tune for parton showering and hadronization. QCD

Sample	σ**(pb)**
tZ(ll)q (4F)	0.0942
ttZ(qq)	0.5297
ttZ(ll)	0.2529
ttW(lν)+jets	0.2043
ttH (no ttbb)	0.2151
WZ+jets	5.26
Z(ll)Z(ll)	1.212
tWll (5F)	0.01123

TABLE 6.3: Simulated processes in the analysis along with the theoretical predictions of their corresponding cross sections at 13 TeV, which are used to obtain the normalization of each of the samples.

Monte Carlo generators have parameters that can be adjusted or *tuned* to control the modeling of the properties of the events. A specified set of such parameters adjusted to fit certain prescribed aspects of the data is referred to as a *tune*. CUETP8M1 (where CUET stands for "CMS underlying event Tune", P8 is for "Pythia8" and is labeled M for the *Monash Tune* [91]) is used in Pythia8 for all samples. The underlying event (UE) consists of particles from the hadronization of beam-beam remnants (BBR), of multiple-parton interactions (MPI), and their associated initial and final state radiation (ISR and FSR, respectively).

The events are simulated in final states that include decays to electrons, muons, and τ leptons. A top quark mass of 172.5 GeV is assumed. Multiple minimum-bias events generated with Pythia are added to each simulated event to mimic the presence of pileup, with weights that reproduce the measured distribution of the number of pileup vertices in data. Finally, the simulation of the passage of particles through the detector material is performed using the GEANT4 package [92].

The full list of MC simulated processes considered is given in table 6.3. These simulated samples are normalized to their corresponding cross sections, obtained from NLO calculations for all samples except for the tWZ(ll) sample, where the calculation is done at LO. These values are shown in table 6.3.

6.4.1 Splitting of the WZ+jets sample

The flavour content of the WZ+jets simulated sample is not reliable. That means that the proportion of jets originated by b-, c-, or light-partons in the sample is not properly simulated. To improve the modelling of the WZ+jets background in the analysis, the contributions from WZ+b, WZ+c and WZ+light-flavour jets in the sample are separated, and treated as independent background sources in all steps of the analysis. The separation is using the generated information of the

sample, namely the flavour of the hadrons from which the reconstructed jets in the event originate. If a hadron containing a b quark (B hadron) is found, the event is collected in the WZ+b sample. Events that do not contain B hadrons, but instead have at least one hadron coming from a c quark (C hadron) are stored in the WZ+c sample. If no B or C hadrons are found, all jets are assumed to be originated by u,d,s, quarks or gluons (udsg), and the event kept in the WZ+light sample. Even when one B hadron is found, additional jets in the event, classified as WZ+b, may be originated by c or udsg partons. Likewise, the WZ+c sample may contain, in addition to the jet including C hadrons, other jets arising from udsg. The jet flavour composition of the three WZ+jets subsamples is shown in table 6.4. Light jets are dominant in all of them, accounting also for 80% of the flavour content of the WZ+b and WZ+c samples.

Sample	**WZ+b**	**WZ+c**	**WZ+light**
% b jets	19	0	0
% c jets	4	20	0
% udsg jets	77	80	100

TABLE 6.4: Flavour content of the three WZ samples after the splitting has been applied.*udsg* jets are those initiated by light quarks (u,d,s) or gluons.

Since these contributions are treated separately, they are initially normalised with the same cross-section, but in the final fit (Chapter 7) they are left to vary independently, which provides a different postfit normalization factor for each sample.

6.5 Event and object selection

An efficient selection procedure has to be defined in order to reduce the contamination of other processes which resemble the final state under study, and increase the sensitivity to signal events. The current section describes both the kind of physical object candidates (electrons, muons, jets) and event requirements used in the analysis.

6.5.1 Object selection

Physical objects entering the analysis are required to satisfy different criteria in order to reduce contribution from background processes. These requirements are summarized in table 6.5.

Electrons

Selected electrons are PF electrons with a GSF track (recall section 5.1.2), required to have a $p_T > 25$ GeV and lie within a pseudorapidity coverage of $|\eta| < 2.5$. Moreover, they are required to pass the tight cut-based identification criteria (described in 5.1.2), with a 70% identification efficiency. A relative A^{ρ}_{eff}-based isolation of $I^{e}_{\text{rel}} < 0.059$ (0.057) within a cone of radius $\Delta R < 0.3$ in the endcaps (barrel) is also required. Isolation cuts are encoded in the cut-based electron identification criteria (see table 6.5).

Events with additional electrons with $|p_T|$ >10 GeV within $|\eta|$ <2.5 passing the *veto* cut-based ID are removed from the selection.

Muons

All muons entering the analysis are PF muons that pass the tight working point criteria defined for muon identification, described in 5.1.1, in order to reduce to minimal the level of background from non-prompt muons.

Selected muons are further required to have p_T >25 GeV and lie within $|\eta|$ <2.4. Muon candidates are only considered if they are spatially isolated from electromagnetic and hadronic activity in addition to the tight identification criteria. Muon candidates are required to be isolated with a relative $\Delta\beta$-corrected isolation (described in section 5.1.1) of $I^{\mu}_{\text{rel}} < 0.15$ within a cone of $\Delta R < 0.4$.

Events with additional muons identified as *loose* PF (loose muons are either global or tracker muons without further requirements) with p_T >10 GeV that lie within $|\eta|$ <2.4, satisfying I^{μ}_{rel} <0.20 are vetoed.

The lepton selection used for the NPL background sample differs from the one described above, and details are given in section 6.7.

Jets

Jets are selected if they pass the cuts defined by the *loose* identification criteria working point (WP) defined for physics analysis in CMS, in order to reject fake, badly reconstructed and noisy jets while retaining 98-99% of the real ones. The various WPs give information about the particle composition of the jets (neutral and charged hadron fractions, muon fraction, total number of constituents, etc.) and the requirements usually depend on the η region considered.

The definition of the loose WP is as follows. In the central regions ($|\eta|$ <2.7):

- p_T^{jet} >10 GeV
- Charged hadron fraction > 0.0
- Neutral hadron fraction < 0.99
- Charged track multiplicity > 0.0
- Charged EM fraction < 0.99
- Neutral EM fraction < 0.99

For the *charged* variables the pseudorapidity coverage is restricted to $|\eta| < 2.4$, since there is no tracker coverage outside of this region. In contrast, the *neutral* variables requirements extend to the whole $|\eta| < 2.7$ region. For 2.7< $|\eta|$ <3.0,

- Neutral hadron fraction < 0.98
- Neutral electromagnetic fraction > 0.01
- Neutral particle multiplicity > 2

In the case of 3.0< $|\eta|$ <5.0 we have:

- Neutral electromagnetic fraction < 0.90
- Neutral particle multiplicity > 10

Apart from satisfying the previous conditions, encapsulated in the *loose* identification flag, jets are selected if they have a p_T >30 GeV and lie within $|\eta|$ <4.5 (to account for the forward jets corresponding to the recoiling quark in single top processes). Jets are discarded if a selected lepton lies within a cone of radius $\Delta R = 0.4$ around the jet (lepton-jet separation), and all events containing jets in the regions 2.69< $|\eta|$ <3.0 with 30< p_T <50 GeV are vetoed. This corresponds to a problematic region around the HE-HF transition for which jet energy corrections are not well described, resulting in two unphysical bumps in the η distribution of low-p_T jets ($p_T < 50$ GeV).

Finally, b jets are tagged with the *loose* working point of the CSVv2 discriminant described in 5.4. At the chosen operating point, the CSVv2 algorithm has an efficiency of about 83% to correctly tag b jets and a probability of 10% for mistagging gluons and light quarks.

Electron selection		
	Electron candidates	**Veto electrons**
p_T	> 25 GeV	> 10 GeV
$\|\eta\|$	< 2.5	< 2.5
Electron cut-based ID	Tight	Veto
$\sigma^{5\times5}_{i\eta i\eta}$	< 0.010(b)/0.029(e)	<0.012(b)/0.037(e)
$\Delta\phi_{in}$	< 0.082(b)/0.039(e)	< 0.228(b)/0.213(e)
$\Delta\eta_{in}$	< 0.003(b)/0.006(e)	< 0.00749(b)/0.00895(e)
H/E	< 0.041(b)/0.064(e)	< 0.356(b)/0.211(e)
$\|1/E - 1/p\|$	< 0.013(b)/0.013(e)	< 0.299(b)/0.150(e)
Expected missing inner hits	1	2(b)/3(e)
Pass conversion veto	yes	yes
$I^{\mathrm{e}}_{\mathrm{rel}}(\Delta\mathrm{R} = 0.3)$	0.059 (e) 0.057 (b)	0.159 (e) 0.175 (b)
Muon candidate selection		
	Muon candidates	**Veto muons**
p_T	> 25 GeV	> 10 GeV
$\|\eta\|$	< 2.4	
Muon ID	Tight	Loose
Muon reconstruction type	PF global	PF global or tracker
$I^{\mu}_{\mathrm{rel}}(\Delta\mathrm{R} = 0.4)$	< 0.15	< 0.20
χ^2 of the global-muon track fit	< 10	-
$N_{chamberhits}$ in global muon track fit	> 0	-
Number of matched stations	> 1	-
$\mathrm{d_{xy}}$ of tracker track	< 2 mm	-
$\mathrm{d_z}$ of tracker track	< 5 mm	-
Number of pixel hits	> 0	-
Number of tracker layers with hits	> 5	-
Jet selection		
	Jet candidates	**b jets**
p_T	> 30 GeV	> 30 GeV
$\|\eta\|$	< 4.5	< 2.5
$\Delta\mathrm{R}(\mathrm{lepton}, \mathrm{jet})$	> 0.4	> 0.4
CSVv2 discriminant value	-	> 0.460 (loose WP)

TABLE 6.5: Selection of PF objects used in the analysis. Barrel (b) electrons go up to $|\eta| \leq 1.479$. Endcap (e) electrons range from this value up to $|\eta| < 2.5$.

6.5.2 Event topology and selection

Events are selected if they meet the requirements of tZq final states

$$tZq \to WbZq \to \ell\nu b\ell^+\ell^- q$$

where ℓ stands for either electrons or muons. The additional quark is emitted forward, so the corresponding jet is searched for in a wide pseudorapidity region. All objects used in the analysis are PF objects (see section 5.2). Candidate events must pass the trigger selection criteria defined in section 6.3 and at least one reconstructed primary vertex must be present. Data events are applied a series of *cleaning filters* to remove events containing anomalous detector effects that compromise the integrity of the recorded data and that are either not present or impossible to include in simulation.

The baseline selection in the analysis consists of events with exactly three leptons, two of which have to be compatible with a Z boson decay (opposite sign, same flavour, and their reconstructed mass compatible with the Z boson nominal mass within a mass window of 15 GeV). Events passing the baseline selection, shown in table 6.6, will be further cathegorized according to their jet and b jet multiplicities (this will be described in section 6.8). Selected objects are required to fulfill the criteria specified in the previous section (6.5.1).

Event cleaning
Trigger selection: tri-, bi- and single lepton paths
Event filters
Baseline trilepton selection
Exactly 3 high-p_T isolated leptons (e, μ)
Two OSSF leptons with $m_{\ell\ell} \in [m_Z - 15\,\text{GeV}, m_Z + 15\,\text{GeV}]$
Veto on any additional leptons

TABLE 6.6: Event cleaning and baseline trilepton selection. Events passing this first selection will be further categorized in terms of their jet multiplicities.

6.6 Correction to simulations

Simulated samples do not perfectly describe what is observed in data. In order to obtain solid predictions in an analysis, different types of corrections are applied to the simulated samples to fit the distributions from data. These might be introduced as scale factors used to provide MC weights or smearing of the different physics objects momenta, making use of tag and probe methods, among others. The

corrections used in the different MC samples entering the analysis will be reviewed in this section.

- **Pileup reweighting**

  Because of the high instantaneous luminosity, multiple proton-proton interactions take place during a single bunch crossing. The products of these collisions interact with the detector at the same time and complicate the measurement of the particles originating from the collision of interest. This effect is referred to as *pileup*. The number of proton-proton interactions per bunch crossing quantifies the amount of pileup and is proportional to the number of primary vertices reconstructed in the event. Of these reconstructed vertices, the one with the highest reconstructed energy is considered the primary vertex that is relevant for the analysis. Particles produced by the other reconstructed vertices should not be included on the event reconstruction.

  The effect of pileup on the generated samples is simulated with minimum-bias interactions overlaid on top of the hard scattering event, following a Poisson distribution. However, the number of minimum-bias events overlaid in digitization does not exactly match the actual data taking conditions for pileup. As a consequence, an event weight is derived from the ratio of the distributions of pileup interactions in data and simulation.

  The number of interactions in data is proportional to the instantaneous luminosity $\mathcal{L}$ times the total inelastic pp cross section σ_{pp}:

  $$\langle N_{\text{pileup}} \rangle = \sigma_{pp} \times \mathcal{L} \tag{6.3}$$

  The estimated value for the inelastic pp cross section at 13 TeV pp is 69.2±4.6 mb. In 2016, for a bunch crossing rate of 25 ns, an average of 27 pileup interactions has been observed in 13 TeV pp collisions at the IP of CMS (see figure 6.3). In order to do the reweighting, an initial *unnormalized* weight is obtained from the ratio of the pileup distributions in data and simulation

  $$\omega_0 = \frac{N_{\text{data}}^{\text{pileup}}}{N_{\text{MC}}^{\text{pileup}}} \tag{6.4}$$

  A normalization factor κ_{pileup} is obtained by comparing the original and the weighted MC samples, as in:

  $$\kappa_{\text{pileup}} = \frac{\text{yield in the unweighted sample}}{\text{yield in weighted sample}} \tag{6.5}$$

  from which the final event weight used on the simulated events

  $$\omega_{\text{pileup}} = \kappa_{\text{pileup}} \cdot \omega_0(N_{\text{vtx}}) \tag{6.6}$$

  is derived, where N_{vtx} is the number of vertices in the event. Figure 6.4 shows the distribution of the estimated number of interactions (pileup) in the

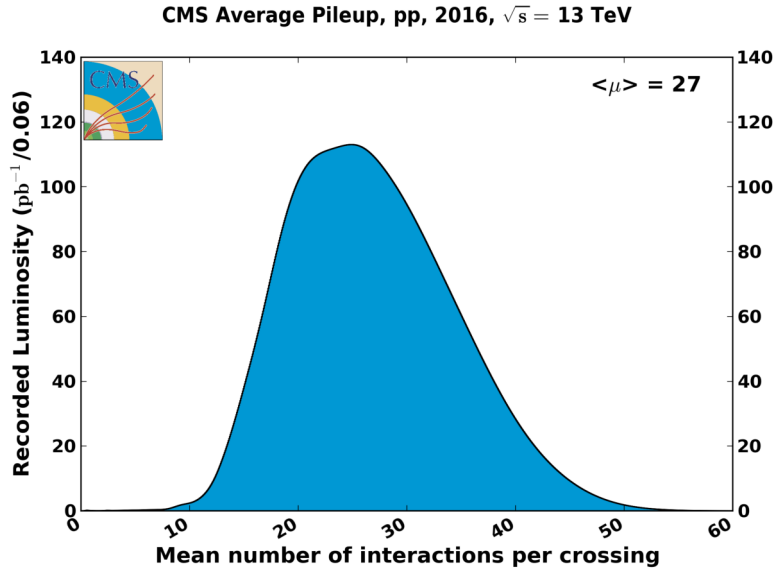


FIGURE 6.3: Mean number of interactions per bunch crossing for the 2016 pp run at 13 TeV. Information taken from https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults

run period used for the current analysis in data (dotted) and simulation, both before (dashed red line) and after (blue line) the reweighting has been applied. A significant improvement in the agreement between data and MC is found after the reweighting is performed.

- **B tag efficiency** The b tagging efficiency usually depends on the p_T and η of the jet, and is defined as

$$\epsilon_f(p_T, \eta) = \frac{N_f^{\text{tagged}}(p_T, \eta)}{N_f^{\text{total}}(p_T, \eta)} \tag{6.7}$$

where f stands for the flavour of the jet (if it has been initiated by a light, c or b quark). These efficiencies usually differ in data and MC, and scale factors need to be applied on the simulated samples to correct for these differences. Different methods can be used to apply the calculated scale factors ($\text{SF} = \varepsilon_{\text{DATA}}/\varepsilon_{\text{MC}}$) to MC simulated events. These are grouped into two general categories, depending on whether they involve event reweighting or not. In our analysis, we use a method that aims to correct the shape of the b-tag discriminator distribution, referred to as *discriminant reshaphing method* (or *event reweighting using discriminator-dependent scale factors*). The scale factors for jets in MC simulation are calculated as

$$SF(CSV, p_T, \eta) = \frac{\text{DATA} - \text{MC}_A}{\text{MC}_B}$$

where A/B = light/heavy flavour for heavy flavour SF or A/B = heavy/light

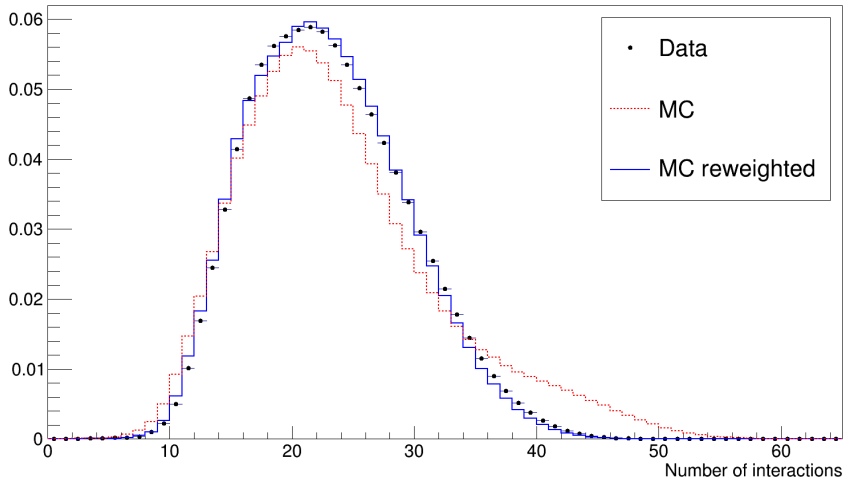


FIGURE 6.4: Distribution of the number of interactions in data (dotted) and simulation, before (red dashed line) and after (blue line) reweighting has been applied.

flavour for light flavour SF, thus accounting for different flavour contamination in each case. The *IterativeFit* [78] method is used to measure b tagging efficiency. This method is based on the calibration of the full b tagging discriminator shape and is designed to meet the needs of analyses in which the full distribution of the b tagging discriminator values is used.

A scale factor is applied to each jet, depending on its flavour and p_T and $|\eta|$ values, and the event weight is calculated as

$$SF(\text{total}) = \prod_{i}^{N_{\text{jets}}} SF_{jet_i}$$

where i goes through all the jets passing the selection in the event.

- **Muon identification and isolation** Muon identification efficiencies are estimated using the *tag and probe* method using $Z \to \mu^+\mu^-$ events. In the selected events, one muon is required to pass the identification criteria (*tag*). It is then measured in how many instances the other muon fulfills the identification criteria as well (*probe*) to infer the efficiency. The difference between the efficiencies measured in data and MC are corrected in simulation by applying (p_T, η)-dependent scale factors $(\epsilon_{\text{data}}/\epsilon_{\text{MC}})$ to simulated events.

  The scale factors were calculated separately for two different running periods, which had different data taking conditions. The scale factors for each of the two running periods are estimated as

$$SF_{\mu,i} = \epsilon^{ID}_{\mu,i} \cdot \epsilon^{ISO}_{\mu,i} \tag{6.8}$$

  where i stands for the run period under assumption. ϵ^{ID} describes the identification efficiency and ϵ^{ISO} stands for the isolation efficiency. The total scale

factor is calculated as the weighted sum

$$SF_{\mu} = \frac{\sum_i SF_{\mu,i} \cdot \mathcal{L}_i}{\sum_i \mathcal{L}_i} \tag{6.9}$$

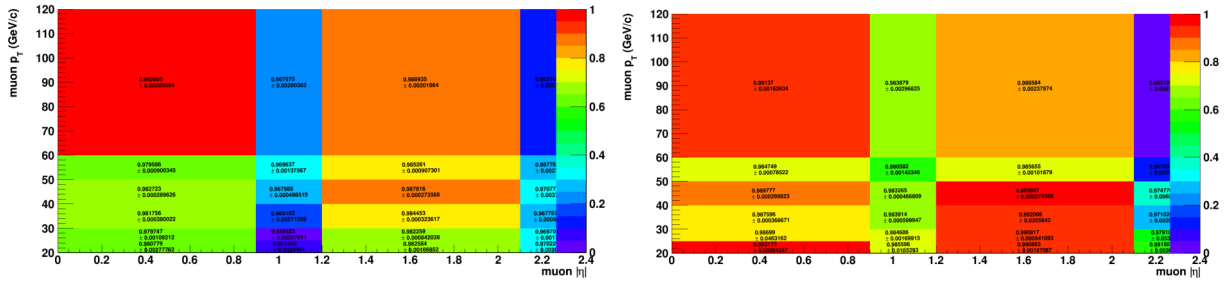
where $\mathcal{L}_i$ refers to the associated luminosity for each of the data taking periods. The uncertainties associated to these scale factors are calculated as:

$$\Delta SF_{\mu,i} = \sqrt{\left(\Delta\epsilon^{ID}_{\mu,i} \cdot \epsilon^{ISO}_{\mu,i}\right)^2 + \left(\epsilon^{ID}_{\mu,i} \cdot \Delta\epsilon^{ISO}_{\mu,i}\right)^2} \tag{6.10}$$

and the total uncertainty will therefore be calculated using the previous values for each running period as

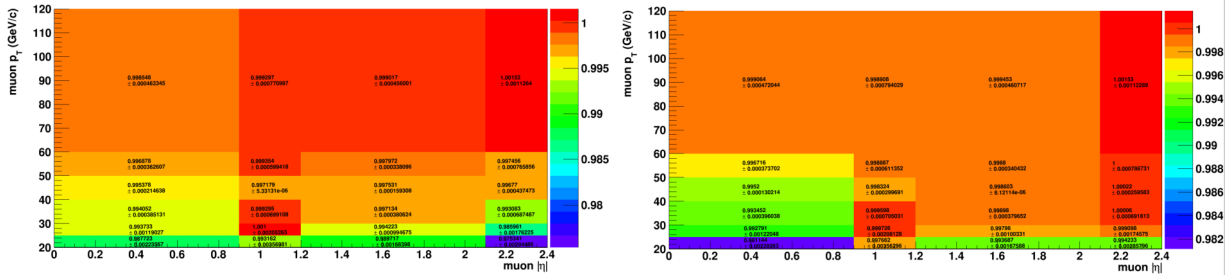
$$\Delta SF_{\mu} = \frac{\sqrt{\sum_i \left(\Delta SF_{\mu,i} \cdot \mathcal{L}_i\right)^2}}{\sum_i \mathcal{L}_i} \tag{6.11}$$

The values of the efficiency values used in the scale factor calculations are presented in figures 6.5 and 6.6 for the identification and isolation efficiencies, respectively.



(A) (B)

FIGURE 6.5: Muon tight identification efficiencies for running periods 1 (left) and 2 (right).



(A) (B)

FIGURE 6.6: Muon isolation efficiencies for running periods 1 (left) and 2 (right).

- **Electron identification and isolation** The scale factors for electron tight cut-based identification criteria are given in terms of the electron p_T and η values, and are presented in figure 6.7. The presented scale factors already contains combined information from isolation and identification efficiencies.

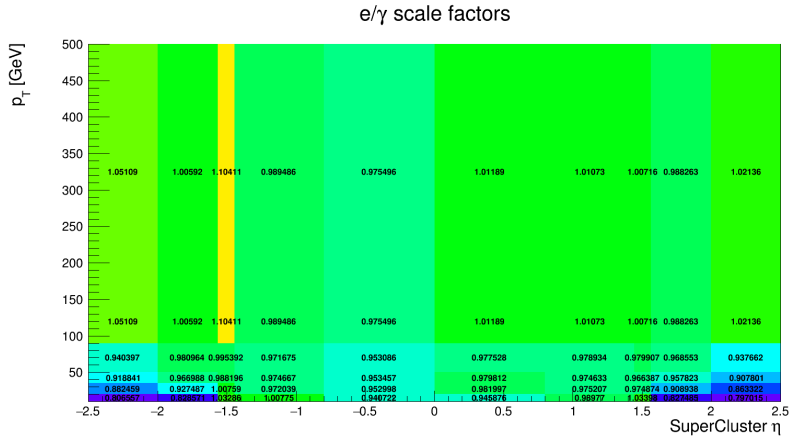


FIGURE 6.7: Electron cut-based ID tight working point scale factors.

- **Electron energy regression**: This correction relies on the information from the Monte Carlo generator to improve the reconstructed p_T response. The calculation is done using a multivariate regression technique in order to improve the determination of the electron (or other object) momentum with respect to the reconstructed value, allowing to get an additional correction beyond the standard CMS energy corrections. Regression is essentially a multi-dimensional calibration to the particle level, going from the reconstructed energy to the generated one, taking into account geometrical and cluster shape variables.

  The MVA regression is trained to predict the true energy (E_{true}) of the object under consideration, given the uncorrected supercluster energy (E_{raw}). This uncorrected energy is taken as the sum of individual crystal energies in a supercluster. After training, the regression predicts the full probability density function (pdf) for the inverse response E_{true}/E_{raw}, improving the electron energy resolution.

- **Electron energy scale and resolution smearing** Electrons are reconstructed using information from the tracker and the electromagnetic calorimeter (see 5.1.2), and their energy scale and resolution are derived using $Z \to e^+e^-$ events. The ECAL energy resolution is extracted from a maximum likelihood fit to the dielectron invariant mass distribution in terms of η of the final-state electrons[3] [93]. The energy resolution itself depends on the amount

[3]Two bins of R9 are also used. R9 is a cluster shape variable, defined as the ratio between the energy in a 3×3 crystal array around the most energetic crystal in the supercluster and the supercluster energy itself. Further information can be found in reference [93]

of material particles traverse before reaching the ECAL and other geometrical effects, and is found to be better in simulation than in data. Although the origin of the disagreement is not fully understood, better tracker material description, an improved clustering algorithm and simulation of the detector response lead to better results.

This mismatch in the energy resolution between data and MC is accommodated in the different analyses by applying an additional Gaussian smearing to the electron energy in MC events. Gaussian functions are often used to smear momentum distributions as they describe random fluctuations. However, if a detector has a particular bias so that the errors introduced are not random, different functions could be used for smearing. As can be seen in figure 6.8, the MC sample with a Gaussian smearing provides a good description of the detector response.

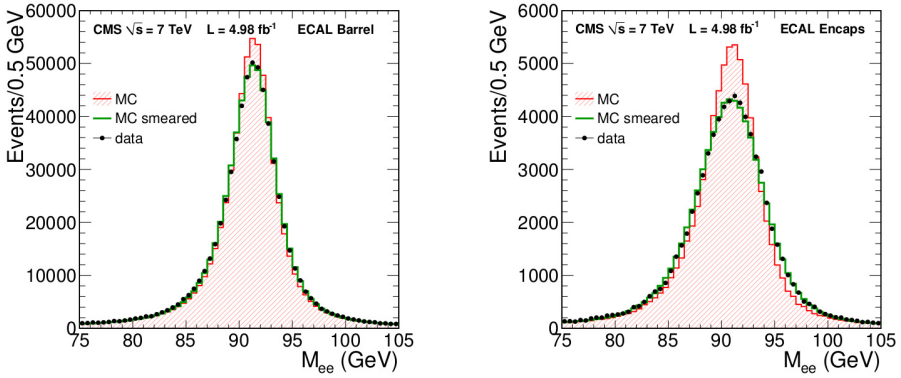


FIGURE 6.8: Distribution of the dielectron invariant mass for the default MC simulation (filled line), for the MC simulation with additional Gaussian smearing (green line), and for the data (dots). The distributions for events with both electrons in EB (left) and in EE (right) are displayed.

- **Jet energy resolution smearing**: The MC events are generated using specific resolution functions which might not be the same as observed in data. Measurements show that the jet energy resolution (JER) in data is worse than in the simulation and the jets in MC need to be smeared to describe the data. Thus, a set of corrections or scale factors derived for this effect are applied to the MC samples so as to coincide with the data, similarly to the electrons' case.

6.6.1 Trigger efficiency

The trigger efficiency is estimated both in data and MC. It is calculated as

$$\epsilon_T = \frac{N_{TRIG}^{3leptons}}{N_{TOT}^{3leptons}} \tag{6.12}$$

where $N_{TOT}^{3leptons}$ is the number of events that pass our offline baseline trilepton selection cuts before the trigger selection is applied, and $N_{TRIG}^{3leptons}$ is the number of events that pass the baseline trilepton selection, including the set of selected trigger paths (see section 6.3). No jet requirements are applied in any of the selections.

In order to perform an unbiased measurement of the trigger efficiency, it is of key importance to use a sample of events that does not contain any cut on the isolation requirement or the p_T threshold of the leptons and any track requirement. With that aim, a sample containing events collected by MET triggers is used to perform the efficiency calculation on data. This sample contains events with high missing transverse energy, but without any specification on the lepton p_T threshold, isolation or track requirements.

The trigger efficiency studies in MC, on the other hand, are based on the signal sample, with applied pileup and luminosity correction scale factors. Thus, the trigger efficiencies in each case are calculated as

$$\epsilon_{DATA} = \frac{N^{\text{Lepton+MET}}}{N^{\text{MET}}} \tag{6.13}$$

and

$$\epsilon_{MC} = \frac{N^{\text{Lepton}}}{N^{\text{Total}}} \tag{6.14}$$

where $N^{\text{Lepton+MET}}$ is the number of events passing lepton and MET triggers, N^{MET} and N^{Lepton} are the number of events passing either MET or lepton triggers, and N^{Total} is the total number of events.

Channel	ϵ_{Data}	ϵ_{MC}	$\epsilon_{\text{Data}}/\epsilon_{\text{MC}}$
$\mu\mu\mu$	$1.0000^{+0.0000}_{-0.0169}$	$0.9998^{+0.0001}_{-0.0002}$	1.000±0.017
$\mu\mu e$	$0.9907^{+0.0077}_{-0.0210}$	$0.9992^{+0.0003}_{-0.0004}$	0.991±0.021
$ee\mu$	$0.9915^{+0.0071}_{-0.0194}$	$0.9988^{+0.0003}_{-0.0004}$	0.993±0.019
eee	$0.9833^{+0.0138}_{-0.0373}$	$0.9974^{+0.0006}_{-0.0008}$	0.986±0.037

TABLE 6.7: Trigger efficiencies after the baseline trilepton selection, for data and MC. The uncertainties displayed are statistical uncertainties.

Then, the efficiencies were estimated as functions of the p_T of the most energetic lepton of the event. Since weights are not allowed in the tool used for the

calculations, no pileup corrections are applied for the MC sample. The efficiencies are shown in figure 6.9 for each channel.

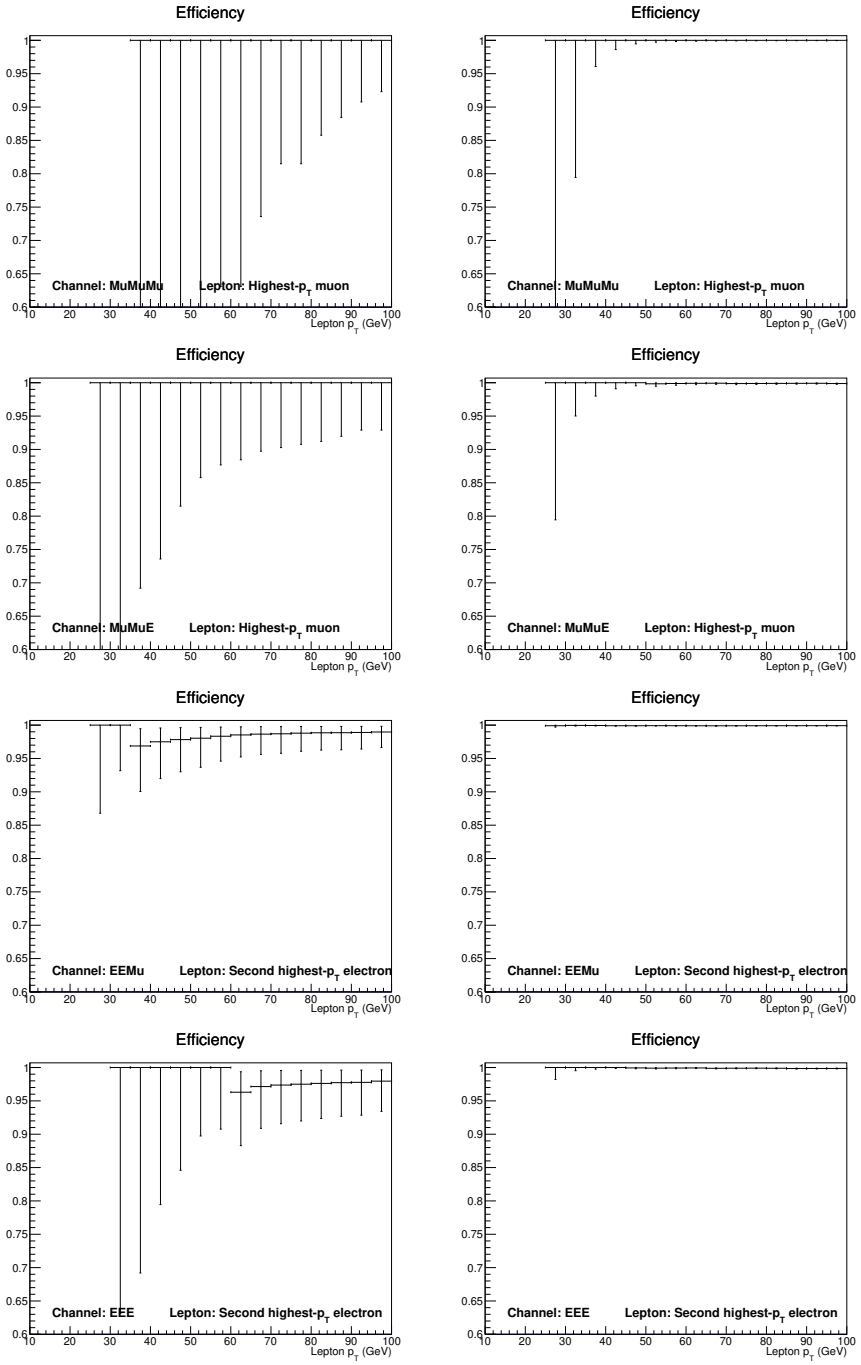


FIGURE 6.9: Trigger efficiencies as functions of the triggering lepton p_T in data (left) and in Monte Carlo (right) for each channel.

6.7 The non-prompt lepton (NPL) background

One of the most important background sources to the tZq signal under study, besides the already mentioned $t\bar{t}Z$ and WZ+jets processes, are events containing at least one non-prompt or *fake* lepton. The origin of fake leptons depends on the flavour of the lepton to be considered. In some cases, they can be objects wrongly reconstructed as leptons, or they might be real leptons themselves, but that come from other hadronic decays. Fake muons come primarily from semi-leptonic decays of heavy flavour (b- or c-) hadrons, whereas fake electrons arise both from hadronic decays or photon conversions. Since they originate from different sources, fake muons and fake electrons are treated separately. In the case in which these leptons pass the selection criteria, events containing fake leptons might be mistaken with signal event candidates.

The non-prompt lepton sample

The background events containing non-prompt leptons originate from, in order of importance, DY+jets production, $t\bar{t}$ events containing two leptons, and WW and tW processes. Each of these background sources contain two prompt and one non-prompt leptons. Given the low probability that a non-prompt lepton is identified as a prompt lepton, the contribution from events with more than one non-prompt lepton is negligible. Non-prompt electron (muon) templates are obtained from events containing exactly one non-prompt electron (muon), and two prompt leptons (either electrons or muons). In the NPL sample, the non-prompt leptons can be associated either with the top quark or with the Z boson candidates.

Fake leptons constitute a source of instrumental background, which means that their contribution is very difficult to be modelled in simulation. As a consequence, identifying and constraining the fake lepton background represents a challenging task. In contrast with the other background sources, which are estimated fully from MC simulation, the determination of the shape and the normalization of the NPL background must be done using data-driven techniques. Even though the presence of fake leptons is very small, the cross sections of the processes that originate these fake leptons are quite large in contrast to the signal production rate, and as a consequence its contribution becomes important.

Non-prompt lepton definition

To select a sample enriched of non-prompt leptons, we establish certain criteria the objects under consideration must fulfill.

- **Non-prompt muon**: it satisfies the same kinematical requirements as a prompt muon, but has instead a *loose* lepton ID (in contrast with *tight* ID from prompt muons) and has a relative isolation above 0.25 (as opposed to prompt muons, for which this value is below 0.15).

- **Non-prompt electron**: whereas prompt electrons are required to pass the *tight* cut-based identification requirements, non-prompt electrons are tagged as *veto* according to these criteria. Moreover, the isolation entering the cut-based selection is reversed ($Iso^{CB,veto}_{barrel} > 0.175$ and $Iso^{CB,veto}_{endcaps} > 0.159$). Finally, to remove photons from this sample, the object isolation is always required to be below 1, and the $1/E - 1/p$ variable is required to have the same cut as in the tight cut-based selection (see table 6.8).

Fake electron selection	
p_T	> 25 GeV
$\|\eta\|$	< 2.5
Electron cut-based ID	Veto
$\|1/E - 1/p\|$	< 0.0129
$I^{e}_{rel}(\Delta R = 0.3)$	> 0.159 but < 1.0 (e) > 0.175 but < 1.0 (b)
Fake muon selection	
p_T	> 25 GeV
$\|\eta\|$	< 2.4
Muon ID	Loose
Muon reconstruction type	PF global or tracker
$I^{\mu}_{rel}(\Delta R = 0.4)$	> 0.25

Table 6.8: Non-prompt electron and muon selection.

These differences are listed in table 6.9. Kinematical requirements are identical to those for prompt leptons. The NPL sample is then constructed from data, and is identical to the signal sample, except for the fact that one of the three leptons considered has to fulfill the fake lepton criteria just mentioned, meaning it satisfies a looser identification criteria, failing at the same time the isolation requirements.

Lepton	**Identification criteria**	**Isolation**
Muon	Tight ID	<0.15
Fake muon	Loose ID	>0.25
Electron	Tight CB	<0.0361 (barrel) & <0.094 (endcaps)
Fake electron	Veto CB	>0.1750 (barrel) & >0.159 (endcaps) <1 in both cases

Table 6.9: Selection criteria differences for prompt and non-prompt leptons.

6.8 Background control

In order to constrain the main background sources ($t\bar{t}Z$, WZ+jets and NPL), the following strategy is adopted. On a first step, events passing the baseline trilepton selection are taken as potential signal candidates. This initial sample contains events with exactly three leptons, two of which are required to be compatible with a Z boson decay. This means they need to have opposite sign, same flavour, and have a reconstructed invariant mass at least 15 GeV around the Z boson nominal mass. Events containing any additional leptons with p_T >10 GeV are vetoed, in order to reduce the contribution from background sources containing four leptons in the final state, such as ZZ, $t\bar{t}Z$ or $t\bar{t}H$, for instance.

This sample is then split in three statistically independent regions: the signal region and two control regions. These two control regions are defined to be as close as possible to the signal region in order to minimize the effect of extrapolating from one region to another. Thus, the only difference between the three subsamples remains in the jet and b jet multiplicities. If we take a look back at section 6.2.1, we see that

- tZq events are expected to have two jets, one of which has to be heavy (b-)tagged.
- $t\bar{t}Z$ events are expected to contain two of these b jets.
- No b jets are to be expected in $WZ + jets$ production.

The initial trilepton sample is then split as follows:

- **"1bjet" (signal enriched region)**: contains events with either two or three jets, and exactly one b jet.
- **"2bjet" ($t\bar{t}Z$ enriched region)**: events with more than one jet and more than one jet would fall in this region.
- **"0bjet" (WZ+jets and NPL enriched region)**: containing events with at least one jet, but without any b jets present.

This is summarized in table 6.10. By using this splitting, we assure that most of signal events will fall in the signal region, whereas the "2bjet" region will contain mostly $t\bar{t}Z$ events, which will make it easier to estimate and thus constrain the contribution from this background in the final fit. The same applies for the definition of the "0bjet" region, which contains mostly events originating from WZ process, along with DY+jets events which do not contain b jets but are nonetheless contributing to the fake lepton background (we consequently say that this region is also enriched in NPL background events).

Control region	Jet multiplicity	b jet multiplicity
1bjet (tZq)	2 or 3	$= 1$
2bjet ($t\bar{t}Z$)	> 1	> 1
0bjet (WZ+jets and NPL)	> 0	$= 0$

TABLE 6.10: Selection criteria used to split the initial 3 lepton sample and define the three statistically independent regions used in the final fit.

6.9 Data-driven estimation of the non-prompt muon and electron samples

NPL background contribution represents an instrumental source of contamination which cannot be modelled accurately by simulation and is therefore fully determined from data in the analysis, both its shape and its normalization. This source of contamination is probably the most challenging to treat.

The two main processes giving rise to these fake leptons are $t\bar{t}$ and Drell Yan production in association with jets (DY+jets). The expectation values of the fraction

$$\frac{N_{t\bar{t}}}{N_{\text{DY+jets}}} \tag{6.15}$$

of $t\bar{t}$ over DY+jets range from 0.06 for the *eee* channel when no jet requirements are made, to a non-negligible 0.77 for the $\mu\mu\mu$ channel when one b jet is required. Even though in DY+jets events the additional lepton (the one assigned as coming from the top quark decay) is more prone to be the non-prompt lepton, this is not the case in $t\bar{t}$ events, in which the non-prompt lepton may be either the additional or one of the opposite-sign same-flavour leptons attributed to the Z boson decay. In fact, the probability that the non-prompt lepton in $t\bar{t}$ events in the $\mu\mu\mu$ channel (where $N_{t\bar{t}} \simeq 0.77 \cdot N_{\text{DY+jets}}$) corresponds to the additional one is just 0.50. This means that associating the non-prompt lepton to the additional lepton from the top quark decay would not provide a good approximation to reality. Hence, the non-prompt leptons can be either associated to the decay of the top quark or to the Z boson candidate.

The NPL sample is thus selected to be identical to the signal sample in what refers to two of the three selected leptons, with the only difference that events are accepted if (only) one of the three leptons satisfies the requirements stated in section 6.7, meaning it has looser identification properties and reversed isolation. The same trigger selection and data samples as in the main sample are also used.

The shape of the distributions used in the multivariate analysis are provided by templates and the normalization is estimated in a two step process:

- **Pre-normalization**: a preliminar fit to the m_T^W distribution in the 0bjet region is performed. The use of this region to provide the relative NPL yields in the four channels is justified by the dominance of the DY process as a source of NPL background events in all three b tagging regions. In this prefit, the normalization of all other background sources is fixed to the SM prediction. This prefit yields a total of eight output normalization factors, corresponding to the four decay channels ($\mu\mu\mu$, $\mu\mu e$, $ee\mu$ and eee) for both the non-prompt electron and muon samples. Figure 6.10 shows the templates of the m_T^W distribution in the four different channels, with the (a priori) expected yields of the non-prompt muon and electron backgrounds. This templates and the output scale factors from the prefit will be used as input to the second step.
- **Final normalization**: The scale factors obtained in the preliminar fit will then be used as input in the final fit. However, even though they might describe well data behaviour in the 0bjet region, they do not assure an accurate description in the other two regions entering the fit. As a consequence, the normalization of the NPL samples is left free to vary in this final fit in order to be readjusted therein when fitting the three regions simultaneously, treating independently muons and electrons.

Further information on how the final fit is performed will be provided in the next chapter.

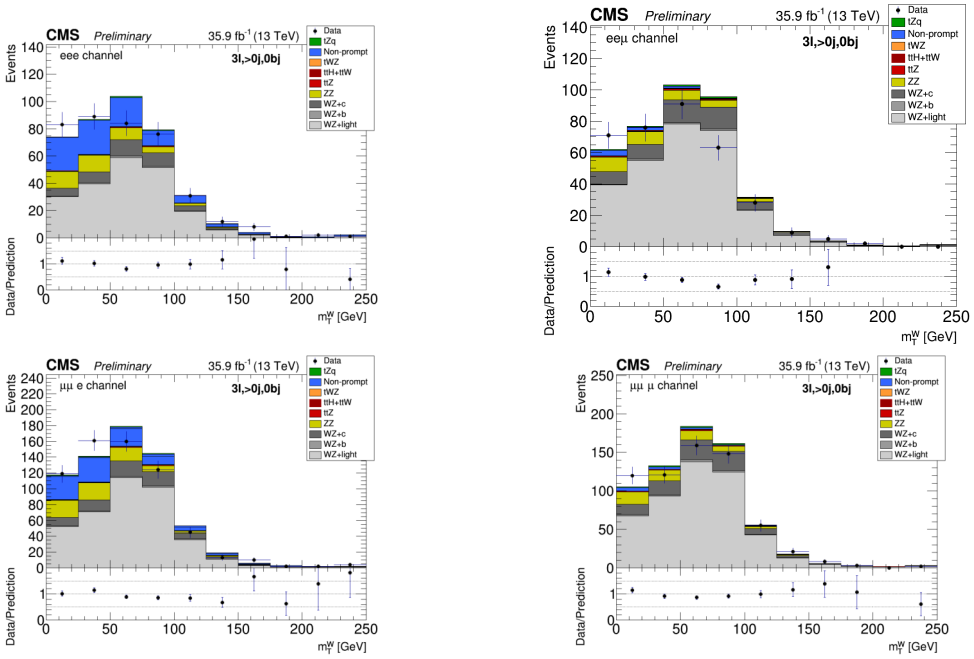


FIGURE 6.10: Prefit normalization of the fake lepton background contamination using the m_T^W variable in the 0bjet region for the four different channels (eee, $ee\mu$, $\mu\mu e$ and $\mu\mu\mu$).

Chapter 7

Shape Analysis

This chapter provides a description and brief introduction to the multivariate and statistical techniques used to obtain the tZq cross section measurement. A description of the boosted decision trees used in the analysis is presented; a dedicated section describes the different input variables used in the multivariate analysis, where the reconstruction of the top quark is also reviewed. A general overview of these multivariate techniques can be found in appendix A. The statistical analysis, implemented within the Combine framework, is also described along the last sections.

7.1 A multivariate analysis approach

The general procedure for identifying events of interest amongst high background environments in high energy physics often consists on applying a set of criteria or *cuts* on the eligible candidates so that the number of background events passing the selection is minimized while maximizing the number of true signal events. In some cases, however, the use of more sophisticated techniques might considerably boost the performance of the analysis. *Multivariate analysis* (MVA) techniques, for example, have proven to be quite successful in experiments characterized by a low signal over background ratio.

Decision trees [94] are one of these multivariate techniques, broadly used in high energy physics and some social sciences. Decision trees basically consist on a set of binary tests arranged in a tree-like structure comprising nodes, branches and leaves. Nodes split the data in two, according to the value of a particular attribute (i.e. if a given variable has a value higher or lower than certain optimized threshold) sequentially forming branches, until a leaf (an end-point in the tree structure) is reached. In the end, the information from all attributes is combined into a final variable that is used to discriminate between two initial hypothesis (i.e. signal or background candidates).

The time for building the decision trees is relatively short, as compared to other

multivariate methods such as neural networks, making them easier to study and develop. Different techniques are used to boost the performance of single decision trees. Appendix A provides more information on how decision trees and boosting work.

7.1.1 BDTs in the analysis

Analyzing rare processes (with small cross section values) generally requires applying more sophisticated techniques than, for instance, a simple counting experiment, in order to achieve a better sensitivity. Even after a selection optimization (see section 6.5), the number of background events is still overwhelming in these cases. Thus, with a view to improving the sensitivity to signal events and increasing signal-to-background separation, two *Boosted Decision Trees* (BDTs) were designed and trained in the analysis. This technique performs well in cases where a high degree of optimization is required. The distribution of the two BDT output discriminants will then be used as templates during the final fit, as will be described later on. These two classifiers are:

- **BDT tZq**: This first BDT is trained in the signal enriched (1bjet) region, and is implemented in order to increase the analysis sensitivity to signal events, enhancing their separation from all different background events.
- **BDT ttZ**: The second BDT entering the analysis is trained in the ttZ enriched (2bjet) region, to discriminate signal from $t\bar{t}Z$ events.

The low statistics of the 0bjet sample from data does not allow the usage of a BDT in that region.

Several user defined parameters (often referred to as *tunable tree parameters*) affect the training and boosting of the decision trees, and as a consequence also the performance of the analysis. These parameters are selected so as to increase the accuracy of the results and prevent from overtraining (appendix section A.3). BDTs in the analysis are implemented in the TMVA framework [95], a toolkit that hosts a large variety of multivariate classification algorithms integrated in ROOT. TMVA offers a set of configuration options to customize the different classifiers. These include:

- **Boosting type**: the type of boosting dictates how the different decision trees are sequentially built in the ensemble. The boosting type used in the analysis is called *gradient boost* (appendix section A.2.1) using the Huber loss function[1] which is implemented by default in TMVA.

[1]Loss function indicates the difference between target and predicted values, and the aim of boosting is to minimize this loss function. Huber loss is one specific type of these functions, and is described in the dedicated appendix.

- **Number of trees**: a large number of trees implies a deeper learning, which can lead to overtrain in the analysis. Therefore, this number has to be optimized to avoid this. The total number of trees used to built the tree ensemble is set to 200 in the analysis.

- **Granularity of the histograms used in variable cut optimization**: the training procedure selects the variable and cut value that optimizes the increase in the separation index (also called *impurity function*, which gives an idea of the quality of the signal-to-background separation achieved at each node) between the parent node and the sum of the indices of the two daughter nodes, weighted by their relative fraction of events.

  These cuts are optimized scanning over each variable range with a *granularity* set by the option `nCuts`. This means that two different histograms with a total of `nCuts+1` bins are built. One of them is filled with signal events, the other one with background events. The range of the two histograms lies between the given variable range, and the width of the bins is uniform (the larger the value of `nCuts`, the narrower the bins). Then, for each bin, the difference between signal and background is separated, and that bin for which this separation is highest dictates the value of the variable cut. In the analysis, this granularity parameter is set to `nCuts = 200`.

- **Shrinkage**: this parameter allows to adjust the learning rate of the algorithm. A small value (0.1-0.3) demands more trees to be grown, but can significantly improve the accuracy of the prediction in difficult settings. However, it comes at the price of increasing computational time both during training and testing, as a lower learning rate requires more iterations. In our analysis this value is set to 0.4.

- **Maximum allowed depth of the tree**: *pruning* is the process of reducing the size of the tree by turning some branch nodes into leaf nodes, and removing the leaf nodes under the original branch. Lower branches may be strongly affected by outliers, and a simpler tree often avoids overtraining. Figure 7.1 shows a tree of depth 3. This parameter is set to 2 in our BDTs.

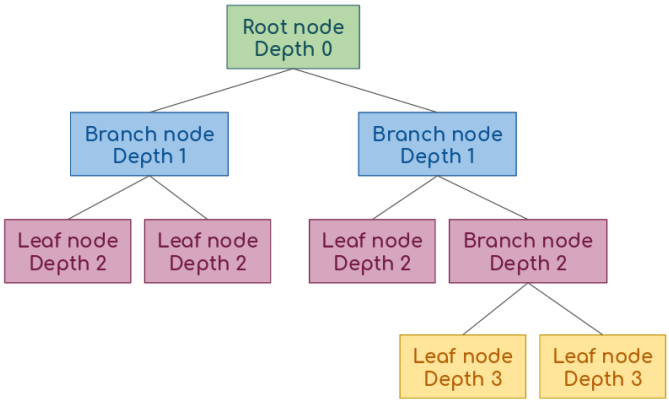


FIGURE 7.1

- **Treatment of negative event weights in training**: in NLO Monte Carlo generators, events with (unphysical) negative weights may occur in some phase space regions. Such events are often troublesome to deal with, and it depends on the concrete implementation of the MVA method, whether or not they are treated properly. In cases where a method does not properly treat events with negative weights, it is advisable to ignore such events for the training (but to include them in the performance evaluation to not bias the results). This can be explicitly requested for each MVA method via a boolean configuration option. In our case, this option was set so that the negative weights were ignore at the training stage.

7.2 Input variables to the BDTs

The selection of the variables used in the MVA has a strong dependence on the characteristics of the analysis. In general, these variables are expected to provide a higher degree of discrimination between signal and backgrounds, and are related to the topology and kinematics of signal events. The estimation of the variables used for the two BDTs is described in the following.

7.2.1 Z boson and top quark reconstruction

Several variables included in the BDT rely on the full reconstruction of the final state Z boson and top quark, which proceeds through the combination of the four momentum of their decay products. A good reconstruction relies on the correct identification of these decay products, and for that, leptons and jets are assigned to either a Z or a top quark according to the following criteria:

Leptons assignment. The Z boson reconstruction is done first, assuming that it decays in two leptons of same flavour and opposite sign (as discussed in 6.5.2). Lepton assignment is therefore trivial in the cases of $ee\mu$ and $\mu\mu e$ channels: the Z boson decay products are trivially chosen as the e^+e^- and $\mu^+\mu^-$, respectively, and the *additional lepton* (μ or e, respectively) is assigned as a decay product of the top quark.

In the channels containing 3 leptons of the same flavor, this choice is not possible, and in this case, the pair of leptons yielding an invariant mass closest to the nominal Z boson mass of 91.2 GeV is chosen as decaying from the Z boson, while the additional (third) lepton is assigned as a decay product of the top quark.

As shown in Fig. 7.2, the top quark reconstruction requires further identification of two other decay products: a neutrino and a b jet.

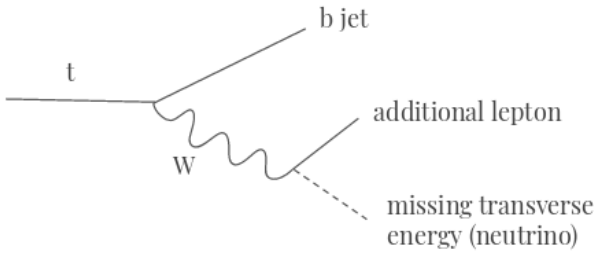


FIGURE 7.2: Top quark reconstruction is performed with the information from the decay particles (the neutrino, the additional lepton and a b jet candidate), imposing constraints on the mass of the W boson and the quark itself.

Neutrino p_Z reconstruction and assignment Momentum conservation allows to infer the neutrino momentum in the transverse plane, based on the energy balance in the calorimeters (see section 5.5): the missing transverse energy (MET) is assigned to the transverse momentum of the neutrino. However, the momentum in the z plane cannot be derived directly due to the unknown momentum of the proton and antiproton remnants that scatter at low angles. In this case, the neutrino p_z can be estimated from the combined information of the measured charged lepton momentum and the inferred neutrino momentum in the transverse direction (MET), constraining the reconstructed W mass to its known value.

For the massless neutrino, we have

$$E_\nu = \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} \tag{7.1}$$

and for the W boson reconstructed from the additional lepton ℓ and the neutrino ν

$$m_W^2 = E_W^2 - \overrightarrow{p}_W^2 = (E_\ell + E_\nu)^2 - (\overrightarrow{p}_{T,\ell} + \overrightarrow{p}_{T,\nu})^2 - (\overrightarrow{p}_{z,\ell} + \overrightarrow{p}_{z,\nu})^2 \tag{7.2}$$

where the subindex T represents the momentum component in the transverse XY plane, whereas z describes the momentum component along the longitudinal direction. This leads to a quadratic equation with different $p_z(\nu)$ solutions. There are two different possibilities:

- **Unique complex solution**: the square root term in the $p_{z,\nu}$ solution is negative. Such an unphysical result can simply indicate that this particular $(\ell + \nu)$ combination is not compatible with the hypothesis of coming from a W boson. That may happen e.g. if the event is actually a background event that does not contain a W boson. But even in the case of signal events, a complex solution may arise from an incorrect lepton assignment, or from instrumental sources, e.g. a lepton or MET resolution effect, producing a spurious imbalance on the measured transverse-plane energy. To avoid losing

any signal, events with unphysical solution are not discarded; instead, the negative square-root term is set to zero, and the resulting $p_{z,\nu}$ (real) solution is used to reconstruct the neutrino.

- **Double real solution**: in this case there are two possible $p_{z,\nu}$ solutions, either taking the positive or the negative square root of the quadratic p_z solution. In this case, the two solutions are tested as valid hypotheses for the top quark reconstruction, and the one yielding a reconstructed top quark mass closest to the reference value is retained.

b jet assignment Once the longitudinal momentum of the neutrino has been estimated, the top quark reconstruction proceeds with the assignment of the last decay product, the b jet. There are three possibilities:

- **0bjets region**: in case there are no b jets in the event, all selected jets are tested as potential decay products of the top quark.
- **1bjet region**: in case there is one b jet in the event, the b jet is assigned to the top quark.
- **2bjets region**: in this case, the two b jets will be considered as top quark decay product candidates.

After testing all possible combinations of the neutrino $p_{z,\nu}$ solutions and b jet candidates described above, together with the additional lepton, the top quark reconstruction used further in the analysis is the one that yields a reconstruction mass

$$m_{top} \equiv M_{(\text{lepton+neutrino+bjet})}$$

closest to the reference value. The nominal values used to constrain the top quark and the W boson masses during $p_z(\nu)$ calculation and jet assignment are 172.5 GeV and 80.38 GeV, respectively. The distribution of the reconstructed system mass is presented in figure 7.3.

7.2.2 The Matrix Element Method

In this section an overview of the Matrix Element Method (MEM), from which computed weights are used as input to the two BDTs used in the analysis is provided. The MEM is a powerful reweighting tool that allows to have an estimate of the probability of each event to be compatible with the signal or the different background hypotheses.

The method, originally designed to study $t\bar{t}$ events at the D0 and CDF experiments, was first introduced at the Tevatron collider to achieve a more precise

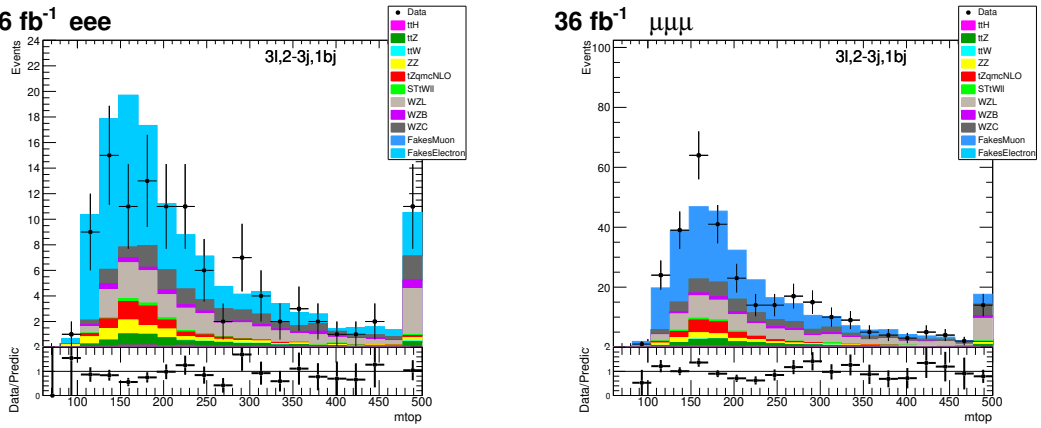


FIGURE 7.3: Reconstructed top quark mass distributions in the *eee* and $\mu\mu\mu$ channels for the signal region (1-bjet).

measurement of the top quark mass [96] and played an important role in the discovery of single top electroweak production [97]. First proposed by Kondo [98, 99] and later by Dalitz and Goldstein [100, 101], it has also proven very useful in analyses where a small signal has to be extracted from large backgrounds.

This technique combines theoretical and experimental information, thus being model dependent. Model-independent methods have the advantage of not having to deal with theoretical uncertainties. However, most of them do not make use of or provide any information on most of the properties of the particles under study, and as a consequence it is often useful to consider complementary and model-dependent tools both in searches of new phenomena and in precision measurements. The MEM is one of such methods, making maximal use of both experimental information and the theoretical model on an event-by-event basis. Matrix Element Method uses the global event information at the reconstruction level, similarly to MVA methods, with the difference that the latter only use a subset of that information. The use of the MEM in our analysis serves as a complement to the multivariate classifiers.

One of the advantages of the method is that it is universal and can be applied to a wide variety of particle processes for which theoretical models have been established. In contrast with multivariate methods, it does not need training, offering a good discrimination between the different hypotheses even with limited statistics, which can be an issue in searches for rare processes. Despite all these advantages, the implementation of the MEM reweighting is not straightforward and is very computationally intensive as it works on an event-by-event basis.

It consists in estimating the probability of an event of being compatible with the signal and background hypotheses, by the computation of the different processes cross sections at a given point of the phase space, corresponding to the reconstructed kinematic properties of the event. Given a theoretical assumption α, a weight $\omega_{i,\alpha}$ is assigned to each event i that quantifies the validity of each theoretical frame for this event. The value of the weight is the probability to observe the event i in the

theoretical frame α. In our analysis, the different hypothesis are either signal (tZq), $t\bar{t}Z$ or WZ+jets. This weights are computed as

$$\omega_{i,\alpha}(\Phi') = \frac{1}{\sigma_\alpha}\int d\Phi_\alpha \cdot \delta^4\left(p_1^\mu + p_2^\mu - \sum_{k\geq 2} p_k^\mu\right) \cdot \frac{f(x_1,\mu_F)f(x_2,\mu_F)}{x_1 x_2 s} \cdot \left|\mathcal{M}_\alpha(p_k^\mu)\right|^2 \cdot W(\Phi'|\Phi_\alpha)$$

where σ_α is the cross section of the process α, Φ' is the 4-momenta of the reconstructed particles in the event, $d\Phi_\alpha$ are the process-dependent integration variables, corresponding to the 4-momenta of all the particles at the vertex in the hypothesis α, the δ symbol represents the momentum conservation between incoming and final state particles, $f(x,\mu_F)$ are the parton density function in the proton, x_1, x_2 are the fraction of proton energy carried by the incoming particles, $\left|\mathcal{M}_\alpha(p_k^\mu)\right|^2$ is the matrix element squared, and W are the transfer functions (see section 7.2.2) relating the energy of particles at the vertex with their energy reconstructed with the detector.

The mass of the W boson from the top decay follows a Breit-Wigner distribution, as does the mass of the virtual Z boson in the $t\bar{t}Z$ hypothesis (interference with $\gamma^\star$ is included in the computation in the matrix element). A narrow-width approximation ($\Gamma \ll m_{top}$) is used for the top quark.

Jet assignment to objects at parton level

Jet assignment used for event reconstruction for the MEM is different to that described in 7.2.1 for top quark reconstruction, as in this case we are not only interested in reconstructing the top quark, but rather in doing so for the complete tZq set.

To compute the MEM weights, selected leptons, jets and b jets need to be accurately associated to the leptons and quarks at parton level. Since the correct assignment is not known a priori, the ME is evaluated for all possible permutations of the selected leptons and jets, and an average weight is computed for each hypothesis, from which the maximum is taken. The assignment in the MEM is done as follows:

- **Signal region (1 b jet, 1 or 2 jets)**:
  - ◇ *Signal hypothesis*: the jet with the highest CSVv2 discriminant value is assigned to the b jet from the top quark. If there is only one additional quark, it is associated to the additional forward quark. If there are two jets, we consider only the one with highest $|\eta|$.
  - ◇ $t\bar{t}Z$ *hypothesis*: a permutation is performed with the highest CSVv2 jet, which is assigned consecutively to the two b quarks arising from the top and antitop decays (the other one is considered to not have been

reconstructed). Another permutation is performed with the other jets, which are assigned to the quarks from the hadronic W decay.

- ◇ *WZjj hypothesis*: the b-tagged and the highest-p_T jets are assigned to the two quarks.

- $t\bar{t}Z$ **enriched region (more than one b jet and more than one jet)**:
  - ◇ *Signal hypothesis*: the two jets with highest CSVv2 discriminator value are selected and permutation over them is performed for the assignment of the b quark. The highest $|\eta|$ jet is associated to the forward quark.
  - ◇ $t\bar{t}Z$ *hypothesis*: again we select the two jets with highest CSVv2 discriminator and permute over them (and associate them to the top and the antitop). Among the remaining jets, we select those with mass closest to the nominal W boson mass, and permute over them to assign them to the first or the second quark from the hadronic W.

Transfer functions

Transfer functions $W(\Phi'|\Phi_\alpha)$ measure the probability of observing the set of physical observables Φ' under the assumption of the phase space point Φ_α at matrix element level. No transfer functions are used for leptons and quarks (they are assumed to be 1) considering that:

- The energy and direction of leptons is assumed to be perfectly measured.
- Direction of quarks is assumed to be perfectly measured from the direction of the reconstructed jets.

These transfer functions are evaluated in simulation, and used only for jets and b jets. If a jet is expected at matrix element level, but has not been reconstructed, its transfer function is set to 0 (1) if the associated quark lies in $|\eta| > 2.4$ (< 2.4). The pdfs are histograms parametrized in terms of the ratio E_{rec}/E_{gen} where E_{rec} is the reconstructed energy of jets after corrections have been applied, and E_{gen} is the energy of the associated quarks at generator level. Transfer functions are also used to constrain the total momentum in the transverse (XY) direction at parton level from the total reconstructed momentum.

MEM discriminants used in the analysis

Some MEM variables are used as input to the BDTs. Four of them are used in the 1bjet (signal) region and two of them in the 2bjet ($t\bar{t}Z$ enriched) region. The different variables used are listed below.

- Only in the signal (1bjet) region:
  - ◇ Log-likelihood ratio of tZq hypothesis versus $t\bar{t}Z$ hypothesis.
  - ◇ Log-likelihood ratio of tZq hypothesis versus $t\bar{t}Z$ hypothesis with $t\bar{t}Z$ and tZq weights rescaled so that their mean values are similar.
  - ◇ Log-likelihood ratio of tZq hypothesis versus $t\bar{t}Z$ + WZ hypothesis.
- Only $t\bar{t}Z$ (2bjet) enriched region:
  - ◇ Logarithm of the MEM score associated to the most probable $t\bar{t}Z$ kinematic configuration.
- Both in the 1bjet and 2bjet regions:
  - ◇ Logarithm of the MEM score associated to the most probable tZq kinematic configuration.

The use of the MEM increases the sensitivity of the analysis by about a 20%. This effect can be seen in figure 7.4. Some of these variables are among the most discriminating variables (the ranking of these variables is shown in table 7.2).

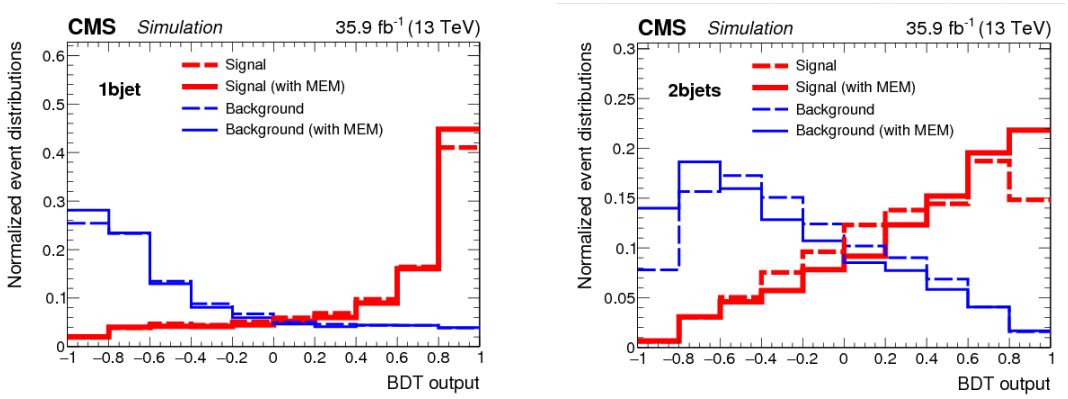


FIGURE 7.4: Normalized distributions of the BDT output for signal (thick lines) and backgrounds (thin lines) from simulation for the tZq (left) and $t\bar{t}Z$ (right) regions. The discriminators including and excluding MEM variables in the BDT training are shown, respectively, as solid and dashed lines. Contributions from the four considered channels are included in the signals and backgrounds.

7.2.3 Complete list of input variables to the BDT

The BDT is optimized to keep a minimal number of significant variables, to keep the algorithm faster and avoid possible overtraining. In this optimization study, the first step was to remove the variables that were strongly correlated than others. From

each pair of strongly correlated variables, the one with slightly better discriminating power was kept (the final correlation matrix is displayed in Fig. 7.5). Then, the variables ranked by the BDT as least discriminant were removed from the BDT for testing, one at a time, and left out if their presence did not improve the BDT performance. The remaining variables, actually used in the construction of the BDTs, include masses, kinematics and angular distributions involving the recoiling (forward) jet, the reconstructed top quark and Z boson and their decay products. The complete list is:

- **b tagging properties**

  The output distribution of the CSVv2 algorithm discriminant of each of the different jets considered in the event is used as input to the BDT (`btagDiscri` or d_{CSV}).

- **Reconstructed Z boson properties**

  The following variables related to the Z boson, reconstructed as described in 7.2.1, are used:

  - η of the Z boson (`ZEta` or η_Z).
  - p_T of the Z boson (`Zpt` or p_T^Z).

- **Reconstructed top quark properties**

  The following variables related to the top quark, reconstructed as explained in 7.2.1, are used in the BDT:

  - Top quark mass (`mtop` or m_{top}).
  - Top quark decay lepton (the additional lepton, associated to the W boson decay) asymmetry. It is defined as the product of the lepton electric charge and its absolute η value: $q_\ell \cdot |\eta_\ell|$ (`AddLepAsym` or $Asym_\ell$).
  - η of the top quark decay lepton (`AddLepETA` or η_ℓ).

- **Recoiling jet (q') properties**

  The recoiling jet is taken as the selected jet with highest p_T value which is not the b jet (if the number of jets in the event is two or more).

  - η of the recoiling jet (`etaQ` or η_Q).
  - p_t of the recoiling jet (`ptQ` or p_T^Q).

- **Other kinematical properties regarding different reconstructed objects**

  - ΔR separation between the jet identified as a b quark and the recoiling jet, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = \sqrt{(\phi_b - \phi_Q)^2 + (\eta_b - \eta_Q)^2}$ (`dRjj` or ΔR_{jj}).

- ΔR separation between the top quark decay lepton and the jet closest to it (`dRAddLepClosestJet` or $\Delta R_{j\ell}$).
- Azimuth angle separation ($\Delta\phi$) between the top quark decay lepton and the Z boson (`dPhiZAddLep` or $\Delta\phi_{Z\ell}$).
- Azimuth angle separation between the top quark decay lepton and the b quark (`dPhiAddLepB` or $\Delta\phi_{b\ell}$).
- ΔR separation between the top quark decay lepton and the recoil jet (`dRAddLepQ` or $\Delta R_{Q,\ell}$).
- ΔR separation between the Z boson and the top quark (`dRZTop` or $\Delta R_{Z,\text{top}}$).
- η of the jet with highest p_T (`LeadJetEta` or η_j).
- Number of jets in the event (`NJets` or N_{jets}).

- **MEM discriminants:**
  - Log-likelihood ratio of the tZq hypothesis against the $t\bar{t}Z$ hypothesis. (MEMvar_0 or $\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$).
  - Logarithm of the MEM score associated to the most probable tZq kinematic configuration. (MEMvar_1 or $\log(\omega_{\text{tZq}})$).
  - Logarithm of the MEM score associated to the most probable $t\bar{t}Z$ kinematic configuration. (MEMvar_2 or $\text{KIN}\omega_{(\text{t}\bar{\text{t}}\text{Z})}$).
  - Log-likelihood ratio of the tZq hypothesis against the $t\bar{t}Z$ hypothesis with $t\bar{t}Z$ and tZq weights rescaled such that their mean values are similar. (MEMvar_3 or $\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$).
  - Log-likelihood ratio of the tZq hypothesis against the $t\bar{t}Z$ + WZ hypothesis. (MEMvar_8 or $\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z}-\text{WZ})}$).

Table 7.1 lists all these variables just described, indicating which ones are used in each of the BDTs. The shapes of the distributions of these input variables are shown in figures B.1 and B.2 of Appendix B, for the BDTs of 1bjet and 2bjet regions, respectively. The distributions are shown for both signal and background in the $\mu\mu\mu$ channel.

The correlations between the different variables are shown in figure 7.5 and the overtraining test and background rejection are shown on figure 7.6. The importance of each variable is estimated by removing (one at a time) the variables from the training and calculating the variation of the expected significance with respect to the case were all the variables are included. Since this process is very CPU-consuming, this calculation did not include all systematic uncertainties described in Section 7.4, and only the dominant ones were kept. The ranking of the five most discriminating variables in both BDTs for the different channels is shown in table 7.2. The complete ranking of variables is presented in Appendix B.

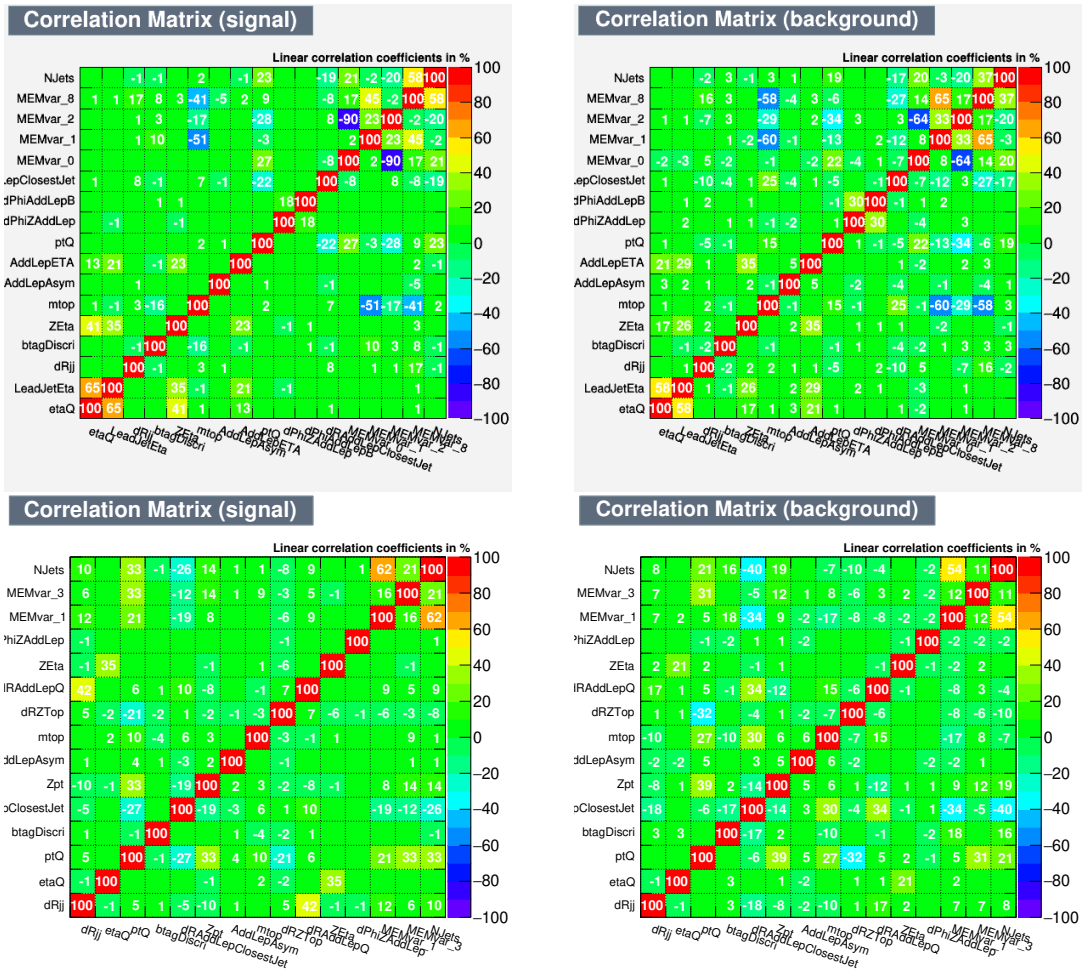


FIGURE 7.5: Correlation matrices for the signal (left) and the backgrounds (right), in the $\mu\mu\mu$ channel for the BDT trained in the 1bjet region (top) and in the 2bjet region (bottom).

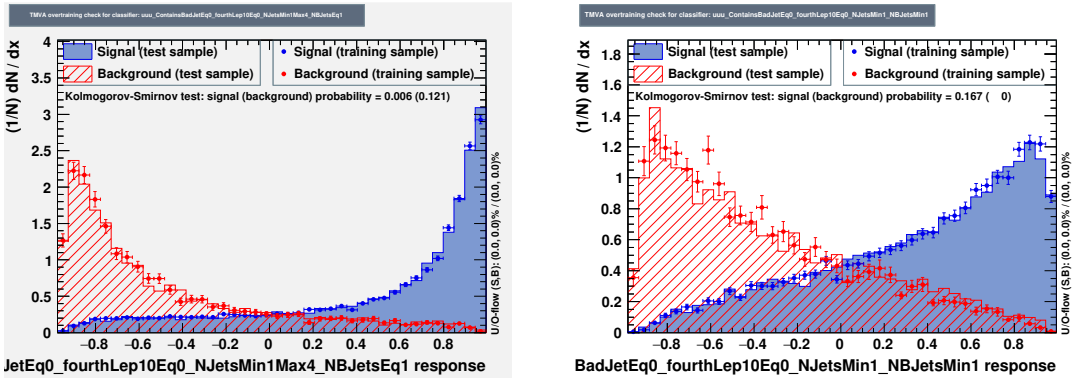


FIGURE 7.6: Overtraining test for the two BDTs used in the analysis: the one trained in the 1bjet region (left) and in the 2bjet region (right), both figures showing the results obtained in the $\mu\mu\mu$ channel.

	Variable description	1bjet	2bjet
1 d_{CSV}	CSVv2 algorithm discriminant	Y	Y
2 ΔR_{jj}	ΔR separation between the jet identified as a b quark and the recoiling jet	Y	Y
3 η_Q	η of the recoiling jet	Y	Y
4 p_T^Q	p_t of the recoiling jet	Y	Y
5 η_Z	η of the Z boson	Y	Y
6 m_{top}	Top quark mass	Y	Y
7 $\Delta R_{j\ell}$	ΔR separation between the top quark decay lepton and the jet closest to it	Y	Y
8 $Asym_\ell$	Top quark decay lepton asymmetry	Y	Y
9 $\Delta\phi_{Z\ell}$	Azimuth angle separation between the top quark decay lepton and the Z boson	Y	Y
10 $\Delta\phi_{b\ell}$	Azimuth angle separation between the top quark decay lepton and the b quark	Y	N
11 η_ℓ	η of the top quark decay lepton	Y	N
12 η_j	η of the jet with highest p_T	Y	N
13 $\Delta R_{Q,\ell}$	ΔR separation between the top quark decay lepton and the recoil jet	N	Y
14 $\Delta R_{Z,\text{top}}$	ΔR separation between the Z boson and the top quark	N	Y
15 p_T^Z	p_T of the Z boson	N	Y
16 N_{jets}	Number of jets	N	Y
17 $\log(\omega_{\text{tZq}})$	Logarithm of the MEM score associated to the most probable tZq kinematic configuration	Y	Y
18 $\text{KIN}\omega_{(\text{t}\bar{\text{t}}\text{Z})}$	Logarithm of the MEM score associated to the most probable $\text{t}\bar{\text{t}}\text{Z}$ kinematic configuration	N	Y
19 $\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	Log-likelihood ratio of the tZq hypothesis against the $\text{t}\bar{\text{t}}\text{Z}$ hypothesis	Y	N
20 $\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	Log-likelihood ratio of the tZq hypothesis against the $\text{t}\bar{\text{t}}\text{Z}$ hypothesis with $\text{t}\bar{\text{t}}\text{Z}$ and tZq weights rescaled such that their mean values are similar	Y	N
21 $\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z}-\text{WZ})}$	Log-likelihood ratio of the MEM weights for $\text{t}\bar{\text{t}}\text{Z}$ against $\text{t}\bar{\text{t}}\text{Z}$ + WZ hypothesis	Y	N

TABLE 7.1: Description of the variables used in the BDTs. The symbol Y (N) in the third and fourth columns indicates that the variable was (was not) used in the 1bjet and 2bjet BDTs.

Ranking	eee	$ee\mu$	$e\mu\mu$	$\mu\mu\mu$
BDT 1bjet (tZq)				
1	ΔR_{jj}	ΔR_{jj}	p_T^Q	d_{CSV}
2	η_j	d_{CSV}	η_j	m_{top}
3	$Asym_\ell$	η_Q	ΔR_{jj}	ΔR_{jj}
4	p_T^Q	m_{top}	m_{top}	$\Delta R_{j\ell}$
5	η_ℓ	η_j	d_{CSV}	p_T^Q
BDT 2bjet ($\text{t}\bar{\text{t}}\text{Z}$)				
1	m_{top}	N_{jets}	N_{jets}	N_{jets}
2	N_{jets}	$Asym_\ell$	m_{top}	$\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$
3	d_{CSV}	d_{CSV}	$Asym_\ell$	m_{top}
4	η_Q	$\log(\omega_{\text{tZq}})$	ΔR_{jj}	$Asym_\ell$
5	$Asym_\ell$	m_{top}	d_{CSV}	$\Delta R_{j\ell}$

TABLE 7.2: Ranking of the five most discriminating variables in the BDT for the 1bjet and 2bjet regions, in the four different channels.

7.2.4 Control plots

The prefit data-to-prediction comparison for some of the most discriminant variables that enter the BDTs is shown in figure 7.7 for the 1bjet region. These distributions include events from all four channels combined. The distributions for all variables of Table 7.1, both for combined and individual channels, are shown in appendix B.2.

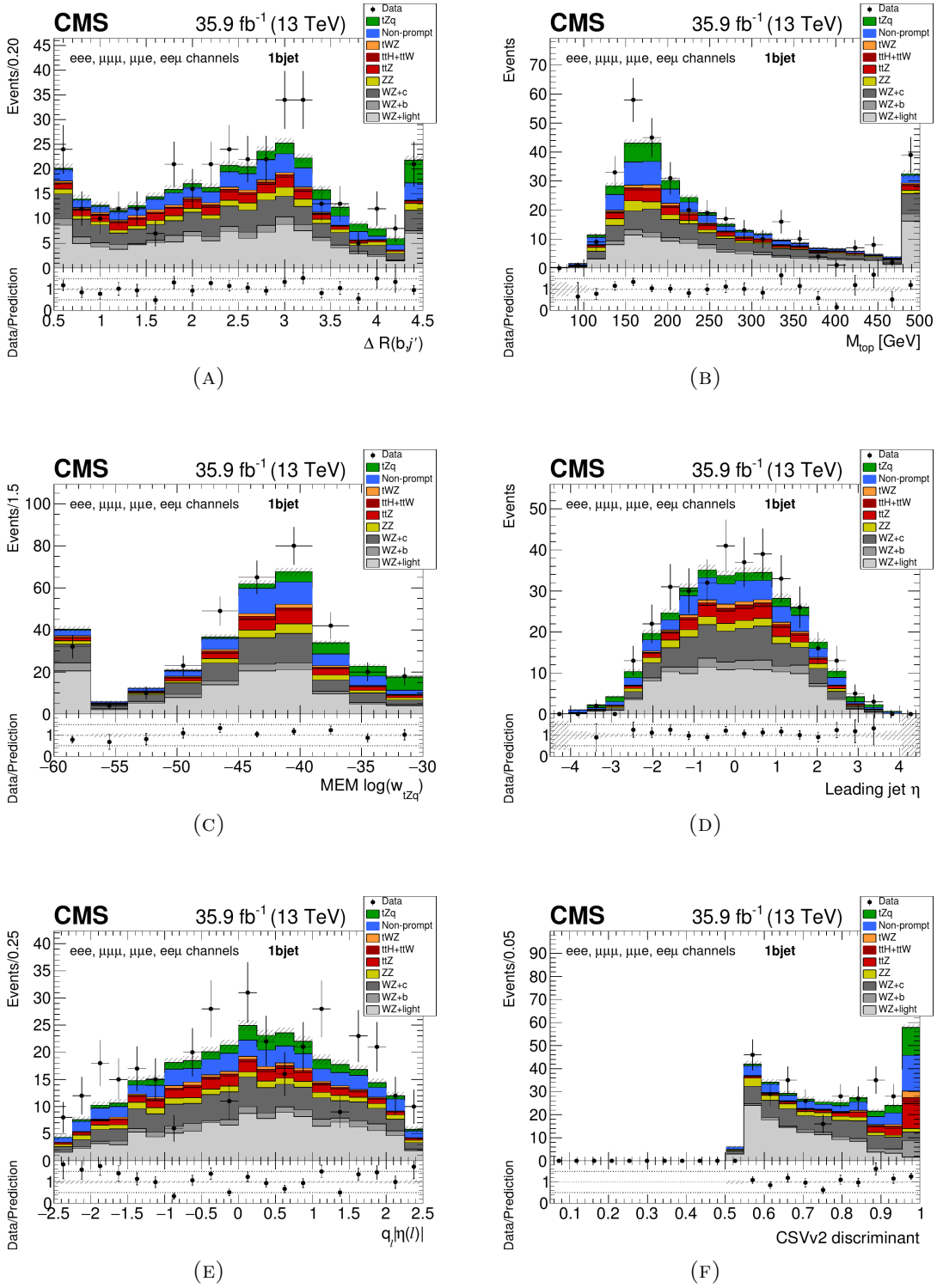


FIGURE 7.7: Data-to-prediction comparisons for some of the most discriminating variables in the 1bjet region (signal region) for all the channels combined. The distributions include events from all final states. Underflows and overflows are shown in the first and last bins, respectively. The predictions correspond to the normalizations obtained before the final fit (*prefit*). The hatched bands include the total uncertainty on the background and signal contributions. The pulls in the distributions are shown in the bottom panels.

7.3 Shape analysis

The tZq cross section is measured using the data collected with the CMS experiment during the whole 2016 period, corresponding to an integrated luminosity of 35.9 fb^{-1}. The statistical procedure used for the signal extraction is presented in this section.

7.3.1 Inputs for the shape analysis: templates

The analysis relies on a binned maximum likelihood fit, using the RooStats-based Combine framework described in section 7.3.3. The fit is performed simultaneously on the three previously defined, statistically independent regions (0bjet, 1bjet and 2bjet), for the four leptonic channels (eee, e$\mu\mu$, $\mu\mu\mu$ and eeμ), yielding therefore a simultaneous fit of 12 templates. The templates distributions in each region are defined as:

- **BDT-tZq**: the output discriminant distribution from the BDT trained in the 1bjet (signal) region, to optimize the separation between the tZq signal and the different background sources.
- **BDT-t$\bar{t}$Z**: the discriminant distribution from the BDT trained in the 2bjet (t$\bar{t}$Z enriched) region, to better constrain the contribution from this background.
- $\mathbf{m_W^T}$: the distribution of the transverse W mass in the 0bjet region, used to constrain the WZ+jets and the NPL backgrounds.

A scheme of the templates usage according to the regions and leptonic channels is displayed in figure 7.8.

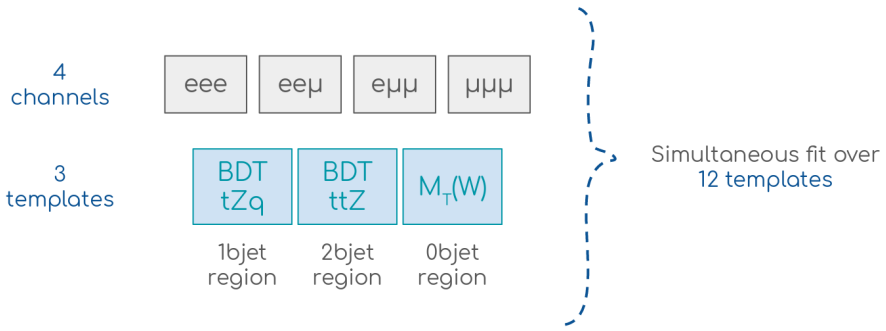
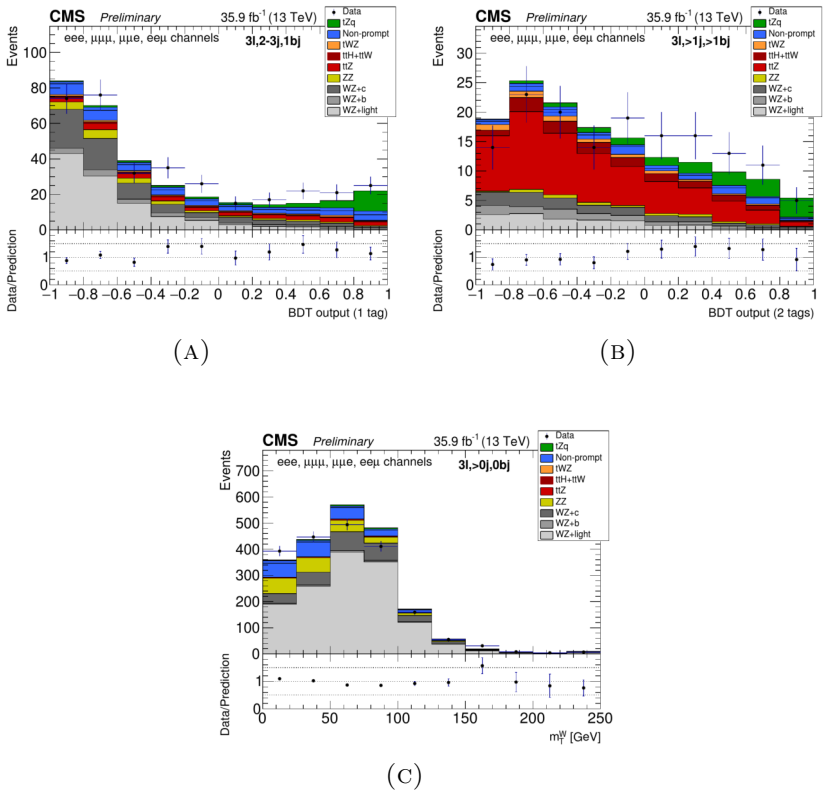


FIGURE 7.8: Scheme of how the final simultaneous fit is performed.

The templates distributions combining events from the four leptonic channels are shown in figure 7.9. The individual template distributions separated for each channel, as used as inputs for the Combine framework, can be found in section C.2 of the appendices.



(A) (B)

(C)

FIGURE 7.9: Prefit data-to-prediction comparison plots for the BDT discriminant in the 1bjet (signal) region, the BDT discriminant in the 2bjet (tt̄Z enriched) region and the m_T^W distribution in the 0bjet (WZ+jets enriched) region.

7.3.2 Likelihood model

For each bin i of the different distributions, the expected number of events can be parametrized as

$$\langle n_i \rangle (\mu, \theta) = \mu s_i(\theta) + b_i(\theta) \tag{7.3}$$

where s_i and b_i represent respectively the expected amount of signal and background in that bin, normalized according to the predicted cross sections and taking into account all sources of systematic uncertainties, represented by θ, as nuisance parameters. In Eq.7.3, the μ parameter determines the signal strength for a given theoretical model. This parameter is defined as the ratio of the production cross

section of the observed signal (σ_{obs}) with respect to the theoretical prediction,

$$\mu = \frac{\sigma_{obs}}{\sigma_{SM}}. \tag{7.4}$$

The SM prediction corresponds by construction to $\mu = 1$, the value $\mu = 0$ corresponding to the absence of signal, or background-only hypothesis. In order to take into account the contribution coming from the NPL background yields, which will be denoted by $B_i^\ell(\theta)$ for each lepton flavour ($\ell = \mu, e$), two additional terms are added to the previous expression (7.3):

$$\langle n_i \rangle (\mu, \theta) = \mu s_i(\theta) + b_i(\theta) + \alpha_e B_i^e(\theta) + \alpha_\mu B_i^\mu(\theta) \tag{7.5}$$

where the two factors α_ℓ determine the normalization of the non-prompt electron and muon backgrounds. These parameters are let to vary free in the fit.

The likelihood associated to an observation of $N = \sum_i n_i$ events in data compatible with a given hypothesis is given by the product of the Poisson probabilities of each bin i:

$$\mathcal{L}(data|\mu, \theta) = \prod_i \frac{\mu s_i(\theta) + b_i(\theta) + \alpha_e B_i^e(\theta) + \alpha_\mu B_i^\mu(\theta)}{n_i!} \exp\left[-\mu s_i(\theta) - b_i(\theta) - \alpha_e B_i^e(\theta) - \alpha_\mu B_i^\mu(\theta)\right] \tag{7.6}$$

The simultaneous fit to the data templates (BDT discriminators or the transverse W boson mass, depending on the region) in the four channels maximizes this likelihood, from which the measured (*observed*) cross section is extracted according to its relation to the signal strength (7.4), this parameter left to vary free in the fit. The systematic uncertainties enter the fit as *nuisance parameters*.

In our analysis, the nuisance parameters are associated with log-normal priors [2]. A log-normal distribution represents a random variable whose logarithm follows a normal distribution, its standard form being as follows:

$$\rho(\theta, \hat{\theta}) = \frac{1}{\sqrt{2\pi}\ln\kappa}\frac{1}{\theta}\left(-\frac{(\ln(\theta/\hat{\theta})^2}{2(\ln\kappa)^2}\right) \tag{7.7}$$

which means that $\Theta = \ln\theta$ follows a Gaussian distribution

$$\rho(\Theta, \hat{\Theta}) = \frac{1}{\sqrt{2\pi}\sigma}\left(-\frac{(\Theta - \hat{\Theta})^2}{2\sigma^2}\right) \tag{7.8}$$

with mean value $\hat{\Theta} = \ln\Theta$ and standard deviation $\sigma = \ln\kappa$. Only the uncertainty related to the NPL background normalization is associated with a Uniform prior (meaning that in this case $\rho(\theta)$ is constant). The different sources of systematic uncertainties that enter the analysis are detailed in the following section (7.4).

[2] A prior probability distribution of an unknown quantity is the probability distribution that would express one's beliefs about this quantity before some evidence is taken into account.

7.3.3 Combine

The statistical analysis is based on a RooStats-based software tool used for statistical analysis developed for the combination oh Higgs measurements between ATLAS and CMS, usually referred to as Combine. The input to Combine, is given in the form of plain ASCII files (referred to as *datacards*) that contain the information that defines the details of the experiment. This same format is used whether the experiment is a simple counting experiment or a shape analysis.

Some information must be available in the datacards, such as the number of channels (12 in our case), the number of background sources and the number of nuisance parameters (`kmax`). The datacards also contain information regarding the different source of systematic uncertainties, how they are treated in the fit (the kind of prior distribution that will be used: in our analysis either log-normal or uniform priors) and the relative effect of the systematic uncertainty on the rate of each process in each channel. Information on whether the different systematics affect the normalization or the shape of the distributions is also contained in the datacards. The tool allows the fit to be performed either simultaneously in all three regions for the four possible leptonic channels, or separately for each leptonic channel.

In our implementation of the tool, the signal strength is computed using a Maximum Likelihood Fit while the expected and observed significances are computed from a simple profile likelihood approximation.

7.4 Systematic uncertainties

Systematic uncertainties affect the multivariate in different ways: they can alter the shape and/or the normalization of the template distributions; systematic uncertainties modifying the values of the variables used as selection criteria can have an impact on the sample acceptance. In the latter case, as an effect of the systematic variations, final state objects (e.g. selected jets, b-jets, leptons) can be added to or removed from the event, potentially modifying the classification of the whole event according to the selection region. The different sources of systematic uncertainties considered in the analysis are presented below.

- **Luminosity** The instantaneous luminosity is measured at CMS using the silicon pixel detector, the closest detector to the beam pipe. The method for measuring the luminosity relies in the fact that the instantaneous luminosity is proportional to the number of reconstructed clusters in the pixel detector for each bunch crossing. It is measured by comparing the event rate with the visible cross section as $\mathcal{L} = R/\sigma_{vis}$. Calibration of the detectors is performed with dedicated Van der Meer (VdM) scans during the data taking. The overall uncertainty in the measured integrated luminosity recorded in 2016 by the CMS experiment is estimated to be 2.5% [102].
- **Pileup** Pileup reduces the accuracy of the jet energy measurement and might as well produce additional jets that do not originate from the hard interaction.

As a consequence, pileup degrades the resolution of the measured MET, affecting therefore the top quark reconstruction. To reduce possible bias in our analysis, the number of simulated pileup events is corrected to match the number of pileup events measured in data. Its uncertainty is obtained by varying the minimum bias cross-section used to perform the pileup simulation. The uncertainty on the total inelastic cross-section is taken to be 4.6%, considered only in the shapes of the distributions.

- **Trigger efficiency** The trigger efficiency is measured to be 100% both in data and MC. The uncertainty related to the trigger efficiency is estimated by varying the data-to-simulation normalization by a $\pm 1\%$ to predict the yields in the $\mu\mu\mu$ and $ee\mu$ channels, and a $\pm 2\%$ for the $e\mu\mu$ and eee ones, in order to account for residual differences in the efficiency measurements between data and simulation.

- **Lepton identification and isolation efficiency** The correction factors applied on simulation so that lepton reconstruction, identification and isolation in data are well described are obtained using tag-and-probe methods. Uncertainties on these factors are another source of systematic uncertainties, and their impact on the analysis is assessed by varying the factors for muons and electrons independently, within one standard deviation.

- **Jet energy scale and resolution** The jet energy scale systematic comes from the non-linear response of the hadronic calorimeter in the energy measurement. Jet energy scale and resolution corrections are both varied within their corresponding uncertainties, the observed changes propagated to all related kinematic quantities, including the missing transverse momentum.

- **b tagging efficiency** As detailed in the previous chapter, the difference between b tagging efficiency in data and simulation is accounted for by the application of discriminator-dependent scale factors to simulated events via an event reweighting method. Given that the CSV distributions and other jet variables are input variables to the BDT algorithm, a significant systematic effect is expected to arise from this correction. Variations that account for different effects on the b tagging and mistagging efficiency scale factors are taken into account. Variations due to the uncertainties on the light flavour contamination (*lf*) and linear and quadratic statistical fluctuations (*hfstats1* and *hfstats2*) are applied to heavy flavour jets; heavy flavour contamination (*hf*) and linear and quadratic statistical fluctuations, (*lfstats1* and *lfstats2*), are applied to light flavour jets; linear and quadratic uncertainties (*cferr1* and *cferr2*) are applied to charm flavour jets. Uncertainties on the b tagging efficiency due to JES uncertainties (*jes*) affect both light and heavy flavour jets, and they are studied coherently with the standard jet energy scale systematic variations of the previous bullet. Each variation is assigned a different, independent, nuisance parameter. The scale factors themselves are varied within one standard deviation.

- **Normalization of simulated backgrounds** The uncertainty on the normalization of the different background contributions is conservatively taken to be 30%. One nuisance parameter is associated to each background source, and they are let to vary independently from one another.

- **NPL background estimation** The shape-related uncertainties on the NPL backgrounds are estimated by varying the isolation criteria used to determine the NPL

sample. The shape variations of the NPL muon and electron backgrounds involve different nuisance parameters.

- **Theory uncertainties only applied on signal (tZq) modeling** The effect of the parton showering (PS) scale in the Pythia8 generator is evaluated using dedicated samples with the PS scale varied by factors 2 and 1/2. This shift in the PS scale is equivalent to modifying the value of α_S. This variation is only taken into account for the signal modelling.

- **Theory uncertainties applied on signal and background modeling** These uncertainties affect the shape of the signal as well as the shape and normalization of the simulated background distributions, except for tWZ events, for which only normalization uncertainties from scales and PDF are considered.

  - ◇ *Matrix Element factorization and renormalization scales* The renormalization and factorization scales, at the matrix element level, are set to an identical value, which depends on the event generator and on the simulated processes. In particular, the scales for the simulated signal are set to $\sum \sqrt{m^2 + p_T^2}/2$, where the sum runs over all particles in the final state. The scales are varied up and down by a factor of 2.

  - ◇ *Factorization and renormalization scales at parton shower level* These are identical to the matrix element scales, and are also varied by factors of 0.5 and 2. This uncertainty is only evaluated for the signal sample.

  - ◇ *Parton Distribution Function* The uncertainty from the choice of the PDF is determined by reweighting the simulation using the 100 different variations available in the NNPDF set, the total uncertainty taken as the RMS of these variations, following the PDF4LHC recommendations.

- **Limited size of the samples** The uncertainties due to the limited size of the samples used to build the templates are accounted as additional nuisance parameters. One independent nuisance parameter is included for each bin of the templates with a statistical uncertainty above 5%.

The different systematic uncertainties affect the measurement differently. They can have an impact on the sample acceptance, or on the normalization or shape of the templates. This is described in table 7.3.

Source	Variation	Type
Non-prompt muon rate	free (fit together with signal)	norm.
Non-prompt electron rate		norm.
Non-prompt muon shape		shape
Non-prompt electron shape		shape
WZ+ h.f. jets scale	Q^2 ×2,1/2	shape +acceptance
WZ+ h.f. jets norm.	30%	norm.
WZ+ l.f. jets scale	Q^2 ×2,1/2	shape +acceptance
WZ+ l.f. jets norm.	30%	norm.
ZZ	30%	norm.
ZZ	Q^2 ×2,1/2	shape +acceptance.
$t\bar{t}$ +Z	30%	norm.
$t\bar{t}$ +Z	Q^2 ×2,1/2	shape +acceptance.
$t\bar{t}$ +H	30%	norm.
tWZ	30%	norm.
$t\bar{t}$ +H	Q^2 ×2,1/2	shape +acceptance.
Trigger	±1% ,±2%	norm.
Lept. sel.	1%	norm.+shape
JES	$\pm 1\sigma(p_T, \eta)$	shape+acceptance
JER	$\pm 1\sigma(p_T, \eta)$	shape+acceptance
b tagging	$\pm 1\sigma(p_T, \eta)$	shape+acceptance
pileup	$\pm 1\sigma$ of the min.bias σ	shape+acceptance
PDF	PDF4LHC recipe	shape+acceptance (S and B)
signal scales ME	Q^2 ×2,1/2	shape+acceptance
signal scales PS	Q^2 ×2,1/2	shape+acceptance
lumi	2.5%	norm.

TABLE 7.3: Sources of systematic effects, along with their uncertainty values, introduced as nuisance parameters in the likelihood fit. The column labelled *Type* represents how the uncertainty affects the measurement.

Chapter 8

Results and interpretation

This chapter summarizes the main results obtained in the analysis. The SM tZq trilepton production cross section is reported, and relevant postfit results are also presented in the chapter, such as the final yields of the different processes in the three considered regions and the final templates. The effect of the systematic uncertainties is also reviewed.

8.1 Postfit results: yields and data-to-simulation plots

In this section a detailed description of the yields and templates obtained after the final fit is presented. These results can be compared with the prefit yields and templates presented in appendix C.

8.1.1 Postfit yields

The expected number of events (prefit yields) of the different SM processes composing the selected samples is obtained from simulation, using their cross section values (see table 6.3). The exceptions are the two NPL components, for which data-driven methods were used, as explained in section 6.9. In the fitting process, the different nuisance parameters are adjusted in order to maximize the likelihood function. The postfit normalization of the signal and background processes are the result of this adjustment.

Table 8.1 presents the expected (prefit) and observed (postfit yields) of the different processes considered in the analysis in the 1bjet or signal enriched region. Columns two to five of the table show the predicted yields obtained from the final fit, considering one leptonic channel at a time (eee, $e\mu\mu$, $\mu\mu\mu$ and $ee\mu$ from left to right), along with their statistical uncertainties for the different SM processes (listed in the first column). The sixth column presents the expected yields for all the four channels combined, and the last column shows the ratio between the postfit and prefit predicted yields. The prefit yields disclosed by leptonic channel can be found in table C.1. The two last rows in the table

show the total predicted yields after the fit ("Total") and the observed yields in data ("Observed"), respectively.

Process	**eee**	**e$\mu\mu$**	**$\mu\mu\mu$**	**eeμ**	All channels	$\frac{\mathbf{N^{postfit}}}{\mathbf{N^{prefit}}}$
tZq	5.0±1.5	8.5±2.5	12.3±3.6	6.6±1.9	32.3±5.0	—
t$\bar{\text{t}}$Z	3.7±0.7	6.1±1.2	8.0±1.5	4.7±0.9	22.4±2.2	0.9±0.2
t$\bar{\text{t}}$W	0.3±0.1	0.7±0.2	0.6±0.2	0.3±0.1	1.9±0.3	1.0±0.2
ZZ	4.8±1.3	9.0±2.5	7.8±2.2	3.2±0.9	24.7±3.6	1.3±0.3
WZ+b	3.0±0.9	4.6±1.4	5.5±1.7	3.4±1.1	16.6±2.6	1.0±0.2
WZ+c	9.0±2.4	18.0±4.9	24.2±6.5	13.7±3.7	64.8±9.3	1.0±0.2
WZ+light	12.2±1.6	22.4±2.8	29.1±3.4	16.6±2.0	80.3±5.1	0.7±0.1
ttH	0.6±0.2	1.0±0.3	1.5±0.4	0.9±0.3	4.0±0.6	1.0±0.2
tWZ	1.0±0.3	1.7±0.5	2.4±0.7	1.3±0.4	6.5±1.0	1.0±0.2
NPe	19.2±3.1	17.9±2.8	—	0.6±0.1	37.7±4.2	—
NPμ	—	31.1±9.9	15.3±4.9	7.2±2.3	53.6±11.3	—
Total	58.8±4.8	121±12	107±10	58.4±5.5	345±18	
Data	56	104	125	58	343	

TABLE 8.1: Observed and expected postfit yields for each production process in the 1bjet (signal) region. The yields of columns 2 to 5 correspond to each channel, and column 6 displays the total for all channels. The last column displays the ratio between postfit and prefit yields.

Tables 8.2 and 8.3 show the postfit yields for the 2bjet and 0bjet control regions, respectively, which also contain a fraction of tZq events. A non-negligible 12.7% over the total of events in the 2bjet region corresponds to tZq events and about a 1.1% of the events in the 0bjet region correspond to signal.

The fit constrains the normalization of the different background processes. The last columns of the three tables show that the prefit and postfit values are relatively close to each other in most cases, except for the ZZ component and the WZ+light background. In the WZ+light case the yield deviates from the SM prediction by about a 30%.

This postfit yield for the WZ+light component was verified in different ways, and its impact on the analysis evaluated. Firstly, the predicted shapes of the kinematic variables relevant to the analysis were controlled in the WZ+light enriched region (see section B.2.3), and verified to describe the data. The analysis was then repeated with the WZ+light normalization relative uncertainty increased from 30% to 50%. In this case, the results of the cross section measurement changed only about a half a percent. As a final cross check, a direct measurement of the WZ+light cross section was performed. In this case, tZq enters the fit as a background source, with its normalization fixed to its SM prediction, and an associated nuisance parameter to account for a 30% uncertainty on the tZq cross section. The final fit in this case is performed only in the 0bjet region, yielding a signal strength value for the WZ+light measurement of

$$\mu(\text{WZ} + \text{light}) = 0.73 \pm 0.11$$

Process	eee	e$\mu\mu$	$\mu\mu\mu$	eeμ	All channels	$\frac{N^{postfit}}{N^{prefit}}$
tZq	3.0±0.9	5.4±1.6	7.2±2.2	3.8±1.2	19.4±3.1	1.3±0.2
t$\bar{t}$Z	10±2	16±3	21±4	13±2	60±5	0.9±0.2
t$\bar{t}$W	0.4±0.1	0.8±0.2	0.8±0.2	0.6±0.2	2.6±0.4	0.9±0.2
ZZ	0.9±0.3	1.4±0.4	1.5±0.4	0.8±0.2	4.5±0.7	1.4±0.3
WZ+b	1.5±0.5	1.8±0.6	2.8±0.9	1.9±0.6	8.0±1.3	0.9±0.2
WZ+c	2.1±0.7	2.7±0.8	4.6±1.5	2.2±0.7	11.6±1.9	0.9±0.3
WZ+light	1.6±0.3	2.3±0.5	3.5±0.7	1.7±0.3	9.1±0.9	0.7±0.2
t$\bar{t}$H	1.4±0.4	2.9±0.8	3.7±1.1	2.0±0.6	10.0±1.5	0.9±0.2
tWZ	0.9±0.3	1.3±0.4	1.9±0.6	1.1±0.3	5.3±0.8	1.0±0.2
NPe	4.0±0.6	5.0±0.8	—	0.1±0.0	9.1±1.0	1.1±17.1
NPμ	—	7.5±2.4	4.0±1.3	2.5±0.8	14.1±2.8	3.8±58.4
Total	25.8±2.3	46.7±4.3	50.7±4.9	29.6±2.9	152.8±7.5	
Data	25	38	51	37	151	

TABLE 8.2: Observed and expected (postfit) yields for each production process in the 2bjet (t$\bar{t}$Z enriched) region. The yields of columns 2-5 correspond to each channel, and column 6 displays the total number of all channels summed. The last column displays the ratio between postfit and prefit yields.

compatible with the value derived in the main analysis. It is worth noticing that prefit yields of the three WZ+jets subsamples are estimated from the reference cross section for the inclusive WZ production. However, Figure 2 of reference [103] shows that the jet multiplicity of a direct WZ cross section is not well described by this SM prediction. The requirement of at least one or two jets on the analysis may therefore be the cause of the larger difference between the SM prediction and observed yields in this case. This feature, however, has a low impact on the tZq cross section measurement, as verified in several steps of the analysis (discussed also in section 8.2).

8.1.2 Postfit templates

The distributions of the templates in each channel are modified according to the likelihood estimations of the nuisance parameters, implying on a reshaping and renormalization of the templates of the different processes. Figures 8.1, 8.2 and 8.3 show the postfit data-to-prediction comparison plots for the output discriminant distribution of the main BDT in the 1bjet region, the discriminant of the second BDT trained in the 2bjet control region and the m_T^W variable in the 0bjet control region, respectively. These can be compared with the prefit plots shown in appendix C.

As observed in figure 8.1, the tZq signal (shown in green) populates the high region of the BDT discriminant, in the signal region, as expected. The contribution of the NPL background is distributed along all BDT values, and the lower values are populated mostly by events from double-boson processes, specially WZ production in association with light and c jets (shown in light and dark grey, respectively). In channels where the

Process	**eee**	**e$\mu\mu$**	**$\mu\mu\mu$**	**eeμ**	**All channels**	$\frac{\mathbf{N^{postfit}}}{\mathbf{N^{prefit}}}$
tZq	3.6±1.1	6.2±2.0	8.5±2.7	4.8±1.5	23.0±3.9	1.3±0.2
t$\bar{t}$Z	2.3±0.5	3.7±0.9	5.0±1.2	2.9±0.7	13.9±1.7	1.0±0.2
t$\bar{t}$W	0.1±0.0	0.3±0.1	0.3±0.1	0.2±0.1	1.0±0.2	0.9±0.2
ZZ	55±16	101±28	73±21	41±12	269±40	1.4±0.3
WZ+b	4.3±1.3	6.8±2.1	9.2±2.8	3.9±1.2	24.2±3.9	1.0±0.2
WZ+c	42±12	73±20	95±27	54±16	264±39	1.0±0.2
WZ+light	161±16	304±27	383±33	217±20	1064±50	0.8±0.1
t$\bar{t}$H	0.3±0.1	0.5±0.1	0.7±0.2	0.4±0.1	1.9±0.3	1.0±0.2
tWZ	0.8±0.2	1.3±0.4	1.7±0.5	1.0±0.3	4.8±0.7	1.0±0.2
NPe	106±17	118±19	—	1.0±0.2	225±25	1.2±17.2
NPμ	—	40±13	43±14	29±9	112±21	3.8±53.9
Total	375±30	654±50	620±50	355±30	2004±82	
Data	387	640	637	345	2009	

TABLE 8.3: Observed and expected (postfit) yields for each production process in the 0bjet (WZ+jets and fakes enriched) region. The yields of columns 2-5 correspond to each channel, and column 6 displays the total number of all channels summed. The last column displays the ratio between postfit and prefit yields.

lepton associated to the top quark decay is an electron (eee and $e\mu\mu$), the data points are below the simulation for bins situated near the upper tail of the BDT discriminant (this effect is more relevant in the $e\mu\mu$ channel). Comparing with the prefit templates of the BDT discriminant in the 1bjet region (figure C.1), it can be observed that the data-to-prediction ratio in these bins is already close to one. In order to better describe the data, the fit increases the yield of the NPL component, leading the total predicted yields to overestimate the data in the signal region.

This difficult interplay between NPL and signal is the limiting factor on the significance of the measurement. One potential solution for this problem would have been to train an additional BDT against the NPL background. If that would be possible, the contribution of the NPL to the high-BDT region would decrease, and as a consequence the significance of the analysis would be better. However, the low statistics in the NPL sample does now allow the training of a dedicated BDT, preventing also the NPL contribution to be well constrained using the current templates.

The postfit distribution of the BDT output discriminant in the 2bjet region is presented in figure 8.2. A significant increase in the NPL component is observed in general, the effect being specially noticeable in the $e\mu\mu$ channel (see table 8.2). Due to the lack of statistics in this region, the agreement between data and simulation is poorer than in the other regions. Finally, figure 8.3 shows the postfit distribution of the m_T^W variable in the 0bjet region. This region is basically populated by the double boson processes (WZ+jets and ZZ production) and the NPL background components. By requiring no b jets in the event, the contribution from the signal is almost negligible with respect to the total.

8.1.3 Cross section and significance extraction

The simultaneous binned fit to the twelve templates described at the beginning of the chapter maximizes the binned likelihood function given in 7.6, from which the measured cross section $\sigma(\mathrm{t}\ell^+\ell^-\mathrm{q})$ is extracted according to its relation to the signal strength μ.

The maximum likelihood fit yields an observed tZq signal strength of:

$$\mu = 1.31^{+0.35}_{-0.33}(\mathrm{stat})^{+0.31}_{-0.25}(\mathrm{syst})$$

from which, using the reference NLO cross section ($\sigma(tZq \to Wb\ell^+\ell^-q) = 94.2^{+1.9}_{-1.8}(scale) \pm 2.5(PDF)$ fb), the measured cross section in found to be

$$\sigma(t\ell^+\ell^-q) \ = 123^{+33}_{-31}(\mathrm{stat})^{+29}_{-23}(\mathrm{syst})\ \mathrm{fb},$$

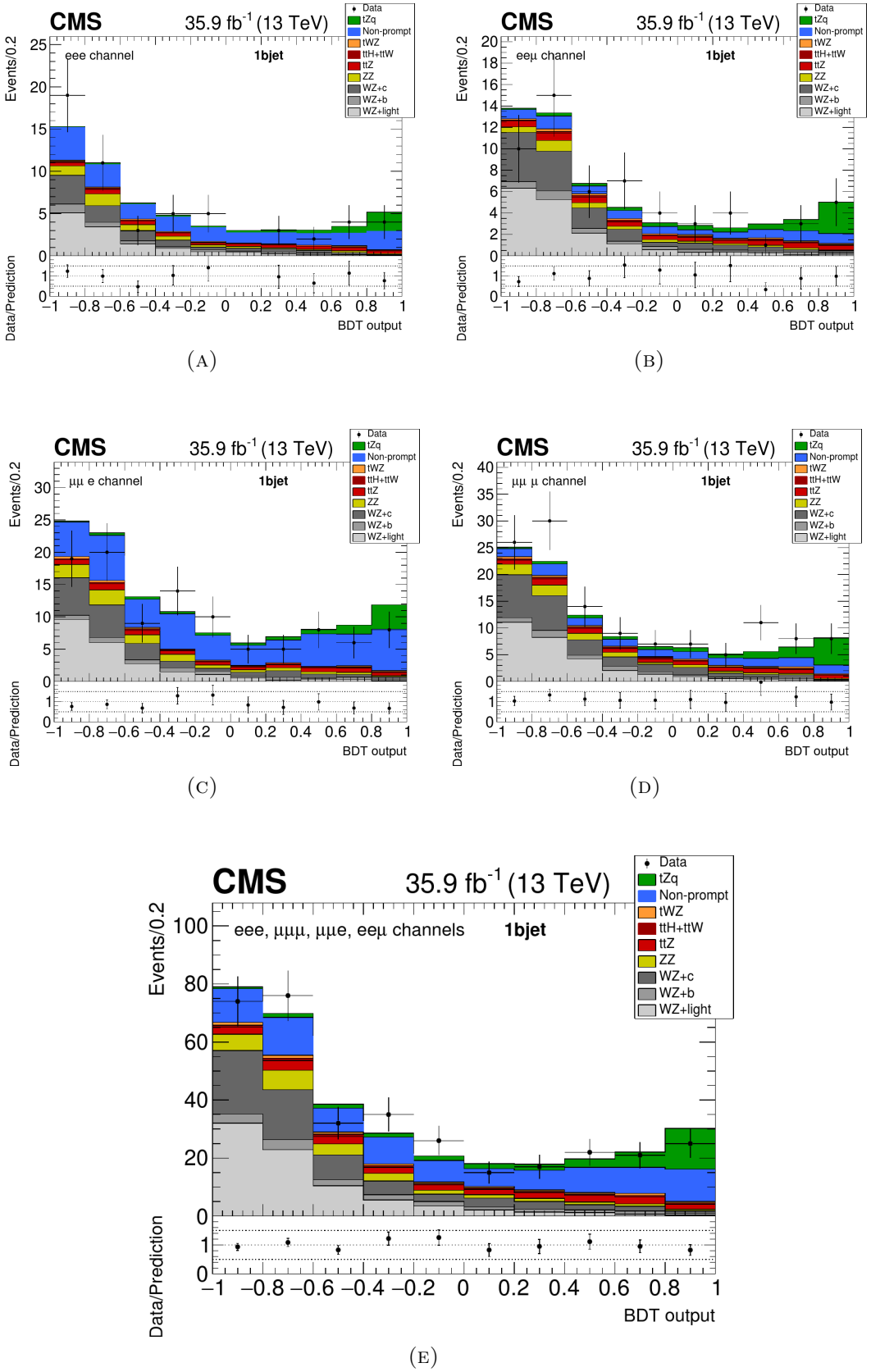
for $m_{\ell\ell} > 30$ GeV, where ℓ stands for electrons, muons, and τ leptons. [1] The precision of the measurement is limited by the statistical uncertainty. This effect will be further reviewed in section 8.2.

The observed and expected significances are extracted from a profile likelihood. The expected significance is obtained using simulation after the fit, and the observed one is derived from data. The obtained significances are

$$\text{Observed} = 3.7\,\text{SD} \qquad \text{Expected} = 3.1\,\text{SD}$$

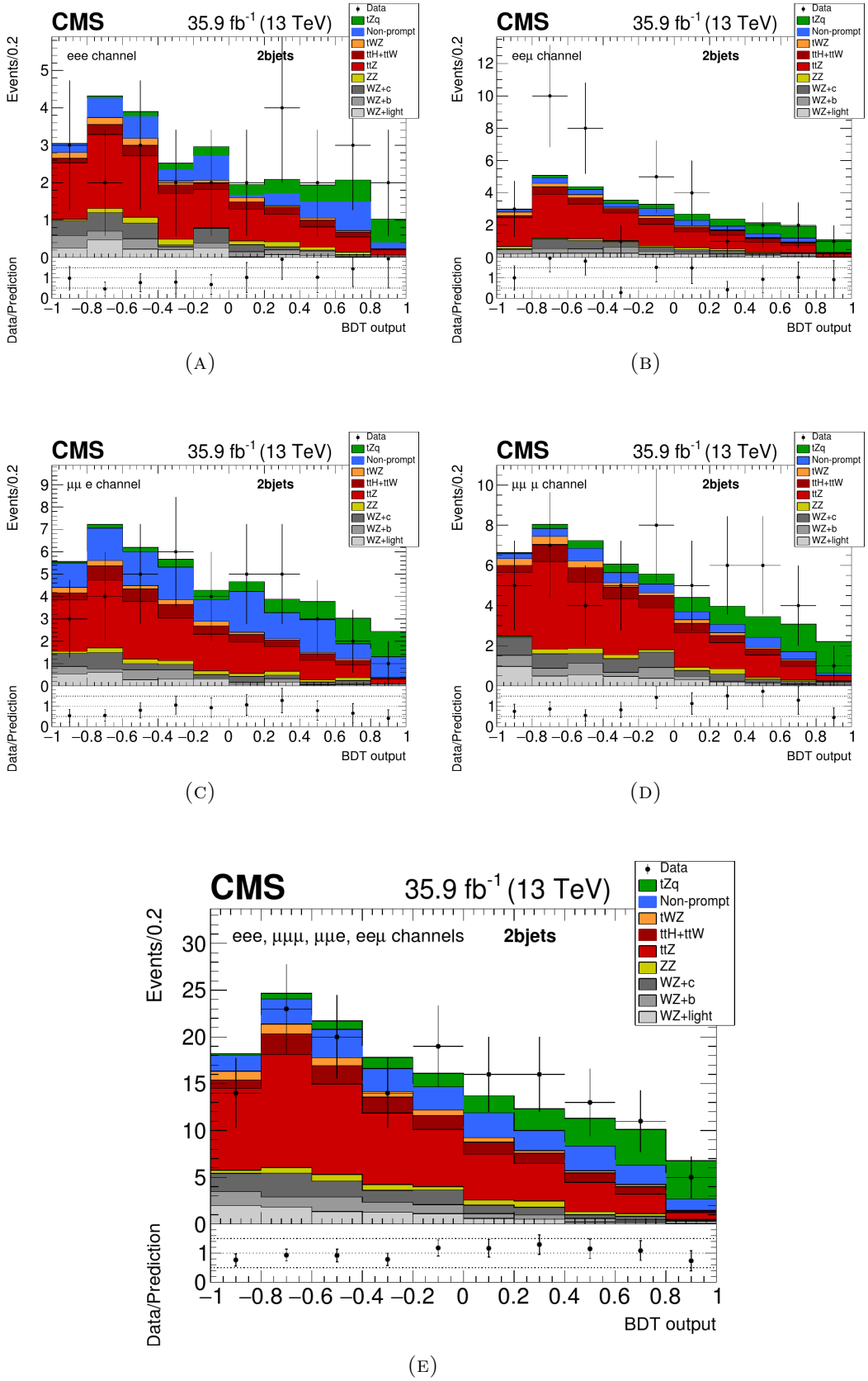
where SD stands for *standard deviations*.

[1] The fit is redone without including the systematic uncertainties, to evaluate the statistical uncertainty of the result. The quoted systematic uncertainty is then calculated as the difference in quadrature between the 68% CL intervals obtained in the nominal fit and in the fit without systematic uncertainties.



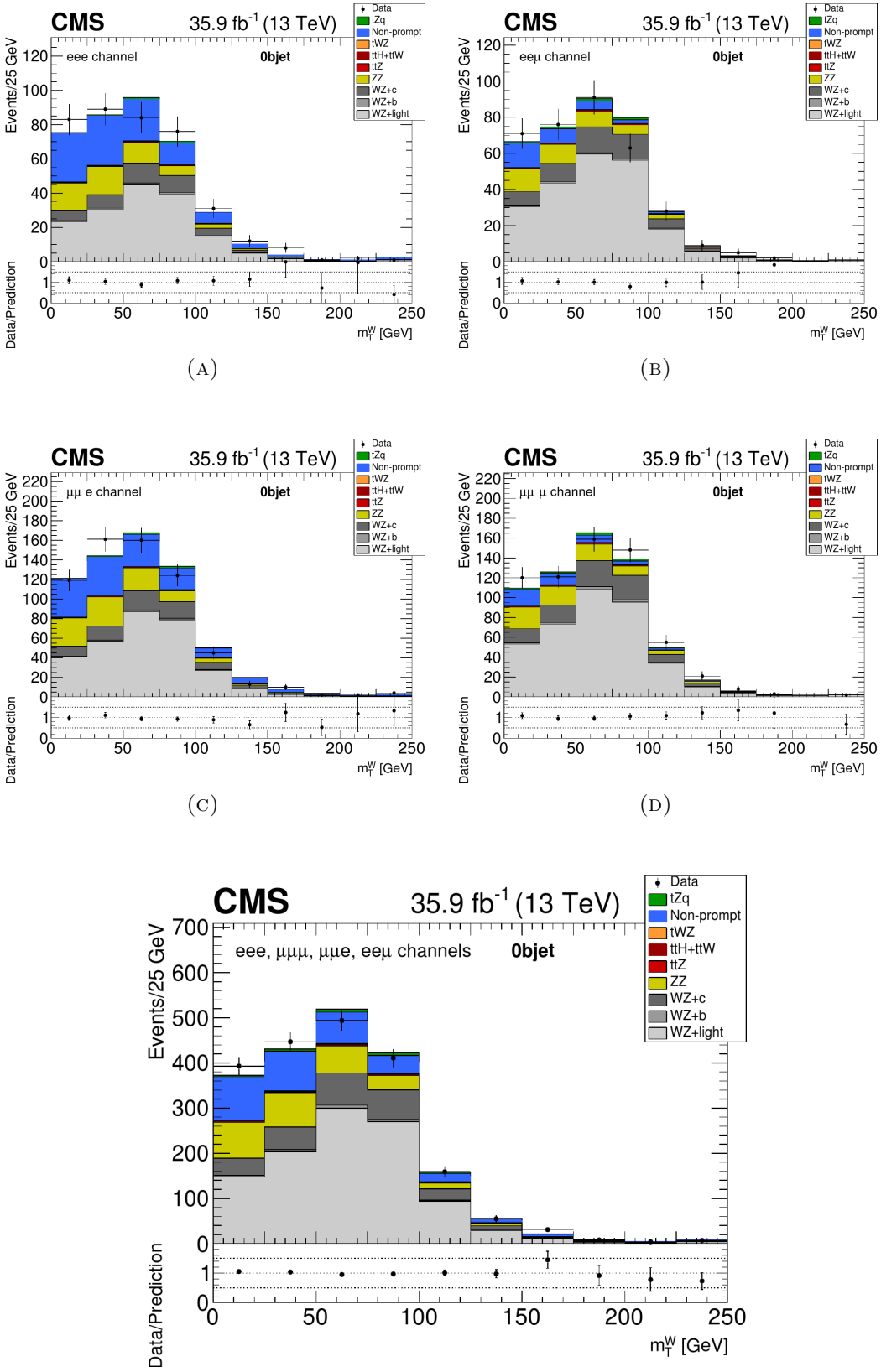
(A) (B) (C) (D) (E)

FIGURE 8.1: Postfit data-to-prediction comparison plots for the BDT discriminant in the 1bjet (signal) region, computed for the four channels individually and with all of them summed (last plot).



(A) (B) (C) (D) (E)

FIGURE 8.2: Postfit data-to-prediction comparison plots for the BDT discriminant in the 2bjet ($t\bar{t}Z$ enriched) region, computed for the four channels individually and with all of them summed.



(A) (B) (C) (D) (E)

FIGURE 8.3: Postfit data-to-prediction comparison plots for the m_T^W variable in the 0bjet region (where the contribution from WZ+jets and NPL backgrounds is most important), computed for the four channels individually and with all of them summed.

8.2 Postfit systematic uncertainties and NPL contribution

Systematic uncertainties in the analysis may affect only the overall normalization of the samples, or the shapes of the distributions used to categorize the events and to build the final discriminants, or both at the same time. Table 7.3 in the previous chapter shows how the different sources of systematic uncertainty affect the measurement. This effect is quantified in the analysis by a set of associated nuisance parameters; they are given initial values which are adjusted in the fit, the postfit values being those that maximize the likelihood function.

The study of the postfit values of these parameters allows to know how the tZq cross section measurement is affected by the different sources of systematic uncertainties. The only sources of systematic uncertainty which are treated in a slightly different manner in the fit are the shape and rate of the NPL backgrounds (remind that NP electrons and muons are treated separately). In contrast with the rest of systematic uncertainties, the parameters associated with the normalization and shape of these backgrounds are left free in the fit, and are fitted together with the signal.

Let us first briefly remind how the uncertainties related to the shape and normalization of the NPL background are treated in the analysis. Shape related uncertainties on the data-driven backgrounds involving NP leptons are estimated by varying the isolation criteria used to determine the NP sample. The shape variations of NP electrons and muons are associated two different nuisance parameters. Another two parameters account for the uncertainties on the normalization estimation of these two samples. They are associated a log-normal prior with a very large uncertainty and are left to float free in the fit, together with the signal.

The effect of the different sources of systematic uncertainties and the NPL background normalization on the measurement can be studied using *impact plots*. The *impact* of a nuisance parameter θ on the parameter of interest μ is defined as the shift ($\Delta\mu$) induced when θ is brought to its $\pm 1\sigma$ postfit values:

$$\Delta\mu^{\pm} = \hat{\mu}(\hat{\theta} \pm \Delta\theta) - \hat{\mu}(\hat{\theta}) \tag{8.1}$$

with all other parameters profiled as normal. This is effectively a measure of the correlation between the nuisance parameters and the parameter of interest, and is useful for determining which nuisance parameters have the largest effect on the uncertainty of the parameter of interest.

Figure 8.4 shows the impact of each systematic variation on the signal strength measurement in the visible phase space after performing the fit with the real data. The systematic uncertainties are listed in decreasing order of their impact on μ. The postfit value and uncertainty of the signal strength observed in figure 8.4 may differ from the numbers presented in section 8.1.3 as in the calculation shown in figure 8.4 all nuisance

parameters are treated as uncorrelated. The left part of the plot shows the *pull*

$$pull(\theta) = \frac{\hat{\theta} - \theta_0}{\Delta\theta} \tag{8.2}$$

of each nuisance parameter, that quantifies how far from its expected value θ_0 we had to pull the parameter while finding the maximum likelihood estimate.

The summed postfit uncertainties are shown for the combination of the 4 leptonic channels in Figure 8.5.

The uncertainty in the tZq measurement coming from the normalization of the NPL background is by far the most important (particularly the rate of the NP muon component, `FakeRateMu`), followed by the Q^2 scale variations at parton shower in the signal sample (`PSscale`), the uncertainty in the b tagging efficiency related to the light flavour contamination in heavy-flavour tagged jet (`LFcont`) from all simulated samples, PDF uncertainty (`pdf`) and the $t\bar{t}Z$ and ZZ normalization uncertainties (`ttZ_rate` and `ZZ_rate`, respectively).

It is reassuring to observe that the uncertainty associated to the normalization of the WZ+light sample (`WZl_rate`), 30% away from the prefit prediction as discussed in section 8.1.1, has a rather small impact on the measurement.

An alternative way to visualize this is to study the impact of each systematic on the likelihood function by scanning the negative log-likelihood ($-2\Delta \ln \mathcal{L}$) versus the signal strength (in the plots referred to as r). This is shown in figure 8.6 for some specific sources of uncertainty. To view the effect of a given systematic, the fit must be redone freezing the corresponding nuisance parameter. The larger the distance to the total likelihood (black solid line), the larger the impact of the parameter under study. The 1 and 2 standard deviation limits are indicated by the intersections of the horizontal lines at 1 and 4, respectively, with the log-likelihood scan curves.

Figure 8.6a shows the impact of the total statistical uncertainty, obtained performing the fit with all systematics frozen. The likelihood obtained this way is shown in a red dashed line. The statistical uncertainty is the limiting factor on the precision of the measurement, slightly larger than the associated to the sum of the different systematic uncertainties.

Figure 8.6b shows the likelihood obtained when freezing the rate of the NP muon background (blue dashed line), which the dominant systematic uncertainty (as can be seen in figure 8.4). In figure 8.6c, the likelihood presented is obtained freezing all associated theoretical uncertainties: the scale of the parton shower (`PSscale`), the factorization and renormalization scales at matrix element level (`Q2`), and the parton distribution function associated nuisance parameters (`pdf`). The effect of all theoretical uncertainties superposed is comparable to that of the rate of the NP muon background alone. In fact, the PS scale and PDF uncertainties have respectively the second and fifth positions in the ranking shown in figure 8.4.

Figure 8.6d shows the contribution to the likelihood when the nuisance parameters

associated to all experimental sources of uncertainty are frozen. These include the lepton and trigger efficiencies, pileup, luminosity, jet energy scale and resolution, along with the different uncertainties associated to the tagging of b jets.

Figures 8.6e and 8.6f show the contribution from the background estimation and the $t\bar{t}Z$ associated uncertainty. $t\bar{t}Z$ has the largest impact among all background sources (excluding the contribution from the NPL background), but still it is much lower than the statistical and the NP muon uncertainties.

Different contributions can be estimated in the same way, but most of them are quite small in comparison and the likelihood distributions appear very close to the total (*observed*) likelihood. The associated errors of the presented systematic uncertainties are displayed in the respective legends. In each plot, the error contribution from all systematic uncertainties excluding the corresponding frozen nuisance parameters are displayed as well. These are summed in quadrature to obtain the total error.

It is also interesting to study how the fit affects each of the systematics by comparing the prefit and postfit values of the nuisance parameters and their associated uncertainties. Figure 8.7 shows how each of the considered parameters change using as reference their input (prefit) values. It is straightforward to see here that the parameters that suffer larger differences after the fit are the ZZ rate and the fake muon rate (this can also be seen in table 8.4).

Figure 8.8 shows the ratio between the output (postfit) and input (prefit) uncertainty values of all the nuisance parameters associated to each of the systematics, $\sigma_\theta/\sigma_\theta(\text{prefit})$. These results can be compared with those from figure 8.7, where the increase in the uncertainty of the nuisance parameter associated to the jet energy resolution (JER) is also noticeable, which means this uncertainty was underestimated initially. In contrast, the fit constrains the uncertainties associated to the fake lepton rates, both for electrons and muons, whose contributions were left free to vary in the fit along with the signal strength.

Nuisance parameter	**b-only fit**		**s+b fit**		**rho**
	shift	σ	shift	sigma	
FakeRateEl	+0.10	0.06	+0.05	0.08	-0.15
FakeRateMu	+0.70	0.07	+0.54	0.12	-0.39
FakeShapeEl	+0.12	0.74	+0.18	0.61	+0.01
FakeShapeMu	-0.56	0.87	-0.50	0.99	+0.01
JER	-0.09	1.02	-0.09	1.50	-0.01
JES	-0.68	0.54	-0.72	0.51	-0.05
Q2	+0.44	0.84	+0.78	0.85	+0.09
WZc rate	-0.13	0.87	-0.04	0.87	+0.00
WZl rate	-1.03	0.46	-0.91	0.47	+0.01
ZZ rate	+0.83	0.93	+1.29	0.96	+0.11
pdf	+0.43	0.92	+0.53	0.90	+0.13
ttZ rate	+0.42	0.72	-0.06	0.75	-0.16

TABLE 8.4: Shift in the value and the postfit uncertainty of each nuisance parameter, both normalized to the input values. The last column shows the linear correlation between each parameter and the signal strength.

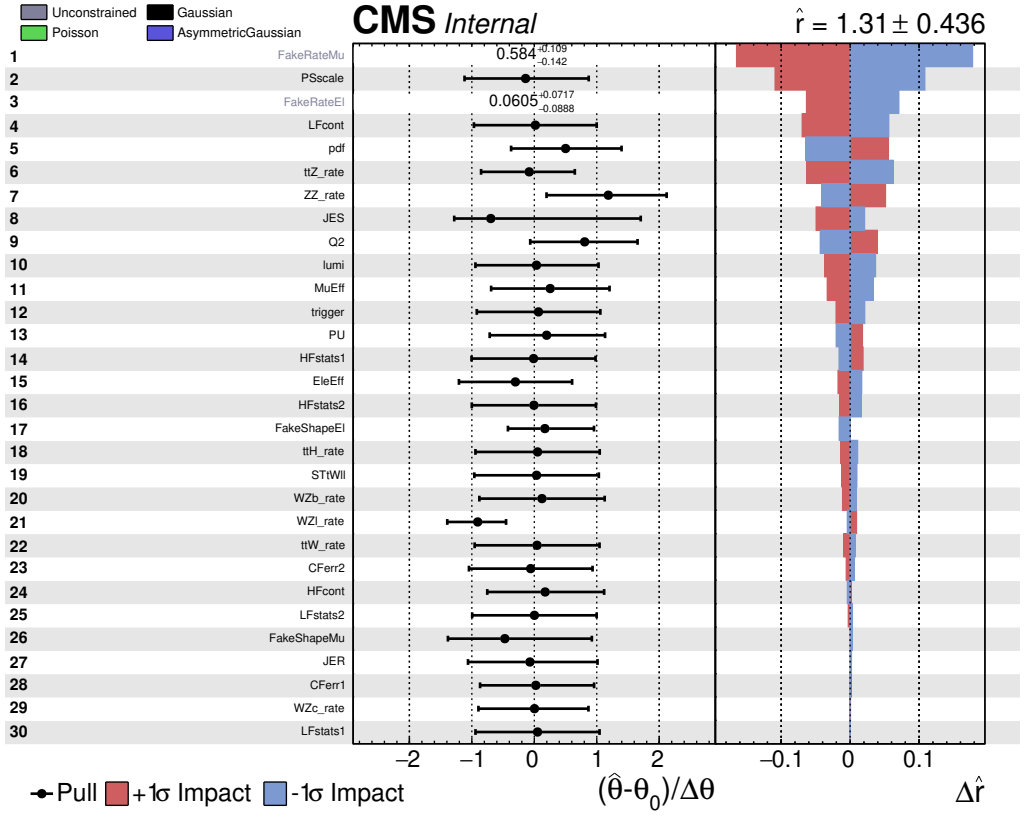


FIGURE 8.4: Postfit central values and uncertainties of the different nuisance parameters present in the fit. The black circles show the deviations of the postfit nuisance parameters $\hat{\theta}$ from their nominal values θ_0 expressed in terms of standard deviations with respect to their nominal uncertainties $\Delta\theta$. The associated error bars show the postfit uncertainties of the nuisance parameters, relative to their nominal uncertainties. The right part of the plot shows the impact of each nuisance parameter on the parameter of interest. The boxes in red (blue) represent the variations on μ when fixing the corresponding individual nuisance parameter θ to its postfit value $\hat{\theta}$, modified upwards (downwards) by its postfit uncertainty, and repeating the fit. The different systematic uncertainties are listed in decreasing order regarding their impact on μ. In grey are shown the uncertainties of the normalization parameters that are freely floating in the fit (NP electron and muon background rates). The quoted value of the postfit parameter of interest and its uncertainty are obtained treating all sources of systematic uncertainties are uncorrelated.

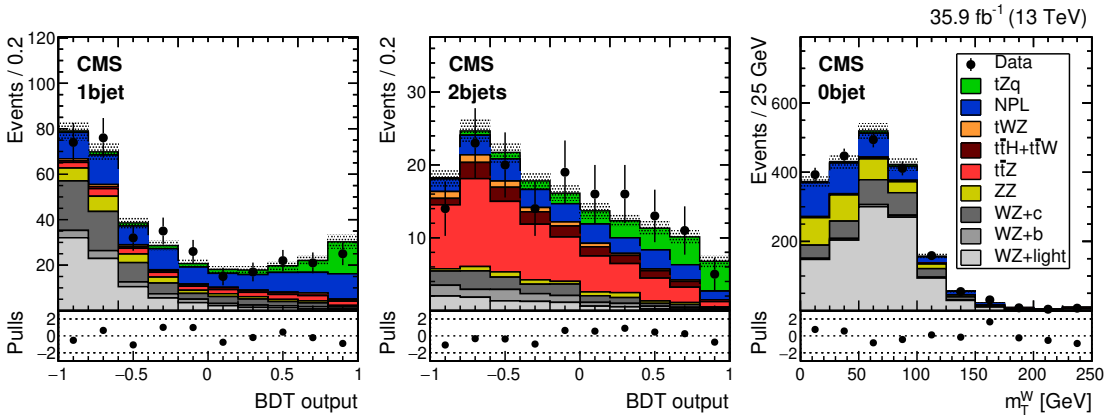
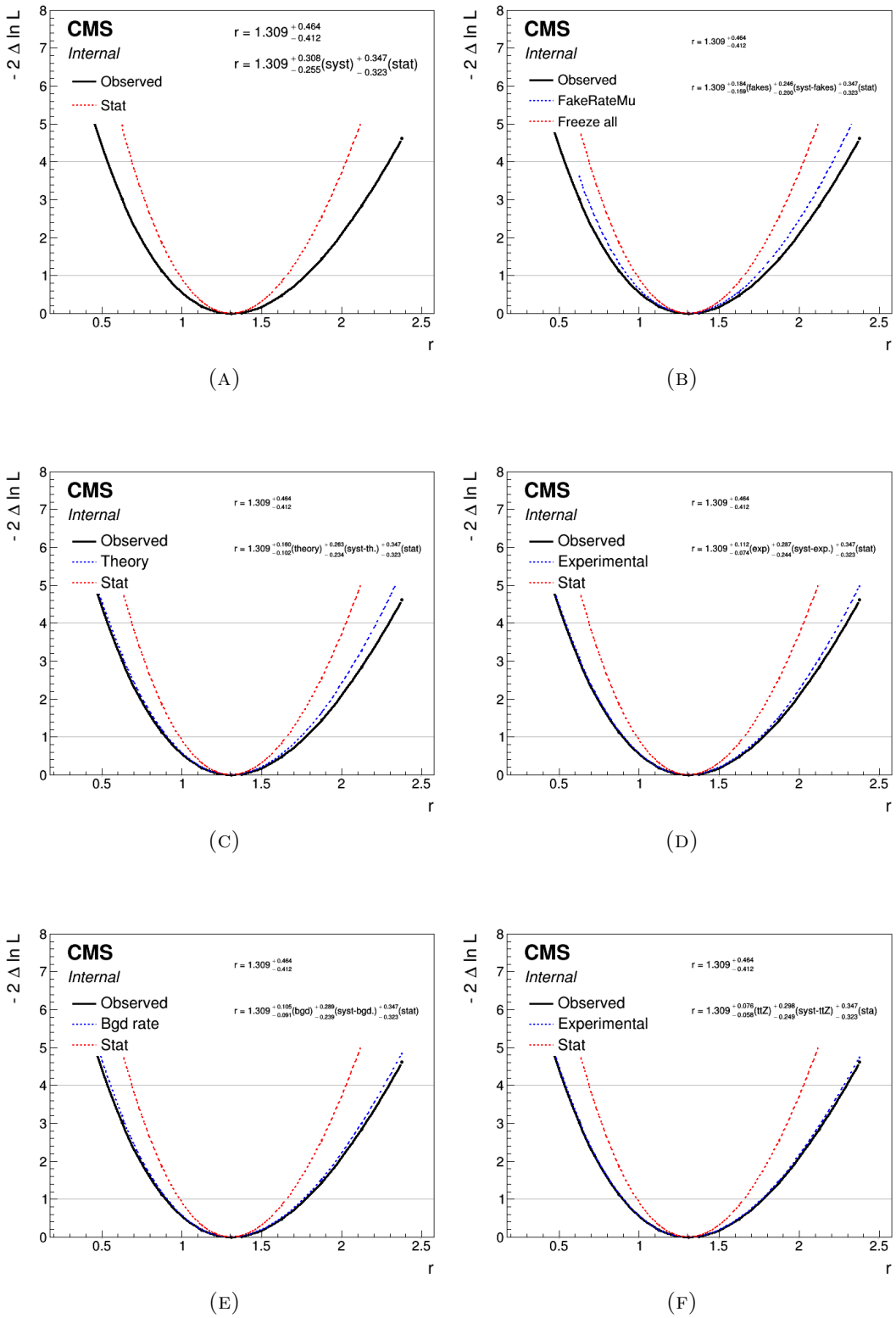


FIGURE 8.5: Postfit template distributions and uncertainties. Left: BDT discriminator in the tZq region;centre: BDT output in the $t\bar{t}Z$ control region; right: m_T^W in the WZ control region. The distributions include events from all final states. Underflows and overflows are shown in the first and last bins, respectively. The hatched bands include the total uncertainty on the background and signal contributions. The pulls in the distributions are shown in the bottom panels.



(A) (B) (C) (D) (E) (F)

FIGURE 8.6: Likelihood distribution as a function of the signal strength (the parameter of interest) in the visible phase space when freezing (a) all systematics, (b) the rate of the fake muon background, (c) all theoretical systematic sources, (d) all experimental systematics, (e) all background sources associated systematic uncertainties and (f) the $t\bar{t}Z$ background rate. The red line superimposed in all figures shows the effect of the statistical uncertainty, to make it easier to compare it to the rest of the shown uncertainties. The legends include the numerical contribution of each uncertainty to the total error.

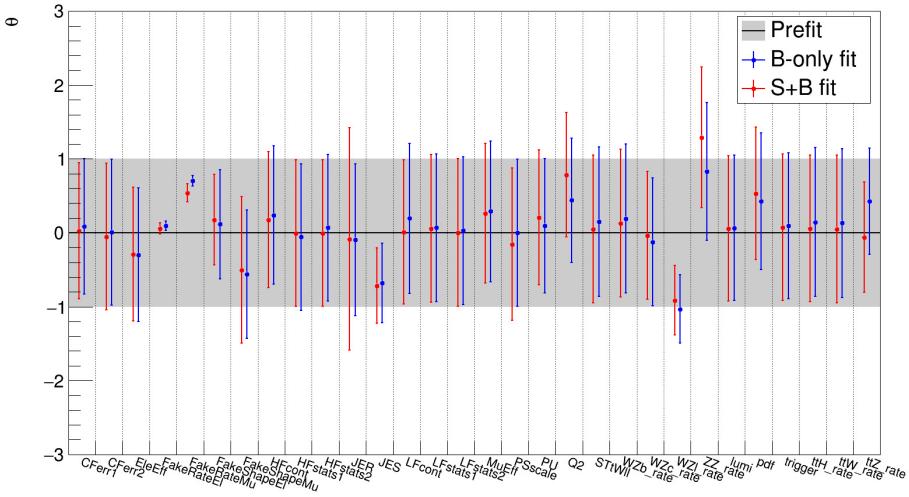


FIGURE 8.7: Changes in the nuisance parameter values and uncertainties, relative to their initial (prefit) values.

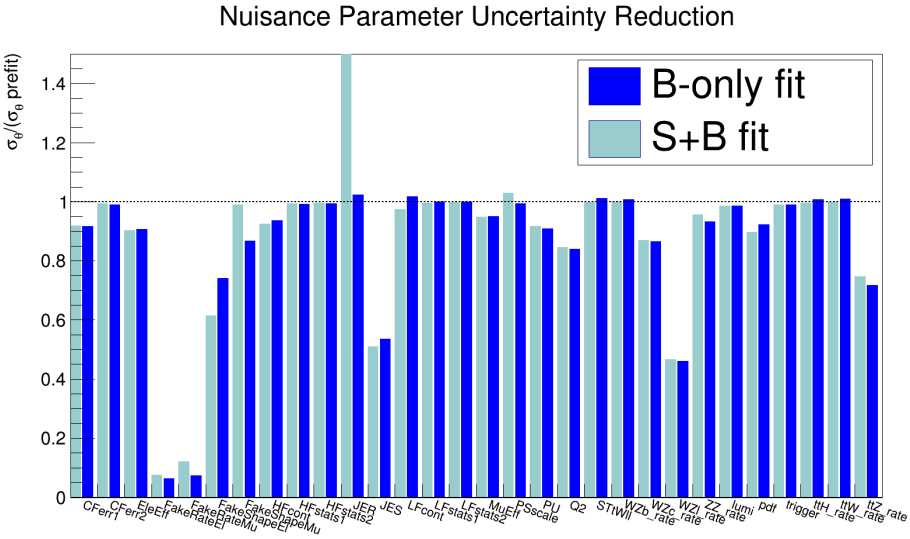


FIGURE 8.8: Ratio between the postfit and prefit uncertainties of each nuisance parameter, for the background-only hypothesis (dark blue) and the signal-background hypothesis (light blue).

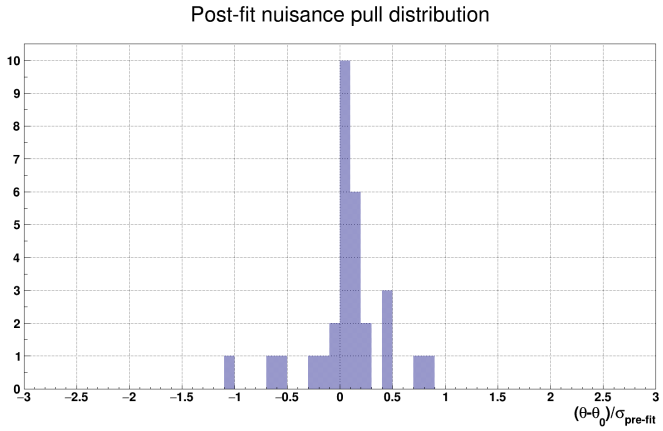


FIGURE 8.9: Nuisance parameter pull distribution. Here, θ_0 and σ_{prefit} stand for the initial values of each nuisance parameter and its uncertainty, respectively.

8.3 Stability of the results

The stability of the analysis was tested against some potential sources of bias.

Cross section by channel

The fit is repeated individually to get the cross section in the four different channels (eee, $e\mu\mu$, $ee\mu$ and $\mu\mu\mu$). This is done performing the fit on just three templates, corresponding to the two BDT distributions and the m_T^W, for each of the channels separately. The results are displayed in table 8.5, along with the expected and observed significances. The results from the four channels are compatible with each other, with the maximum significance observed in the $\mu\mu\mu$ channel. The impact of the NPL background is more strongly perceived in the $ee\mu$ channel, where the signal yields are artificially reduced to accommodate the NLP contribution.

Channel	Signal strength	Observed significance	Expected significance
eee	$1.38^{+1.07}_{-0.99}$	1.60	1.17
$ee\mu$	$0.67^{+0.78}_{-0.63}$	1.08	1.48
$e\mu\mu$	$0.13^{+0.95}_{-0.13}$	0.15	1.11
$\mu\mu\mu$	$1.27^{+0.75}_{-0.64}$	2.18	1.98

TABLE 8.5: Cross section measurement and significance results with one decay channel considered at a time. For the final measurement, a combined fit to all final states is performed.

Event selection: b tagging

The distribution of the CSVv2 b tagging discriminant is used as input variable to the two BDTs used in the analysis. It is one of the most discriminating variables (see table 7.2 or B.1 and B.2 from the appendices) and uncertainties related to the b tagging of jets (in particular the light flavour contamination in heavy flavour tagged jet samples) are amongst the nuisance parameters that have higher impact in the tZq cross section measurement (see figure 8.4). But in addition, the b tagging discriminant also enters the analysis as a selection criterion for the classification of a jet as a "b jet". The stability of results against the choice CSVv2 working point was verified. This was done switching from *loose* to the *medium* the working point of the CSVv2 algorithm (described in section 5.4, with the specific cut discriminant values given in Table 5.1). As the yield of the WZ sample decreased considerably in the signal region using the medium working point, the splitting of the sample in WZ+c, WZ+b and WZ+light was not used in this case. A slight decrease in the observed significance was found when switching to the medium working point, while the signal strength remained stable within about 1%. The decrease in the observed significance could be related to the fact that there are less events passing the signal selection (there are less b-tagged jets when switching to the medium working point of the tagger).

Stability of the $t\bar{t}Z$ background

The uncertainty in the rate of the $t\bar{t}Z$ sample is also amongst the ones that have a largest impact on the measurement. Thus, the stability of the results against the input $t\bar{t}Z$ cross section was also checked by measuring the tZq and $t\bar{t}Z$ cross sections simultaneously. The results obtained this way were very similar to the default analysis: the signal strength increased by less than a 1%, whereas the observed and expected significances decreased by about a 1%.

Stability of the NPL background

The stability of the result was also checked against the NPL rate, whose related uncertainties have the highest impacts in the measurement. For this, the NP muon and electron normalization factors were set to their input values (the way in which the input factors are derived is described in section 6.7) and allowed to vary in the fit as Gaussian constraints of 100% uncertainty. In this case, both the tZq signal strength and the expected and observed significances increase by about 10%, while the uncertainties on the signal strength increase by about 5%.

Select and count analysis

As a final check, the results were verified in a counting analysis, using the yields observed in control regions selected using similar criteria to those of the 1bjet, 2bjet, and 0bjet regions. The simulated backgrounds were normalized according to their SM predictions, while the normalization of the NPL contribution follows the procedure described in section 6.7 as the first step of the NPL normalization in the shape analysis. The results from the counting and shape analyses are in agreement.

In conclusion, the checks reveal that the measurement is robust and stable against background modelling and event selection.

Chapter 9

Summary and conclusions

This chapter summarizes the main results obtained in the measurement of the SM tZq associated production. Some prospects for future studies are also presented.

Since its discovery in 1995, the top quark remains the heaviest known elementary particle within the context of the SM. Due to its large mass, the top quark plays a special role not only in the understanding of the electroweak symmetry breaking mechanism but also in the search for new physics beyond the SM. Physics beyond the SM can manifest itself by altering the expected properties of the top quark and special effort has been put on studying this particle ever since its prediction in the early 1970s.

The high centre-of-mass proton-proton collision energy of 13TeV at the LHC, together with large integrated luminosities, allows the study of processes with very small cross sections that were not accessible at lower energies such as the production of a single top quark in association to a vector boson.

This dissertation presents an analysis originally devoted to the search of the SM associated production of a single top quark along with a Z boson and an additional quark (tZq), which conducted to one of the first measurements of the cross section for this process at an energy of $\sqrt{s} = 13$TeV at the LHC. The study was carried out by members of the CMS collaboration working at Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) in Madrid and the Institut Pluridisciplinaire Hubert Curien (IPHC) from Strasbourg, leading to a paper publication in Physics Letters B in december of 2017 [104].

The data used for the analysis was recorded by the CMS detector during 2016, corresponding to an integrated luminosity of 35.9fb^{-1}. The measurement is performed in the full leptonic configuration, where both the Z boson ($Z \to l^+l^-$) and the top quark ($t \to Wb \to l\nu_l b$) decay into leptons. Here l denotes either electrons or muons, resulting in four possible final-state leptonic combinations: eee, $\mu\mu\mu$, $ee\mu$ and $e\mu\mu$. A small contribution from τ leptons is expected from those decaying to muons or electrons, but they are not specifically targeted at selection. Data samples have been collected using a trigger logic specifically designed to keep this type of events, including three-lepton, dilepton and single-lepton triggers, achieving a trigger efficiency of nearly 100%. Each event is then reconstructed using the Particle Flow (PF) algorithm, which allows to identify all the

particles produced in the collision combining information from all CMS subdetectors.

Events are required to pass an initial trilepton baseline selection, demanding the presence of exactly three isolated high-p_T leptons, of which two need to be compatible with arising from a Z boson decay. In order to reduce the impact of the background determination uncertainties on the tZq measurement, the baseline selection is subdivided into three regions of interest according to their jet and b-tagged jet multiplicities: one of them enriched in tZq events and two control regions specifically designed to be populated by events from the main backgrounds ($t\bar{t}Z$ in one, and WZ+jets and NPL in the other, respectively). Most backgrounds are determined from MC simulation, except for the NPL background, which is derived using data-driven techniques.

The strength of the analysis relies on the usage of multivariate techniques to improve signal to background separation. Two boosted decision trees were optimized and trained in two of the statistically independent analysis regions, the first to separate tZq signal from the backgrounds estimated from simulation, and the second to distinguish $t\bar{t}Z$ background from the tZq signal. Up to 21 variables, mostly uncorrelated, were employed build the two trees. These variables describe the most diverse aspects of the topology and dynamics of the tZq events, and had strong discriminating power about most backgrounds.

Given the low statistics of the NPL samples, the boosted decision trees could not be trained against that background, which, as a consequence, could not be as well constrained as the other backgrounds, degrading the signal to background separation. In spite of that limitation, a clear tZq signal could be seen above the background.

The cross section is extracted from a binned maximum likelihood fit performed simultaneously on these three different regions, so that the normalization of the main backgrounds are better constrained. It is performed using the Higgs Combine Tool. Two boosted decision trees (BDT) are used to enhance the signal to background separation. The obtained cross section is then extrapolated to the whole leptonic phase space to account for the contribution from τ leptons too.

For the cross section calculation, different sources of systematic uncertainty have been taken into account. Those affecting the measurement include luminosity, pileup, trigger efficiency, jet energy scale and resolution, b-tagging efficiency, lepton isolation and identification efficiencies, normalization of the different background sources, and scale and PDF uncertainties for simulated signal and background processes. These are implemented in the final fit as a set of nuisance parameters. Two free parameters in the fit are used to estimate the contribution from the NPL backgrounds, which are the sources of the dominant systematic uncertainties on the analysis, as well the limiting factor, besides the statistical uncertainty, on the tZq cross section measurement significance.

Evidence for tZq production is found with an observed (expected) significance of 3.7 (3.1) standard deviations. The cross section is measured to be

$$\sigma(t\ell^+\ell^- q) \ = 123^{+33}_{-31}(\text{stat})^{+29}_{-23}(\text{syst}) \text{ fb},$$

for $m_{\ell\ell} > 30$ GeV. This result is an extrapolation to include the contribution from τ

leptons (thus, in the previous result $\ell = \{e, \mu, \tau\}$). This value is compatible with the next-to-leading-order SM prediction of 94.2 ± 3.1 fb. The corresponding observed (expected) significance against the background-only hypothesis is 3.7 (3.1) standard deviations, with an observed statistical p-value of 0.0001. The 68% CL interval of the expected significance is 1.4–5.9.

This measurement consist on yet another stringent test of the standard model, with no clear evidence of new physics. Nevertheless, the measurement is still relevant on its pursue, given that the inclusion of tZq cross sections in combined Effective Field Theory fits provides new constrains to the available phase space where new physics can manifest itself, increasing the chances of finding it.

Chapter 10

Resumen y conclusiones

Este capítulo resume los principales resultados obtenidos en el análisis de la producción asociada tZq en el Modelo Estándar.

Casi tres décadas después de su descubrimiento en 1995, el quark top sigue siendo la partícula elemental más pesada dentro del marco del Modelo Estándar. Debido al elevado valor de su masa, el quark top juega un papel fundamental tanto en la comprensión detallada del mecanismo de ruptura espontánea de la simetría electrodébil como en la búsqueda de fenómenos de física más allá del Modelo Estándar. Estos fenómenos pueden manifestarse como modificaciones de las propiedades del quark top tal y como se esperan en el Modelo Estándar, por lo que, tras la predicción de su existencia en los años 70, se ha desarrollado un importante esfuerzo experimental para su observación y caracterización.

El aumento de la energía en el centro de masas de las colisiones protón-protón a 13 TeV, así como la gran luminosidad integrada proporcionada en los últimos años por el LHC, han permitido estudiar procesos con secciones eficaces muy pequeñas e inaccesibles a energías inferiores, como es la producción asociada de un quark top y un bosón Z.

Esta tesis presenta una de las primeras medidas de la sección eficaz de producción de un quark top en asociación con un bosón Z y un quark adicional (tZq) en colisiones pp a una energía centro de masas de 13 TeV en el LHC. El análisis, inicialmente focalizado a obtener la confirmación experimental de este proceso, hizo también la primera medida en CMS de su sección eficaz de producción. El estudio fue llevado a cabo por miembros de la colaboración CMS del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) en Madrid y del Intituto Pluridisciplinar Hubert Curien (IPHC) de Estrasburgo (Francia). Los resultados obtenidos se publicaron en la revista científica Physics Letters B en diciembre de 2017 [104].

Los datos analizados en este trabajo fueron recogidos por el experimento CMS durante el año 2016. La muestra corresponde a una luminosidad integrada de 35.9fb^{-1}. La medida se llevó a cabo en el canal leptónico, donde tanto el bosón Z ($Z \to l^+l^-$) como el bosón W procedente de la desintegración del quark top ($t \to Wb \to l\nu b$) se desintegran en estado finales leptónicos. Aquí, l se refiere tanto a electrones como a muones, dando lugar a cuatro posibles combinaciones (o canales) de leptones en el estado final: eee, $\mu\mu\mu$, $ee\mu$ y $e\mu\mu$. La muestra seleccionada incluye una pequeña contribución procedente de la

desintegración de leptones τ en muones o electrones, aunque no eran objetivo prioritario para el diseño de la estrategia de análisis. Los datos se tomaron utilizando una selección *online* con un combinación de *triggers* diseñada específicamente para este tipo de sucesos, e incluye *triggers* que seleccionan eventos con uno, dos o tres leptones, obteniendo así una eficiencia cercana al 100%. Los eventos seleccionados se reconstruyen usando un algoritmo específico de CMS llamado *Particle Flow* (PF) que utiliza información de todos los subdetectores de CMS para identificar las partículas generadas en la colisión.

En una primera preselección de eventos, se toman aquellos que tienen exactamente tres leptones aislados con un alto p_T, dos de ellos compatibles con proceder de la desintegración de un bosón Z (dos leptones del mismo sabor y cargas opuestas). Para reducir el impacto en la medida de las incertidumbres en la determinación de los procesos de fondo procedentes de fuentes diversas, la selección establece tres regiones de interés en función del número de jets y b-jets presentes en el evento. La primera, la región de señal, contiene mayoritariamente eventos tZq y las otras dos regiones (de control) fueron diseñadas de manera específica para controlar las principales fuentes de contaminación ($t\bar{t}Z$ por un lado, y las contribuciones de WZ+jets y *non-prompt leptons, NPL*, por otro). La contribución de la mayoría de las fuentes de contaminación se estima a partir de muestras de sucesos simuladas por técnicas de MonteCarlo, excepto en el caso de la contribución de los NPL, que se deriva con métodos que emplean otras muestras de datos.

Uno de los pilares del análisis es la utilización de técnicas de análisis multivariable para optimizar la separación entre señal y fondo, lo que incluye la optimización y el entrenamiento de dos *boosted decision trees* (árboles de decisión o BDT) en regiones de análisis estadísticamente independientes. La primera BDT se diseñó para separar la señal de tZq de los fondos estimados a partir de muestras simuladas, y la segunda BDT, para separar eventos tZq de $t\bar{t}Z$. Se usaron hasta un total de 21 variables, esencialmente no correlacionadas entre sí, para construir las dos BDT. Las variables elegidas son representativas de las características topológicas y dinámicas de los sucesos tZq y tienen un gran poder de discriminación entre la señal y la mayoría de los fondos.

No fue posible entrenar una BDT específica para conseguir una separación óptima de la señal frente al fondo de *non-prompt leptons* debido a la estadística reducida de las muestras de control recogidas para este proceso. Pese a ello, se consiguió la observación de una señal clara de sucesos tZq sobre el fondo procedente de otros procesos.

La medida de la sección eficaz de producción se obtiene a partir de un ajuste simultáneo sobre las tres regiones previamente mencionadas, de manera que la normalización de las principales fuentes de fondo quede constreñida en el ajuste. Para ello se utilizaron las herramientas estadísticas desarrolladas en CMS (*Higgs Combine Tool*). El resultado se extrapola a todo el espacio de fases leptónico incluyendo también la contribución a la sección eficaz de los procesos con leptones τ.

Se ha estimado cual es la contribución de las diferentes fuentes sistemáticas de incertidumbre en la medida de la sección eficaz. Se han incluido los efectos de la incertidumbre en la determinación de la luminosidad, el *pileup*, la eficiencia del *trigger*, la resolución y escala de la energía de los jets, la eficiencia de identificación de b-jets, las eficiencias de aislamiento e identificación de los leptones, la normalización de los procesos de fondos, así como las incertidumbres en las funciones de distribución de partones (PDF) y de escala

en las simulaciones de los procesos de señal y fondo. Todos estos efectos se implementan en forma de parámetros (*nuisance parameters*) en el ajuste. Los parámetros asociados a la normalización del fondo NPL se dejan como parámetros libres en el ajuste. La incertidumbre en esta contribución, así como la incertidumbre estadística son los elementos principales que condicionan la significancia de la señal tZq observada.

Se ha obtenido evidencia experimental de la producción tZq con un valor de la significancia observada (esperada) de 3.7 (3.1) desviaciones estándar y un p-value de 0.0001.

Se ha medido la sección eficaz de producción:

$$\sigma(t\ell^+\ell^-q) \ = 123^{+33}_{-31}(\text{stat})^{+29}_{-23}(\text{syst}) \text{ fb},$$

con $m_{\ell\ell} > 30$ GeV. Este resultado es la extrapolación que incluye la contribución de los leptones τ (es decir, en el anterior resultado $\ell = \{e, \mu, \tau\}$). Este valor es compatible con la predicción del Modelo Estándar 94.2 ± 3.1 fb, calculada a *next-to-leading-orden* en teoría de perturbaciones.

Esta medida supone un test exigente de la validez del Modelo Estándar y los resultados obtenidos no ofrecen indicios claros de nueva física. A pesar de ello, esta medida de la sección eficaz de producción de tZq es relevante para este propósito ya que su inclusión en los ajustes combinados que se desarrollan en el contexto de Teorías de campo efectivas (EFT) proporciona nuevas cotas al espacio de fases en el que fenómenos de física más allá del Modelo Estándar podrían manifestarse, aumentando por tanto las posibilidades de encontrarla.

Appendix A

Decision Trees

A.1 Decision Trees

One of the most popular MVA algorithms is the so-called *Decision Tree* (DT) technique which, in contrast with other MVA methods, has a straightforward interpretation. A decision tree is a machine-learning technique which combines several classifiers (variables) to create a more powerful multivariate discriminant. It is a way of organizing and choosing the cuts applied to a candidate depending on whether it passed or failed the previous cuts.

A decision tree is a structure that employs sequential cuts (the way in which they are determined and optimized will be explained next) organized into *nodes*. Nodes are decision points within the tree structure at which an optimized cut on a certain variable entering the decision tree is applied, and the data is split into events either passing or failing the cut. This determines which node the candidate will encounter next. The primary node receiving all signal and background events is often referred to as *root node* which branches off to two secondary nodes, the rest of the sequence following a structure as that shown in figure A.1.

Each cut path eventually stops at some terminal node or *leaf* with a classifier value which will be assigned to the candidate. Any event that fails a certain cut will not be thrown away immediately as background, but will rather continue to be analyzed starting with the assigned classifier value. In the end, all events are given a decision tree score between 0 and 1.

Decision trees must be trained with dedicated samples containing simulated events of the signal and background processes, each with a certain weight ω_i, to build a tree structure of cuts node by node. The function used to quantify the separation between signal and background at any node is called *Gini index*, which is defined as

$$\text{Gini} = 2p(1-p) \tag{A.1}$$

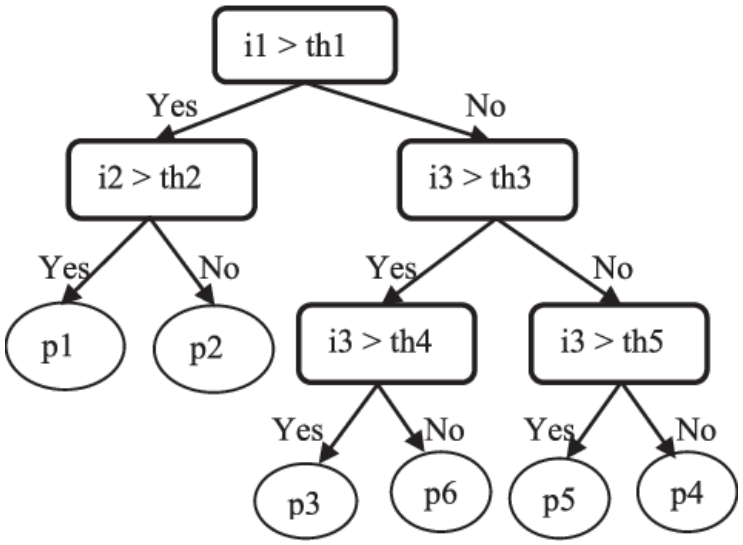


FIGURE A.1: Basic scheme of a decision tree.

where p is the *purity*, defined as

$$p = \frac{S}{S+B} = \frac{\sum\limits_{S} \omega_s}{\sum\limits_{S} \omega_s + \sum\limits_{B} \omega_b} \tag{A.2}$$

S(B) is the weighted total number of signal (background) events which landed on the node during the training. As a consequence, the Gini index for a node is 0 (minimal) when the purity is either 1 or 0 (corresponding to pure signal or pure background, respectively) and maximized when the purity is 0.5 (corresponding to a maximally mixed sample).

Beginning from all events at the root node, the variable and the corresponding cut which maximize the separation are calculated and executed. The maximum separation is defined as the maximum change in the Gini index between the mother node and the two daughter nodes, as:

$$\Delta\text{Gini} = \text{Gini}_{\text{mother}} - f_{\text{daughter1}} \cdot \text{Gini}_{\text{daughter1}} - f_{\text{daughter2}} \cdot \text{Gini}_{\text{daughter2}} \tag{A.3}$$

where f_{daughter} is the weighted fraction of events in each daughter node. The cut corresponding with the highest ΔGini is selected. This same process is performed recursively to construct the tree. This process of node splitting is performed recursively until a given terminal criterion is satisfied, when there is either no possible reduction in impurity or when the available number of events is too small to proceed further in the classification. Nodes that are not split are the *leafs* previously mentioned. Samples with high statistics shall be used to achieve an optimal performance of the decision tree. Test events then go through all the tree conditions starting from the root node until they land on a terminal node. The particular choice of the type and number of variables used depend on the analysis. Further details on the variables used in this analysis will be given in the next sections.

Despite all advantages they offer, decision trees have certain limitations, such as their

instability with respect to the training sample (slight differences in training samples may result in significantly different trees). These limitations are overcome using procedures such as *boosting*.

A.2 Boosted decision trees

Boosting is a way of enhancing the classification performance with respect to single trees by sequentially applying a multivariate algorithm (a decision tree in this case) to reweighted (*boosted*) versions of the training data. It also helps smoothing distributions in cases where a specific training sample has limited statistics. Boosting uses the training results of the first tree to increase the weights of candidates that were misclassified. New trees are trained using these first weights. Boosting reweights candidates that the previous tree classified incorrectly in order to increase their importance during the next training. Terminal leaves are labelled either background or signal leaves according to a set of purity threshold (often 0.5). Misclassification occurs when a candidate of one type (signal or background) ends on a leaf of opposite classification.

The goal of boosting is to combine simple decision algorithms (the decision trees) characterized by the fact that their individual separation power is small, into a new, more stable classifier, with a smaller error rate and better performance. Due to the large number of outputs from the single trees being averaged together, the output value of the boosted classifier appears as a continuous variable. It is then possible to choose the working point of the selector by imposing a cut on this BDT output variable. A working point is a choice of the cut corresponding to a specific compromise between signal efficiency and signal purity.

There are different boosting algorithms available, such as *AdaBoost* (adaptive boosting), *gradient boosting* or *XGBoost*. Gradient boosting is the algorithm used for tree construction in this analysis.

A.2.1 Gradient boosting

The different boosting algorithms define how the classifier ensemble is built. In each step, a new sub-model (i.e. a single decision tree) is added, that tries to compensate the errors made by the previous set of sub-models.

The first classifier gives an output, from which *residuals* are computed. These residuals are nothing more than the difference between the target values and the ones predicted by the current classifier. Then, a second classifier is built that tries to model these residuals, and is added to the ensemble. This process is repeated iteratively, in a process in which the residuals are minimized (or randomly distributed around zero, figure A.2 illustrates how this is performed). The algorithm just described can be seen as a form of *gradient descent*. In general, the algorithm tries to optimize the boosted ensemble with respect to a *loss function*. The loss function is a method to quantify how well the algorithm

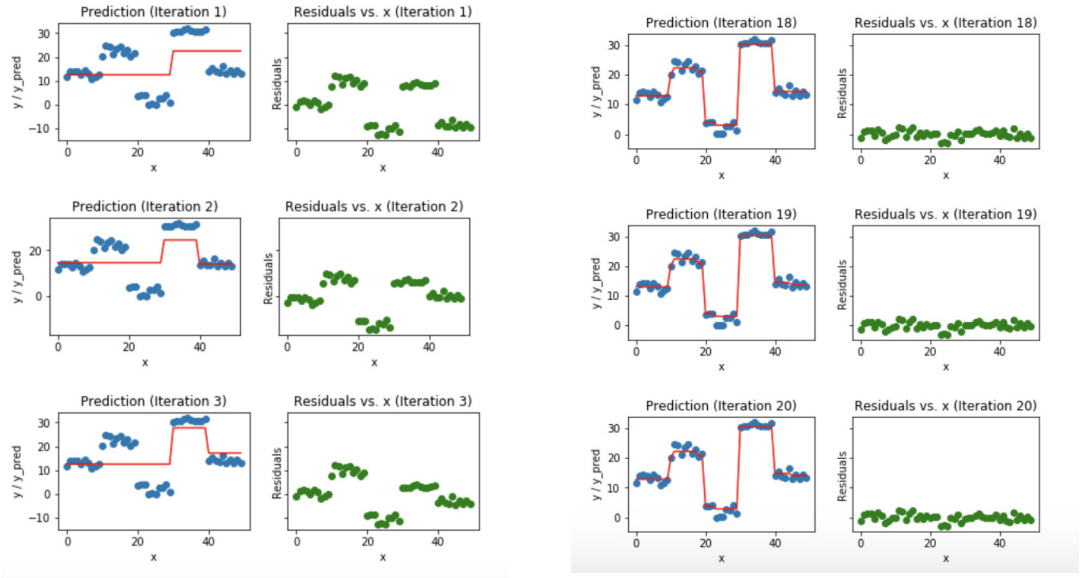


FIGURE A.2: Visualization of gradient boosting predictions. Images are taken from [105].

models the data. If predictions do not fit the data, the loss function will output a high value. Otherwise, if predictions are good, the output value will be low. Gradient descent represents a process in which this loss function is minimized applying the following update rule repeatedly:

$$x = x - \eta \nabla_{\text{Loss}}(x)$$

where $\nabla_{\text{Loss}}(x)$ is the gradient of the loss function the ensemble aims to optimize, and η is the step length (the learning rate, which can be controlled to reduce the risk of overtraining (see section A.3), which means that the model becomes too specific and will not perform well in more generic cases). The negative gradient of the squared error loss

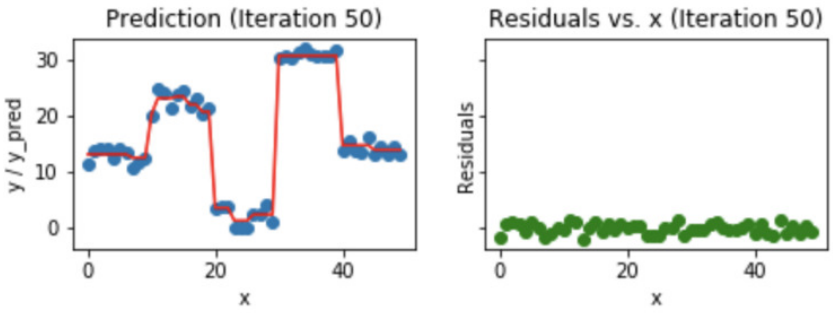


FIGURE A.3: Visualization of gradient boosting predictions: for 50 iterations, the model is overfitting the predictions. Images are taken from [105].

function is the same as the residual (multiplied by a factor 2):

$$\text{Loss}(y_i, \hat{y}) = (y_i - \hat{y})^2 \qquad -\nabla_{\text{Loss}}(\hat{y}) = 2 \cdot (y_i - \hat{y})$$

In such a way, the boosting algorithm can be seen as a gradient descent that minimizes the squared error loss function. At each step, a new sub-model is added that tries to mimic the negative gradient of this loss (the learning rate η in equation A.2.1 would be 0.5 in this case). The boosting algorithm iteratively adds new sub-models to the ensemble in order to minimize the loss function. This can be seen pictorially in figure A.4.

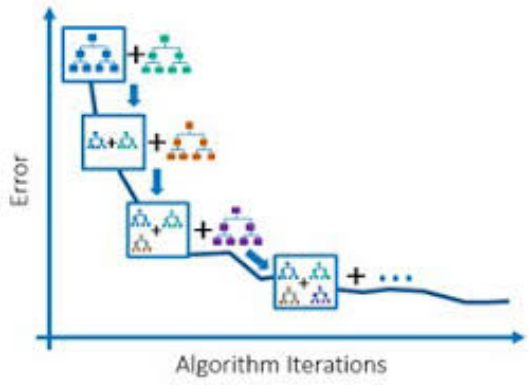


FIGURE A.4: Gradient boosting algorithm for minimization.

This recipe can be generalized for different loss functions. At each step, the gradient of the loss function is calculated. In the general case, the gradient no longer represents the residuals, but they are rather referred to as *pseudo-residuals*. The next model to be added to the ensemble is then trained to imitate this gradient. This is reproduced sequentially until the final ensemble is formed.

There is not a single loss function that works for all kind of data. It depends on a number of factors such as the presence of outliers, the algorithm chosen, time efficiency of gradient descent, ease of finding the derivatives or the confidence of the predictions. If the outliers represent anomalies that are important and should be targeted, then squared error loss represents a suitable loss function.

The loss function implemented by default for gradient boosting in the MVA tool used for the analysis is the so-called (convex and differentiable) *Huber loss*, which is less sensitive to outliers in data than the squared error loss and is defined as

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & \text{for } |y - f(x)| \leq \delta, \\ \delta\,|y - f(x)| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases} \tag{A.4}$$

The parameter $\delta > 0$ describes where the transition from quadratic to linear takes place. This can be seen in figure A.5. Squared error loss can be seen as an specific case of Huber loss with constant $\delta = \infty$. In general, the Huber loss is equivalent to the squared loss for points which are well-fit by the model (few standard deviations), but reduces the loss contribution of outliers (high standard deviation values). For example, a point 10 standard deviations from the fit in figure A.5 has a squared loss of 50, but a c=1 Huber loss of just over 10. The choice of δ is critical because it determines what the analysts will to consider as an outlier.

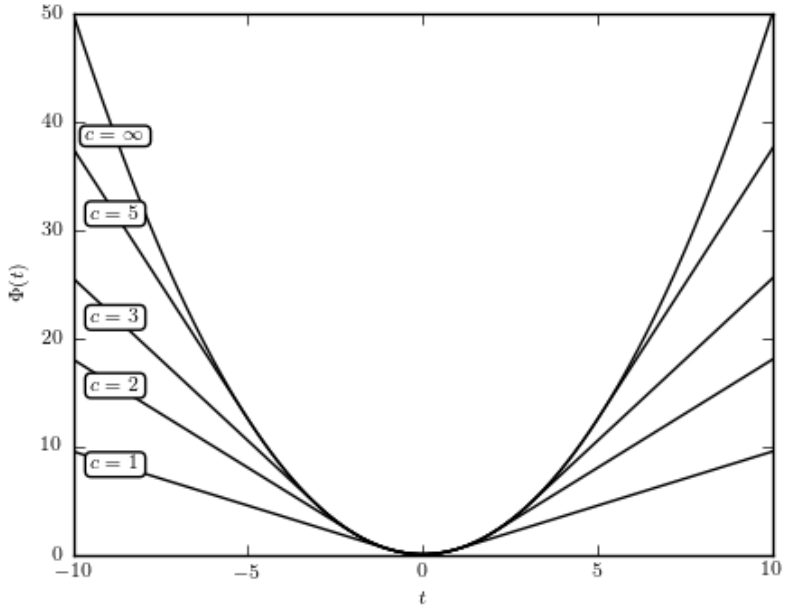


FIGURE A.5: Huber loss distributions for different values of constant δ (here denoted as c) as a function of the standard deviations.

A.3 Overtraining

One issue regarding multivariate classifiers is that they might be *overtrained*, meaning they are too adapted to a specific training sample. It occurs for instance when decisions are made on statistical fluctuations of the training sample, meaning that the learning did not pick up actual signal or background properties, but rather statistical fluctuations. The performance of the classifier is then better on the training sample (the algorithm learns more), while the actual performance on data becomes worse. Overtraining can lead to inefficiencies in the separation of the samples, and the magnitude of the effect can be seen by comparing the BDT output of the training sample with that of the test sample, using a Kolmogorov-Smirnov test (see figure 7.5). The Kolmogorov-Smirnov test quantifies the agreement between two given samples, being closer to 1 (0) when they are more (less) prone to have the same parent distribution.

To better illustrate how overtraining works, an example is given in figure A.6. In this picture, the red line shows the signal to background separation in terms of two generic values x_1 and x_2 after training. The diagram on the right picture shows clear dependence on the training sample, and it will be less representative when used over different samples: we say it is overtrained. It is clear that the boundaries on the picture on the left will have a better performance on a statistically independent samples than the one on the right.

Figure A.7 shows how the error rate (efficiency) on the training sample decreases (increases) with learning, whereas it goes up (down) when the same BDT is used on any test sample. The algorithm has to be set (or tuned) so that the learning stops before this turning point is reached.

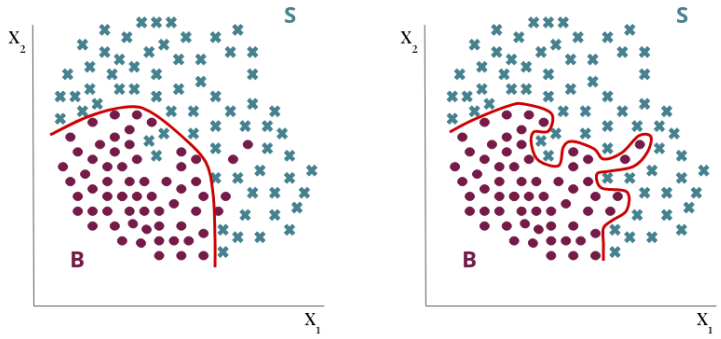


FIGURE A.6: Depiction of how overtraining works. The boundary on the right diagram is overly suited to the training sample. Such a specific profile offers an optimal performance on the this sample, but will most certainly not offer an effective signal to background separation in a more general case, in contrast to the boundary on the left diagram. We say that there is *overtraining* in the right diagram.

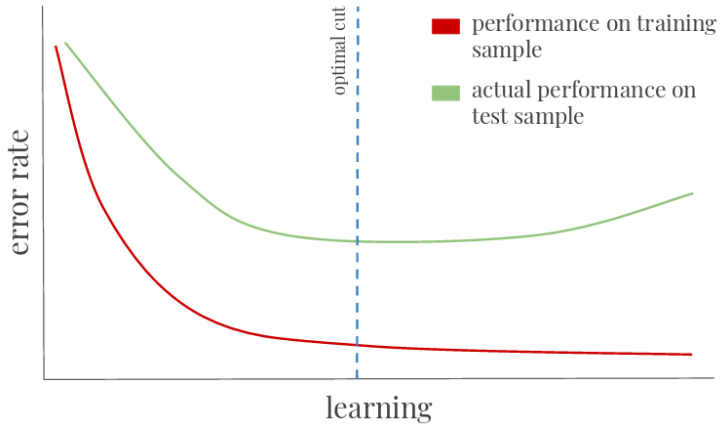


FIGURE A.7: Predicted error rate of the MVA with respect to the learning of the method. In this graph one can identify the point where the model (consisting of a certain number of successive trees in the case of a BDT) begins to overfit the data. The prediction error for the training data (red line) steadily decreases as more and more additive terms (trees) are added to the model. There is a point at which the performance for independently sampled testing data (green line) actually begins to deteriorate, indicating the point where the model begins to overfit the data. The dashed blue line shows the optimal point at which to truncate the learning process in order to avoid overtraining.

Overtraining is a generic property of multivariate classifiers, and it might be overcome with different techniques. Increasing the available statistics during the training allows to mitigate a bit this effect, but sometimes it does not completely solve the problem. For instance, in figure A.7 one can see that the error rate on test sample reaches a minimum and then rises. A possible way to address this problem would be to stop boosting at the minimum. In the current analysis, this is done by adjusting the number of trees in the BDT in order to avoid overtraining.

Keeping the system free from over-training is not easy, as there are different aspects that may affect it. General and empirical rules indicate that keeping the systems simple (for instance, by growing small trees), and the training process just long enough, forces them to understand the data rather that to learn it, thus avoiding the overtraining problem. Several user defined parameters affect the training and boosting of the decision trees (and thus the performance of the analysis), and they are tuned to increase the accuracy of the method and prevent from overfitting. Some specific parameters in decision tree boosting are, for instance, the maximum depth of the trees (total number of nodes), the number of trees in the BDT, or the learning rate of the algorithm (through a parameter known as *shrinkage*). These will be reviewed later.

Appendix B

Input variables to the BDT

B.1 Ranking

BDT 1bjet (tZq)				
Ranking	eee	$ee\mu$	$e\mu\mu$	$\mu\mu\mu$
1	ΔR_{jj}	ΔR_{jj}	p_T^Q	d_{CSV}
2	η_j	d_{CSV}	η_j	m_{top}
3	$Asym_\ell$	η_Q	ΔR_{jj}	ΔR_{jj}
4	p_T^Q	m_{top}	m_{top}	$\Delta R_{j\ell}$
5	η_ℓ	η_j	d_{CSV}	p_T^Q
6	d_{CSV}	$\log(\omega_{\text{tZq}})$	$\Delta R_{j\ell}$	$\log(\omega_{\text{tZq}})$
7	$\Delta\phi_{Z\ell}$	$Asym_\ell$	$\Delta\phi_{b\ell}$	η_Z
8	m_{top}	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z}-\text{WZ})}$	$\Delta\phi_{Z\ell}$	$Asym_\ell$
9	$\log(\omega_{\text{tZq}})$	p_T^Q	η_Q	η_Q
10	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z}-\text{WZ})}$	$\Delta R_{j\ell}$	η_Z	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z}-\text{WZ})}$
11	η_Z	$\text{KIN}\omega_{(\text{t}\bar{\text{t}}\text{Z})}$	$\log(\omega_{\text{tZq}})$	$\Delta\phi_{b\ell}$
12	η_Q	$\Delta\phi_{b\ell}$	$Asym_\ell$	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$
13	$\Delta R_{j\ell}$	η_Z	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z}-\text{WZ})}$	η_ℓ
14	$\Delta\phi_{b\ell}$	η_ℓ	η_ℓ	η_j
15	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	N_{jets}	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	$\text{KIN}\omega_{(\text{t}\bar{\text{t}}\text{Z})}$
16	N_{jets}	$\Delta\phi_{Z\ell}$	N_{jets}	$\Delta\phi_{Z\ell}$
17	$\text{KIN}\omega_{(\text{t}\bar{\text{t}}\text{Z})}$	$\mathcal{LR}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	$\text{KIN}\omega_{(\text{t}\bar{\text{t}}\text{Z})}$	N_{jets}

TABLE B.1: Ranking of all the input variables in the BDT for the 1bjet region, in the four different channels.

BDT 2bjet ($t\bar{t}Z$)				
1	m_{top}	N_{jets}	N_{jets}	N_{jets}
2	N_{jets}	$Asym_{\ell}$	m_{top}	$\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$
3	d_{CSV}	d_{CSV}	$Asym_{\ell}$	m_{top}
4	η_{Q}	$\log(\omega_{\text{tZq}})$	ΔR_{jj}	$Asym_{\ell}$
5	$Asym_{\ell}$	m_{top}	d_{CSV}	$\Delta R_{j\ell}$
6	$\log(\omega_{\text{tZq}})$	η_{Q}	$\log(\omega_{\text{tZq}})$	ΔR_{jj}
7	p_T^{Q}	$\Delta R_{j\ell}$	η_{Q}	η_{Q}
8	$\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	$\Delta R_{Z,\text{top}}$	$\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	d_{CSV}
9	ΔR_{jj}	$\mathcal{LR}^{\text{rescaled}}_{(\text{tZq}-\text{t}\bar{\text{t}}\text{Z})}$	η_{Z}	p_T^{Q}
10	p_T^{Z}	p_T^{Z}	p_T^{Q}	$\Delta R_{Z,\text{top}}$
11	η_{Z}	η_{Z}	$\Delta R_{j\ell}$	p_T^{Z}
12	$\Delta R_{j\ell}$	ΔR_{jj}	$\Delta R_{Z,\text{top}}$	$\log(\omega_{\text{tZq}})$
13	$\Delta R_{Q,\ell}$	$\Delta R_{Q,\ell}$	p_T^{Z}	η_{Z}
14	$\Delta R_{Z,\text{top}}$	p_T^{Q}	$\Delta\phi_{Z\ell}$	$\Delta\phi_{Z\ell}$
15	$\Delta\phi_{Z\ell}$	$\Delta\phi_{Z\ell}$	$\Delta R_{Q,\ell}$	$\Delta R_{Q,\ell}$

TABLE B.2: Ranking of all the different input variables in the BDT for the 2bjet region, in the four different channels.

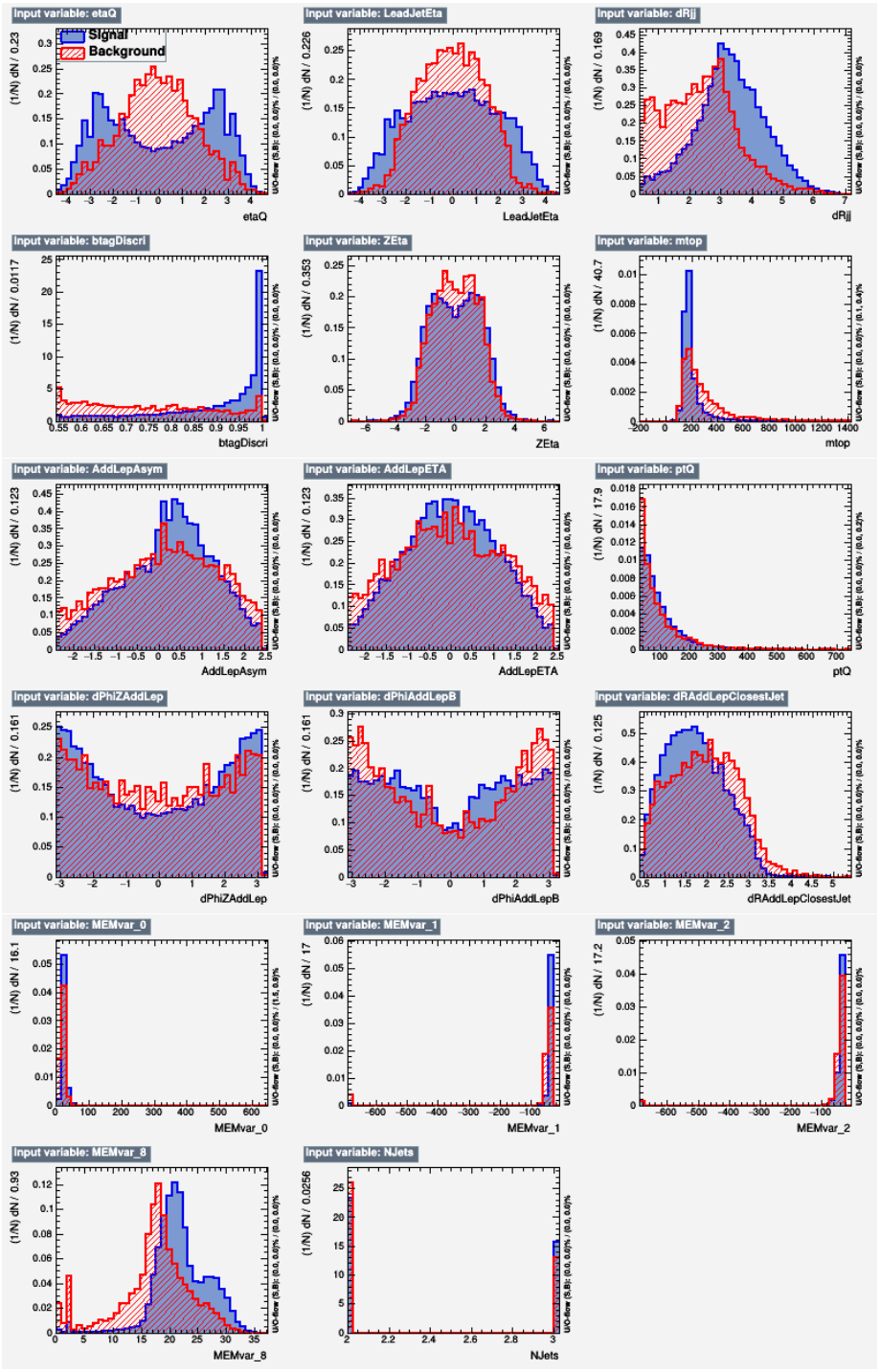


FIGURE B.1: Distributions of the variables used as input to the BDT training for tZq (1bjet) in the $\mu\mu\mu$ channel. The jet multiplicity is also shown, even though it is not used as input to the BDT.

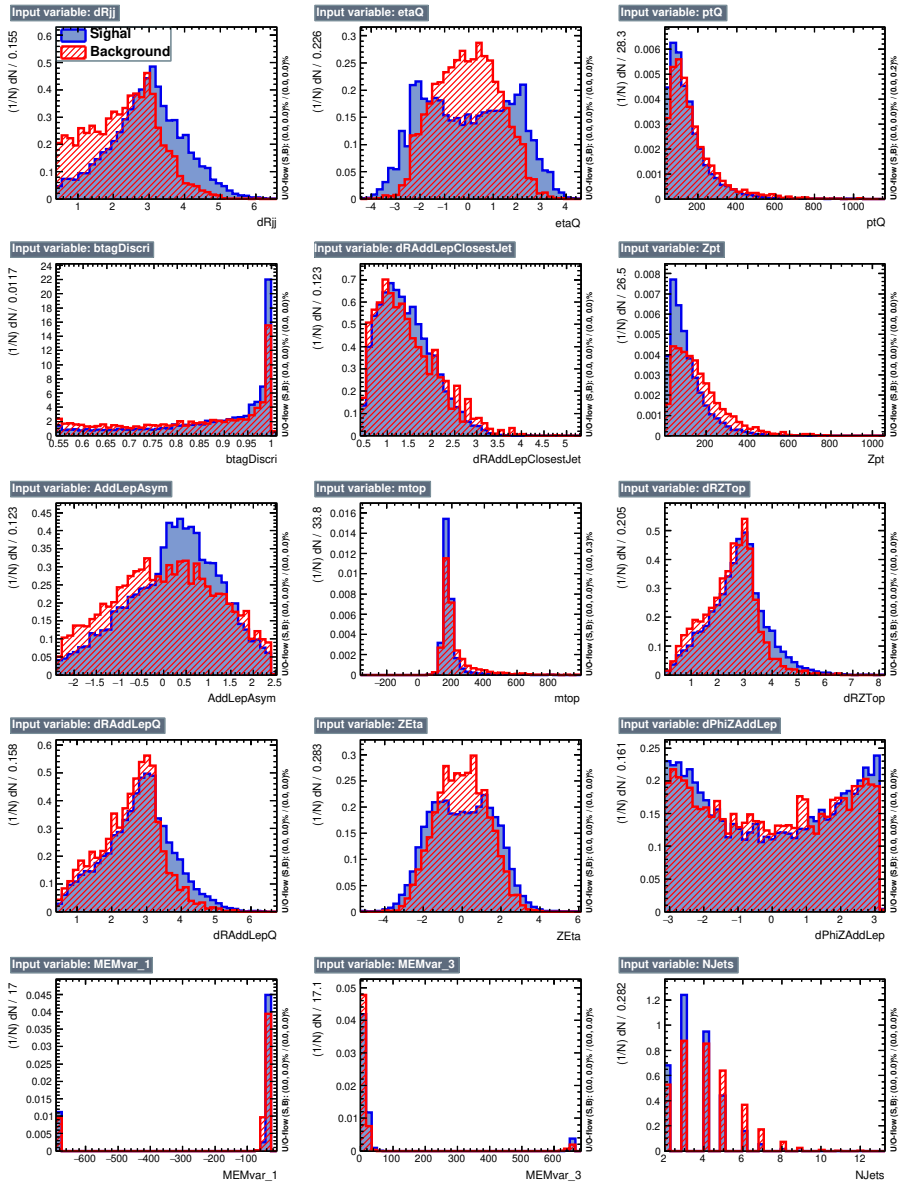


FIGURE B.2: Distributions of the variables used as input to the BDT training for $t\bar{t}Z$ (2bjet) in the $\mu\mu\mu$ channel. The jet multiplicity is also shown, even though it is not used as input to the BDT.

B.2 Control Plots

B.2.1 Signal region (1bjet)

B.2.2 $t\bar{t}Z$ region (2bjet)

B.2.3 WZ region (0bjet)

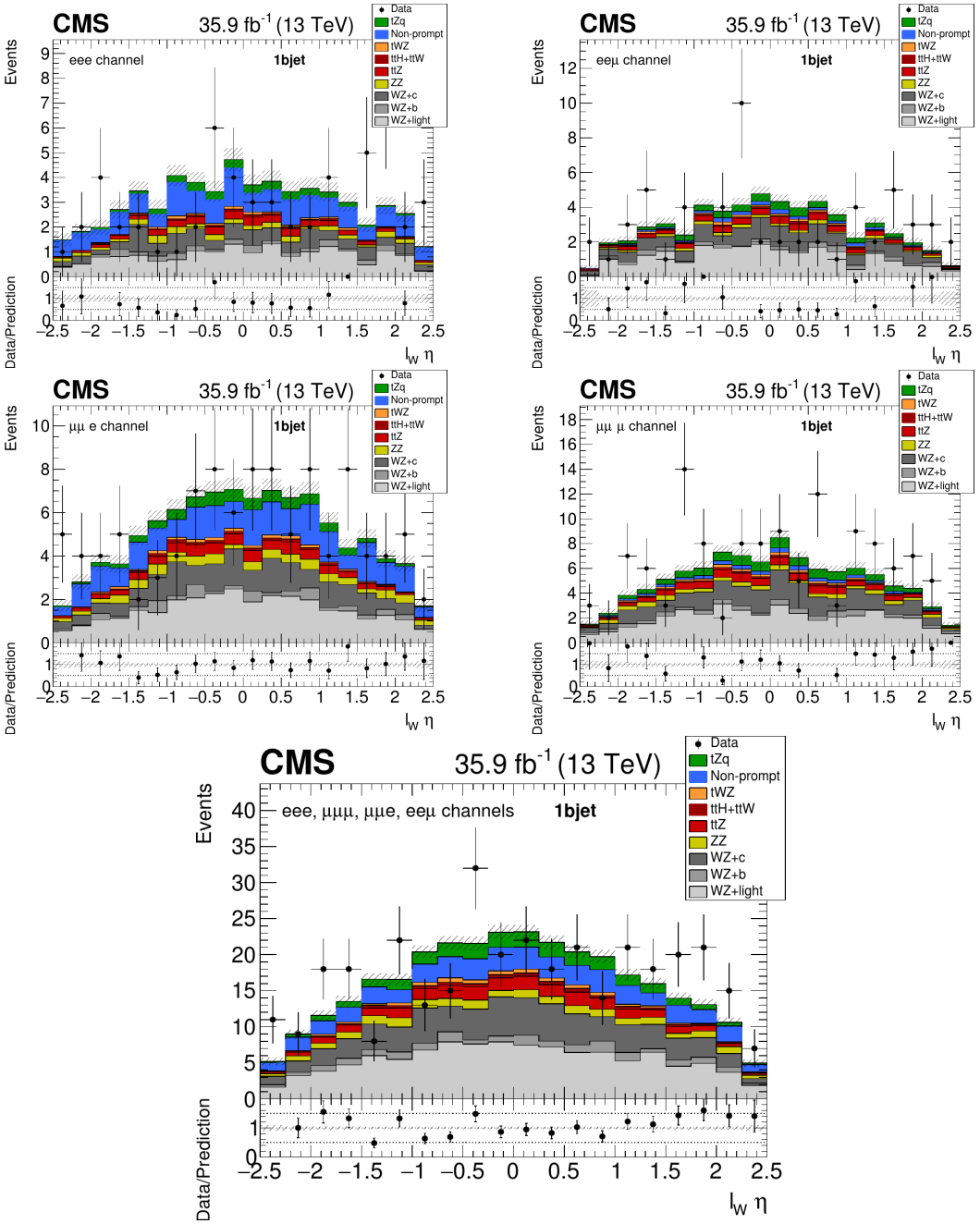


FIGURE B.3: Data-to-simulation comparison plots of the η distribution of the lepton associated to the top quark decay for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

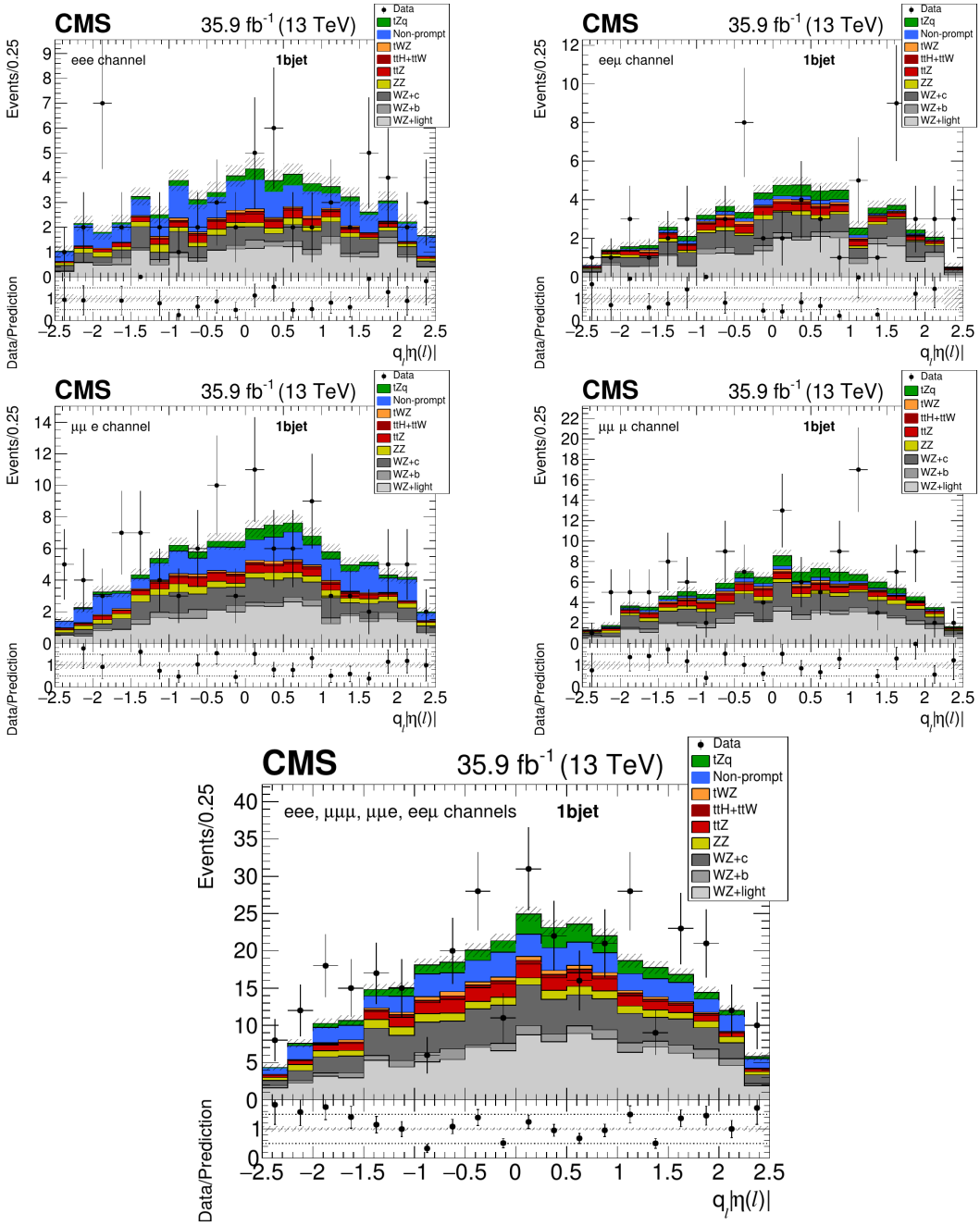


FIGURE B.4: Data-to-simulation comparison plots of the asymmetry distribution of the lepton associated to the top quark decay for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

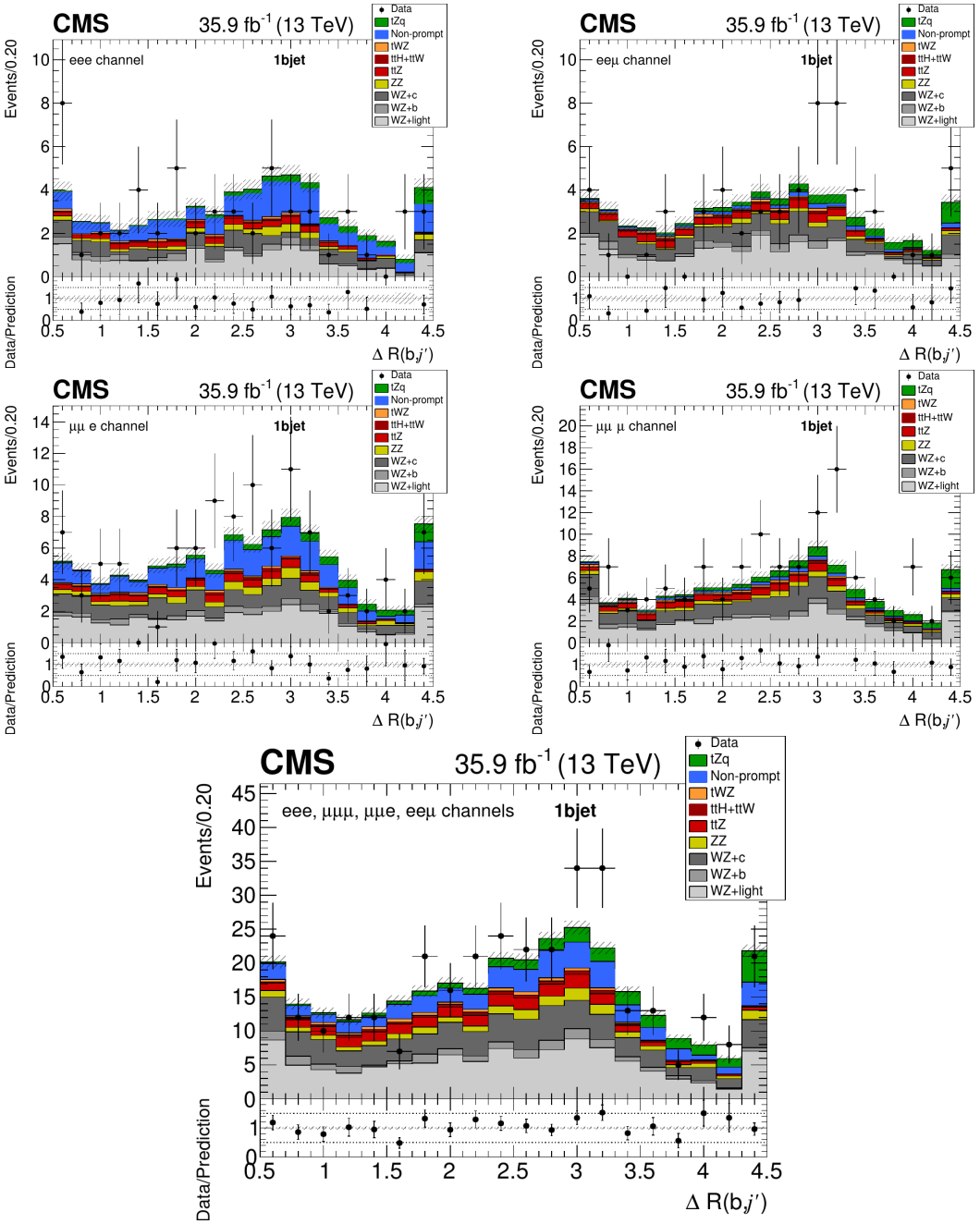


FIGURE B.5: Data-to-simulation comparison plots of the ΔR separation between the forward and the b-tagged jet, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

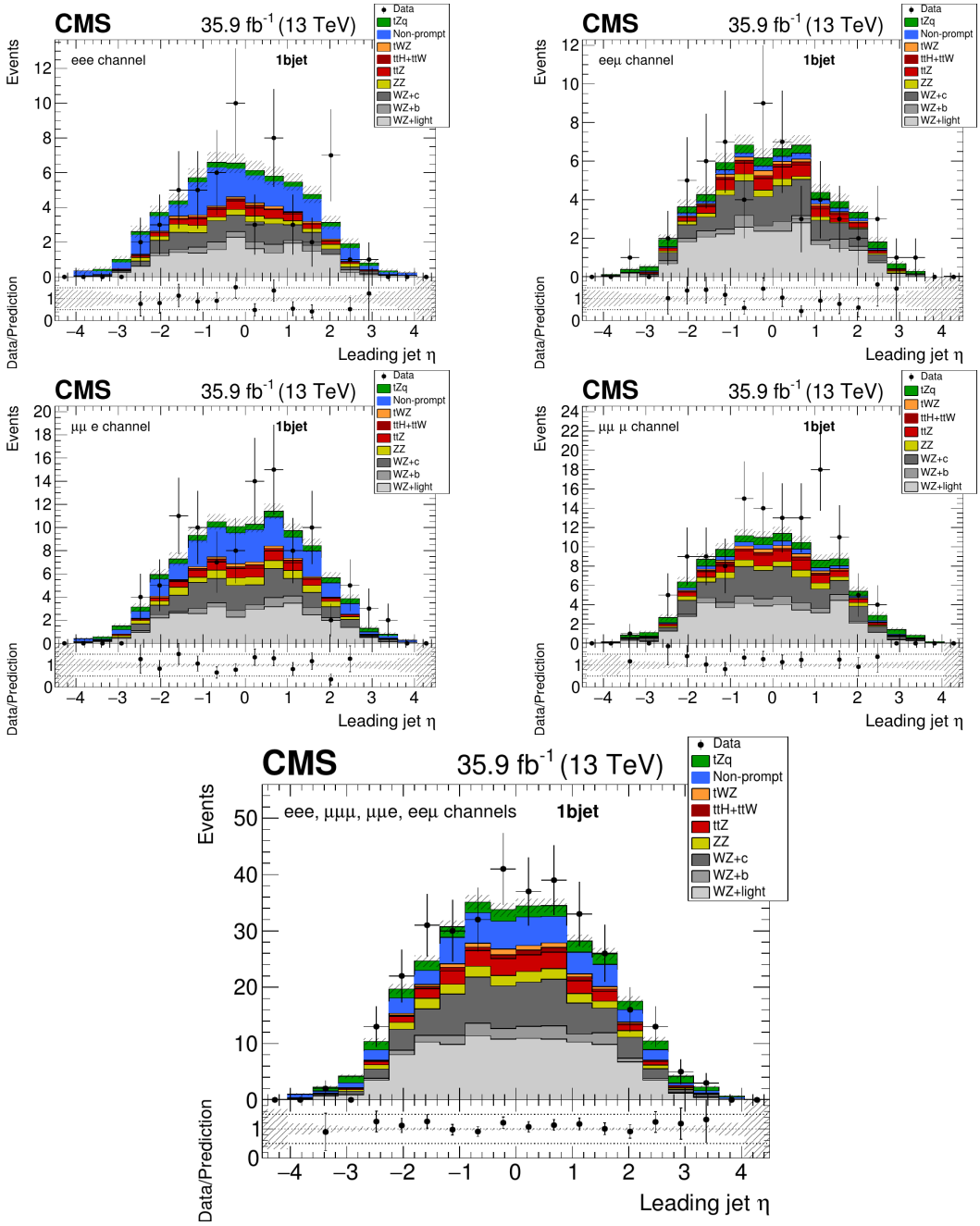


FIGURE B.6: Data-to-simulation comparison plots of the η distribution of the leading jet, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

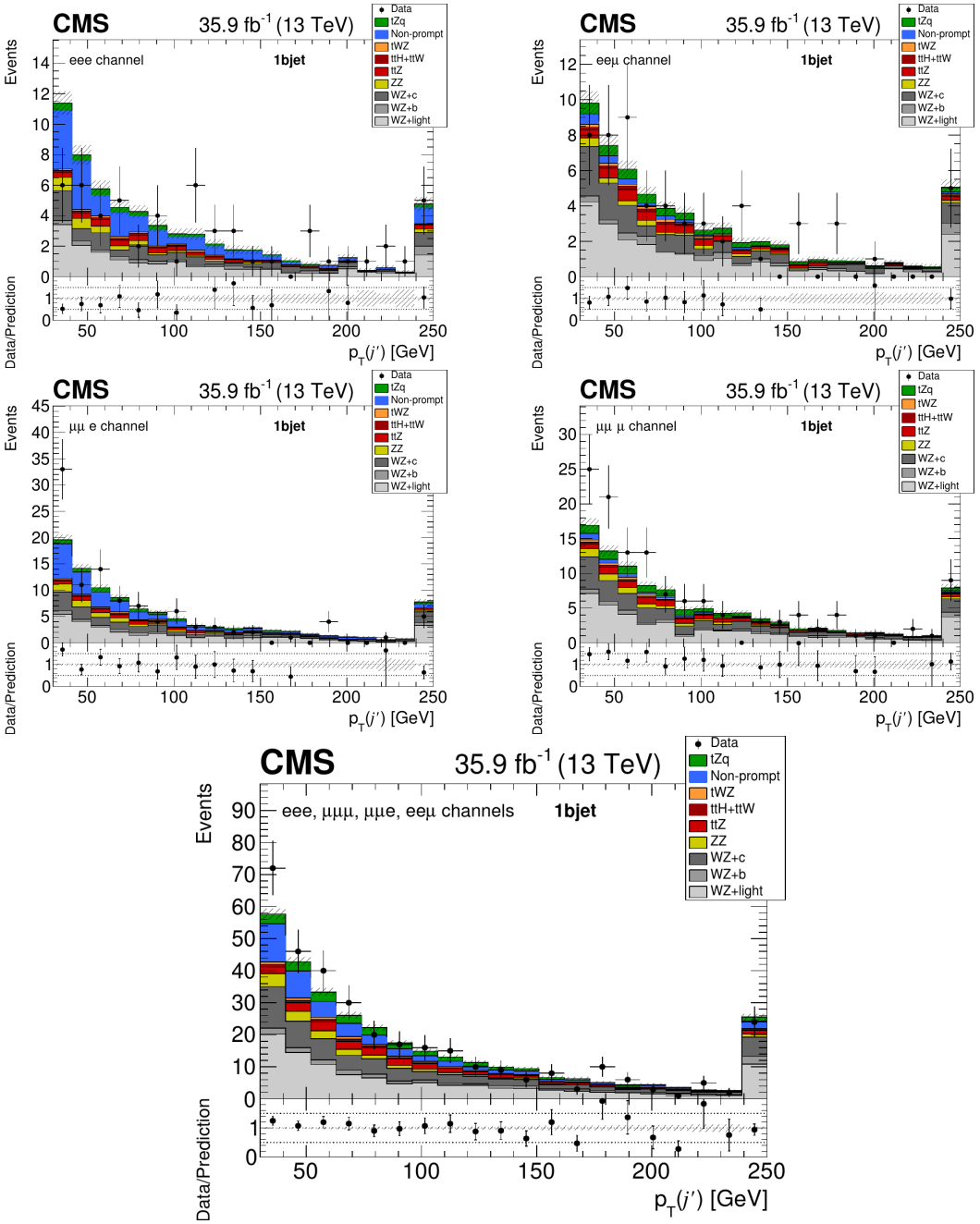


FIGURE B.7: Data-to-simulation comparison plots of the p_T distribution of the forward jet, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

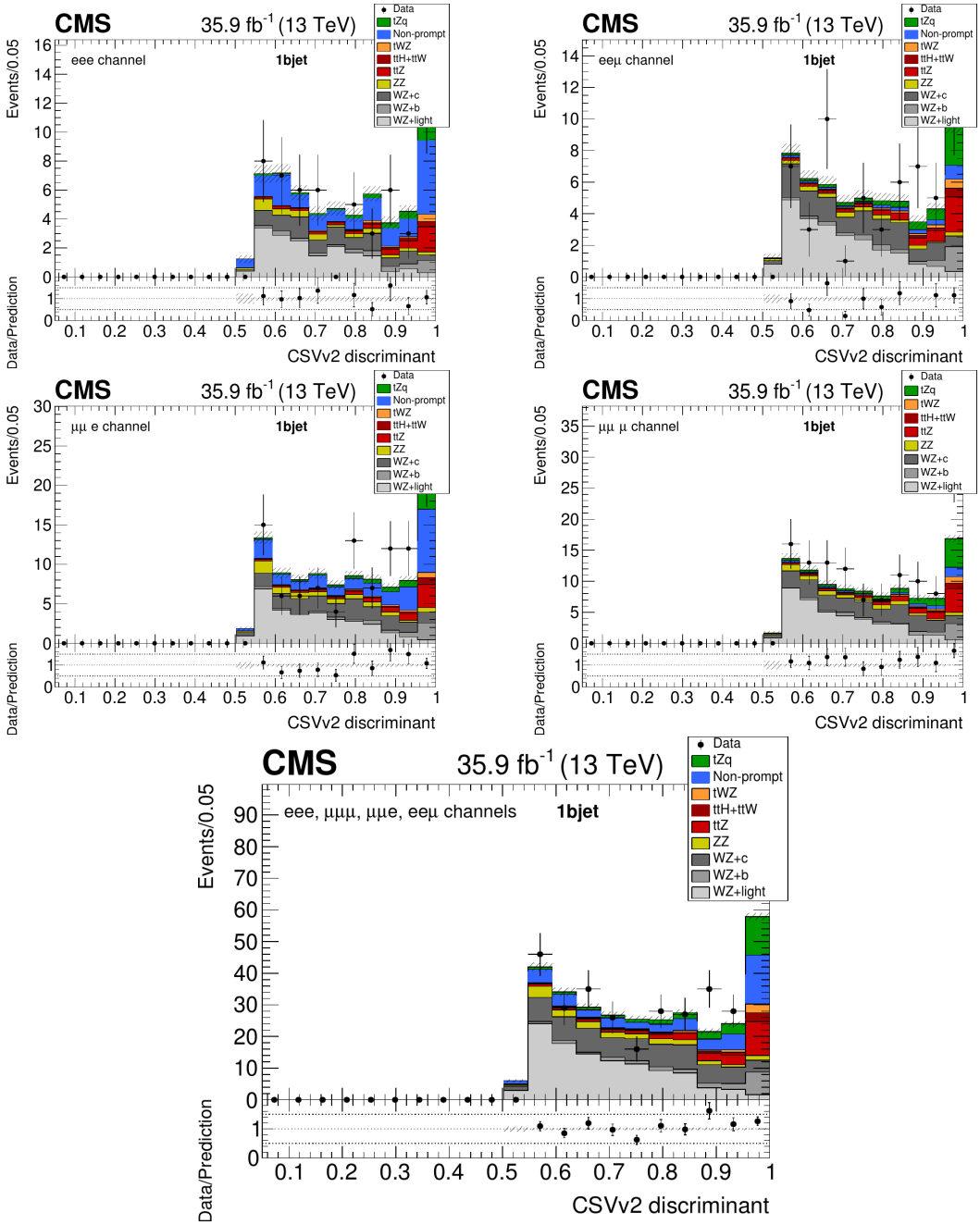


FIGURE B.8: Data-to-simulation comparison plots of the largest CSVv2 discriminant value among all selected jets, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

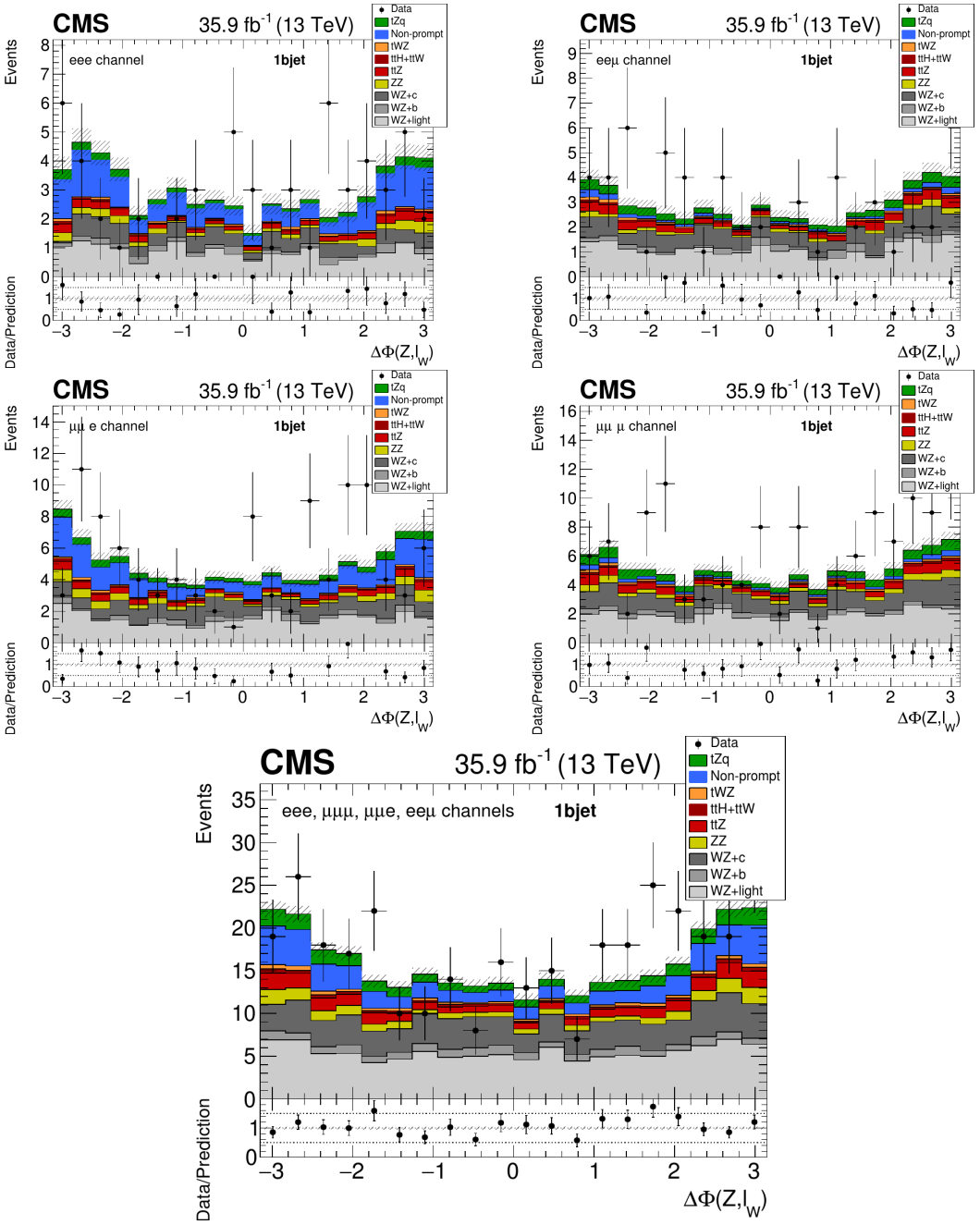


FIGURE B.9: Data-to-simulation comparison plots of the azimuthal separation between the Z boson and the lepton associated to the top quark decay, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

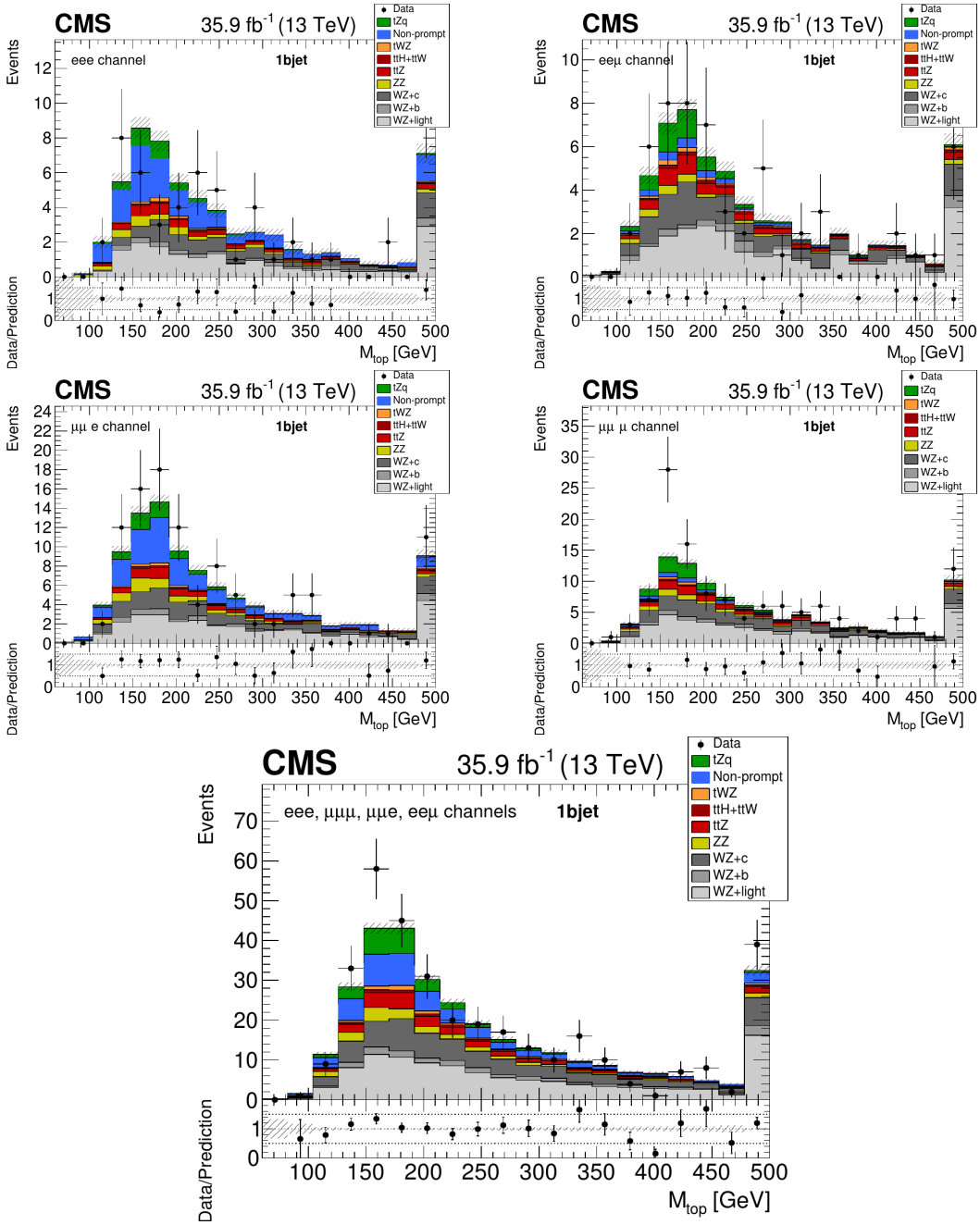


FIGURE B.10: Data-to-simulation comparison plots of the reconstructed top quark mass, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

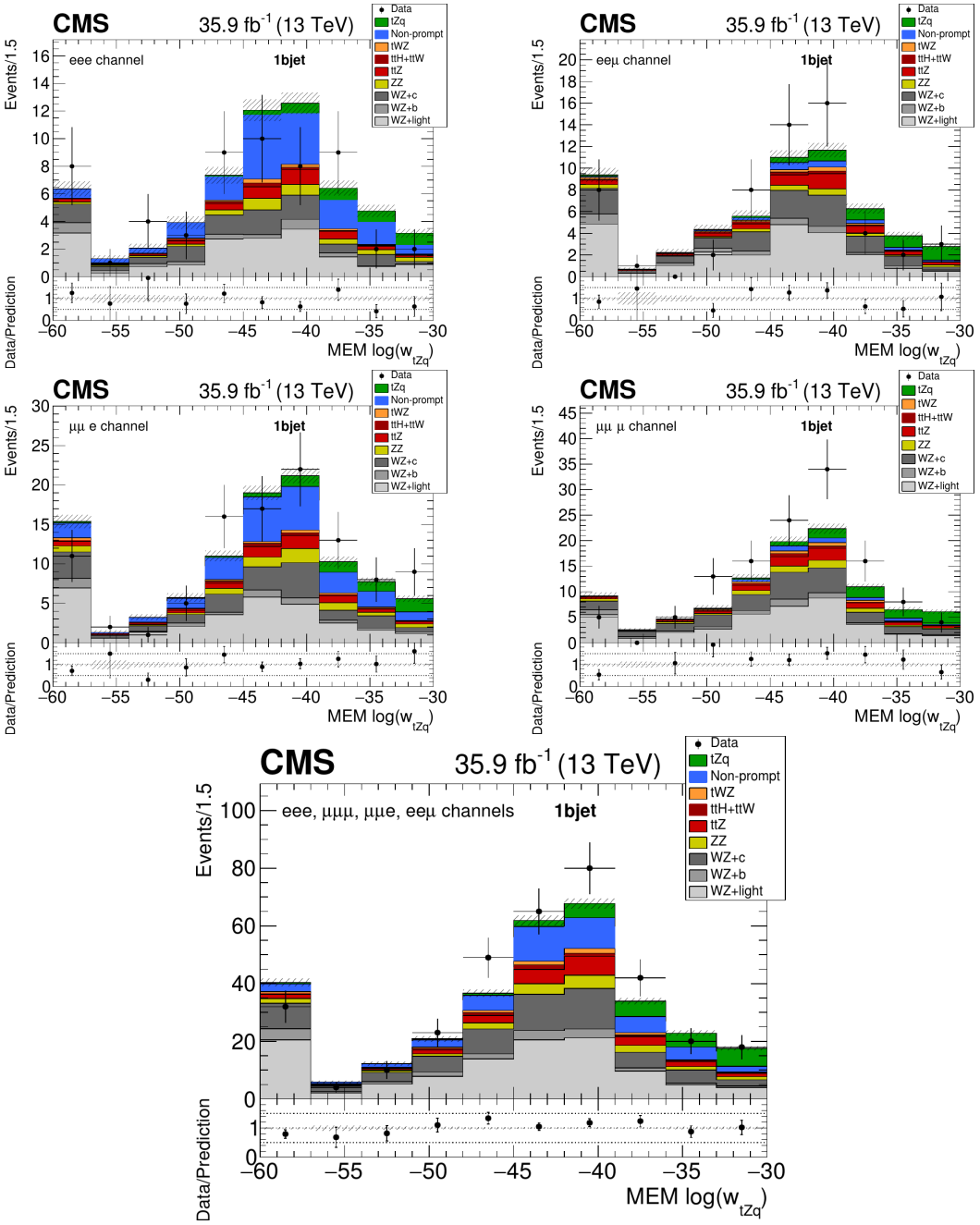


FIGURE B.11: Data-to-simulation comparison plots of the logarithm of the MEM score associated to the most probable tZq kinematic configuration, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

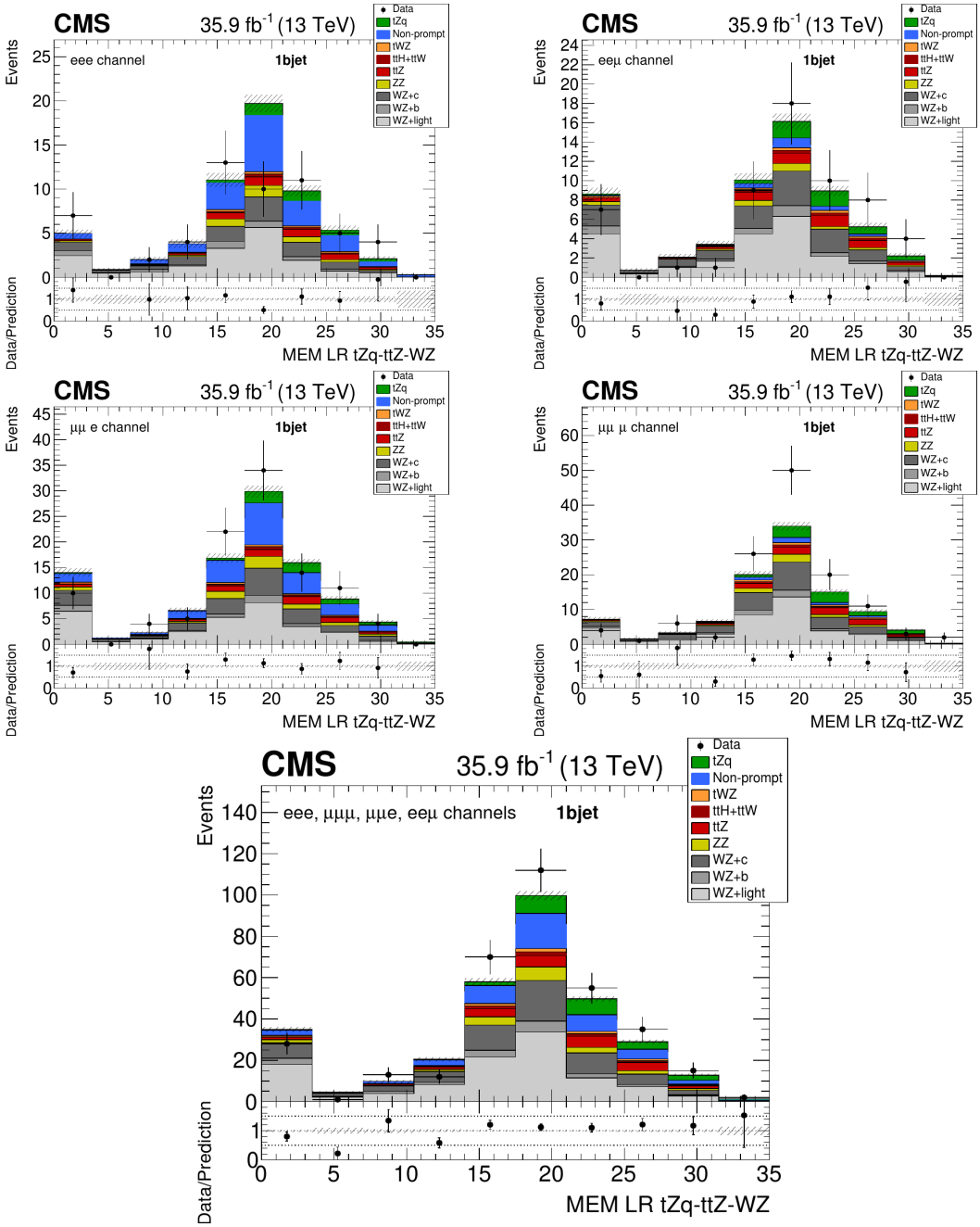


FIGURE B.12: Data-to-simulation comparison plots of the log-likelihood ratio of the tZq hypothesis agains the $t\bar{t}Z$ + WZ hypothesis, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

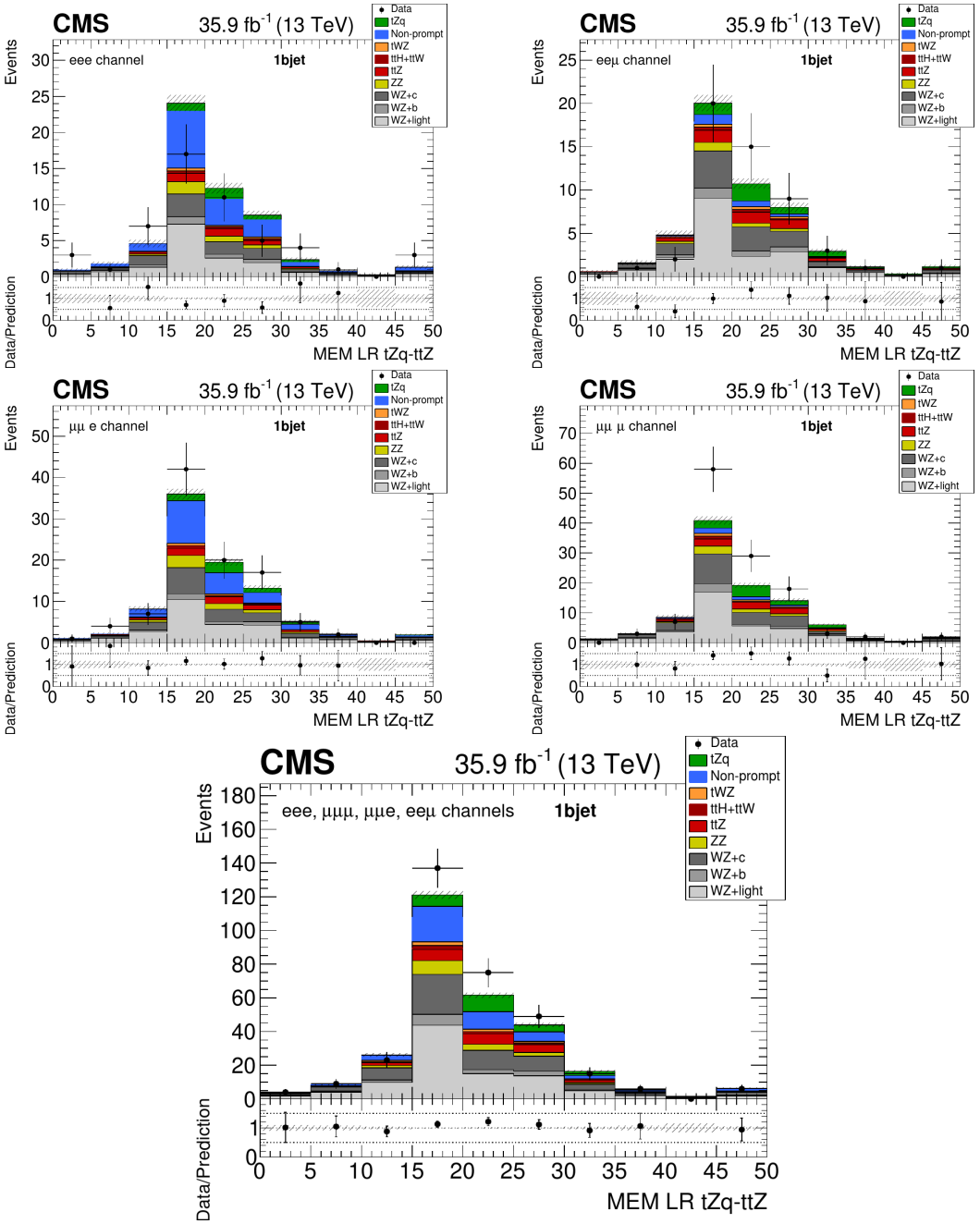


FIGURE B.13: Data-to-simulation comparison plots of the log-likelihood ratio of the tZq hypothesis agains the $t\bar{t}Z$ hypothesis, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

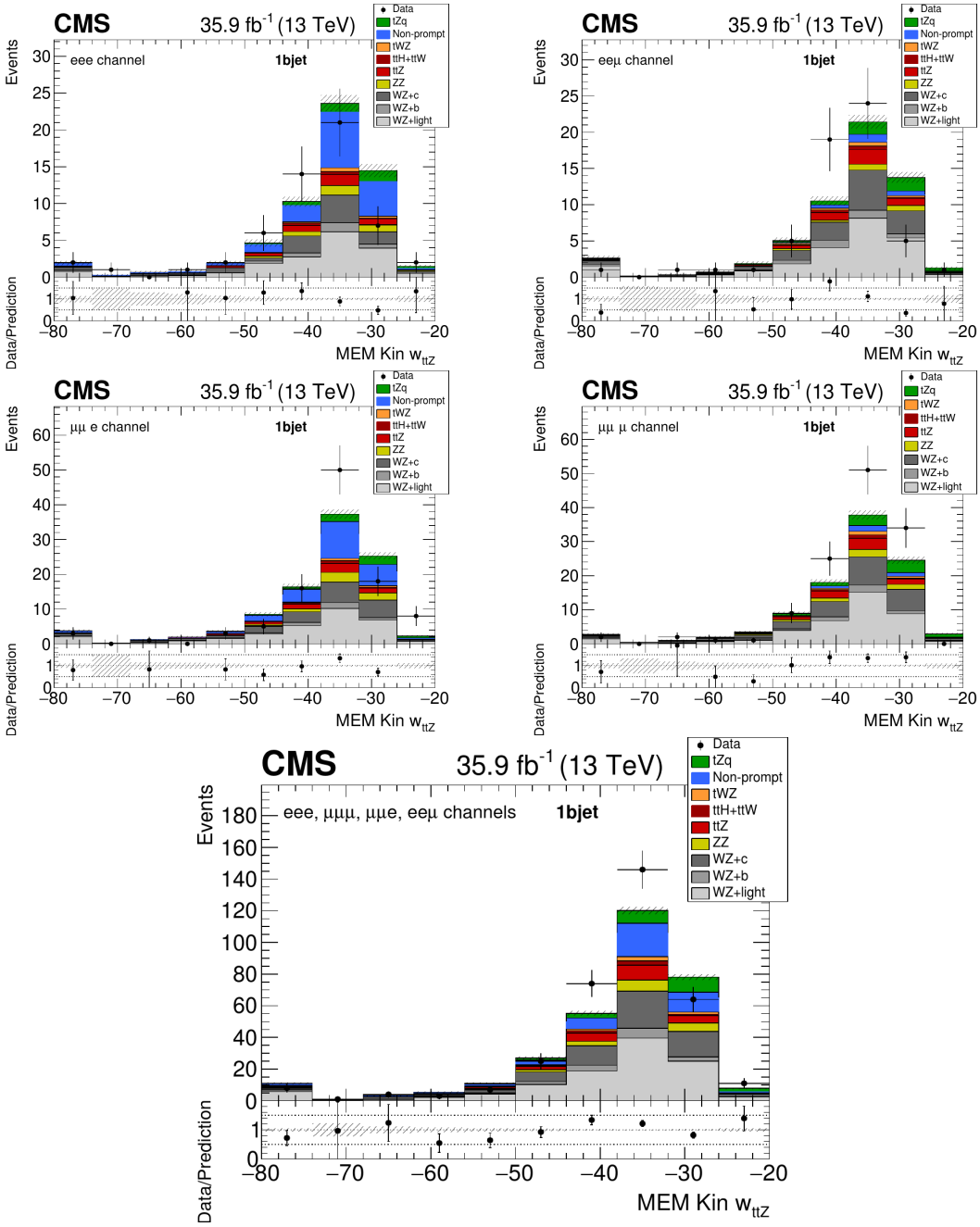


FIGURE B.14: Data-to-simulation comparison plots of the logarithm of the MEM score associated to the most probable $t\bar{t}Z$ kinematic configuration, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

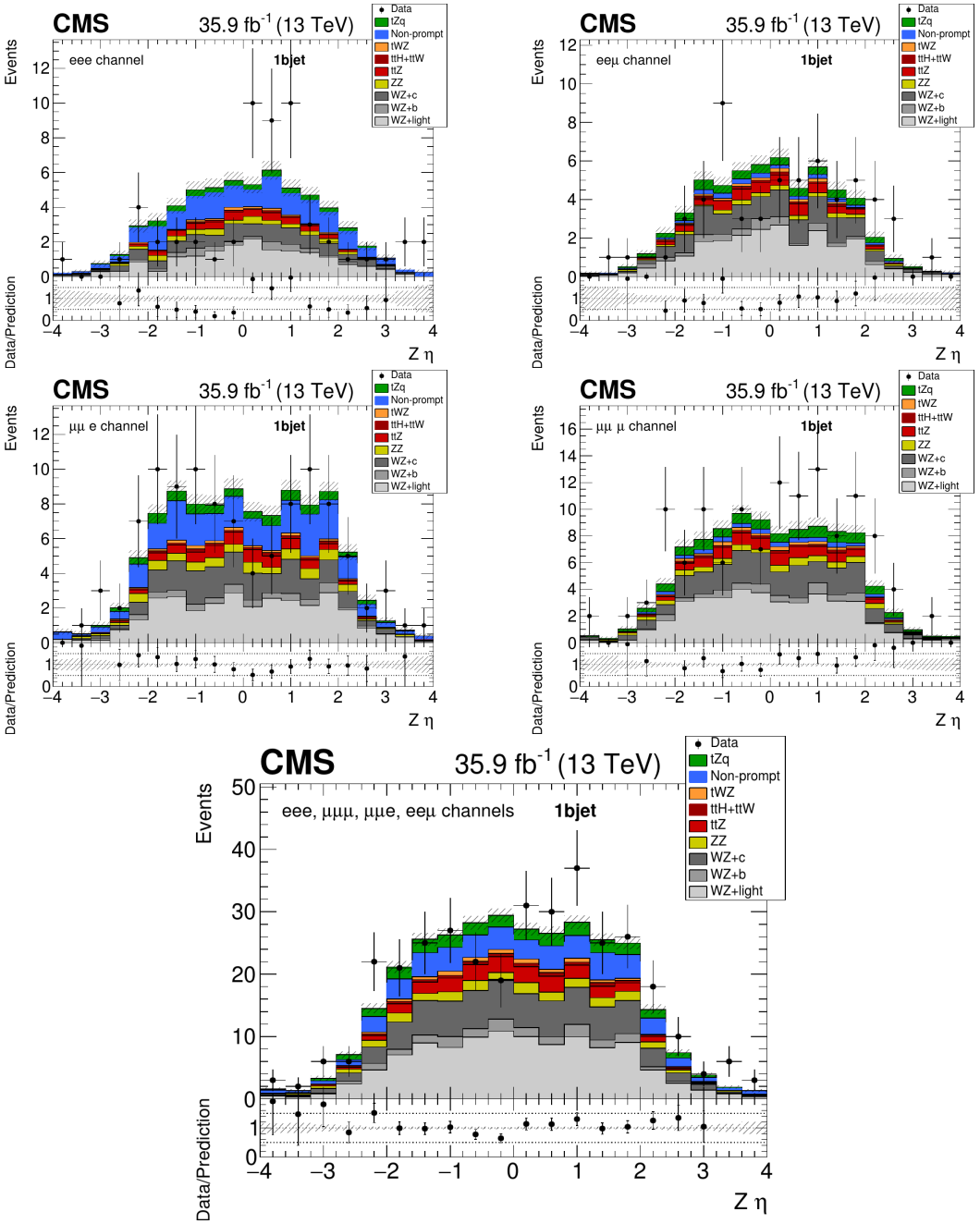


FIGURE B.15: Data-to-simulation comparison plots of the η distribution of the reconstructed Z boson, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

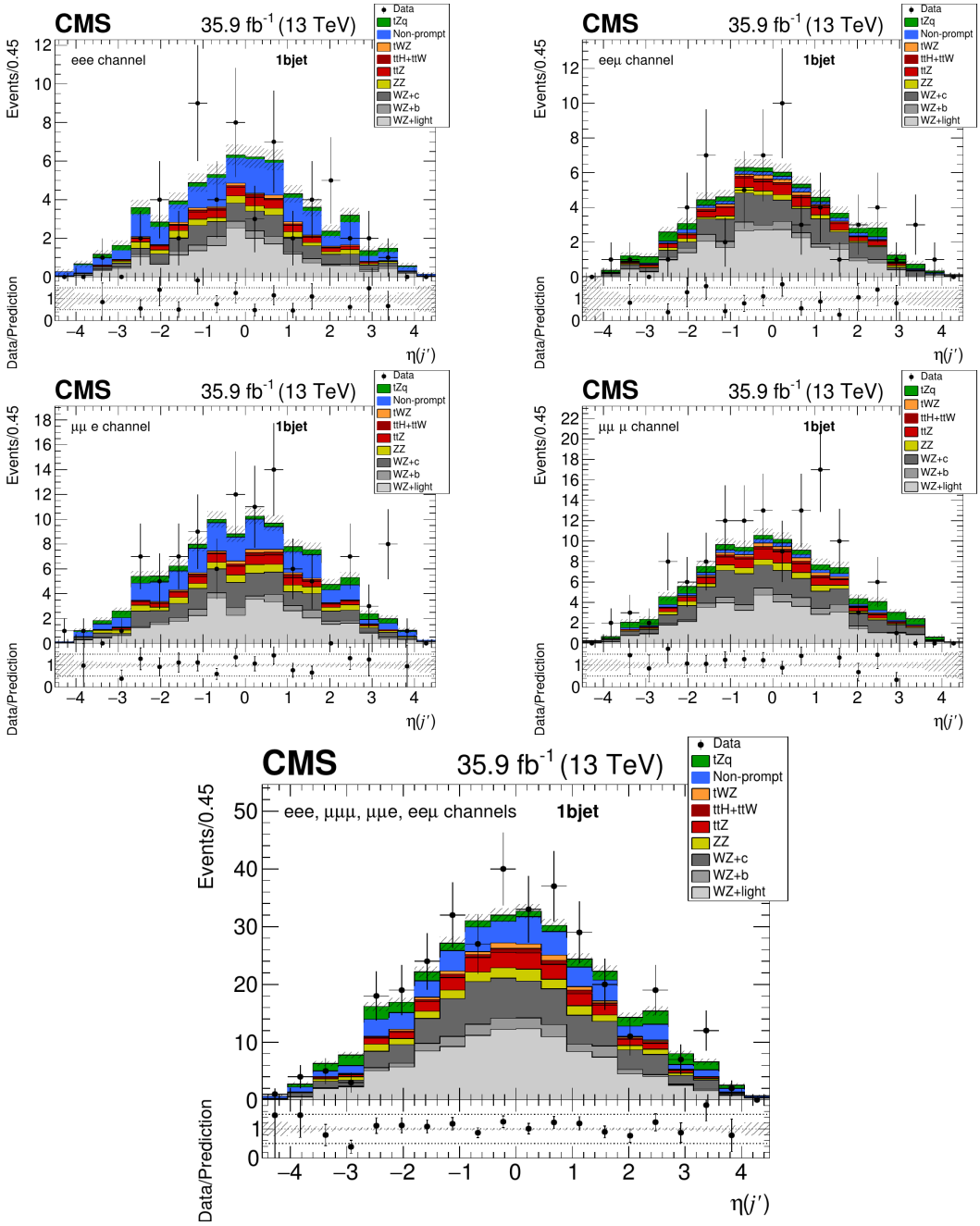


FIGURE B.16: Data-to-simulation comparison plots of the η distribution of the forward jet, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

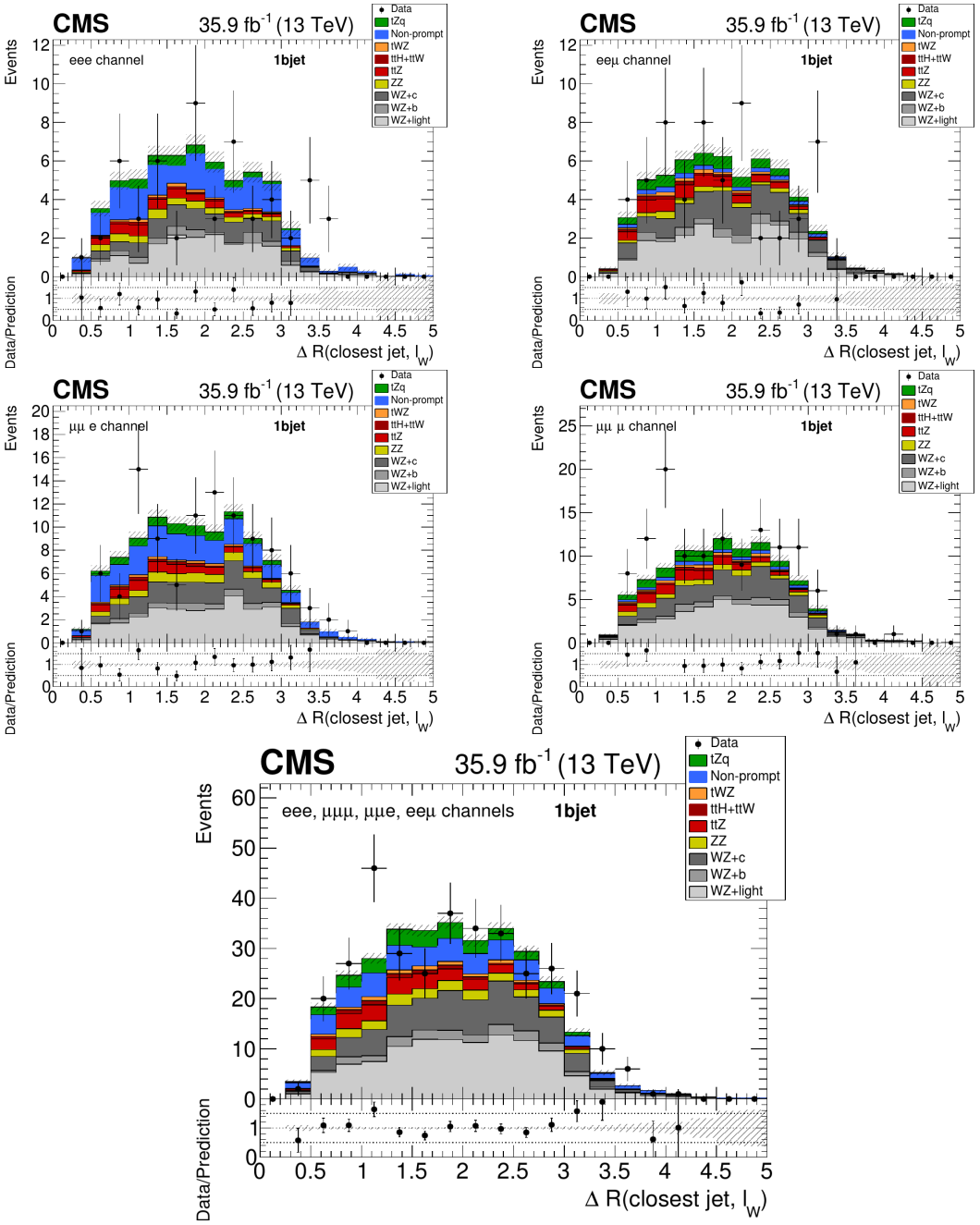


FIGURE B.17: Data-to-simulation comparison plots of the ΔR separation between the lepton associated to the top quark decay and its closest jet, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

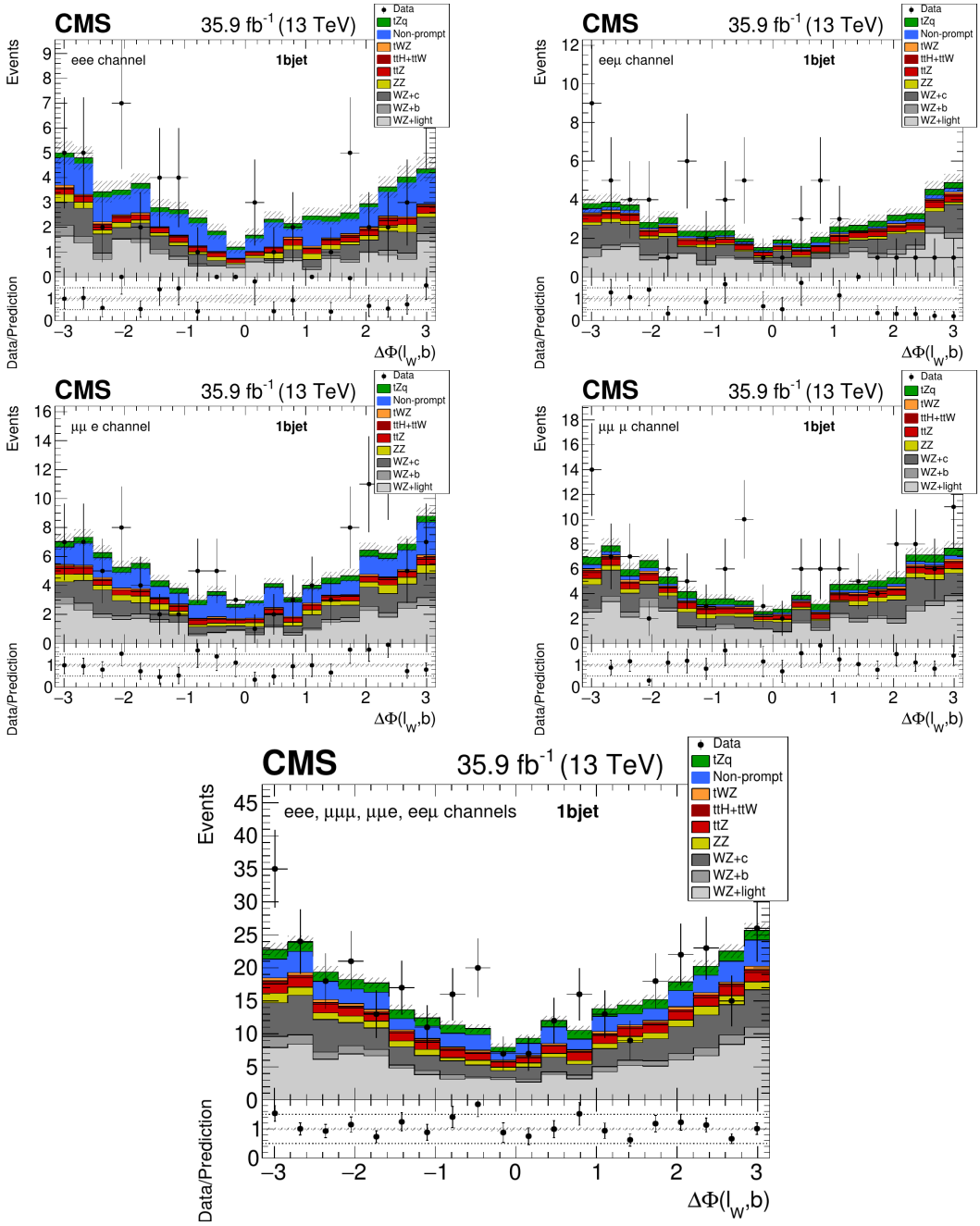


FIGURE B.18: Data-to-simulation comparison plots of the azimuthal separation between the lepton associated to the top quark decay and the b jet, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

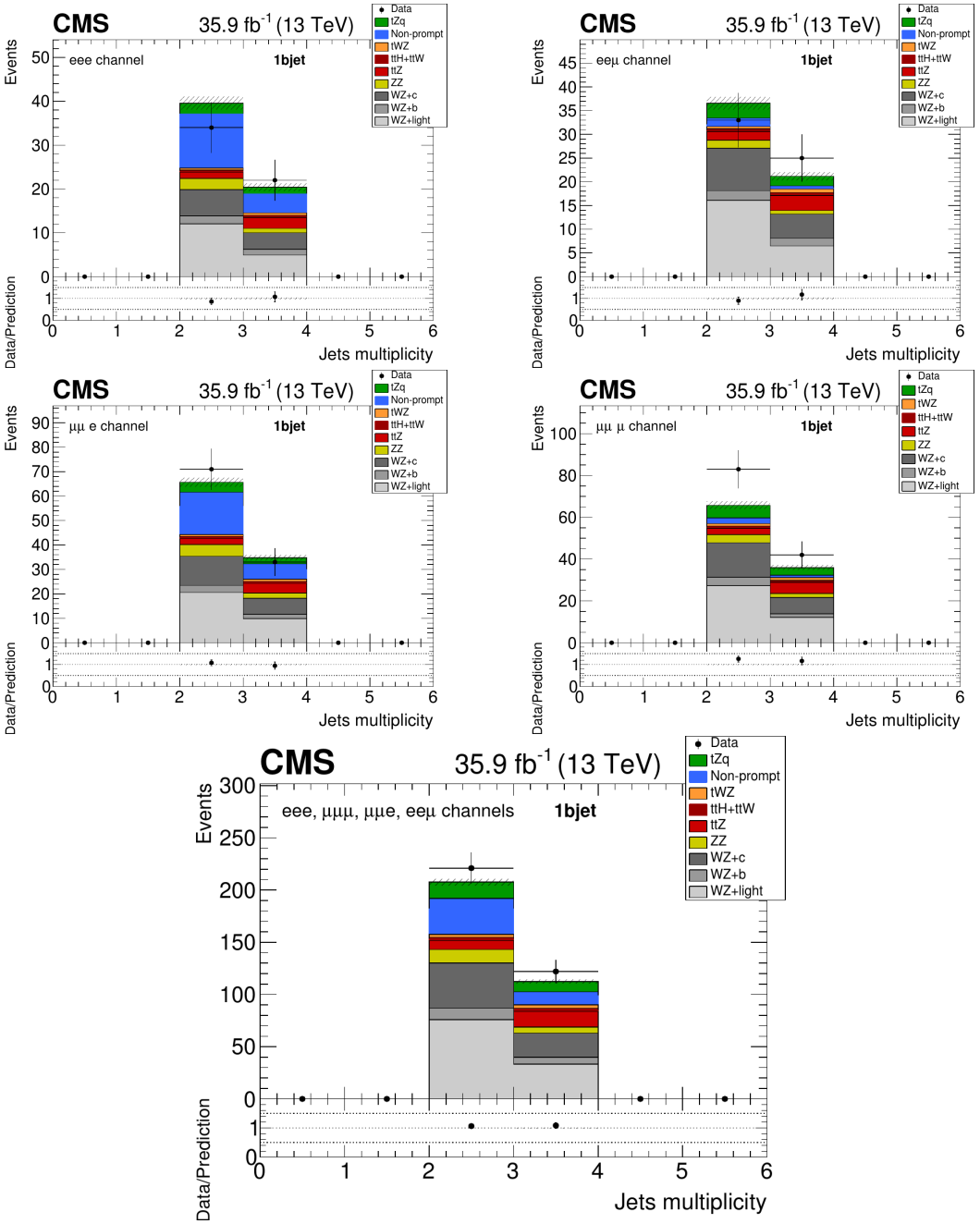


FIGURE B.19: Data-to-simulation comparison plots of the number of selected jets in the event, for the different decay channels in the 1bjet region. The last plot contains the distribution for all channels combined.

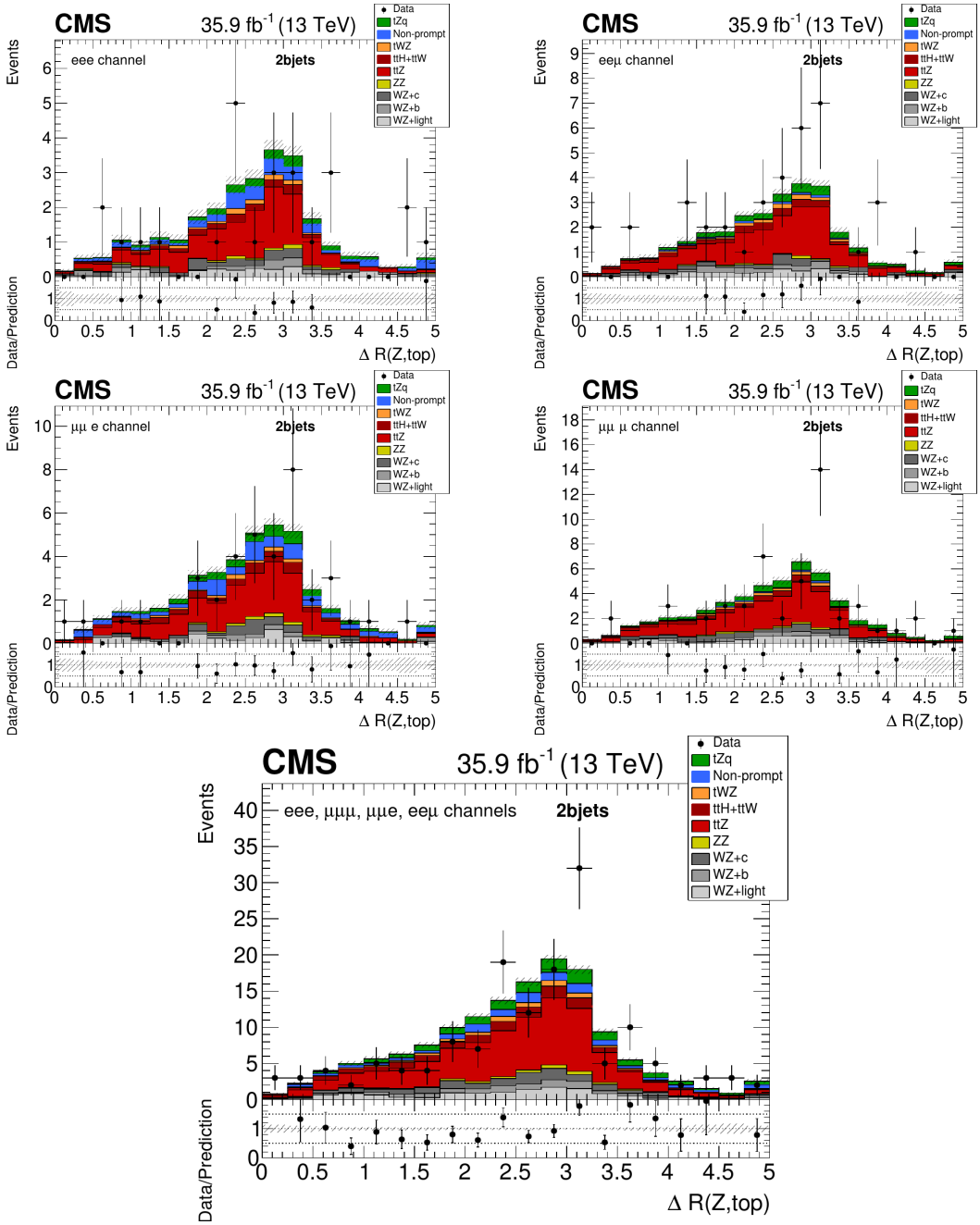


FIGURE B.20:)
Data-to-simulation comparison plots of the ΔR separation between the reconstructed Z boson and the top quark, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

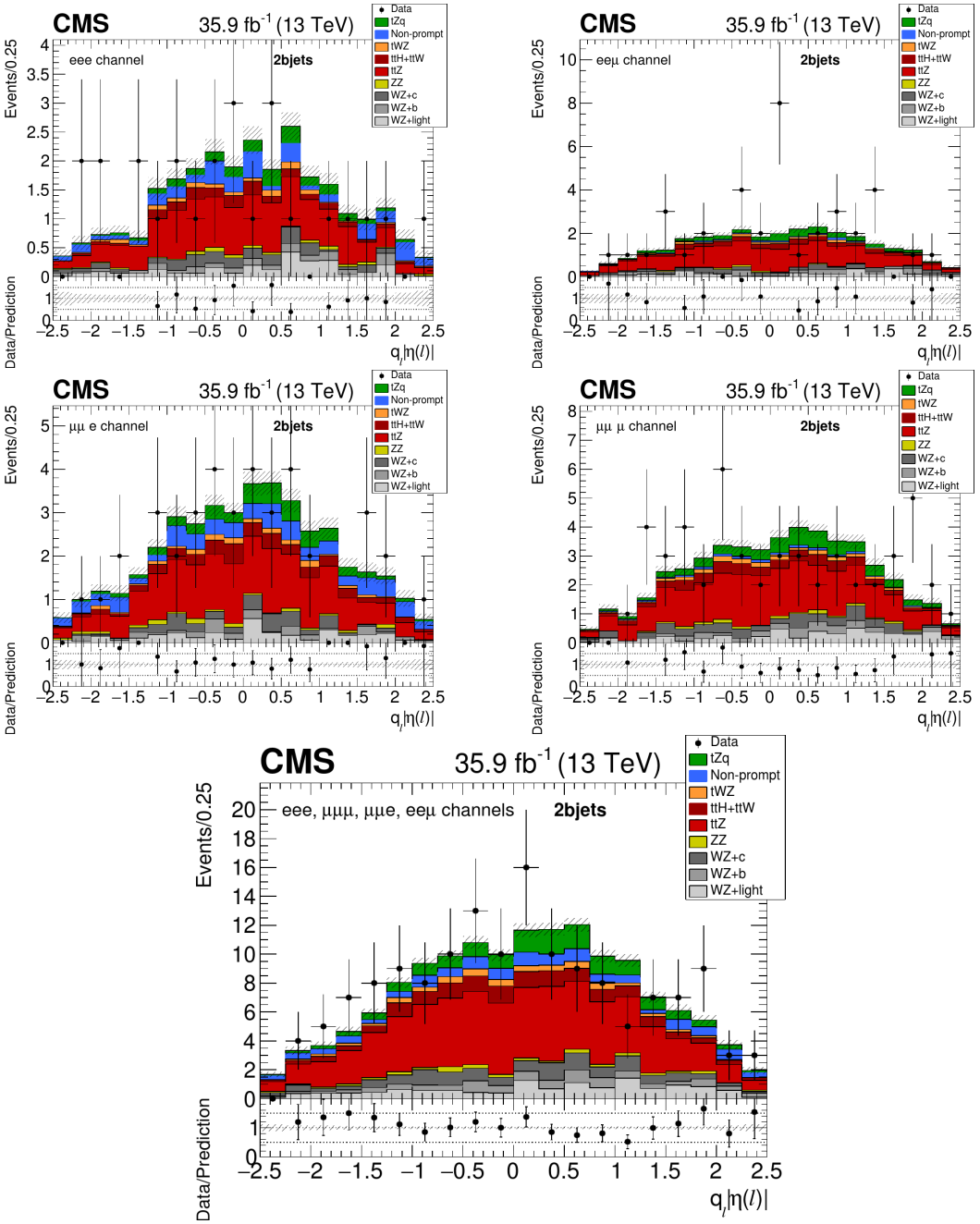


FIGURE B.21:)

Data-to-simulation comparison plots of the asymmetry distribution of the lepton associated to the top quark decay for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

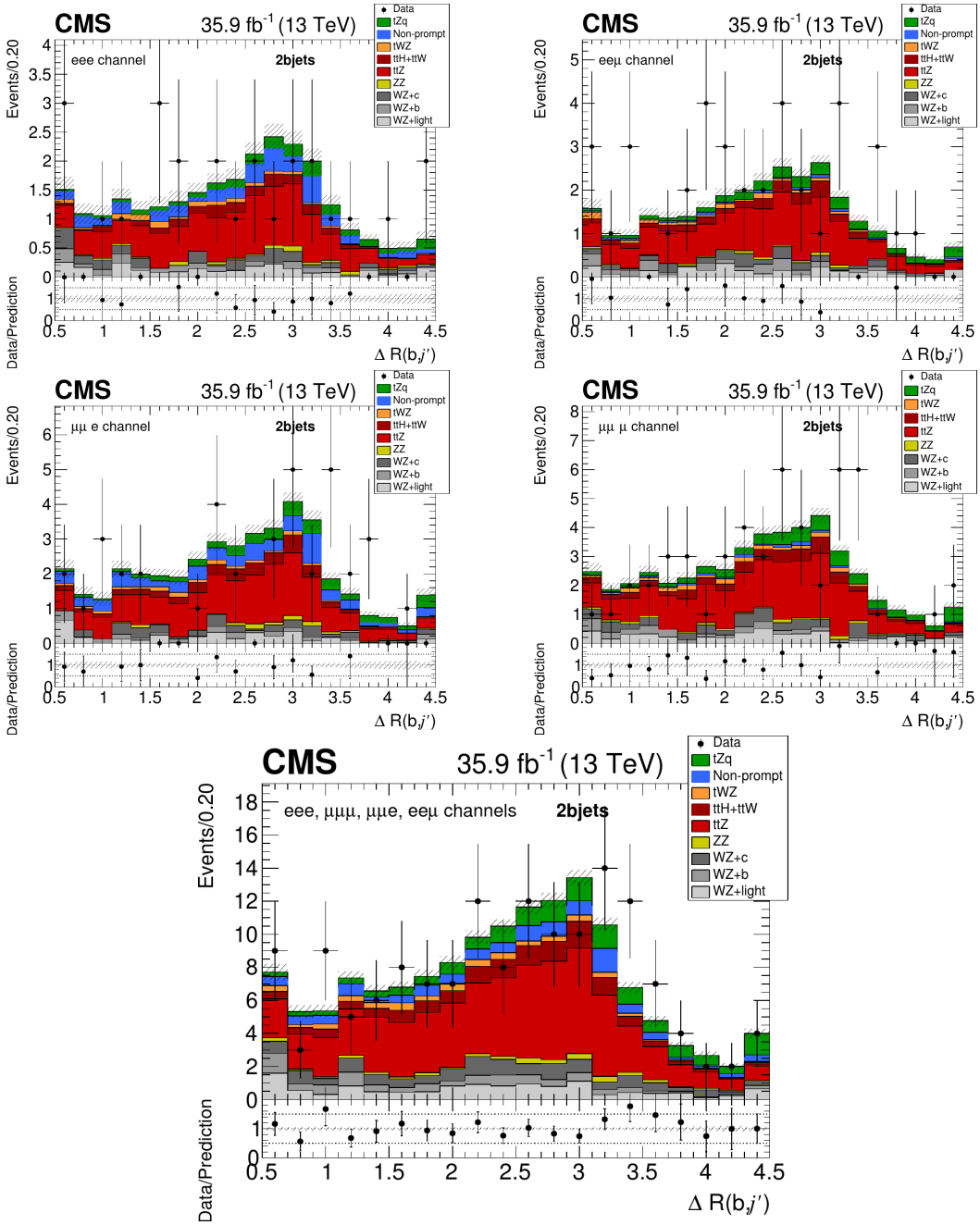


FIGURE B.22:)
Data-to-simulation comparison plots of the ΔR separation between the forward and the b-tagged jet, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

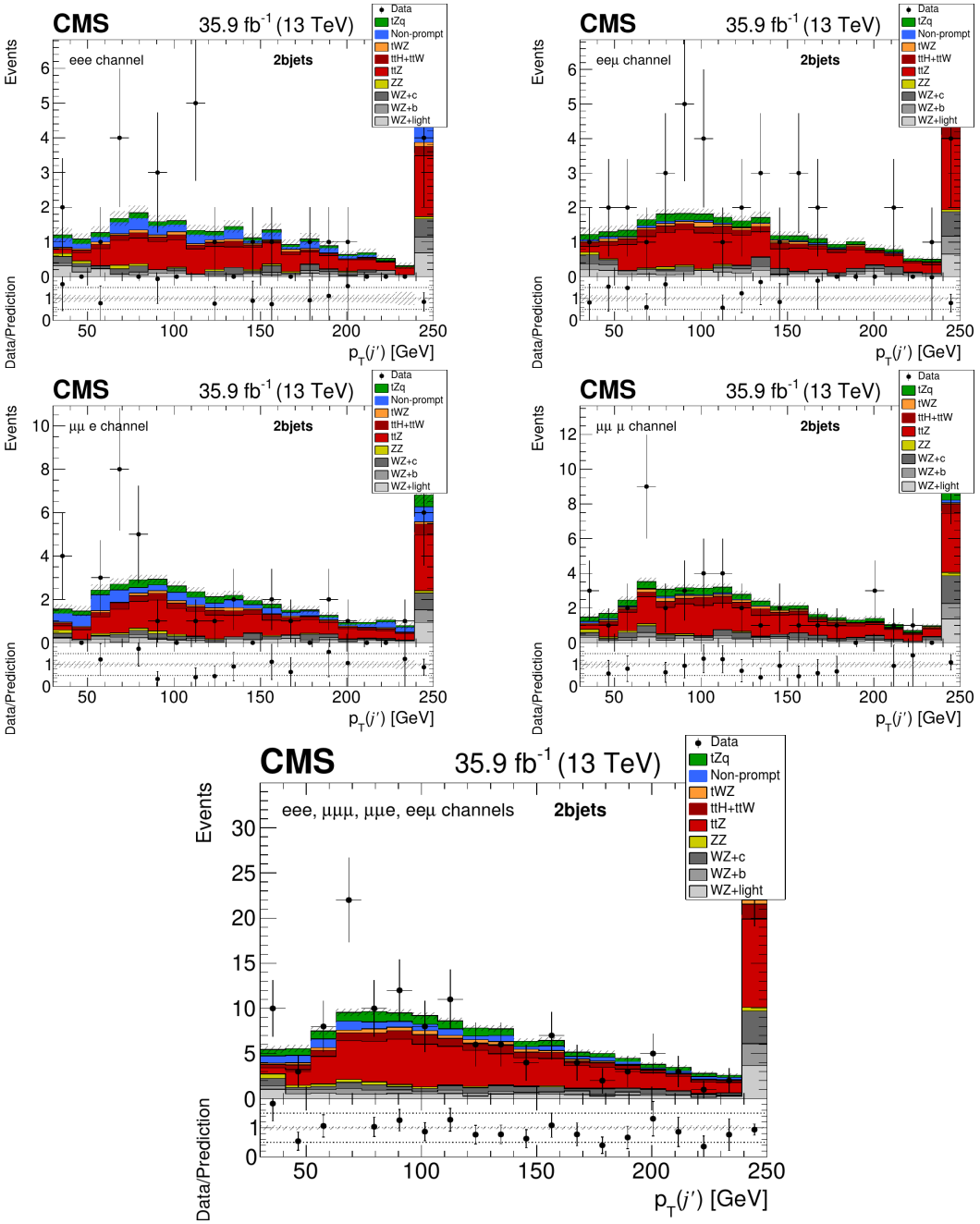


FIGURE B.23:)
Data-to-simulation comparison plots of the p_T distribution of the forward jet, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

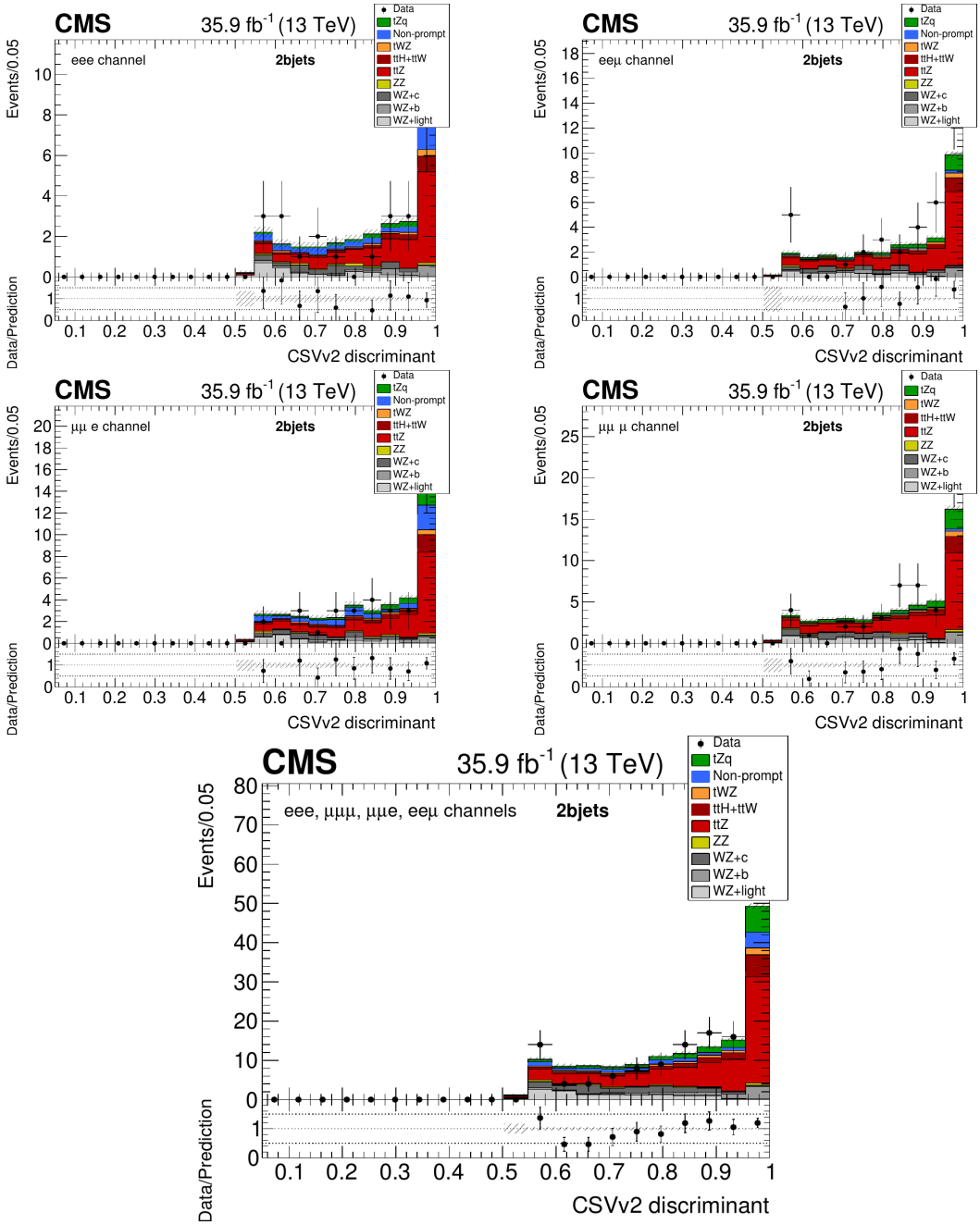


FIGURE B.24:)
Data-to-simulation comparison plots of the largest CSVv2 discriminant value among all selected jets, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

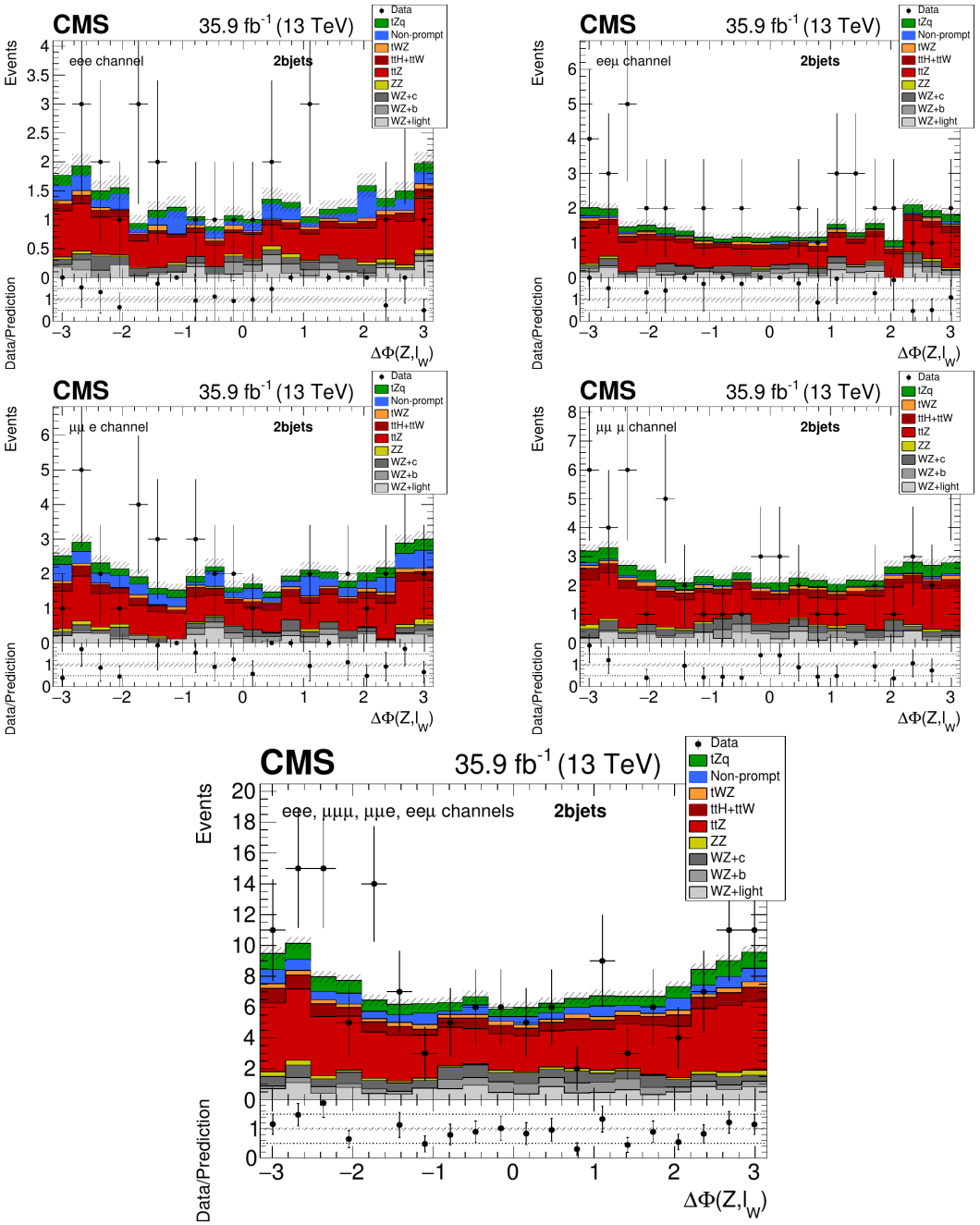


FIGURE B.25:)

Data-to-simulation comparison plots of the azimuthal separation between the Z boson and the lepton associated to the top quark decay, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

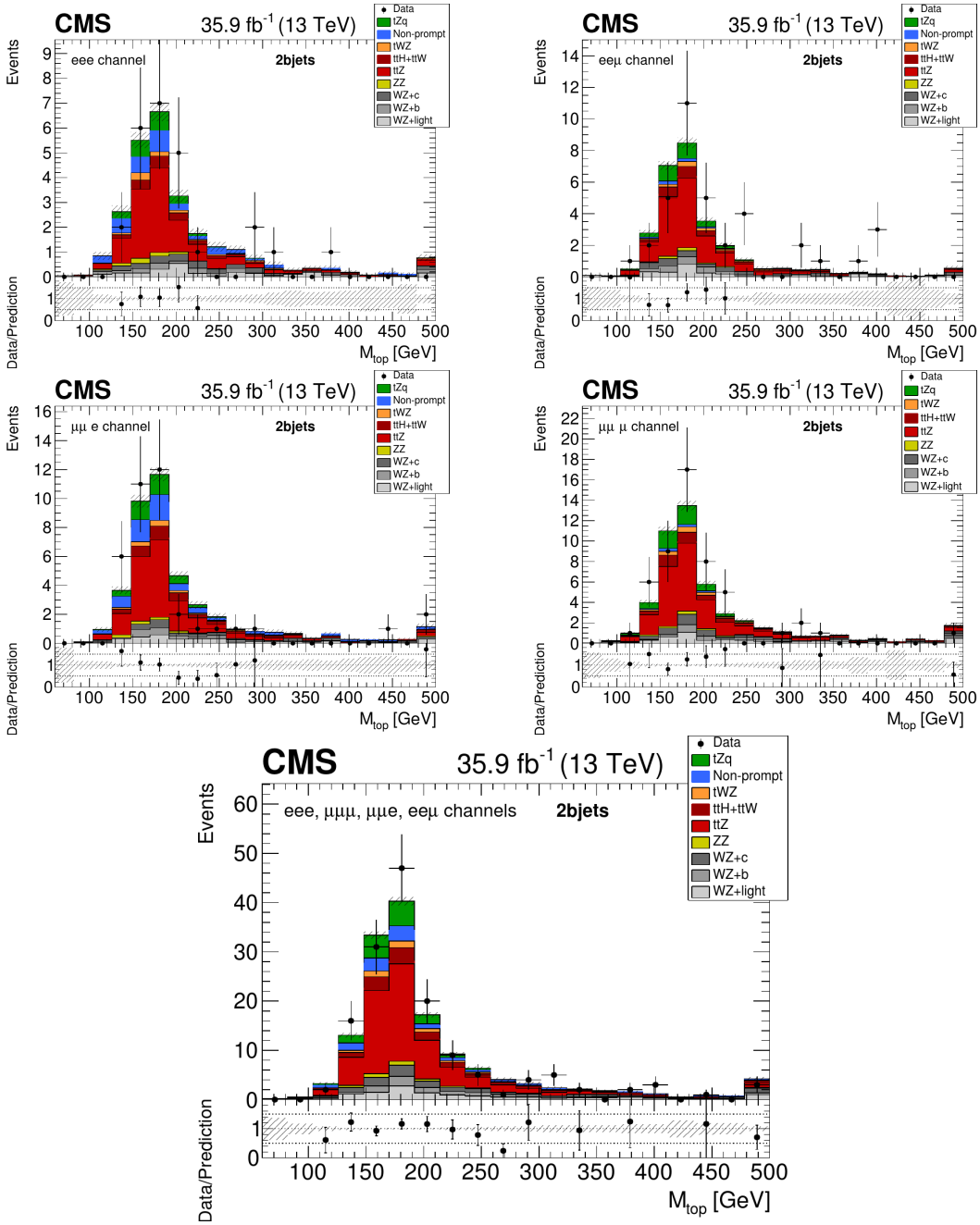


FIGURE B.26:)
Data-to-simulation comparison plots of the reconstructed top quark mass, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

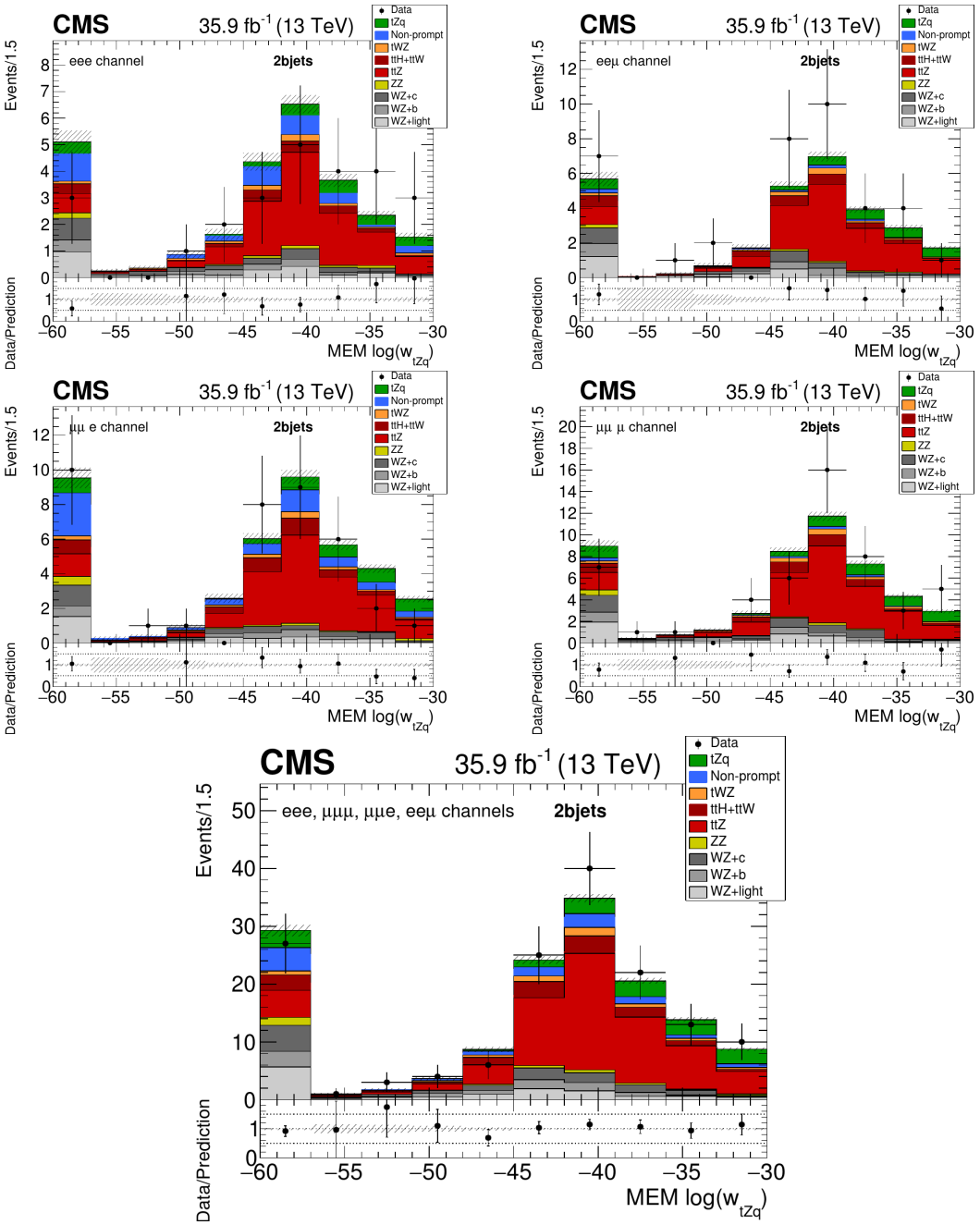


FIGURE B.27:)
Data-to-simulation comparison plots of the logarithm of the MEM score associated to the most probable tZq kinematic configuration, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

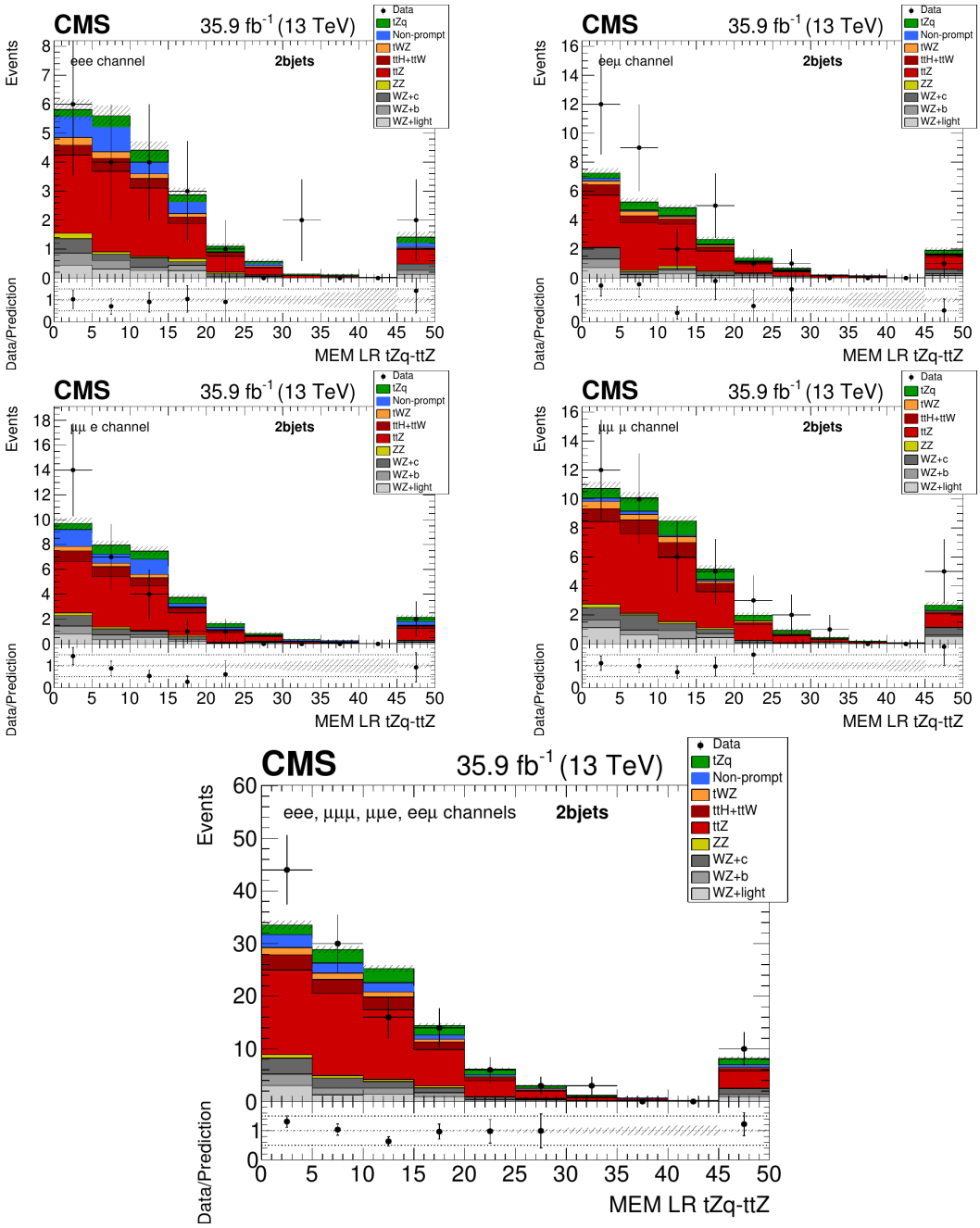


FIGURE B.28:)

Data-to-simulation comparison plots of the log-likelihood ratio of the tZq hypothesis agains the $t\bar{t}Z$ hypothesis (with $t\bar{t}Z$ and tZq weights rescaled such that their mean values are similar), for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

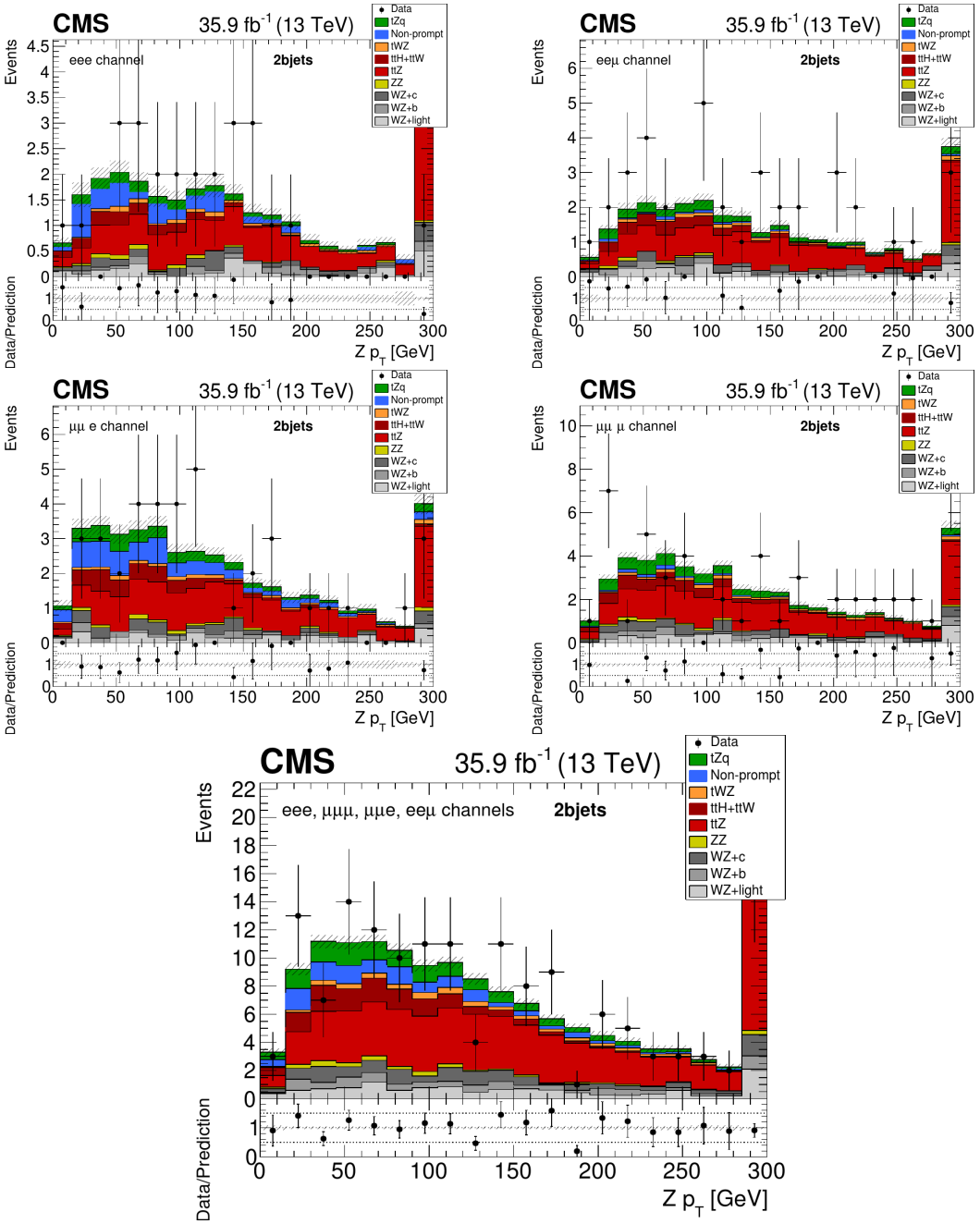


FIGURE B.29:)

Data-to-simulation comparison plots of the p_T of the reconstructed Z boson, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

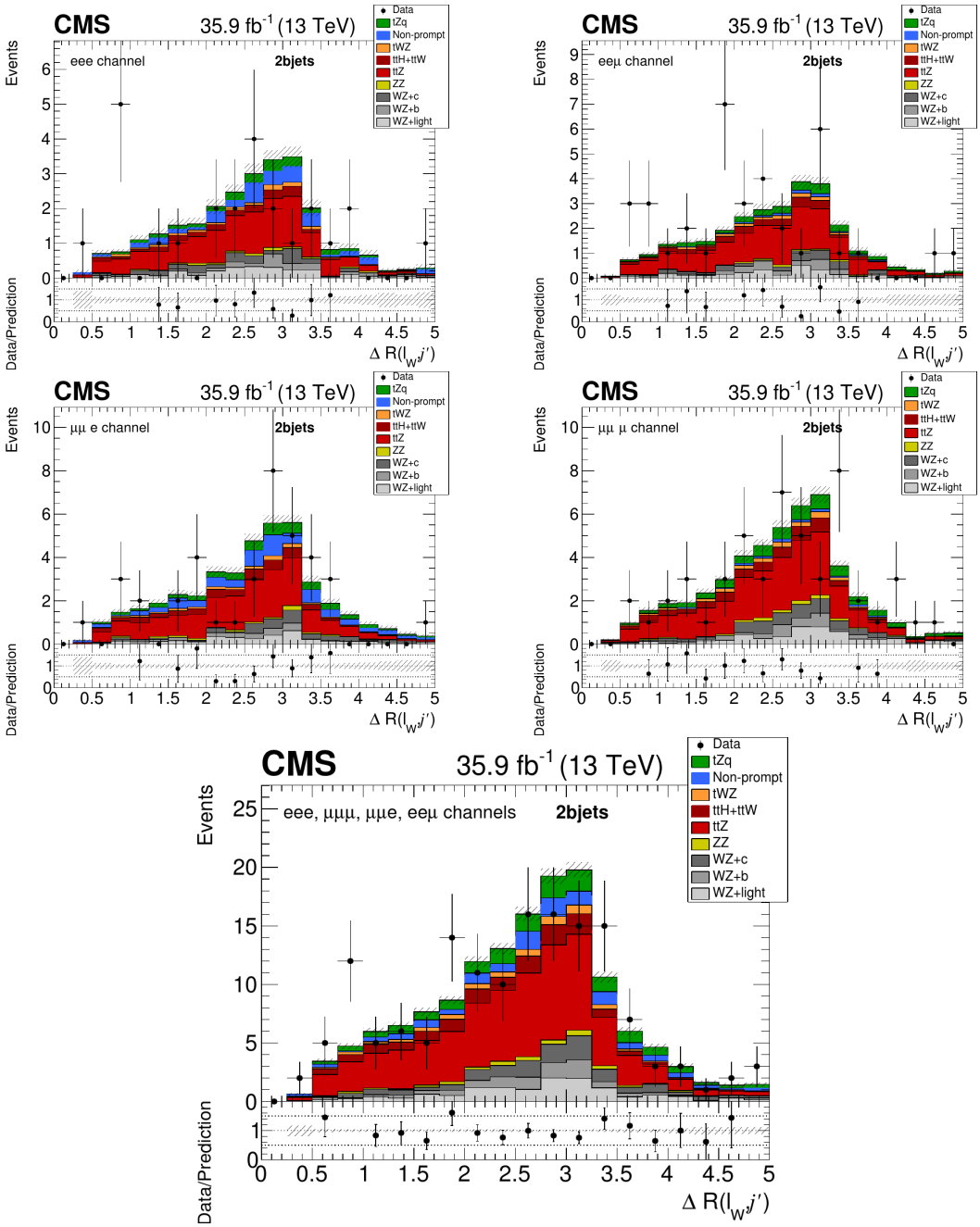


FIGURE B.30:)
Data-to-simulation comparison plots of the ΔR separation between the lepton associated to the top quark decay and the forward jet, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

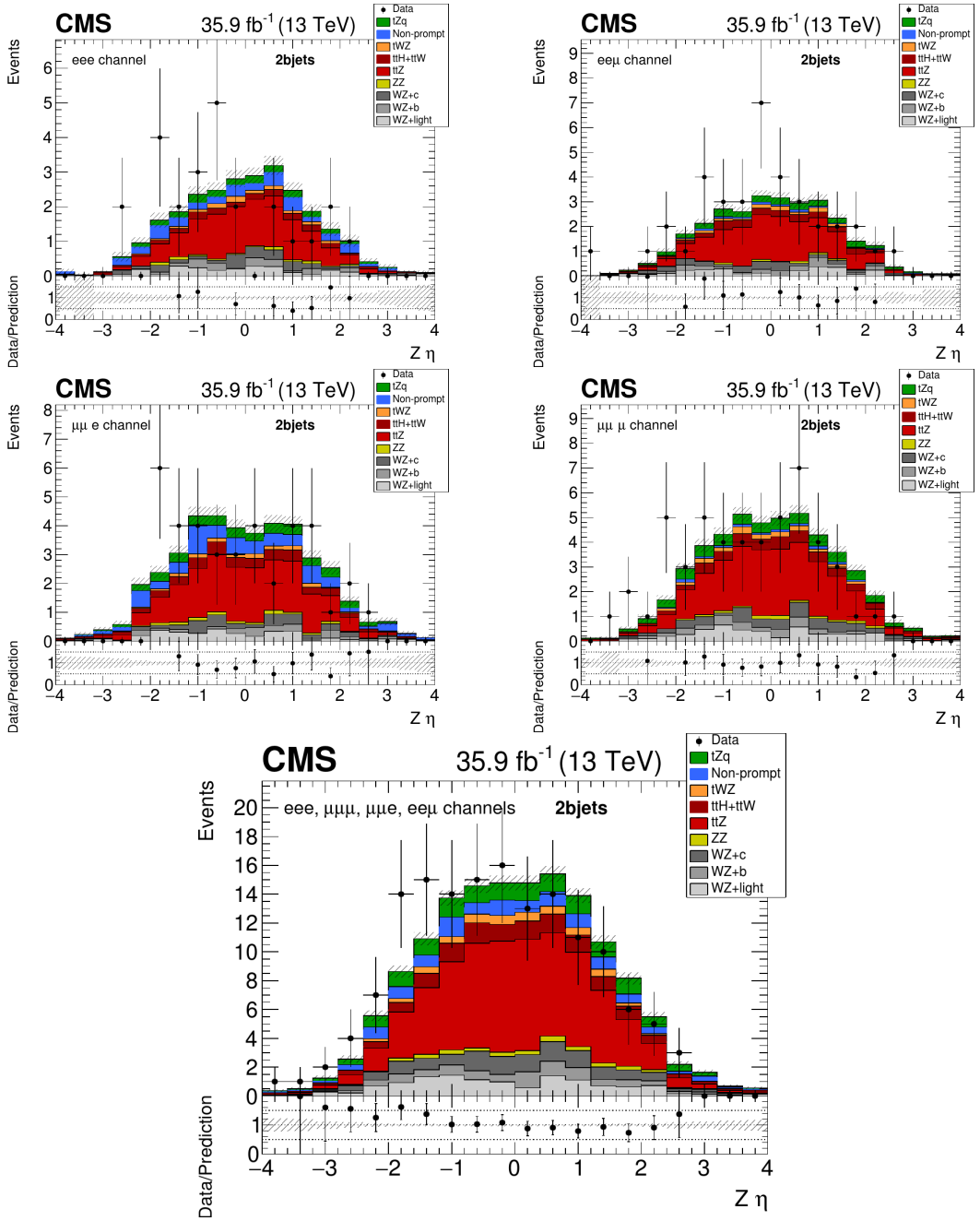


FIGURE B.31:)

Data-to-simulation comparison plots of the η distribution of the reconstructed Z boson, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

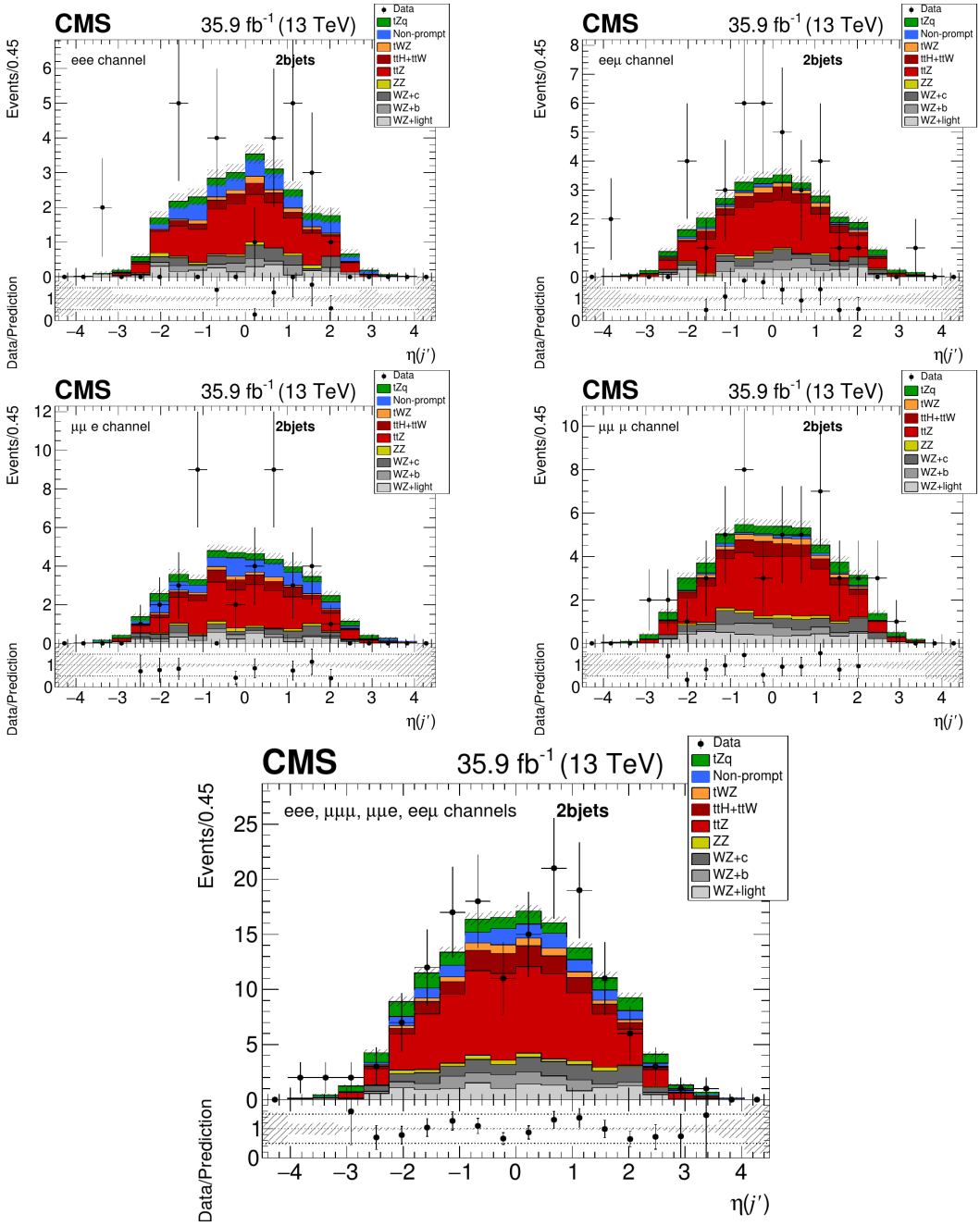


FIGURE B.32:)
Data-to-simulation comparison plots of the η distribution of the forward jet, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

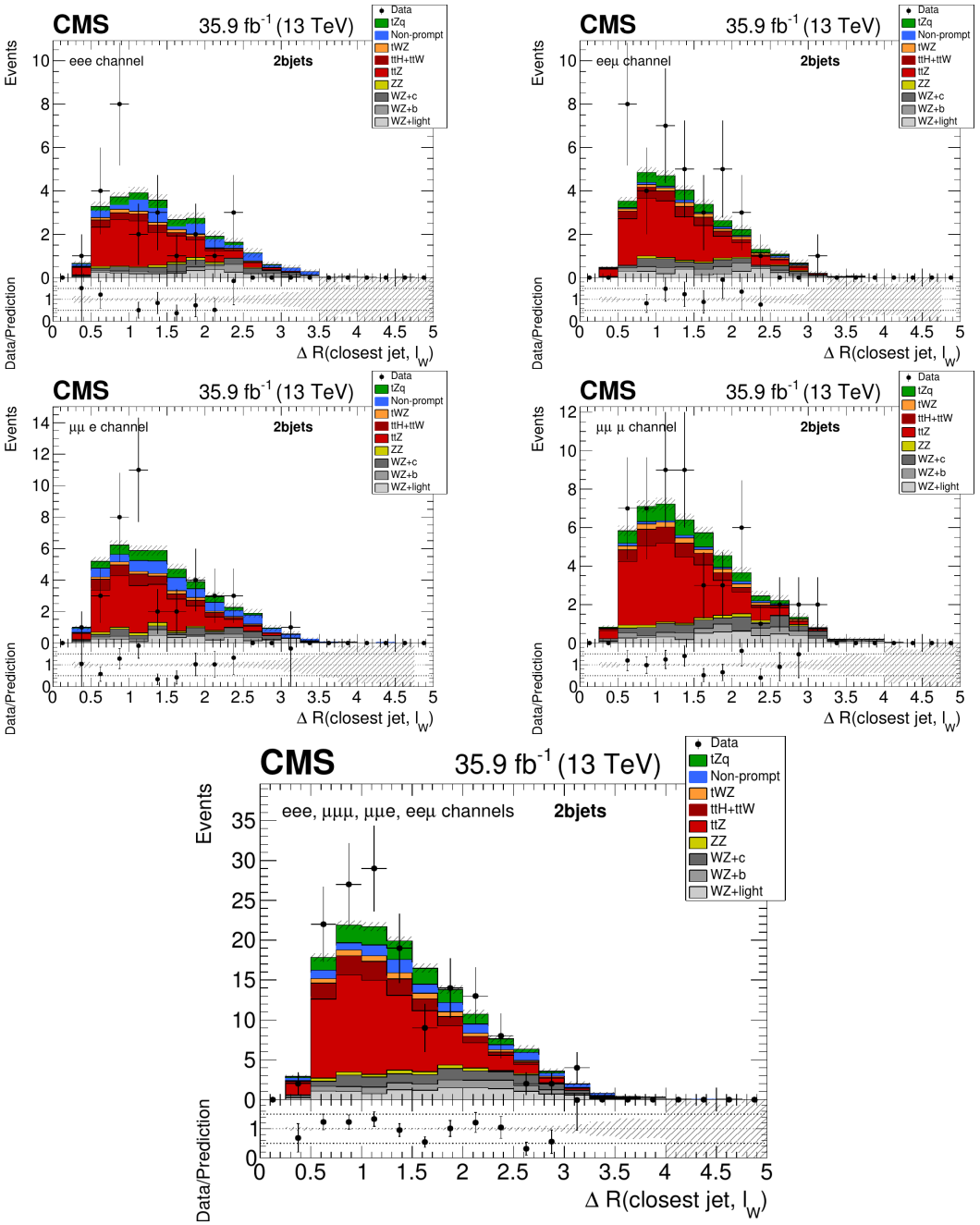


FIGURE B.33:)

Data-to-simulation comparison plots of the ΔR separation between the lepton associated to the top quark decay and its closest jet, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

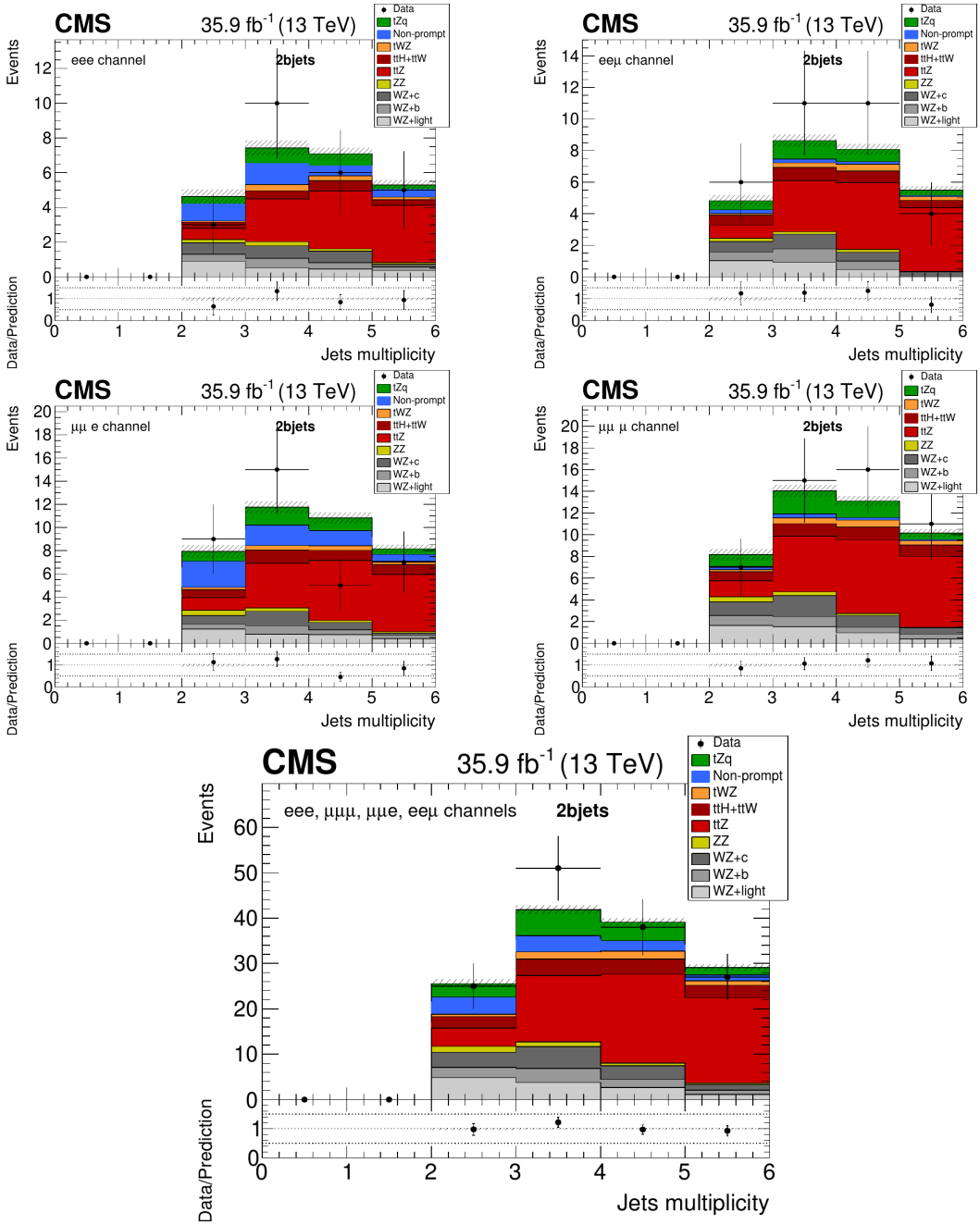


FIGURE B.34:)
Data-to-simulation comparison plots of the number of selected jets in the event, for the different decay channels in the 2bjet region. The last plot contains the distribution for all channels combined.

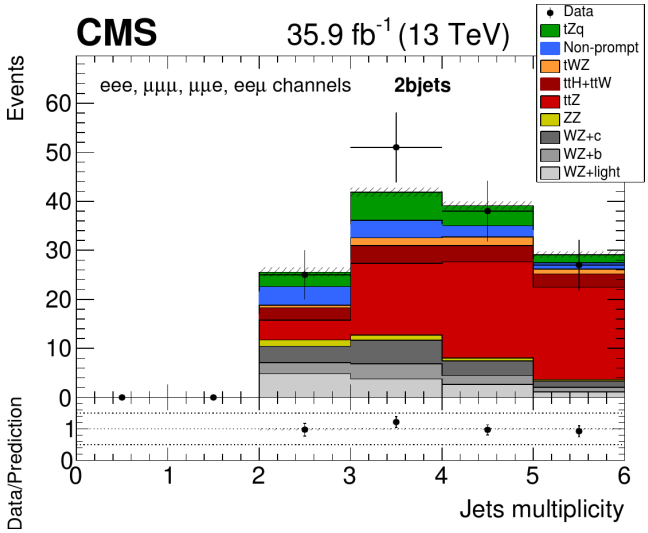


FIGURE B.35:)
Data-to-simulation comparison plots of the number of selected jets in the event, for the combination of all different decay channels in the 0bjet region

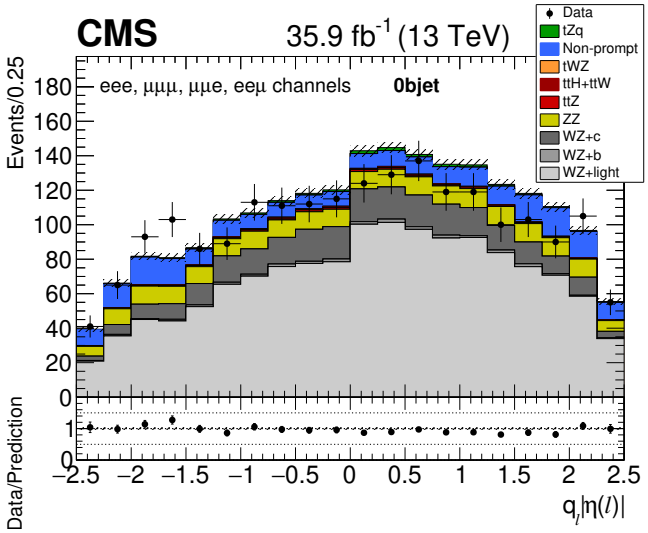


FIGURE B.36:)
Data-to-simulation comparison plot of the top quark decay lepton asymmetry for the combination of all different decay channels in the 0bjet region.

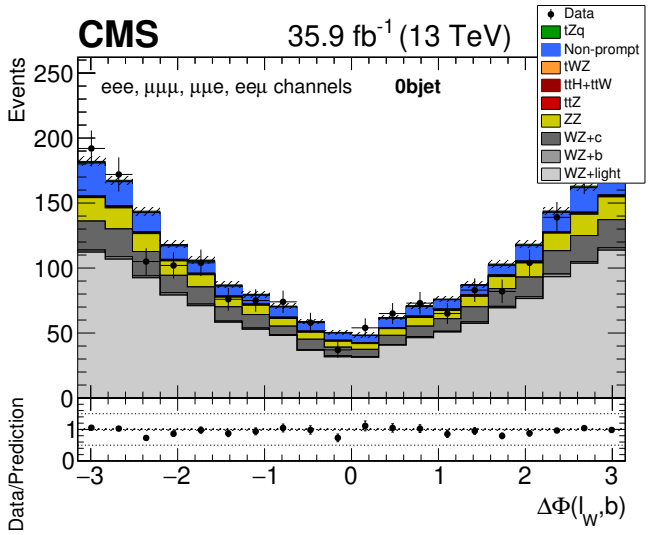


FIGURE B.37:)
Data-to-simulation comparison plot of the azimuth angle separation between the top quark decay lepton and the b quark for the combination of all different decay channels in the 0bjet region.

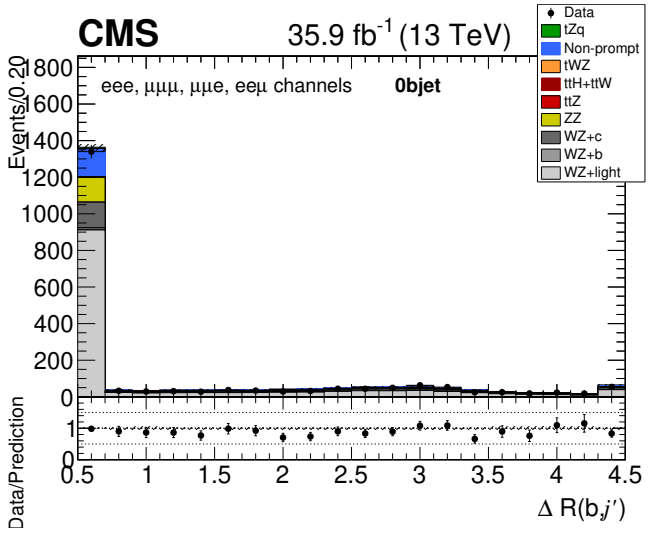


FIGURE B.38:)
Data-to-simulation comparison plot of the ΔR separation between the b jet and the recoiling jet for the combination of all different decay channels in the 0bjet region.

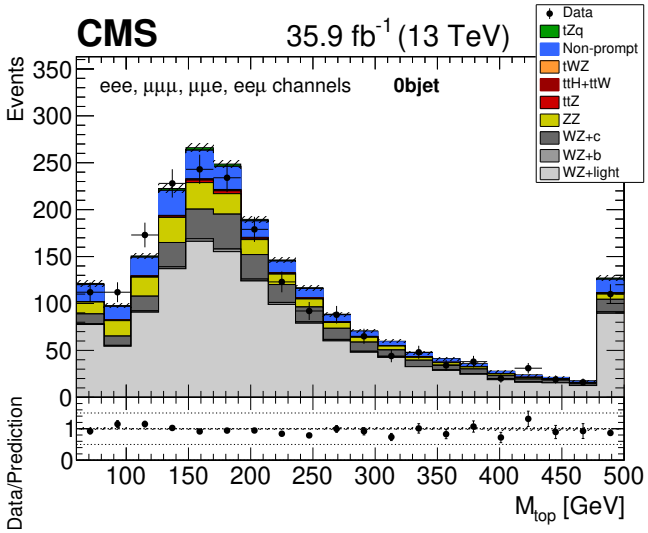


FIGURE B.39:)
Data-to-simulation comparison plot of the top quark mass for the combination of all different decay channels in the 0bjet region.

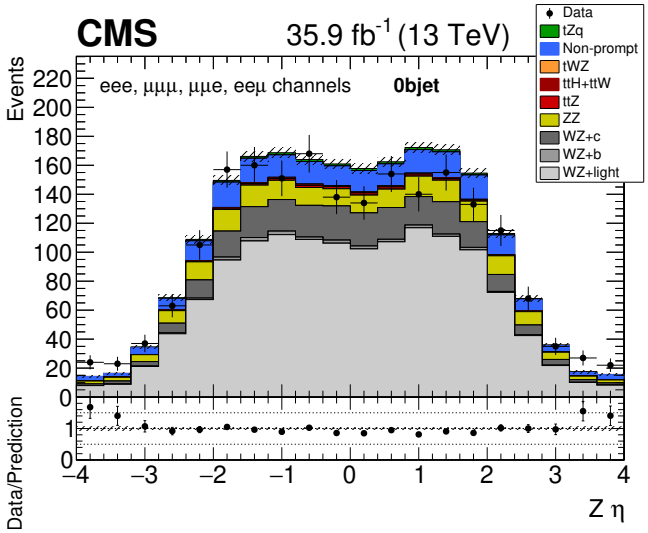


FIGURE B.40:)
Data-to-simulation comparison plot of the η of the Z boson for the combination of all different decay channels in the 0bjet region.

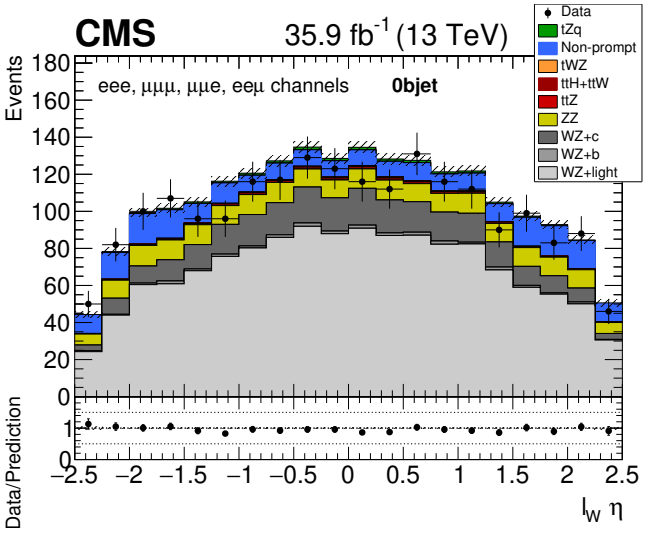


FIGURE B.41:)
Data-to-simulation comparison plot of the η of the top quark decay lepton for the combination of all different decay channels in the 0bjet region.

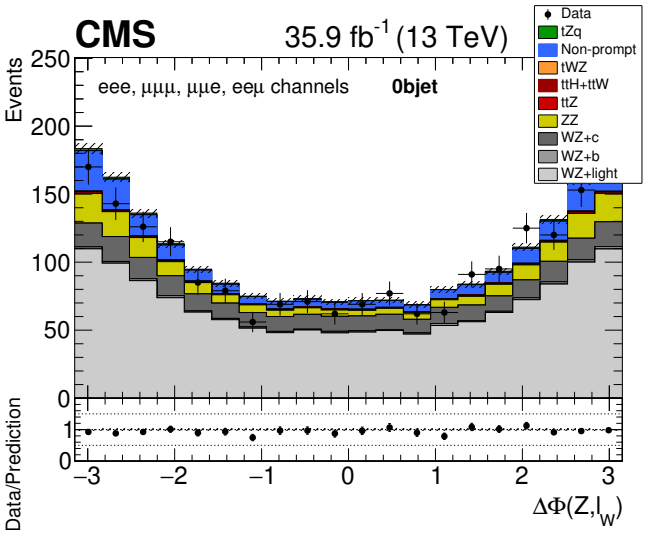


FIGURE B.42:)
Data-to-simulation comparison plot of the azimuthal separation between the top quark decay lepton and the Z boson for the combination of all different decay channels in the 0bjet region.

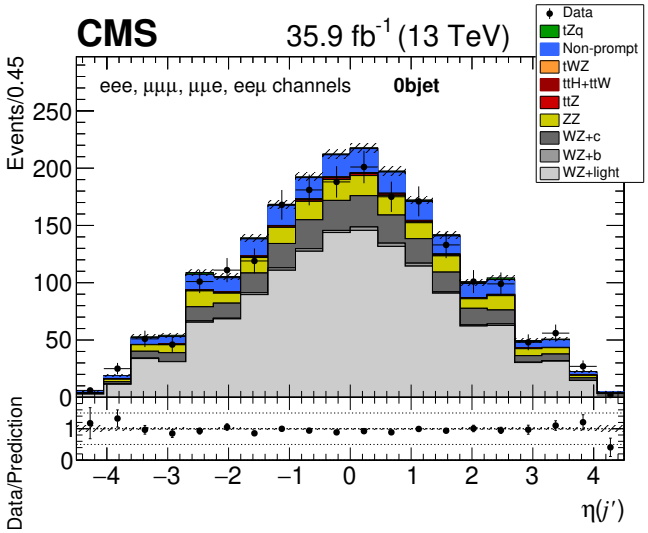


FIGURE B.43:)
Data-to-simulation comparison plot of the η of the recoiling jet for the combination of all different decay channels in the 0bjet region.

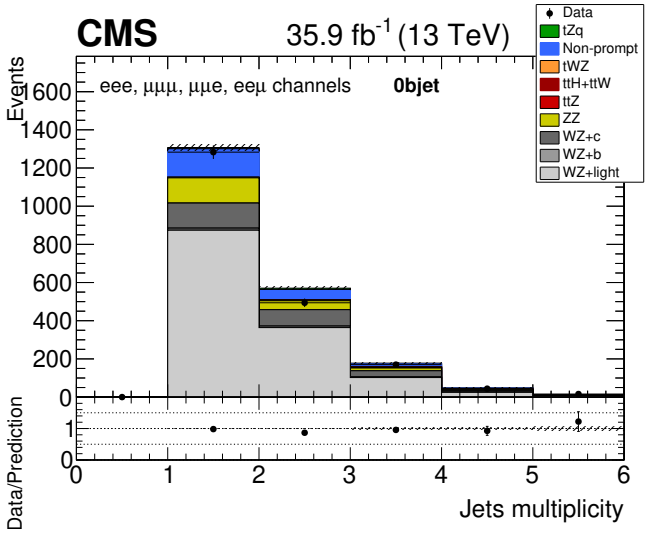


FIGURE B.44:)
Data-to-simulation comparison plot of the number of jets in the event for the combination of all different decay channels in the 0bjet region.

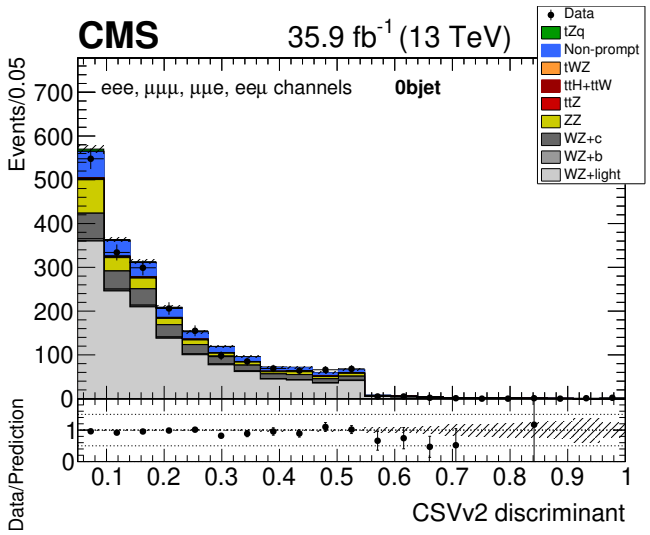


FIGURE B.45:)
Data-to-simulation comparison plot of the CSVv2 algorithm discriminant for the combination of all different decay channels in the 0bjet region.

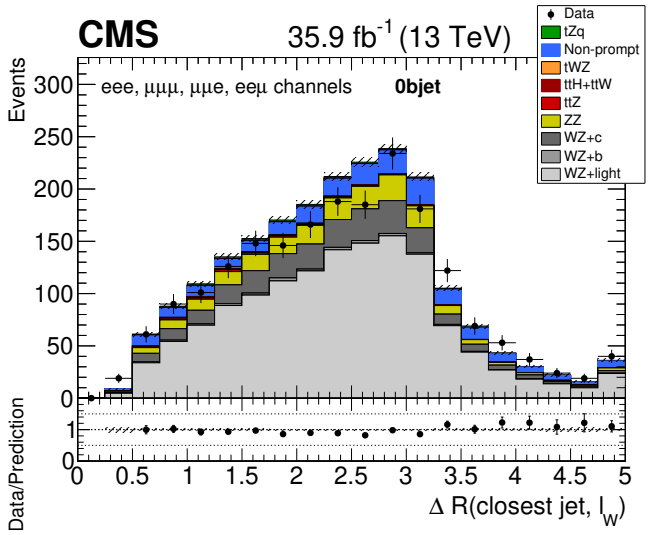


FIGURE B.46:)
Data-to-simulation comparison plot of the ΔR separation between the top quark decay lepton and the jet closest to it for the combination of all different decay channels in the 0bjet region.

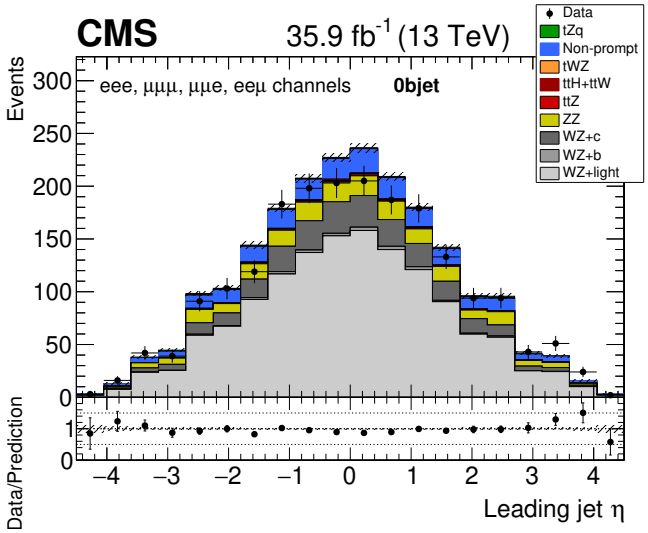


FIGURE B.47:)
Data-to-simulation comparison plot of the η of the leading jet for the combination of all different decay channels in the 0bjet region.

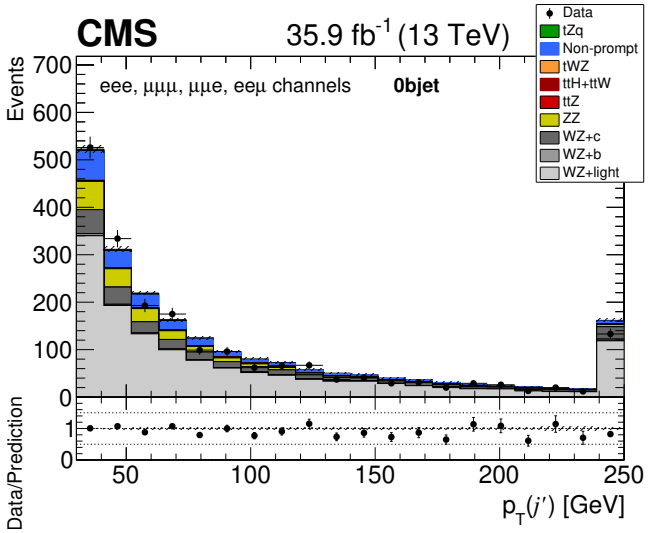


FIGURE B.48:)
Data-to-simulation comparison plot of the p_T of the recoiling jet for the combination of all different decay channels in the 0bjet region.

Appendix C

Prefit results

C.1 Prefit yields

Tables C.1, C.2 and C.3 show the signal and the different background processes yields. The expected yields from almost all backgrounds are predicted according to their standard model theoretical cross sections displayed in table 6.3. The normalization from the NPL background for electrons and muons is estimated from data as described in chapter 6, from a prefit to the m_T^W variable using data in the 0bjet region.

Process	eee	$e\mu\mu$	$\mu\mu\mu$	$ee\mu$	All channels
tZq	4.0	6.6	9.6	5.2	25.4
ttZ	3.9	6.4	8.4	5.0	23.6
ttW	0.3	0.7	0.6	0.3	1.9
ZZ	3.5	6.8	5.8	2.4	18.6
WZ+b	3.1	4.7	5.7	3.7	17.2
WZ+c	9.8	18.5	24.3	14.0	66.6
WZ+light	17.0	30.4	39.5	22.6	109.4
ttH	0.7	1.1	1.5	0.9	4.1
tWZ	1.0	1.6	2.4	1.2	6.3
Nonprompt electrons	16.7	15.5	0.0	0.6	32.8
Nonprompt muons	0.0	8.1	4.0	1.9	14.0
Total	60.0	100.5	101.7	57.8	319.9
Data	56	104	125	58	343

TABLE C.1: Observed and expected prefit yields for each production process in the 1bjet (signal) region. The yields of columns 2 to 5 correspond to each channel, and column 6 displays the total for all channels.

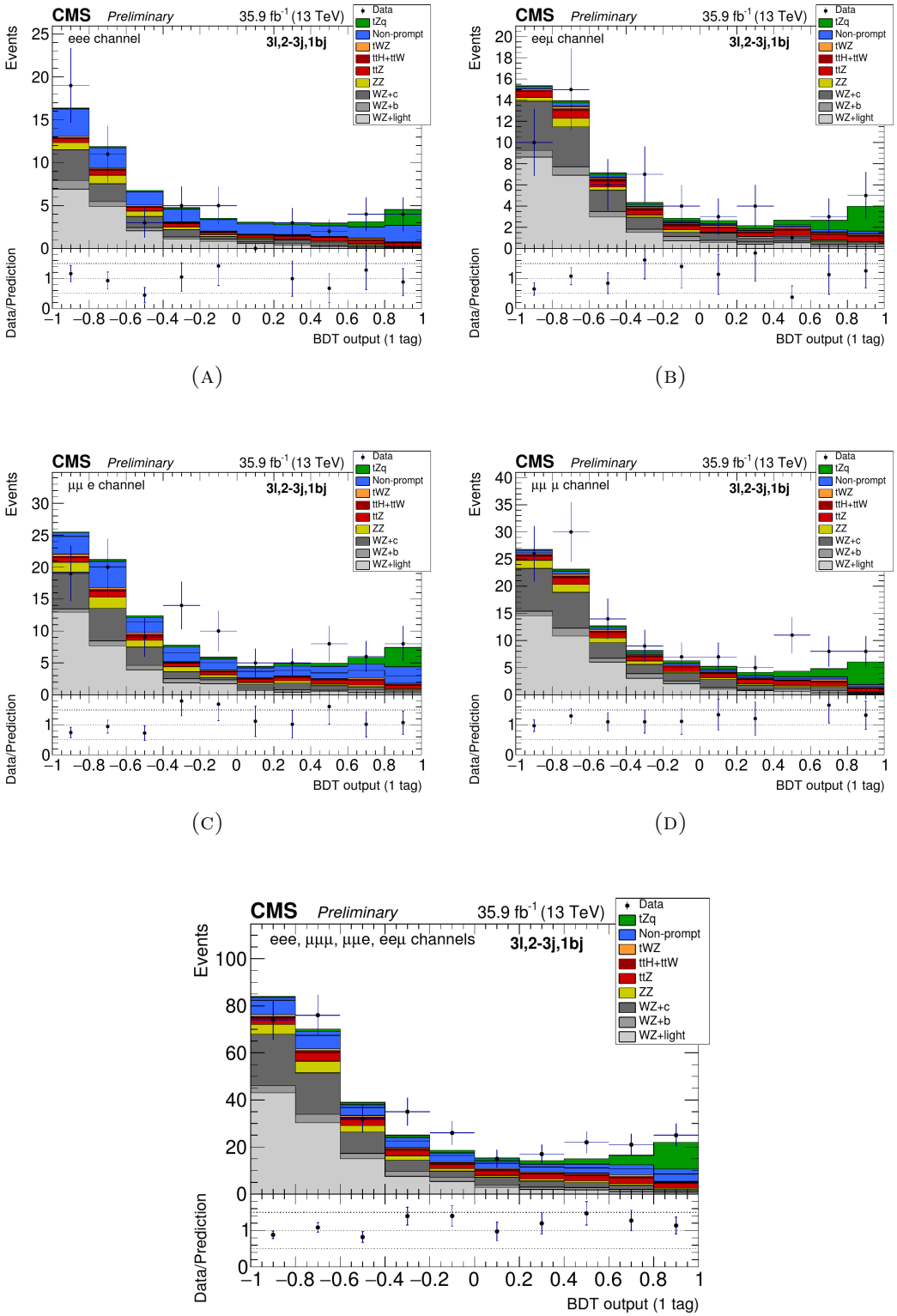
Process	*eee*	*eμμ*	*μμμ*	*eeμ*	All channels
tZq	2.3	3.0	4.1	5.5	15.0
$t\bar{t}Z$	11.1	13.9	17.0	22.4	64.5
$t\bar{t}W$	0.4	0.7	0.9	0.8	2.8
ZZ	0.6	0.6	1.0	1.1	3.3
WZ+b	1.6	2.1	1.9	2.9	8.4
WZ+c	2.3	2.4	3.0	4.7	12.4
WZ+light	2.3	2.5	3.2	4.6	12.6
ttH	1.5	2.1	3.1	3.9	10.6
tWZ	1.0	1.1	1.3	1.9	5.3
Nonprompt electrons	3.5	0.1	4.4	0.0	7.9
Nonprompt muons	0.0	0.7	2.0	1.0	3.7
Total	26.6	29.1	41.8	48.8	146.4
Data	25	38	51	37	151

TABLE C.2: Observed and expected prefit yields for each production process in the 2bjet ($t\bar{t}Z$ enriched) region. The yields of columns 2-5 correspond to each channel, and column 6 displays the total number of all channels summed.

Process	*eee*	*eμμ*	*μμμ*	*eeμ*	All channels
tZq	2.8	4.7	6.5	3.6	17.5
$t\bar{t}Z$	2.4	4.0	5.2	3.0	14.6
$t\bar{t}W$	0.1	0.4	0.4	0.2	1.1
ZZ	40.3	73.2	53.6	30.3	197.4
WZ+b	4.2	6.5	8.8	3.6	23.2
WZ+c	41.6	73.1	94.0	53.8	262.4
WZ+light	209.3	391.8	488.9	281.9	1371.8
ttH	0.3	0.5	0.7	0.4	1.9
tWZ	0.8	1.3	1.6	1.0	4.7
Nonprompt electrons	92.6	102.5	0.0	0.8	195.9
Nonprompt muons	0.0	10.3	11.2	7.6	29.2
Total	394.5	668.0	670.9	386.2	2119.7
Data	387	640	637	345	2009

TABLE C.3: Observed and expected prefit yields for each production process in the 0bjet (WZ+jets and fakes enriched) region. The yields of columns 2-5 correspond to each channel, and column 6 displays the total number of all channels summed.

C.2 Prefit templates

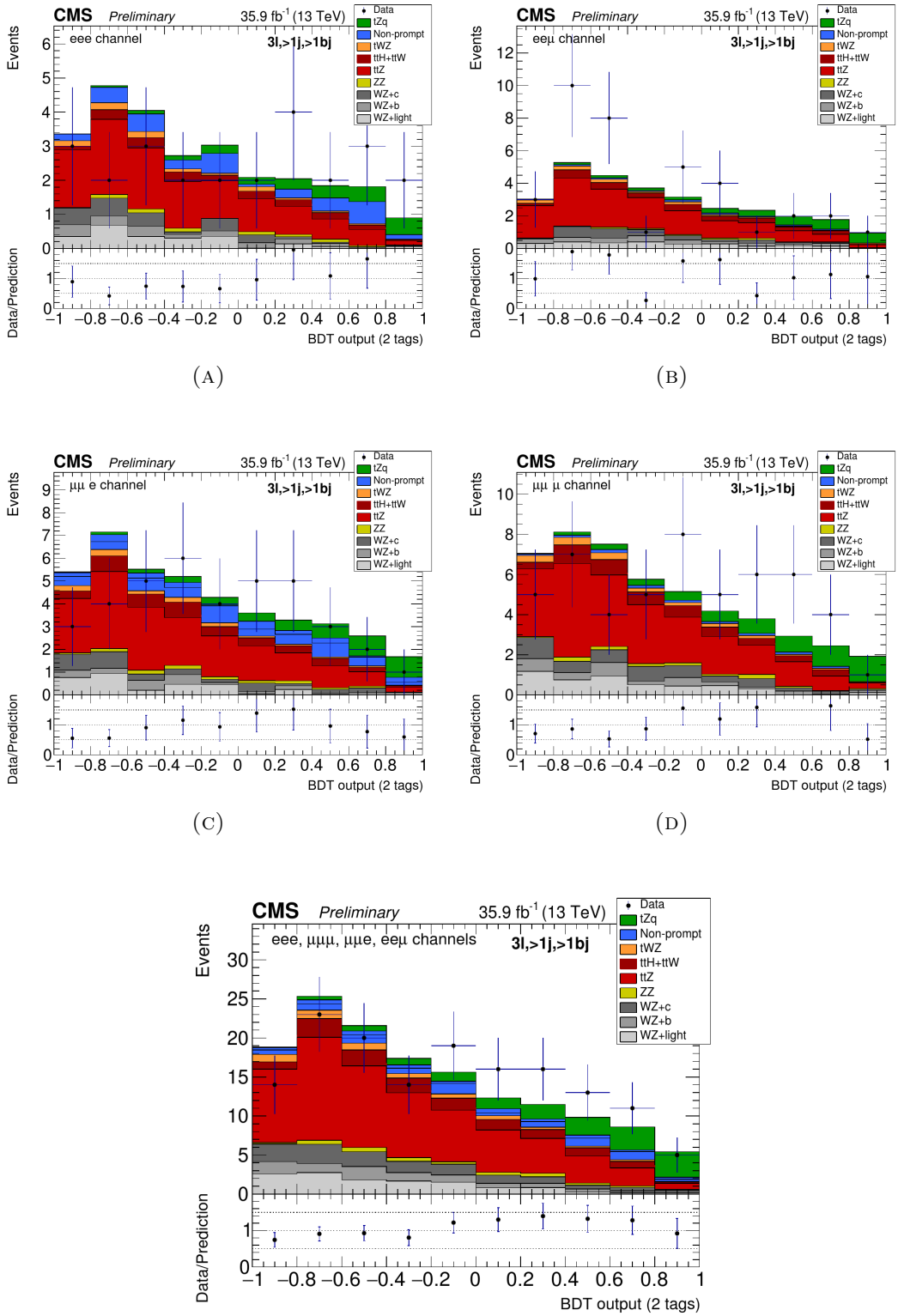


(A) (B)

(C) (D)

(E)

FIGURE C.1: Prefit data-to-prediction comparison plots for the BDT discriminant in the 1bjet (signal) region, computed for the four channels individually and with all of them summed (last plot).

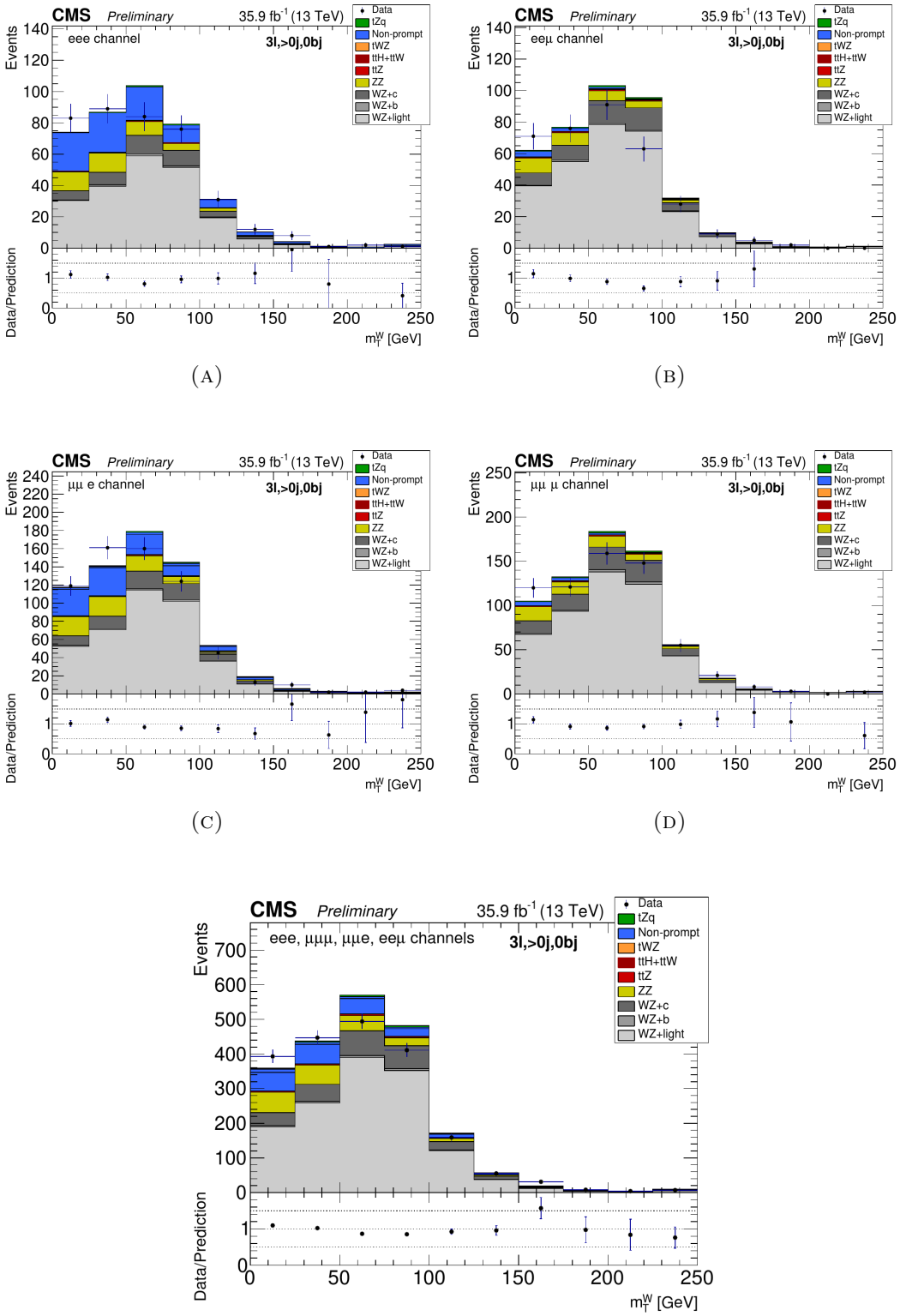


(A) (B)

(C) (D)

(E)

FIGURE C.2: Prefit data-to-prediction comparison plots for the BDT discriminant in the 2bjet ($t\bar{t}Z$ enriched) region, computed for the four channels individually and with all of them summed.



(A) (B) (C) (D) (E)

FIGURE C.3: Prefit data-to-prediction comparison plots for the m_T^W variable in the 0bjet region (where the contribution from WZ+jets and NPL backgrounds is most important), computed for the four channels individually and with all of them summed.

Bibliography

[1] **Gfitter** Collaboration, M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer, "The global electroweak fit at NNLO and prospects for the LHC and ILC," *Eur. Phys. J. C* **74** (2014) 3046, arXiv:1407.3792 [hep-ph].

[2] **Particle Data Group** Collaboration, M. Tabanashi *et al.*, "Review of particle physics," *Phys. Rev. D* **98** (2018) 030001.

[3] M. L. Perl *et al.*, "Evidence for anomalous lepton production in e^+e^- annihilation," *Phys. Rev. Lett.* **35** (1975) 1489–1492.

[4] **CDF** Collaboration, F. Abe *et al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.* **74** (1995) 2626–2631, arXiv:hep-ex/9503002 [hep-ex].

[5] **D0** Collaboration, S. Abachi *et al.*, "Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. Lett.* **74** (1995) 2422–2426, arXiv:hep-ex/9411001 [hep-ex].

[6] **D0** Collaboration, S. Abachi *et al.*, "Observation of the top quark," *Phys. Rev. Lett.* **74** (1995) 2632–2637, arXiv:hep-ex/9503003 [hep-ex].

[7] **CDF** Collaboration, T. Aaltonen *et al.*, "Observation of electroweak single top-quark production," *Phys. Rev. Lett.* **103** (2009) 092002, arXiv:0903.0885 [hep-ex].

[8] **D0** Collaboration, V. Abazov *et al.*, "Observation of single top-quark production," *Phys. Rev. Lett.* **103** (2009) 092001, arXiv:0903.0850 [hep-ex].

[9] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett. B* **716** no. 1, (2012) 30–61, arXiv:1207.7235 [hep-ex].

[10] **ATLAS** Collaboration, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B* **716** no. 1, (2012) 1–29, arXiv:1207.7214 [hep-ex].

[11] **CMS** Collaboration, A. M. Sirunyan *et al.*, "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," *JHEP* **07** (2017) 003, arXiv:1702.01404 [hep-ex].

[12] D. J. Griffiths, *Introduction to Elementary Particles*. TextBook Physics. Wiley, New York, NY, 1987.

[13] F. Halzen and A. D. Martin, *Quarks and leptons: An introductory course in modern particle physics*, vol. 53. Wiley, 1984.

[14] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.* **22** (1961) 579–588.

[15] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[16] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental Test of Parity Conservation in Beta Decay," *Phys. Rev.* **105** (1957) 1413–1414.

[17] T. D. Lee and C. N. Yang, "Question of parity conservation in weak interactions," *Phys. Rev.* **104** (1956) 254–258.

[18] **UA1** Collaboration, G. Arnison *et al.*, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$ = 540-GeV," *Phys. Lett. B* **122** (1983) 103–116.

[19] **UA2** Collaboration, M. Banner *et al.*, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN anti-pp collider," *Phys. Lett. B* **122** (1983) 476–485.

[20] **UA1** Collaboration, G. Arnison *et al.*, "Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider," *Phys. Lett. B* **126** (1983) 398–410.

[21] **UA2** Collaboration, P. Bagnaia *et al.*, "Evidence for $Z^0 \to e^+e^-$ at the CERN anti-pp collider," *Phys. Lett. B* **129** (1983) 130–140.

[22] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[23] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.* **10** (1963) 531–533.

[24] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of the top quark mass with lepton+jets final states using pp collisions at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J. C* **78** no. 11, (2018) 891, arXiv:1805.01428 [hep-ex].

[25] M. Jeżabek and J. Kühn, "QCD corrections to semileptonic decays of heavy quarks," *Nuclear Physics B* **314** no. 1, (1989) 1–6.

[26] J. R. Espinosa, "Implications of the top (and Higgs) mass for vacuum stability," *PoS* **TOP2015** (2016) 043, arXiv:1512.01222 [hep-ph].

[27] F. Bezrukov and M. Shaposhnikov, "Why should we care about the top quark Yukawa coupling?," *J. Exp. Theor. Phys.* **120** (2015) 335–343, arXiv:1411.1923 [hep-ph]. [Zh. Eksp. Teor. Fiz.147,389(2015)].

[28] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the t-channel single-top-quark production cross section and of the $| V_{tb} |$ CKM matrix element in pp collisions at $\sqrt{s}$= 8 TeV," *JHEP* **06** (2014) 090, arXiv:1403.7366 [hep-ex].

[29] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of $\mathrm{t\bar{t}}$ normalised multi-differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV, and simultaneous determination of the strong coupling strength, top quark pole mass, and parton distribution functions," *Eur. Phys. J. C* **80** no. 7, (2020) 658, arXiv:1904.05237 [hep-ex].

[30] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the mass difference between top quark and antiquark in pp collisions at $\sqrt{s} = 8$ TeV," *Phys. Lett. B* **770** (2017) 50–71, arXiv:1610.09551 [hep-ex].

[31] Q.-H. Cao, S.-L. Chen, and Y. Liu, "Probing Higgs Width and Top Quark Yukawa Coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ Productions," *Phys. Rev. D* **95** no. 5, (2017) 053004, arXiv:1602.01934 [hep-ph].

[32] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for Z' resonances decaying to $t\bar{t}$ in dilepton + jets final states in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. D* **87** no. 7, (2013) 072002, arXiv:1211.3338 [hep-ex].

[33] N. Kidonakis, "Charged Higgs production with a top quark at the LHC," in *Proceedings, 12th International Workshop on Deep Inelastic Scattering (DIS 2004): Strbske Pleso, Slovakia, April 14-18, 2004*.

[34] N. Kidonakis, "Charged Higgs production with a W boson or a top quark," in *2017 European Physical Society Conference on High Energy Physics (EPS-HEP 2017) Venice, Italy, July 5-12, 2017*.

[35] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV," *Eur. Phys. J. C* **80** no. 2, (2020) 75, arXiv:1908.06463 [hep-ex].

[36] **CMS** Collaboration, A. M. Sirunyan *et al.*, "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," *JHEP* **06** (2018) 102, arXiv:1712.02399 [hep-ex].

[37] **CMS** Collaboration, A. M. Sirunyan *et al.*, "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," *JHEP* **07** (2017) 003, arXiv:1702.01404 [hep-ex].

[38] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for anomalous single top quark production in association with a photon in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP* **04** (2016) 035, arXiv:1511.03951 [hep-ex].

[39] **ATLAS** Collaboration, G. Aad *et al.*, "Search for single top-quark production via flavour-changing neutral currents at 8 TeV with the ATLAS detector," *Eur. Phys. J. C* **76** no. 2, (2016) 55, arXiv:1509.00294 [hep-ex].

[40] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for top quark decays via Higgs-boson-mediated flavour-changing neutral currents in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP* **02** (2017) 079, arXiv:1610.04857 [hep-ex].

[41] J.-L. Agram, J. Andrea, E. Conte, B. Fuks, D. Gelé, and P. Lansonneur, "Probing top anomalous couplings at the LHC with trilepton signatures in the single top mode," *Phys. Lett. B* **725** (2013) 123–126, arXiv:1304.5551 [hep-ph].

[42] J. M. Yang, B.-L. Young, and X. Zhang, "Flavour-changing top quark decays in R-parity-violating supersymmetric models," *Phys. Rev. D* **58** (1998) 055001, arXiv:hep-ph/9705341 [hep-ph].

[43] G.-r. Lu, F.-r. Yin, X.-l. Wang, and L.-d. Wan, "The rare top quark decay $t \to cV$ in the top-colour-assisted technicolor model," *Phys. Rev. D* **68** (2003) 015002, arXiv:hep-ph/0303122 [hep-ph].

[44] J. A. Aguilar-Saavedra, "Effects of mixing with quark singlets," *Phys. Rev. D* **67** (2003) 035003, arXiv:hep-ph/0210112 [hep-ph]. [Erratum: Phys. Rev.D69,099901(2004)].

[45] J. A. Aguilar-Saavedra, "Top flavor-changing neutral interactions: Theoretical expectations and experimental detection," *Acta Phys. Polon. B.* **35** (2004) 2695, arXiv:hep-ph/0409342 [hep-ph].

[46] D. Atwood, L. Reina, and A. Soni, "Phenomenology of two Higgs doublet models with flavor-changing neutral currents," *Phys. Rev. D* **55** (1997) 3156–3176. https://link.aps.org/doi/10.1103/PhysRevD.55.3156.

[47] J. J. Cao, G. Eilam, M. Frank, K. Hikasa, G. L. Liu, I. Turan, and J. M. Yang, "SUSY-induced FCNC top-quark processes at the large hadron collider," *Phys. Rev. D* **75** (2007) 075021, arXiv:hep-ph/0702264 [hep-ph].

[48] G. Eilam, A. Gemintern, T. Han, J. Yang, and X. Zhang, "Top-quark rare decay $t \to ch$ in R-parity-violating SUSY," *Phys. Lett. B* **510** no. 1, (2001) 227 – 235, arXiv:hep-ph/0102037.

[49] K. Agashe, G. Perez, and A. Soni, "Collider Signals of Top Quark Flavor Violation from a Warped Extra Dimension," *Phys. Rev. D* **75** (2007) 015002, arXiv:hep-ph/0606293 [hep-ph].

[50] K. Agashe and R. Contino, "Composite Higgs-Mediated FCNC," *Phys. Rev. D* **80** (2009) 075016, arXiv:0906.1542 [hep-ph].

[51] L. Evans and P. Bryant, "LHC Machine," *JINST* **3** (2008) S08001.

[52] **ALICE** Collaboration, K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *JINST* **3** (2008) S08002.

[53] **ATLAS** Collaboration, G. Aad *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST* **3** (2008) S08003.

[54] **CMS** Collaboration, S. Chatrchyan *et al.*, "The CMS Experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[55] **LHCb** Collaboration, J. Alves *et al.*, "The LHCb Detector at the LHC," *JINST* **3** (2008) S08005.

[56] **LHCf** Collaboration, O. Adriani *et al.*, "The LHCf detector at the CERN Large Hadron Collider," *JINST* **3** (2008) S08006.

[57] **TOTEM** Collaboration, G. Anelli *et al.*, "The TOTEM experiment at the CERN Large Hadron Collider," *JINST* **3** (2008) S08007.

[58] **MoEDAL** Collaboration, J. Pinfold *et al.*, "Technical Design Report of the MoEDAL Experiment,". https://cds.cern.ch/record/1181486.

[59] **CMS** Collaboration, *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. http://cds.cern.ch/record/368412.

[60] **CMS Tracker Group** Collaboration, W. Adam *et al.*, "The CMS Phase-1 Pixel Detector Upgrade," *JINST* **16** no. 02, (2021) P02027, arXiv:2012.14304 [physics.ins-det].

[61] **CMS** Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/349375.

[62] **CMS** Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/357153.

[63] **CMS** Collaboration, *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. http://cds.cern.ch/record/343814.

[64] **CMS** Collaboration, V. Khachatryan *et al.*, "The CMS trigger system," *JINST* **12** no. 01, (2017) P01020, arXiv:1609.02366 [physics.ins-det].

[65] **CMS Trigger, Data Acquisition Group** Collaboration, W. Adam *et al.*, "The CMS high level trigger," *Eur. Phys. J. C* **46** no. 3, (2006) 605–667, arXiv:hep-ex/0512077.

[66] **CMS** Collaboration, D. Bonacorsi, "The CMS Computing Model," *Nucl. Phys. B Proc. Supp.* **172** (2007) 53 – 56.

[67] I. Antcheva *et al.*, "ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization," *Comput. Phys. Commun.* **180** (2009) 2499–2512, arXiv:1508.07749 [physics.data-an].

[68] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$ =13 tev," *JINST* **13** no. 10, P10005, arXiv:1809.02816 [hep-ex].

[69] **CMS** Collaboration, V. Khachatryan *et al.*, "Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS," *JINST* **11** no. 01, (2016) P01019, arXiv:1510.07488 [physics.ins-det].

[70] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV," *JINST* **13** no. 06, (2018) P06015, arXiv:1804.04528 [physics.ins-det].

[71] **CMS** Collaboration, S. Chatrchyan *et al.*, "Performance of CMS Muon Reconstruction in *pp* Collision Events at $\sqrt{s} = 7$ TeV," *JINST* **7** (2012) P10002, arXiv:1206.4071 [physics.ins-det].

[72] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc," *Journal of Physics G: Nuclear and Particle Physics* **31** no. 9, (2005) N9–N20.

[73] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois, "Electron reconstruction in CMS," *Eur. Phys. J. C* **49** (2007) 1099–1116.

[74] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Particle-flow reconstruction and global event description with the CMS detector," *JINST* **12** (2017) P10003, arXiv:1706.04965 [physics.ins-det].

[75] **CMS** Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," tech. rep., CERN, Geneva, Apr, 2009. https://cds.cern.ch/record/1194487.

[76] W. Adam, B. Mangano, T. Speer, and T. Todorov, "Track Reconstruction in the CMS tracker," tech. rep., CERN, Geneva, Dec, 2006. https://cds.cern.ch/record/934067.

[77] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP* **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[78] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV," *JINST* **13** no. 05, (2018) P05011, arXiv:1712.07158 [physics.ins-det].

[79] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector," *JINST* **14** no. 07, (2019) P07004, arXiv:1903.06078 [hep-ex].

[80] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in Summer 2011," *CMS-NOTE-2011-005; ATL-PHYS-PUB-2011-11* (2011) . https://cds.cern.ch/record/1379837.

[81] **ATLAS** Collaboration, G. Aad *et al.*, "Measurements of the top quark branching ratios into channels with leptons and quarks with the ATLAS detector," *Phys. Rev. D* **92** no. 7, (2015) 072005, arXiv:1506.05074 [hep-ex].

[82] J. Campbell, R. K. Ellis, and R. Röntsch, "Single top production in association with a Z boson at the LHC," *Phys. Rev. D* **87** (2013) 114006, arXiv:1302.3856 [hep-ph].

[83] 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," *JHEP* **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[84] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP* **11** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

[85] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP* **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[86] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP* **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[87] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," *Eur. Phys. J. C* **71** (2011) 1547, arXiv:1009.2450 [hep-ph].

[88] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions," *JHEP* **09** (2009) 111. [Erratum: JHEP02,011(2010)].

[89] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG BOX," *JHEP* **11** (2011) 078, arXiv:1107.5051 [hep-ph].

[90] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2" *Comput. Phys. Commun.* **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[91] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune," *Eur. Phys. J. C* **74** no. 8, (2014) 3024, arXiv:1404.5630 [hep-ph].

[92] **GEANT4** Collaboration, S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth. A* **506** (2003) 250–303.

[93] **CMS** Collaboration, S. Chatrchyan *et al.*, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in *pp* Collisions at $\sqrt{s} = 7$ TeV," *JINST* **8** (2013) P09009, arXiv:1306.2016 [hep-ex].

[94] L. Breiman, J. Friedman, C. Stone, and R. Olshen, *Classification and Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series.

[95] A. Hocker *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," arXiv:physics/0703039 [physics.data-an].

[96] **D0** Collaboration, V. M. Abazov *et al.*, "A precision measurement of the mass of the top quark," *Nature* **429** (2004) 638–642, arXiv:hep-ex/0406031 [hep-ex].

[97] **CDF** Collaboration, T. Aaltonen *et al.*, "Observation of single top quark production and measurement of $|V_{tb}|$ with CDF," *Phys. Rev. D* **82** (2010) 112005, arXiv:1004.1181 [hep-ex].

[98] K. Kondo, "Dynamical likelihood method for reconstruction of events with missing momentum. 1: Method and Toy Models," *J. Phys. Soc. Jap.* **57** (1988) 4126–4140.

[99] K. Kondo, "Dynamical likelihood method for reconstruction of events with missing momentum. 2: Mass spectra for 2→2 processes," *J. Phys. Soc. Jap.* **60** (1991) 836–844.

[100] R. H. Dalitz and G. R. Goldstein, "Decay and polarization properties of the top quark," *Phys. Rev. D* **45** (1992) 1531–1543.

[101] G. R. Goldstein, K. Sliwa, and R. H. Dalitz, "On observing top quark production at the Tevatron," *Phys. Rev. D* **47** (1993) 967–972, arXiv:hep-ph/9205246 [hep-ph].

[102] **CMS** Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period,". http://cds.cern.ch/record/2257069.

[103] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13$ TeV," *Phys. Lett. B* **766** (2017) 268–290, arXiv:1607.06943 [hep-ex].

[104] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV," *Phys. Lett. B* **779** (2018) 358–384, arXiv:1712.02825 [hep-ex].

[105] Prince Grover, "Gradient boosting from scratch: simplifying a complex algorithm." https://medium.com/mlreview/gradient-boosting-from-scratch-1e317ae4587d

Index