Universidad de Oviedo

Universidad de Oviedo

Estudio de procesos con dos leptones aislados en el experimento CMS como posible indicación de nueva física en colisiones protón-protón en el LHC.

Memoria de tesis presentada por

Santiago Folgueras Gómez

para optar al grado de Doctor

Dirigida por

Dr. Francisco Javier Cuevas Maestro

Física Fundamental y Aplicada

Oviedo, Noviembre de 2015

Universidad de Oviedo

Universidad de Oviedo

Study of processes with two isolated leptons in the CMS experiment as a possible indication of new physics in proton-proton collisions at LHC

Memoria de tesis presentada por

Santiago Folgueras Gómez

para optar al grado de Doctor

Dirigida por

Dr. Francisco Javier Cuevas Maestro

Física Fundamental y Aplicada

Oviedo, Noviembre de 2015

"Aprender no es saber, hay eruditos y hay sabios, la memoria hace a los primeros y la filosofía a los segundos."

Abate Faria, *El Conde de Monte Cristo*

"To learn is not to know; there are the learners and the learned. Memory makes the one, philosophy the others."

Abbe Faria, *The Count of Monte Cristo*

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma: ESTUDIO DE PROCESOS CON DOS LEPTONES AISLADOS EN EL EXPERIMENTO CMS COMO POSIBLE INDICACIÓN DE NUEVA FÍSICA EN COLISIONES PROTÓN-PROTÓN EN EL LHC	Inglés: STUDY OF PROCESSES WITH TWO ISOLATED LEPTONS IN THE CMS EXPERIMENT AS A POSSIBLE INDICATION OF PHYSICS IN PROTON-PROTON COLLISIONS AT LHC

2.- Autor	
Nombre: **SANTIAGO FOLGUERAS GÓMEZ**	DNI/Pasaporte/NIE:
Programa de Doctorado: FÍSICA FUNDAMENTAL Y APLICADA	
Órgano responsable: UNIVERSIDAD DE OVIEDO	

RESUMEN (en español)

Los resultados presentados en esta memoria de tesis usan el conjunto de datos completo recogido por el experimento CMS en colisiones protón-protón a energía en centro de masas de 7 y 8 TeV, así como los primeros datos proporcionados por el LHC a 13 TeV en el verano de 2015. En esta memoria se ha estudiado tanto el Modelo Estándar como escenarios de nueva física, en particular, la Supersimetría, que ha sido puesta a prueba usando todos los datos disponibles. Rara vez se encuentran procesos en el Modelo Estándar que produzcan dos leptones de la misma carga eléctrica en el estado final, sin embargo, aparecen de modo natural en diferentes escenarios de nueva física. Es por ello que, dicho estado final se ha utilizado para explorar dichos escenario y en particular, producción fuerte y electro-débil de partículas supersimétricas. Por otro lado, se han usado sucesos con un electrón y un muón de distinta carga eléctrica para medir con precisión la sección eficaz de producción de pares de quarks top-antitop. Utilizando un método similar se ha realizado la primera medida de dicha sección eficaz a una energía en centro de masas de 13 TeV.

En el desarrollo de la memoria, se presenta una descripción de los modelos utilizados, del experimento CMS y del acelerador LHC, junto con una descripción de la reconstrucción e identificación de observables que se usarán en los diferentes estudios presentados en este trabajo.

A continuación, el procedimiento que se ha llevado a cabo para estudiar los procesos con dos leptones de la misma carga eléctrica en el estado final y que pueden producirse en procesos derivados de modelos más allá del Modelo Estándar. Los resultados obtenidos se interpretan en términos de la compatibilidad con las predicciones del Modelo Estándar. Esto ha permitido excluir la presencia de partículas supersimétricas, en particular gluinos, con masas por debajo de 1050 GeV/c^2 (JHEP, 1401:163, 2014) . En este contexto, los datos se han utilizado para comprobar la posible producción electrodébil de partículas supersimétricas, cuyas secciones eficaces son sustancialmente más bajas que las que corresponden a la producción

fuerte de este tipo de partículas. Al no observar ninguna desviación con respecto a las predicciones del Modelo Estándar se ha establecido un límite en las masas de las partículas supersimétricas tipo neutralino por debajo de 600 GeV/c^2 (Eur.Phys.J., C74(9):3036, 2014).

El quark top es la partícula más pesada conocida y se acopla fuertemente al bosón de Higgs, por ello, cualquier proceso de nueva física puede mostrarse como una pequeña desviación en la sección eficaz de producción de dichas partículas. La producción de pares de quarks top-antitop es también el fondo dominante en las búsquedas descritas en este trabajo. Se ha determinado la sección eficaz de la producción de top-antitop usando la misma muestra de datos que el caso anterior en el estado final con un electrón y un muón. Para extraer el valor de la sección eficaz se ha usado un método de clasificación secuencial de los sucesos (CMS-PAS-TOP-13-004, enviado a publicación). Así mismo se ha realizado una primera medida de la sección eficaz de este proceso con los primeros datos recogidos a 13 TeV en 2015 obteniendo un resultado en acuerdo con las predicciones NNLO+NNLL de QCD para esta energía en centro de masas (enviada a PRL, CMS-PAS-TOP-15-003)..

En el último capítulo, se han explorado las posibilidades que ofrece el Run II del LHC a esta nueva energía en centro de masas y se han obtenido los primeros resultados de la posible extensión de dichos estudios con respecto a los realizados en el Run I del LHC.

RESUMEN (en Inglés)

This thesis presented here uses proton-proton collision data collected with the CMS experiment at a centre-of-mass energy of 7 and 8 TeV. The four different analysis presented in these pages have allowed to both test the Standard Model as well as explore new physics scenarios, such as SUSY, that has been pushed to its limits given the available luminosity and energy at the time this work has been developed. Same-sign dilepton events, a very rare signature in the SM, have been used to search for new physics and explore new territory, in particular, strong and electroweak production of supersymmetric particles has been probed using pp collision data at 8 TeV. Events with opposite-sign, different flavour lepton pairs have been used to accurately measure the top-antitop production cross-section using the full luminosity collected at 7 and 8 TeV. A first measurement at 13 TeV using the first data collected at the higher centre-of-mass has been also presented.

A general overview of the CMS experiment and the LHC is firstly presented, together with a brief description of the objects reconstruction that will be used in the several analysis covered in this work.

A first search for new physics beyond the Standard Model with same-sign dilepton events is described. The data are analysed using a set of exclusive search regions, targeting strong production of supersymmetry probing signatures both with and without

third generation squarks. No significant deviation from the Standard Model expectation is observed and limits on several new physics models are calculated. With the available luminosity and energy, gluinos with masses up to 1050 GeV/c^2 are probed. A second search using the same final state is designed focusing on the electroweak production of supersymmetric particles based on the same dataset. The smaller cross section motivates a strategy based on multiple dedicated analysis aiming at different regions of the phase space. Results are compatible with the Standard Model and allow to probe chargino and neutralino masses up to 600 GeV/c^2.

The top quark is the heaviest known particle and it couples heavily to the Higgs Boson and therefore any new physics phenomena may show up deviating the cross-section value from the one predicted by QCD. top-antitop production constitutes the dominant source of background for the search covered in this thesis. Thus, a measurement of the inclusive top-antitop production cross section in proton-proton collisions at LHC using data samples of 5.0 fb^{-1} at $\sqrt{s}$ = 8 TeV and 19.7 fb^{-1} at $\sqrt{s}$ = 8 TeV in the eμ final state is presented as part of this work. A cut-and-count technique is used to extract the cross section value. As a continuation of this analysis, a first measurement of the tt production cross section at 13 TeV is also presented without using b-tagging information. Results are in good agreement with recent NNLO QCD calculations.

Finally, having explored extensively the Standard Model together with many new physics models with supersymmetric particle production and having found no evidence of physics, a complete re-design of the new physics search described in this work seems necessary to improve the sensitivity of this search and maximise the reach at higher energies. Several developments are already presented in this document, and some of the described methods have been already tested and commissioned using the first data collected at 13 TeV. Other developments are still under study and they will be included in a publication describing the search using the data collected during 2015 at a centre-of-mass energy of 13 TeV.

SR. DIRECTOR DE DEPARTAMENTO DE FÍSICA/
SR. PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO EN FÍSICA FUNDAMENTAL Y APLICADA

Acknowledgements

Antes de nada, me gustaría agradecer a Javier todo lo que me ha ayudado a llevar la tesis a buen puerto con sus consejos, su dedicación y su pasión por la física. Han sido siempre una gran ayuda y una fuente de inspiración y admiración. No puedo describir la suerte que he tenido al tenerlo como director de tesis y sólo espero que dentro de poco tiempo vuelva a ser mi jefe. Él me dio la oportunidad de empezar a trabajar con él cuando aún no había acabado la carrera, ya entonces tenía la ilusión de trabajar en el CERN y a partir de ese momento comenzó una etapa muy emocionante en mi vida.

Muchas gracias también a todas aquellas compañeros y amigos que me han acompañado y guiado a lo largo de estos cuatro años, sin los cuales, seguramente no lo hubiera conseguido. De todos ellos he aprendido cosas sin las cuales no hubiera podido escribir esta tesis. En orden aleatorio, y pidiendo disculpas porque quizá se me olvide alguno, gracias a Javi, Lara, Chiqui, Filip, Boris, Marc, Benjamin, Chus, Alicia, Celso, Cristina, Ivan, Silvia, Giovanni, Marco, Matthieu, Constantin, Jan, Didar, Lesya, Hugues, Giuseppe, Frank, Carlo, Pablo, Javi B., Kike, María, Carmen, Jan, Jeremy, Emmanuele, Chiara, Jonatan, Rebeca, Bárbara, Keith, Claudio, Wolfgang, Rainer, Pieter, Ben y ...

Por último muchas gracias a mi familia, papá, mamá, tito, tita y María, porque siempre me habéis apoyado y ayudado. Gracias a Teresa, por estar siempre ahí. Y gracias a mis amigos y a mi nueva familia, sin ellos tampoco lo habría conseguido: Parra, Elisa, Lolo, Chus, Andrés, Pocket, Marcos y Pochi.

Contents

List of Figures

List of Tables

A Teresa

Chapter 1

Introduction

What is the world made of and what binds it together? The answer to this question has been sought by philosophers and scientists for centuries.

One of the most successful answer to this question if the Standard Model of particle physics, it describes all the known elementary particles and their interactions. It has been able to explain remarkably well the experimental data over the last 40 years. However it may not be the end of the story.

The Large Hadron Collider (LHC) and its experiments constitute a unique facility for testing the Standard Model and beyond. Since its start back in 2010 at a centre-of-mass energy of 7 TeV, the LHC has performed remarkably well and delivered a proton-proton collision dataset of about 25fb^{-1} up to 2012. CMS is a general-purpose experiment designed and built to study high-energy collisions. The LHC is designed to be a discovery machine and there are many reasons to believe that new physics resides within its energy reach. This thesis probes some of the new physics phenomena which may be present at the TeV-scale.

The search described in this thesis uses the full 2012 dataset collected by CMS at a centre-of-mass energy of 8 TeV, as well as the first data collected in 2015 at 13 TeV. Events containing two leptons (electrons and muons) with the same electric charge are selected. This signature offers a prime search channel for new physics as the expected background from the Standard Model is very small and many new physics scenarios predict an enhanced rate of such events. The search is designed to remain as model-independent as possible.

Two different search strategies were followed in the search described in this work for the 8 TeV dataset. First, a search including jets in the final state was performed as the data was arriving and multiple public results were released and presented in conference,

once the full luminosity was collected, the search was refined and published. Secondly, a more dedicated search was performed targeting electroweak production of supersymmetric particles. This second search was combined with dedicated multi lepton searches performed by CMS. This combination has also been published.

Having found no evidence of new physics, understanding the Standard Model and measuring precisely any potential deviation, provides an indirect search for new physics. The top quark is the heaviest known particle and it couples heavily to the Higgs Boson and therefore any new physics phenomena may show up deviating the cross-section value from the one predicted by QCD. Also, $t\bar{t}$ production constitutes the dominant source of background for the search covered in this thesis and thus a precise determination of the $t\bar{t}$ production cross-section is presented in this work.

In 2015, the LHC has restarted and the centre-of-mass energy has been increased to 13 TeV, opening up a great new scenario for new physics searches as higher particle masses could be reached. A first measurement of the $t\bar{t}$ production was performed using the first data collected. Using the same dataset, MC simulation and background estimation methods for the same-sign search were commissioned.

This work is structured as follows: Chapter 2 describes briefly the theoretical background. Chapter 3 describes the experimental setup, both the LHC and CMS that was used for taking the data used in this thesis. Chapter 4 covers the event reconstruction and a description of the observables that can be measured in CMS. In Chapter 5 the same-sign analysis techniques used at 8 TeV are described, while the search strategy and results are shown in Chapters 6 for the search targeting strong production of SUSY and in Chapter 7 for the search targeting electroweak production. To conclude the 8 TeV part, a precise determination of the $t\bar{t}$ production cross-section is described in 8 using full dataset at 7 and 8 TeV. A first measurement of the $t\bar{t}$ production cross-section is also described in this chapter. Chapter 9 describes several new developments for the same-sign search at 13 TeV search and the commissioning of some of this new methods using the first data. Finally the conclusions of this work are drawn in Chapter 10.

Conventionally, the energy in particle physics is not measured in Joules, but rather "electronvolts", eV. It is defined as the energy gathered by a unit charge in an electric potential difference of 1 Volt. The LHC physics deals with energies in the range of GeV to TeV. In addition to this special unit of energy, particle physicists use so called natural units, in which the Planck constant and the speed of light are set to unity, i.e. $\bar{h} = c = 1$.

Chapter 2

Motivation

Particles Physics is a branch of physics that studies the elementary constituents of matter and radiation and the interactions between them. It is also called High Energy Physics, because many elementary particles do not appear under normal circumstances in nature, but can be created and detected during energetic collisions of other particles, as is done in particle accelerators. The development of Quantum Mechanics at the beginning of the 20th century, Quantum Field Theory some decades later and finally, the Standard Model of particle physics, formulated in the sixties, represent major breakthroughs in our way to understand matters constituents and their interactions. After about 30 years of extensive testing, the Standard Model is one of the best theories of modern physics and it has been tested up to a very high precision over a large range of energies and an agreement at the per mil level with the theoretical predictions has been found. The Standard Model is briefly described in this chapter, together with its limitations and resulting need for beyond the Standard Model physics.

2.1 The Standard Model of Particle Physics

The Standard Model (SM), which describes elementary particles and their interactions, is a successful theory experimentally tested to high precision over the past decades. However, many unresolved questions suggest it cannot be the ultimate theory of nature. A wide variety of new physics models have been developed over the year to address some of the limitations of the Standard Model. Some of these models predict final states with same-sign dileptons as an experimental signature, which forms the basis of the search described in this dissertation.

Elementary particles in the SM are divided into two categories: spin-1/2 matter particles (*leptons* and *quarks*) and spin-1 force carriers (*gauge bosons*). Among the leptons there

are three charged leptons: electron (e), muon (μ) and tau (τ) and another three, nearly massless, neutrinos (ν_e, ν_μ, ν_τ). Quarks can be categorised, as well, based on their electric charge as they can be up-type ($Q = +2/3$) and down-type ($Q = -1/3$). The up-type quarks include the up (u), charm (c) and top (t) quarks while the down-type include the down (d), strange (s) and bottom (b) quarks. All matter particles are organised into three generations as ($e\,\nu_e\,u\,d$), ($\mu\,\nu_\mu\,s\,c$) and ($\tau\,\nu_\tau\,b\,t$). The only distinction between the three generation is the increasingly heavier masses.

The second category of fundamental particles, the gauge bosons, are the force carriers. The neutral and massless photon (γ) is the carrier of the electro-magnetic force. The weak interaction which is carried by the heavy $W^\pm$ and Z^0 bosons. As the W bosons carry electrical charge, they interact with each other, as well as with the Z and γ. At higher energies the electro-magnetic and weak forces become indistinguishable and we speak of the "electroweak" force. Understanding this transformation, referred to as Electroweak Symmetry Breaking (EWSB), is a major challenge in high energy physics today. Eight massless gluons (g) which are electrically neutral but carry color charge are responsible for the strong interaction. Quarks have color charge, allowing them to interact through the strong interaction in addition to the electromagnetic and weak interaction.

The last particle in the Standard Model is the Higgs boson, a neutral scalar particle which is necessary to give mass to the heavy gauge bosons and the rest of the particles. The Higgs boson has been discovered recently by the ATLAS and CMS Collaborations in July 2012 with a mass of 125.09 GeV [1–3].

The particle content of the Standard Model is summarised in Table 2.1 In addition to the listed particles, charged leptons, quarks, and the charged gauge bosons have associated antiparticles with opposite electric charge, nearly doubling the particle content. The neutral gauge bosons are their own antiparticles.

2.1.1 Theoretical Formulation

This section gives a summary of the most important concepts in theoretical formulation of the Standard Model a more wide description can be found here [5].

The Standard Model is formulated as a quantum field theory whose interactions are described as the result of local symmetries. Enforcing invariance of a Lagrangian under local symmetries necessarily leads to the introduction of covariant derivatives that contain additional fields, identified with the interaction bosons from above. The three

TABLE 2.1: Summary of the particle content of the Standard Model, including their mass, spin, and electric charge [4].

Type	Name	Mass	Charge [e]
LEPTONS (spin=1/2)	electron (e)	0.511 MeV	-1
	electron neutrino (ν_e)	< 2 eV	0
	muon (μ)	106 MeV	-1
	muon neutrino (ν_μ)	<0.19 MeV	0
	tau (τ)	1.78 GeV	-1
	tau neutrino (ν_τ)	<18.2 MeV	0
QUARKS (spin=1/2)	up (u)	2.3 MeV	+2/3
	down (d)	4.8 MeV	-1/3
	charm (c)	1.28 GeV	+2/3
	strange (s)	95 MeV	-1/3
	top (t)	174 GeV	+2/3
	bottom (b)	4.18 GeV	-1/3
GAUGE BOSONS (spin=1)	photon (γ)	0	0
	$W^{\pm}$	80.4 MeV	± 1
	Z^0	91.2 GeV	0
	gluons (g)	0	0
HIGGS BOSON (spin=0)	Higgs (H)	125.09 GeV	0

fundamental interactions, excluding gravity, contained in the Standard Model are introduced as a local gauge symmetry of the group $SU(3)_C \times SU(2)_L \times U(1)_Y$.

As any relativistic quantum field theory it follows Poincare's symmetry including translations in time and space, rotations in space, and transformations connecting two uniformly moving systems (Lorentz boosts), which leads to the conservation of energy and momentum and the dynamics of special relativity.

The Strong Interaction

The strong interaction between quarks and gluons is characterised by *Quantum Chromodynamics* (QCD). Is described by an $SU(3)_C$ gauge theory rising three color charges (blue, green and red) and eight independent gluon fields which maintain the local color invariance. Gluons carry themselves color charge and therefore interact with quarks as well as with other gluons.

The strength of its coupling constant (α_s) depends with the energy scale of the interaction, or to be more precise, on the momentum transfer Q involved in the process:

$$\alpha_s(Q^2) \approx \frac{1}{\ln(Q^2/\Lambda)} \tag{2.1}$$

where Λ is the non-perturbative scale of QCD.

A physical implication of this energy dependence is what is called "asymptotic freedom" and "confinement". The former entails the decreased interaction strength with increasing energy, and therefore quarks and gluons can effectively be treated as free particles at very high energies. The latter implies that color charges exhibit anti-screening and the strength of the interaction increases with distance from the charge. Meaning that it is impossible to have free quarks or gluons as it will always be energetically preferable to create new color charges from the vacuum to make them color-neutral. Hence, quarks only show up in color neutral states, either in groups of three (baryons) or as pairs of quarks and anti-quarks (mesons). At the high energies of the LHC, quarks and gluons can be treated as free particles. However, due to confinement, they will not appear as free particles in the detector. Instead, they appear as collections of stable hadrons, known as "jets". The process of forming hadrons from the initial quarks and gluons is called hadronization.

What further distinguishes Quantum Chromodynamics (QCD), in a purely calculative sense, from other field theories are the relatively high coupling constants that are rising with falling energy scales of the interaction. This implies that the perturbation theory, breaks down at low energies and we are unable to use it to calculate physical quantities even though we seem to know the underlying theory very well. In practice, it means that we are able to calculate to a certain degree of precision the hard interaction in a particle collision, but are forced to employ empirical hadronization models to describe the further evolution from intermediately free color charges to bound and stable states of mesons and baryons.

The Electroweak Interaction

The electromagnetic interaction, which acts on all charged particles and it is mediated by the photon, is described as a quantum field theory subjected to a local $U(1)$ gauge symmetry, resulting in the highly successful theory of quantum electrodynamics (QED). The weak interaction is mediated by the $W^{\pm}$ and Z^0 bosons and acts between quarks and leptons and it allows flavour-changing transitions of quarks and leptons. The weak interactions were not fully described until Glashow, Salam, and Weinberg in the 1960s developed the theory of the electroweak interaction, in which the electromagnetic and weak forces are unified.

The electroweak interaction is described by the $SU(2)_L \times U(1)_Y$. Electrically charged fermions appear both left and right-handed states, the weak interaction also distinguish between left and right handed states. Only the left-handed fermions interact with the

TABLE 2.2: Left-handed and right-handed fermions together with their charge (Q), hypercharge (Y) and third component of the hypercharge (T_3).

Generation			Quantum Numbers		
I	II	III	Q	Y	T_3
$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L$	-1 0	$1/2$	$+1/2$ $-1/2$
e_R	μ_R	τ_R	-1	0	0
$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$+2/3$ $-1/3$	$1/2$	$+1/2$ $-1/2$
u_R	c_R	t_R	$+2/3$	0	0
d_R	s_R	b_R	$-1/3$	0	0

$SU(2)_L$ gauge fields, they are organized in $SU(2)$ doublets, while the right-handed states are singlets, as shown in Table 2.2.

Electroweak Symmetry Breaking

For many years, one of the main problems of the SM was how to introduce the mass of the standard particles into the theory. The weak vector bosons, for example, are massive particles, but putting terms into the Lagrangian that directly give them a mass would break the underlying symmetry that was so elegantly used to introduce them in the first place. The solution is to break the electroweak symmetry in a spontaneous way. The Lagrangian remains invariant but the ground state, the vacuum, is not invariant under the gauge symmetry.

Spontaneous electroweak symmetry breaking is governed by the Higgs mechanism. An additional complex scalar field (ϕ) doublet is introduced. The symmetry breaking occurs as the field ϕ adquires a vacuum expectation value (vev):

$$\langle \phi \rangle = \frac{1}{2} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

Once the Higgs field acquires a vev, the gauge bosons gain mass through interactions with the Higgs field. There are now three vector fields representing massive gauge bosons, $W_\mu^\pm$ and Z_μ^0, and a fourth vector field, A_μ, orthogonal to Z_μ^0 which remains massless and can be interpreted as the photon. For the same reason, quarks and charged leptons also gain mass when interacting with the Higgs boson.

The complex scalar field doublet which is added to the SM Lagrangian can be written in terms of four real scalar fields. Three of the fields are absorbed to generate mass to the heavy gauge bosons while the fourth emerges as a new massive scalar boson, the Higgs boson. The mass of the Higgs boson is not predicted by the theory and has to be experimentally determined. After the Higgs discovery by CMS [3] and ATLAS Collaborations [1], both collaborations have been working together to conclude that the Higgs boson has a mass of 125.09 $\pm$ 0.31 GeV [2].

2.1.2 Limitations of the Standard Model

The SM is a very successful theory that has been tested experimentally very precisely. However, many fundamental questions remain unsolved.

The more evident shortcoming of the SM as a complete theory of all elementary particles and fundamental interactions is that it does not include gravity. General relativity successfully describes gravitational dynamics on astrophysical scales, however is seems negligible at the distance scales studied in particle physics. Attempts have been made to incorporate gravity inside at quantum level, so far without success.

Another open question comes from the nature of the electroweak symmetry breaking. Do multiple Higgs bosons and charged Higgs bosons exist? Also, the mechanism of electroweak symmetry breaking is related to a particularly severe problem of the SM from the theoretical point of view: the "hierarchy problem". The basis is that in the SM, when expanding its contributions in perturbation theory one encounters divergent corrections from loop diagrams [6]. However, in order to get a mass compatible with recent measurements, these corrections have to cancel mutually to a great extent. The SM does not provide any mechanism by which to protect the Higgs mass from these corrections, so we are left with the unsatisfactory assumption that the parameters just happen to be very finely tuned to yield a low value for the mass, which seems fairly unnatural [7, 8].

Other intriguing question the Standard Model does not provide an answer is what causes the matter/antimatter asymmetry observed in the universe. Three conditions were postulated by Sakharov [9] to be fulfilled to allow for this asymmetry: baryon number violation, C- and CP-symmetry violation, and interactions having the previous two features that fall out of thermal equilibrium. The Standard Model provides no source for CP-violation which is strong enough to explain the observed asymmetry.

Another unsolved questions come from the field of cosmology. From astrophysical observations, there are strong evidence for dark matter [10]. The most common explanation

consistent with observations is that dark matter is made of stable neutral particles. From cosmic microwave background measurements the density of the dark matter can be inferred as well as its non-relativistic nature (cold dark matter). The only stable, neutral particles within the standard model are the neutrinos and they are not compatible with these observations. Therefore, a new theory beyond the SM would be needed to provide a weakly interacting massive particle (WIMP) that would serve as a candidate for dark matter.

2.2 Physics Beyond the Standard Model

Over the years, many models have been developed starting from the Standard Model to try to explain as many of these open questions discussed before. Many theoretical models have been postulated that could potentially solve some of these problems where the Standard Model has failed to provide an answer to. With the LHC restart in Jun 2015, the quest for such new physics models beyond the Standard Model has started.

2.2.1 Supersymmetry

One of the most beloved theory that may cover most of the questions the Standard Model fails to solve is the so called Supersymmetry. A very brief description is covered here, for a more detailed and rigorous presentation of the theory, please refer to the primer of Martin [8]. The reader may also want to refer to Pape and Treille's review [11] covering the topic from an experimental perspective.

Taming divergences, solving the hierarchy problem.

While the SM keeps a clear distinction between the matter particles (fermions) and the force carriers (bosons), Supersymmetry [12] introduces a symmetry by associating each fermionic state of the SM to a bosonic state in the same multiplet and vice versa. In the simplest SUSY version, called the minimal supersymmetric standard model (MSSM) [13] a scalar superpartner is associated with each SM fermion and a fermionic superpartner to each vector boson. An important consequence of this doubling of states is that it tames the divergences of quantum corrections to the Higgs mass. In supersymmetry,the contributions from SM fermionic loops are exactly compensated by opposite sign contributions from their bosonic superpartners and vice versa, canceling the quadratic divergences.

SUSY breaking

As it happens in the SM, the symmetry introduced by SUSY must be broken somehow, however, the mechanism which breaks supersymmetry and yields to particle masses is unknown. The most general formulation of SUSY introduces 105 parameters, in addition to the 19 of the SM. Some of them are severely limited, as they would induce flavour changing neutral currents (FCNC) or CP violation at an unacceptable level. More simple models with less parameters have been postulated to try to derive experimental constrains.

Unification

All three coupling "constants" of the SM gauge group have been known with great precision at the Z^0 scale. This couplings vary their strength with the energy scale at which they are evaluated. This energy dependence suggested that a unification of the coupling might be possible at some high energy scale, leading to a grand unification of the gauge groups, unfortunately LEP data ruled out this possibility. On the other hand, if superpartners with masses in the TeV scale are included, the slopes of the evolution are changed accordingly and the couplings meet at a scale of about 2×10^{16} GeV [14].

SUSY Particle Spectrum

The MSSM introduces one superpartner for each SM particle that is only distinguished by its spin. Matter particles (fermions) of spin 1/2, quarks and leptons, get coupled with scalar spin 0 sparticles, called "squarks" and "sleptons". Gauge vector bosons of spin 1 acquire fermion partners of spin 1/2 , called the gauginos.

In the SUSY context we need at least two Higgs doublets to generate the masses, this results in at least four scalar Higgs bosons in the MSSM, two of which are charged. Their mass eigenstates are: h^0, H^0, A^0 and $H^{\pm}$. Their supersymmetric partners are called higgsinos and carry spin 1/2.

Gauginos and Higgsinos are mixed up together to form mass eigenstates: four neutral and four charged fermions, called "neutralinos" and "charginos".

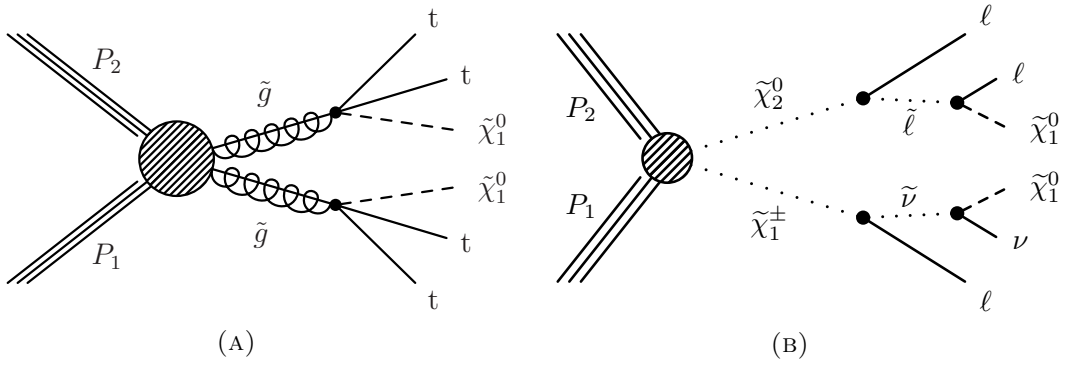
By convention, supersymmetric particles are denoted by a tilde over the corresponding particle symbol, e.g. a smuon would be denoted by $\tilde{\mu}$ and the gluino would be $\tilde{g}$, the neutralinos and charginos are denoted by the letter χ and are numbered in order of ascending masses: $\tilde{\chi}^0_{1,2,3,4}$ and $\tilde{\chi}^{\pm}_{1,2}$.

Dark Matter Candidate

The SM particles are distinguished from their superpartners by a quantum number, called R-parity [15], which is considered to be +1 for ordinary particles and -1 for superpartners. If it is assumed that R-parity is conserved in the interactions, SUSY particles can only be produced in pairs, and their decay will always involve another SUSY particle, and finally, at the end of the decay chain, the lightest supersymmetric particle (LSP) that has to be stable. This provide an excellent candidate for the dark matter in the universe as it is a massive, neutral and weakly interacting particle.

2.2.2 Experimental Signatures

Same-sign dileptons constitute a powerful experimental search channel for discovering supersymmetry [16], this signature, although heavily suppressed in the SM, can appear naturally in many new physics scenario and in SUSY in particular. Events containing same-sign dileptons may be produced for example through a cascade decays of gluino pair production or from chargino-neutralino production as depicted in Figure 2.1.



(A) (B)

FIGURE 2.1: Feynman diagrams illustrating like-sign dilepton production through gluino pair production if two top quarks decay leptonically (2.1a) and direct chargino-neutralino production (2.1b).

Although this work is focused on supersymmetry, there are many alternative models that can produce same-sign dileptons in the final state such as universal extra dimensions [17–19], production of $T_{5/3}$ particles (fermionic partners of the top quark) [20], heavy Majorana neutrinos [21] or same-sign top-quark pair production [22, 23].

2.3 Simplified Models for LHC New Physics Searches

Since the beginning of the data-taking back in 2010, both CMS and ATLAS have been probing new physics beyond the reach of any past experiment. With this early searches

at the LHC it is crucial to determine whether certain classes of new physics can evade existing searches, but still be detected with newer techniques. It is also important to understand the physical implications of the new-physics searches, either with or without evidence.

Simplified models [24], involving relatively few particles and decay modes, provide a foundation for assessing the impact of existing searches, and how they can be extended or better optimized. In addition, they could be also a useful starting point for characterizing any evidence for new physics.

Every model of new physics is defined by a effective Lagrangian at the TeV-scale that describes particles and interactions. A simplified model is designed to involve only a few new particles and interactions. In many cases they can be considered as limits of more general new physics scenarios. Simplified models are described by a small number of parameters directly related to collider physics observables: particle masses (and their decay widths, which can sometimes be neglected), production cross-sections, and branching fractions.

Simplified models are still model-dependent, however they avoid some of the drawbacks of model-dependence. The sensitivity of any new-physics search can be studied and presented as a function of the few parameters of the simplified model and in particular over the full range of new particle masses. Though defined within a simplified model, these topology-based limits also apply to more general models giving rise to the same topologies.

Chapter 3

The LHC and the CMS experiment

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [25] is the world's largest superconducting circular proton-proton accelerator and collider, built by the European Centre of Nuclear Research (CERN) and hosted in a 27 km underground tunnel it was design to collide two beams of protons at a centre-of-mass energy of 14 TeV with an instantaneous luminosity of $10^{34} cm^{-2} s^{-1}$.

The first operations of the LHC machine started back in 2009. In spring 2010 operations at 3.5 TeV per beam started and by the end of the year CMS was able to collect 36pb^{-1} of data; the operation at that energy continued until the end of 2011 and during such period the LHC delivered 4.8 fb^{-1} of physics data to the experiments. In 2012 the beam energy was increased to 4 TeV per beam, with an instantaneous luminosity close

Table 3.1: Key design parameters of the LHC machine

Parameter	Value
Circumference	26.659 km
Maximum beam energy	7 TeV
Dipole magnet field (at 7 TeV)	8.33 T
Dipole electrical current (at 7 TeV	11850 A
Peak luminosity	$10^{34} cm^{-2} s^{-1}$.
Number of bunches	2808
Minimum bunch spacing	25 ns
Number of protons per bunch	1.15×10^{11}
Number of pp collisions per bunch crossing	20
Magnet operating temperature	1.9 K

to $6.5 \times 10^{33} cm^{-2} s^{-1}$, allowing the detector to deliver more than 23 fb^{-1} by the end of the year. In Figure 3.1 the delivered luminosity versus time for 2010, 2011 and 2012 pp data is shown.

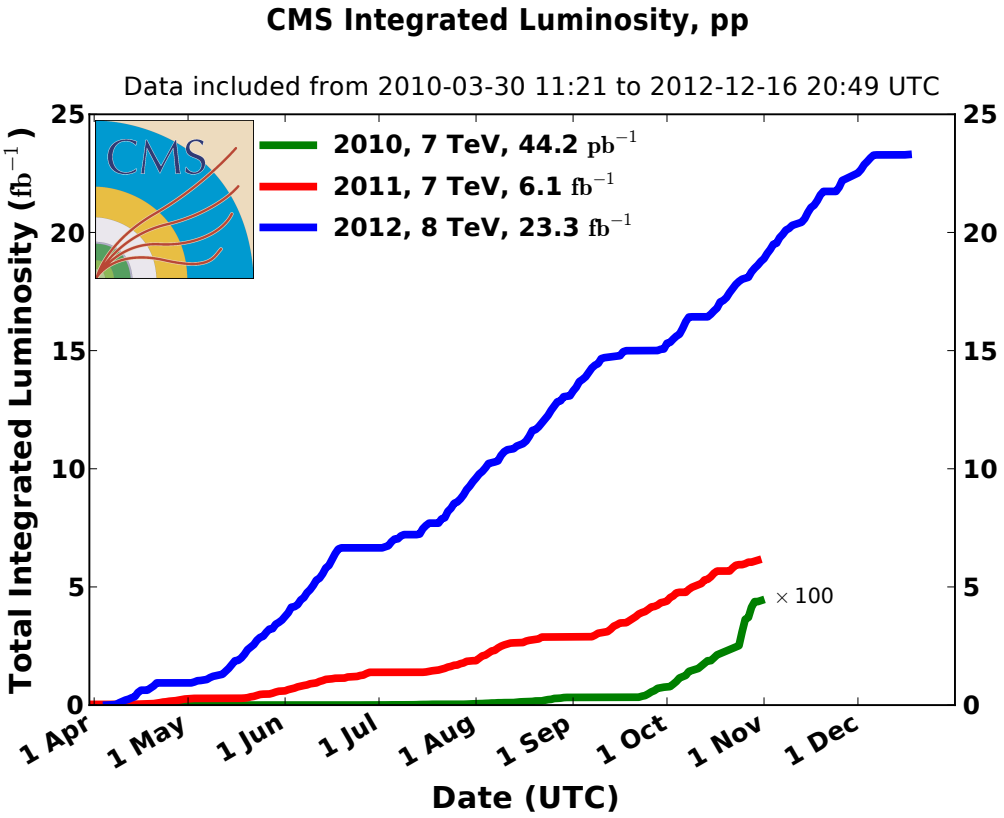


FIGURE 3.1: Delivered Luminosity versus time for 2010, 2011, 2012 (p-p data only). Cumulative luminosity versus day delivered to CMS during stable beams and for p-p collisions. This is shown for 2010 (green), 2011 (red) and 2012 (blue) data-taking.

In 2011 and 2012 the LHC was operated a bunch spacing of 50 ns, and therefore a capacity of 1400 colliding bunches. Considering instantaneous luminosity values and the cross section for inelastic pp collisions, one would expect 10-36 interactions per bunch crossing. This pile-up of multiple interactions poses significant challenges to the trigger, event reconstruction and analysis, as it increases significantly the complexity of the events, for example, an event with 15 simultaneous interactions yield to a total flux of particles similar to a $t\bar{t}$ event.

In 2015, the LHC started its operation, firstly with a bunch spacing of 50 ns, and then with a 25ns bunch spacing and with a nominal instantaneous luminosity of $5 \cdot 10^{33} \mathrm{cm}^{-2}\mathrm{s}^{-1}$. Delivering around 1 fb^{-1} to both CMS and ATLAS by the 25^{th} of September as seen in Figure 3.2.

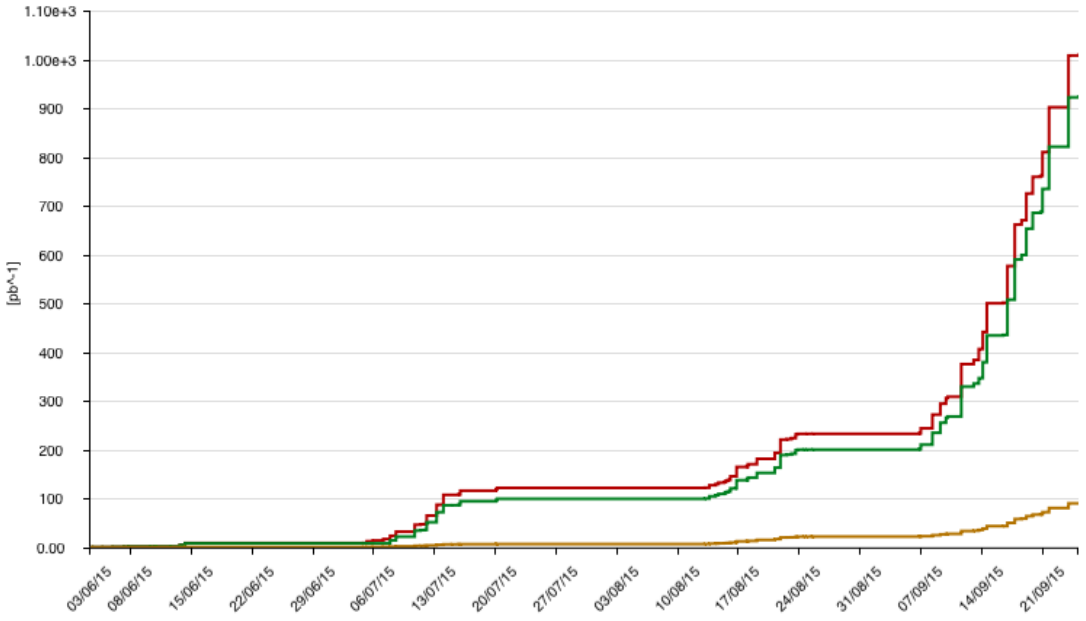


FIGURE 3.2: Delivered Luminosity versus time for 2015 (pp data only). Cumulative luminosity versus day delivered to ATLAS (red), CMS (green) and LHCb (brown) during stable beams and for p-p collisions.

3.2 The CMS experiment

The Compact Muon Solenoid (CMS) detector [26] is a general-purpose experiment installed at Point 5 of the LHC ring. CMS is designed to fully exploit the physics reach of the LHC. After the discovery of the Higgs Boson the CMS physics program has been focused in exploring the new energy and luminosity regime to search for new physics that may help to complete the picture of elementary particle physics.

The CMS detector meets the following requirements: good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx$ 1% at 100 GeV), and the ability to determine unambiguously the charge of muons with p $<$ 1 TeV; good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ's and b-jets, requiring pixel detectors close to the interaction region; good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx$ 1% at 100GeV), wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation at high luminosities; good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and with fine lateral segmentation.

To achieve the above goals, CMS uses a large superconducting solenoid magnet that envelopes a large silicon tracker and both electromagnetic (ECAL) and hadronic (HCAL) calorimeters. Outside the solenoid there is an iron yoke to control the magnetic flux

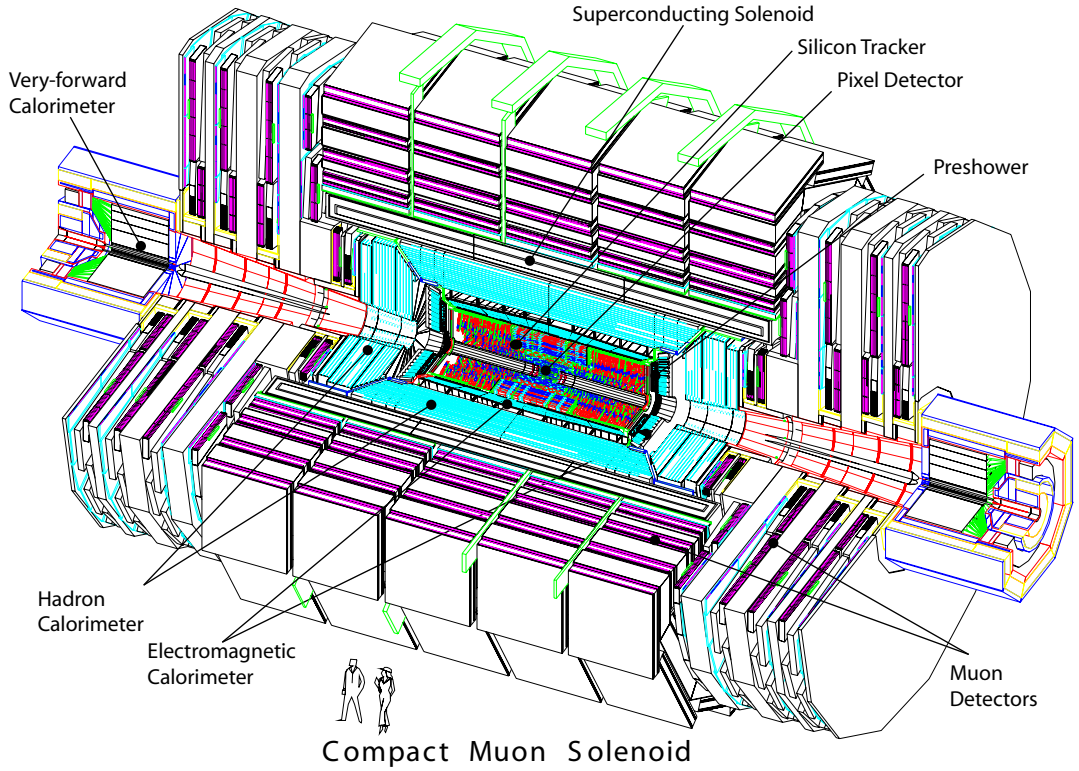


FIGURE 3.3: A perspective view of the CMS detector

return with multiple layers of muon detectors installed in it. A schematic view of the CMS detector is presented in Figure 3.3. In the rest of this chapter I will describe briefly each of the subdetectors. A more detailed description can be found in [26] and [27].

Coordinate System. The coordinate system adopted by CMS has the origin centred at the nominal collision point inside de experiment, the y-axis points vertically upward, and the x-axis points radially inward toward the centre of the LHC. Thus, the z-axis points along the beam direction toward the Jura mountains from LHC Point 5. The azimuthal angle ϕ is measured from the x-axis in the x-y plane and the radial coordinate in this plane is denoted by r. The polar angle θ is measured from the z-axis. Pseudo-rapidity is defined as $\eta = -\ln \tan(\theta/2)$. Thus, the momentum and energy transverse to the beam direction, denoted by p_T and E_T, respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss}.

3.2.1 Tracking System

The tracker is the innermost sub-detector. It allows CMS to reconstruct the trajectories and momenta of the charged particles that emerge from the interaction point as well as the position of the primary and displaced vertices for rejecting additional interactions and to reconstruct objects with significant lifetimes such as b-jets. The high granularity and the large number of tracker layers allow to provide an efficient reconstruction of every single collision. A schematic cross section of the inner detector is shown in Figure 3.5.

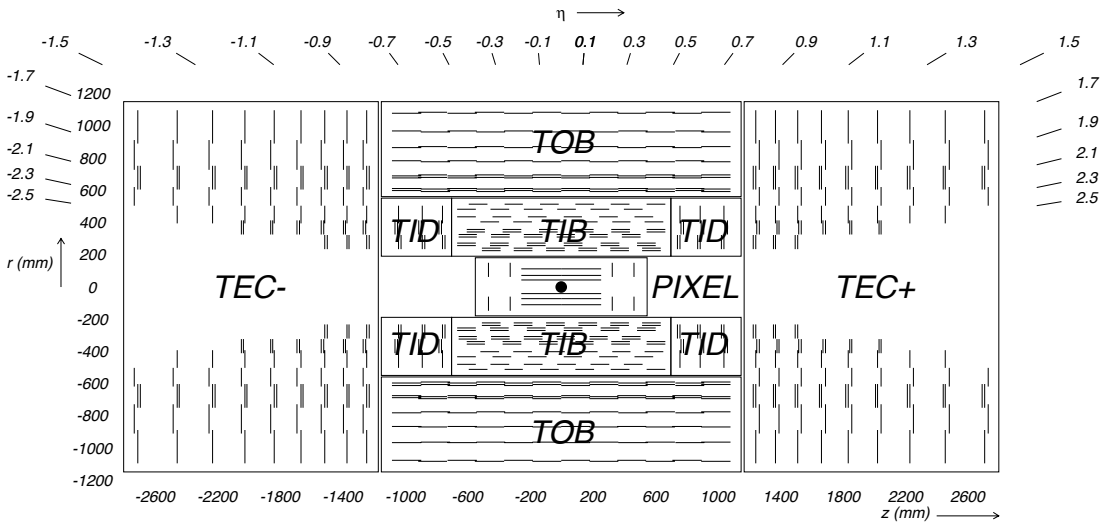


FIGURE 3.4: Schematic cross section of the CMS silicon tracker. Each line corresponds to a detector module while double lines correspond to back-to back modules

Pixel detectors Closest to the interaction point, the tracks are measured with silicon pixel detectors, arranged in three barrel layers and two endcap discs. These layout provides CMS with 2D points with a resolution of 10-15 μm, allowing to precisely determine the position of the interaction point as well as other vertices.

Silicon Strip detectors Outside the pixel detector, the particle flow is much lower so silicon strip detectors are used, arranged in 10 barrel layers and 12 endcap disks. In the barrel, strips are parallel to the z axis, while in the endcaps they are along the radial coordinate, to provide in both cases a precise measurement of the $r-\phi$ coordinate (about 20–50 μm). In order to measure also the z coordinate with a precision better than the strip length (about 10 cm), some tracker layers contain an additional set of sensors tilted by 100 mrad with respect to z. Matching the hits on the $r-\phi$ detectors to those on the tilted detectors, it is possible to reconstruct 2D points with a z resolution of 200–500 μm.

3.2.2 Electromagnetic calorimeter

The electromagnetic Calorimeter (ECAL) is a hermetic, homogeneous calorimeter comprising 61200 lead tungstate ($PbWO_4$) scintillating crystals mounted in the central barrel part, closed by 7324 crystals in each of the 2 endcaps providing a pseudo-rapidity coverage $|\eta| < 3.0$. A pre-shower detector is installed at the face of each of the two endcaps. Avalanche photo-diodes (APDs) are used as photo-detectors in the barrel and vacuum photo-triodes (VPT) in the endcaps.

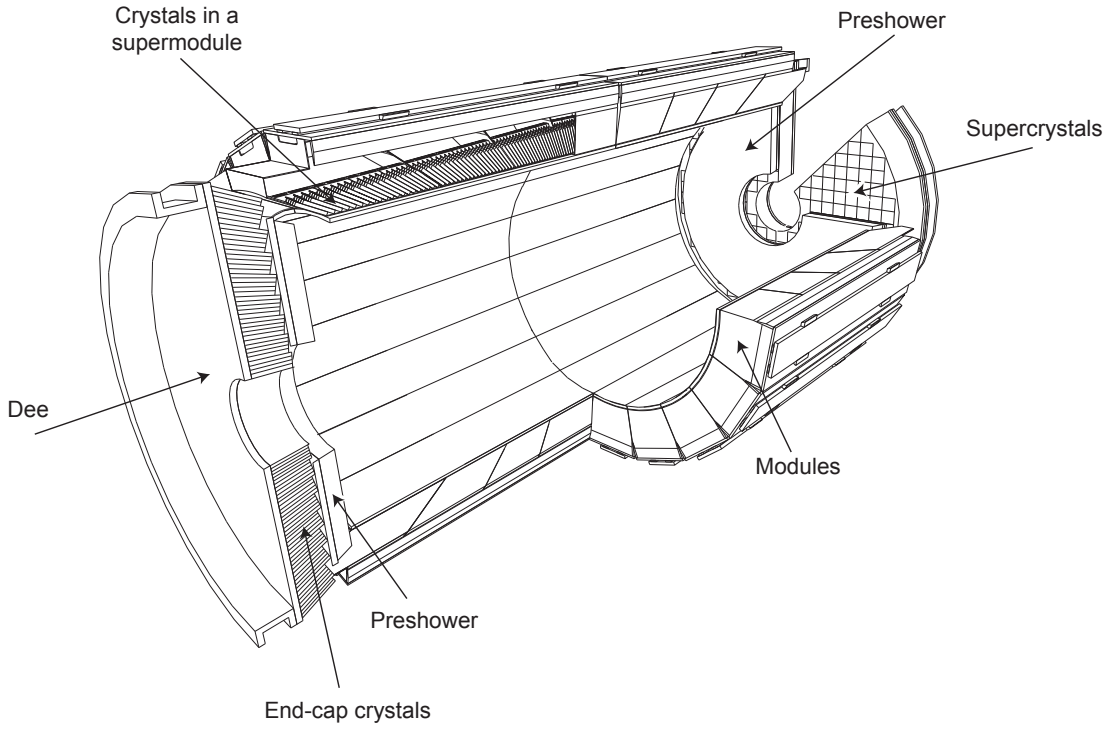


FIGURE 3.5: Layout of the CMS Electromagnetic calorimeter.

Lead tungstate is characterised by a short radiation length ($X_0 = 0.89$cm), allowing a very good longitudinal shower containment with a detector depth of about 26 X_0 in the barrel, and about (3 + 25) X_0 in the pre-shower+endcaps. The position resolution is optimised by requiring small transverse size of the electromagnetic shower. The transverse profile of electromagnetic showers is expressed by the Moliere radius (R_m). The Moliere radius of the lead tungstate crystals is 2.2 cm and combined with a barrel crystal front face of 2.2×2.2 cm^2 ensures that the core of the shower will be included in a 2×2 crystal area. Lead tungstate provides optimal scintillation time since about 80% of the light is emitted in 25 ns which is the nominal beam crossing time separation.

The use of $PbWO_4$ crystals has thus allowed the design of a compact calorimeter inside the solenoid that is fast, has fine granularity, and is radiation resistant.

3.2.3 Hadronic calorimeter

The design of the hadron calorimeter (HCAL) is strongly influenced by the choice of magnet parameters since most of the CMS calorimetry is located inside the magnet coil and surrounds the ECAL system. An important requirement of HCAL is to minimise the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity for the E_T^{miss} measurement. Hence, the HCAL design maximises material inside the magnet coil in terms of interaction lengths. This is complemented by an additional layer of scintillators, referred to as the hadron outer (HO) detector, lining the outside of the coil.

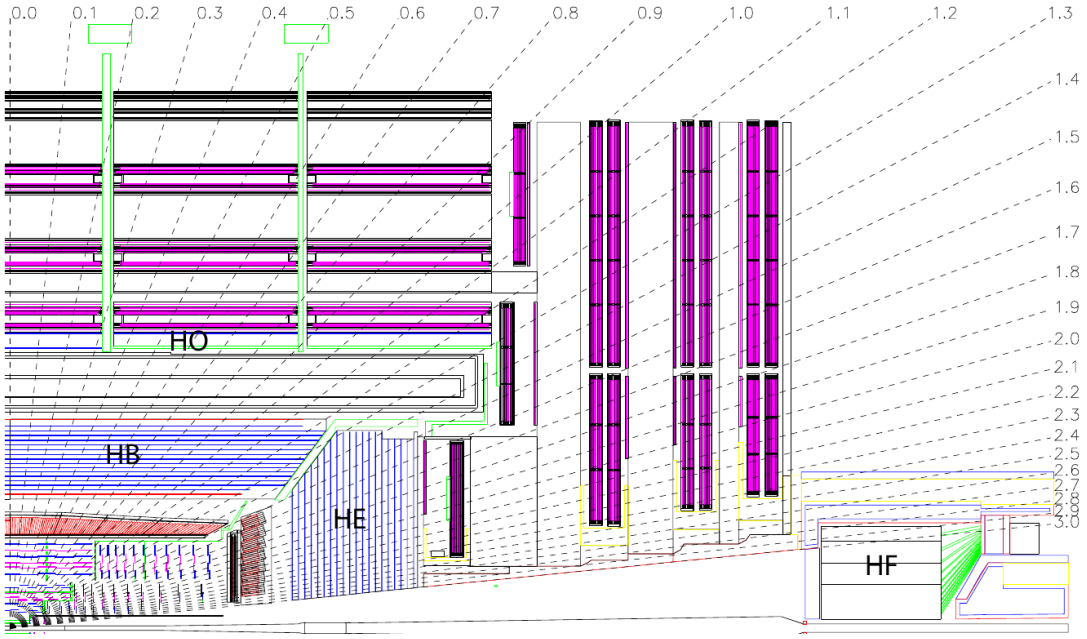


FIGURE 3.6: Layout of the CMS Hadronic calorimeter.

Brass has been chosen as absorber material as it has a reasonably short interaction length, is easy to machine and is non-magnetic. Maximising the amount of absorber before the magnet requires keeping to a minimum the amount of space devoted to the active medium. The tile/fibre technology makes for an ideal choice. It consists of plastic scintillator tiles read out with embedded wavelength-shifting (WLS) fibres. The WLS fibres are spliced to high-attenuation-length clear fibres outside the scintillator that carry the light to the readout system. The overall assembly enables the HCAL to be built with essentially no cracks or dead areas in ϕ. The gap between the barrel and the endcap HCAL, through which the services of the ECAL and the inner tracker pass, is inclined at 53° and points away from the centre of the detector.

3.2.4 Muon System

An efficient and precise muon identification is one of the main requirement that influenced CMS design. Muons are the only particles that cross the full detector before decaying, leaving a unique trace in the detector: as it is a charged particle it will leave a track in the tracking system, a very small deposit in the calorimeters and a number of hits in the muon system.

The muon detectors are placed inside the iron yoke to exploit the return field that provides bending of the particles for momentum measurement. Three types of gaseous detector are used to identify and measure muons. The choice of the detector technologies has been driven by the very large surface to be covered and by the different radiation environments In the barrel region ($|\eta| < 1.2$), the muon rate is low and the residual magnetic field in the chambers is low, drift tube (DT) chambers are used. In the endcaps, the muon rate is high, and the magnetic field is also high, cathode strip chambers (CSC) are deployed and cover the region up to $|\eta| < 2.4$. In addition to this, resistive plate chambers (RPC) are used in both the barrel and the endcap regions. These RPCs are operated in avalanche mode to ensure good operation at high rates, providing a fast response with good time resolution but with a coarser position resolution than the DTs or CSCs. A layout of the muon system is depicted in Figure 3.7.

Drift Tube chambers In the CMS barrel, where the muon rate is under control, four layers of muon stations are used, occupied by Drift Tube(DT) Chambers covering up to $|\eta| < 1.2$. Drift Tube chambers consist of individual drift tube cells that contain a 50μm-diameter anode wire and two electrode plates that create the drift electric field. The walls of the cell are grounded, acting as cathodes. The cells are filled with a gas mixture of 85% Ar and 15% CO_2 and the wire and electrodes are operated with a voltage difference of about 1.8 kV. Picking a transverse dimension for the cell of 21 mm to optimise drift time, gain and number of channels, all the above design conditions provide a gain of 105, resulting in a drift time of 380 ns and a linear relationship between drift time and drift path which is essential for the chamber to provide triggering capabilities. Each DT chamber is made of three (or two) super-layers (SL) where each super-layer is made of four layers of drift cells staggered by half a cell. The outer SLs have wires aligned parallel to the beam line providing a measurement in the transverse bending plane while in the inner SL, the wires are perpendicular to the beam line to provide a measurement of the z position of the track.

Cathode Strip Chambers The Cathode Strip Chambers (CSCs) are installed in the endcaps, providing full muon coverage up to $|\eta| < 2.4$. The CSCs are multi-wire

proportional chambers consisting of six planes of anode wires interleaved among seven cathode panels. Wires run azimuthally, defining the tracks radial component, while strips are milled on cathode panels and run lengthwise at a constant $\Delta\phi$ width. The nominal gas mixture is 40% Ar, 50% CO_2 and 10% CF_4. The addition of CF_4 is used to avoid polymerisation of the wires. The CSCs can operate at high rates and in large and non-uniform magnetic fields without requiring precise monitoring of gas, pressure or temperature and can provide trigger and precision position measurement in the same device. The wires give very fast signals that provide very good time resolution while the development of the avalanche on the strips gives very good position resolution.

Resistive Plate Chambers Resistive Plate Chambers (RPC) are gaseous parallel plate detectors than combine adequate position resolution with very high speed. RPCs are able to tag the presence of an ionising particle in much less than a bunch crossing, which makes RPCs an ideal trigger device since they can associate the correct bunch crossing with the muon. The CMS RPC chamber consists of two gaps operated in avalanche mode with read-out strips in between. The total induced signal is the sum of the induced signal in both gaps. RPCs need intensive monitoring of temperature, humidity and pressure to ensure stability of conditions for proper operation. The RPC system spans both barrel and endcap and for the initial CMS operation chambers have been installed up to $|\eta| < 1.6$.

3.2.5 Trigger System

The LHC bunch crossing rate of 40 MHz leads to about 10^9 interactions per second at design luminosity, however only about 1000 crossings per second can be written to archival media. The CMS trigger system consists of 2 stages, the Level-1 trigger (calorimeter, muon, and global) and the High-Level Trigger (HLT) which allow for a rejection factor of nearly 10^6.

3.2.5.1 Level-1 trigger

Custom hardware processors form the Level-1 decision. The L1 triggers involve the calorimetry and the muon systems, as well as some correlation information between these systems. The L1 decision is based on the presence of "trigger primitives" objects such as photons, electrons, muons and jets above certain energy threshold. It also employs global sums of E_T and $\mathrm{E_T^{miss}}$. The L1 rate is limited at 100 kHz, which is set by the average time to transfer the full detector information through the readout system.

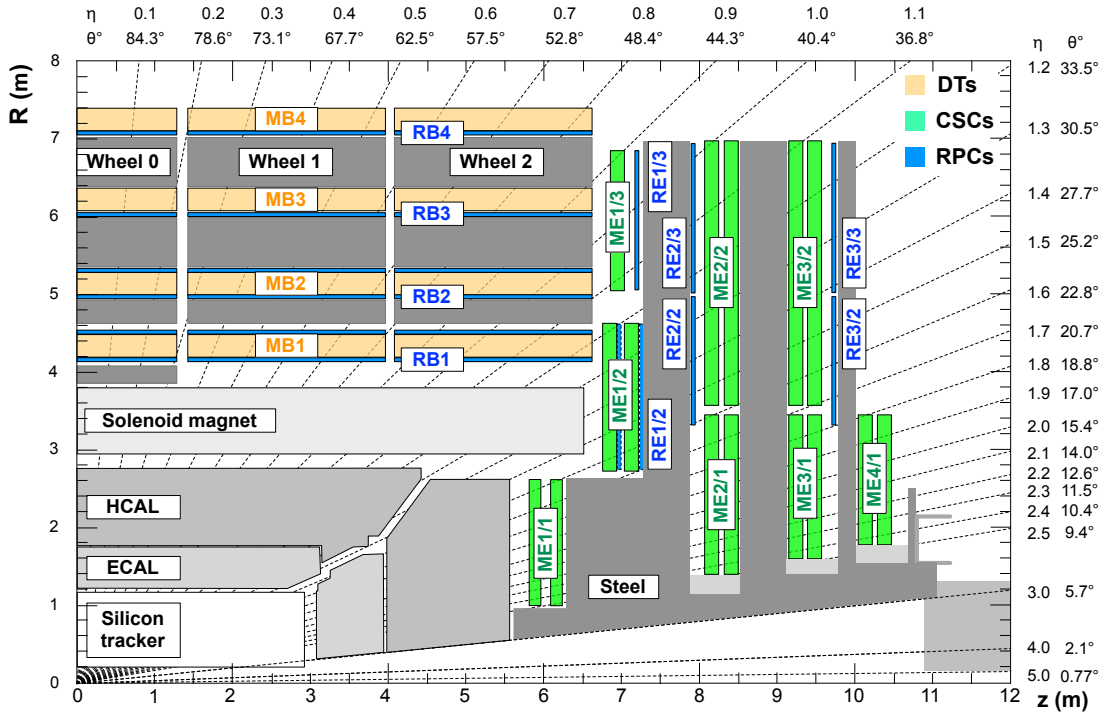


FIGURE 3.7: An R–z cross section of a quadrant of the CMS detector with the axis parallel to the beam (z) running horizontally and radius (R) increasing upward. The interaction point is at the lower left corner. Shown are the locations of the various muon stations and the steel disks (dark grey areas). The 4 drift tube (DT, in light orange) stations are labeled MB ("muon barrel") and the cathode strip chambers (CSC, in green) are labeled ME ("muon endcap"). Resistive plate chambers (RPC, in blue) are in both the barrel and the endcaps of CMS, where they are labeled RB and RE, respectively.

During the Level-1 decision-making period, all the high-resolution data is held in pipelined memories. Computer processors make subsequent decisions using more detailed information from all of the detectors in more and more sophisticated algorithms that approach the quality of final reconstruction.

3.2.5.2 High-Level trigger

Upon receipt of a L1 trigger, the data is transferred to the HLT farm, where each processor runs the same HLT software to reduce the L1 output of 100kHz to 1kHz for offline processing and storing.

Since the HLT processing is much slower than the L1, rather than reconstruct all possible objects in an event, whenever possible only those objects and regions of the detector that are actually needed are reconstructed. Events are discarded as soon as possible to optimise the time each node takes to process a event: simpler algorithms run first and basic requirements are applied before running more complex algorithms that take longer

time. If an event is rejected in the first steps, complex algorithms will not run and the processing node will be ready for the next event. The algorithms used at the HLT are very similar (or identical) to the ones used for the offline reconstruction of the event.

Chapter 4

Event Reconstruction

Once the data is collected, the collision information from every sub-detector is stored in "RAW" format for offline event reconstruction. Remember than only about 1000 events per second are stored as the rest of the events have been rejected by the trigger system. The reconstruction process is a complex set of algorithms used to identify particle candidates and higher level quantities to be used later for analysing the events.

4.1 Track and Vertex Reconstruction

The CMS inner tracker system is used to detect charged particles and measure their momentum through the bending in the magnetic field. A good performance of the inner detector and tracking is crucial to perform any new physics search.

The analysis described in this thesis relies on leptons, for which information from the inner detector is used to determine the momentum, together with information from ECAL for electrons and from the muon system for muons. Also the tracker is used to determine the impact parameter of the leptons, that helps identifying whether the lepton was promptly produced or not. The lepton charge is also measured in the tracker by looking at the bending of the tracks due to the magnetic field.

The reconstruction of charged hadrons is fundamental for the particle flow event reconstruction, since they constitute two thirds of the total particles produced in the collision. Charge hadron tracks are also very useful to select isolated leptons from signal processes rejecting the backgrounds from heavy quark decays or misidentified particles.

Charge hadron reconstruction, precise impact parameter determination and vertex identification are also the main ingredients to tag jets from heavy flavour decays.

4.1.1 Track reconstruction

Tracks are reconstructed in CMS [27] using multiple iterations of the same algorithm, that consists of three steps: seeding, patter recognition and final fit. After each iteration, tracks are passed through a quality filter to reject fakes, and hits corresponding to good tracks are removed from the input lists of the subsequent iterations to simplify the pattern recognition.

Seeding algorithm

Seed generation provides the initial trajectory candidates for the full track reconstruction. It searches for pairs or triplets of hits compatibles with the hypothesis of a track coming from the interaction region. The beamspot information is used in the seeding step on hit pairs. In order to cope with large pile-up, two techniques are used to reduce combinatorics: ignore hits whose charge distribution is not compatible with the incidence angle and using pixel-only vertex reconstruction to reduce the combinatorics of hit pairs.

Pattern recognition

The pattern recognition is based on a combinatorial Kalman filter method [28]. The filter proceeds iteratively from the seed layer, starting from a coarse estimate of the track parameters provided by the seed, and including the information of the successive detection layers one by one. On each layer, i.e. with every new measurement, the track parameters are known with a better precision, up to the last point, where they include the full tracker information.

First, a dedicated navigation component determines which layers are compatible with the initial seed trajectory. The trajectory is then extrapolated to these layers according to the equations of motion of a charged particle in a constant magnetic field, accounting for multiple scattering and energy loss in the traversed material.

Since several hits on the new layer may be compatible with the predicted trajectory, several new trajectory candidates are created, one per hit. In addition, one further trajectory candidate is created, in which no measured hit is used, to account for the possibility that the track did not leave any hit on that particular layer. This fake hit is called an invalid hit.

Each trajectory is then "updated" with the corresponding hit according to the Kalman filter formalism. This update can be seen as a combination of the predicted trajectory

state and the hit in a weighted mean, as the weights attributed to the measurement and to the predicted trajectory depend on their respective uncertainties.

All resulting trajectory candidates are then grown in turn to the next compatible layer(s), and the procedure is repeated until either the outermost layer of the tracker is reached or a "stopping condition" is satisfied. In order not to bias the result, all trajectory candidates are grown in parallel. To avoid an exponential increase of the number of trajectory candidates the total number of candidates is truncated at each layer.

4.2 Electron Reconstruction

The reconstruction of electrons in CMS uses information from the pixel detector, the tracker and the ECAL. One is looking for a single track coming from the interaction point that matches an electromagnetic supercluster.

Electrons in CMS are reconstructed using either ECAL or tracker seeds. ECAL seeded reconstruction starts from an electromagnetic deposit in the ECAL and matches this deposit with a track to build an electron candidate [29]. Track seeded reconstruction starts from tracks and associates bremsstrahlung radiation to reconstruct electron objects and it's more suitable for low momentum electrons or electrons inside jets.

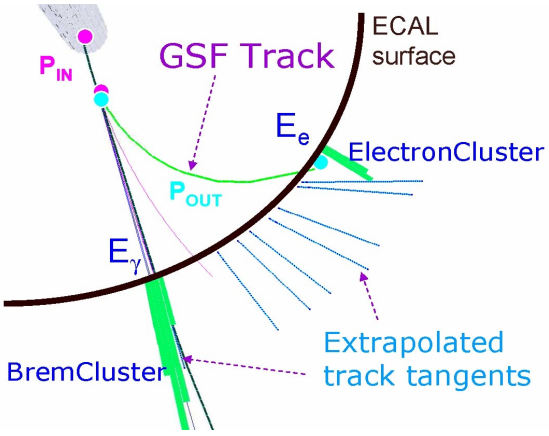


FIGURE 4.1: Illustration of an electron undergoing a large radiative energy loss in the material, depicted in the transverse plane: the local curvature of the GSF track close to the interaction point provides an estimate of the initial momentum, while that close to the ECAL surface allows to estimate the final one. The tangents to the trajectory are used to collect the bremsstrahlung clusters in the particle flow.

The reconstruction in the ECAL starts from the clustering of the calorimeter hits into clusters. Due to the large amount of tracker material before the calorimeter, most of the electrons will radiate photons when going through the tracker. Because of the large magnetic field the energy arrives to the clusters spread in ϕ. To account for this

energy spread, basic clusters are summed into superclusters which are extended in the ϕ direction. Superclusters with $E_T > 4\ GeV$ are selected and matched to track seeds in the inner tracker layers to built electron tracks from those seeds. Trajectories are reconstructed using a dedicated model of the electron energy loss and fitted using a Gaussian Sum Filter(GSF) algorithm [30]. Finally the momentum assignment for the electron is performed by combining the energy measured in the ECAL with the momentum assigned to the track by the GSF algorithm.

The reconstruction algorithm is completed by a particle flow based one where the energy deposits are searched along the tangents of the tracks. This is illustrated in Figure 4.1

4.2.1 Charge estimation

The measurement of the electron charge is affected by bremsstrahlung followed by photon conversions. In particular, when the bremsstrahlung photons convert upstream in the detector, they lead to very complex hit patterns, and the contributions from conversions can be wrongly included in the fit of the electron track.

The natural choice for a charge estimate would be the sign of the GSF track curvature, which unfortunately can be altered by the misidentification probability in presence of conversions. This is improved by combining two other charge estimates, one based on the associated Kalman-Filter track matched to a GSF track, and the second one that is evaluated using the supercluster position, and defined as the sign of the difference in ϕ between the vector joining the beamspot to the supercluster position and the vector joining the beamspot and the first hit of the electron GSF track. In CMS, electron charge is defined by the sign shared by at least two of the the estimates, however, in this dissertation, the agreement of the three estimates will be required to reduce the charge misidentification to a minimum as it will be detailed later in Chapter 5.

4.2.2 Electron identification

Several strategies are used in CMS to identify prompt isolated electrons (signal) [31], and to separate them from background sources, mainly originating from photon conversions, jets misidentified as electrons, or electrons from semileptonic decays of b and c quarks. Simple and robust algorithms have been developed to apply sequential selections on a set of discriminants.

Variables that provide discriminating power are grouped into three main categories: Observables that compare measurements obtained from the ECAL and the tracker (track–cluster matching, including both geometrical as well as supercluster energy–track

momentum matching). Purely calorimetric observables used to separate genuine electrons (signal electrons or electrons from photon conversions) from misidentified electrons (e.g., jets with large electromagnetic components), based on the transverse shape of electromagnetic showers in the ECAL and exploiting the fact that electromagnetic showers are narrower than hadronic showers. Energy fractions deposited in the HCAL (expected to be small, as electromagnetic showers are essentially fully contained in the ECAL), as well as the energy deposited in the pre-shower in the endcaps are also used. Finally, tracking observables employed to improve the separation between electrons and charged hadrons, exploiting the information obtained from the GSF-fitted track, and the difference between the information from the Kalman-Filter and GSF-fitted tracks.

The distance $\Delta\eta$ ($\Delta\phi$), defined as the distance between the supercluster position in η (ϕ) and the track η (ϕ) extrapolated from the innermost track position and direction to the position of closest approach to the supercluster, reduces the misidentification probability while preserving the highest efficiency $\Delta\eta$ and $\Delta\phi$ increases with the amount of bremsstrahlung.

The lateral extension of the shower along the η direction is expressed in terms of the variable $\sigma_{\eta\eta}$, which is defined as $(\sigma_{\eta\eta})^2 = [\sum(\eta_i - \eta)^2 w_i]/\sum w_i$. The sum runs over the 5x5 matrix of crystals around the highest $\mathrm{E_T}$ crystal of the SC, and w_i is a weight that depends logarithmically on the contained energy. The positions η_i are expressed in units of crystals, which has the advantage that the variable-size gaps between ECAL crystals (in particular at modules boundary) can be ignored. The discrimination power of the $\sigma_{\eta\eta}$ is greater than the analogous variable in ϕ, because bremsstrahlung strongly affects the pattern of energy deposition in the ECAL along the ϕ direction.

Finally $1/E - 1/p$, where E is the supercluser energy and p the track momentum at the point of the closest approach to the vertex. While signal electrons tend to have values very close to zero, background electrons have negative values.

The identification selection applies requirements on five identification variables among those discussed previously: $\Delta\eta$, $\Delta\phi$, H/E, $\sigma_{\eta\eta}$ and $1/E - 1/p$. In addition, a selection is also applied on the variables used to reject converted photons. Finally, the impact parameters of the electron, d_0 and d_z, are required to be small for the electron to originate from the vertex of interest. Four working points were designed by CMS and were optimised separately for electrons in the ECAL barrel and endcaps: loose, medium, tight, and a very loose point for analyses aiming at vetoing electrons. The selection that will be used in this work is very similar to the medium working point and will be described later in Section 5.2.2. The identification efficiency as a function of the $\mathrm{p_T}$ of the electron can be seen in Figure 4.2

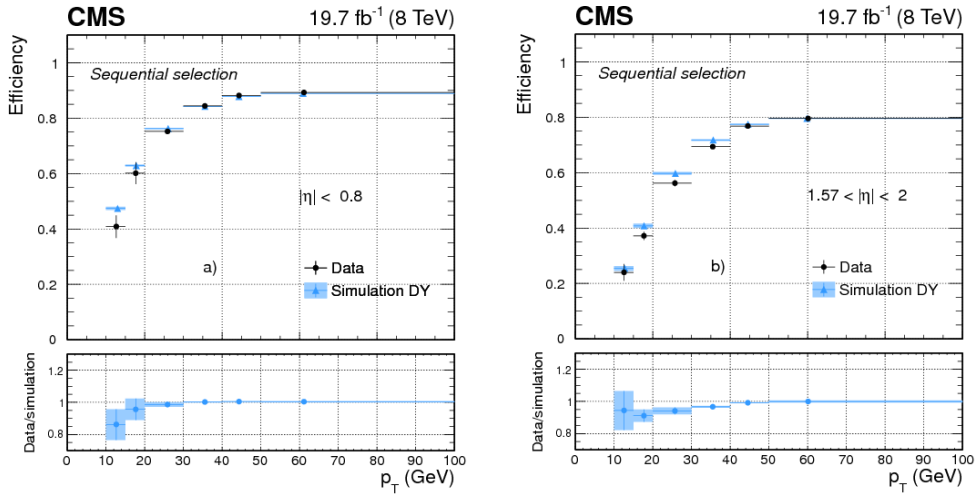


FIGURE 4.2: Efficiency as a function of electron p_T for dielectron events in data (dots) and DY simulation (triangles), for the medium working point of the sequential selection in a) $|\eta| < 0.8$, and b) $1.57 < |\eta| < 2$. The corresponding data-to-simulation scale factors are shown in the bottom panels of each plot.

.

4.2.3 Electron isolation

A significant fraction of background to isolated primary electrons is due to misidentified jets or to genuine electrons within a jet resulting from semileptonic decays of b or c quarks. In both cases, the electron candidates have significant energy flow near their trajectories, and requiring electrons to be isolated from such nearby activity greatly reduces these sources of background. The isolation requirements are separated from electron identification, as the interplay between them tends to be analysis-dependent. Moreover, the inversion of isolation requirements, independent of those used for identification, provides control of different sources of such backgrounds in data.

Two isolation techniques are used at CMS. The simplest one is referred to as detector-based isolation, and relies in the sum of the energy depositions in the ECAL and HCAL in the electron trajectory plus the scalar sum of the p_T of all tracks reconstructed from the collision vertex within a cone radius of $\Delta R = 0.3$ around the electron direction. This procedure, which has good performance in rejecting jets misidentified as electrons, is used by the HLT.

The second approach is using the particle flow [32] (PF) algorithm to define isolation quantities. Rather than using energy measurement, the isolation is defined using the PF candidates reconstructed with a momentum direction within some cone of isolation. In this way, the correct calibration of the energy can be used, and possible double-couting of energy is avoided.

The PF isolation is defined as:

$$I_{PF}^{rel} = \left(\sum_{ch} p_T + \max(0, \sum_{ne} E_T + \sum_{\gamma} E_T - p_T^{PU}) \right) / p_T^{ell} \tag{4.1}$$

where the sums run over the charged PF candidates, neutral hadrons and photons, within a chosen ΔR cone around the electron direction. The charged candidates are required to originate from the vertex of the event of interest, and p_T^{PU} is a correction related to event pile-up. The isolation is one of the most sensitive observables to the extra energy from pile-up interactions (either occurring in the same or earlier bunch crossings), which spoils the isolation efficiency when there are many interactions per bunch crossing.

The contribution from pile-up in the isolation cone, which must be subtracted, is computed using the average energy density per unit area, ρ, assuming $p_T^{PU} = \rho \cdot A_{eff}$. The dependence of ρ on pile-up is shown in Figure 4.3a along with the dependence of both the neutral and charge components of the isolation. The charged component remains independent of pile-up as only the candidates compatible with the primary vertex are considered. For both the ρ and the neutral component the dependence is almost linear. The effective area A_{eff} is then defined as by $(\Delta R)^2$, scaled by the ratio of the slopes for ρ and the specific isolation component. Once the correction is applied to the neutral components, the dependence with the number of vertices is much reduced, as it can be seen in Figure 4.3b. The plots refer to electrons with $|\eta| < 1$, but similar conclusions hold in any range of pseudorapidity.

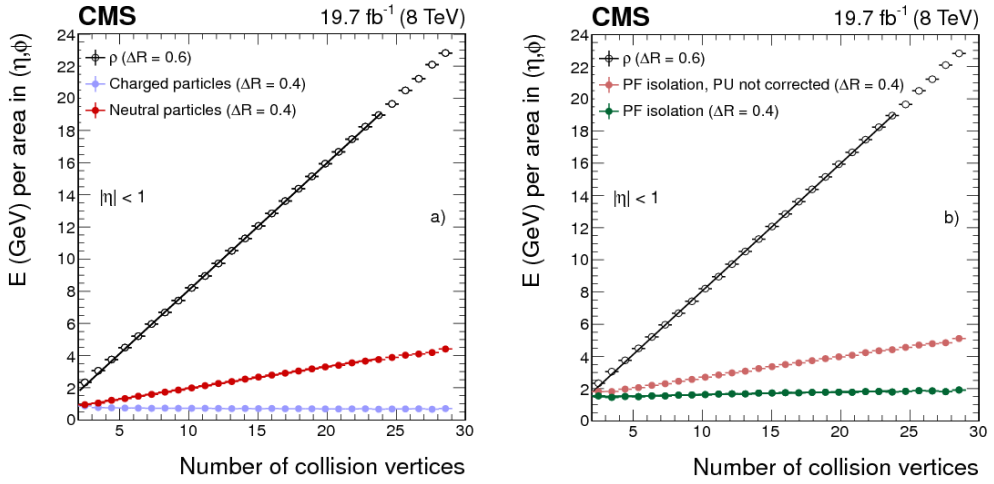


FIGURE 4.3: Average energy density as a function of the number of reconstructed proton-proton collision vertices, for electron candidates with $p_T > 20$ GeV and $|\eta| < 1$ from data dominated by $Z \to e^+e^-$ events. The energy density is shown a) for the event ρ (open dots) and separated components of the particle isolation: neutral particles (red dots) and charged particles associated with the vertex (blue dots), and b) after the correction for pile-up on PF isolation.

4.3 Muon Reconstruction

Muons are reconstructed using combined information from the tracker and the muon detectors. The first step is the reconstruction using the muon system alone, to produce track segments of full tracks which can be matched to the tracks reconstructed in the inner detector. The algorithm reconstruct first segments in the individual stations and then combine them to produce tracks.

Global muons (outside-in) For each standalone-muon track, a matching tracker track is found by comparing parameters of the two tracks propagated onto a common surface. A global-muon track is fitted combining hits from the tracker track and standalone-muon track, using the Kalman-filter technique [28]. At large transverse momenta, $p_T > 200$ GeV , the global-muon fit can improve the momentum resolution compared to the tracker-only fit.

Tracker Muon reconstruction (inside-out) In this approach, all tracker tracks with $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV are considered as possible muon candidates and are extrapolated to the muon system taking into account the magnetic field, the average expected energy losses, and multiple Coulomb scattering in the detector material. If at least one muon segment (i.e., a short track stub made of DT or CSC hits) matches the extrapolated track, the corresponding tracker track qualifies as a Tracker Muon. Track-to-segment matching is performed in a local (chamber) coordinate system, where local x is the best-measured coordinate (in the $r - \phi$ plane) and local y is the coordinate orthogonal to it. The extrapolated track and the segment are considered to be matched if the distance between them in local x is less than 3 cm or if the value of the pull for local x is less than 4, where the pull is defined as the difference between the position of the matched segment and the position of the extrapolated track, divided by their combined uncertainties

Tracker Muon reconstruction is more efficient than the Global Muon reconstruction at low momenta, $p < 5$ GeV, because it requires only one muon segment in the muon system, whereas Global Muon reconstruction is designed to have high efficiency for muon penetrating more than one muon station and typically requires segments in at least two muon stations.

Owing to the high efficiency of the tracker-track reconstruction and the very high efficiency of reconstructing segments in the muon system, about 99% of muons produced in pp collisions within the geometrical acceptance of the muon system and having sufficiently high momentum are reconstructed either as a Global Muon or a Tracker Muon,

and very often as both. Candidates found both by the Global Muon and the Tracker Muon approaches that share the same tracker track are merged into a single candidate.

4.3.1 Muon identification

The combination of different algorithms provides robust and efficient muon reconstruction. Physics analyses can set the desired balance between identification efficiency and purity by applying a selection based on various muon identification variables.

The tracker muon segment matching is a powerful tool to reject hadron punch-through, requiring matched segments in at least two stations, drastically reduces the probability for a punch-through to be identified as a muon as well as rejecting muons from light-quark decays. The impact parameter is a useful variable for identifying muons coming from the primary vertex or from heavy-flavour decays. The $\chi^2/d.o.f$ of the global muon fit is a good discriminant to suppress muons from decays in flight. Furthermore, particle flow algorithm [32] can be also used for identification and reconstruction.

A combination of these muon identification variables is used in this work and the detailed selection will be specified later in the analysis chapters.

An example selection could be defined as follows: the $\chi^2/d.o.f$ of the global muon track fit is required to be smaller than 10, as well as asking the muon candidate to have a hit in the muon chamber. To increase the purity of the selection, the corresponding tracker track is required to be matched to muon segments in at least two muon stations (in this way, the muon would be reconstructed both as Global and Tracker Muon). If the muon candidate has a certain number of tracker / pixel hits also increases the purity of the selected muons as well an impact parameter compatible with the primary vertex. This selection, is described as the "Tight-muon" identification [33] and reduces the rate of muons from decays in flight significantly, at the price of a few percent loss in efficiency loss for signal ("prompt") muons. The efficiency of the reconstruction and identification can be seen in Figure 4.4

4.3.2 Muon isolation

Requiring that a muon is isolated, meaning that the energy flow in its vicinity is below a certain threshold, is very useful when trying to distinguish prompt-muons, i.e. muons from decays of W or Z bosons, from muons produced from QCD processes.

Particle-Flow algorithm is used for computing the muon isolation. The discriminating variable is the sum of the p_T of all the charge hadrons, the transverse energies E_T of

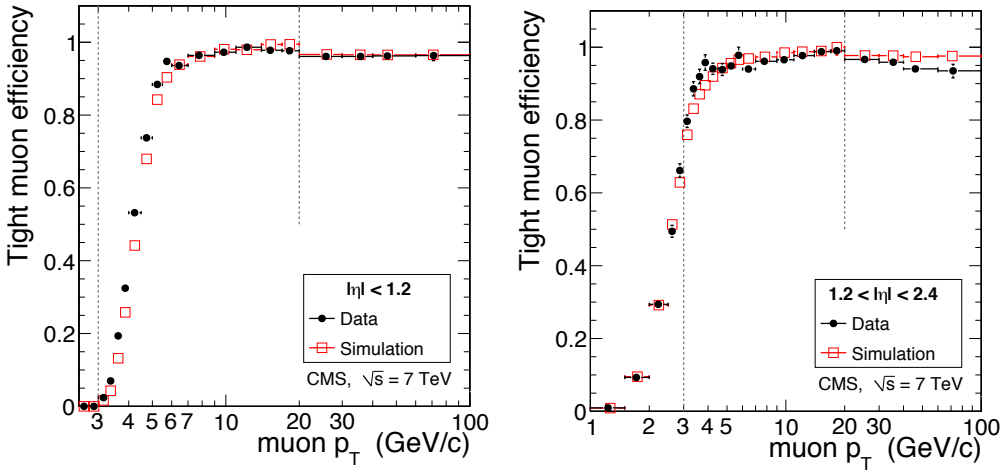


FIGURE 4.4: Reconstruction and identification muon efficiency in data compared to simulation. Given that a tracker track exists, the plots show the efficiency as a function of muon p_T for the example-selection discussed in the text in the barrel (left) and endcaps (right). The measurement is made using $J/\Psi \to \mu^+\mu^-$ for $p_T < 20$ GeV and $Z \to \mu^+\mu^-$ for $p_T > 20$ GeV.

all the photons and neutral particles reconstructed by the particle flow algorithm [32] within a cone of radius $\Delta R < 0.3$ centred on the muon-track direction, divided by the muon p_T .

$$I_{PF}^{rel} = \left(\sum_{ch} p_T + \sum_{ne} E_T + \sum_{\gamma} E_T \right) / p_T^{\mu} \tag{4.2}$$

The size of the cone and the cut value can and should be optimised for each analysis. Different strategies were used for 8 and 13 TeV analysis as it will be described later.

4.3.2.1 Pile-up corrections

The efficiency of any isolation algorithm features a dependence on the muon pseudo-rapidity, and this dependence is expected to become more pronounced as the number of pile-up collisions increases. Methods to mitigate the impact of pile-up on the performance of isolation algorithms have been developed. One such technique is based on the measurement, event-by-event, of the average transverse momentum per unit area ρ added to the event by minimum-bias pile-up collisions [34]. Another technique uses reconstructed tracks and primary vertices to compute a correction factor β to be applied to all or part of the numerator of the isolation variables described above. In both cases, the ρ and β variables allow the energy in the isolation cone due to particles produced in pile-up collisions to be estimated.

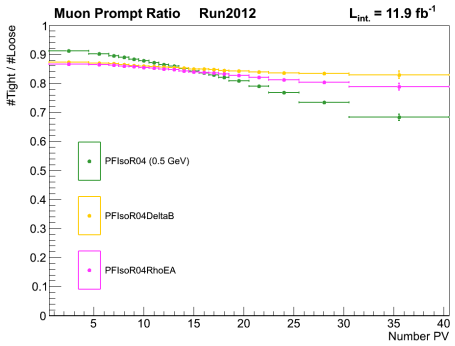


FIGURE 4.5: Dependence with the number of vertex of the various isolation algorithms (with and without correction). Green depitcs the non-corrected isolation algorithm, pink the correction using effective areas and yellow the correction using $\Delta\beta$ corrections.

The first approach uses the average transverse momentum per unit area, ρ, and the effective area of the cone, for correcting the isolation variable. The effective area is calculated as the ratio between the slope of the ρ and the isolation variable as a function of the number of vertices.

$$I^{rel}_{EA} = I^{rel}_{PF} - \rho \cdot A_{eff} \tag{4.3}$$

The second approach, uses the momenta of all charged particles not associated to the primary vertex to estimate the neutral contribution that otherwise is impossible to separate, as the charged component of a jet is estimated to be twice as big as the neutral part, only half of the energy is subtracted:

$$I^{rel}_{PF-\Delta\beta} = \left(\sum_{ch} p_T + \max(0, \sum_{ne} E_T + \sum_{\gamma} E_T - \frac{1}{2} \sum_{PUch} p_T) \right) / p^{\mu}_T \tag{4.4}$$

Figure 4.5 shows the significant reduction of the pile-up dependence after introducing the described corrections.

SUSY Muon POG contact and DQM developer Since 2013 I've been appointed as SUSY Muon POG contact person, being in charge of propagating the latest developments on muon identification and reconstruction to the SUSY group. Revising the usage of muons across the different as well as the measured isolation and identification efficiencies across several SUSY analysis. At the same time I have been in charge of the Offline Data Quality Monitoring (DQM) for muons.

4.4 Particle flow reconstruction

The particle flow event reconstruction [32] allows to reconstruct and identify all stable particle in the event,i.e., electrons, muons, photons, charged hadrons and neutral hadrons. An optimised combination of all the information from the different subdetectors is used to estimate direction, energy and type of each particle. This list of individual particles is then used, as if it came from a Monte-Carlo event generator, to build jets (from which the quark and gluon energies and directions are inferred), to determine the missing transverse energy $\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ (which gives an estimate of the direction and energy of the neutrinos and other invisible particles), to reconstruct and identify taus from their decay products, to quantify charged lepton isolation with respect to other particles, to tag b jets, etc.

4.5 Jet Reconstruction

Jet reconstruction at CMS is performed at particle level by clustering particles identified with the particle flow reconstruction. The usage of the particle flow approach provides a much better determination of the jet energy and direction resolution with respect to the traditional calorimetric-only approaches, as charged hadron and photons (that are reconstructed with excellent precision) account for about 90% of the energy in jets. High reconstruction efficiency for low momentum constituents is also achieved with this approach.

Particles are clustered into jets using the anti-k_T algorithm [35]. The main feature of this jet cluster algorithm is that it is infrared and collinear (IRC) safe, which means that it is robust against the presence of soft particles radiated by the partons and that it is capable of recombining collinear partons into the original one. This is crucial to have well defined theoretical calculation at parton level leading order.

Objects are clustered hierarchically, starting from the pair with smallest distance:

$$d_{ij} = \min\left(p_T^{-2}(i), p_T^{-2}(j)\right) \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2} \tag{4.5}$$

and clusters are then promoted to final jets when their p_T^{-2} is smaller than any remaining distance.

By using this algorithm, soft particles will tend to cluster with hard particles before clustering among themselves, as the distance is downscaled by the squared of the $\mathrm{p_T}$ of the hard object. Since the distance does not depend on the $\mathrm{p_T}$ of the soft object, this

algorithm produces conical jets of radius equal to the distance parameter, which introduces an experimental advantages simplifying the event interpretation and pile-up corrections.

As distance parameter, 8 TeV analysis use $R = 0.5$, while 13 TeV analysis use $R = 0.4$ to unify jet algorithms criteria with ATLAS collaboration.

4.5.1 Jet Energy Calibration

CMS uses a factorised approach to jet energy calibration [36]. The purpose of the jet energy calibration is to relate, on average, the energy measured for the detector jet to the energy of the corresponding true particle jet. A true particle jet results from the clustering of all stable particles originating from the fragmenting parton, as well as of the particles from the underlying event (UE) activity. The correction is applied as a multiplicative factor $\mathscr{C}$ to each component of the raw jet four-momentum vector p_{raw}:

$$p_{\mu}^{cor} = \mathscr{C} \cdot p_{\mu}^{raw} \tag{4.6}$$

The correction factor $\mathscr{C}$ is composed of the offset correction C_{offset}, the MC calibration factor C_{MC}, and the residual calibrations C_{rel} and C_{abs} for the relative and absolute energy scales, respectively.

The offset correction is the first in the chain of factorised correction, it allows to remove the contamination from pile-up and the underlying event, this is achieved by computing the per-event median energy density by using jet areas [34]. Then, the MC correction removes the bulk of the non-uniformity in η and the non-linearity in $\mathrm{p_T}$. Finally, the residual corrections account for the small differences between data and simulation. The various components are applied in sequence as described by the equation below:

$$\mathscr{C} = C_{offset}(p_T^{raw}) \cdot C_{MC}(p'_T, \eta) \cdot C_{rel}(\eta) \cdot C_{abs}(p''_T) \tag{4.7}$$

where p' is the $\mathrm{p_T}$ of the jet after applying the offset correction and p'' is the $\mathrm{p_T}$ of the jet after all previous corrections. In the following sections, each component of the jet energy calibration will be discussed separately.

4.6 Missing Transverse Energy

A good reconstruction of the missing transverse energy ($\mathrm{E_T^{miss}}$) is crucial when studying processes with invisible particles in the final state, such as neutrinos, but also the lightest supersymmetric particle (the lightest neutralino, or LSP). Missing transverse energy is defined as the negative vector sum of all the particles in the event:

$$\vec{\not{E}}_T = -\sum_{particles} \vec{p}_T \tag{4.8}$$

4.7 Identification of b-jets

Jets that arise from bottom quark hadronization and decay (b-jets) are present in a wide range of physics processes such as the decay of top quarks, Higgs bosons and various supersymmetric particles. The ability of accurately identifying b jets is vital to reduce the overwhelming background from other processes involving jets from gluons and light-flavour quarks. In this analysis, b-jets are extensively used to reduce the background contamination and enhance the sensitivity for various SUSY models.

A variety of reconstructed objects (tracks, vertices and identified leptons) can be used to build observables that discriminate between b and light jets. These observables are then used in various algorithms of different complexity. Each algorithm yields a single discriminator value for each jet. The minimum thresholds on these discriminators define loose ("L"), medium ("M"), and tight ("T") working points with a misidentification probability for light-parton jets of close to 10%, 1%, and 0.1%, respectively, at an average jet $\mathrm{p_T}$ of about 80 GeV.

The CMS b-jet identification starts from the reconstructed particle flow jets described in this section. High quality track are selected inside the jet requiring a transverse momentum of at least 1 GeV, at least eight hits associated to the track and a good track fit, with $\chi^2/ndof < 5$. Since track measurements in the vicinity of the interaction vertex contain most of the discriminating power at least two hits are required in the pixel system. A loose selection on the track impact parameters is used to further increase the purity and to reduce the contamination by decay products of long-lived particles like K_s^0. The transverse and longitudinal impact parameters, d_{xy} and d_z, are defined as the transverse and longitudinal distance to the primary vertex at the point of closest approach to the beam line. Their norms have to be smaller than 0.2 cm and 17 cm, respectively. At the point of closest approach the distance to the jet axis has to be smaller than 700 μm and this point has to be within 5 cm of the primary vertex. This

sample of associated tracks is the basis for all algorithms which use impact parameters for discrimination that are fully described in the reference [37].

4.7.1 Combined Secondary Vertex

In this work a discriminant built using secondary vertex information combined with track-based lifetime information was used for discriminate between b and non-b jets.

In order to enhance b purity, secondary-vertex candidates must fulfil the following requirements: share less than 65% of their associated track with the primary vertex and the significance of the radial distance between the two vertices has to exceed 3σ; if the radial distance is bigger than 2.5 cm with respect to the primary vertex, and the vertex mass is compatible with the mass of the K^0 or larger than 6.5 GeV the vertex-candidate is rejected to reduce the contamination by vertices corresponding to the interaction of particles with the detector and by decays of long-lived mesons; the flight direction of each vertex candidate has to be within a cone of $\Delta R < 0.5$ around the jet direction.

The *Combined Secondary Vertex* (CSV) is then computed using the following set of varaibles with high discriminating power and low correlations: the vertex category; the flight distance significance in the transverse plane; the vertex mass; the number of tracks at the vertex; the ratio of the energy carried by the tracks at the vertex with respect to all tracks in the jets; the pesudorapidities of the tracks at the vertex with respect to the jet axis; the 2D impact parameter significance of the first track that raises the invariant mass above the charm threshold of 1.5 GeV; the number of tracks in the jet; and the 3D impact parameter significance of each track in the jet.

Two likelihood ratios are built from these variables. They are used to discriminate between b and c jets and between b and light-parton jets. They are combined with prior weights of 0.25 and 0.75, respectively. The distribution of the CSV discriminator for a di-jet sample and a $t\bar{t}$ sample is depicted in Figure 4.6.

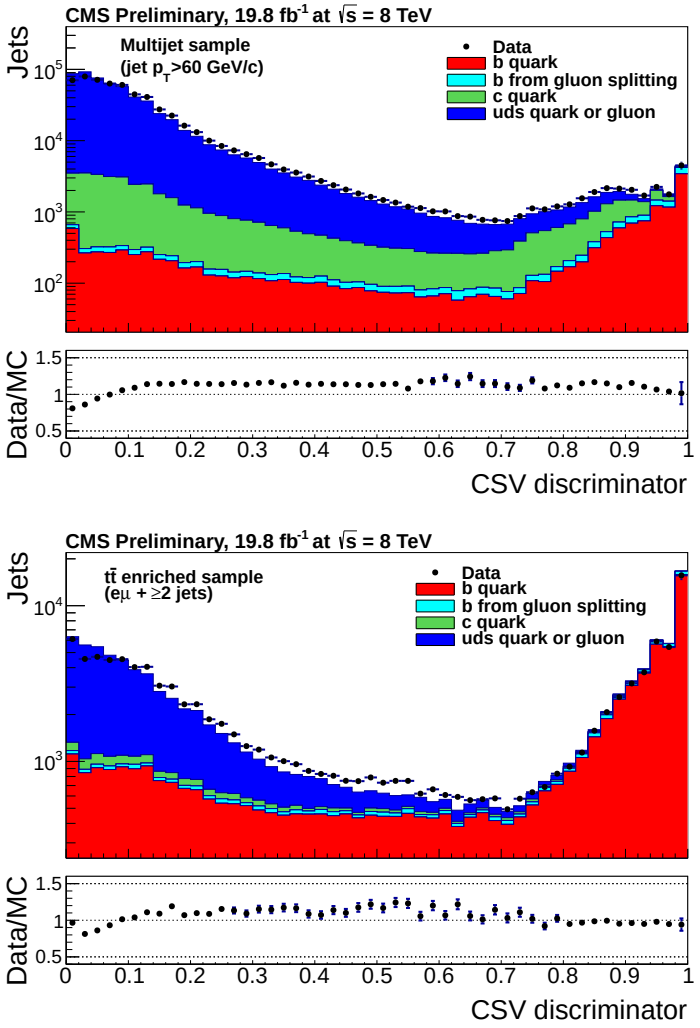


FIGURE 4.6: Shape of the CSV discriminator for the inclusive multijet sample (top) and for an enriched $t\bar{t}$ sample (bottom). Overflows are added to the last bin. The shape

Chapter 5

Same-Sign dilepton searches at $\sqrt{s} = 8$ TeV

5.1 Introduction

In the SM, proton-proton collision events having a final state with isolated leptons of the same sign are extremely rare, however such final states appear naturally in many new physics scenarios. These include supersymmetry [38–40], universal extra dimensions [41], pair production of $T_{5/3}$ [20], heavy Majorana neutrinos [21], and same-sign top-quark production [22, 23]. In SUSY, for example, same-sign dileptons occur naturally in the production of gluino pairs, when each gluino decays to a top quark and a top anti-squark, with the anti-squark further decaying into a top anti-quark and a neutralino.

The search for new physics in events with a pair of same-sign leptons (ee, $\mu\mu$ and $e\mu$) will be described in the following chapters. Two different search strategies are defined depending on the following considerations:

Strong production of SUSY New physics signal with larger cross section are likely to be produced by strong interactions and, thus, one expects significant hadronic activity accompanying the two leptons. Furthermore, astrophysical evidence for dark matter [42] suggests considering SUSY models with R-parity conservation, which provides an excellent dark matter candidate: a stable lightest supersymmetric particle (LSP) that escapes detection and that will be perceived in the detector as E_T^{miss}. Nevertheless, to achieve sensitivity to SUSY models with R-parity violation (RPV), signatures without a significant amount of E_T^{miss} will also be included. The results on this search targeting coloured production of SUSY are described in Chapter6.

Electroweak production of SUSY Heavy squarks and gluinos may favour models with direct electroweak production of charginos ($\widetilde{\chi}_1^{\pm}$) and neutralinos ($\widetilde{\chi}_2^0$), mixtures of the SUSY partners of gauge and Higgs bosons, and of sleptons ($\tilde{\ell}$). The smaller cross sections of direct electroweak SUSY production (see Figure 5.1) require dedicated searches targeting the wide variety of possible signal topologies. Depending on the mass spectrum, the charginos and neutralinos can have significant decay branching fractions to leptons or W, Z, and Higgs bosons (H), yielding final states with one, two or three isolated leptons. Similarly, slepton pair production gives rise to final states with two leptons. In all these cases, and under the assumption of R-parity conservation [43], two stable, LSP ($\widetilde{\chi}_1^0$) are produced, which are presumed to escape without detection, leading to significant missing transverse energy. The same-sign search described here will be combined with a three-lepton search to provide some sensitivity to these models. Very little hadronic activity is expected in such final states. Results on this search are described in Chapter 7.

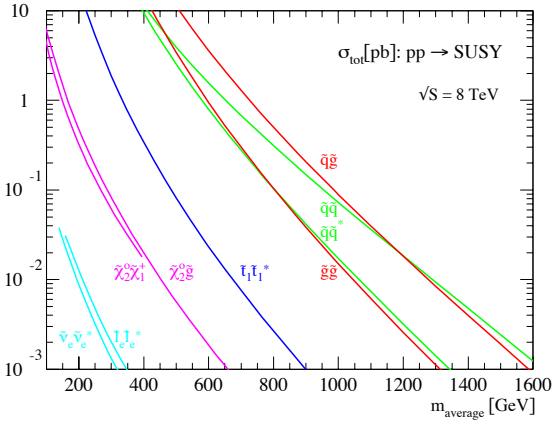


FIGURE 5.1: Production cross section for several SUSY processes at proton-proton collisions at a center of mass energy of 8 TeV.

The event selection and background estimation methods used in both searches are described in the following sections.

5.2 Event Selection and objects definition

The events used in these searches are selected using a set of dilepton triggers (two electrons, two muons or and electron and a muon), requiring the first (second) highest-p_T lepton to have $p_T > 17(8)$ GeV at the HLT. In the case of muons no further requirements are applied in the online selection, while for electrons, some basic identification and isolation cuts are applied to control the rate of the triggers paths.

In order to prevent any possible bias from the online selection, the offline selection has to be tight enough to ensure the trigger efficiency is close to 100%. Therefore, the looser offline selection of the objects is driven by the online selections in the trigger algorithms.

5.2.1 Muons

Muons are reconstructed in the pseudo-rapidity range between $-2.4 < \eta < 2.4$ and are required to have a transverse momentum greater than 20 GeV. The associated reconstructed track is required to have at least one hit in the pixel detector, more than 5 layers in the silicon tracker and at least one hit in one of the muon chambers. An extra requirement in the χ^2/N_{dof} of the global fit ensures the quality of the fit. To suppress muons from pile-up vertices and from heavy flavour decays, the reconstructed muon is required to have a transverse (longitudinal) impact parameter with respect to the selected primary vertex smaller than 50 μm (1 cm). In addition, the muon is required to be identified by the particle flow algorithm and by the GlobalMuon algorithm. Extra requirements on to veto ECAL and HCAL deposits are also required to reject background with minimal signal loss.

Finally, the muons are required to be isolated. This isolation, as explained in Section 4.3.2, is computed using the particle flow algorithm and corrected using the $\Delta\beta$ corrections to suppress the contamination from pile-up.

As prompt-muons are expected to be very isolated the requirement on the muon isolation is $Iso^{PF}_{rel} < 0.1$. That means that only 10% of the energy of the muon is allowed in the isolation cone. Nevertheless, as we will see later in Section 5.4, in some occasions this isolation criteria will be relaxed to $Iso^{PF}_{rel} < 1.0$. The muon object selection is summarised in table 5.1

TABLE 5.1: Muon object selection requirements

p_T	> 20 GeV
$\lvert\eta\rvert$	< 2.4
χ^2/N_{dof} (globalFit)	< 10
$\lvert d_0\rvert$	< 0.05 mm (measured w.r.t. first DA vertex)
$\lvert d_z\rvert$	< 1.00 cm (measured w.r.t. first DA vertex)
Number of Layers (Si track)	> 5
Number of valid StandAlone Hits	> 0
Number of valid Pixel Hits	> 0
`ecalVetoDep`	≤ 4 GeV
`hcalVetoDep`	≤ 6 GeV
Reconstructed with both `PFMuon` and `GlobalMuon` algorithms.	
Iso^{PF}_{rel}	0.1 (1.0)

5.2.2 Electrons

As for muons, electrons are reconstructed in the same kinematical range of $p_T > 20$ GeV and $|\eta| < 2.4$, excluding the pseudo-rapidity range of 1.444 - 1.566, to avoid the crack between the barrel and endcap modules.

To assess the quality of the selected electrons, and help discriminating between mis-reconstructed hadronic jets and real electrons, several selection cuts related to the shape of the energy deposit and the matching of the track and energy cluster are applied. The supercluster η width, as taken from the covariance matrix using logarithmic weights, is required to be smaller than 0.01 for electrons in the barrel ($|\eta| < 1.479$) and smaller than 0.03 for electrons in the endcap ($|\eta| > 1.479$). The distance between the electron track from the vertex to the supercluster position has to agree within $|\Delta\eta| < 0.004(0.007)$ and $|\Delta\phi| < 0.06(0.03)$ for barrel (endcap) electrons. Furthermore, the hadronic activity in HCAL cells behind the electrons supercluster has to be smaller than 10 (7.5) % of the supercluster energy. Finally, the quantity $|1/E - 1/p|$ has to be smaller than 0.05.

Electron pairs may be produced in interactions of hard photons with the detector material, such electrons will have longer transverse distances from the beam spot than prompt electrons. The first valid hit of the track for a photon converted electron may not necessarily be located in the innermost tracker layer, so extrapolating this track back to the beam line, one could find detector layers which do not have hits compatible with the track. In order to veto these converted electrons, a conversion vertex is performed with two tracks, one being compatible with an electron. The vertex fit probability is required to be greater than 10^{-6}. Also, the electron candidates must not have missing expected hits in their track.

To prevent electrons arising from pile-up events, the electron track is required to have transverse and longitudinal impact parameters with respect to the primary vertex smaller than 10 μm and 0.1 cm respectively.

Additionally, to reduce the charge mis-identification rate to a minimum (see Section 5.5 for further explanations), the measured charge has to be consistent among the three different algorithms used for charge assignment: the standard GSF electron tracking, the CTF tracking algorithm and the sign of the supercluster $\phi-$position with respect of the initial track direction.

To control the triggering rate, the HLT paths involving electrons apply rudimentary, detector-based cuts on isolation values that have to be reflected in the offline selection. These cuts are applied separately for the p_T sum of the tracks, ECAL energy deposits, and HCAL energy deposits, and restrict the sum of each of the contributions to be less

than 20% of the electron's transverse momentum. The loosest possible combined relative isolation cut for electrons is therefore at the level of 60% of the electron-p_T.

Finally, the electrons are required to be isolated. This isolation, as explained in Section 4.2.3, is computed using the particle flow algorithm and uses the effective area (EA) correction to suppress the contamination from pile-up.

As prompt-electrons are expected to be very isolated the requirement on the electrons isolation is $Iso_{rel}^{PF} < 0.09$. That means that only 9% of the energy of the electron is allowed in the cone. Nevertheless, as we will see later in Section 5.4, in some occasions this isolation criteria will be relaxed to match the trigger isolation. The full ID and isolation selection on signal electrons can be then found in Table 5.2.

TABLE 5.2: Electron object selection cuts

p_T	> 20 GeV
$\lvert\eta\rvert$	$< 2.4, \notin [1.444, 1.566]$
$\lvert d_0\rvert$	< 0.01 mm (measured w.r.t. first DA vertex)
$\lvert d_z\rvert$	< 0.10 cm (measured w.r.t. first DA vertex)
$\sigma_{i\eta i\eta}$	< 0.01 (barrel), < 0.03 (endcap)
$\lvert\Delta\phi_{In}\rvert$	< 0.06 (barrel), < 0.03 (endcap)
$\lvert\Delta\eta_{In}\rvert$	< 0.004 (barrel), < 0.007 (endcap)
H/E	< 0.01 (barrel), < 0.075 (endcap)
$\lvert 1/E - 1/p\rvert$	< 0.05 (barrel), < 0.05 (endcap)
Conversion rejection	
No missing pixel hits Vertex fit propability: $< 1 \times 10^{-6}$	
Charges of CTF, GSF and SuperCluster in agreement	
RelIso	< 0.09

As opposed to muons, where only the isolation criteria is relaxed to define "loose" object, in the case of electrons, the ID is relaxed as much as possible to gain statistics and selection cuts matching the trigger ID and isolation are imposed on the reconstructed electrons. The details on the cut variables and values can be found on Table 5.3.

Finally, in the rare cases where an electron candidate is within $\Delta R < 0.1$ of a muon passing all selection cuts, it is rejected to avoid selecting electrons that originate from a muon radiating a photon.

5.2.3 Jets and E_T^{miss}

Jets and missing transverse energy (E_T^{miss}) are reconstructed using the particle flow algorithm. The jet momenta is corrected to have a response that is flat in p_T and η

TABLE 5.3: Loose Electron selection

p_T	> 20 GeV
$\|\eta\|$	$< 2.4, \notin [1.444, 1.566]$
$\|d_0\|$	< 0.01 cm (measured w.r.t. first DA vertex)
$\|d_z\|$	< 0.20 cm (measured w.r.t. first DA vertex)
$\sigma_{i\eta i\eta}$	< 0.011 (barrel), < 0.031 (endcap)
$\|\Delta\phi_{In}\|$	< 0.15 (barrel), < 0.10 (endcap)
$\|\Delta\eta_{In}\|$	< 0.01 (barrel), < 0.01 (endcap)
H/E	< 0.10 (barrel), < 0.075 (endcap)
Charges of CTF, GSF and SuperCluster agree	
RelPFIso	< 1.0

(`L2L3`), as well as for contributions from pile-up events. (`L1FastJet`). In addition, `L2L3` residual corrections are applied on data.

The hadronic activity in the event is then characterised as the scalar sum of all selected jet transverse momenta:

$$H_T = \sum_i p_T(Jet_i)$$

The full Jet selection is detailed on Table 5.4.

TABLE 5.4: Jet selection

`ak5PFJets`
$p_T > 40$ GeV $\|\eta\| < 2.4$
`L1FastJet` corrected
`L2L3` corrected
`L2L3 residual` corrections in data
No selected lepton within cone of $\Delta R < 0.4$ Jets should pass the loose PF jet ID

Eventually we will require such jets to be identified as b-jets. For that we use the recommended combined-secondary-vertex (CSV) tagging algorithm with the medium working point for the discriminator of 0.679, which corresponding to an identification efficiency for b jets of about 70% and a misidentification probability of about 10% for charm jets and 0.1% for light-flavour jets (u,d,s and gluons).

For the E_T^{miss} Particle-Flow with `L1` corrections will be used. These corrections for the E_T^{miss} are applied to all jets with $p_T > 10$ GeV.

5.2.4 Basic event selection

The standard event selection for the following chapters will require two leptons (electrons and muons) with the same electric charge, resulting in three channels: ee, $e\mu$ and $\mu\mu$. By convention, priority to the number of muons is given when selecting the pair and then to the sum of the p_T of the objects. Both leptons will be required to have p_T greater than 20 GeV making the dileptonic trigger paths fully efficient. Finally, to suppress contributions from low-mass resonances with mis-reconstructed charges, the invariant mass of the selected pair is required to be greater than 8 GeV.

Events with a leptonically decaying Z boson, or with low-mass resonances are suppressed by applying a veto on events containing opposite-sign same-flavour pairs with an invariant mass within 15 GeV of the Z boson mass or smaller than 12 GeV. As the lepton selection described before is optimised for high background rejection, a looser object selection is applied for increasing the rejection power. The p_T requirement is dropped to 10(5) GeV and both identification and isolation requirements are also loosened, more details on the exact selection can be found in Tables 5.5 and 5.6.

TABLE 5.5: Veto muon selection

p_T	$> 10/5$ GeV for the Z/low-mass veto
$\lvert\eta\rvert$	< 2.4
RelPFIso	< 0.2
Recoed as `PFMuon` and (`GlobalMuon` || `TrackerMuon`)	

TABLE 5.6: Veto electron selection

p_T	$> 10/5$ GeV for the Z/low-mass veto
$\lvert\eta\rvert$	< 2.4, with $\lvert\eta_{SC}\rvert \notin [1.444, 1.566]$
$\lvert d_0\rvert$	< 0.49 mm (measured w.r.t. first DA vertex)
$\lvert d_z\rvert$	< 2.00 mm (measured w.r.t. first DA vertex)
$\sigma_{i\eta i\eta}$	< 0.01 (barrel), < 0.03 (endcap)
$\lvert\Delta\phi_{In}\rvert$	< 0.80 (barrel), < 0.70 (endcap)
$\lvert\Delta\eta_{In}\rvert$	< 0.007 (barrel), < 0.010 (endcap)
H/E	< 0.01 (barrel), $< \infty$ (endcap)
RelPFIso	< 0.20

5.3 Background Composition

Understanding the background composition is, together with a reliable estimation of the background processes, the most important aspect of every physics analysis. Specially

for searches, as only when understanding and properly estimating the background contribution to our search regions we would be able to be sensitive to any potential "signal" that may appear.

In a *pp* collision, leptons can be produced "promptly", in the decay of a W or a Z boson or "non-promptly", in the decay of heavy flavour hadrons or simply hadronic jets mis-reconstructed as leptons. In addition, other sources of non-prompt leptons could be muons from light-meson decays in flight, or electrons from unidentified photon conversions.

Events with two same-sign leptons in the final state are considered, the selected samples would be dominated by events with one prompt lepton and one non-prompt lepton, mostly from W + jets , Z + jets and semileptonic $t\bar{t}$ processes. A very important contribution will come from rare SM processes yielding prompt leptons of like-sign in the final state. Among these processes one can find direct production of dibosons (WZ or ZZ), direct production of two same-sign W bosons ($W^{\pm}W^{\pm}$), production of a $t\bar{t}$ pair in association of a vector boson ($t\bar{t}$ +W and $t\bar{t}$ +Z) and production of three or more vector bosons. In addition, the sample will contain some events with an opposite-sign pair of leptons (from Z + jets and dileptonic $t\bar{t}$) where one of the leptons has a wrongly assigned charge. This last process only affects significantly to electrons, where the charge mis-identification may arise from a hard bremsstrahlung and subsequent asymmetric conversion to an electron of the opposite charge. The charge assignment for muons is more reliable as they hardly interact with detector material and also thanks to the possibility of a second, independent measurement in the muon chambers.

Events with two fake or non-prompt leptons from QCD multijet production can play a role in event selections with low requirements on hadronic activity and missing transverse energy. However, the suppression achieved with isolation and identification cuts helps making this background almost negligible. Table 5.7 shows the contribution of different processes estimated from MC for a minimal selection of two same-sign lepton events.

Estimations from Monte Carlo simulation could be used to derive the contribution to the background. However, backgrounds involving non-prompt leptons and mis-identified charges are estimated using data-driven techniques extrapolating from control regions defined in data, as the simulation of the processes leading to their production is not always reliable. On the other hand, events with prompt-leptons can be simulated more reliably and the background yields will be taken directly from the Monte Carlo.

TABLE 5.7: Yields from different processes in simulation, compared to observed yields in data, for a selection of two same-sign leptons.

Process	$\mu\mu$	$e\mu$	ee	Sum
$t\bar{t}$	28.42 ± 2.01	58.40 ± 2.92	36.00 ± 2.31	122.82 ± 4.23
Single t	4.74 ± 20.17	7.24 ± 20.21	1.96 ± 20.15	13.93 ± 34.94
W+jets	31.58 ± 72.27	69.31 ± 67.09	104.07 ± 100.73	204.96 ± 140.96
Z+jets	0.00 ± 4.21	57.28 ± 12.67	228.31 ± 23.10	285.59 ± 26.68
WW	0.10 ± 0.13	1.54 ± 0.33	1.18 ± 0.30	2.82 ± 0.47
WZ	137.90 ± 1.06	269.67 ± 1.53	113.42 ± 1.00	520.98 ± 2.11
ZZ	9.40 ± 0.07	17.18 ± 0.10	6.55 ± 0.06	33.12 ± 0.14
Vγ+jets	35.88 ± 2.01	30.33 ± 1.94	0.56 ± 0.44	66.77 ± 2.83
$t\bar{t}$H	12.09 ± 0.47	23.62 ± 0.67	9.55 ± 0.43	45.26 ± 0.92
$t\bar{t}$W	17.94 ± 0.60	31.28 ± 0.81	12.61 ± 0.53	61.83 ± 1.14
$t\bar{t}$Z	3.61 ± 0.26	7.43 ± 0.37	3.33 ± 0.25	14.38 ± 0.52
$t\bar{t}\gamma$	0.00 ± 1.09	5.15 ± 2.18	2.61 ± 1.76	7.76 ± 3.01
$W^{\pm}W^{\pm}$	19.73 ± 0.84	35.25 ± 1.14	14.11 ± 0.75	69.09 ± 1.60
VVV	6.02 ± 0.21	11.59 ± 0.31	5.14 ± 0.23	22.74 ± 0.44
Sum	317.45 ± 75.23	641.01 ± 71.36	545.04 ± 105.35	1503.50 ± 147.81
Data	450	1051	896	2397

Lepton and trigger scale factors

The changing trigger conditions during Run I data-taking lead to a non-accurate simulated HLT-menu present in the MC samples that were used for this analysis. Trigger efficiencies are measured in data using a weakly correlated sample collected with HLT paths that are orthogonal to the ones used to the analysis described in Section 5.2. Trigger efficiencies are therefore computed per run period as a function of p_T and η to derive scale factors, Table 5.8 shows the computed scale factors that are then used to correct MC simulation.

TABLE 5.8: Scale factors from trigger inefficiencies applied to Monte Carlo predictions of the *irreducible* backgrounds. For some channels, scale factors are parametrised by the trailing lepton p_T or $|\eta|$. The quoted uncertainties are both statistical and systematic.

Channel	Scale Factor
ee, $p_T < 30$	0.92 ± 0.05
ee, $p_T > 30$	0.96 ± 0.06
$e\mu$	0.93 ± 0.06
$\mu\mu$, $\|\eta\| < 1$	0.90 ± 0.05
$\mu\mu$, $1.0 < \|\eta\| < 2.4$	0.81 ± 0.05

Offline lepton selection efficiencies are measured both in data and simulation using Z-boson events to derive simulation-to-data correction factors. A tag-and-probe method is

used for measuring these scale factors [33, 44, 45]. The "tag" lepton is required to pass a fairly tight trigger requirement along with all other ID and isolation requirements, while the "probe" leptons is required to pass a much looser selection. The efficiency is taken as the ratio of the counts of the probes passing the full lepton selection over the total number of probes. The contribution from Z events in data is extracted by fitting the invariant mass range 60 to 120 GeV simultaneously with various background and signal hypothesis. Due to the difference on the background kinematics the fit is performed on a bin-by-bin basis as a function of p_T and $|\eta|$.

The resulting correction factors, applied to simulation only, are detailed in Tables 5.10 and 5.9. The uncertainty of the total efficiency is 5% (3%) for electrons (muons). An additional systematic uncertainty is assigned to account for potential mismodelling of the lepton isolation efficiency due to varying hadronic activity in signal events. This uncertainty is 3% for all leptons except muons with $p_T < 30$ GeV, for which it is 5%. For the trigger efficiencies, a 6% systematic uncertainty is considered to absorb the maximum deviation among the bins from the central efficiency value.

TABLE 5.9: Electron efficiencies measured using the tag-and-probe method. The uncertainties are only statistical.

p_T , GeV		$20-30$	$30-40$	$40-50$	$50-200$
$\|\eta\|$	$0.0-0.8$	0.954 ± 0.002	0.960 ± 0.001	0.972 ± 0.001	0.969 ± 0.001
	$0.8-1.4442$	0.923 ± 0.003	0.935 ± 0.001	0.955 ± 0.001	0.956 ± 0.001
	$1.566-2.0$	0.921 ± 0.005	0.924 ± 0.002	0.950 ± 0.001	0.995 ± 0.002
	$2.0-2.4$	0.993 ± 0.004	0.959 ± 0.003	0.968 ± 0.002	0.969 ± 0.004

TABLE 5.10: Muon efficiencies measured using the tag-and-probe method. The uncertainties are only statistical.

p_T , GeV		$20-30$	$30-40$	$40-50$	$50-200$
$\|\eta\|$	$0.0-1.2$	0.964 ± 0.001	0.971 ± 0.000	0.978 ± 0.000	0.974 ± 0.001
	$1.2-2.4$	0.959 ± 0.001	0.978 ± 0.001	0.984 ± 0.000	0.977 ± 0.001

The yields derived from Monte Carlo are weighted by the corresponding simulation-to-data correction factors (or scale factors) described in this section. All the yields within this chapter are re-scaled using these numbers.

5.4 Backgrounds Involving Non-Prompt Leptons

Processes involving non-prompt leptons: either coming from heavy flavour decays, photon conversions or from mis-identified jets, constitute the dominant source of background in the same-sign dilepton searches presented here, most of these events come from W + jets , Z + jets or semileptonic $t\bar{t}$.

Although jets mis-reconstructed as leptons and non-prompt leptons are physically very different, they are both produced inside hadronic jets and a tight requirement in ID and isolation is crucial for rejecting such events.

The method described in the following subsection is based on a extrapolation on the isolation sideband and some identification variables, and hence, covers all backgrounds containing non-isolated leptons. To simplify the notation, I will refer to both fake and non-prompt leptons as fake leptons, as opposed to prompt leptons from on-shell electroweak boson decays.

After applying the object identification criteria, most of the remaining fake muons are the result of the decay of heavy flavour hadrons. A very small contribution is expected from "punch-through" muons, where a hadron penetrates the calorimeters and magnet to produce hits in the muon detection system from "decays-in-flight", where light or strange mesons decay to muons before reaching the calorimeter. Therefore, muons can be expected in all processes involving production of b and c quarks.

The electron case is more complicated. As electron reconstruction is based on single tracks pointing to energy deposits in the electromagnetic calorimeter, additional sources can produce fake electrons. They can be the result of simple mis-reconstruction of jets with high electromagnetic energy, e.g. from π^0 mesons with nearby tracks. Based on simulation, about 50% of fake electrons in $t\bar{t}$ are expected to be from heavy-flavour decays, about 30% from γ conversions, and about 20% from light jet fakes.

5.4.1 Method

The estimation of the fake lepton contribution is based on a *loose-to-tight* extrapolation in the isolation and identification variables. Hence, the two sets of lepton selection cuts were defined in Section 5.2: a *tight* set corresponding to the final object selection defined above and a *loose* selection with less stringent requirements on lepton ID and isolation. Note that the loose selection set includes the tight selection set, i.e. every tight lepton is by construction also a loose lepton.

The number of fake leptons passing the tight cuts over the number of fake leptons passing the loose cuts is called the *fake-ratio*, f. The equivalent ratio for prompt leptons is called the *prompt ratio* p. One has to make the assumption on the universality on these ratios between the signal and control regions, that is, assuming that the selection cuts used to defined signal and control regions do not change the shape of the variables used to define loose and tight objects.

In order to make this extrapolation,there must be some correlation between the tight/loose separation and prompt/fake categories. The problem can, therefore, be formulated as a set of linear equations relating the yields of prompt/fake events to yields of tight/loose, for making such extrapolation the probability for either category of leptons to pass/fail the tight cuts after passing the loose ones is used.

First, N_{TT}, N_{TL}, N_{LT} and N_{LL} are defined as the number of events with two leptons passing the tight selection, first one passes tight and second one passes loose, first one passes loose and second one passes tight, and both pass only the loose cuts respectively. The two subindices correspond to the two leptons, and their order identifies the two leptons in the event. For the case of same-flavour events the first index will correspond to the harder of the two; in the case of different-flavour events, the first index refers to the muon and the second to the electron. Hence $e\mu$ events in the N_{LT} category have an electron passing tight cuts and a muon passing loose, but not tight cuts. The N_{TT} region would be our signal region, or the region in which we want to determine the fake lepton background contamination. The event selections where one or two leptons fails the tight selection are exclusive sidebands of the signal region and it is used to extrapolate. We want to estimate the quantities N_{pp}, N_{pf}, N_{fp} and N_{ff}. Remember that this quantities are defined on top of the loose ID/isolation cuts, i.e. N_{pp} is the number of true prompt-prompt events with both leptons passing the loose selection. The notation of the extrapolation factors would be f_1, f_2 for the first and second fake leptons, and p_1, p_2 for the first and second prompt lepton. The set of equations can be written as:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} p_1p_2 & p_1f_2 & f_1p_2 & f_1f_2 \\ p_1(1-p_2) & p_1(1-f_2) & f_1(1-p_2) & f_1(1-f_2) \\ (1-p_1)p_2 & (1-p_1)f_2 & (1-f_1)p_2 & (1-f_1)f_2 \\ (1-p_1)(1-p_2) & (1-p_1)(1-f_2) & (1-f_1)(1-p_2) & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} \quad (5.1)$$

The matrix can be inverted to solve for the estimated prompt-fake yields on the right hand side:

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \frac{1}{(f_1 - p_1)(f_2 - p_2)} \begin{pmatrix} (1-f_1)(1-f_2) & -(1-f_1)f_2 & -f_1(1-f_2) & f_1 f_2 \\ -(1-f_1)(1-p_2) & (1-f_1)p_2 & f_1(1-p_2) & -f_1 p_2 \\ -(1-p_1)(1-f_2) & (1-p_1)f_2 & p_1(1-f_2) & -p_1 f_2 \\ (1-p_1)(1-p_2) & -(1-p_1)p_2 & -p_1(1-p_2) & p_1 p_2 \end{pmatrix} \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} \quad (5.2)$$

and thus the set of equations needed to estimate the yield of events with fake and prompt leptons using the measured fake and prompt ratios and the yields in the sidebands are obtained. Keep in mind that these yields are defined in a way that they sum up to the total number of events passing at least the loose-loose selection:

$$N_{tot} = N_{pp} + N_{pf} + N_{fp} + N_{ff} = N_{TT} + N_{TL} + N_{LT} + N_{LL} \quad (5.3)$$

5.4.2 Implementation

As the fake and prompt ratios depend on the lepton kinematics, the method is implemented to apply a per-event weight. Instead of using the integrated yields (N_{TT}, N_{TL}, etc), each event is associated with a weight, which depends on whether the leptons in the event pass the tight cuts (T) or only the loose ones (L). Each event has four weights associated, each one corresponding to the probability of originating from prompt-prompt, prompt-fake, fake-prompt, or fake-fake processes. Therefore, the fake estimation formulas from equation 5.2 are transformed into a set of sixteen weight formulas, as shown in table 5.11.

TABLE 5.11: Event-by-event weights associated in four tight-loose by four prompt-fake categories. f_i and p_i are the fake and prompt ratios evaluated for lepton i. All entries have to be divided by $(f_1 - p_1)(f_2 - p_2)$.

	TT	TL	LT	LL
w_{pp}/p_1p_2	$(f_1-1)(f_2-1)$	$(f_1-1)f_2$	$f_1(f_2-1)$	f_1f_2
w_{pf}/p_1f_2	$(f_1-1)(1-p_2)$	$-(f_1-1)p_2$	$f_1(1-p_2)$	$-f_1p_2$
w_{fp}/f_1p_2	$(1-p_1)(f_2-1)$	$(1-p_1)f_2$	$-p_1(f_2-1)$	$-p_1f_2$
w_{ff}/f_1f_2	$(1-p_1)(1-p_2)$	$-(1-p_1)p_2$	$-p_1(1-p_2)$	p_1p_2

Then the predictions for N_{pp}, N_{pf}, N_{fp} and N_{ff} are obtained by summing up all the corresponding weights for all events, e.g. $N_{pp} = \sum w_{pp}$ etc. Note that there is no distinction between tight-loose and loose-tight respectively between prompt-fake and fake-prompt for same-flavour events. For different-flavour events, the first index always

represents the muon and the second one represents the electron, meaning that e.g. TL is a tight muon and a loose electron while LT is the other way around.

5.4.3 Control Regions

For the application of the method as it was described in the previous section, one needs to measure the tight-to-loose ratio for fake and prompt leptons (electrons and muons), hence control regions to measure the ratios are defined. Ideally, the control region will be populated by events emulating the same environment as in the signal region, but with a very pure sample of fake or prompt leptons. Therefore, the extrapolation variables, in particular the lepton isolation, shall be the same for both the control and the signal regions. This ideal situation can be approximated by selecting samples of events dominated by QCD multijets for fakes and leptonic Z boson decays for prompt leptons. However one has to assume the the universality of the fake/prompt ratios.

Measurement of f

The fake-ratio f is measured in a QCD di-jet environment. Such sample can be obtained by selecting events with one lepton while reducing the contamination from sources of prompt-leptons, mainly W + jets and Z + jets due to their high cross-section. An upper cut on $\mathrm{E_T^{miss}}$ at 20 GeV and an upper cut on $m_T(lep, E_T^{miss})$ at 15 (20) GeV for muons (electrons) is enough to suppress events involving W bosons. In addition, events with any extra veto-leptons of $\mathrm{p_T}$ greater than 5 GeV are rejected suppressing the remaining Z + jets .

We further require the presence of one and only one jet, which is well separated from the lepton in the event, i.e. $\Delta\phi(lep, jet) > 2$ with the minimum jet-$\mathrm{p_T}$ at 65 GeV. In this control region, the presence of exactly one loose lepton is also required as the f ratio is calculated using such lepton, i.e. the probability of this lepton passing the loose criteria to also pass the tight criteria.

Definition of the control region The event selection described here was designed to minimise the electroweak contamination and to mimic the isolation distribution of the fake leptons with the selected data in the control region. The original cuts developed for the 7 TeV analysis [46] allowed too much contamination from electroweak processes and due to changes in the auxiliary triggers the isolation distribution in data was not matching the one from the fake leptons, as it can be seen in Figure 5.2.

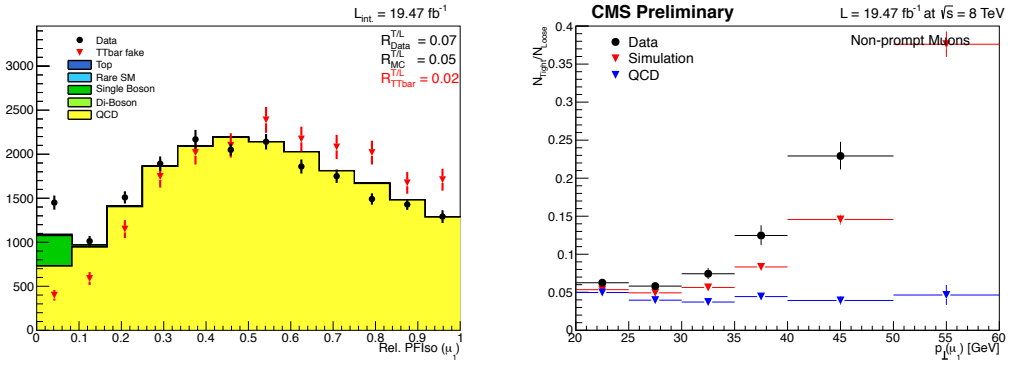


FIGURE 5.2: Isolation distribution of loose muons from the QCD control sample and comparison with the isolation distribution from fake muons from $t\bar{t}$ (left). Muon fake ratio calculated in the QCD control sample (right), a comparison between data, QCD and a MC cocktail is shown.

Since fake-rate shall be estimated from a sample that is very QCD-like, the reduction of the electroweak contamination should be aimed by purifying the sample in QCD-dijet-like events. The first step was to cut on any additional jets in the sample, i.e. requiring the number of jets in the control sample to be exactly one. Another handle on purifying in QCD dijet events is to introduce a $\Delta\phi$ cut between the loose lepton in the event and the jet in the event. One can see in Figure 5.3 the distributions for both those cuts.

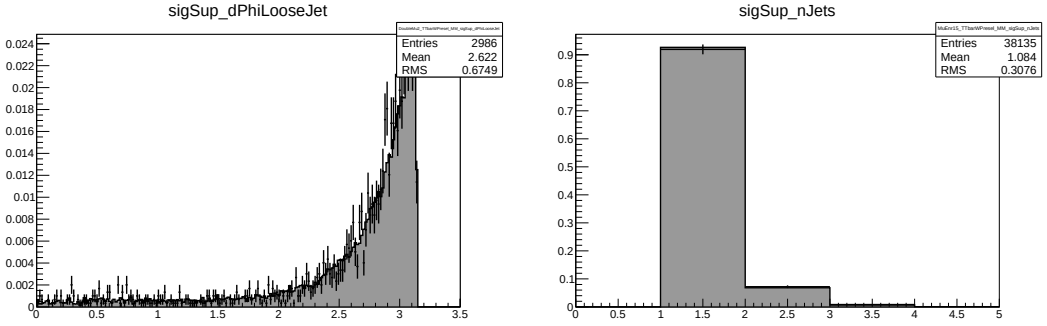


FIGURE 5.3: $\Delta\phi(lep, jet)$ (left) and N_{jets} distribution of the events in the QCD control sample. Most of the events have one jet and in such case it tends to be away from the lepton.

The next thing to be accomplished is to mimic, with the leptons in our control region, the behaviour of the isolation of the fake leptons coming from $t\bar{t}$, as this process is dominating our single-fakes sample. This was achieved by increasing the minimum jet-p_T from 40 GeV to 65 GeV and by tightening even more the m_T cut accepting only events with $m_T < 15$ GeV (for muons). The increase in the jet p_T threshold allowed to match the shape of the isolation distribution, however some discrepancy between data and MC is still present in the most-isolated bin. Tightening the m_T cut does help in

reducing that discrepancy, as the electroweak contamination that the MC is not able to reproduce is smaller in that case.

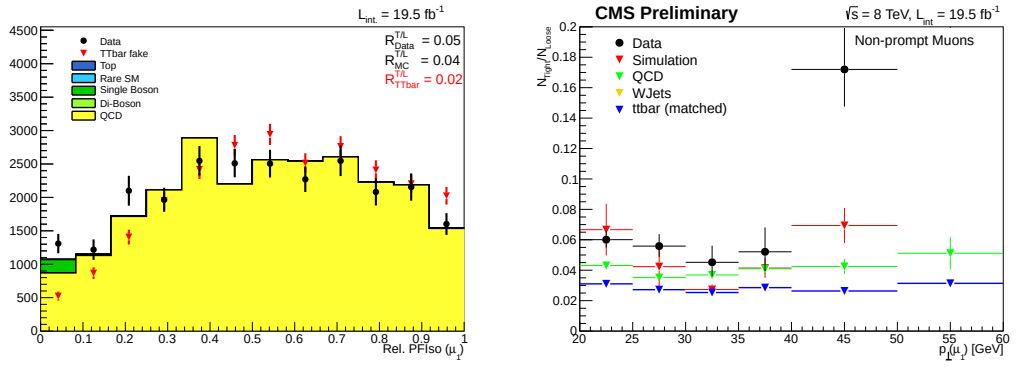


FIGURE 5.4: Isolation distribution of loose muons from the QCD control sample (with away-jet $p_T > 65$ GeV and $m_T < 15$ GeV) compared with the isolation distribution from fake muons from $t\bar{t}$ (left) and corresponding muon fake ratio (right)

For electrons, as it suffers more from poor statistics, and has less electroweak contamination, we decided to keep the m_T cut at 20 GeV. However we decided to apply the jet $p_T > 65$ GeV as it does influence the kinematics of the fakes and in that way the electron and muon sample is more consistent with each other.

Measurement of p

To measure the tight-to-loose ratio on prompt leptons, events with leptonic Z boson decays are used. Events containing two opposite-sign same-flavour loose leptons with an invariant mass within 15 GeV of the Z boson mass are selected. To further mimic the busy environment of the main same-sign background ($t\bar{t}$ + jets), and BSM signal, the presence of two hadronic jets is also required. Finally, to avoid contamination from fakes and non-prompt leptons from $t\bar{t}$ + jets production, an additional upper cut on the missing transverse energy of $E_T^{miss} < 20$ GeV is applied.

5.4.4 Fake Ratios

The data sample used for measure the fake ratio were collected using pre-scaled single-lepton auxiliary triggers. In the case of the muons, two paths are used, one requiring a single muon with $p_T > 8$ GeV and another one requiring a muon with $p_T > 17$ GeV (`HLT_Mu8_v*` and `HLT_Mu17_v*`). For the electron case, four paths are used requiring one electron with the same identification and isolation cut as the dilepton signal triggers (to avoid any possible bias in the measurement) and with a p_T threshold of 8 and 17

GeV, two of the paths also require the presence of an extra jet with $p_T > 30$ GeV. The list of used paths during 8 TeV data-taking is summarised in Table 5.12. All these paths are heavily pre-scaled (i.e. only accepts one out of N selected events).

TABLE 5.12: HLT paths used to select the QCD-enriched measurement region.

Muon	Electron
HLT_Mu8_v*	HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
HLT_Mu17_v*	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
-	HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Jet30_v*
-	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Jet30_v*

The fake ratios for muons as functions of p_T and $|\eta|$ are shown in Fig. 5.5, compared to values obtained from a cocktail of Standard Model MC events. The upward slope present in the fake ratio as a function of the lepton p_T is attributed to prompt lepton contamination from leptons stemming from W + jets events passing the inverted E_T^{miss} and m_T cuts. As there is no further electroweak suppression, the p_T spectrum is cut off at 40 GeV, i.e. every lepton which has a p_T higher than 40 GeV will be weighted by the ratio from the bin up to 40 GeV.

An electroweak subtraction method has been developed based on the normalisation of the m_T shape after inverting the m_T cut. This method is not applied to the search presented here but will be used for future iterations of this analysis.

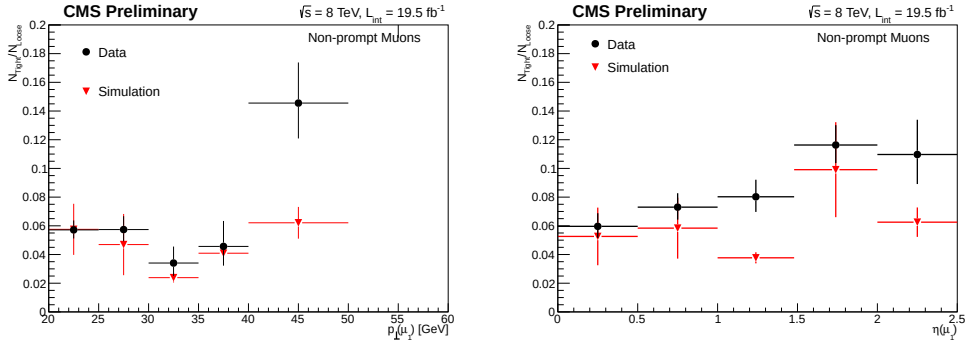


FIGURE 5.5: Fake ratios for muons as a function of p_T (left) and $|\eta|$ (right). The red markers depict the values obtained from a cocktail of Standard-Model Monte-Carlo samples including QCD, top and electro-weak processes.

Despite the dependence on lepton kinematics, no other significant dependence has been noticed. Observables monitored are: the hadronic activity, the missing transverse energy, the number of vertices in the event to check the dependency with pile-up and the p_T of the jet (see Figure 5.7 for muons and Figure 5.8 for electrons). As in previous plots, the

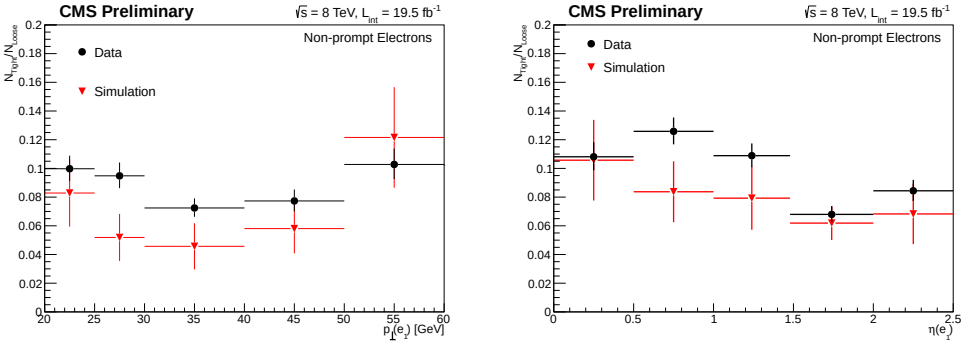


FIGURE 5.6: Fake ratios for electrons as a function of p_T (left) and $|\eta|$ (right). The red markers depict the values obtained from a cocktail of Standard-Model Monte-Carlo samples including QCD, top and electro-weak processes.

values observed in a cocktail of SM Monte-Carlo samples is shown for reference. The agreement with data is good up to the statistical precision of these plots.

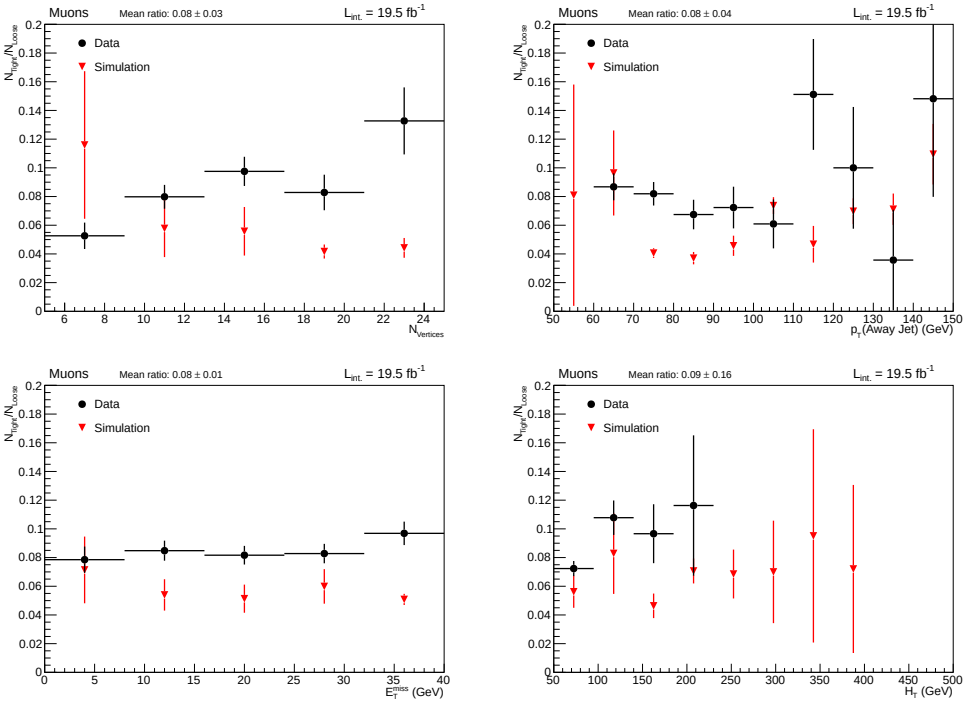


FIGURE 5.7: Fake ratios for muons as a function of the number of vertices (top left), away jet p_T (top right), E_T^{miss} (bottom left) and H_T (bottom right). The red markers depict the values obtained from a cocktail of Standard-Model Monte-Carlo samples including QCD, top and electro-weak processes.

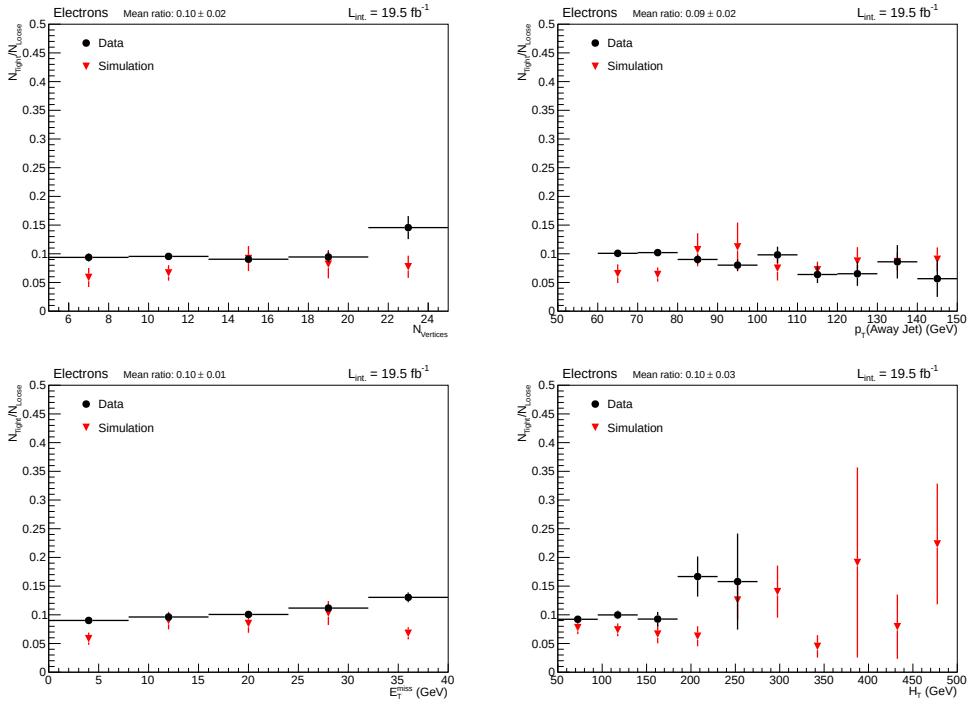


FIGURE 5.8: Fake ratios for electrons as a function of the number of vertices (top left), away jet p_T (top right), E_T^{miss} (bottom left) and H_T (bottom right). The red markers depict the values obtained from a cocktail of Standard-Model Monte-Carlo samples including QCD, top and electro-weak processes.

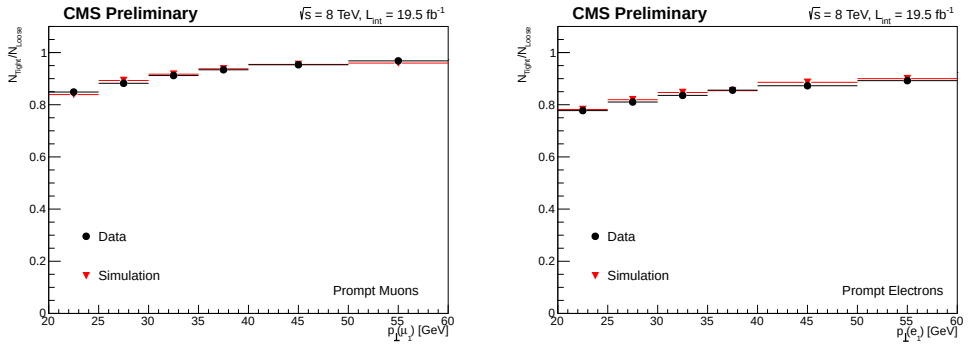


FIGURE 5.9: Prompt ratios for muons (left) and electrons (right) as functions of their p_T. The red markers depict the values obtained from a cocktail of Standard-Model Monte-Carlo samples including QCD, top and electro-weak processes.

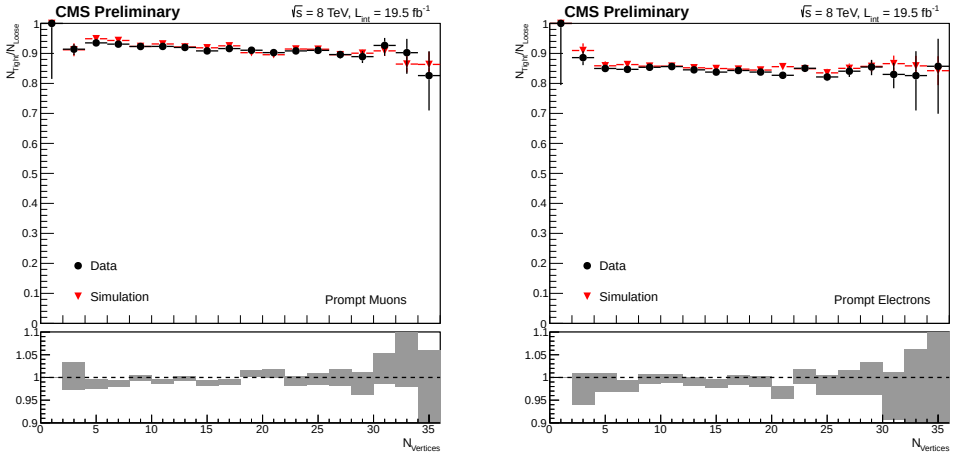


FIGURE 5.10: Prompt ratios for muons (left) and electrons (right) as functions of the number of primary vertices. The red markers depict the values obtained from a cocktail of Standard-Model Monte-Carlo samples including QCD, top and electro-weak processes.

5.4.5 Prompt Ratios

The HLT paths used for the prompt control regions are the same as those used for the signal regions, i.e. the dileptonic triggers. The ratio is measured on the softer leg of the pair. The resulting tight-loose ratios for prompt leptons are shown in Fig. 5.9, as a function of lepton p_T, together with the curve obtained from simulated events. The ratios, overwhelmingly dominated by events from the Z + jets process, as expected, are well reproduced in the simulation. Fig. 5.10 shows the prompt ratio as a function of the number of vertices.

5.5 Background from Charge Misidentification

Opposite-Sign dilepton events from standard-model processes, like (fully-leptonic) $t\bar{t}$ + jets, Z + jets and $W^{\pm}W^{\mp}$ can contribute to the same-sign yields in the case where one of the leptons has a badly reconstructed charge. This affects primarily high-momentum tracks with a small curvature that appear nearly straight in the detector. However as the lepton p_T range of this analysis is way below this point, the only relevant contribution to this background comes from electrons and occurs when an electron radiates a hard photon that then converts in the inner tracker layers into an e^+e^- pair where one of the electrons takes most of the momentum.

Studies from cosmic muons have shown that the probability of misidentifying the charge of a muon is significantly lower than in the electron case, due to the longer lever arm of

measuring the muon track and less interactions of the muon with the detector material. At higher p_T the situation would be different, but the two main processes contributing to the same-sign yield via charge mis-identification ($t\bar{t}$ + jets and Z + jets) rarely produce leptons with $p_T > 100$ GeV. The background contribution from muon charge mis-assignment will therefore be neglected.

The contribution to the final yields from this background can be directly estimated from a simple extrapolation from the opposite-sign yields to the same-sign signal region.

5.5.1 Estimation method

The contribution from intrinsic opposite-sign dilepton pair to the opposite-sign and same-sign dilepton yields can be estimated as:

$$N^{OS} = ((1-p)^2 + p^2) \cdot N^{OS}_{true} \quad (5.4)$$

$$N^{SS} = 2p(1-p) \cdot N^{OS}_{true} \quad (5.5)$$

where p is the probability of electron charge misidentification, N^{OS} is the number of opposite-sign dilepton pair events, N^{SS} is the number of same-sign dilepton pairs and N^{OS}_{true} is the intrinsic opposite-sign dilepton pairs.

Removing the unknown N^{OS}_{true}, the contribution to the same-sign yields from opposite-sign events with an electron charge misidentified can be estimated for both the ee and $e\mu$ channel as:

$$\begin{aligned} N^{SS}_{ee} &= (2p(1-p))/((1-p)^2 + p^2) \cdot N^{OS}_{ee} \\ &\approx 2p \cdot N^{OS}_{ee} \quad (5.6) \\ N^{SS}_{e\mu} &= p/(1-p) \cdot N^{OS}_{e\mu} \\ &\approx p \cdot N^{OS}_{e\mu} \quad (5.7) \end{aligned}$$

5.5.2 Measurement of the charge misidentification probability

The probability of misidentifying the charge of an electron can be extracted from Z events. The method relies in selecting di-electron events with an invariant mass compatible with a Z boson and classify them into same-sign and opposite sign pairs.

Such sample is obtained by selecting events with two electrons passing all the identification and isolation criteria defined in Table 5.2 with an invariant mass between 76 and 106 GeV. In addition an upper cut in the $E_T^{miss} < 30$ GeV and m_T (e, E_T^{miss}) < 25 GeV is also applied to suppress background contributions from W + jets processes.

TABLE 5.13: Per-electron charge misidentification probability, separately in barrel-barrel, endcap-barrel and endcap-endcap pairs.

	Barrel - Barrel	Barrel - Endcap	Endcap - Endcap
Data	0.000063 ± 0.000007	0.00027 ± 0.00002	0.00061 ± 0.00008
MC	0.000060 ± 0.000007	0.0002 ± 0.00002	0.00057 ± 0.00007

The measurement is performed separately for *ee* pairs in barrel-barrel, barrel-endcap, and endcap-endcap , and the results are summarised in Table 5.13. Due to the larger material budget between the interaction point and ECAL for forward electrons, the probability for the endcap sections is higher by an order of magnitude. The probability also rises towards higher p_T as explained at the beginning of the section. However, over the relevant range of electron p_T 's in Z boson decays and in $t\bar{t}$ + jets events the change is not dramatic, it will be neglected for the application of the method. Nevertheless, this decepence will be taken into account when estimating the systematic uncertainties for this background source.

5.6 Irreducible Standard-Model Backgrounds

In addition to backgrounds involving fake leptons and charge misidentified electrons there are also a number of Standard Model processes that can contribute to the same-sign signature with the production of real prompt-prompt leptons of equal charge. Since these yields do not depend on the difficult simulation of the production mechanisms of fake leptons in hadronic showers, their contributions is estimated from Monte Carlo simulation.

All Monte Carlo samples are simulated with the MADGRAPH5 [47] program when available, Pythia 6 [48] is used otherwise. The detector response to the generated events is simulated with the CMS simulation framework based on GEANT4 [49], and the events are reconstructed with the same software used to process the data. Parton showering, hadronization, and the decay of particles, are described using PYHIA 6 [48]. Simulated minimum bias interaction, generated with PYTHIA, are overlaid on the hard scatter events to closely emulate the multiple pp interactions present in the current and in

adjacent bunch crossings (pileup) present in the data. The Monte Carlo samples are re-weighted to fully reproduce the pileup profile in data. The samples are also re-weighted to match trigger and lepton efficiencies with the conditions in data. Table 5.14 displays the generator used for each SM process together with the production cross-section used for normalisation.

TABLE 5.14: Table of Standard Model MC samples used, with the highest order available cross-sections in picobarns.

Process	Generator	$\sigma(pb)$
$W\gamma^* \to \mu$	MADGRAPH	1.914
$W\gamma^* \to \tau$	MADGRAPH	0.336
$W^{\pm}W^{\mp} \to l^{\pm}\nu l^{\mp}\nu$	MADGRAPH	5.8123
$W^{\pm}Z \to l^{\pm}\nu l^{\pm}l^{\mp}$	MADGRAPH	1.0575
$ZZ \to l^{\pm}l^{\mp}l^{\pm}l^{\mp}$	MADGRAPH	0.1769
$t\bar{t}H, WH, ZH(H \to W^+W^-)$	PYTHIA	0.2604
$t\bar{t}H, WH, ZH(H \to ZZ)$	PYTHIA	0.0320
$t\bar{t}H, WH, ZH(H \to \tau^+\tau^-)$	PYTHIA	0.0177
$t\bar{t}W$	MADGRAPH	0.232
$t\bar{t}Z$	MADGRAPH	0.2057
$t\bar{t}\gamma$	MADGRAPH	2.166
$tbZ \to llll$	MADGRAPH	0.0114
$W^{\pm}W^{\pm}jj$	PYTHIA	0.5879
WWZ	MADGRAPH	0.05795
WZZ	MADGRAPH	0.01968
WWG	MADGRAPH	0.528
ZZZ	MADGRAPH	0.005527
WWW	MADGRAPH	0.08058
$t\bar{t}WW$	MADGRAPH	0.002037
W^+W^+	MADGRAPH	0.2482
W^-W^-	MADGRAPH	0.08888

Chapter 6

Search for new physics in events with same-sign dileptons and jets in pp collisions at $\sqrt{s} = 8$ TeV

The search described in this chapter is based on proton-proton collision data at $\sqrt{s} =$ 8TeV collected during 2012 with the CMS detector at the LHC, corresponding to an integrated luminosity of 19.5 fb^{-1} . This analysis was firstly developed for 7 TeV [46, 50] and several extensions and updates were performed during 2012 data taking period [51, 52].

The event selection, and background estimation methods have been described in Chapter 5 and the search strategy and the results will be discussed here.

6.1 Search strategy

The search strategy is a simple cut and count method, that is comparing the number of observed events with the expectation from SM background in different search regions (SR) with different requirements on four discriminating variables: H_T , $\text{E}_\text{T}^\text{miss}$, N_jets and $N_{\text{b-jets}}$. Search regions are defined in bins of the number of jets and b-tagged jets providing broad coverage of strongly produced SUSY particles, including signatures with low hadronic activity as well as signatures involving third-generation squarks. Additionally, as SUSY models with a small mass splitting between the parent sparticle and the LSP may result in low $\text{E}_\text{T}^\text{miss}$, we also define search regions with a looser requirement on $\text{E}_\text{T}^\text{miss}$.

Three different "baseline regions" (BSR0, BSR1 and BSR2) are define with looser event selection requirements, yielding to a background-dominated sample with negligible signal contamination. These will be used to understand the background composition as well as validate the background estimation methods that will be used to predict SM backgrounds. For defining these baseline regions (BSR) events having at least two jets, $H_T > 80$ GeV, $E_T^{miss} > 30$ GeV (only if $H_T < 500$ GeV) and 0 (BSR0), 1 (BSR1), 2 or more (BSR2) b-jets are selected. This selection is summarized in Table 6.1.

TABLE 6.1: Definition of the baseline regions.

Region	H_T (GeV)	E_T^{miss} (GeV)	N_{jets}	$N_{b\text{-jets}}$
BSR0	> 80	> 30 if $H_T < 500$ GeV	≥ 2	$N_{jets} \geq 0$
BSR1	> 80	> 30 if $H_T < 500$ GeV	≥ 2	$N_{jets} = 1$
BSR2	> 80	> 30 if $H_T < 500$ GeV	≥ 2	$N_{jets} \geq 2$

On top of these baseline regions, the event selection criteria are tightened and granularity of the regions is increased to define 24 final SRs described in Table 6.2. All 24 signal regions are mutually exclusive and may therefore be statistically combined. Additional (overlapping) signal regions, listed in Table 6.3, are defined with no or loose E_T^{miss} requirement in order to provide better sensitivity to scenarios such as RPV SUSY models and same-sign top-quark production. Because in RPV SUSY scenarios the LSP decays, mainly into detectable leptons and quarks, such events are not expected to have large E_T^{miss}, but they usually have substantial H_T. Thus, in search regions designed for such models, the E_T^{miss} requirement is removed completely, while a relatively high $H_T > 500$ GeV requirement is applied to reduce the level of SM background. These search regions are labeled as RPV0 and RPV2 for $N_{b\text{-jets}} \geq 0$ and ≥ 2, respectively.

TABLE 6.2: Signal region definitions and nomenclature.

$N_{b\text{-jets}}$	E_T^{miss}	N_{jets}	H_T [200-400]	H_T [> 400]
≥ 0	50-120	2-3	SR01	SR02
		≥ 4	SR03	SR04
	> 120	2-3	SR05	SR06
		≥ 4	SR07	SR08
$= 1$	50-120	2-3	SR11	SR12
		≥ 4	SR13	SR14
	> 120	2-3	SR15	SR16
		≥ 4	SR17	SR18
≥ 2	50-120	2-3	SR21	SR22
		≥ 4	SR23	SR24
	> 120	2-3	SR25	SR26
		≥ 4	SR27	SR28

TABLE 6.3: Signal regions for same-sign top production and RPV SUSY.

N_{jets}	$N_{b\text{-}jets}$	E_T^{miss}	H_T	name	comment
≥ 2	≥ 2	30	80	SR30	++/–
≥ 2	≥ 2	30	80	SR31	++ only
≥ 2	==1	30	80	SR34	++/–
≥ 2	==1	30	80	SR35	++ only
≥ 2	≥ 0	0	500	SR32	RPV1
≥ 2	≥ 2	0	500	SR33	RPV2

6.2 Validation of the backgrounds estimation methods.

Three different sources of backgrounds affect this search as it was described before: "Non-Prompt leptons (Fakes)", "rare SM processes" and "charge mis-identification". A detail description of these backgrounds and the estimation techniques applied in each case is found in Section 5.3.

Prior to apply the background prediction techniques into the signal regions, the performance of these background estimation techniques needs to be assessed. The background dominated baseline signal regions will be used for this purpose.

First, a closure test on MC simulation by deriving the full fake-ratio from the cocktail of Monte Carlo samples is perform. After deriving this fake-ratio, it is applied again to the full cocktail of MC samples to get an estimate for the fake and charge misidentification. The relative difference between the prediction and the observation is of the order of 20 to 30 percent, as can be seen from on Table 6.4. Note that due to limited statistics, especially in the $W - jets$ sample, the errors on the numbers of this closure test are naturally very high.

This discrepancy can be caused by the difference between the parton and the lepton p_T. This difference could, theoretically, be taken into account by computing the fake ratios as a function of the underlying parton p_T, however, it is not obvious how this dependence could be measured, as the determination of the underlying parton p_T is very laborious and suffers from poor resolution. As a result, a conservative 50% is assigned as a flat systematic uncertainty to the fake-lepton background estimation. More sophisticated methods to try to overcome this problem have been developed for 13 TeV analysis and they will be discussed later in the following chapters.

The performance of the background estimation methods is also validated in data using the baseline signal regions defined in Table 6.1, with negligible signal contamination and dominated by background events. The results of the background estimation methods and the yields of the observed events from data for the baseline signal regions for incluse,

TABLE 6.4: Closure test with ratio from MC-cocktail and closure done on MC-cocktail.

	$\mu\mu$	$e\mu$	ee	Sum
Fakes	43.33± 5.87	89.79± 7.68	52.19± 8.72	185.32± 13.02
Charge MisID		3.02± 0.22	10.07± 0.49	13.09± 0.54
Irreducible	36.89± 1.65	72.97± 2.89	31.37± 2.27	141.23± 4.03
WZ Production	17.24± 0.47	37.03± 0.67	18.69± 0.48	72.96± 0.95
Total Pred.	97.46± 6.12	202.81± 8.24	112.31± 9.04	412.59± 13.67
Observed	71.67± 56.92	154.60± 56.98	89.46± 57.03	315.72± 98.68
Pred./Obs.	1.36± 0.99	1.31± 0.43	1.26± 0.70	1.31± 0.37
Pred.-Obs./Pred	0.26± 0.59	0.24± 0.28	0.20± 0.51	0.23± 0.24

1 b-tag and 2 or more b-tags are shown in Tables 6.5, 6.6 and 6.7. Good agreement between prediction and observed yields and data is seen in all three baseline regions.

TABLE 6.5: Summary of background predictions and observed yields in SR00

	$\mu\mu$	$e\mu$	ee	Total
Double Fakes	0.43 ± 0.25	1.98 ± 1.02	2.31 ± 1.18	4.72 ± 2.39
Single Fakes	48.94 ± 24.81	109.73 ± 55.02	49.34 ± 24.98	208.02 ± 104.25
Charge MisID	-	2.96 ± 0.26	12.38 ± 0.72	15.34 ± 0.77
Rare SM	37.15 ± 18.64	73.63 ± 36.90	31.81 ± 16.03	142.58 ± 71.38
WZ Prod.	17.24 ± 2.61	37.04 ± 5.59	18.70 ± 2.83	72.98 ± 10.98
Total Bkg	103.77 ± 31.18	225.33 ± 67.30	114.54 ± 30.66	443.63 ± 128.70
Observed	111	220	146	477

TABLE 6.6: Summary of background predictions and observed yields in SR10

	$\mu\mu$	$e\mu$	ee	Total
Double Fakes	0.15 ± 0.11	0.45 ± 0.27	0.32 ± 0.19	0.92 ± 0.50
Single Fakes	22.40 ± 11.54	46.75 ± 23.53	19.39 ± 9.97	88.54 ± 44.50
Charge MisID	-	1.39 ± 0.12	1.98 ± 0.12	3.36 ± 0.17
Rare SM	12.11 ± 6.18	26.48 ± 13.42	11.36 ± 5.92	49.95 ± 25.15
WZ Prod.	1.26 ± 0.22	2.77 ± 0.45	1.50 ± 0.26	5.53 ± 0.86
Total Bkg	35.92 ± 13.01	77.84 ± 27.28	34.55 ± 11.62	148.31 ± 51.46
Observed	35	75	42	152

For the non-prompt background estimation three different estimates where combined together for producing the results on the paper [52]. For this combination, the BLUE method [53] was used. BLUE stands for "Best Linear Unbiased Estimator" and it corresponds to an error weighted combination of the three results with correlations taken into account. The correlations are only affecting the statistical part of the total

TABLE 6.7: Summary of background predictions and observed yields in SR20

	$\mu\mu$	$e\mu$	ee	Total
Double Fakes	-0.00 ± 0.01	0.04 ± 0.05	0.03 ± 0.04	0.07 ± 0.07
Single Fakes	2.02 ± 1.32	7.22 ± 3.76	3.06 ± 1.79	12.30 ± 6.36
Charge MisID	-	0.81 ± 0.07	0.74 ± 0.05	1.55 ± 0.09
Rare SM	5.45 ± 2.98	10.36 ± 5.33	4.28 ± 2.45	20.09 ± 10.26
WZ Prod.	0.08 ± 0.04	0.14 ± 0.05	0.04 ± 0.03	0.25 ± 0.07
Total Bkg	7.54 ± 3.20	18.58 ± 6.53	8.15 ± 2.97	34.27 ± 12.06
Observed	16	25	11	52

error. The systematical uncertainties are taken fully correlated among the groups and the value of 50% is taken into account.

6.3 Efficiencies and associated uncertainties

In this section a summary of the experimental and theoretical uncertainties associated to the signal yields is presented. Whenever a particular uncertainty is model-dependent it will be evaluated in-situ for the specific model that would be used for the interpretation.

Lepton and trigger scale factors

Correction factors are applied to simulation only to reduce any potential discrepancy with data as described in Section 5.3. The uncertainty of the total efficiency is 5% (3%) for electrons (muons). An additional systematic uncertainty is assigned to account for potential mismodelling of the lepton isolation efficiency due to varying hadronic activity in signal events. This uncertainty is 3% for all leptons except muons with $p_T < 30$ GeV, for which it is 5%. For the trigger efficiencies, a 6% systematic uncertainty is considered to absorb the maximum deviation among the bins from the central efficiency value.

b-tagging efficiency

The b-tagging efficiency for b-quark jets with $|\eta| < 2.4$, measured in data using samples enriched in $t\bar{t}$ and muon-jet events, has a p_T-averaged value of 0.72. The mis-tagging probability for charm-quark jets is approximately 20%, while for jets originating from light-flavour quarks or gluons it is of the order of 1%. Differences between simulation and data depend of the jet flavour and kinematic, correction factors are applied to account for such differences. Additional correction factors have to be applied to cover the difference

between FullSim and FastSim generated samples. The effect of the b-tagging correction (scale) factors on the $N_{\text{b-jets}}$ distribution as well as the effect of varying the uncertainty up and down can be seen in Figure 6.1

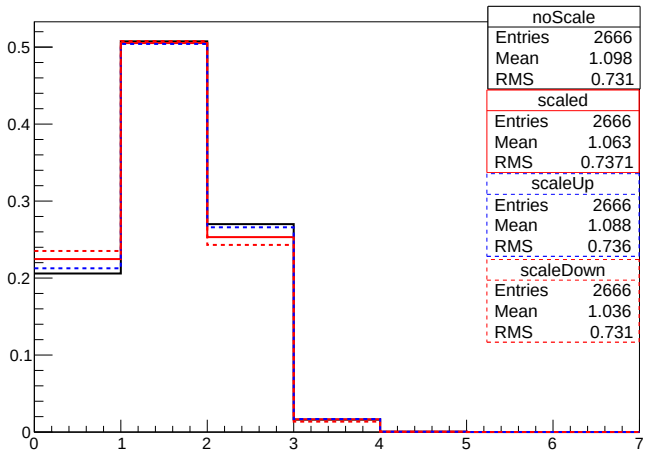


FIGURE 6.1: $N_{\text{b-jets}}$ distribution for $t\bar{t}$ W MC sample to illustrate the effect of the b-tag scale factors on the number of b-tagged jets in a sample.

The total uncertainty of the b-tagging efficiency is determined by simultaneously varying the efficiencies to tag a bottom, charm, or light quark up and down by their uncertainties [37]. The importance of this effect depends on the signal region and the model of new physics. In general, models with more than two b quarks in the final state are less affected by this uncertainty.

Jet Energy Scale and Resolution

Another source of systematic uncertainty is associated with the jet energy scale (JES) correction. This systematic uncertainty varies between 5% and 2% in the p_T range 40-100 GeV for jets with $|\eta| < 2.4$ [54]. It is evaluated on a jet-by-jet basis, and its effect is propagated to H_T , E_T^{miss} , the number of jets and the number of b-tagged jets. The importance of this effect depends on the signal region as well as on the model. Models with high hadronic activity and large E_T^{miss} are less affected by the uncertainty on the jet energy scale.

The effect of the jet-energy-resolution (JER) is evaluated by smearing the jet energy of each jet with a gaussian of width of the jet-energy-resolution in MC, however the effect is much smaller than the JES uncertainty. The uncertainty is calculated separately from the JES one and propagated to H_T , E_T^{miss} and the number of jets, providing a full

knowledge of the correlation of these systematics. This is also taken into account in the limit-setting procedure when combining various signal regions statistically.

The total effect of the JES and JER are of the order of 10% and 3% respectively.

Pile-up modelling in MC

The effect of the possible different pile-up (PU) distributions in data and rare SM background processes were studied on a $t\bar{t}$ W sample. Figure 6.2 compares H_T and E_T^{miss} distributions with and without PU re-weighting. The effect of the PU re-weighting is always smaller than 5.5% and hence negligible compared to the 50% systematic uncertainty added for such processes. Therefore, this procedure is not applied across the range of MC samples employed in this analysis.

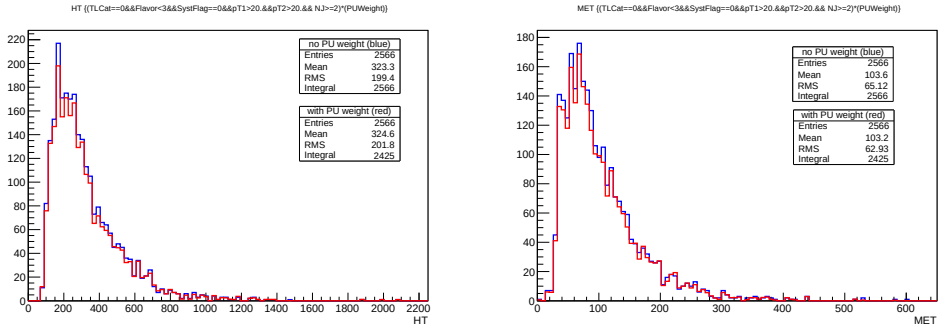


FIGURE 6.2: The results of the pile-up reweighting performed in the $t\bar{t}$ W sample, both H_T (left) and E_T^{miss} (right) distribution show very little variation due to pile-up reweighting.

The signal samples used in this analysis were generated using a pile-up distribution with a slightly higher mean than the one in data. Two tests were performed to asses the effect of this difference. First, the MC samples were re-weighted to have the same pile-up distribution. Then a 5% variation (up and down) of this distribution is implemented to test any possible systematic effect from the lack of knowledge of the pile-up profile, almost no effect on the overall acceptance was seen in any of the signal regions with sufficient. The second test consists on removing completely the PU re-weighting to check its effect in various search regions. The effect is as small as 4-5% for the most populated signal regions. For other search regions, with much smaller statistics the effect is fluctuating up to $\pm$ 20%. The lack of statistics in the sample does not allow for a proper re-weighting procedure.

The possible mismodeling of PU is, therefore, very small. Hence, no pile-up re-weighting is performed and a 5% systematic uncertainty is considered on the signal acceptance.

Other uncertainties

Additional uncertainties due to possible mismodelling of the pile-up conditions or initial-state radiation (ISR) [55] are evaluated and found to be 5% and 3–15%, respectively. The uncertainty of the signal acceptance due to the PDF choice is found to be less than a few percent. Finally, there is a 2.6% uncertainty in the yield of events because of the uncertainty in the luminosity normalisation [56].

Summary of signal acceptance uncertainties

A summary of the systematic uncertainties associated with the acceptance and signal efficiency for this analysis is provided in Table 6.8. While the uncertainties associated with the integrated luminosity, modelling of lepton selection, trigger efficiency, and pile-up are taken to be constant across the parameter space of the new physics models considered in this paper, uncertainties arising from the remaining observables are estimated for each model separately on an event-by-event basis by varying those observables within their uncertainties. The total uncertainty in the computed acceptance is in the 13–25% range. However, these uncertainties are representative and do not characterise the results for extreme kinematic regions, such as those near the diagonal of the parameter space of the SUSY simplified models discussed later, where the particle mass spectra are compressed

TABLE 6.8: Summary of representative systematic uncertainties for the considered signal models.

Source	%
Luminosity	2.6
Modelling of lepton selection (ID and isolation)	10
Modelling of trigger efficiency	6
Pileup modelling	5
Jet energy scale	1–10
Jet energy resolution	0–3
b-jet identification	2–10
ISR modelling	3–15
Total	13–25

6.4 Results

After applying the event selection for the three baseline regions described in Table 6.1, 681 events remain in our selection. The distributions of E_T^{miss} versus H_T for events in the three baseline signal regions are shown in Figure 6.3. The corresponding results for the

four selection variables H_T, E_T^{miss}, N_{jets} and $N_{b\text{-}jets}$ are shown in Figure 6.4, where also the SM background prediction is shown. There are no significant discrepancies observed between the observed yields ad the background prediction for any region.

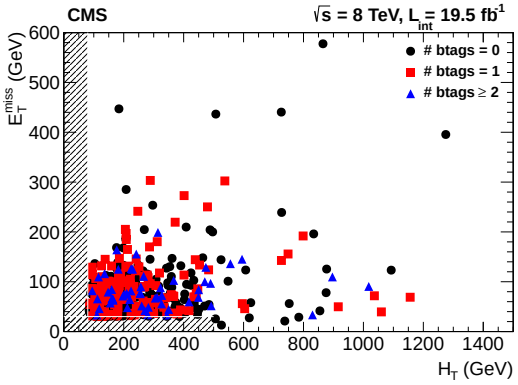


FIGURE 6.3: Distribution of E_T^{miss} versus H_T for the baseline signal regions BSR0 (black), BSR1 (red) and BSR2 (blue). The regions indicated with the hatched area are not included in the analysis.

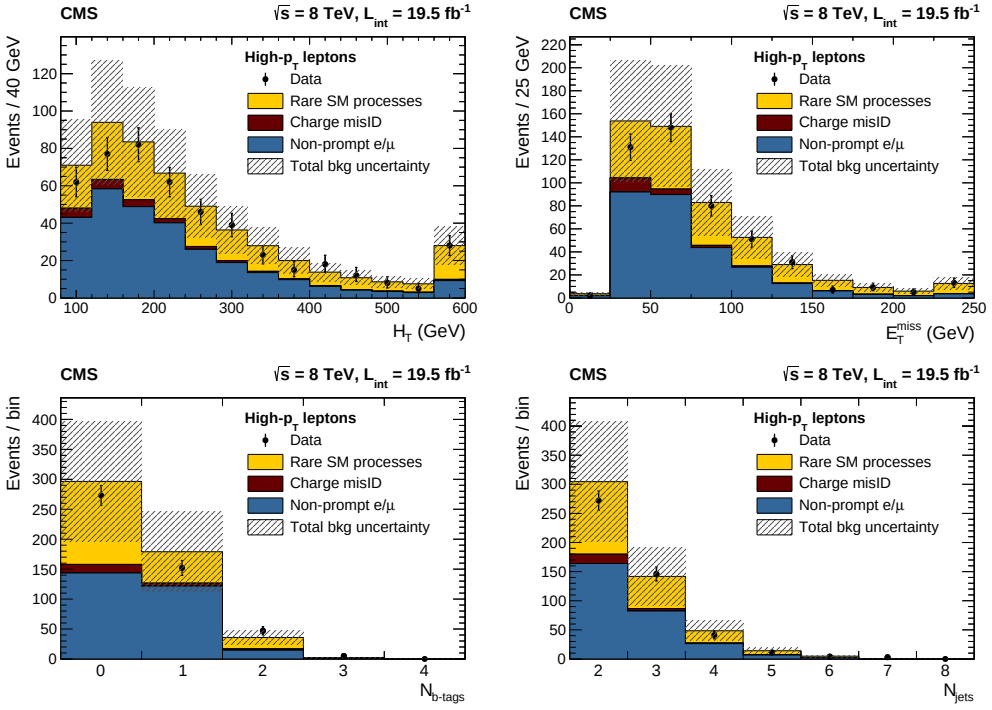


FIGURE 6.4: Distributions of H_T, E_T^{miss}, $N_{b\text{-}jets}$ and N_{jets} for events in the baseline region with no $N_{b\text{-}jets}$ requirement (events selected in BSR0, BSR1 and BSR2). Also show as a histograms is the background prediction. The shaded area represents the total background uncertainty, both statistical and systematic.

The observations of the final signal regions are presented in Tables 6.9 and 6.10 and in Figure 6.5 along with the SM background predictions. The contribution from rare SM processes and non-prompt lepton vary among the different signal regions between 40% and 60%, while the charge misidentification background is almost negligible across most

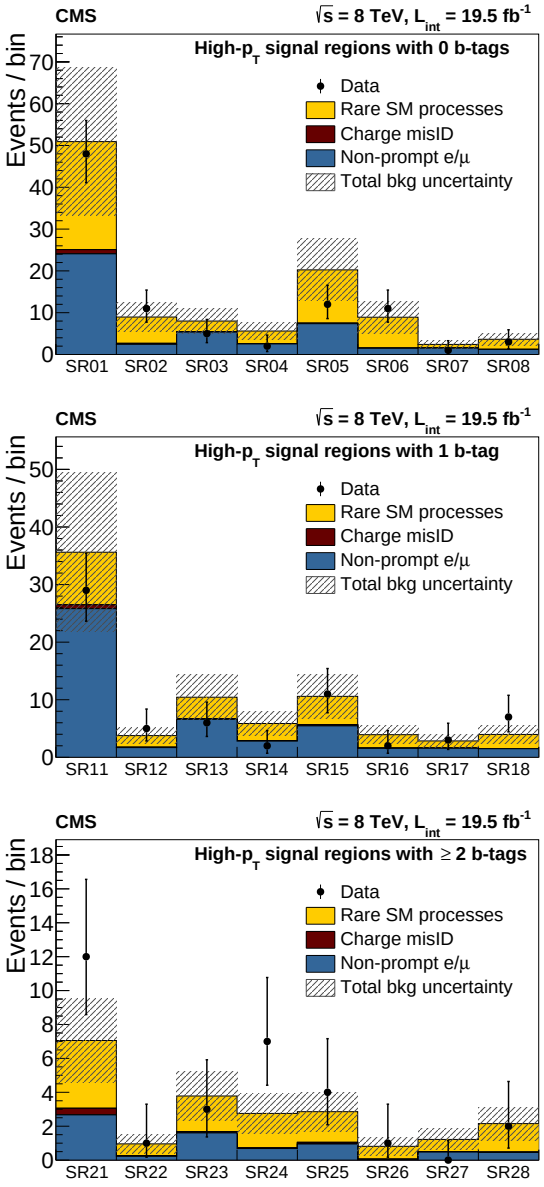


FIGURE 6.5: Summary plots showing the predicted background from each source and observed event yields as a function of the SRs. The shaded area represents the total background uncertainty, both statistical and systematic.

signal regions. The observed yields are consistent with the background expectations within their uncertainties. The p-values [4] for each signal region are shown along with the observed yields in Table 6.9 and they are found to be consistent with a uniform distribution between 0 and 1.

TABLE 6.9: Predicted and observed event yields for the signal regions. The p-value is also shown for each signal region.

Region	Expected	Observed	p-value
SR01	51 ± 18	48	0.56
SR02	9.0 ± 3.5	11	0.35
SR03	8.0 ± 3.1	5	0.79
SR04	5.6 ± 2.1	2	0.92
SR05	20 ± 7	12	0.84
SR06	9 ± 4	11	0.35
SR07	2.4 ± 1.0	1	0.85
SR08	3.6 ± 1.5	3	0.64
SR11	36 ± 14	29	0.67
SR12	3.8 ± 1.4	5	0.34
SR13	10 ± 4	6	0.84
SR14	5.9 ± 2.2	2	0.93
SR15	11 ± 4	11	0.48
SR16	3.9 ± 1.5	2	0.83
SR17	2.8 ± 1.1	3	0.51
SR18	4.0 ± 1.5	7	0.16
SR21	7.1 ± 2.5	12	0.12
SR22	1.0 ± 0.5	1	0.57
SR23	3.8 ± 1.4	3	0.67
SR24	2.8 ± 1.2	7	0.05
SR25	2.9 ± 1.1	4	0.33
SR26	0.8 ± 0.5	1	0.51
SR27	1.2 ± 0.6	0	1
SR28	2.2 ± 1.0	2	0.59

TABLE 6.10: Predicted and observed event yields in the signal regions designed for same-sign top-quark pair production and RPV SUSY models.

SR	Expected	Observed
RPV0	38 ± 14	35
RPV2	5.3 ± 2.1	5
SStop1	160 ± 59	152
SStop1++	90 ± 32	92
SStop2	40 ± 13	52
SStop2++	22 ± 8	25

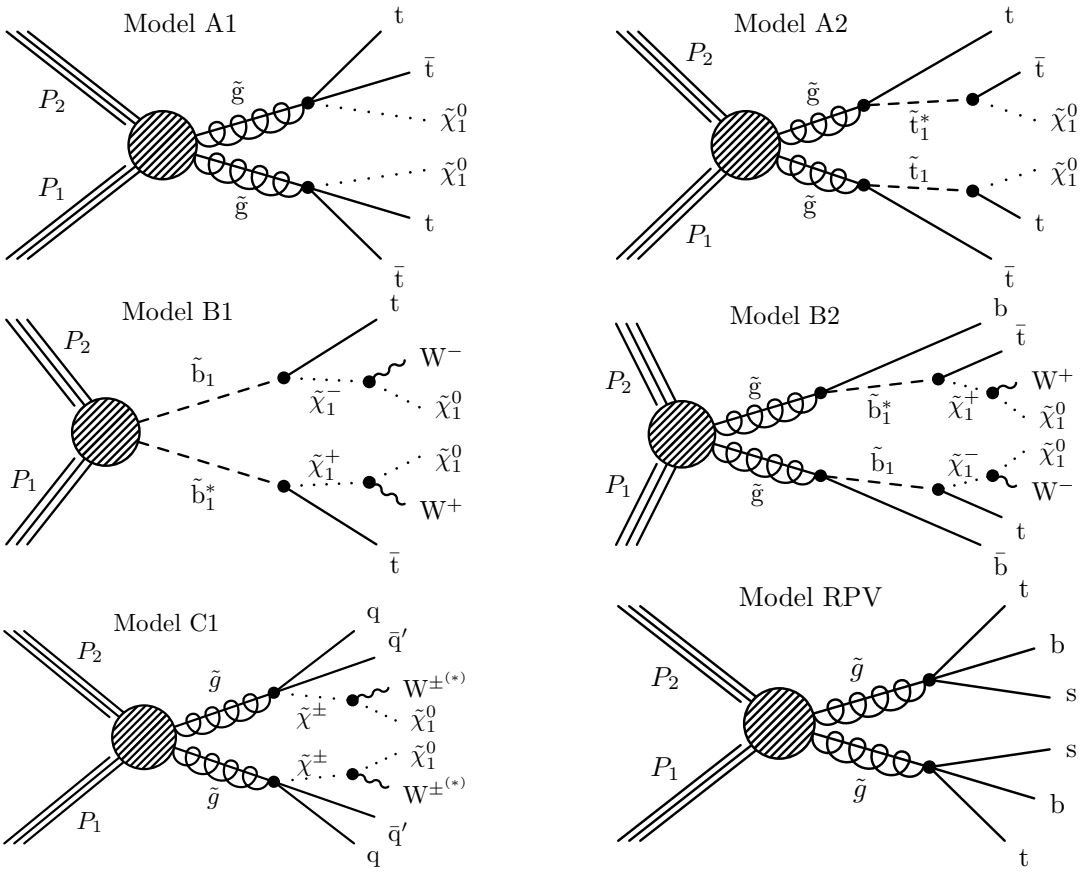


FIGURE 6.6: Diagrams for the six SUSY models considered (A1, A2, B1, B2, C1, and RPV).

6.5 Limits on new physics models and on rare SM processes

Given the lack of a significant excess over the expected SM background, the results on the search are used to derive limits on the parameters of various models of new physics and to derive limits on the cross section of rare SM processes. The 95% confidence level (CL) upper limits on the signal yields are calculated using LHC-type CL_S method [57–59].

The number of events that are expected to satisfy the selection for a given signal model is obtained from simulation. The uncertainties for the event yields are computed as it was described in Section 6.3. For a given signal region, the different sources of uncertainties in the signal acceptance are considered to be uncorrelated, with correlations across signal regions taken into account. The uncertainties in the total background across the signal regions are considered to be fully correlated

TABLE 6.11: Signal regions used for limit setting for the new physics models considered in this analysis.

Model	Constraints on parameters	Signal regions used
A1		21–28
A2	$m_{\widetilde{\chi}_1^0} = 50$ GeV	21–28
B1	$m_{\widetilde{\chi}_1^0} = 50$ GeV	11–18, 21–28
B1	$m_{\widetilde{\chi}_1^0}/m_{\widetilde{\chi}_1^\pm} = 0.5$	11–18, 21–28
B2	$m_{\widetilde{\chi}_1^0} = 50$GeV, $m_{\widetilde{\chi}_1^\pm} = 150$ GeV	21–28
B2	$m_{\widetilde{\chi}_1^0} = 50$GeV, $m_{\widetilde{\chi}_1^\pm} = 300$ GeV	21–28
C1	$m_{\widetilde{\chi}_1^\pm} = 0.5m_{\widetilde{\chi}_1^0} + 0.5m_{\widetilde{g}}$	01–08
RPV		RPV2
$pp \rightarrow t\bar{t}$, $t\bar{t}$		SStop1, SStop2
$pp \rightarrow t\bar{t}$		SStop1++, SStop2++
$pp \rightarrow t\bar{t}t\bar{t}$		21–28

Limits on the parameter spaces of various R-parity-conserving simplified SUSY models [24] are presented. The exclusion contours are obtained with the gluino or bottom-squark pair production cross sections at the NLO+NLL (i.e. next-to-leading-logarithm) accuracy that are calculated in the limit where other sparticles are heavy enough to be decoupled [60, 60–63]. The production of SUSY particles and the decay chains under consideration are shown schematically in Figure 6.6.

Models A1 and A2 represent a scenario with gluino pair production leading to a $t\bar{t}t\bar{t}\widetilde{\chi}_1^0\widetilde{\chi}_1^0$, with $\widetilde{\chi}_1^0$as the lightest neutralino [24, 64–67]. For model A1, the gluino undergoes a three-body decay $\tilde{g} \rightarrow t\bar{t}\widetilde{\chi}_1^0$ mediated by a virtual to squark. In model A2, the gluino decays to a top quark and a top anti-squark, with the on-shell anti-squark decaying into a top anti-quark and the $\widetilde{\chi}_1^0$. Both models produce four on-shell W bosons and four b-quarks. Therefore, search regions SR21–SR28, which require at least two b-tagged jets, are used to derive the limits on the parameters of these models; the region with the best sensitivity is SR28.

The 95% CL upper limits on the cross section times branching fraction, as well as the exclusion contours, are shown in Figure 6.7. For model A1, the results are presented as a function of gluino mass and $\widetilde{\chi}_1^0$ mass, and for model A2 as a function of gluino mass and top squark mass with the $\widetilde{\chi}_1^0$ mass set to 50 GeV. In model A2, the limits do not depend on the top squark or $\widetilde{\chi}_1^0$ masses provided that there is sufficient phase space to produce on-shell top quarks with a moderate boost in the decay of both the gluino and the top squark. This range extends to approximately 600 GeV for the $\widetilde{\chi}_1^0$ mass.

These results extend the sensitivity obtained in the previous analysis [51] on gluino and sbottom masses. For the gluino-initiated models (A1, A2, B2, and C1). Gluino masses

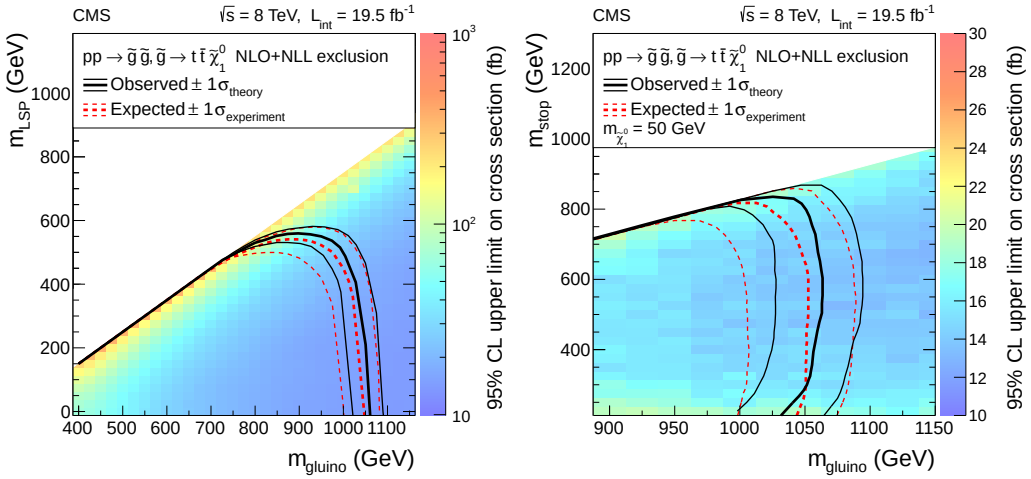


FIGURE 6.7: Exclusion regions at 95% CL in the planes of (left) $m(\widetilde{\chi}_1^0)$ versus $m(\widetilde{g})$ (model A1), and (right) $m(\tilde{t}_1)$ versus $m(\widetilde{g})$ (model A2). The excluded regions are those within the kinematic boundaries and to the left of the curves. The effects of the theoretical uncertainties in the NLO+NLL calculations of the production cross sections [68] are indicated by the thin black curves; the expected limits and their ± 1 standard-deviation variations are shown by the dashed red curves.

are probed up to about 1050 GeV, mainly because the limits are driven by the gluino pari production cross section. In the case of the direct bottom-squark pair production, model B1, the search shows sensitivity for bottom-squark masses up to about 500 GeV.

These models are also probed by other CMS new physics searches in different decay modes. Other searches are usually interpreted in the context of model A1 but not A2, B1, or B2. For model A1, the limits given here are complementary to the limits from the searches presented in Refs. [69–72]. In particular, they are less stringent at low $m(\widetilde{\chi}_1^0)$ but more stringent at high $m(\widetilde{\chi}_1^0)$. A similar conclusion applies to model A2, since the final state is the same. For bottom-squark pair production, limits on $m(\tilde{b}_1)$ of about 600 GeV have been presented [71], but assuming the decay mode $\tilde{b}_1 \to \mathrm{b}\widetilde{\chi}_1^0$ instead of the model B1 mode $\tilde{b}_1 \to \mathrm{t}\widetilde{\chi}_1^-$ considered here. Comparable limits for model A1, as well as for similar models with top and bottom quarks from gluino decays, have been reported by the ATLAS Collaboration [73–76].

A single RPV scenario is considered in this analysis, one in which gluino pair production is followed by the decay of each gluino to three quarks, as is favoured in the SUSY model with minimal flavour violation [77]: $\widetilde{g} \to \mathrm{tbs}(\bar{\mathrm{t}}\bar{\mathrm{b}}\bar{\mathrm{s}})$ (model RPV). Such decays lead to same-sign W-boson pairs in the final state in 50% of the cases. Compared with the decays $\widetilde{g} \to \mathrm{tsd}(\bar{\mathrm{t}}\bar{\mathrm{s}}\bar{\mathrm{d}})$, which also yield same-sign W-boson pairs, the mode considered profits from having two extra b quarks in the final state, resulting in a higher signal selection efficiency. The model is governed by one parameter ($m_{\widetilde{g}}$), which dictates the production cross section and the final state kinematics. The dedicated search region

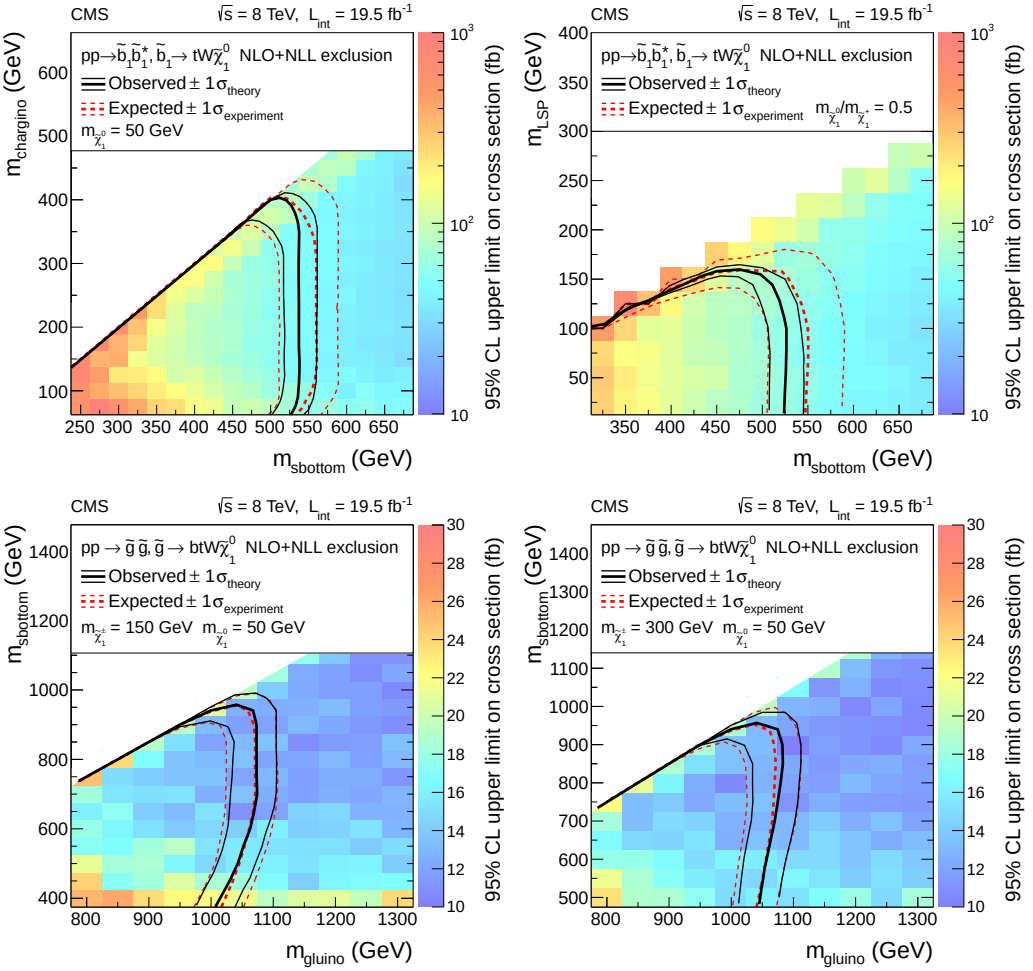


FIGURE 6.8: Exclusion regions at 95% CL in the planes of (top) $m(\widetilde{\chi}_1^{\pm})$ versus $m(\tilde{b}_1)$ and $m(\widetilde{\chi}_1^0)$ versus $m(\tilde{b}_1)$ (model B1), and (bottom) $m(\tilde{b}_1)$ versus $m(\widetilde{g})$ (model B2). The convention for the exclusion curves is the same as in Fig. 6.7.

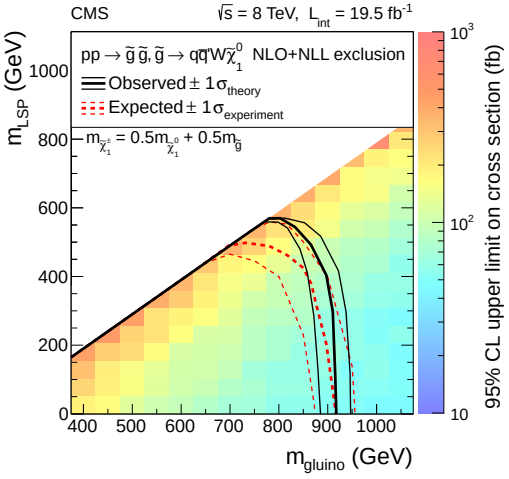


FIGURE 6.9: Exclusion regions at 95% CL in the planes of $m(\widetilde{\chi}_1^0)$ versus $m(\widetilde{g})$ for two different values of chargino mass (model C1). The convention for the exclusion curves is the same as in Fig. 6.7.

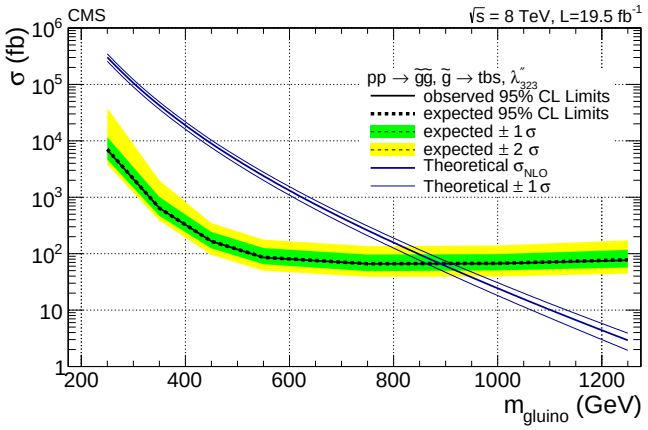


FIGURE 6.10: 95% CL upper limit on the gluino production cross section for an RPV simplified model, $pp \to \widetilde{g}\widetilde{g}, \widetilde{g} \to tbs(\bar{t}\bar{b}\bar{s})$.

RPV2 is used to place an upper limit on the production cross section. The result is shown in Figure 6.10. In this scenario, the gluino mass is probed up to approximately 900 GeV.

The results for the signal regions SStop1, SStop1++, SStop2, and SStop2++ are used to set limits on the cross section for same-sign top-quark pair production, $\sigma(pp \to tt,\ \bar{t}\bar{t})$ from SStop1 and SStop2, and $\sigma(pp \to tt)$ from SStop1++ and SStop2++. Here $\sigma(pp \to tt,\ \bar{t}\bar{t})$ is shorthand for the sum $\sigma(pp \to tt) + \sigma(pp \to \bar{t}\bar{t})$. These limits are calculated using an acceptance obtained from simulated $pp \to t\bar{t}$ events and an opposite-sign selection. This acceptance, including branching fractions, is 0.43% (0.26%) for the SStop1 (SStop2) search region. The relative uncertainty in this acceptance is 14%. The observed upper limits are $\sigma(pp \to tt,\ \bar{t}\bar{t}) < 720$ fb and $\sigma(pp \to tt) < 370$ fb at 95% CL. The median expected limits are 470^{+180}_{-110} fb and 310^{+110}_{-80} fb, respectively.

Similarly, the results from signal regions SR21–SR28 are used to set limits on the SM cross section for quadruple top-quark production. The observed upper limit is $\sigma(pp \to tt\bar{t}\bar{t}) < 49$ fb at 95% CL, compared to a median expected limit of 36^{+16}_{-9} fb. The SM cross section as computed with the MC@NLO program [78] is $\sigma_{SM} = 0.914 \pm 0.005$ fb. The most sensitive signal regions, SR24 and SR28, have a signal acceptance of 0.52% and 0.49%, respectively, with relative uncertainties of 13% and 17%.

6.6 Summary

A search for physics beyond the SM with events with two same-sign leptons has been described in this chapter. The search uses pp collisions at a centre-of-mass energy of 8

TeV corresponding to an integrated luminosity of 19.5 fb^{-1} . A combination of exclusive search regions is used to probe signature both with and without third-generation squarks. No significant deviation from the SM expectation is found.

Using sparticle production cross sections calculated in the decoupling limit, and assuming that gluinos decay exclusively into top or bottom squarks and that the top and bottom squarks decay as $\tilde{t}_1 \to \text{t}\widetilde{\chi}_1^0$ and $\tilde{b}_1 \to \text{t}\widetilde{\chi}_1^-$ ($\widetilde{\chi}_1^- \to W^-\widetilde{\chi}_1^0$), lower limits on gluino and sbottom masses are calculated. Gluinos with masses up to approximately 1050 GeV and bottom squarks with masses up to about 500 GeV are probed. In models where gluinos do not decay to third-generation squarks, sensitivity for gluino masses up to approximately 900 GeV is obtained. A similar reach in the gluino masses is demonstrated in the scope of an R-parity violating model. These results improve by about 100 GeV previous CMS results [51].

The results are used to set upper limits on the same-sign top-quark pair production cross section $\sigma(\text{pp} \to \text{tt},\ \overline{\text{tt}}) < 720$ fb and $\sigma(\text{pp} \to \text{tt}) < 370$ fb at 95% CL. An upper limit at 95% CL of $\sigma(\text{pp} \to \text{tt}\overline{\text{tt}}) < 49$ fb is obtained for the cross section of quadruple top-quark production. This work has been published at the beginning of 2014 at JHEP [52].

Chapter 7

Search for direct production of charginos and neutralinos in the same-sign dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV

While most of LHC searches focus on strong production of SUSY with larger cross sections, electroweak production of SUSY may be the key to new physics, as the squarks and gluinos may be too heavy to be produced at LHC energies. The decay chain will produce sleptons (or W,Z and h bosons) and lead to multiple-lepton final states with little hadronic activity and significant missing transverse energy (E_T^{miss}) from the lightest-supersymmetric particle (LSP).

As the production cross-section is very small, many final states are combined together targeting different production mechanisms to reach enough sensitivity for new physics.

The analysis described here is an extension of the analysis described in the previous chapters. New signal regions are defined to provide enough sensitivity to the targeted model while object selection and background estimation methods remain exactly the same.

7.1 Search strategy

As opposed to the model-independent search described in Chapter 6, this analysis is targeting at chargino-neutralino ($\widetilde{\chi}_1^{\pm}$-$\widetilde{\chi}_2^0$) production, where both the chargino and neutralino decay into leptons, see Figure 7.1.

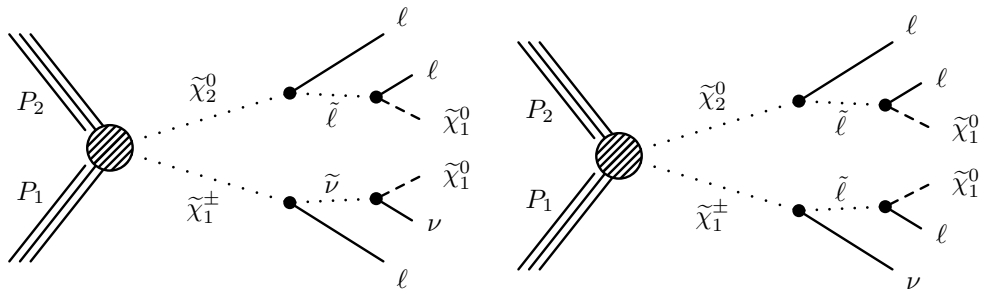


FIGURE 7.1: Chargino-neutralino pair production with decays mediated by sleptons and sneutrinos, leading to a three-lepton final state with missing transverse energy $\mathrm{E_T^{miss}}$

The masses of the new-physics particles are treated as independent parameters. SUSY models with a bino-like $\widetilde{\chi}_1^0$ and wino-like $\widetilde{\chi}_2^0$ and $\widetilde{\chi}^{\pm}$ motivate the simplifying assumption $m_{\widetilde{\chi}} \equiv m_{\widetilde{\chi}^{\pm}} = m_{\widetilde{\chi}_2^0}$ since these two gauginos belong to the same gauge group multiplet. Therefore the slepton mass ($m_{\tilde{\ell}}$) is parametrised as a function of the other supersymmetric particles intervening in the decay:

$$m_{\tilde{\ell}} = m_{\widetilde{\nu}} = m_{\widetilde{\chi}_1^0} + x_{\tilde{\ell}}\,(m_{\widetilde{\chi}} - m_{\widetilde{\chi}_1^0}),$$

where $0 < x_{\tilde{\ell}} < 1$. We present results for $x_{\tilde{\ell}} = 0.05$ and 0.95, i.e., the slepton mass close to the LSP or $\widetilde{\chi}_2^0$,$\widetilde{\chi}_1^{\pm}$ mass, respectively.

Although the final state predicted by such models naturally contains three leptons, under certain conditions such as very small mass splittings, events may elude the three-lepton analysis acceptance. This acceptance can be partially recovered by requiring only two same-sign leptons (to avoid dealing with large SM background). This search, therefore complements the three-lepton search performed by CMS. Both searches are described together in a more extended paper covering a wide variety of searches for electro-weak production of SUSY [79].

Very little hadronic energy is expected for this signature, except for some some residual contamination from initial state radiation. This feature will be exploited when designing the analysis. The two $\widetilde{\chi}_1^0$ will escape undetected leaving a high $\mathrm{E_T^{miss}}$ signature.

7.1.1 Third lepton veto

As mentioned before, this search complements the three-lepton search fully described here [79]. For facilitating the combination of the two analysis an extra third-lepton veto

(e, μ, τ) is applied on the leptons passing the identification criteria summarised in Tables 7.1 and 7.2. This third lepton veto will be applied only when combining the results with the three lepton analysis.

TABLE 7.1: Electron and muon selection for 3rd lepton veto.

p_T	> 10 GeV
$\|\eta\|$	< 2.4
Tight WP	see section 5.2
RelPFIso	< 0.15
$\Delta R_{\mu e}$	> 0.1

TABLE 7.2: Tau selection for 3rd lepton veto.

HPS PF tau	
p_T	> 20 GeV
$\|\eta\|$	< 2.3
DecayModeFinding	> 0.5
AgainstMuonTight	> 0.5
AgainstElectronMVA	> 0.5
Comb. DB Loose isolation	> 0.5
$\Delta R_{\tau e, \tau \mu}$	> 0.1

7.1.2 Search regions

The search strategy is a simple cut-and-count method with much two search regions defined using only two discriminating variables: $\mathrm{E_T^{miss}}$ and $\mathrm{N_{jets}}$. This provides enough sensitivity of electroweakly produced SUSY particles for which low-hadronic activity and high-$\mathrm{E_T^{miss}}$ is expected.

A "baseline region" is defined with looser event selection requirements, yielding to a background-dominated sample with negligible signal contamination. These will be used to understand the background composition as well as validate the background estimation methods that will be used to predict SM backgrounds. For defining these baseline regions (BSR) events having at least two jets, $\mathrm{H_T} > 80$ GeV, $\mathrm{E_T^{miss}} > 30$ GeV (only if $\mathrm{H_T} < 500$ GeV) and 0 (BSR0), 1 (BSR1), 2 or more (BSR2) b-jets are selected. This selection is summarised in Table 6.1.

On top of the baseline selection, two fully exclusive signal regions are defined for which the background contribution is expected to be low. This regions exploit the signal topology described before:

- $120 < \mathrm{E_T^{miss}} < 200$ GeVand $N_{jets} = 0$.

- $E_T^{miss} > 200$ GeV.

The analysis was firstly designed without the jet-veto region, as the signal MC available at that time was simulated using PYTHIA [48]. Different studies showed that the ISR was poorly simulated in such samples, and thus any cut on the number of jet couldn't be trusted, a b-jet veto was implemented to try to further reduce background in the lower-E_T^{miss} region. Later in the year, a MC sample was generated using MADGRAPH5 [47] and the full jet-veto was implemented probing much higher sensitivity than the one achieved with the b-jet veto.

7.2 Validation of the background estimation methods

Three different sources of backgrounds affect this search as it was described before: "Non-Prompt leptons (Fakes)", "rare SM processes" and "charge mis-identification". A detail description of these backgrounds and the estimation technique applied in each case is found in Section 5.3. A particular case shall be considered under the "rare SM processes": WZ production, although marginal in the previously described analysis, plays an important role here and will be considered separately when quoting yields.

Prior to apply the background prediction techniques into the signal regions, the performance of these background estimation techniques needs to be assessed. The background dominated baseline region will be used for this purpose. The performance of the non-prompt lepton estimation has been shown in Section 6.2.

7.2.1 WZ validation

As it was stated before the WZ process constitutes the most relevant source of irreducible background to this analysis. It is therefore crucial to make sure that this process is well reproduced in the baseline region by the simulation, and associate a systematic uncertainty to account for the observed differences with respect to what is directly extracted from simulation.

The E_T^{miss} distribution and the data/MC agreement of a WZ-enriched sample, orthogonal to the signal selection, where the shapes and integrated yields can be compared. For defining this control regions, the Z-veto described in Section 5.2.4 is inverted to select events with three leptons making an opposite-sign same-flavour pair, furthermore, events with two same-sign leptons $E_T^{miss} > 40$ GeV and no identified b-jets are selected.

A comparison between the integrated yields obtained purely from MC and the observed region in the WZ enriched samples is shown in Table 7.3. The difference between observed yields and predictions goes from 3% to 20% depending on the considered channel. A 15% systematic uncertainty is considered to account for the observed difference between data and MC in each of the channels.

TABLE 7.3: Comparison of predicted and observed yields in a WZ dominated region with $E_T^{miss} > 40$ GeV and no b-jets. Systematic errors of the MC-estimated backgrounds are not taken into account.

	$\mu\mu$	$e\mu$	ee	Total
W/Z/WW/ZZ	39.91 ± 18.24	54.63 ± 24.91	26.87 ± 12.95	121.40 ± 54.86
Rare SM (Sum)	9.08 ± 6.43	14.65 ± 10.37	5.48 ± 3.89	29.21 ± 20.66
QCD/Top	0.98 ± 0.82	1.77 ± 1.34	0.78 ± 0.64	3.53 ± 2.60
WZ Prod	288.88 ± 1.55	444.03 ± 1.99	168.60 ± 1.23	901.52 ± 2.80
Observed	399	529	186	1114
Background	338.85 ± 19.42	515.08 ± 27.09	201.74 ± 13.60	1055.66 ± 58.74
Obs. - Bkg./ WZ	1.21 ± 0.07	1.03 ± 0.06	0.91 ± 0.08	1.06 ± 0.07
Obs. - Pred./ WZ	0.21 ± 0.07	0.03 ± 0.06	-0.09 ± 0.08	0.06 ± 0.07

The well simulation of E_T^{miss} distribution must be also considered for this validation procedure. Given than the E_T^{miss} shape is well simulated by our MC (within errors), as shown in Figure 7.2, therefore we will not add any systematic error to account for that.

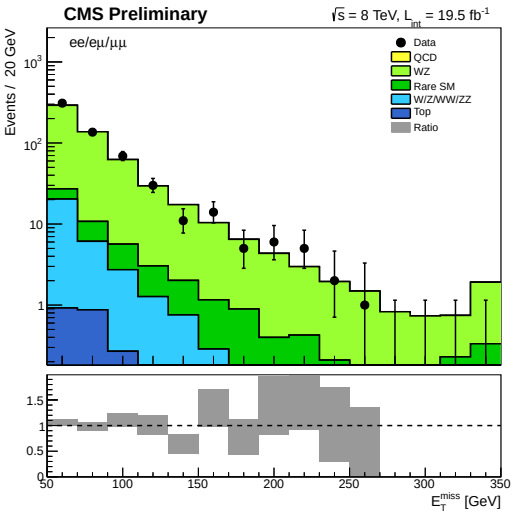


FIGURE 7.2: Distribution of observed missing transverse momentum in the Z-veto inverted, WZ dominated region, compared to pure simulation normalized to the integrated luminosity. As can be seen, WZ constitutes about 90 % of the yield towards the high end tail of the distribution.

7.2.2 Prediction for baseline region

The performance of the background estimation methods is also validated in data using the baseline signal region defined previously, with negligible signal contamination and dominated by background events. The results of the background estimation methods and the yields of the observed events from data for the baseline signal region is shown in Table 7.4. Good agreement between prediction and observed yields and data is seen.

TABLE 7.4: Summary of background predictions and observed yields in the E_T^{miss} >120 GeV baseline region

	$\mu\mu$	$e\mu$	ee	Total
Double Fakes	0.01 ± 0.03	0.03 ± 0.05	0.04 ± 0.05	0.08 ± 0.09
Single Fakes	4.29 ± 2.46	16.55 ± 8.41	8.20 ± 4.35	29.04 ± 14.72
Charge MisID	-	0.56 ± 0.05	0.51 ± 0.03	1.08 ± 0.06
Rare SM	12.28 ± 6.28	22.31 ± 11.24	9.81 ± 5.09	44.40 ± 22.33
WZ Prod.	6.24 ± 0.97	14.42 ± 2.19	6.45 ± 1.00	27.11 ± 4.10
Total Bkg	22.82 ± 6.76	53.86 ± 14.22	25.02 ± 6.71	101.71 ± 27.04
Observed	27	50	17	94

7.3 Results.

After applying the event selection for the baseline region, 94 events remain in our selection. The distributions of E_T^{miss} versus H_T for the selected events are shown in Figure 7.3. The corresponding results for the selection variable E_T^{miss} is shown in Figure 7.4 (left), where also the SM background prediction is shown. There are no significant discrepancies observed between the observed yields ad the background prediction for any region.

The observations of the final signal regions are presented in Tables 7.5 and in Figure 7.4 (right) along with the SM background predictions. The contribution from rare SM processes and non-prompt lepton vary among the different signal regions between 40% and 60%, while the charge misidentification background is almost negligible across most signal regions. The observed yields are consistent with the background expectations within their uncertainties. Note that both the results with and without the 3^{rd} lepton veto are shown, the former will be used for combining the results with the three-lepton analysis.

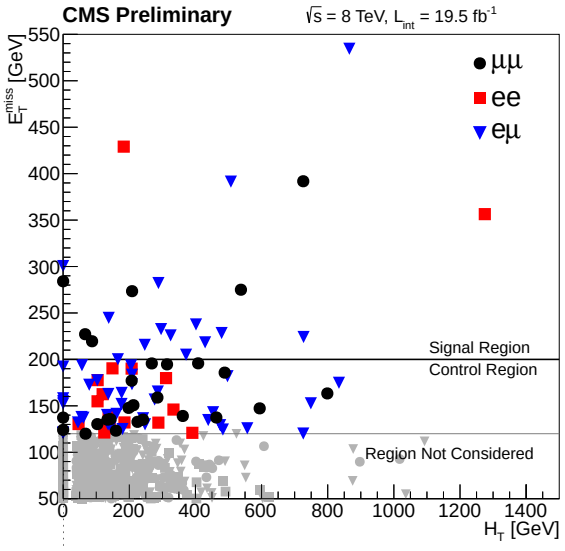


FIGURE 7.3: Scatter plot in the H_T vs. E_T^{miss} plane of the observed events.

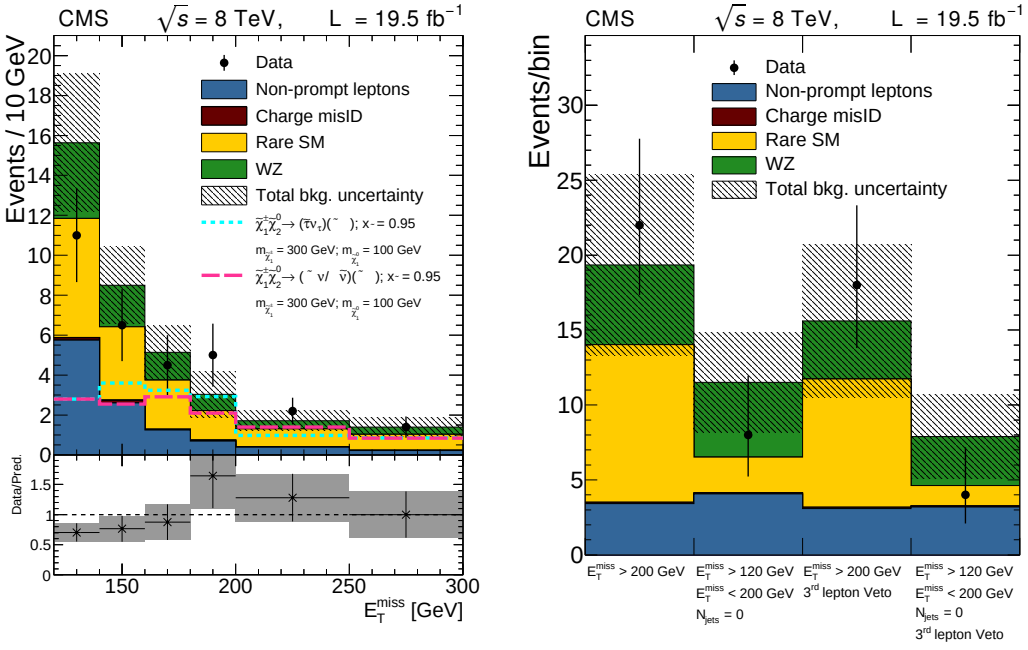


FIGURE 7.4: **Left:** E_T^{miss} distribution for same-sign dilepton candidates in comparison with the SM expectations. The bottom panel shows the the ratio and corresponding uncertainty of the observed and total SM expected distributions. The third lepton veto is not applied. The distributions of example signal scenarios are overlaid. **Right:** Observed yields and expected backgrounds for the different search regions. In both plots, events with E_T^{miss} >120 GeV are displayed, and the hashed band shows the combined statistical and systematic uncertainties of the total background.

TABLE 7.5: Summary of background predictions and observed yields for search regions: 120<E_T^{miss} <200 GeV, $N_{jets} \leq 2$ and $N_{bjets} = 0$, and E_T^{miss} >200 GeV, with and without third lepton veto applied (including taus). Uncertainties include statistical and systematic contributions.

	E_T^{miss} >200 GeV	120<E_T^{miss} <200 GeV $N_{jets} = 0$	E_T^{miss} >200 GeV 3^{rd} lepton veto	120<E_T^{miss} <200 GeV $N_{jets} = 0$ 3^{rd} lepton veto
Fakes	3.4 ± 1.9	4.1 ± 2.2	3.1 ± 1.7	3.2 ± 1.7
Charge MisID	0.09 ± 0.01	0.08 ± 0.01	0.09 ± 0.01	0.07 ± 0.01
Rare SM	10.5 ± 5.7	2.4 ± 2.4	8.6 ± 4.8	1.4 ± 2.1
WZ Prod.	5.3 ± 0.8	5.0 ± 0.8	3.9 ± 0.6	3.3 ± 0.5
Total Bkg	19.4 ± 6.0	11.5 ± 3.3	15.6 ± 5.1	7.9 ± 2.8
Data	22	8	18	4

7.4 Limits on new physics models

With the absence of a signal, the obtained results are interpreted in terms of simplified SUSY models with direct $\widetilde{\chi}_1^{\pm}$- $\widetilde{\chi}_2^0$ production with light sleptons depicted in Figure 7.1, combining with the results of the three-lepton search. Three different scenarios are considered, with different assumptions about the nature of the sleptons, which affect the number of τ leptons in the final state. These interpretations depend on whether the sleptons are the SUSY partners $\tilde{\ell}_L$ or $\tilde{\ell}_R$ of left-handed or right-handed leptons. We consider two limiting cases. In one case, $\tilde{\ell}_R$ does not participate while $\tilde{\ell}_L$ and $\widetilde{\nu}$ do: then both diagrams in Figure 7.1 exist, and the chargino and neutralino decay to all three lepton flavours democratically. Furthermore, two additional diagrams in which the decay $\widetilde{\chi}_2^0 \to \ell\,\tilde{\ell} \to \ell\,\ell\,\widetilde{\chi}_1^0$ is replaced by $\widetilde{\chi}_2^0 \to \widetilde{\nu}\,\nu \to \nu\,\nu\,\widetilde{\chi}_1^0$ reduce the fraction of three-lepton final states by 50%. In the second case, in which $\tilde{\ell}_R$ participates while $\tilde{\ell}_L$ and $\widetilde{\nu}$ do not, only the diagram on the right of Figure 7.1 exists, and there is no reduction in the three-lepton final states. Because the $\tilde{\ell}_R$ couples to the chargino via its higgsino component, chargino decays to $\tilde{\ell}_R$ strongly favor production of a τ lepton. We thus consider two flavour scenarios:

- the "flavour-democratic" scenario: the chargino ($\widetilde{\chi}_1^{\pm}$) and neutralino ($\widetilde{\chi}_2^0$) both decay with equal probability into all three lepton flavours, as expected for $\tilde{\ell}_L$;
- the "τ-enriched" scenario: the chargino decays exclusively to a τ lepton as expected for $\tilde{\ell}_R$, while the neutralino decays democratically;

Signal samples are generated with the MADGRAPH 5.1.5.4 generator [47] including up to two additional partons at the matrix element level. Parton showering, hadronization, and the decay of particles, including SUSY particles, are described with the PYTHIA

6.4.26 [48] program. Signal cross sections are calculated at NLO+NLL using the RESUMMINO [80–82] calculation, where NLL refers to the next-to-leading-logarithmic precision.

For the flavour democratic scan with x=0.05, i.e. the model in which both the chargino and neutralino decay democratically into electrons, muons, and taus, efficiency x acceptance maps are shown in Figure 7.5 for both search regions with and without the third lepton veto. Figure 7.6 shows efficiency x acceptance maps for an x value of 0.95.

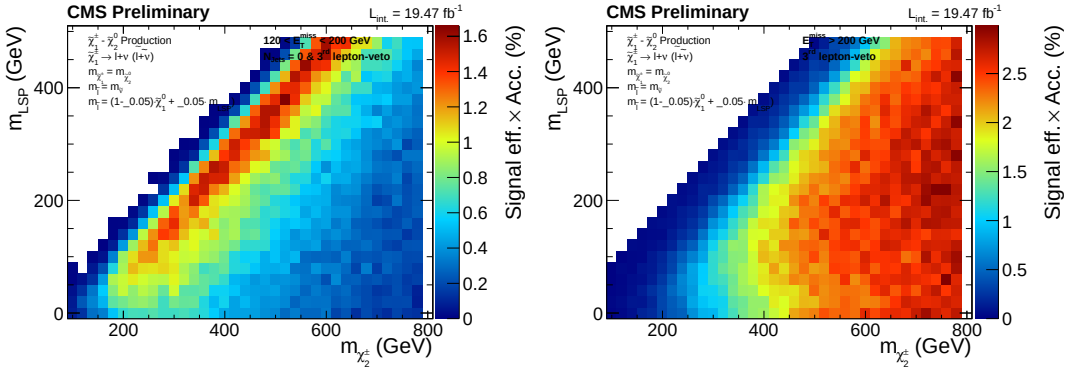


FIGURE 7.5: Efficiency times acceptance maps for the flavour democratic model with x = 0.05, calculated for both signal regions with the 3^{rd} lepton veto applied.

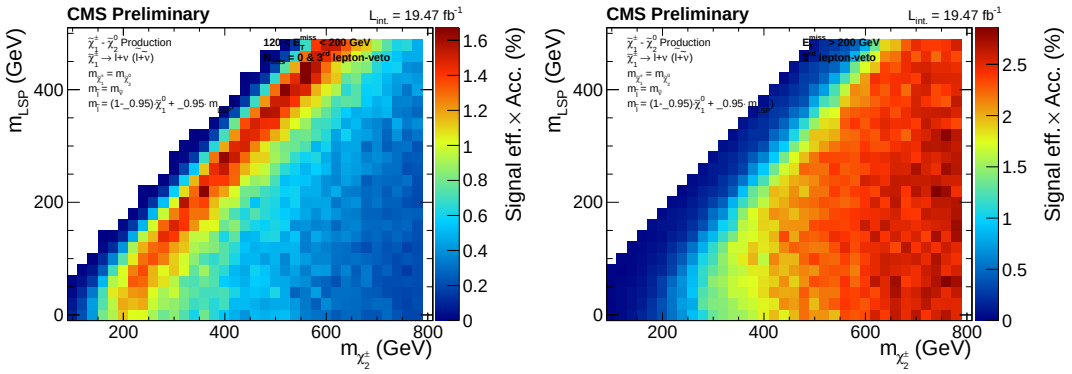


FIGURE 7.6: Efficiency times acceptance maps for the flavour democratic model with x = 0.95, calculated for both signal regions with the 3^{rd} lepton veto applied.

For the τ-enriched scan with x=0.95, i.e. the model in which both the chargino decays to stau, neutrino only ($\widetilde{\chi}_1^{\pm} \to \widetilde{\tau}\nu$) and neutralino decay democratically, efficiency maps are shown in Figure 7.7 for both search regions with the third lepton veto. Figure 7.8 shows the same for an x value of 0.05. Almost no sensitivity is expected for the x=0.05 scenario as one of the two expected same-sign lepton is very soft due to the mass splitting between the slepton and chargino/neutralino.

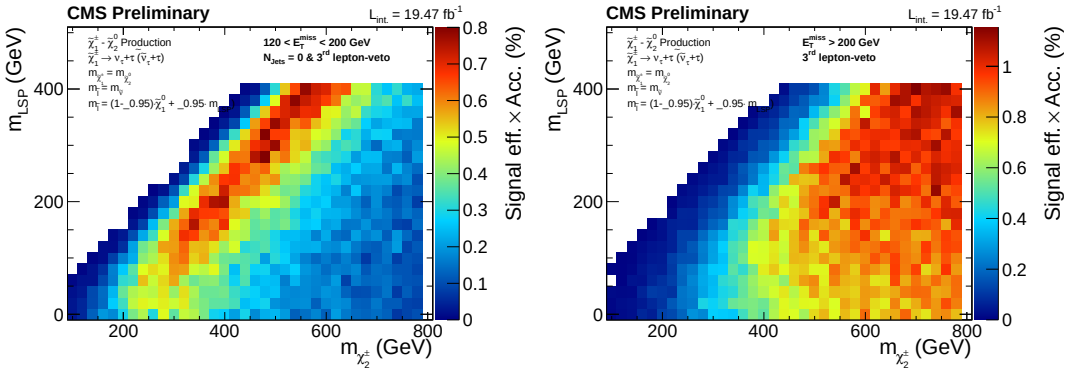


FIGURE 7.7: Efficiency times acceptance maps for the τ-enriched model with x = 0.95, calculated for both signal regions with the 3^{rd} lepton veto applied.

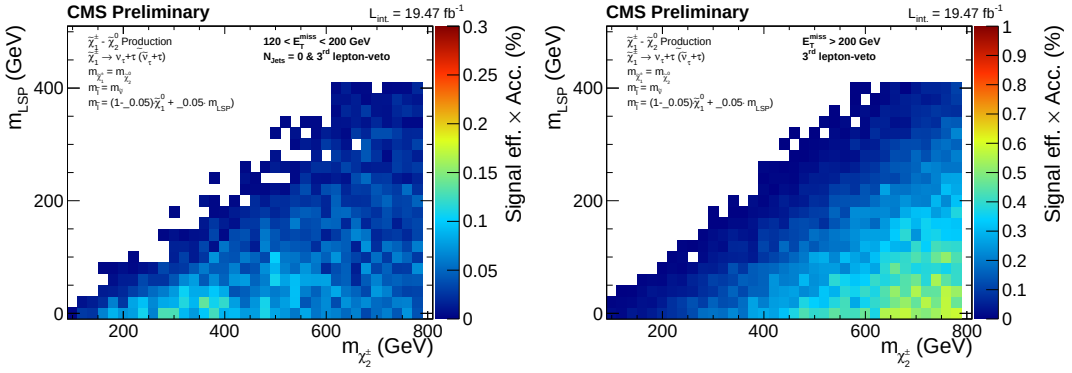


FIGURE 7.8: Efficiency times acceptance maps for the τ-enriched model with x = 0.05, calculated for both signal regions with the 3^{rd} lepton veto applied.

Generally speaking, efficiencies are of the order of a few percent and tend to be higher towards large mass splittings in the $E_T^{miss} > 200$ GeV region while efficiencies are highest close to the diagonal for the $E_T^{miss} > 120$ GeV region.

When applying a third lepton veto, a significant decrease in efficiency in the bulk of the plane is observed. One interesting feature, however, is the fairly stable efficiency close to the diagonal in the $E_T^{miss} > 120$ GeV signal region. This means that the same-sign dilepton channel is able to contribute the most in this region of the parameter space. Since efficiencies are low towards the diagonal for the $E_T^{miss} > 200$ GeV region, the application of the 3^{rd} lepton veto essentially renders this region insensitive to the discussed model.

A detailed description of the number of sources of systematic uncertainties and how to estimate them for the signal scan has been described in Section 6.3 and the same approach has been applied for these models. Table 7.6 shows the different sources of uncertainties that are being considered as well as the systematic uncertainty introduced

by each of them. Please note that the uncertainty varies from region to region, the average uncertainty is 20% for all the regions.

TABLE 7.6: Systematic uncertainties considered for the signal scan. The maximum uncertainty introduced by each source is also shown.

Source	Uncertainty (%)
Luminosity	4.4
Trigger Scale Factor	6
Lepton ID/Iso Scale Factor	10
b-tagging Scale (up/down)	0.5
MET Scale (up/down)	7
Jet Energy Scale (up/down)	2
Theoretical	7
Avg. Total Syst. uncertainty	20

Figure 7.9 displays the combined limits from the SS dilepton and three-lepton searches, interpreted in the flavour-democratic scenario for two values of $x_{\tilde{\ell}}$ (0.05 and 0.95). The figure depicts the 95% CL upper limit on the cross section times branching fraction in the $m_{\widetilde{\chi}_1^0}$ versus $m_{\widetilde{\chi}_2^0}$ $(= m_{\widetilde{\chi}_1^\pm})$ plane. The 50% branching fraction to three leptons is taken into account. The upper limit on the cross section times branching fraction generally becomes more stringent with the increasing mass difference between the chargino or heavy neutralino and the LSP. A drop in sensitivity is observed in the region where this mass difference leads to dilepton pairs with invariant masses close to that of the Z boson, and is caused by a higher rate for the WZ background.

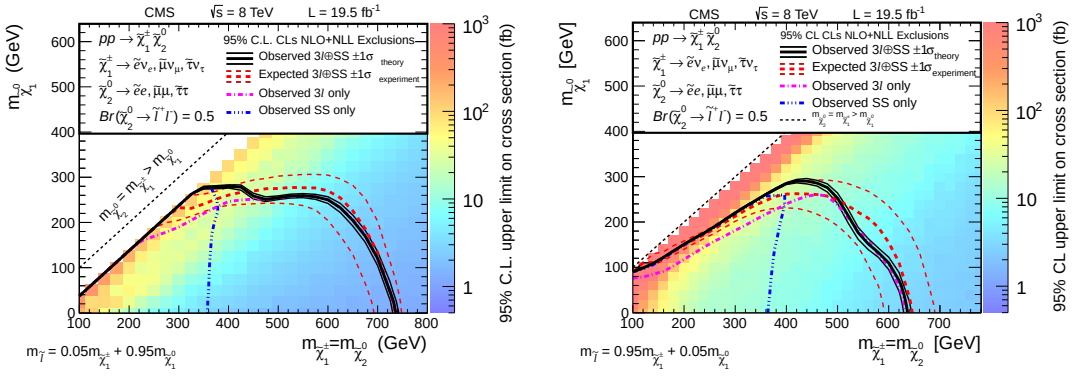


FIGURE 7.9: Interpretation of the results of the three-lepton search, the same-sign dilepton search, and their combination, in the flavour-democratic signal model with $x_{\tilde{\ell}} = 0.05$ (left) and $x_{\tilde{\ell}} = 0.95$ (right). The shading indicates the 95% CL upper limits on the cross section times branching fraction, and the contours the excluded regions assuming the NLO+NLL signal cross sections.

Figure 7.10 presents the corresponding limits for the τ-enriched scenario, for which the same-sign contribution is much less significant than for the flavour democratic scenarios.

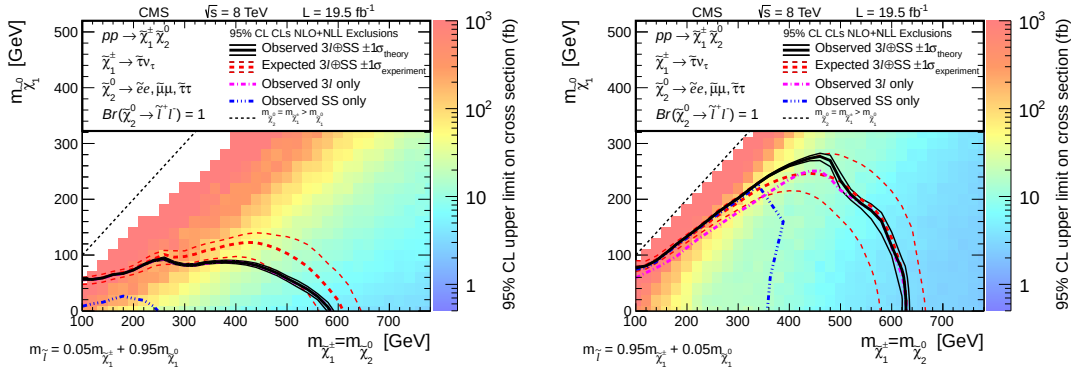


FIGURE 7.10: Interpretation of the results of the three-lepton search, the same-sign dilepton search, and their combination, for the τ-enriched signal model with $x_{\tilde{\ell}} = 0.05$ (left) and $x_{\tilde{\ell}} = 0.95$ (right). The shading indicates the 95% CL upper limits on the cross section times branching fraction, and the contours the excluded regions assuming the NLO+NLL signal cross sections.

As it has been described at the beginning of this chapter, the same-sign helps the three-lepton search in the the experimentally challenging region with very small mass splitting.

For the models with $x_{\tilde{\ell}} = 0.05$ the decay $\tilde{\tau} \to \tau\widetilde{\chi}_1^0$ is not kinematically allowed for signal scenarios with $m_{\widetilde{\chi}_2^0} - m_{\widetilde{\chi}_1^0} < 20m_\tau$. Therefore, in this region, the decay $\widetilde{\chi}_1^\pm \to \tilde{\tau}\nu_\tau$ is suppressed. Similarly, in the models with $x_{\tilde{\ell}} = 0.95$ the decay $\widetilde{\chi}_2^0 \to \tilde{\tau}\tau$ is not kinematically allowed in the region with $m_{\widetilde{\chi}_2^0} - m_{\widetilde{\chi}_1^0} < 20m_\tau$.

7.4.1 Summary

A search for weakly produced SUSY has been presented in this chapter using events with two same-sign leptons in the final state. This analysis uses, essentially, the same background methods as the analysis described in Chapter 6. Data agrees well with MC expectations and limits are set on four different models of chargino-neutralino production by combining the results with an analysis searching in the final state with three leptons, a third-lepton-veto is implemented for this purpose.

The same-sign analysis is able to probe chargino/neutralino masses up to 350 GeV, while the combination with the three-lepton search improves this result up to 700 GeV, the same-sign provides better sensitivity near the diagonal while the three-lepton final state is able to probe higher masses. This result improves significantly previous CMS publication and has been published in EPJ [79].

Chapter 8

Measurement of the $t\bar{t}$ production cross section in the $e\mu$ channel in pp collisions at $\sqrt{s} = 7$, 8 and 13 TeV

8.1 Introduction

The study of top quark ($t\bar{t}$) production provides one of the most crucial tests to the SM. The top quark is the heaviest known elementary particle and plays a crucial role in the electroweak symmetry breaking scale. At the LHC energies, the dominating production mechanism is gluon fusion and the total production cross-section, $\sigma_{t\bar{t}}$ can be accurately predicted by QCD calculations at NNLO. The precise experimental determination can therefore be extremely sensitive to potential physics beyond the Standard Model. For instance, in supersymmetry, $t\bar{t}$ pairs can be produced in the decays of supersymmetric particles increasing the $t\bar{t}$ production yields.

Studies of $t\bar{t}$ production have been performed in recent years by the ATLAS [83] and CMS [84] collaborations. All results are consistent with the SM prediction.

The most precise measurement of the total $t\bar{t}$ production cross section with the CMS experiment is described in this chapter at a centre-of-mass energy of $\sqrt{s} = 7$ and 8 TeV using total integrated luminosities of 5.0 fb^{-1} and 19.7 fb^{-1} respectively. The measurement is performed in the $e\mu$ final state, where both W bosons arising from the top quark decay decays leptonically. A more precise treatment of the systematic uncertainties have been perform with respect to the previously published paper. The first

measurement of the $t\bar{t}$ production cross-section at a centre-of-mass energy of $\sqrt{s} = 13$ TeV using the $e\mu$ final state is also presented and the end of this chapter.

8.2 Event simulation and theoretical calculations

As opposed to the searches described before, this analysis is, essentially a background free analysis where the great majority of the selected events correspond to $t\bar{t}$ events. Therefore, a very precise simulation of the detector and the different physical processes that play a role in this measurement is crucial.

The $t\bar{t}$ sample is simulated using the LO MADGRAPH event generator [47], which implements the relevant matrix elements up to three additional partons. The MADSPIN package [85] is used to incorporate spin correlation effects with matrix elements for up to three additional partons. The value of the top quark mass is fixed to $m_t = 172.5$ GeV and the proton structure is described by the parton distribution functions (PDF) CTEQ6L1 [86]. The generated events are then decayed using PYTHIA 6 [48] for parton showering and hadronization. Decays of τ leptons are handled with TAUOLA [87]. The CMS detector response is simulated using GEANT 4 [49].

Standard model background processes are simulated using MADGRAPH whenever available. W + jets , Z + jets and $t\bar{t}$V proces ses are simulated using MADGRAPH with up to two additional partons in the final state. The POWHEG [88] generator is used to simulate single top quark production, while PYTHIA is used for simulating diboson production. Although they could be generated with a lower precision generator, each sample is then normalised to the higher order available cross section calculation. These are taken from NNLO (W + jets and Z + jets), NLO+NNLL (single top quark tW-channel [89]), NLO (VV [90], $t\bar{t}W$ [91] and $t\bar{t}Z$ [92]) Correction factors are applied to nail down differences between simulation and data.

The $t\bar{t}$ simulated events are normalised using the best theoretical calculation available, NNLO accuracy in perturbative QCD, including resummation of NNLL soft gluon terms [93]. Assuming a top quark mass of $m_t = 172.5$ GeV, the predicted cross sections at different centre-of-mass energies are,

$$\sigma_{t\bar{t}(7\text{TeV})} = 177.3^{+4.7}_{-6.0}\,(\text{scale}) \pm 9.0\,(\text{PDF} + \alpha_s)\,\text{pb} \quad (8.1)$$

$$\sigma_{t\bar{t}(8\text{TeV})} = 252.9^{+6.4}_{-8.6}\,(\text{scale}) \pm 11.7\,(\text{PDF} + \alpha_s)\,\text{pb} \quad (8.2)$$

$$\sigma_{t\bar{t}(13\text{TeV})} = 832^{+20}_{-29}\,(\text{scale}) \pm 35\,(\text{PDF} + \alpha_s)\,\text{pb} \quad (8.3)$$

The first uncertainty is an estimate of the effect of missing higher order corrections and is determined by independent variations of the factorisation and renormalisation scales, μ_F and μ_R, by factors of two up and down from their default values. The second uncertainty is associated to variation in the PDF and α_S following the PDF4LHC recommendations [94].

8.3 Event selection

Events are selected using dilepton triggers that requires the presence of an electron and a muon, one of them having a $p_T > 17$ GeV and the other a $p_T > 8$ GeV. In the case of muons no further requirements are applied in the online selection, while for electrons, some basic identification and isolation cuts are applied to control the rate of the triggers paths. The value of the trigger efficiency for events passing the full selection criteria is measured in data using a set of orthogonal triggers, and it is found to be approximately 93%. Simulation events are corrected are re-weighted by a data-to-simulation scale factor that corrects from the differences in simulation with respect to data.

Events with an opposite sign electron-muon pair are selected if both leptons emerge from the same interaction vertex. If more than than one $e\mu$ pair passing the selection criteria described in the following, the pair with the largest value of the scalar sum of the transverse momenta of the leptons is selected. Events with τ leptons are consider only if they decay leptonically. Both leptons are required to have $p_T > 20$ GeV and pseudorapidity $|\eta| < 2.4$.

Lepton candidates are required to fulfill similar (but slightly looser) identification requirements as the ones described in Section 5.2. Furthermore, the leptons are required to be isolated from other particles in the event. This isolation, as explained in Section 4.3.2, is computed using the particle flow algorithm and corrected using the $\Delta\beta$ correction to suppress the contamination from pile-up. The isolation requirement is $Iso^{PF}_{rel} < 0.15$. That means that only 15% of the energy of the lepton is allowed in the cone.

Lepton identification and isolation efficiency is measured in data from Z boson events. The measured values for the combined isolation and identification efficiency are around 90% for muons and 80% for electrons. Simulation efficiencies are corrected to match the measured efficiency in data.

Jets are reconstructed using the particle flow algorithm [32]. The jet momenta is corrected to have a response that is flat in p_T and η (`L2L3`), as well as for contributions from pile-up events. (`L1FastJet`). In addition, `L2L3` residual corrections are applied on data. Jets are selected if they have $p_T > 30$ GeV and $|\eta| < 2.4$ and the angular distance

to leptons passing all selection criteria fulfils ΔR(jet, lepton) > 0.5. Events are selected if they contain at least two jets, this cut helps reducing the contribution from single-top background.

$t\bar{t}$ events are expected to contain jets arising from the hadronization of b quarks, thus requiring the presence of b jets in the event can reduce background contamination. Jets are identified as b-jets using the POG recommended combined-secondary-vertex (CSV) tagging algorithm with the medium working point for the discriminator of 0.679, which corresponding to an identification efficiency for b jets of about 70% and a misidentification probability of about 10% for charm jets and 0.1% for light-flavour jets (u,d,s and gluons). The last piece of the event selection for 7 and 8 TeV analysis, is requiring the presence of at least one b-tagged jet and it furthers reduce DY and diboson background.

8.4 Cross section extraction

After applying these cuts, the resulting selected sample is already very pure in $t\bar{t}$ events (close to 95%) A simple cut and count approach is followed for extracting the cross section for $t\bar{t}$ production:

$$\sigma_{t\bar{t}} = \frac{N_{\text{obs}} - N_{\text{bkg}}}{\epsilon \cdot \mathcal{L}} \tag{8.4}$$

where, N_{obs} is the number of selected events in data, N_{bkg} is the number of expected background events that could contaminate our data sample. ϵ accounts for the signal efficiency, acceptance and branching ratios of the selection and $\mathcal{L}$ is the total integrated luminosity used for the measurement.

8.5 Background estimation methods

Backgrounds arising from single-top quark, diboson events and associated production of $t\bar{t}$ and a vector boson, in which at least two prompt leptons are produced, are determined from MC. Other background sources, such as Z + jets background (DY) and processes with at lease one non-prompt lepton (arising from a misidentified jet or from heavy flavour decays) are derived using a data-driven technique as the simulation does not provide a very good description of these processes.

8.5.1 Drell-Yan background estimation

The Drell-Yan (DY) background contamination arises from the following processes: $Z/\gamma^* \to \tau\tau \to \mathrm{e}\mu\nu_e\nu_\tau\nu_\mu\nu_\tau$.

The contribution from this background, although very small after requiring the presence of a b-tagged jets, is estimated using the so called $\mathrm{R_{out/in}}$ method. The method is based on the measurement of the ratio of DY events outside and inside the dilepton invariant mass window ($76 < m_{ll} < 106$ GeV). This ratio is used to estimate the expected number of events outside the Z mass window and then to extract a data-to-simulation correction factor to normalize the simulation.

Two control regions selected using events passing all the selection cuts but with a same-flavour pair of leptons are used for estimating this background. This control region will contain mostly DY events, but also some contribution from flavour symmetric processes.

Although this control region will be very pure in DY events, some contamination from flavour symmetric backgrounds can still be present inside the Z mass window. $e\mu$ events, scaled accordingly to account for the different reconstruction efficiencies of electron and muons are used to subtract this contamination from the event yields in the same-flavour dileptonic sample that is used for measuring this background.

The number of events outside the Z-veto can be estimated from data as:

$$N_{out}^{l^+l^-,obs} = R_{out/in}^{l^+l-}(N_{in}^{l^+l^-} - 0.5N_{in}^{e\mu}k_{ll}) \tag{8.5}$$

where $ll = \mu\mu$ or ee and $R_{out/in}$ is the ratio of the number of events outside/inside the Z mass window from a DY MC sample: $R_{out/in} = \frac{N_{DY\,MC}^{out}}{N_{DY\,MC}^{in}}$, k is a correction factor to take into account the differences between electron and muon reconstruction, it is calculated using the events in the Z peak region passing the standard dilepton and jet selections, and can be expressed as:

$$k_{ee} = \sqrt{\frac{N_{in}^{e^+e^-}}{N_{in}^{\mu^+\mu^-}}}$$

$$k_{\mu\mu} = \sqrt{\frac{N_{in}^{\mu^+\mu^-}}{N_{in}^{e^+e^-}}}$$

As the $R_{out/in}$ factor relies on MC, we studied its dependence in MC as a function of the applied cuts. Figure 8.1 shows the variation of the $R_{out/in}$ factor as a function of the different cut levels applied, the maximum difference, 50% is assigned as systematic uncertainty that is propagated to the data-driven estimate.

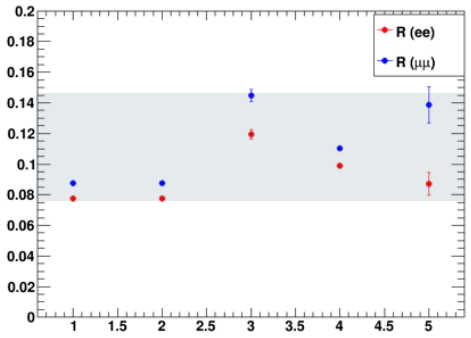


FIGURE 8.1: Evolution of the Ratio for the different cuts applied in the analysis between the number of events outside/inside the Z mass window, measured in MC events for electrons and muons. The maximum difference across different cut levels, 50%, is assign as systematic uncertainty that will be propagated to the final estimate.

TABLE 8.1: Derived correction factor to be applied to Monte Carlo using the $R_{out/in}$ method.

	ee	$\mu\mu$	$e\mu$
N_{in}	3944.0 ± 62.8	5265.0 ± 72.6	
k_{ll}	0.890 ± 0.011	1.15 ± 0.011	
$R_{out/in}$	0.086 ± 0.044	0.139 ± 0.071	
N_{out} (estimation)	165.6 ± 2.5	377.7 ± 5.0	
SF (DD/MC)	1.479 ± 0.162	1.652 ± 0.131	1.563 ± 0.148

The estimation extracted from this method is then compared to the one from the simulation and a data-to-simulation correction factor is applied to the MC to match the data-driven estimation for both the ee and $\mu\mu$ control samples. The data-to-simulation correction factor for the $e\mu$ channel is simply the square root of the product of the correction factor from ee and $\mu\mu$ control samples. As shown Table 8.1, SF_{DY}^{emu} is then 1.563 ± 0.148 (stat) with a systematic variation of the order of 15-20% obtained from the variation of $R_{out/in}$ at the different levels of selection in the analysis (2l, 2l+2j and 2l+2j+1btag) propagated to the SF.

It has also to be noted that DY here is around 10% of the total background, dominated by tW (67% of the total), and compared to the signal is of the order of 0.7% (so, this is basically its contribution to the measurement) while in [84] it was 22% of the background, and 2% with respect to the signal. This difference comes basically from the use of the tighter b-tagging working point that reduces such background.

8.5.2 Non-prompt lepton background estimation

Non-W/Z backgrounds arise from processes with one prompt-lepton (decaying from a W or a Z boson) and one non-prompt lepton that passes the isolation and identification criteria. The latter will be also referred as a "fake lepton". Most of the contribution to such backgrounds is due to W+jets production and most likely, semi-leptonic $t\bar{t}$ production.

Non-W/Z backgrounds have been substantially reduced already by the tight isolation and identification criteria. A data-driven technique inspired in the searches described in Chapters 5-7.

As it has been described before, when requiring events with two same-sign leptons, most of the selected events are yield from instrumental backgrounds (mainly non-prompt leptons and leptons with charge-misidentification) with some small contribution from real prompt-leptons from processes such as $t\bar{t}V$. These instrumental backgrounds are totally independent of the charge of the lepton pair and hence, a similar contribution of such events is expected in the opposite-sign sample.

The data control sample used for estimating this background, is built by simply inverting the opposite-sign requirement. Background contribution to the same-sign yields such as charge mis-identified electrons and processes with two prompt-leptons are subtracted from the same-sign yields in data. This number is then corrected by the ratio of opposite-sign over same-sign events coming from non-prompt lepton backgrounds, which is measured purely from MC, to account for the possible differences in the charge composition of the non-prompt lepton sample (as the major contribution to the non-prompt lepton background arises from semileptonic $t\bar{t}$, this value is expected to be close to unity). Table 8.2 shows the background composition of the same-sign control sample for ee, $\mu\mu$ and $e\mu$ events as well as the final data-driven estimation for the 19.7 fb^{-1} at 8 TeV.

As this procedure is meant to estimate the instrumental backgrounds there should be no dependence with the lepton flavour neither with the selection cuts. The variation of this ratio as a function of the lepton flavour and the cut level is shown in Figure 8.2, this should be very well absorbed by a 30% systematic uncertainty that we add to this background estimation. Needless to say that the impact on the cross-section extraction of this uncertainty is almost negligible.

TABLE 8.2: Background composition and non-prompt lepton background estimation. Table also shows the ratio of opposite-sign and same-sign events that is used to correct the estimation.

Source	ee	$\mu\mu$	$e\mu$
$t\bar{t}$ dilepton	54.8 ± 1.9	3.2 ± 0.5	98.2 ± 2.6
Drell-Yan	0.0 ± 0.0	0.0 ± 0.0	4.2 ± 3.0
Single top quark	2.5 ± 0.1	0.1 ± 0.0	4.9 ± 0.2
Dibosons	0.9 ± 0.2	1.1 ± 0.2	3.6 ± 0.4
$t\bar{t}V$	10.8 ± 0.4	13.2 ± 0.5	35.0 ± 0.8
Total background	68.9 ± 2.0	17.6 ± 0.7	145.9 ± 4.1
Data	135	71	346
SS data - bkg	66.1 ± 2.0	53.4 ± 0.7	200.1 ± 4.1
Non-prompt lep (SS)	18.4 ± 2.4	16.1 ± 1.2	62.6 ± 3.1
Non-prompt lep (OS)	22.9 ± 2.5	20.4 ± 1.4	69.7 ± 2.4
R (OS/SS)	1.24 ± 0.21	1.27 ± 0.13	1.11 ± 0.07
Non-prompt estimation	82.0 ± 14.1	67.9 ± 1.0	222.7 ± 19.1

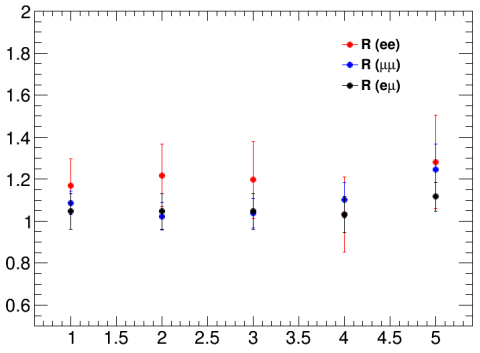


FIGURE 8.2: Evolution of the Ratio between OS/SS events with a non-prompt lepton as a function of the cut level and lepton flavour

8.6 Systematic uncertainties

An exhaustive treatment of all the systematic sources of uncertainties is crucial to achieve a very precise measurement of the $t\bar{t}$ cross-section. The better knowledge on the observables and the theoretical uncertainties the more precise measurement.

Table 8.3 summarises the magnitude of the systematic uncertainties on the $\sigma_{t\bar{t}}$ measurement.

In the remaining part of the section, a brief description of the various uncertainties and how to calculate them will be covered:

TABLE 8.3: Summary of the individual contributions to the systematic uncertainty on the $\sigma_{t\bar{t}}$ measurement. The uncertainties are given in percentage.

Source	Uncertainty [%]	
	7 TeV	8 TeV
Trigger efficiencies	1.3	1.14
Lepton efficiencies	1.1	1.44
Lepton energy scale	0.2	0.06
Jet energy scale	1.9	1.97
Jet energy resolution	0.7	0.01
b-tagging/mistag	1.4	1.33
Pileup	0.7	0.86
Top p_T	0.4	1.39
PDF	1.0	1.10
CR	0.4	0.27
Scale of QCD (μ)	0.9	1.17
Matching partons to showers	0.9	0.98
Drell-Yan DD	0.1	0.11
Non-W/Z	0.3	0.18
tW	1.3	1.31
VV	0.1	0.05
Rare	—	0.11
Total Systematic	3.75	4.13
Statistical	1.5	0.66
Integrated luminosity	2.2	2.6
Total	4.6	4.9

Trigger and lepton efficiencies and lepton momentum scale

Correction factors are applied to simulation only to reduce any potential discrepancy with data. Lepton efficiencies are measured using a Tag-and-Probe technique using a data sample enriched in Z events. These correction factors are derived as a function of p_T and η. A conservative flat uncertainty of 1% per lepton is considered to cover both the error of the measurement itself and any kinematic difference between Z events used for the efficiency measurement and $t\bar{t}$ events.

Trigger efficiencies are measured using a control sample selected by a set of orthogonal triggers. Correction factors are applied to correct simulation as a function of the η of the leptons. A conservative 1.14% is assigned to account for this effect.

The relative bias $\Delta(p_T)/p_T$ in reconstructed lepton transverse momentum with respect to its true value that could be caused by imperfect knowledge of the magnetic field is generally constant as a function of momentum. Similarly, inaccuracies in the modelling of the energy loss (dependent on the material distribution) produce relative biases that are essentially independent of the momentum. On the other hand, alignment effects produce

relative biases that generally increase linearly with momentum. A two percent variation (up/down) is applied to the momentum of every lepton to account for any systematic effect not accounted by the reconstructions. The overall impact is very small, of the order of 0.03%.

Jet energy scale and resolution

Another source of systematic uncertainty is associated with the jet energy scale (JES) correction. This systematic uncertainty varies between 5% and 2% in the p_T range 40-100 GeV for jets with $|\eta| < 2.4$ [54]. It is evaluated on a jet-by-jet basis, and its effect is propagated through all the used observables in the analysis.

The effect of the jet-energy-resolution (JER) is evaluated by smearing the jet energy of each jet with a gaussian of width of the jet-energy-resolution in MC, however the effect is much smaller than the JES uncertainty. The uncertainty is calculated separately from the JES one and propagated to all observables in the analysis, providing a full knowledge of the correlation of these systematics.

The total effect of the JES and JER are of the order of 2% and 0.03% respectively.

B-tagging uncertainties

The b-tagging efficiency for b-quark jets with $|\eta| < 2.4$, measured in data using samples enriched in $t\bar{t}$ and muon-jet events, has a p_T-averaged value of 0.72. The mis-tagging probability for charm-quark jets is approximately 20%, while for jets originating from light-flavour quarks or gluons it is of the order of 1%.

The total uncertainty of the b-tagging efficiency is determined by simultaneously varying the efficiencies to tag a bottom, charm, or light quark up and down by their uncertainties [37], the uncertainty is calculated on a jet-by-jet basis and separately for real b-jets (quoted as b-tagging uncertainty) and for light flavour jets (mis-tagging). The total effect is measured to be 1.30% and 0.03% respectively.

pile-up re-weighting

The samples used in this analysis were generated using a pile-up distribution with a slightly higher mean than the one in data. Hence, the MC samples are re-weighted to have the same pile-up distribution. To assess the impact of any possible systematic effect from the lack of knowledge of the pile-up profile, a 5% variation (up and down)

of this distribution is implemented. The overall impact in the analysis is smaller than 0.05%.

Other experimental uncertainties

For each background source derived from MC, a conservative uncertainty of 30% is assumed. In the case of the single-top background, this uncertainty covers the theory uncertainty on the absolute rate including uncertainties due to PDFs.

The uncertainty on the luminosity ("lumi") measurement is 2.6%.

Top p_T re-weighting

In the normalised differential top-quark-pair cross section analysis [95] the shape of the p_T spectrum of the individual top quarks in data was found to be softer than predicted by the various simulations while the available NNLO prediction delivers a reasonable description. Based on this measurements, event-by-event correction factors have been derived to test the potential impact of the modelling of the top quark p_T spectrum on the analysis. The origin of difference between data and simulation is not clear at the moment. It might be due to higher order QCD corrections, EWK corrections, non-resonant production of $t\bar{t}$ -like final states (e.g. $gg \rightarrow b\,\bar{b}\,\mu\,\nu\,e\,\nu$) or something completely unknown.

As this effect is very small in an inclusive cross-section measurement, no re-weight is applied in this analysis and the observed difference when applying the re-wright and when not applying it is taken as systematic uncertainty, which is evaluated to be 0.76%.

QCD scale and matrix-element uncertainties

The uncertainty on modelling of the hard-production process ("Q2 scale") is assessed through changes in the renormalisation and factorisation scales in the MADGRAPH sample by factors of two and 0.5 relative to their common nominal value, which is set to the Q of the hard process. In MADGRAPH, Q is defined by $Q^2 = m_t^2 + \sum p_T^2$, where the sum is over all additional final state partons in the matrix element. The impact of this uncertainty in the analysis is quite substantial, of the order of 1.3%.

The impact of the choice of the scale that separates the description of jet production through matrix elements or parton shower ("ME/PS matching") in MADGRAPH is studied by changing its reference value of 20 GeV to 40 GeV and to 10 GeV. The effect is measured to be 0.88%.

Other modelling uncertainties

The differences in results between using POWHEG [88] for the $t\bar{t}$ simulation instead of MADGRAPH is taken as an additional modelling uncertainty ("MG+PY → PH+PY").

The uncertainties from ambiguities in modelling colour reconnection effects ("Color reconnection") are estimated by comparing simulations of an underlying event tune including colour reconnection to a tune without it.

The uncertainty from the choice of PDFs ("PDF") is determined by re-weighting the sample of simulated $t\bar{t}$ events according to the 52 CT10 error PDF sets [94], at 90% confidence level.

8.7 Results

Table 8.4 shows the total number of events observed in data and the number of signal and background events expected from simulation or estimates from data. A good agreement between data and expected number of events is observed.

TABLE 8.4: Total number of events observed in data and the number of signal and background events expected background events. Both statistical and systematic errors are shown.

Source	Number of $e^{\pm}\mu^{\mp}$ events	
	7 TeV	8 TeV
DY	$22.1 \pm 3.1 \pm 3.3$	$173.3 \pm 25.1 \pm 26.0$
Non-W/Z	$51.0 \pm 0.7 \pm 15.3$	$145.9 \pm 14.8 \pm 43.8$
Single top quark (tW)	$204.0 \pm 3.1 \pm 61.2$	$1033.6 \pm 2.9 \pm 313.8$
VV	$6.9 \pm 0.6 \pm 2.1$	$35.4 \pm 1.9 \pm 11.1$
$t\bar{t}V$	$--$	$83.6 \pm 1.3 \pm 25.5$
Total background	$284.0 \pm 16.0 \pm 63.2$	$1471.7 \pm 46.7 \pm 319.1$
$t\bar{t}$ dilepton signal	$5008.2 \pm 15.4 \pm 188.0$	$24439.6 \pm 43.6 \pm 956.4$
Data	4970	25441

Figure 8.3 shows the b-jet multiplicity in events passing the full event selection but before the b-jet requirement.

The cross section $\sigma_{t\bar{t}}$ is determined by subtracting from the number of observed data events the estimated number of events from all background sources and dividing the resulting number by the total efficiency ϵ_{total} and the integrated luminosity of the data sample.

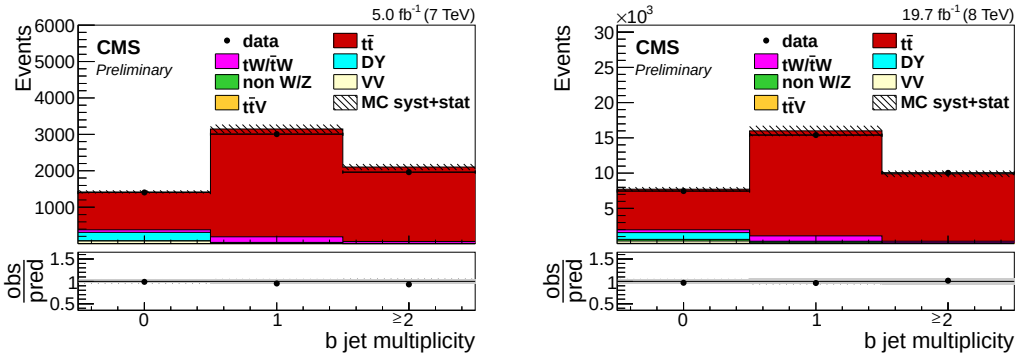


FIGURE 8.3: Data-MC comparisons of the b-jet multiplicity distributions in the $e\mu$ channel after the 2 jets selection requirement. Uncertainties are statistical only.

The total efficiency ϵ_{total} is the product of event acceptance, selection efficiency and branching fraction of the selected $t\bar{t}$ final state, as estimated from simulation for a top-quark mass of 172.5GeV. It also takes into account the contributions from $W \to \tau\nu_\tau$, with leptonic τ decays. The estimated values are 0.506 ± 0.020 for the 8 TeV dataset and 0.559 ± 0.021 for the 7 TeV dataset.

The measured cross section is therefore:

$$\sigma_{t\bar{t}}\ (7\ \text{TeV}) = 165.9 \pm 2.5\ (\text{stat}) \pm 6.2\ (\text{syst}) \pm 3.6\ (\text{lumi})\ \text{pb} \quad (8.6)$$

$$\sigma_{t\bar{t}}\ (8\ \text{TeV}) = 247.6 \pm 1.6\ (\text{stat}) \pm 9.7\ (\text{syst}) \pm 6.3\ (\text{lumi})\ \text{pb} \quad (8.7)$$

which agrees with the recent NNLO+NNLL calculation [93] within uncertainties. Being the most precise measure of the $t\bar{t}$ cross-section measured by CMS, the total uncertainty, including systematics, statistical and luminosity is below 5%.

8.8 First measurement of the top quark pair production cross section in proton-proton collision at $\sqrt{s} = 13$ TeV

In this section the first measurement of the $t\bar{t}$ production cross section using data corresponding to an integrated luminosity of 42 pb^{-1} recorded by the CMS experiment in pp collisions at a centre-mass-energy of 13 TeV is described. This analysis is a natural continuation of the one previously described in this Chapter as it uses the very same methods described before.

8.8.1 Monte Carlo simulation and event selection

Monte Carlo simulation

Several MC event generators are used to simulate signal and background events. The NLO POWHEG(v2) [88] generator is used for $t\bar{t}$ events, assuming a top quark mass of 172.5GeV. These events are interfaced to PYTHIA(v8.2) [96] to simulate parton showering, hadronization, and the underlying event. An alternative sample is obtained by showering the events with HERWIG++ [97]. Another sample of $t\bar{t}$ events is generated using MG5_aMC@NLO [98] and MADSPIN [85], as well as PYTHIA for parton showering.

The MG5_aMC@NLO generator is also used to simulate W + jets events and Z + jets. The normalisation is taken from data as described below. Single top quark events are simulated using POWHEG(v1) [88] and PYTHIA, and are normalised to the approximate next-to-next-to-leading order (NNLO) cross sections [99]. The contributions from VV are simulated with PYTHIA, and are normalised to the next-to-leading order (NLO) cross sections [90]. All other backgrounds are estimated from control samples extracted from collision data. The simulated samples include additional interactions per bunch crossing (pile-up), with distributions that are corrected to match the observed data.

The $t\bar{t}$ simulated events are normalised using the best theoretical calculation available, NNLO accuracy in perturvative QCD, including resummation of NNLL soft gluon terms [93]. Assuming a top quark mass of $m_t = 172.5$ GeV, the predicted cross section is,

$$\sigma_{t\bar{t}(13\text{TeV})} = 832^{+20}_{-29}\,(\text{scale}) \pm 35\,(\text{PDF} + \alpha_s)\,\text{pb} \tag{8.8}$$

The first uncertainty is an estimate of the effect of missing higher order corrections and is determined by independent variations of the factorisation and renormalisation scales, μ_F and μ_R, by factors of two up and down from their default values. The second uncertainty is associated to variation in the PDF and α_S following the PDF4LHC recommendations [94].

Event selection

The event selection is very similar to the one described in Section 8.3. Events are selected by a trigger that requires the presence of one electron and one muon, the leading lepton

having a $p_T > 17$ GeV and the trailing electron (muon) having $p_T > 12$ (8) GeV. The efficiency of these triggers is measured in data using triggers based on p_T imbalance in the event and it is measured to be 0.91 ± 0.05. A correction factor is applied to simulation to match measured values in data.

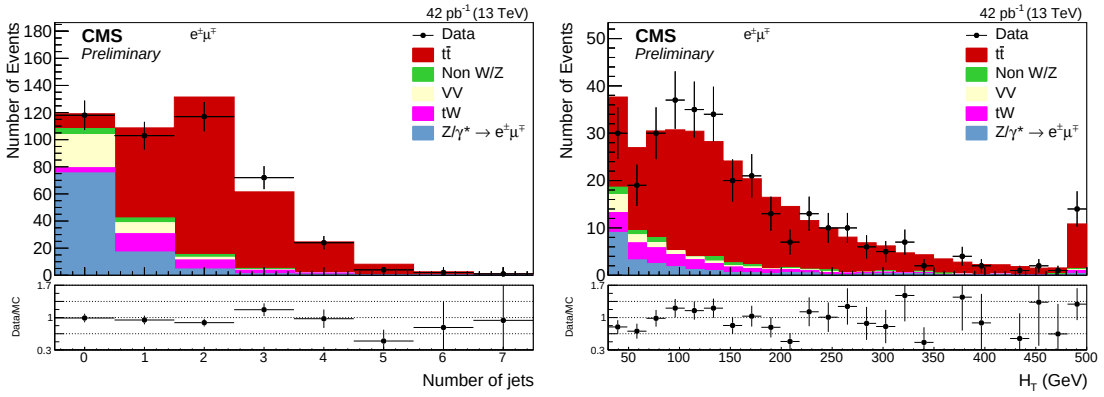
The efficiency of the lepton selection is measured using a tag-and-probe method in same-flavour dilepton events enriched in Z boson candidates. The measured values for the combined identification and isolation efficiencies are typically 0.92 for muons and 0.77 for electrons. Based on a comparison of lepton selection efficiencies in data and simulation, the event yield in simulation is corrected using p_T - and η-dependent correction factors to provide consistency with data.

The most outstanding difference with respect to the previously described result is the non usage of the b-tagging information to further discriminate between signal and background. Events with two opposite-sign, different flavour leptons (e, μ), with an invariant mass greater than 20 GeV and two identified jets with p_T greater than 30 GeV are considered.

8.8.2 Background determination and control plots

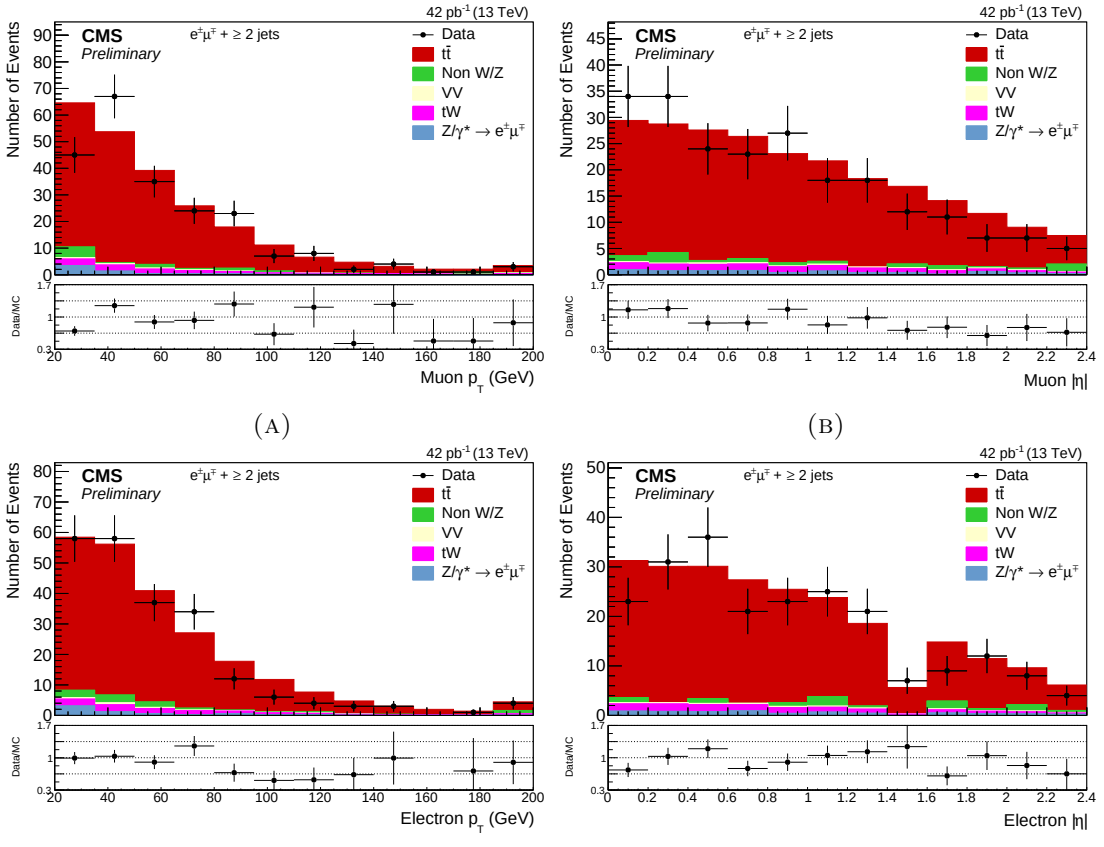
Backgrounds are determined using the same methods as described in Section 8.5. Yields from single top quark and diboson is estimated from simulation while DY background and non-prompt background is derived from the previously described data-driven techniques. The data-to-simulation factor for the DY background is found to be 1.06 ± 0.17. A same-sign control sample is used to derive the non-W/Z-lepton background using the method described in Section 8.5.2.

Figure 8.4 (a) shows the multiplicity of jets, and (b) the the scalar sum the transverse momenta of all jets (H_T) for events passing the dilepton criteria. After requiring at least two jets, Figs. 8.5 show the p_T and η distributions of the highest-p_T muons and electrons, respectively, and Figs. 8.6 show the p_T and η distributions of the highest- and second highest-p_T jets, respectively. The ratios of the data to the sum of simulations and data-based predictions for the signal and backgrounds are shown in the bottom panels of all figures. Agreement between data and the predictions for signal and background is observed.



(A) (B)

FIGURE 8.4: The distributions for 8.4a the jet multiplicity and 8.4b H_T in events passing the dilepton criteria. The expected distributions for $t\bar{t}$ signal and individual backgrounds are shown after data-based corrections are applied; the last bin contains the overflow events. The ratios of data to the sum of the expected yields are given at the bottom of each panel.



(A) (B)

(C) (D)

FIGURE 8.5: The distributions for 8.5a p_T and 8.5b η for the highest-p_T muon, and for 8.5c p_T and 8.5d η for the leading electron after all selections. The expected distributions for $t\bar{t}$ signal and individual backgrounds are shown after data-based corrections are applied; for the left plots (a,c) the last bin contains the overflow events. The ratios of data to the sum of the expected yields are given at the bottom of each panel.

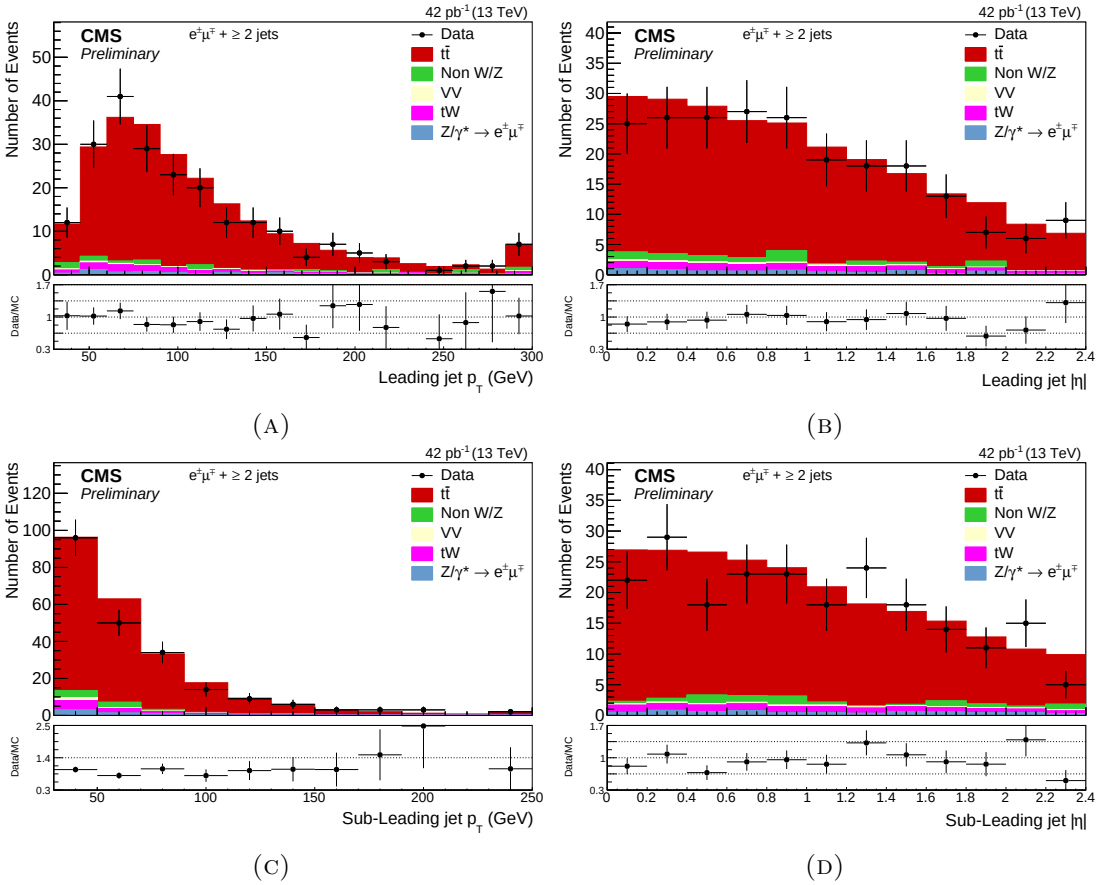


FIGURE 8.6: The distributions for 8.6a p_T and 8.6b η for the highest-p_T jet, and for 8.6c p_T and 8.6d η for the second highest-p_T jet after all selections. The expected distributions for $t\bar{t}$ signal and individual backgrounds are shown after data-based corrections are applied; for the left plots 8.6a,8.6c the last bin contains the overflow events. The ratios of data to the sum of the expected yields are given at the bottom of each panel.

8.8.3 Systematic uncertainties

Systematic uncertainties are assessed following the same methods described in Section 8.6. Table 8.5 summarises the statistical and systematic uncertainties on the $t\bar{t}$ production cross section arising from different sources.

The uncertainties on the correction factors applied to simulation to account for data-mc differences for trigger and electron (muon) efficiencies are 5% and 4% (1-3%) respectively. The modeling on the lepton energy scale is estimated to be of 1% for electrons and 0.5% for muons.

The impact of the uncertainty on the JES and JER are estimated by evaluating the effect of varying the jet momenta by 4% for the JES and by η-dependent variation on the JER scale factors. The uncertainty assigned to the number of pile-up events in simulation is obtained by changing the inelastic cross section by $\pm 5\%$

The potential bias related to the missing high-order diagrams in POWHEG are estimated by changing the renormalisation and factorisation scale simultaneously up and down by a factor two. Furthermore, the prediction of two NLO generators (POWHEG and MG5_aMC@NLO) for $t\bar{t}$ production and compared using PYTHIA for hadronization and showering. The hadronization uncertainty is accounted by comparing two samples generated by POWHEG, where the hadronization is modelled either by PYTHIA or HERWIG++. Finally, the uncertainty from the choice of PDF is determined by re-weighting the sample of simulated $t\bar{t}$ events according to the 100 NNPDF3.0 error PDF sets [100].

Concerning background estimation and normalisation, same values from 7 and 8 TeV are assumed. The biggest uncertainty for this first measurement arises from the uncertainty on the total integrated luminosity, which is estimated to be of the order of 12% after a preliminary x-y beam scan performed in July 2015.

TABLE 8.5: Summary of the individual contributions to the systematic uncertainty on the $\sigma_{t\bar{t}}$ measurement. The uncertainties are given in pb and as relative uncertainties. The statistical uncertainty and the total uncertainty on the result are also given.

Source	$\Delta\sigma_{t\bar{t}}$ (pb)	$\Delta\sigma_{t\bar{t}}/\sigma_{t\bar{t}}$ (%)
Data statistics	60	7.7
Trigger efficiencies	39	5.0
Lepton efficiencies	33	4.3
Lepton energy scale	< 1	≤ 0.1
Jet energy scale	20	2.6
Jet energy resolution	< 1	≤ 0.1
Pileup	2.8	0.4
Scale (μ_F and μ_R)	1.5	0.2
$t\bar{t}$ NLO generator	15	1.9
$t\bar{t}$ hadronization	14	1.8
PDF	12	1.5
Single top quark	14	1.8
VV	3.5	0.5
Drell–Yan	3.9	0.5
Non-W/Z leptons	8	1.0
Total systematic (no integrated luminosity)	62	8.0
Integrated luminosity	93	12
Total	126	16.4

8.8.4 Results

Table 8.6 shows the total number of events observed and the number of signal and background events expected. A good agreement between data and expected number of events is observed.

Figure 8.7a shows the distribution of the invariant dilepton mass $M_{e\mu}$ which can probe, for example, the existence of a new heavy object decaying into a top quark pair. Figure 8.7b shows the difference in azimuthal angle between the two selected leptons $\Delta\Phi(e^{\pm}, \mu^{\mp})$ and explores the correlation between the top and antitop quark spins [101, 102]. For both distributions data are in agreement with the SM predictions.

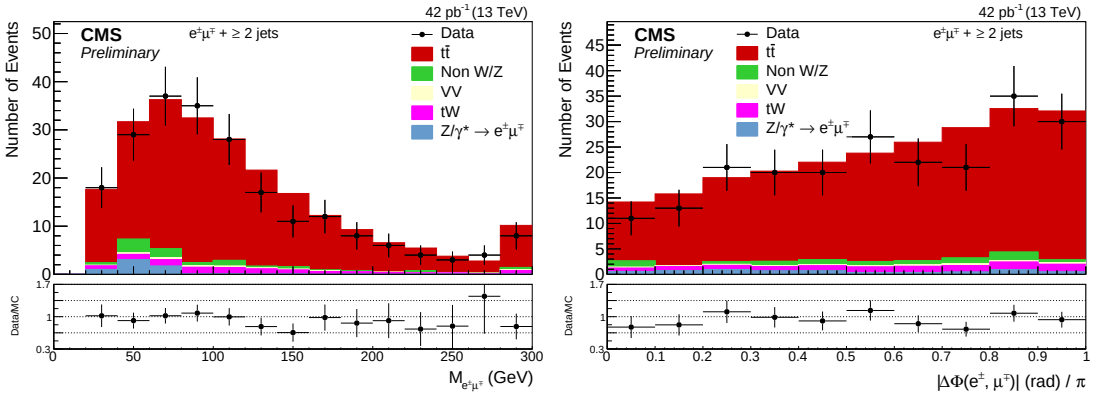
The total efficiency ϵ_{total} is the product of event acceptance, selection efficiency and branching fraction of the selected $t\bar{t}$ final state, as estimated from simulation for a top-quark mass of 172.5 GeV is $(0.60 \pm 0.04)\%$, including statistical and systematic uncertainty. The extracted cross section value is:

$$\sigma_{t\bar{t}}\ (13\ \text{TeV}) \quad = \quad 772 \ \pm \ 60\ (\text{stat}) \ \pm \ 62\ (\text{syst}) \ \pm \ 93\ (\text{lumi})\ \text{pb} \tag{8.9}$$

in good agreement with the most accurate theoretical calculation 8.8.

TABLE 8.6: Number of dilepton events obtained after applying the full selection. The results are given for the individual sources of background, $t\bar{t}$ signal with a top quark mass of 172.5 GeV and $\sigma_{t\bar{t}}^{\text{NNLO+NNLL}} = 832^{+40}_{-46}$ pb, and data. The uncertainties correspond to the statistical and systematic components added in quadrature.

Source	Number of $e^{\pm}\mu^{\mp}$ events
Drell–Yan	6.4 ± 1.2
Non-W/Z leptons	8.5 ± 4.3
Single top quark	10.6 ± 3.4
VV	2.6 ± 0.9
Total background	28.1 ± 5.7
$t\bar{t}$ dilepton signal	207 ± 16
Data	220



(A) (B)

FIGURE 8.7: Distributions in 8.7a the dilepton invariant mass, and 8.7b the difference in the azimuthal angle between the two selected leptons, after all selections. For the first plot the last bin contains the overflow events. The ratios of data to the sum of the expected yields are given at the bottom.

8.9 Summary

A measurement of the inclusive $t\bar{t}$ production cross section in proton-proton collisions at the LHC using the full 2011 and 2012 data samples of 5.0fb^{-1} at $\sqrt{s}$ = 7 TeV and 19.7fb^{-1} at $\sqrt{s}$ = 8 TeV has been described in this Chapter. A first measurement using the first 42pb^{-1} at $\sqrt{s}$ = 13 TeV is also presented. The analysis is performed in the dilepton $t\bar{t} \to e\mu\nu\bar{\nu}b\bar{b}$ decay channel. The cross sections are extracted using a cut-and-count approach, assuming a top quark mass of m_t = 172.5 GeV, the results are:

$$\begin{aligned}
\sigma_{t\bar{t}}\ (7\ \text{TeV}) &= 165.9 \pm 2.5\ (\text{stat}) \pm 6.2\ (\text{syst}) \pm 3.6\ (\text{lumi})\ \text{pb} \\
\sigma_{t\bar{t}}\ (8\ \text{TeV}) &= 247.6 \pm 1.6\ (\text{stat}) \pm 9.7\ (\text{syst}) \pm 6.3\ (\text{lumi})\ \text{pb} \\
\sigma_{t\bar{t}}\ (13\ \text{TeV}) &= 772 \pm 60\ (\text{stat}) \pm 62\ (\text{syst}) \pm 93\ (\text{lumi})\ \text{pb}
\end{aligned}$$

in good agreement with recent NNLO QCD calculations. This measurement is also in agreement with a recent measurement from the ATLAS Collaboration [103]. The 7 and 8 TeV measurements will be publish together with a binned likelihood fit to the p_T of the non b-tagged jet with the lowest p_T in the event, using different categories of b-tagged and additional non-b-tagged jets [104]. The 13 TeV measurement will be sent to PRL soon [105].

Figure 8.8 shows the latest results for $\sigma_{t\bar{t}}$ for both dilepton [105] and lepton+jets channel [106] together with previous CMS results at $\sqrt{s}$ = 7 TeV and 8 TeV [84, 107–109], the LHC combination at 8 TeV [110], and the combination of the Tevatron [111]. The data are compared to the NNLO+NNLL predictions as a function of centre-of-mass energy for $p\bar{p}$ and pp collisions [93]. A beautiful agreement with SM predictions is observed.

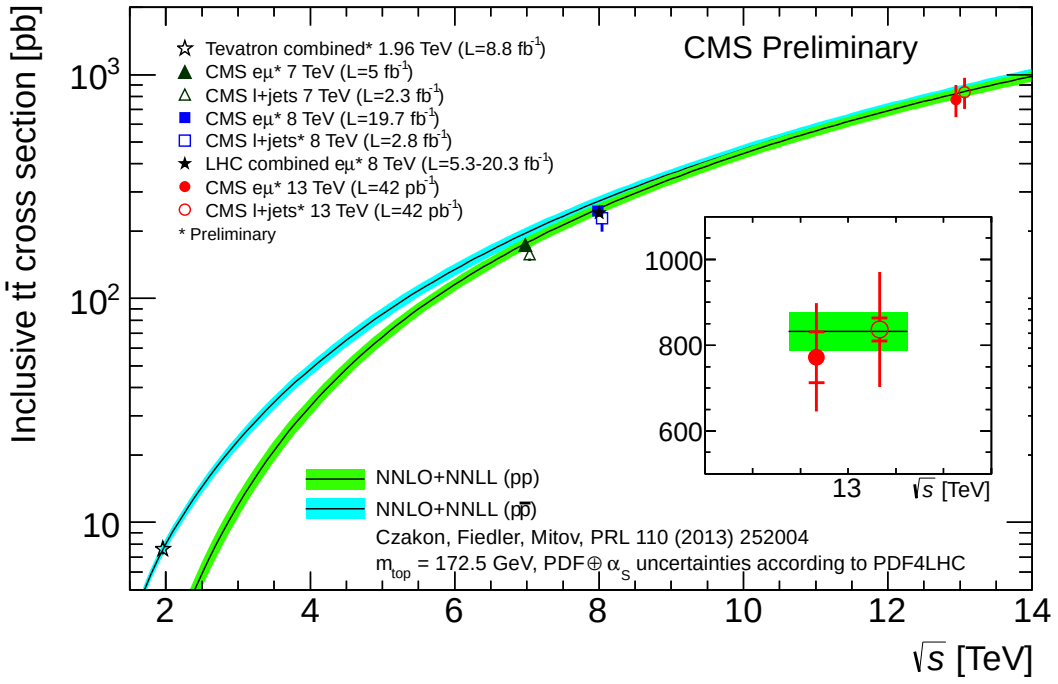


FIGURE 8.8: Top quark pair production cross section in $p\bar{p}$ and pp collisions as a function of center-of-mass energy. The Tevatron combination at $\sqrt{s} = 1.96$ TeV are displayed, as well as CMS results at 7, 8 and 13 TeV in the dilepton and semileptonic ("l+jets") channels. For the $e\mu$ channel, the LHC combination at 8 TeV is shown. The measurements are compared to the NNLO+NNLL theory predictions.

Chapter 9

Run II preparation and commissioning

The Long ShutDown I (LS I) of the LHC lasted for almost two years, many interventions were done to both the LHC and the experiments to prepare them for the increase of energy that was foreseen for 2015. On the analysis side, several improvements were developed to increase the expected sensitivity and favour a potential discovery during Run II.

9.1 Trigger Studies for Fake Rate Measurement in 2015 Data

The fake ratio is measured in a QCD-enriched region, that from now on, will call "measurement region". To select such events in data one needs to define an appropriated trigger strategy. A new trigger strategy has been developed for measuring the fake ratio in 2015.

In 2012, the strategy was to use one (or two) paths to populate the whole p_T range: from 10 to 50 GeV, although in most cases, only leptons with $p_T > 20$ GeV were used. Due to the exponentially falling spectrum of the p_T of the leptons, the number of selected events with higher p_T was very small, increasing the statistical uncertainty of the method.

During LSI a new strategy has been developed, that consists of a set of prescaled single-lepton triggers with same identification and isolation criteria as the signal triggers that be used by the same-sign analysis (and in most dilepton analysis). These triggers were commissioned with the first data at 13 TeV with 50ns bunch spacing.

9.1.1 General strategy

As a first approximation, the total number of events in the measurement region in a certain $p_T - \eta$ bin could be estimated by:

$$N_{p_T-\eta} = R \times \Delta t \times \epsilon_{sel} \times \epsilon_{p_T-\eta} \tag{9.1}$$

where, R is the total rate of the HLT paths used for filling the measurement region (0.3 Hz); Δt is the total interval of data-taking (2 months of data-taking assuming 8h per day); ϵ_{sel} is the (offline) selection efficiency, and $\epsilon_{p_T-\eta}$ is the efficiency of a certain lepton to fall under a certain $p_T - \eta$ bin, measured in a QCD MC sample.

A first estimation using L1 emulation quickly showed that a configuration with 4 paths with different p_T thresholds would be enough to cover the whole p_T spectrum and would allow us to have enough statistics to provide a fake-ratio measurement with a statistical uncertainty between 10-20% across the whole 2D map as it is shown on Figure 9.1.

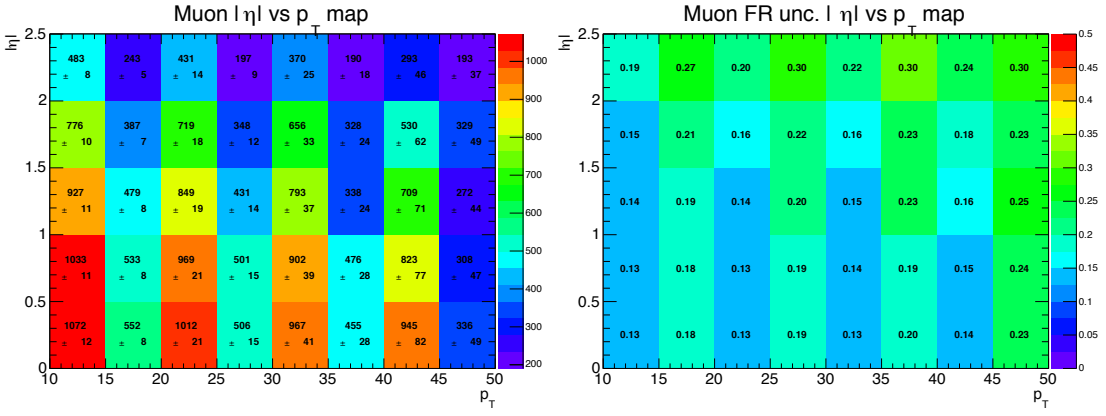


FIGURE 9.1: First estimation of the total number of expected events in the measurement region using a configuration with four paths with a total rate of 0.3 Hz after 2 months of data-taking, a 5% offline selection efficiency was assumed for this calculation. For estimating the uncertainty on the fake-ratio a constant value of 10% was considered.

The very same reconstruction, identification and isolation sequence that was implemented in for the double lepton signal triggers has been used. The paths were originally designed for the 40 pile-up and 25ns scenario ($L_{inst} = 1.4 \times 10^{34}\ cm^{-2}s^{-1}$), prescales will be adapted to keep constant rate for smaller luminosities.

Originally, the total allotted bandwidth for such auxiliary triggers is $4Hz$ (in the 40 pile-up and 25ns scenario), half of the bandwidth will be used by the muon paths and the other half by the electron paths. Some overlap between the different triggers is also expected but the prescales will avoid such problem in most of the cases.

9.1.2 Muon paths

Four muon HLT paths were designed to be identical to the 2015 signal double muon path (HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v1), except for requiring only one muon and the p_T cut that might be different for some paths. Four extra paths without isolation but with the same identification requirement are also included in the set of auxiliary paths. The reconstruction, identification and isolation is meant to be exactly the same as the one present in the signal triggers so these triggers could be used also for monitoring the per-leg efficiency.

Table 9.1 summarises the expected rate for each path, after applying the L1 prescale, and the desired HLT prescale to control the rate for a luminosity scenario of $L_{inst} = 1.4 \times 10^{34}\ cm^{-2}s^{-1}$.

TABLE 9.1: HLT rates for the muon paths, the rate is estimated using a mixture of QCD, Drell-Yan and W samples. HLT prescales are set to 1. $L_{inst} = 1.4{\times}10^{34}\ cm^{-2}s^{-1}$ and L1 prescales taken from the L1Menu_Collisions2015_25ns_v2.

Path	L1 seed	L1 prescale	Rate after L1 PS (Hz)
HLT_Mu8_TrkIsoVVL_v1	L1_SingleMu5	50000	0.20
HLT_Mu17_TrkIsoVVL_v1	L1_SingleMu12	3000	0.28
HLT_Mu24_TrkIsoVVL_v1	L1_SingleMu16	1000	0.21
HLT_Mu34_TrkIsoVVL_v1	L1_SingleMu20	700	0.10
HLT_Mu8_v1	L1_SingleMu5	50000	0.22
HLT_Mu17_v1	L1_SingleMu12	3000	0.31
HLT_Mu24_v1	L1_SingleMu16	1000	0.25
HLT_Mu34_v1	L1_SingleMu20	700	0.10

The composition of the selected events by these triggers have also been studied. A high purity on QCD events is very well desired for the application of the method. A 99% of the total bandwidth of the HLT_Mu8_TrkIsoVVL_v1 is coming from QCD events, as the p_T threshold increases the contribution from QCD gets reduced down to a 70% for the HLT_Mu34_TrkIsoVVL_v1. An increase on the electroweak contamination explain such reduction. An extra consideration must be taken into account: only 30% of the QCD rate for the lower p_T path will be kept when applying the selection of the measurement region, that contains a high-p_T jet. For the rest of the paths the loss when applying selection efficiency of the measurement region is below 30%.

To try to further increase the QCD purity of the events and also increase the offline selection efficiency, the addition of a jet to these paths was also studied. The rate reduction was significant in the case of the Mu8 (50%) and quite small (15%) for the Mu34 path, also the electroweak contamination was reduced by less than a 20% for all the paths with the exception of the Mu8, where no gain was observed. The addition of such jet requirement reduced the flexibility in the offline jet selection for systematic

studies so it was decided to discard such addition, as it is not needed to control the rate / contamination.

The startup scenario would be much more friendly than the one used here for the calculations, this will imply a reduction on the L1 prescales and therefore the total number of events selected by these triggers would be sufficient to perform the measurement for the first round of publications.

9.1.3 Electron paths

The same strategy was followed for the electron side, four HLT paths with the very same reconstruction, identification and isolation sequence as the one present in the double electron and muon-electron signal triggers, except for requiring only one lepton with a different p_T in some cases and the additional requirement of an away jet ($|\Delta R(lep, jet)| > 1$) with p_T greater than 30 GeV. Five extra paths were also considered with same identification criteria as the di-electron+H_T cross triggers.

Table 9.2 summarises the rate for each path, before and after applying the L1 prescale, and the desired HLT prescale to control the rate.

TABLE 9.2: HLT rates for the electron paths, the rate is estimated using a mixture of QCD, Drell-Yan and W samples. HLT prescales are set to 1. $L_{inst} = 1.4\times10^{34}\ cm^{-2}s^{-1}$ and L1 prescales taken from the L1Menu_Collisions2015_25ns_v1.

Path	L1 seed	L1 prescale	Rate after L1 PS (Hz)
HLT_Ele12_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	L1_SingleEG10	20000	0.11
HLT_Ele18_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	L1_SingleEG10	20000	0.10
HLT_Ele23_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	L1_SingleEG20	4000	0.26
HLT_Ele33_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	L1_SingleEG20	4000	0.22
HLT_Ele12_CaloIdM_TrkIdM_PFJet30_v1	L1_SingleEG5||L1_SingleEG10	20000	0.09
HLT_Ele12_CaloIdM_TrkIdM_PFJet30_v1	L1_SingleEG10	20000	0.07
HLT_Ele18_CaloIdM_TrkIdM_PFJet30_v1	L1_SingleEG10	20000	0.06
HLT_Ele23_CaloIdM_TrkIdM_PFJet30_v1	L1_SingleEG20	4000	0.15
HLT_Ele33_CaloIdM_TrkIdM_PFJet30_v1	L1_SingleEG20	4000	0.11

For this case, 99% of the total rate is coming from QCD events, that means that the QCD purity of the events selected by these trigger would be very high as it is desired for performing the fake-ratio measurement. Furthermore, low p_T QCD events only contribute to 20% of the rate in the worse-case scenario which means that the loss of events when applying the measurement selection will be smaller than in the case of the muons.

Other configurations were also tested and discarded. An HLT path with a p_T threshold of 8 GeV as it is used in the muon case was tested but the L1 prescale of the relevant EG path was so high the expected rate was ridiculously small making the path totally useless for our purpose. Not requiring a jet in the paths only increased slightly the rate but with (almost) no change in the QCD/electroweak composition, therefore, we

decided to keep the jet requirement as it helps increasing the selection efficiency without changing the composition of the selected sample.

As in the case of the muons, the startup scenario would be much more friendly than the one used here for the calculations, this will imply a reduction on the L1 prescales and therefore the total number of events selected by these triggers would be sufficient to perform the measurement for the first round of publications.

9.1.4 Commissioning with 50ns data

A set of HLT auxiliary paths have been developed for fake-ratio measurements. As the first data at 13 TeV arrived, rate and purity measurements have been carried out to check the performance of these triggers in real data in preparation for the bigger run at 25ns bunch spacing.

The LHC restart at 13 TeV, started with a short run with 50ns bunch spacing that delivered about 100 pb^{-1} , CMS collected about 80 pb^{-1} , as it can be seen in Figure 9.2. Only 40 pb^{-1} can be used for physics.

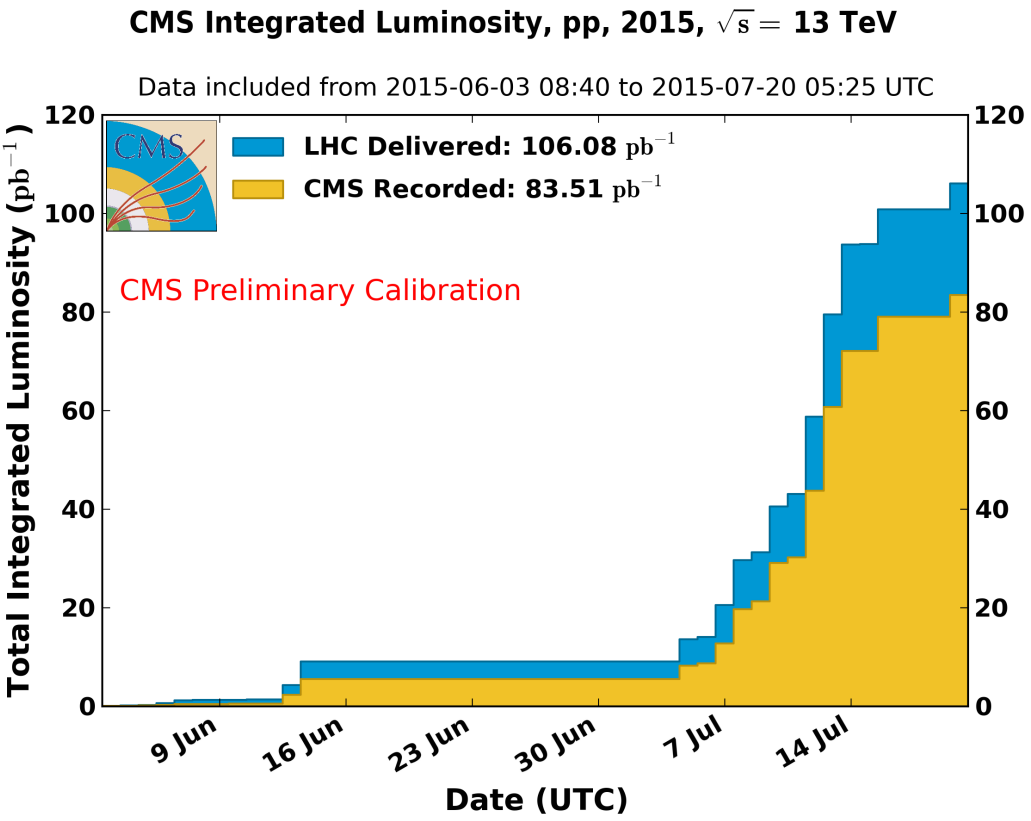


FIGURE 9.2: LHC delivered and CMS collected luminosity for the 50ns bunch-spacing run. Due to several problems with the CMS restart, only 40 pb^{-1} could be used for physics.

Rates

During this run, the L1/HLT prescales have been varying a lot run-by-run complicating even more the rate measurement, after some adjustments, the rates measured in MC and data are shown in Table 9.3. Although the MC rates are within the expected and design values, the result in data for certain paths is smaller than MC expectation, making some of the triggers useless with the small amount of data that was collected, in particular the high-p_T paths for muons almost did not collect any data as well as the low-p_T electrons paths, due to heavy prescales.

These discrepancies were corrected for the 25ns run and the total rate was increased for the lower luminosity scenarios for the sake of collecting enough data to compute the fake ratios despite a slow ramp-up of the machine, also when running at design luminosity, one can take profit of the last part of the run when luminosity decreases and the lower cross-sections also allow for an increase in rate.

TABLE 9.3: HLT rates for the auxiliary paths used for fake-ratio measurements. MC rate is estimated using a mixture of QCD, DY and W samples and are calculated for a scenario of $L_{inst} = 5 \times 10^{33}\ cm^{-2}s^{-1}$ (nominal luminosity for the 50ns Run). These rates correspond to the frozen 50ns menu, the prescales were adjusted after the very first measurements in data.

Path	Rate in MC (Hz)	Rate in data (Hz)
HLT_Ele12_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	0.18 ± 0.07	0.44 ± 0.31
HLT_Ele18_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	0.15 ± 0.07	0.22 ± 0.22
HLT_Ele23_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	0.16 ± 0.07	0.22 ± 0.22
HLT_Ele33_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	0.07 ± 0.04	0
HLT_Ele8_CaloIdM_TrkIdM_PFJet30_v1	0.11 ± 0.06	0.22 ± 0.22
HLT_Ele12_CaloIdM_TrkIdM_PFJet30_v1	0.11 ± 0.06	0
HLT_Ele18_CaloIdM_TrkIdM_PFJet30_v1	0.09 ± 0.05	0
HLT_Ele23_CaloIdM_TrkIdM_PFJet30_v1	0.37 ± 0.10	0.44 ± 0.31
HLT_Ele33_CaloIdM_TrkIdM_PFJet30_v1	0.19 ± 0.07	0
HLT_Mu8_TrkIsoVVL_v1	1.34 ± 0.11	1.77 ± 0.63
HLT_Mu17_TrkIsoVVL_v1	0.47 ± 0.06	0.44 ± 0.31
HLT_Mu24_TrkIsoVVL_v1	0.09 ± 0.03	0.22 ± 0.22
HLT_Mu34_TrkIsoVVL_v1	0.07 ± 0.02	0
HLT_Mu8_v1	1.35 ± 0.11	1.77 ± 0.63
HLT_Mu17_v1	0.81 ± 0.08	1.33 ± 0.54
HLT_Mu24_v1	0.09 ± 0.03	0
HLT_Mu34_v1	0.13 ± 0.03	0

Purities

Another important check is the purity of these paths, hence, how many events, of the total selected by the trigger fulfil the offline requirements that are applied for the fake-ratio measurement.

The offline selection, requires the presence of a lepton, passing the full identification criteria of the analysis, one fully identified away jet with p_T greater than 40 GeV. Furthermore, an extra requirement on $E_T^{miss} < 20$ GeV and $m_T < 20$ GeV to suppress electroweak contamination is applied.

Assuming this selection, the global purity of the muon paths goes from 1.7% up to 4%, while the purity of the electron paths ranges from 5% to 8%. The difference relies exclusively in the jet requirement that is present on the electron paths but not in the muon ones.

In the muon paths, requiring the offline jet kills between 94% and 73% of the events, the E_T^{miss} requirement kills between 60-70% of the events, while the m_T kills between 30-60% of the events. This level of rejection is similar as the one seen in 2012 data. For the electron paths, the situation is a bit better, the jet requirement only kills between 45-20% of the events. E_T^{miss} and m_T requirement has the same level of rejection as for muon paths.

Final menu

The changing conditions of the first months of running at the LHC and the change in the total expected integrated luminosity at the end of the year forced to introduce some changes to the presented menu here. The need to accomodate such paths to the request of other physics group have also motivated this change.

As the total expected integrated luminosity at the end of the year will be much smaller than the original expectations as well as the peak instantaneous luminosity that will be reached is factor three smaller, an increase in the rate is foreseen for being able to collect enough statistics at the end of the year to compute the fake ratio estimation with a 10% statistical uncertainty.

The final trigger menu that is being used for 2015 data taking at 25 ns, together with an estimation of the total rate per path is detailed in Table 9.4.

9.2 Lepton isolation and identification

Several changes for 13 TeV have been developed concerning the lepton identification and isolation. The new algorithms and techniques provide similar or better performance despite the worsening of the expected pile-up conditions that were expected for the Run II.

TABLE 9.4: HLT rates for the auxiliary paths used for fake-ratio measurements. These rates correspond to the frozen 25ns menu, for .

Path	Rate in data (Hz)
HLT_Ele12_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	8.11
HLT_Ele23_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	0.53
HLT_Ele33_CaloIdL_TrackIdL_IsoVL_PFJet30_v1	0.90
HLT_Ele8_CaloIdM_TrkIdM_PFJet30_v1	0.86
HLT_Ele12_CaloIdM_TrkIdM_PFJet30_v1	10.20
HLT_Ele23_CaloIdM_TrkIdM_PFJet30_v1	1.24
HLT_Ele33_CaloIdM_TrkIdM_PFJet30_v1	1.42
HLT_Mu8_TrkIsoVVL_v1	8.28
HLT_Mu17_TrkIsoVVL_v1	25.50
HLT_Mu8_v1	5.02
HLT_Mu17_v1	39.40

In this section, a short summary is presented, the preliminary performance studies as well as the commissioning of the variables used for the identification and isolation with the first data.

9.2.1 Lepton identification

Muons For muons, a new identification algorithm has been developed starting from the POG Loose Muon (i.e. a muon that is identified by the particle flow algorithm that has to be either a global or an arbitrated tracker muon) using additional track-quality and muon-quality requirements. Table 9.5.

TABLE 9.5: Muon medium working identification

Pass POG Muon Loose ID	
Valid fraction of tracker hits	> 0.8
Pass Global Muon ID	
χ^2/N_{dof} (globalFit)	< 3
track Kink finder	< 20
Segment compatibility	> 0.303
OR	
Segment compatibility	> 0.415

This identification provides slightly higher efficiency than 2012 tight working point with same fake rejection. This can be seen in Figure 9.3 The variables used to form this ID will be commissioned with the first 50ns data.

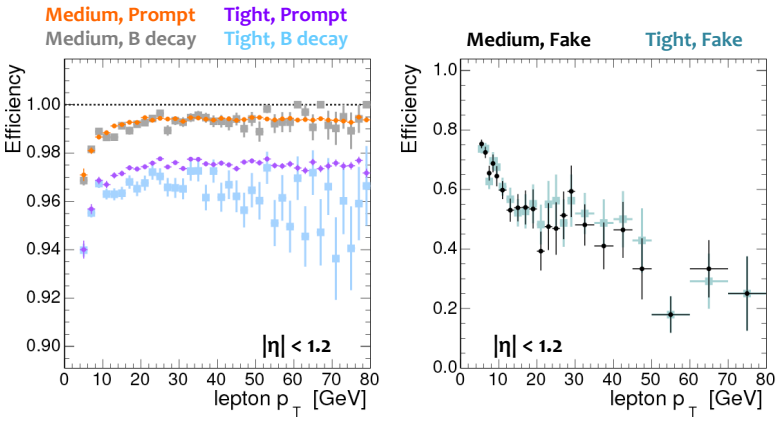


FIGURE 9.3: Medium ID performance in MC, showing performance for prompt muons, muons coming from b-decays as well as fake muons. Medium ID shows better performance in selecting prompt leptons while rejecting same amount of fake muons.

Electrons The electrons are reconstructed by associating ECAL-clusters and GSF tracks. Only electrons within the tracker and ECAL acceptance are considered: $|\eta| < 1.4442$ or $1.566 < |\eta| < 2.4$. A new electron identification using a multivariate discriminant has been developed built with shower-shape variables ($\sigma_{i\eta i\eta}$, $\sigma_{i\phi i\phi}$, the cluster circularity, widths along η and ϕ, R_9, H/E, E_{inES}/E_{raw}), track-cluster matching variables (E_{tot}/p_{in}, E_{Ele}/p_{out}, $\Delta\eta_{in}$, $\Delta\eta_{out}$, $\Delta\phi_{in}$, $1/E - 1/p$) and track quality variables (χ^2 of the KF and GSF tracks, the number of hits used by the FK/GSF filters, fbrem). Two different identification working points, summarized in the Table 9.6, are used in this analysis. Those working points depend of the pseudorapidity of the electrons.

TABLE 9.6: Electron identification working points used in this analysis.

pseudorapidity region	loose WP	tight WP
$0 < \|\eta\| < 0.8$	-0.11	0.73
$0.8 < \|\eta\| < 1.479$	-0.35	0.57
$1.479 < \|\eta\| < 2.5$	-0.55	0.05

9.2.2 Lepton isolation

Although the relative isolation that was used at 8 TeV was corrected to mitigate pile-up effects, under certain conditions it might still be under performant. If the event is very busy, and many tracks and jets populate the event, it would be very difficult to find an isolated lepton. New techniques have been developed to recover some of this events.

With this new approach, the lepton isolation is constructed using three different variables:

Mini isolation The first one is called I_{mini}. It uses a variable cone size that allows to isolate a lepton even in boosted topologies or very busy environments, it is defined as follows:

$$R = \begin{cases} 0.2, & p_T \leq 50GeV \\ \frac{10}{p_T}, & p_T \in (50, 100)\ GeV \\ 0.02 & p_T > 100\ GeV \end{cases}$$

Requiring I_{mini} below a given threshold ensures that the lepton is locally isolated, even in boosted topologies. Despite this variable cone size it is defined as a relative isolation your favourite pile-up correction, in this case the effective area corrections that were described in Section 4.3.2.1, that has been recalculated for 13 TeV (see Table 9.7):

$$I_{\text{mini}} = \frac{\sum_R p_T(h^{\pm}) - \max(0, \sum_R p_T(h^0) + p_T(\gamma) - \rho\mathcal{A}\left(\frac{R}{0.3}\right)^2}{p_T(\ell)}. \tag{9.2}$$

$\lvert\eta\rvert$ range	$\mathcal{A}$(e)	$\mathcal{A}(\mu)$
$0.0 - 0.8$	0.1013	0.0913
$0.8 - 1.3$	0.0988	0.0765
$1.3 - 2.0$	0.0572	0.0546
$2.0 - 2.2$	0.0842	0.0728
$2.2 - 2.5$	0.1530	0.1177

TABLE 9.7: Effective area recomputed with preliminary MC at 13 TeV, for muons and electrons

Ratio of the lepton / jet The second variable is simply the ratio between the p_T of the lepton and the p_T of the jet matched to the very sample lepton:

$$p_T{}^{\text{ratio}} = \frac{p_T(\ell)}{p_T(\text{jet})} \tag{9.3}$$

This jet is matched geometrically to the lepton, and in most of the case is the jet containing the lepton. If no jet is clustering the lepton, then the closest one is chosen. The use of $p_T{}^{\text{ratio}}$ is a simple way to identify leptons in quite boosted topologies, without any jet re-clustering.

ptrel The last variable used is the $p_T{}^{\text{rel}}$ variable, defined as the projection of the lepton p_T towards the closest jet:

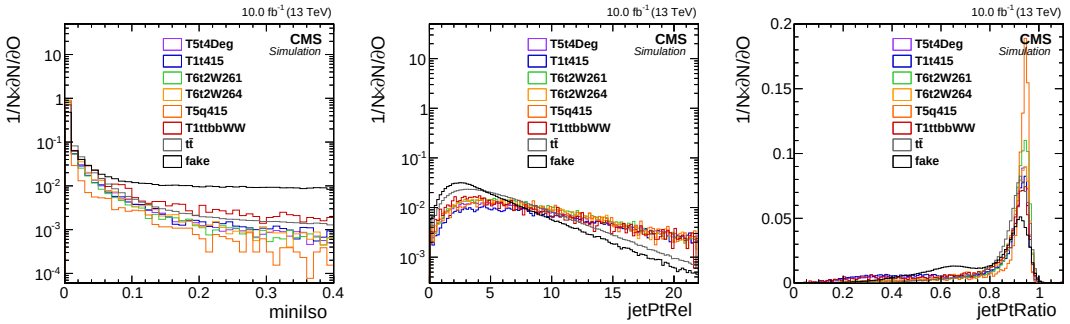


FIGURE 9.4: Isolation variables for leptons to be used in the analysis: I_{mini} (left), ${p_T}^{\text{rel}}$ (middle) and ${p_T}^{\text{ratio}}$ (right). The $t\bar{t}$ process corresponds to leptons originating from W decays. The fake process corresponds to leptons found in $t\bar{t}$ events, not produced by W bosons. Other distributions are for signal-like leptons

$$ {p_T}^{\text{rel}} = \frac{(\vec{p}(\text{jet}) - \vec{p}(\ell)) \cdot \vec{p}(\ell)}{||\vec{p}(\text{jet}) - \vec{p}(\ell)||}. \tag{9.4} $$

This variable allows to recover leptons from accidental overlaps with jets.

The distributions of those three variables, for leptons produced by electroweak boson decays and jets misidentified as leptons in $t\bar{t}$ events, are illustrated on Fig. 9.4.

Multi-isolation Using those three variables, a lepton is considered isolated if the following condition is respected:

$$ I_{\text{mini}} < I_1 \wedge ({p_T}^{\text{ratio}} > I_2 \vee {p_T}^{\text{rel}} > I_3) \tag{9.5} $$

The $I_i, i = 1, 2, 3$ values depends of the flavour of the lepton: as the probability to misidentify a jet is higher for the electrons, tighter isolation values are used. The loose lepton isolation is significantly relaxed, as well as an extra definition (fakable) used for the fake lepton background estimation. The different values are summarised in the Table 9.8. The logic beyond that isolation is a relaxing of the local isolation, compensated by the requirement that the lepton carries the major part of the energy of the corresponding jet, or if not, if the lepton is considered as accidentally overlapping with a jet. The performance of the different working points of the multi-isolation variable for signal-like leptons is shown in Figure 9.5.

9.2.3 Commissioning with 50ns data

To validate the properties of the prompt-lepton isolation and identification variables, two control regions in data with high purity of $Z \to \mu^- \mu^+$ and $Z \to e^- e^+$ are used. Figures

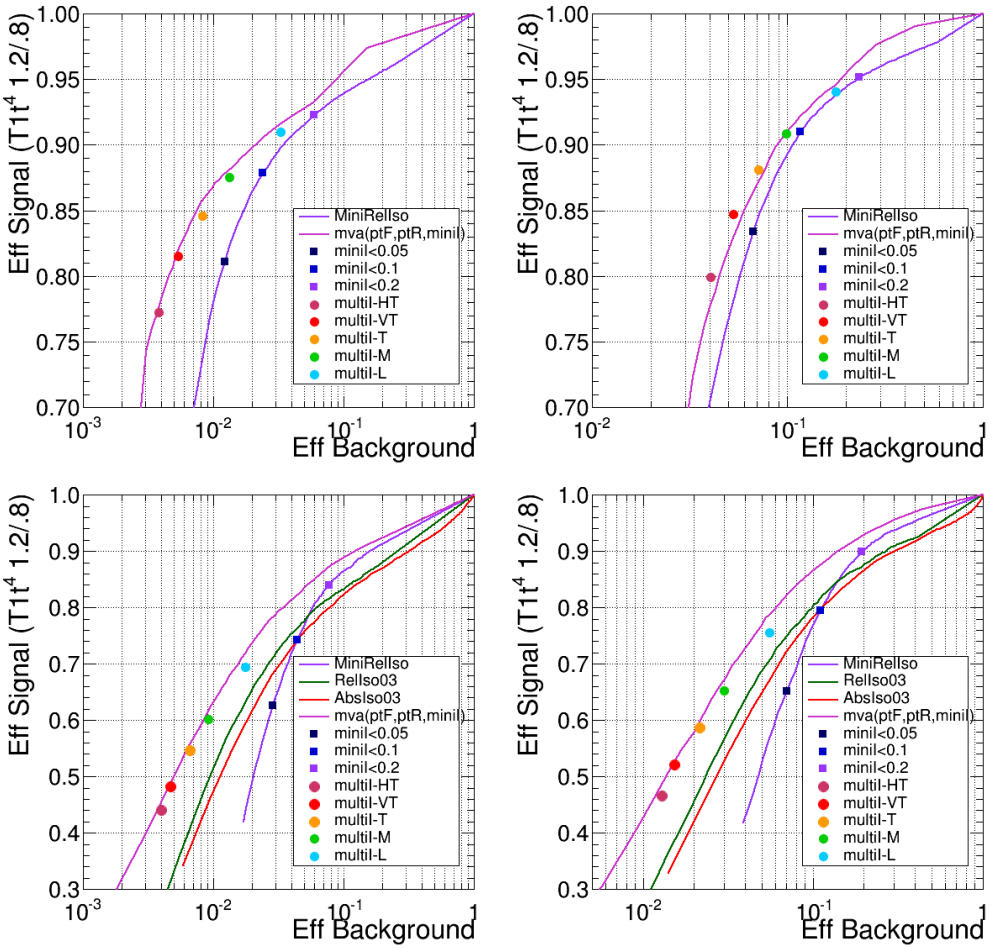
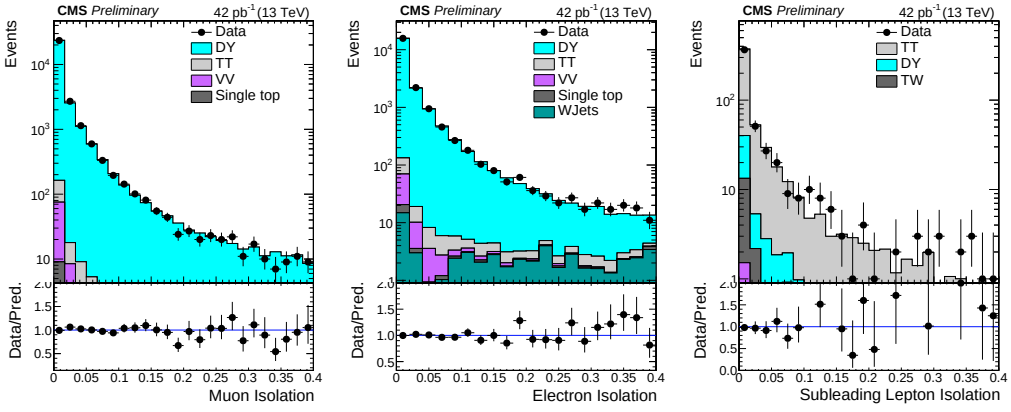


FIGURE 9.5: Prompt lepton selection efficiency versus misidentified lepton efficiency for electrons (left) and muons (right) with $p_T > 25$ GeV (top) and $10 < p_T < 25$ GeV (bottom) in T1tttt (1.2/0.8) events. The tight working point corresponds to the orange full circle, and the loose working point to the green full circle. The standard isolation is indicated by the black point and the pure mini isolation by the violet line.



(A) (B) (C)

FIGURE 9.6: Data to simulation comparisons for lepton identification performance studies. From left to right the probe lepton I_{mini} for $Z \to \mu^-\mu^+$ (9.6a), $Z \to e^-e^+$ (9.6b) and $t\bar{t}$ (9.6c) control region

TABLE 9.8: Isolation working points

isolation value	loose WP	fakable e (μ) WP	μ (Medium) WP	e (Tight) WP
I_1	0.4	0.4	0.14	0.10
I_2	0	0.68 (0.70)	0.68	0.70
I_3	0	6.7 (7)	6.7	7

9.6a and 9.6b show distributions for electron and muon I_{mini} in such control regions. A third control region is defined to check the properties of prompt leptons arising from $t\bar{t}$ by requiring two leptons, with an invariant outside the Z mass, two jets and at lest one b-tagged jet. 9.6c shows distributions for electron and muon I_{mini} in such control region.

To assess the performance of such variables for background leptons samples enriched in leptons from b-hadron decay are selected using two Z+ℓ and W+ℓ control regions. This is show in Figure 9.7.

Variables including jet energies, such as ${p_T}^{\text{ratio}}$ and ${p_T}^{\text{rel}}$, could not be fully commissioned during the 50ns run as those variables are extremely sensitive to the jet energy corrections and those were not fully understood by the time in which these plots were made.

9.3 Signal extraction strategy

A new signal strategy has been designed towards repeating the analysis described in Chapter 6 at 13 TeV. The search is intended to be as wide as possible to be prepare for any mass scheme. Searches performed at 8 TeV did not cover totally some scenarios, such as models with very small mass differences between s-particles, that provide very small phase space for SM particles in the decay cascade. These scenarios provide the appearance of low boosted objects in the final state.

In order to provide maximal sensitivity, three exclusive lepton selection are defined:

- high-high (HH) selection: two nominal leptons with $\text{p}_\text{T} > 25$ GeV,
- high-low (HL) selection: one nominal lepton with $10 < p_T < 25$ GeV and the other with $\text{p}_\text{T} > 25$ GeV,
- low-low (LL) selection: two nominal leptons with $10 < p_T < 25$ GeV.

Figure 9.8 compares the approach used in Chapter 6 with the one proposed here. The HH selection is designed for the search of signal with boosted leptons, while the HL

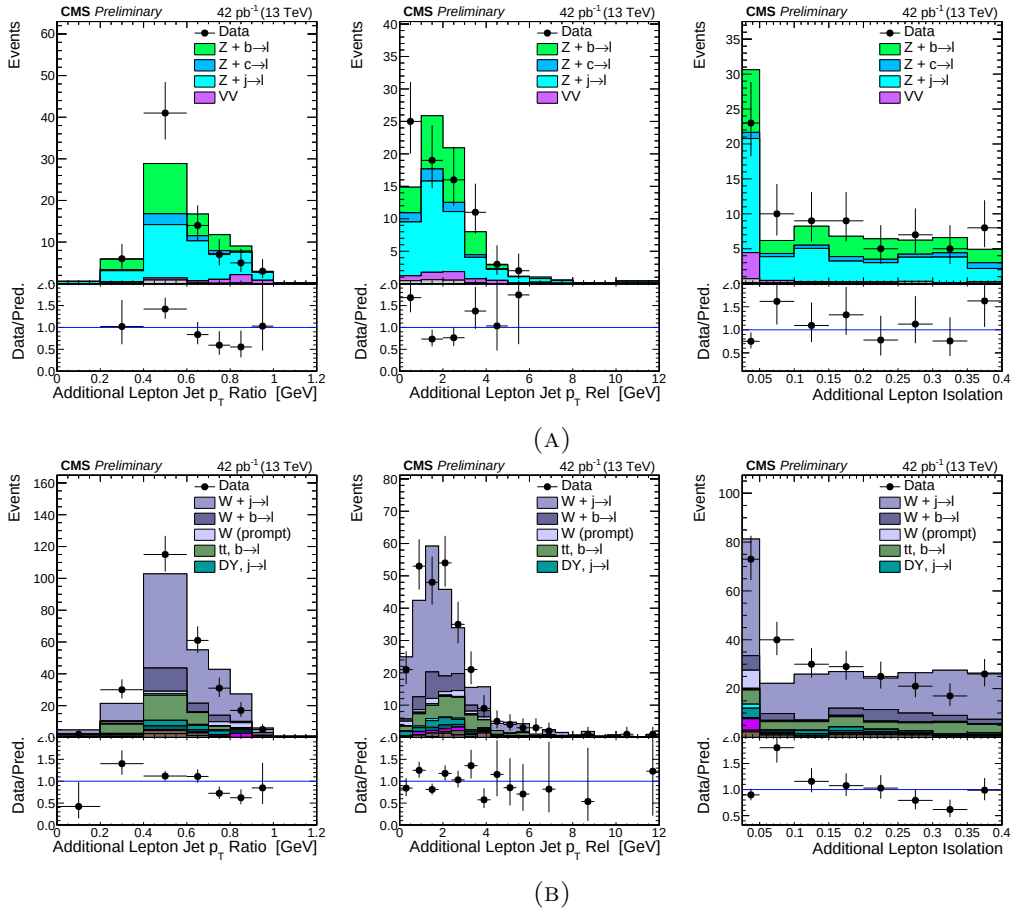
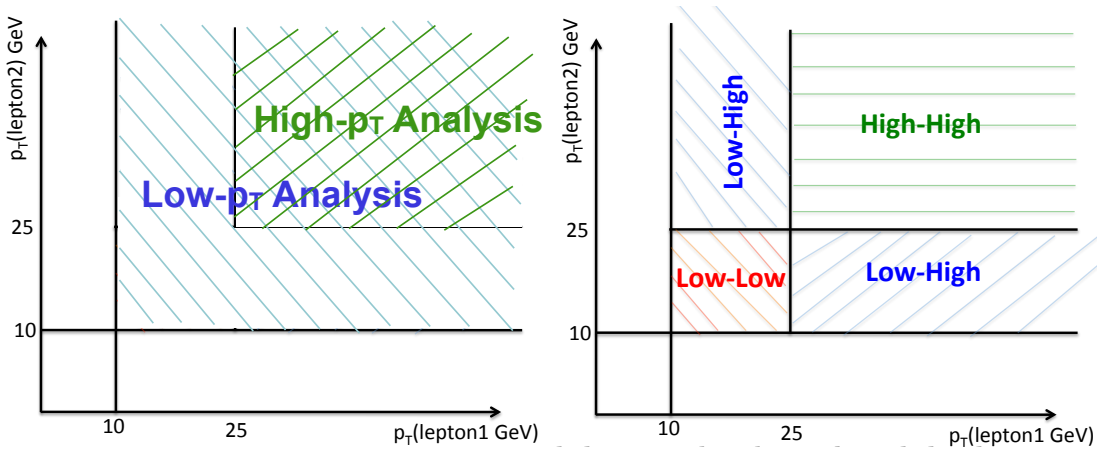


FIGURE 9.7: Data to simulation comparisons for lepton identification performance studies. From left to right the probe lepton ${p_T}^{\text{ratio}}$, ${p_T}^{\text{rel}}$ and I_{mini}. Z+ℓ control region (Fig. 9.7a) is displayed on top and W+ℓ (Fig. 9.7b) on the bottom row.

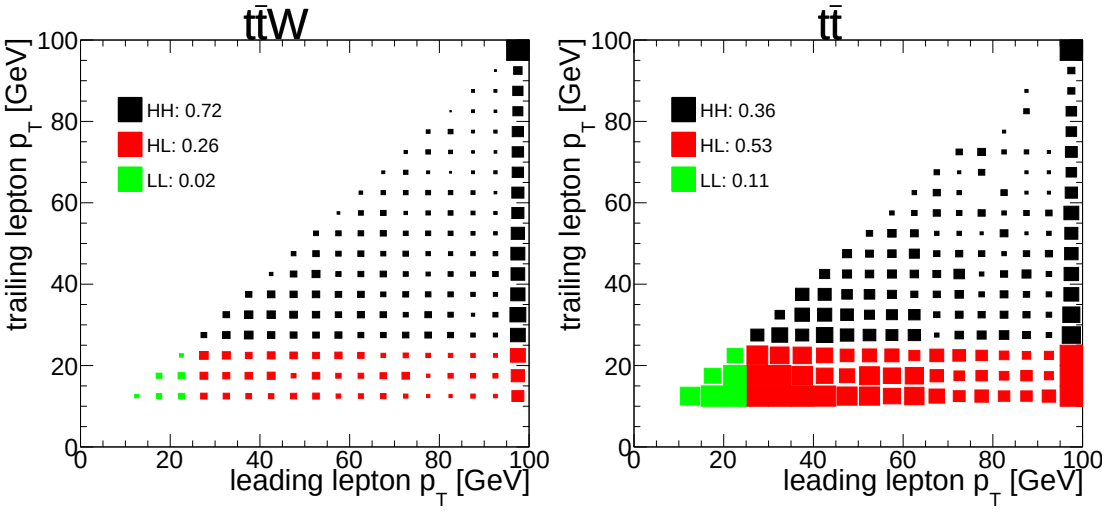
and LL are more sensitive to signal producing low energetic leptons. This division is not only lead by potential signal, but also by the different background composition on these regions. High lepton p_T threshold suppresses the contribution fron non-prompt leptons and therefore HH region is mainly populated by irreducible SM background (see Figure 9.9a). The non-prompt lepton background is largely collected in the HL region (see Figure 9.9b). Additional handles to control the reducible background are applied like M_T, or more strict ID requirements. The LL region benefits from both main contributions being suppressed and therefore could profit from the reduced further kinematical binning in it.

Independently of the choice of lepton selection, three "baseline regions" are defined to reject a significant fraction of the background while allowing backgrounds to be measured or controlled. Baseline regions split events into categories as a function of the number of b jets reconstructed in the event. On top of the baseline selection, the "signal selection" is applied to categorize the events by topology. Categories are defined to discriminate the background from the different signal signatures.



(A) 8 TeV binning (B) 13 TeV binning

FIGURE 9.8: Search regions design for low- and high-pt analyses which are used at 9.8a 8 TeV and 9.8a 13 TeV analysis.



(A) $t\bar{t}$ W (B) $t\bar{t}$

FIGURE 9.9: The distribution of the SM $t\bar{t}$ and $t\bar{t}$ W processes in leading-trailing lepton p_T plane. Numbers indicate fraction of acceptance for each process in the LL, HL, HH regions.

9.3.1 Baseline selection

The baseline region is defined to reject a significant fraction of the reducible background, while allowing to extract a data-driven estimation of the SM background.

Events are selected if a pair of same-sign leptons with p_T thresholds defined following the lepton selection considered is present. Events showing a pair of leptons with an invariant mass smaller than 8 GeV are rejected. Drell-Yan production is controlled by rejecting events in case a loosely identified muon (electron) with $p_T > 5(7)$ GeV is present. Also if

TABLE 9.9: Baseline region definitions

region	$N_{\text{b jets}}$
BR0	$= 0$
BR1	$= 1$
BR2	$= 2$
BR3	≥ 3

one of the nominal leptons forms a same flavour, opposite sign pair with an extra loose lepton with an invariant mass within the Z mass value or below 12 GeV.

A loose $\mathrm{H_T}$ requirement ($H_T > 80$ GeV), a minimal jet multiplicity criteria ($\mathrm{N_{jets}} > 2$) and some requirement on $\mathrm{E_T^{miss}}$ ($\mathrm{E_T^{miss}} > 30$ GeV only if $H_T < 500$ GeV) is are also applied. Finally a categorisation as a function of number of b-jets is applied following the criteria in Table 9.9.

9.3.2 Signal selection

A similar approach to the one used in Chapter 6 is used: search regions are defined using $\mathrm{H_T}$, $\mathrm{E_T^{miss}}$, $\mathrm{N_{jets}}$ and $N_{\text{b-jets}}$. To further discriminate between SM background and new physics, a new variable $M_{\mathrm{T}}^{\mathrm{min}}$ is introduced:

$$M_{\mathrm{T}}^{\mathrm{min}} = min(m_{\mathrm{T}}(\ell_1, \not{E}_{\mathrm{T}}), m_{\mathrm{T}}(\ell_2, \not{E}_{\mathrm{T}})).$$

In case of a same-sign lepton pair in the $\mathrm{t\bar{t}}$ process, this variable has a cut-off at the W mass. Therefore, forming search regions with an upper boundary on $M_{\mathrm{T}}^{\mathrm{min}}$ of 120 GeV allows to contain fake lepton background in dedicated search regions where further kinematic binning allows to discriminate against the $\mathrm{t\bar{t}}$ process. Subsequently, search regions with $M_{\mathrm{T}}^{\mathrm{min}} > 120$ GeV profit from reduced $\mathrm{t\bar{t}}$ background.

The final signal regions SRn(A/B/C) can be subdivided into three groups corresponding to the HH (SRnA), HL (SRnB) and LL (SRnC) selections. The summary of the selection is described in Tables 9.10, 9.11 and 9.12. All SRn(A/B/C) regions are exclusive and can be statistically combined. The regions are configured in such a way that some of the regions are dominated by background event while in other regions are designed to contain 1 background events. These regions were designed for an expected recorded luminosity of 10 $\mathrm{fb^{-1}}$, and thus with the expected conditions for 2015 some of these regions shall be redefined.

TABLE 9.10: Signal region definitions for the HH lepton selection.

$N_{\text{b jets}}$	$M_{\text{T}}^{\text{min}}$	$\not{E}_{\text{T}}$	N_{jets}	$H_T < 300$	$300 < H_T < 1600$	$H_T > 1600$
0 b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets	SR1	SR2	SR32
			5+ jets	SR3	SR4	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets		SR5	
			5+ jets		SR6	
	$M_{\text{T}}^{\text{min}} > 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets		SR7	
			5+ jets		SR8	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets			
			5+ jets			
1 b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets	SR9	SR10	
			5+ jets	SR11	SR12	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets		SR13	
			5+ jets		SR14	
	$M_{\text{T}}^{\text{min}} > 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets		SR15	
			5+ jets		SR16	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets			
			5+ jets			
2 b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets	SR17	SR18	
			5+ jets	SR19	SR20	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets		SR21	
			5+ jets		SR22	
	$M_{\text{T}}^{\text{min}} > 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets		SR23	
			5+ jets		SR24	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets			
			5+ jets			
3+ b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2+ jets	SR25	SR26	
		$200 < \not{E}_{\text{T}} < 500$	2+ jets	SR27	SR28	
	$M_{\text{T}}^{\text{min}} > 120$	$50 < \not{E}_{\text{T}} < 200$	2+ jets	SR29	SR30	
		$200 < \not{E}_{\text{T}} < 500$	2+ jets			
inclusive	inclusive	$\not{E}_{\text{T}} > 500$	2+ jets	–	SR31	

TABLE 9.11: Signal region definitions for the HL lepton selection.

$N_{\text{b jets}}$	$M_{\text{T}}^{\text{min}}$	$\not{E}_{\text{T}}$	N_{jets}	$H_T < 300$	$300 < H_T < 1600$	$H_T > 1600$
0 b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets	SR1	SR2	SR26
			5+ jets	SR3	SR4	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets		SR5	
			5+ jets		SR6	
1 b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets	SR7	SR8	
			5+ jets	SR9	SR10	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets		SR11	
			5+ jets		SR12	
2 b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2-4 jets	SR13	SR14	
			5+ jets	SR15	SR16	
		$200 < \not{E}_{\text{T}} < 500$	2-4 jets		SR17	
			5+ jets		SR18	
3+ b jets	$M_{\text{T}}^{\text{min}} < 120$	$50 < \not{E}_{\text{T}} < 200$	2+ jets	SR19	SR20	
		$200 < \not{E}_{\text{T}} < 500$	2+ jets	SR21	SR22	
inclusive	$M_{\text{T}}^{\text{min}} > 120$	$50 < \not{E}_{\text{T}} < 500$	2+ jets	SR23	SR24	
inclusive	inclusive	$\not{E}_{\text{T}} > 500$	2+ jets	–	SR25	

TABLE 9.12: Signal region definitions for the LL lepton selection. Additional baseline selection of $H_T > 300$ GeV is applied in all search regions.

$N_{\text{b jets}}$	$M_{\text{T}}^{\text{min}}$	$50 < \not{E}_{\text{T}} < 200$	MET > 200
0 b jets	$M_{\text{T}}^{\text{min}} < 120$	SR1	SR2
1 b-tag	$M_{\text{T}}^{\text{min}} < 120$	SR3	SR4
2 b jets	$M_{\text{T}}^{\text{min}} < 120$	SR5	SR6
3+ b jets	$M_{\text{T}}^{\text{min}} < 120$	SR7	
inclusive	$M_{\text{T}}^{\text{min}} > 120$	SR8	

9.4 Background estimation

Backgrounds for the same-sign dilepton final state can be divided in three categories:

- **Fake leptons:** "Non-Prompt" or "Fake" leptons are leptons from heavy-flavour decays, misidentified hadrons, muons from light-meson decays in flight, or electrons from unidentified photon conversions. Depending on the signal regions, this background is dominated by $t\bar{t}$ and W+jets processes.
- **Rare SM processes:** Rare SM processes yield same-sign leptons, mostly from ttW, ttZ, and diboson production. We also include the contribution from the SM Higgs boson produced in association with a vector boson or a pair of top quarks in this category of background. With the exception of WZ, rares are estimated from simulation.
- **Charge flips:** Charge misidentification, i.e. events with opposite-sign isolated leptons where the charge of one of the leptons is misidentified because of severe bremsstrahlung in the tracker material. This background, is relevant only for electrons and is negligible for muons.

9.4.1 Fake leptons

Background from fake leptons is estimated with the "fake rate" method. The number of events in the sample with at least one lepton that passes a loose selection but fails the full set of tight identification and isolation requirements (application region) is weighted using the "tight-to-loose" ratio, i.e. the probability that a loosely identified non-prompt lepton also passes the full set of requirements. This probability is measured as a function of lepton p_T and η, as well as event kinematics, in a control sample of QCD multijet events that are enriched in non-prompt leptons (measurement region). Such region is triggered by the auxiliary triggers defined in Section 9.1 and requires only one denominator lepton in the event, one recoling jet with $\Delta R(jet, lep) > 1.0$ and low

MET and MT to suppress the contribution from W and Z. Main systematic effect is the non-universality of the "tight-to-loose" ratio, particularly due to the dependency from the mother parton p_T and the flavor composition of the sample. In the 8 TeV analysis the corresponding uncertainty was estimated to be 50%.

The new ideas developed in the context of the SUSY Fake Lepton working group [112], have been deployed in the present analysis. The main goal of these developments is to reduce the dependency from the effects described above. The mother parton p_T dependence is accounted for by parametrizing the fake rate vs the lepton p_T corrected by the energy in the isolation cone according to the formula:

$$\begin{aligned} if\ p_T^{rel} > I_3: \quad p_T \rightarrow \quad & p_T \cdot (1 + max(0, I_m - I_1)) \\ else: \quad p_T \rightarrow \quad & max(p_T, p_T(jet) \cdot I_2) \end{aligned} \tag{9.6}$$

where variables are defined as in Section 9.2.1. This definition leaves unchanged the p_T of leptons passing the isolation cut and modifies the p_T of those failing the cut so that it is better proxy of mother parton p_T and results in a flatter fake rate as a function of the mother parton p_T. The cone correction significantly improves the closure of the method (Fig. 9.10).

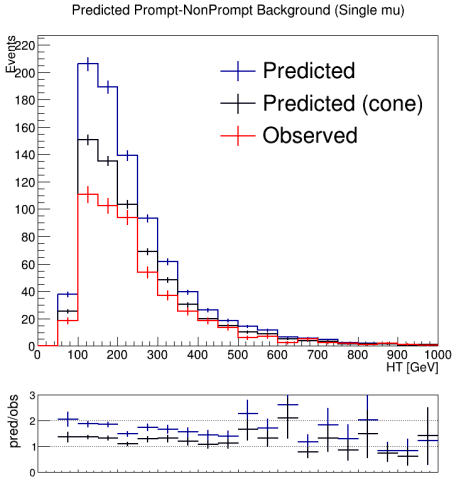


FIGURE 9.10: Muon fake rate closure for QCD measurement with 10/fb as a function of H_T with and without cone correction.

The flavour dependence for electrons is reduced by extrapolating both on isolation and on the MVA ID, variables mostly sensitive to heavy and light flavour decays respectively. The procedure starts from a fakable electron definition with the same cut on the MVA ID as the tight electrons; this cut is relaxed on the fakable lepton definition until the value of the fake rate from light flavour matches the one from heavy flavour. On data

the stability of the fake rate can be verified varying the flavour composition of the measurement sample by making use of b-tagging as a handle to enrich or deplete sample in heavy flavour fakes. The measured closure of the method for the different baseline region is roughly at the 20% level.

An alternative method for the fake rate measurement has been developed following an approach similar to what was used in an ATLAS publication [113]. The main idea is measure the fake rate "in situ", i.e. in an inclusive dilepton baseline region with inverted transverse impact parameter cut ($SIP_{3D}>4$) on one dilepton leg. As the isolation variable used here and the transverse impact parameter are not (or very weakly) correlated, the measured fake rate can be applied - as in the standar method- to the analysis signal regions with $SIP_{3D}<4$. The fakable lepton definition for the "in situ" measurement also employs the cone correction described above and is similar to the one used for the QCD measurement, except that the tight MVA ID cut is used for electrons and, both for electron and muons, the cut on $I_{\text{mini}}<0.4$ is replaced by a cut on multi-iso equation 9.5. The main limitation of this method is the larger statistical uncertainty, especially for luminosities smaller than 5 fb^{-1} (typical uncertainties of ∼40% at 3 fb^{-1}); with 10 fb^{-1} the statistical uncertainties on the fake rate become comparable to the QCD measurement (∼20%). The closure obtained with the "in situ" fake rate measurement, except for a few regions with large statistical uncertainty, is also at 20-30% level.

This method will be used as a cross check if enough statistical precision is achieved and to derive systematic uncertainties: preliminary studies show that the two method agree at 40% level or better. This value is higher than the typical closure on the two separate method because in some cases the non-closure occur in opposite directions.

Measurement of the lepton fake-rate with first data at 50ns bunch spacing

The "fake ratio" for muons and electron (tight-to-fakable ratio) is measured as a function of p_T and η as well as the event kinematics using a QCD-enriched control sample that uses pre-scaled triggers requiring a lepton (with the same ID and isolation as the signal triggers) and one jet. Figure 9.11 shows the muon and electron fake-ratios as a function of the cone-corrected p_T measured in data and compared to simulated QCD multijet events. The measured values are corrected for contamination from prompt leptons from W +jets and Z +jets processes. Uncertainties on this correction is propagated to the fake-ratio as systematic uncertainty.

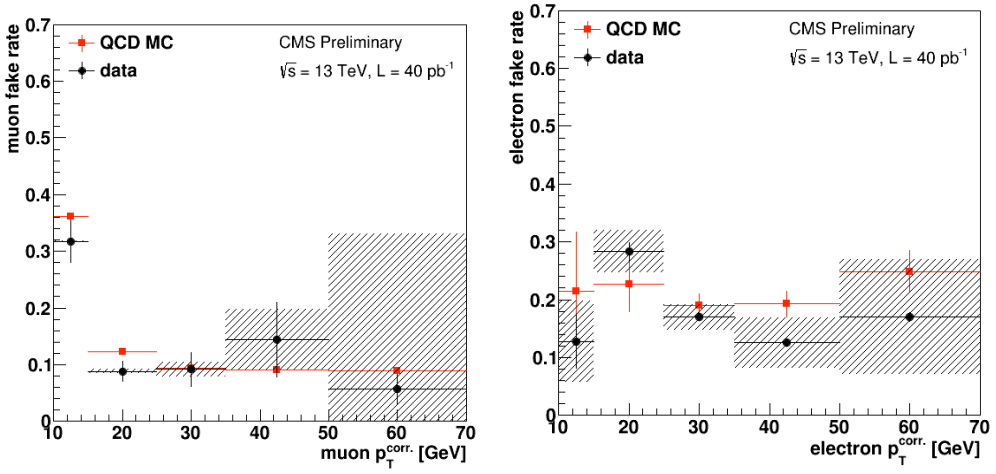


FIGURE 9.11: The muon (left) and electron (right) fake-rate as a function of the cone-corrected muon and electron *pt*.

9.4.2 Rare SM processes

In many signal regions, rare SM processes represent the largest background source. Their contribution is estimated from MC simulation after applying corrections to account for the difference between object selection efficiencies in data and simulation. A conservative 50% systematic uncertainty was assigned in Chapter 6 to this background. Thanks to the new MG5_aMC@NLO [98], more accurate predictions and theoretical systematic uncertainties can be estimated such as QCD scale and PDF uncertainties. These new samples will allow to reduce the uncertainty to the level of 10% in the most favorable regions.

9.4.3 WZ production

WZ production is one of the main backgrounds for the regions with no identified b-jets. While in the 8 TeV analysis it was estimated from MC, a data-driven approach is now foreseen. The WZ normalisation will be estimated from data, in an inclusive orthogonal control region obtained by requiring a third lepton in the event and inverting the Z veto, plus some extra requirements on H_T ($H_T > 80$ GeV), N_{jets} ($N_{jets} > 2$), $N_{b\text{-jets}}$ ($N_{b\text{-jets}} = 0$) and E_T^{miss} ($E_T^{miss} > 50$ GeV). An 85% purity on WZ events is achieved with this control region. After deriving the normalisation factor, the WZ contribution in each signal region is estimated using efficiencies from MC, so additional systematics for the extrapolations will be added. One of the main contribution for the extrapolation to regions requiring at lease one b-tag is the b-tagging mistag rate; previous analysis show than this uncertainty could be of the order of 15%.

Unfortunately, the data collected at 50ns bunch spacing was not sufficient to commission this method.

9.4.4 Charge misidentification

One last source of background to be considered is the electron charge misidentification in opposite-sign events in which one of the electron's charge is badly identified and the event enters into the same-sign selection. This background is estimated by selecting opposite-sign ee or $e\mu$ events passing the full kinematic selection and then weighting them by the electron charge misidentification probability. This probability can be estimated either from MC or from a data control sample. A first measurement with the first data at 13 TeV has been performed:

Measurement of the electron charge misidentification probability with first data at 50ns bunch spacing

The charge misidentification probability for electrons varies in the 0.03-0.3% range depending on the electron p_T and η. Since the momentum of high p_T electrons can be measured reliably by the calorimeter even when the charge is misidentified, it is possible to measure the charge misidentification probability in data by selecting same-sign di-electron events with an invariant mass close to that of an on-shell Z bosons. Figure 9.12 (left) illustrates this by comparing the di-electron mass in opposite-sign and same-sign events. The measured number of same-sign Z bosons is somewhat higher than simulated, but still quite small. This measurement is performed in bins of p_T and η.

An alternative method uses truth information from a MC mixture of $t\bar{t}$ and Z+jets to determine the charge misidentification probability. A closure test is performed with this method by applying the measured charge misidentification probability (from MC truth) to a opposite-sign events in a Z+jets sample in data and comparing them to the observed same-sign events. This is shown in Figure 9.12 (right). Not very good closure is achieved with the data available but further studies are ongoing to try to understand the discrepancies.

9.5 Kinematic distributions with first data at 50ns bunch spacing

Selecting events with two loosely identified leptons and an invariant mass outside the Z mass window, two jets with $p_T > 40$ GeV and at least one reconstructed central jet with

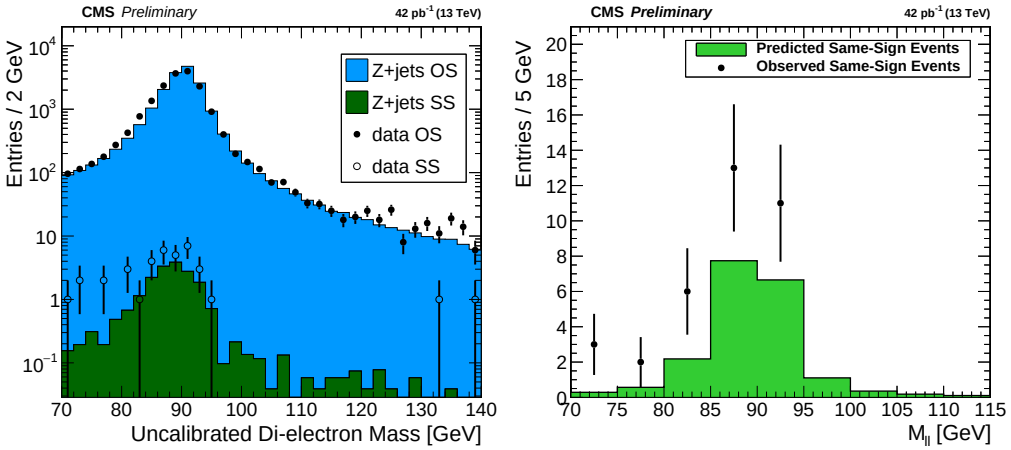


FIGURE 9.12: Left: Opposite-sign and same-sign di-electron events that are used to extract the charge mis-identification probability. Right: Closure test on the charge misidentification probability. The measured charge misidentification probability is applied to a opposite-sign events in a Z +jets sample and comparing them to the observed same-sign events in the same sample without truth-matching

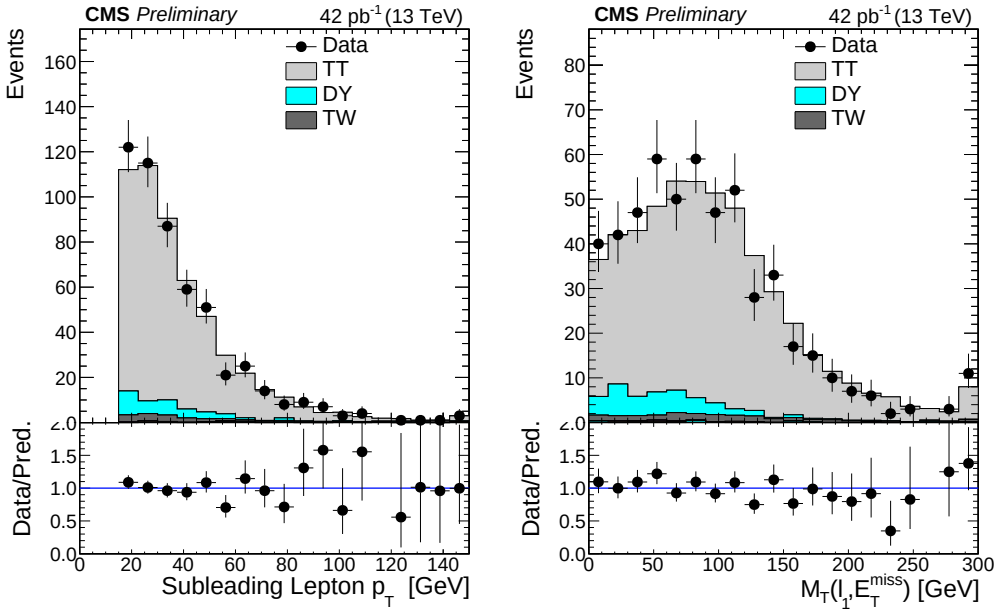


FIGURE 9.13: Data to simulation comparison in the $t\bar{t}$ control region. From left to right: the p_T of the trailing lepton and the transverse mass of the $W \to \ell\nu$.

$p_T > 25$ GeV satisfying the medium working point of the CSV b-tagging discriminator, a sample dominated by dileptonic $t\bar{t}$ will be selected, this sample will be kinematically very similar to the baseline regions and allows to check MC modelling of some significant kinematic distributions. A data/MC comparison is performed in such region, as shown in Figure 9.13.

9.6 Summary

Some developments performed for the search for new physics using same-sign leptons at 13 TeV have been described in this Chapter. Further developments on these methods are currently ongoing in parallel with the understanding of the data as it comes. The changing conditions of the LHC and the change in the expected delivered luminosity during this year will force to modify and adapt some of the methods as the statistical uncertainty will be bigger than the one expected when designing the analysis. The goal is to release a publication by the end of the year with the available recorded luminosity. As a continuation of my thesis work, I will be involved after the defence of this work until the publication.

The LHC restarted its activity with a short running period at 13 TeV but with 50ns bunch-spacing, within this period, 40 pb^{-1} of data were collected by CMS. This dataset has been used to commission the performance of the lepton identification and isolation variables that would be used in the analysis. Auxiliary triggers have been also commissioned and some changes to ensure performance have been included in the 25ns menu. Finally, the performance of the background estimation methods has been also tested and the first measurement of the fake-ratio and the charge mis-identification probability has been carried out with this dataset. Despite the many changes in software reconstruction and the new energies, data is being well described by the simulation. Some of these results have been included in the SUSY performance document [114] which has been recently released.

Chapter 10

Conclusions

This thesis presented here uses proton-proton collision data collected with the CMS experiment at a centre-of-mass energy of 7 and 8 TeV. The four different analysis presented in these pages have allowed to both test the Standard Model as well as explore new physics scenarios, such as SUSY, that has been pushed to its limits given the available luminosity and energy at the time this work has been developed. Same-sign dilepton events, a very rare signature in the SM, have been used to search for new physics and explore new territory, in particular, strong and electroweak production of supersymmetric particles has been probed using pp collision data at 8 TeV. Events with opposite-sign, different flavour lepton pairs have been used to accurately measure the $t\bar{t}$ production cross section using the full luminosity collected at 7 and 8 TeV. A first measurement at 13 TeV using the first data collected at the higher centre-of-mass has been also presented.

A first search for new physics beyond the Standard Model with same-sign dilepton events has been described [52]. The data are analysed using a set of exclusive search regions, targeting strong production of supersymmetry probing signatures both with and without third generation squarks. No significant deviation from the Standard Model expectation is observed. Using sparticle production cross sections calculated in the decoupling limit, and assuming that gluinos decay exclusively into top or bottom squarks and that the top and bottom squarks decay exclusively into top or bottom squarks, lower limits on gluino and sbottom masses are calculated. Gluinos with masses up to 1050 GeV and bottom squarks masses up to 500 GeV are probed. In models where gluinos do not decay to third generation squarks, sensitivity for gluino masses up to 900 GeV is obtained. A similar reach in the gluino masses is demonstrated in the scope of an R-parity violating model.

Secondly, an extension of this search has been designed focusing on the electroweak production of supersymmetric particules using the same dataset [79]. The smaller cross sections motivates a strategy based on multiple dedicated analysis aiming at different regions of the phase space, in particular the results are combined with a three-lepton search that provides higher sensitivity for large mass-splittings while the same-sign search provides extra sensitivity for very small mass difference of the produced particles in the decay chain. Results are compatible with the Standard Model and allow to probe chargino and neutralino masses up to 620 GeV and 720 GeV for the τ-enriched and flavour democratic scenario respectively.

The top quark is the heaviest known particle and it couples heavily to the Higgs Boson and therefore any new physics phenomena may show up deviating the cross-section value from the one predicted by QCD. Furthermore, $t\bar{t}$ production constitutes the dominant source of background for the search covered in this thesis. Thus, a measurement of the inclusive $t\bar{t}$ production cross section in proton-proton collisions at LHC using data samples of 5.0 fb^{-1} at $\sqrt{s} = 7$ TeV and 19.7 fb^{-1} at $\sqrt{s} = 8$ TeV in the $e\mu$ final state has been also presented [104] as part of this work. A cut-and-count technique has been used to extract the cross section value. As a continuation of this analysis, a first measurement of the $t\bar{t}$ production cross section at 13 TeV has been presented [105] without using b-tagging information. Assuming a top quark mass of $m_t = 172.5$ GeV, the measured cross sections are:

$$\begin{aligned}
\sigma_{t\bar{t}}\ (7\ \text{TeV}) &= 165.9 \pm 2.5\ (\text{stat}) \pm 6.2\ (\text{syst}) \pm 3.6\ (\text{lumi})\ \text{pb} \\
\sigma_{t\bar{t}}\ (8\ \text{TeV}) &= 247.6 \pm 1.6\ (\text{stat}) \pm 9.7\ (\text{syst}) \pm 6.3\ (\text{lumi})\ \text{pb} \\
\sigma_{t\bar{t}}\ (13\ \text{TeV}) &= 772 \pm 60\ (\text{stat}) \pm 62\ (\text{syst}) \pm 93\ (\text{lumi})\ \text{pb}
\end{aligned}$$

in good agreement with recent NNLO QCD calculations. The measurements constitute the most precise CMS results on $\sigma_{t\bar{t}}$ and are compatible with recent ATLAS results [83, 103].

Finally, having explored extensively the Standard Model together with many new physics models with supersymmetric particle production and having found no evidence of physics, a complete re-design of the new physics search described in this work seems necessary to improve the sensitivity of this search and maximise the reach at higher energies. Several developments have been already presented in this document, and some of the described methods were already tested and commissioned using the first data collected at 13 TeV

[114]. Other developments are still under study and they will be included in a publication describing the search using the data collected during 2015 at a centre-of-mass energy of 13 TeV.

To conclude, the LHC is entering a new era in particle physics, the increase in the centre-of-mass energy will allow to access the production of heavier particles if they exist. In the context of the search for new physics, the same-sign dilepton final state remains one of the most interesting channels for searching for new physics, both in the context of SUSY as it has been shown in this thesis, but also in other new physics scenarios. Despite the search of new physics, many other interesting processes yield to similar final state such as the associated production of a $t\bar{t}$ pair and a Higgs boson. The increase both in energy and in the integrated luminosity for the upcoming years will facilitate the reach of this final state providing an extraordinary tool to confirm any potential discovery.

Chapter 11

Conclusiones

Los resultados presentados en esta memoria de tesis usan el conjunto de datos completo recogido por el experimento CMS en colisiones protón-protón a energía en centro de masas de 7 y 8 TeV, así como los primeros datos proporcionados por el LHC a 13 TeV en el verano de 2015. En esta memoria se ha estudiado tanto el Modelo Estándar como escenarios de nueva física, en particular, la Supersimetría, que ha sido puesta a prueba usando todos los datos disponibles. Rara vez se encuentran procesos en el Modelo Estándar que produzcan dos leptones de la misma carga eléctrica en el estado final, sin embargo, aparecen de modo natural en diferentes escenarios de nueva física. Es por ello que, dicho estado final se ha utilizado para explorar dichos escenario y en particular, producción fuerte y electro-débil de partículas supersimétricas. Por otro lado, se han usado sucesos con un electrón y un muón de distinta carga eléctrica para medir con precisión la sección eficaz de producción de pares de quarks top-antitop. Utilizando un método similar se ha realizado la primera medida de dicha sección eficaz a una energía en centro de masas de 13 TeV.

En el desarrollo de la memoria, se presenta una descripción de los modelos utilizados, del experimento CMS y del acelerador LHC, junto con una descripción de la reconstrucción e identificación de observables que se usarán en los diferentes estudios presentados en este trabajo. A continuación, el procedimiento que se ha llevado a cabo para estudiar los procesos con dos leptones de la misma carga eléctrica en el estado final y que pueden producirse en procesos derivados de modelos más allá del Modelo Estándar. Los resultados obtenidos se interpretan en términos de la compatibilidad con las predicciones del Modelo Estándar. Esto ha permitido excluir la presencia de partículas supersimétricas, en particular gluinos, con masas por debajo de 1050 GeV [52].

En este contexto, los datos se han utilizado para comprobar la posible producción electrodébil de partículas supersimétricas, cuyas secciones eficaces son sustancialmente más

bajas que las que corresponden a la producción fuerte de este tipo de partículas. Al no observar ninguna desviación con respecto a las predicciones del Modelo Estándar se ha establecido un límite en las masas de las partículas supersimétricas tipo neutralino por debajo de 600 GeV [79].

El quark top es la partícula más pesada conocida y se acopla fuertemente al bosón de Higgs, por ello, cualquier proceso de nueva física puede mostrarse como una pequeña desviación en la sección eficaz de producción de dichas partículas. La producción de pares de quarks top-antitop es también el fondo dominante en las búsquedas descritas en este trabajo. Se ha determinado la sección eficaz de la producción de top-antitop usando la misma muestra de datos que el caso anterior en el estado final con un electrón y un muón. Para extraer el valor de la sección eficaz se ha usado un método de clasificación secuencial de los sucesos ([104], enviado a publicación). Así mismo se ha realizado una primera medida de la sección eficaz de este proceso con los primeros datos recogidos a 13 TeV en 2015 (enviada a PRL, [105]). Asumiendo una masa del quark top de 172.5 GeV, las secciones eficaces medidas son:

$$\begin{aligned}
\sigma_{t\bar{t}}\ (7\ \text{TeV}) &= 165.9 \pm 2.5\ (\text{stat}) \pm 6.2\ (\text{syst}) \pm 3.6\ (\text{lumi})\ \text{pb} \\
\sigma_{t\bar{t}}\ (8\ \text{TeV}) &= 247.6 \pm 1.6\ (\text{stat}) \pm 9.7\ (\text{syst}) \pm 6.3\ (\text{lumi})\ \text{pb} \\
\sigma_{t\bar{t}}\ (13\ \text{TeV}) &= 772 \pm 60\ (\text{stat}) \pm 62\ (\text{syst}) \pm 93\ (\text{lumi})\ \text{pb}
\end{aligned}$$

en buen acuerdo con las predicciones NNLO+NNLL de QCD para esta energía en centro de masas. Estas han sido las medidas más precisas hasta la fecha de esta sección eficaz con el experimento CMS y son compatibles con los resultados publicados por la colaboración ATLAS [83, 103].

En el último capítulo, se han explorado las posibilidades que ofrece el Run II del LHC a esta nueva energía en centro de masas y se han obtenido los primeros resultados de la posible extensión de dichos estudios con respecto a los realizados en el Run I del LHC [114].

Para finalizar, el LHC esta entrando en una nueva era en física de partículas, ya el aumento en la energía en centro de masas permitirá producir partículas mas pesadas (en el caso de que existan). En el contexto de las búsquedas de nueva física, los procesos con dos leptones de la misma carga eléctrica en el estado final sigue siendo uno los canales preferidos para un posible descubrimiento. Además, este mismo estado final aparece en otros procesos sumamente interesantes tales como la producción asociada de un par top-antitop y un bosón de Higgs. En esta nueva era, el aumento de energía y de la

luminosidad recogida en los próximos años proporcionará una herramienta extraordinaria para confirmar cualquier descubrimiento potencial utilizando el estado final antes mencionado.

Bibliography

[1] Georges 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.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.

[2] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys.Rev.Lett.*, 114:191803, 2015. doi: 10.1103/PhysRevLett.114.191803.

[3] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[4] Particle Data Group. Review of particle physics*. *Phys. Rev. D*, 86:010001, Jul 2012. doi: 10.1103/PhysRevD.86.010001.

[5] D. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, New York, USA, 1987.

[6] Chris Quigg. Unanswered Questions in the Electroweak Theory. *Ann.Rev.Nucl.Part.Sci.*, 59:505–555, 2009. doi: 10.1146/annurev.nucl.010909.083126.

[7] Hitoshi Murayama. Supersymmetry phenomenology. pages 296–335, 2000.

[8] Stephen P. Martin. A Supersymmetry primer. *Adv.Ser.Direct.High Energy Phys.*, 21:1–153, 2010. doi: 10.1142/9789814307505_0001.

[9] A.D. Sakharov. Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh.Eksp.Teor.Fiz.*, 5:32–35, 1967. doi: 10.1070/PU1991v034n05ABEH002497.

[10] P.A.R. Ade et al. Planck 2013 results. XVI. Cosmological parameters. *Astron.Astrophys.*, 571:A16, 2014. doi: 10.1051/0004-6361/201321591.

[11] Luc Pape and Daniel Treille. Supersymmetry facing experiment: much ado (already) about nothing (yet). *Reports on Progress in Physics*, 69(11):2843, 2006. URL http://stacks.iop.org/0034-4885/69/i=11/a=R01.

[12] J. Wess and B. Zumino. Supergauge Transformations in Four-Dimensions. *Nucl.Phys.*, B70:39–50, 1974. doi: 10.1016/0550-3213(74)90355-1.

[13] Savas Dimopoulos and Howard Georgi. Softly Broken Supersymmetry and SU(5). *Nucl.Phys.*, B193:150, 1981. doi: 10.1016/0550-3213(81)90522-8.

[14] John R. Ellis, S. Kelley, and Dimitri V. Nanopoulos. Probing the desert using gauge coupling unification. *Phys.Lett.*, B260:131–137, 1991. doi: 10.1016/0370-2693(91)90980-5.

[15] Luis E. Ibanez and Graham G. Ross. Electroweak breaking in supersymmetric models. 1992.

[16] R. Michael Barnett, John F. Gunion, and Howard E. Haber. Discovering supersymmetry with like sign dileptons. *Phys.Lett.*, B315:349–354, 1993. doi: 10.1016/0370-2693(93)91623-U.

[17] Nima Arkani-Hamed, Savas Dimopoulos, and G.R. Dvali. Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity. *Phys.Rev.*, D59:086004, 1999. doi: 10.1103/PhysRevD.59.086004.

[18] Lisa Randall and Raman Sundrum. An Alternative to compactification. *Phys.Rev.Lett.*, 83:4690–4693, 1999. doi: 10.1103/PhysRevLett.83.4690.

[19] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys.Rev.Lett.*, 83:3370–3373, 1999. doi: 10.1103/PhysRevLett.83.3370.

[20] Roberto Contino and Géraldine Servant. Discovering the top partners at the lhc using same-sign dilepton final states. *Journal of High Energy Physics*, 2008(06):026, 2008. URL http://stacks.iop.org/1126-6708/2008/i=06/a=026.

[21] F.M.L. Almeida Jr., Y.A. Coutinho, J.A. Martins Simões, P.P. Queiroz Filho, and C.M. Porto. Same-sign dileptons as a signature for heavy majorana neutrinos in hadron-hadron collisions. *Physics Letters B*, 400(3–4):331 – 334, 1997. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/S0370-2693(97)00143-3. URL http://www.sciencedirect.com/science/article/pii/S0370269397001433.

[22] Yang Bai and Zhenyu Han. Top-antitop and top-top resonances in the dilepton channel at the cern lhc. *Journal of High Energy Physics*, 2009(04):056, 2009. URL http://stacks.iop.org/1126-6708/2009/i=04/a=056.

[23] Edmond Berger, Qing-Hong Cao, Chuan-Ren Chen, Chong Li, and Hao Zhang. Top quark forward-backward asymmetry and same-sign top quark pairs. *Phys. Rev. Lett.*, 106:201801, May 2011. doi: 10.1103/PhysRevLett.106.201801. URL http://link.aps.org/doi/10.1103/PhysRevLett.106.201801.

[24] LHC New Physics Working Group. Simplified models for lhc new physics searches. *Journal of Physics G: Nuclear and Particle Physics*, 39(10):105005, 2012. URL http://stacks.iop.org/0954-3899/39/i=10/a=105005.

[25] Oliver S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, et al. LHC Design Report. 1. The LHC Main Ring. 2004.

[26] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[27] G.L. Bayatian et al. CMS physics: Technical design report. 2006.

[28] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl.Instrum.Meth.*, A262:444–450, 1987. doi: 10.1016/0168-9002(87)90887-4.

[29] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, et al. Electron reconstruction in CMS. *Eur.Phys.J.*, C49:1099–1116, 2007. doi: 10.1140/epjc/s10052-006-0175-5.

[30] Wolfgang Adam, R Frühwirth, Are Strandlie, and T Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. Technical Report CMS-NOTE-2005-001, CERN, Geneva, Jan 2005.

[31] The CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. Technical Report arXiv:1502.02701. CERN-PH-EP-2015-004. CMS-EGM-13-001, CERN, Geneva, Feb 2015. URL http://cds.cern.ch/record/1988091. Comments: Submitted to JINST.

[32] The CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. 2009.

[33] The CMS collaboration. Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7$ tev. *Journal of Instrumentation*, 7(10):P10002, 2012. URL http://stacks.iop.org/1748-0221/7/i=10/a=P10002.

[34] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys.Lett.*, B659:119–126, 2008. doi: 10.1016/j.physletb.2007.09.077.

[35] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008. doi: 10.1088/1126-6708/2008/04/063.

[36] The CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. *Journal of Instrumentation*, 6(11):P11002, 2011. URL http://stacks.iop.org/1748-0221/6/i=11/a=P11002.

[37] The CMS collaboration. Identification of b-quark jets with the cms experiment. *Journal of Instrumentation*, 8(04):P04013, 2013. URL http://stacks.iop.org/1748-0221/8/i=04/a=P04013.

[38] R. Michael Barnett, John F. Gunion, and Howard E. Haber. Discovering supersymmetry with like-sign dileptons. *Physics Letters B*, 315(3–4):349 – 354, 1993. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/0370-2693(93)91623-U. URL http://www.sciencedirect.com/science/article/pii/037026939391623U.

[39] Manoranjan Guchait and D. Roy. Like-sign dilepton signature for gluino production at the cern lhc including top quark and higgs boson effects. *Phys. Rev. D*, 52:133–141, Jul 1995. doi: 10.1103/PhysRevD.52.133. URL http://link.aps.org/doi/10.1103/PhysRevD.52.133.

[40] Howard Baer, Chih-hao Chen, Frank Paige, and Xerxes Tata. Signals for minimal supergravity at the cern large hadron collider. ii. multilepton channels. *Phys. Rev. D*, 53:6241–6264, Jun 1996. doi: 10.1103/PhysRevD.53.6241. URL http://link.aps.org/doi/10.1103/PhysRevD.53.6241.

[41] Hsin-Chia Cheng, Konstantin Matchev, and Martin Schmaltz. Bosonic supersymmetry? getting fooled at the cern lhc. *Phys. Rev. D*, 66:056006, Sep 2002. doi: 10.1103/PhysRevD.66.056006. URL http://link.aps.org/doi/10.1103/PhysRevD.66.056006.

[42] Jonathan L. Feng. Dark Matter Candidates from Particle Physics and Methods of Detection. *Ann.Rev.Astron.Astrophys.*, 48:495–545, 2010. doi: 10.1146/annurev-astro-082708-101659.

[43] Glennys R. Farrar and Pierre Fayet. Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Physics Letters B*, 76(5):575 – 579, 1978. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/0370-2693(78)90858-4. URL http://www.sciencedirect.com/science/article/pii/0370269378908584.

[44] Performance of electron reconstruction and selection with the cms detector in proton-proton collisions at $\sqrt{s} = 8$ tev. *Journal of Instrumentation*, 10(06):P06005, 2015. URL http://stacks.iop.org/1748-0221/10/i=06/a=P06005.

[45] Vardan Khachatryan et al. Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 01:080, 2011. doi: 10.1007/JHEP01(2011)080.

[46] Serguei Chatrchyan et al. Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy. *Phys.Rev.Lett.*, 109:071803, 2012. doi: 10.1103/PhysRevLett.109.071803.

[47] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph 5: going beyond. *Journal of High Energy Physics*, 2011(6):128, 2011. doi: 10.1007/JHEP06(2011)128. URL http://dx.doi.org/10.1007/JHEP06%282011%29128.

[48] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. doi: 10.1088/1126-6708/2006/05/026.

[49] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506: 250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[50] Serguei Chatrchyan et al. Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC. *JHEP*, 1106:077, 2011. doi: 10.1007/JHEP06(2011)077.

[51] Serguei Chatrchyan et al. Search for new physics in events with same-sign dileptons and b jets in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 1303:037, 2013. doi: 10.1007/JHEP03(2013)037,10.1007/JHEP07(2013)041.

[52] Serguei Chatrchyan et al. Search for new physics in events with same-sign dileptons and jets in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 1401:163, 2014. doi: 10.1007/JHEP01(2014)163.

[53] Louis Lyons, Duncan Gibaut, and Peter Clifford. How to combine correlated estimates of a single physical quantity. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 270(1):110 – 117, 1988. ISSN 0168-9002. doi: http://dx.doi.org/10.1016/0168-9002(88)90018-6. URL http://www.sciencedirect.com/science/article/pii/0168900288900186.

[54] The CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. *Journal of Instrumentation*, 6(11):P11002, 2011. URL http://stacks.iop.org/1748-0221/6/i=11/a=P11002.

[55] Serguei Chatrchyan et al. Search for top-squark pair production in the single-lepton final state in pp collisions at $\sqrt{s} = 8$ TeV. *Eur.Phys.J.*, C73(12):2677, 2013. doi: 10.1140/epjc/s10052-013-2677-2.

[56] CMS Collaboration. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. 2013.

[57] A L Read. Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002. URL http://stacks.iop.org/0954-3899/28/i=10/a=313.

[58] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 434(2–3):435 – 443, 1999. ISSN 0168-9002. doi: http://dx.doi.org/10.1016/S0168-9002(99)00498-2. URL http://www.sciencedirect.com/science/article/pii/S0168900299004982.

[59] Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug 2011. URL https://cds.cern.ch/record/1379837.

[60] W. Beenakker, R. Höpker, M. Spira, and P.M. Zerwas. Squark and gluino production at hadron colliders. *Nuclear Physics B*, 492(1–2):51 – 103, 1997. ISSN 0550-3213. doi: http://dx.doi.org/10.1016/S0550-3213(97)80027-2. URL http://www.sciencedirect.com/science/article/pii/S0550321397800272.

[61] A. Kulesza and L. Motyka. Threshold resummation for squark-antisquark and gluino-pair production at the lhc. *Phys. Rev. Lett.*, 102:111802, Mar 2009. doi: 10.1103/PhysRevLett.102.111802. URL http://link.aps.org/doi/10.1103/PhysRevLett.102.111802.

[62] A. Kulesza and L. Motyka. Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the lhc. *Phys. Rev. D*, 80:095004, Nov 2009. doi: 10.1103/PhysRevD.80.095004. URL http://link.aps.org/doi/10.1103/PhysRevD.80.095004.

[63] Wim Beenakker, Silja Brensing, Michael Krämer, Anna Kulesza, Eric Laenen, and Irene Niessen. Soft-gluon resummation for squark and gluino hadroproduction. *Journal of High Energy Physics*, 2009(12):041, 2009. URL http://stacks.iop.org/1126-6708/2009/i=12/a=041.

[64] Bobby Samir Acharya, Phill Grajek, Gordon L. Kane, Eric Kuflik, Kerim Suruliz, et al. Identifying Multi-Top Events from Gluino Decay at the LHC. 2009.

[65] Gordon L. Kane, Eric Kuflik, Ran Lu, and Lian-Tao Wang. Top Channel for Early SUSY Discovery at the LHC. *Phys.Rev.*, D84:095004, 2011. doi: 10.1103/PhysRevD.84.095004.

[66] Rouven Essig, Eder Izaguirre, Jared Kaplan, and Jay G. Wacker. Heavy Flavor Simplified Models at the LHC. *JHEP*, 1201:074, 2012. doi: 10.1007/JHEP01(2012)074.

[67] Michele Papucci, Joshua T. Ruderman, and Andreas Weiler. Natural SUSY Endures. *JHEP*, 1209:035, 2012. doi: 10.1007/JHEP09(2012)035.

[68] Michael Krämer, Anna Kulesza, Robin van der Leeuw, Michelangelo Mangano, Sanjay Padhi, Tilman Plehn, and Xavier Portell. Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV. 2012.

[69] CMS Collaboration. Interpretation of searches for supersymmetry with simplified models. Submitted to *Phys. Rev. D*, 2013.

[70] Serguei Chatrchyan et al. Search for supersymmetry in final states with missing transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions using the variable α_{T}. *JHEP*, 01:077, 2013. doi: 10.1007/JHEP01(2013)077.

[71] Serguei Chatrchyan et al. Search for supersymmetry in hadronic final states with missing transverse energy using the variables α_{T} and b-quark multiplicity in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J. C*, 73:2568, 2013. doi: 10.1140/epjc/s10052-013-2568-6.

[72] Serguei Chatrchyan et al. Search for gluino mediated bottom- and top-squark production in multijet final states in pp collisions at 8 TeV. *Phys. Lett. B*, 725:243, 2013. doi: 10.1016/j.physletb.2013.06.058.

[73] Georges Aad et al. Search for Gluinos in Events with Two Same-Sign Leptons, Jets and Missing Transverse Momentum with the ATLAS Detector in pp Collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 108:241802, 2012. doi: 10.1103/PhysRevLett.108.241802.

[74] Georges Aad et al. Search for supersymmetry in pp collisions at $\sqrt{s} = 7$ TeV in final states with missing transverse momentum and b-jets with the ATLAS detector. *Phys. Rev. D*, 85:112006, 2012. doi: 10.1103/PhysRevD.85.112006.

[75] Georges Aad et al. Search for top and bottom squarks from gluino pair production in final states with missing transverse energy and at least three b-jets with the ATLAS detector. *Eur. Phys. J. C*, 72:2174, 2012. doi: 10.1140/epjc/s10052-012-2174-z.

[76] Georges Aad et al. Multi-channel search for squarks and gluinos in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS detector. *Eur. Phys. J. C*, 73:2362, 2013. doi: 10.1140/epjc/s10052-013-2362-5.

[77] Csaba Csaki, Yuval Grossman, and Ben Heidenreich. Minimal flavor violation SUSY: A Natural Theory for R-Parity Violation. *Phys. Rev. D*, 85:095009, 2012. doi: 10.1103/PhysRevD.85.095009.

[78] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002. doi: 10.1088/1126-6708/2002/06/029.

[79] Vardan Khachatryan et al. Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and W, Z, and Higgs bosons in pp collisions at 8 TeV. *Eur.Phys.J.*, C74(9):3036, 2014. doi: 10.1140/epjc/s10052-014-3036-7.

[80] Benjamin Fuks, Michael Klasen, David R. Lamprea, and Marcel Rothering. Revisiting slepton pair production at the Large Hadron Collider. *JHEP*, 01:168, 2014. doi: 10.1007/JHEP01(2014)168.

[81] Benjamin Fuks, Michael Klasen, David R. Lamprea, and Marcel Rothering. Precision predictions for electroweak superpartner production at hadron colliders with Resummino. *Eur. Phys. J. C*, 73:2480, 2013. doi: 10.1140/epjc/s10052-013-2480-0.

[82] Benjamin Fuks, Michael Klasen, David R. Lamprea, and Marcel Rothering. Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV. *JHEP*, 10:081, 2012. doi: 10.1007/JHEP10(2012)081.

[83] Georges Aad et al. Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector. *Eur.Phys.J.*, C74(10):3109, 2014. doi: 10.1140/epjc/s10052-014-3109-7.

[84] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 1402:024, 2014. doi: 10.1007/JHEP02(2014)024,10.1007/JHEP02(2014)102.

[85] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *JHEP*, 1303:015, 2013. doi: 10.1007/JHEP03(2013)015.

[86] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 0207:012, 2002. doi: 10.1088/1126-6708/2002/07/012.

[87] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was, and Z. Was. Universal Interface of TAUOLA Technical and Physics Documentation. *Comput.Phys.Commun.*, 183:821–843, 2012. doi: 10.1016/j.cpc.2011.12.009.

[88] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 1006:043, 2010. doi: 10.1007/JHEP06(2010)043.

[89] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys.Rev.*, D82:054018, 2010. doi: 10.1103/PhysRevD.82.054018.

[90] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 1107:018, 2011. doi: 10.1007/JHEP07(2011)018.

[91] John M. Campbell and R. Keith Ellis. $t\bar{t}W^{+-}$ production and decay at NLO. *JHEP*, 1207:052, 2012. doi: 10.1007/JHEP07(2012)052.

[92] M.V. Garzelli, A. Kardos, C.G. Papadopoulos, and Z. Trocsanyi. t $\bar{t}$ W^{+-} and t $\bar{t}$ Z Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects. *JHEP*, 1211:056, 2012. doi: 10.1007/JHEP11(2012)056.

[93] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput.Phys.Commun.*, 185: 2930, 2014. doi: 10.1016/j.cpc.2014.06.021.

[94] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, et al. The PDF4LHC Working Group Interim Recommendations. 2011.

[95] Serguei Chatrchyan et al. Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C73(3):2339, 2013. doi: 10.1140/epjc/s10052-013-2339-4.

[96] Torbjorn Sjostrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191: 159–177, 2015. doi: 10.1016/j.cpc.2015.01.024.

[97] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008. doi: 10.1140/epjc/s10052-008-0798-9.

[98] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:29, 2002.

[99] Nikolaos Kidonakis. Top Quark Production. In *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)*, pages 139–168, 2014. doi: 10.3204/DESY-PROC-2013-03/Kidonakis. URL http://inspirehep.net/record/1263209/files/arXiv:1311.0283.pdf.

[100] Federico Demartin, Stefano Forte, Elisa Mariani, Juan Rojo, and Alessandro Vicini. The impact of PDF and α_s uncertainties on Higgs Production in gluon fusion at hadron colliders. *Phys. Rev. D*, 82:014002, 2010. doi: 10.1103/PhysRevD.82.014002.

[101] Gregory Mahlon and Stephen J. Parke. Spin Correlation Effects in Top Quark Pair Production at the LHC. *Phys. Rev.*, D81:074024, 2010. doi: 10.1103/PhysRevD.81.074024.

[102] Werner Bernreuther and Zong-Guo Si. Top quark spin correlations and polarization at the LHC: standard model predictions and effects of anomalous top chromo moments. *Phys. Lett.*, B725:115–122, 2013. doi: 10.1016/j.physletb.2013.06.051, 10.1016/j.physletb.2015.03.035. [Erratum: Phys. Lett.B744,413(2015)].

[103] Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 13$ TeV using $e\mu$ events with b-tagged jets. Technical Report ATLAS-CONF-2015-033, 2015.

[104] Measurement of the ttbar production cross section in the emu channel in pp collisions at 7 and 8 TeV. Technical Report CMS-PAS-TOP-13-004, CERN, Geneva, 2015. URL https://cds.cern.ch/record/2048642.

[105] Measurement of the top quark pair production cross section in proton-proton collisions at sqrt(s)=13 TeV with the CMS detector. Technical Report CMS-PAS-TOP-15-003, CERN, Geneva, 2015. URL https://cds.cern.ch/record/2044918.

[106] Measurement of the inclusive and differential tt production cross sections in lepton + jets final states at 13 TeV. Technical Report CMS-PAS-TOP-15-005, CERN, Geneva, 2015. URL http://cds.cern.ch/record/2048622.

[107] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV with lepton + jets final states. *Phys. Lett.*, B720:83–104, 2013. doi: 10.1016/j.physletb.2013.02.021.

[108] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 11:067, 2012. doi: 10.1007/JHEP11(2012)067.

[109] CMS Collaboration. Top pair cross section in e/μ +jets at 8 TeV. CMS Physics Analysis Summary CMS-PAS-TOP-12-006, 2012.

[110] Combination of ATLAS and CMS top quark pair cross section measurements in the $e\mu$ final state using proton-proton collisions at 8 TeV. ATLAS and CMS Conference Notes ATLAS-CONF-2014-054, CMS-PAS-TOP-14-016, 2014.

[111] Timo Antero Aaltonen et al. Combination of measurements of the top-quark pair production cross section from the Tevatron Collider. *Phys. Rev.*, D89(7):072001, 2014. doi: 10.1103/PhysRevD.89.072001.

[112] CMS SUSY fake-leptons working group. Studies of methods to estimate the non-prompt lepton background to searches for new physics. *CMS Physics Analysis Note*, AN-2014/261, 2015. URL http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2014_261_v2.pdf.

[113] G. Aad et al. Search for anomalous production of prompt same-sign lepton pairs and pair-produced doubly charged Higgs bosons with $\sqrt{s} = 8$ TeV pp collisions using the ATLAS detector. *JHEP*, 1503:041, 2015. doi: 10.1007/JHEP03(2015)041.

[114] CMS Collaboration. Validation of key observables used in SUSY searches with first data at $\sqrt{s} = 13$ TeV. 2015.