Università degli Studi di Napoli "Federico II"

Scuola Politecnica e delle Scienze di Base
Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Magistrale in Fisica

Search for a new vector boson W’ coupling to the third generation quarks in leptonic final states with the CMS detector at LHC

Tesi Sperimentale

Relatori:
Dott. Alberto Orso Maria Iorio
Prof. Luca Lista
Correlatore:
Dott. Stefano Morisi

Candidato:
Andrea Piccinelli
Matricola:
N94/432

Anno Accademico 2018/2019

To everyone who supported me in this journey

There are the rushing waves
mountains of molecules
each stupidly minding its own business
trillions apart
yet forming white surf in unison

Ages on ages
before any eyes could see
year after year
thunderously pounding the shore as now.
For whom, for what?
On a dead planet
with no life to entertain.

Never at rest
tortured by energy
wasted prodigiously by the Sun
poured into space.
A mite makes the sea roar.

Deep in the sea
all molecules repeat
the patterns of one another
till complex new ones are formed.
They make others like themselves
and a new dance starts.
Growing in size and complexity
living things
masses of atoms
DNA, protein
dancing a pattern ever more intricate.

Out of the cradle
onto dry land
here it is
standing:
atoms with consciousness;
matter with curiosity.

Stands at the sea,
wonders at wondering: I
a universe of atoms
an atom in the Universe.

(R. P. Feynman)

Contents

Introduction

The Large Hadron Collider, briefly LHC, is the largest hadron accelerator in the world, designed to provide proton-proton collisions at a centre of mass energy of $\sqrt{s} = 14$ TeV, and with a luminosity of 10^{34} cm^{-2} s^{-1}. Around the collision points, the LHC is equipped with four main experiments, namely ALICE, ATLAS, CMS and LHCb, whose goal is to provide further proofs on the validity of the Standard Model (SM) of Particle Physics and, moreover, to search for new physics at the TeV energy scale.

The SM is the theory that currently best explains the fundamental interactions and particles, receiving several experimental confirmations over the years. Despite its extraordinary success, the SM fails to give an explanation for many other phenomena, with some notable problems still unresolved. Several new physics beyond SM theories have been consequently formulated and, among them, some predict the existence of new charged gauge bosons, W'. It arises from extensions of the electroweak sector of the SM, it has spin 1 and electric charge ± 1 and can be directly detected with pp collisions at the LHC in its leptonic decay channel, $W' \to l\nu_l$, or in its hadronic one, $W' \to tb$, but also indirectly through its effects on low-energy processes, due to interference involving the SM W boson.

The goal of this thesis is to search for a new W' resonance, investigating events with final states compatible with the chain decay $W' \to tb \to l\nu_l bb$. We performed the analysis of two different collections of proton-proton collision data, delivered by the LHC at a centre of mass energy of $\sqrt{s} = 13$ TeV, the first one of 41.53 fb^{-1}, and second one of 59.74 fb^{-1}. They were recorded by the CMS experiment in 2017 and 2018, respectively.

This thesis is organized as follows:

- **chapter 1**, a description about the LHC accelerator machine, the CMS experiment and their operation;
- **chapter 2**, a brief introduction of the Standard Model and a description of its unsolved issues;
- **chapter 3**, concerning the theoretical and experimental state of the art about the W' boson;
- **chapter 4**, about signal characterization, final state, and description of the algorithms used to identify and reconstruct physical objects of interest for the presented physics analysis;
- **chapter 5**, displaying the analysis strategy with description of the datasets and the simulated samples used, of the the event selection and of the fit procedure with its results.

Chapter 1

An overview about the LHC and the CMS experiment

The tremendous tragedy of the Second World War left Europe in ruins, with deep ripercussions on all aspects of society, which includes of course the fundamental research. European science hobbled with respect to the American and Sovietic ones, with the main scientific facilities located in USA and in USSR [62]. The scientists of the Old World felt the desire to build a cutting-edge research center. With the goal of translating that wish into reality, in 1952 twelve European countries [1] assembled a council of scientists that was named, in French, *Conseil Européen pour la Recherche Nucléaire* (European Council for Nuclear Research, in English), called briefly as CERN. After two years, the project of the European research center became reality and the *Organisation Europenne pour la Recherche Nuclaire* (European Organization for Nuclear Research) was born, retaining the same acronym of the next-to-dissolve *Conseil*.

The headquarter of the organization is situated on the border between Switzerland and France, on the north-western outskirts of the city of Geneva. Nowadays, it has 22 member states, joined by some observers-state, including non-European countries. CERN is currently the largest research center in high energy physics, fulfilling the original dream of the founders, and provides the necessary tools to explore the intimate nature of matter and the forces that govern it, making it possible to study the origin of the Universe and its evolution. Among these instruments there are particle accelerators, whose purpose is to accelerate atoms, ions, and subnuclear particles, close to the speed of light, and to let them collide. In the point of the collisions between the particles there are the detectors, which make us to observe these events under controlled conditions.

At CERN operates a network of accelerators, built one after the other in various periods of the history of the institute, starting from the its foundation. Already from the beginning, it has planned that every new and more powerful machine would have used the previous ones as “injectors”. In this way, it has been possible to create a chain of accelerators that brings a beam of particles to increasingly higher energies.

In order to let the chain to work, all the devices are coordinated by a single

[1] namely: Belgium, Denmark, France, the Federal Republic of Germany, Greece, Italy, the Netherlands, Norway, Sweden, Switzerland, the United Kingdom and Yugoslavia.

reference signal, generated by an atomic clocks system and distributed throughout the installation, with an accuracy of the order of the nanosecond. The latest accelerator to be added to this chain is the Large Hadron Collider (LHC), which is the most powerful particle accelerator, and it's able to lead the energy of the particles to values of the order of TeVs. We will give more details about LHC in the next sections.

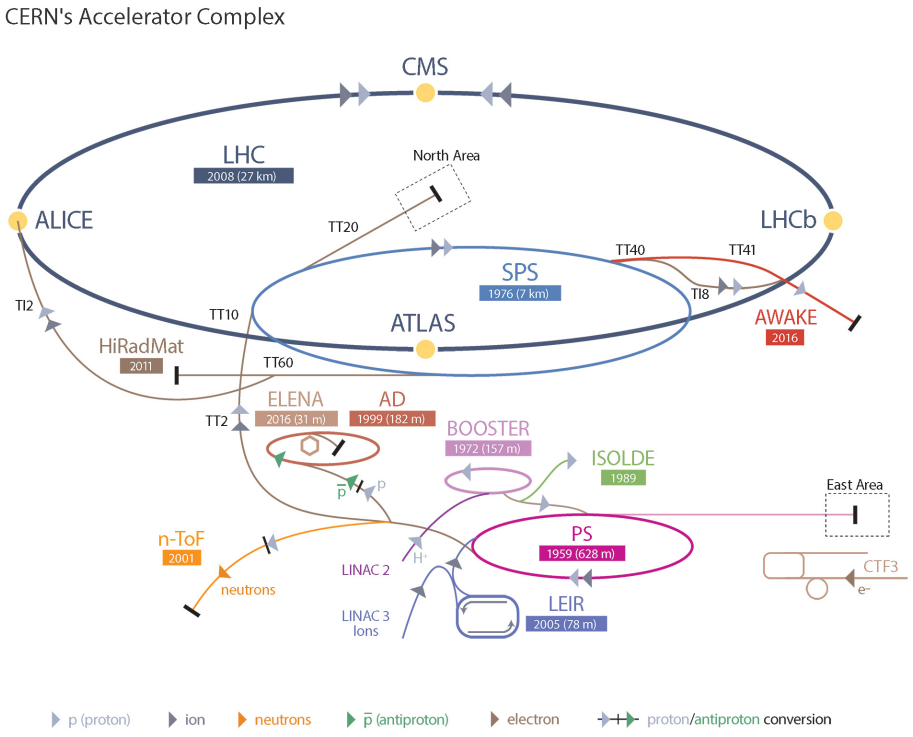


Figure 1.1: The accelerator chain at the CERN.

LHC is currently in the so called phase "Long Shutdown 2" (LS2), and any operations are stopped to let people to implement the upgrades planned for the operative phase "Run 3". Before the LS2, the following machines were at work:

- **Two linear accelerators (LINAC 2 and LINAC 3)**
  Linear accelerators use radiofrequency cavities to create some electromagnetic fields inside the accelerating structure. The particles pass through the cavities, which are alternately charged positive or negative. The conductors behind the beam push them, the conductors ahead pull, making the particles to accelerate. Superconducting magnets spread all over the accelerator ensure particle beams are focalised. LINAC 2 accelerated protons, obtained from hydrogen molecules spilling off their electrons thanks to an electric field which they pass through, to 50 MeV into the Proton Synchrotron Booster (PSB). At the beginning of LS2, the place of LINAC 2 has been taken by LINAC 4. This device will be able to accelerate protons up to 160 MeV, and it will be a key element in the project to increase the luminosity of the LHC during the Run 3.

  LINAC 3, instead, as happened during the previous Runs, will provides heavy ions, specificaly lead ions obtained by stripping away the electrons

from the molecules, at 4.2 MeV for nucleon into the Low Energy Ion Ring (LEIR). LINAC 3 is expected to be in use until at least 2022.

- **Proton Synchrotron Booster (PS Booster)**
  It is made up of four superimposed synchrotron rings, that received beams of protons from the accelerator, up to Run 2, from LINAC 2 at 50 MeV. Then, it accelerated them to 1.4 GeV for injection into the Proton Synchrotron (PS). Before the Booster received its first beams on 26 May 1972, protons were injected directly from the LINAC into the PS, where they were accelerated to 25 GeV. The low injection energy of 50 MeV limited the number of protons the PS could accept. So, inserting the Booster at this stage of the chain allowed the PS to accept over 100 times more protons, which greatly enhanced the use of the beams for experiments.

- **Low Energy Ion Ring (LEIR)** This instrument accelerated ions from the LINAC 3 to the Proton Synchrotron (PS), in order to provide ions for collisions within the Large Hadron Collider (LHC). LEIR received long pulses of lead ions from LINAC 3 and splitted each one of them into four shorter bunches, each containing 22108 lead ions. It takes about 2.5 seconds for LEIR to accelerate the bunches, two for every cycle, from 4.2 MeV to 72 MeV. The ions were then at a suitable energy to be passed to the PS.

- **Proton Synchrotron (PS)** This was CERN first synchrotron, beginning its operation in 1959. With a circumference of 628 metres, the PS has 277 conventional (i.e. at room temperature) electromagnets, including 100 dipoles to bend the (knees) beams, placed all around the ring. The accelerator operated at up to 25 GeV. Currently the PS is a key component in CERN accelerator complex, as it usually accelerates either protons delivered by the PS Booster or heavy ions from the LEIR.

- **Super Proton Synchrotron (SPS)**
  It is a circular accelerator with a diameter of 2 kilometres built in a tunnel and started its operations in 1976. It was designed to deliver an energy of 300 GeV and was gradually upgraded to 450 GeV, has 1317 conventional electromagnets, including 744 dipoles to bend the beams round the ring. Besides it being used for fixed-target experiments, like COMPASS and NA62, it has been operated as a proton-antiproton collider (named in that form SppS collider), and for accelerating high energy electrons and positrons which were injected into the Large Electron Positron Collider (LEP). Since 2008, it has been used to inject protons and heavy ions into the Large Hadron Collider (LHC). It also started to feed the AWAKE expriment, which aims to test new techniques for accelerating particles. Research using SPS beams has probed the inner structure of protons, investigated natures preference for matter over antimatter, looked for matter as it might have been in the first instants of the universe and searched for exotic forms of matter. While the SPS was running as a SppS, a major highlight came in 1983 with the Nobel-prize-winning discovery of W and Z bosons.

- **Large Hadron Collider (LHC)** [27] It is the largest and most powerful accelerator of particles in the world. It was built by the CERN on the border

between France and Switzerland between 1998 and 2008, thanks to the participation of over 10000 scientists and engineers, countless research institutes and universities. Its aim is trying to answer several fundamental questions in the physics of elementary particles, e.g. the origin of mass and the asymmetry between matter and antimatter, to perform precision measurements testing the validity of the Standard Model (MS) and researching new physics.

- **On-Line Isotope Mass Separator (ISOLDE)** This is a facility dedicated to the production of a large variety of radioactive ion beams for many different experiments in the fields of nuclear and atomic physics, solid-state physics and materials science. It permits the study of the vast territory of atomic nuclei, including the most exotic species. The high intensity proton beam from the Proton Synchrotron Booster is directed into specially developed thick targets, yielding a large variety of atomic fragments. Different devices are used to ionize, extract and separate nuclei according to their mass, forming a low-energy beam that is delivered to severak experimental stations. This beam can be further accelerated, making possible to do various studies on nuclear reactions.

- **Antiproton Decelerator (AD)** This instrument is a storage ring that produces low-energy antiprotons for studies of antimatter, and creates antiatoms. A proton beam that comes from the PS is fired into a block of metal. These collisions create a multitude of secondary particles, including lots of antiprotons having too much energy to be useful for making antiatoms. They also have different energies and move randomly in all directions. The job of the AD is to tame these particles and turn them into a useful, low-energy beam that can be used to produce antimatter. The AD is composed of bending and focussing magnets that keep the antiprotons on the same track, while strong electric fields slow them down. The spread in energy of the antiprotons and their deviation from their track is reduced by a technique known as cooling. Antiprotons are subjected to several cycles of cooling and deceleration, until they are slowed down to around a tenth of the speed of light. They are then ready to be ejected into the antimatter experiments.

In the future, LHC will face a massive upgrade, that will bring luminosity at values never reached before for hadronic colliders. Such phase is known as High Luminosity LHC (HL-LHC) and will be able to achieve instantaneous luminosities a factor of five larger than the LHC nominal value, thereby enabling the experiments to enlarge their data sample by one order of magnitude compared with the LHC baseline programme.

After this stage, in a hopefully-not-so-far some projects of new colliders at CERN will become reality, as FCC or CLIC, in order to overcome the present experimental limits of the LHC and enlarge our knowledge of Nature and its secrets that still lie beyond what we can probe with our current instruments.

1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular accelerator of hadrons, specifically of protons and heavy ions. The task of this accelerator is to bring the hadrons at close to the speed of light, reaching a set energy, and then let them to collide in four interaction points. In every point a detector is located, and each one of them observes the results of aforementioned collisions.

According to the machine's original design specifics, the energy range that can be reached spans from 450 GeV per beam, i.e. the injection energy of SPS, up to a maximum of 7 TeV per beam, resulting in a centre of mass energy reaching up to 14 TeV.

Let's shift, for a moment, our focus on an important parameter of any accelerator, namely the **luminosity**. We can define the **instantaneous luminosity** $\mathcal{L}(t)$ as the proportionality factor between the cross section of a process σ and the number of events observed per unit of time in the collision R, i.e. the rate,

$$R = \mathcal{L}\sigma$$

So, it has the dimensions of a flow. Integrating in the time $\mathcal{L}(t)$ we can also define the **integrated luminosity** L,

$$L = \int_0^T \mathcal{L}(t)dt$$

where T is the time of observation.

The instantaneous luminosity can be measured taking into account both the machine parameters taking and the geometric and kinematic characteristics of the beam. Assuming a Gaussian profile of the beams and a head-on collision, the instantaneous luminosity is given by

$$\mathcal{L} = \frac{N_p^2 n_b f \gamma}{4\pi\varepsilon_n \beta^*} F$$

where:

- N_p is the number of particles per bunch;
- n_b is the number of bunches per beam;
- f is the revolution frequency;
- γ is the relativistic Lorentz factor $\left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$;
- ε_n is the normalized transverse beam emittance;
- β^* is the beta function at the collision point, and it is a measure of how narrow the beam is at the interaction point;
- F is the geometric luminosity reduction factor due to the crossing angle at the interaction point.

The geometric luminosity reduction factor can be written as

$$F = \sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2}$$

with θ_c the full crossing angle of the beams at the interaction point, σ_z the bunch length and σ^* transverse RMS beam size at the interaction point.

The Table 1.1 shows the values of the parameters above for the operating period of 2017 and 2018.

Parameter	2017 Value	2018 Value
N_p	1.25×10^{11}	1.1×10^{11}
n_b	$2556 - 1868$	2556
f	40 MHz	40 MHz
γ	4260	4260
ε_n	$1.8 - 2.2$ μm	$1.8 - 2.2$ μm
β^*	$40 - 30$ cm	$30 - 25$ cm
θ_c	$300 - 140$	$320 - 260$
σ_z	9.4 cm	9.4 cm
σ^*	19 μm	19 μm

Table 1.1 Values of the LHC machine parameters for the operating period of 2017 and 2018.

The machine started a first run of data taking (called Run 1) in 2010 with a centre of mass energy of 7 TeV, brought to 8 Tev in 2012. After the end of Run 1, in 2013, the LHC has a technical break, known as Long Shutdown 1 (LS1), to undergo accelerator as well as detector upgrade and mantainance operations. LHC restarted its operations in 2015, reaching this time a centre of mass energy of 13 TeV and operating at a instantaneous luminosity of 2×10^{-34} cm^{-2} s^{-1}. This data taking period is called Run 2 and it ended in 2018. Afterwards, a nee technical break called Long Shutdown 2 has begun, and crucial upgrades of both the accelerators and detectors has started to be implented. As told before, LINAC 4 took the place of LINAC 2, and it's going to accelerate the protons, up to 160 MeV during the Run 3, making the lumonisity to increase up to an expected value of 140 fb^{-1}. Detectors are going to face important upgrades of their hardware and software structures, in order to keep the pace of the increase of the luminosity of the next data taking period.

As regards the more engineering hand, the LHC is built inside a 27 km long underground tunnel, at 100 m depth on average, and it's located on the border between France and Switzerland, in a region between the Geneva airport and the Jura mountains. This tunnel originally was excavated to host the Large Electron-Positron Collider (LEP).

The LHC consists of the following parts:

- 1232 dipole magnets, each one of them 15 me long, that have the task of keeping the orbit of the beam circular and inside the accelerator;
- 392 quadrupole magnets, each 57 metres long, that collimate the beam;

- different radiofrequency cavities that accelerate the beam;
- hexapolar, ocular and superior magnets that further correct the orbit of the particles.

The magnets are superconductors based on the niobium-titaniun (NbTi) Rutherford cables technology, which need temperatures of about 1.9 K to work properly. This allows currents of the order of 35 kA, capable of producing magnetic fields of 8.3 T, to circulate.

The entire circuit is cooled down by the largest refrigeration system in the world, with its 96 tons of superfluid He4. Because the space in the tunnel is limited, a twin-bore design is used, in such a way that it is possible to use only one cryogenic structure with proton rings in the same cryostat. Anyway, this solution requires the presence of oppositely oriented magnetic fields, in order to permit the coexistence of two proton beams along the same circumference. The two beams are kept on parallel orbits and are brought together in a single beam pipe only near the interaction point.

LHC also requires vacuum systems for the insulation of the cryomagnets, for the helium distribution and the required vacuum conditions required for the transportation of the beams. The typical vacua at cryogenic temperatures in the interaction point require a pressure included about in the range $10^{-10} - 10^{-11}$ mbar.

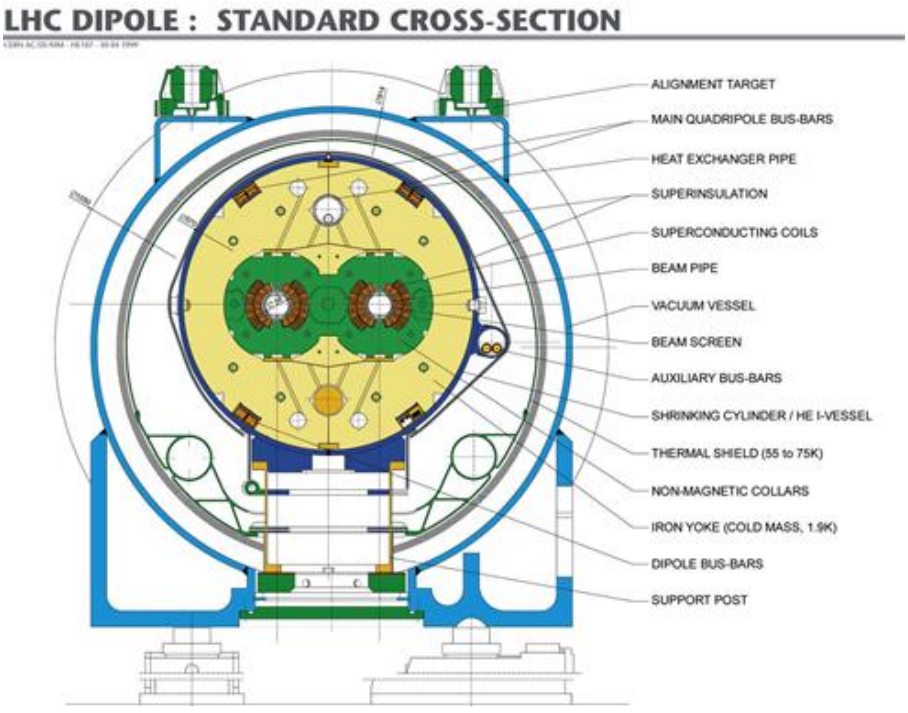


Figure 1.2: A cross section of the LHC dipole.

As fow what concerns the particles that are used in this accelerator, they are mainly protons coming from hydrogen gas, produced by using an electric field that deprives the hydrogen atoms of their electrons. Until the end of Run 2, the proton beams, before being injected into the LHC, passed through a complex of accelerators that increased the energy of the beam up to 450 GeV. The first

accelerator was LINAC 2, that generated protons of 50 MeV which were placed in PSB which, in its turn, brought the energy of the beam to 1.4 GeV and, then, sent them into the PS, which had the task to increase the energy up to 25 GeV. The next step was the the SPS, which carried the protons to 450 GeV, ready, finally, to be placed both clockwise and anticlockwise direction in the main ring. This last step, that took the time of about 20 minutes, brought the particles to reach the regime energy.

At this final stage, the protons were injected into the LHC as bunches of 1.25×10^{11} protons (in 2017 and 2018). Each beam had 2448 circulating proton bunches, which were arranged in groups of 3 and 4 trains of 72 bunches (questa struttura dei branch l'ho presa da Rosalba, ma non sono sicuro che sia valita anche per il 2017), with 25 ns spacing within the train corresponding to 8 empty bunches between two trains. At every bunch crossing occoured the collisions between the beams, so the resulting maximum collisions rate is 40 MHz (anche qui, non sono sicuro che sia così).

The proton beams rotate for many hours in the LHC pipes before colliding in the four points, each one with a main experiment. These ones are:

- **ALICE (A Large Ion Collider Experiment)** It is dedicated to the study of collisions of heavy nuclei (especially Pb-Pb) in a range of very high density. In these conditions ,it is expected to observe the existence of the quark and gluon plasma (QGP), i.e. a new state of matter in which, under the conditions of high density and temperature, quarks and gluons are not anymore confined in hadrons.

- **ATLAS (A Toroidal LHC ApparatuS)** and **CMS (Compact Muons Solenoid)** They are general purpose dectors, i.e. detectors designed to identify a large variety of experimental signatures, in order to allow a wide-ranging study of processes that could happen at LHC. Among the objectives of the experiments are the study of the top and beauty quarks and the Higgs boson, the search for supersymmetric particles or other particles beyond those provided by the SM. This two experiments differ mainly in the configuration of the magnets for the muon detection system, as in ATLAS is toroidal, and in CMS is solenoid. That distinction affects also the magnetic field dictating the trajectory of muons, reflecting, in this way, upon reconstruction techniques.

- **LHCb (Large Hadron Collider beauty experiment)** has been designed and optimized to study the properties of quark b and matter-antimatter asymmetry. Between the four experiments, LHCb is the only one in which the two colliding beams don't have the same energy; in fact, one is at regime energy (7 Tev), the other is at injection energy (450 GeV).

1.2 The CMS experiment

1.2.1 Detector

The CMS detector is built around a huge solenoid magnet, from which it is named after [63]. This magnet has the form of a cylindrical coil of superconducting cable that generates a field of 3.8 T. The field is confined by a steel yoke that forms the bulk of the detectors and it's 14000 tons weight. Instead of being built *in situ*, like the other giant detectors of the LHC experiments, the CMS detector was constructed in fifteen sections at ground level before being lowered into an underground cavern near Cessy, in France, and reassembled.

The detector, in its complete form, is 21 metres long, 15 metres wide and 15 metres high. The structure of the CMS was thought in order to reveal particles deriving from the collision of hadrons and to measure their masses, momenta, energies and charges. The detector, in fact, is made up of several sub-detectors that allow to identify different particles and reconstruct their characteristics with an high momentum resolution on a wide energy and angular coverage range.

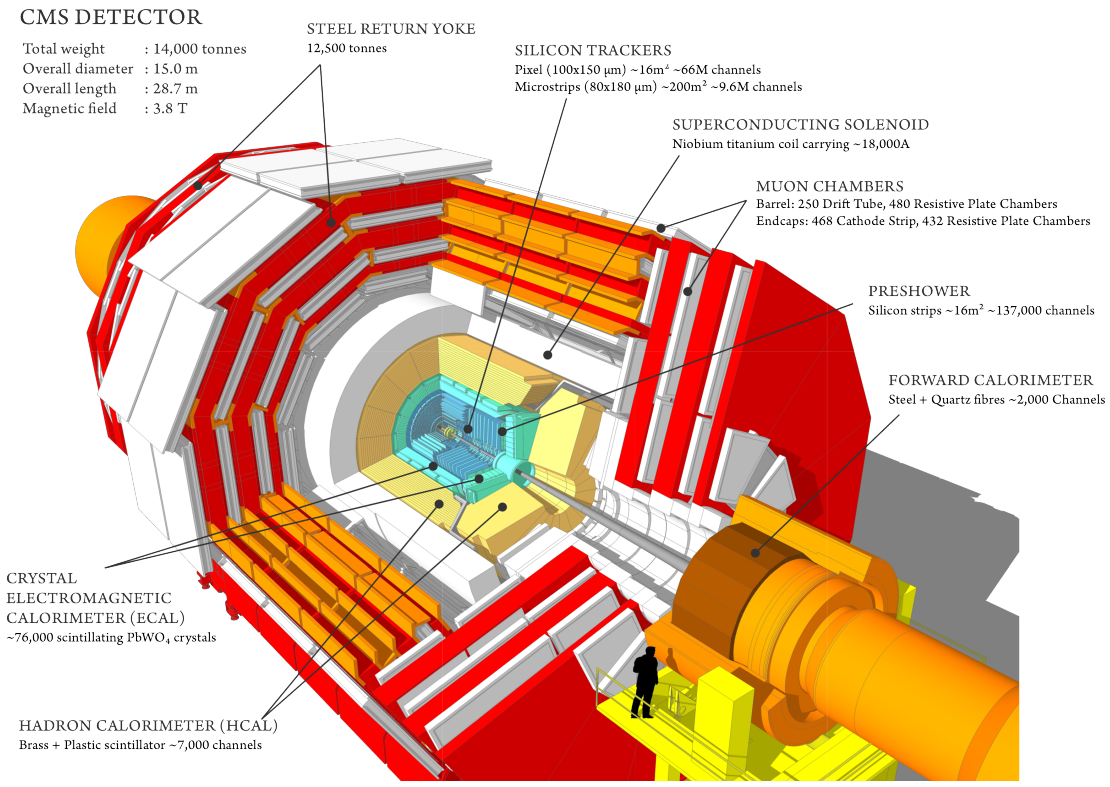


Figure 1.3: A global view of CMS detector.

Moreover, as working at high luminosities, the detector has been designed to operate in a high radiation environment, maintaining good performances over the course of several years of data taking, and optimised to be able to distinguish processes of interest from backgrounds.

Furthermore, at these energies and luminosity a big limit to the physical observations is represented by the so-called **pile up**, that is the overlap between

different particles of the same event or coming from interactions in the same bunch-crossing.

In order to decrease the impact of the pile-up on the quality of the recorded data, it is required that the detector has high granularity, i.e. high spatial resolution. In addiction, good time resolution is required to avoid the response of the detector (or the readout electronics) to be slower than the nominal time interval between two bunch-crossings.

One of the most important requirements of a detector like CMS is the tightness, and because of that CMS has a cylindrical structure (called barrel), that covers the central region, and two caps (endcaps), covering the regions closest to the beams. In this way, it makes possible the cylindrical symmetry around the beam axis to be respected.

For the sake of the measurements, we can as the coordinate system of CMS a right-handed Cartesian frame centred at the nominal interaction point and oriented in the following way:

- the x axis points towards the center of the LHC ring and represents the horizontal coordinate;
- the y axis points upwards, perpendicular to the LHC plane, and represents the vertical coordinate;
- the xy plane represents the transverse plane with respect to the collision axis;
- the z axis is directed along the anticlockwise direction of the beam with respect to xy and represents the longitudinal direction.

Given the cylindrically-symmetric geometry of the CMS detector, the coordinate system often used to reconstruct the tracks of particles is a cylindrical coordinate system, and it has the following structure:

- the radial distance r from the z axis;
- the azimuth angle φ, that is the angle of rotation around the z axis, with value equals to zero when r lies on the x axis and increasing clockwise when one looks in the positive direction of the z axis;
- the polar angle θ is defined as the rotation angle around the x axis with origin on the z axis and increasing clockwise looking in the positive direction of the x axis. A variable often used in place of θ is the pseudorapidity η, defined as

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

  This variable has the advantage that, in relativistic approximation, transforms linearly under boosts along the z axis, implying that differences of this quantity between relativistic particles along the z axis is invariant. Obviously, as the angle increases from zero, the pseudorapidity decreases from infinity.

Using these coordinates, the angular distance between two particles directions can be written as another Lorentz invariant variable:

$$\Delta R = \sqrt{(\Delta\varphi)^2 + (\Delta\eta)^2} \tag{1.2.1}$$

Referring to the Cartesian system, other very useful variables are the transverse momentum p_T and the transverse energy E_T, defined as follows:

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{1.2.2}$$

$$E_T = E \sin\theta \tag{1.2.3}$$

1.2.2 The CMS system of subdetectors

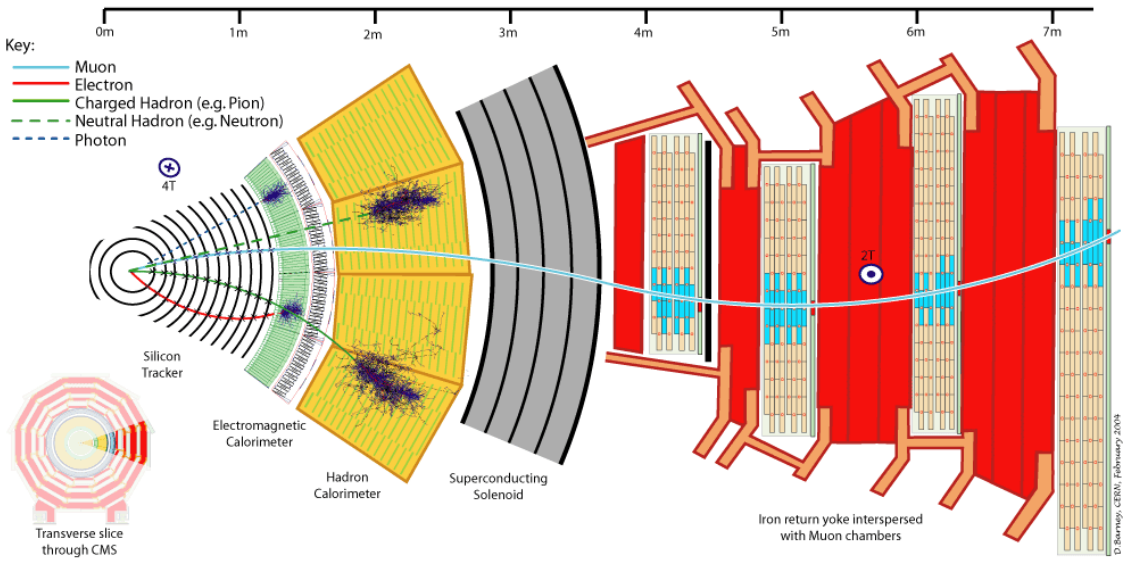


Figure 1.4: A slice view of the structure of CMS.

We can analyse the structure of the CMS detector starting from the interaction point and proceeding towards the outside, as shown in fig. 1.4. Thus, the CMS subdetectors layout is the following:

- **Tracking system**
  [40, 61] It accurately measures the positions of passing charged particles, allowing to reconstruct their tracks. The CMS magnetic field spirals the paths of the charged particles, in order to measure their momenta. It has a radius of 1.2 m and a length of 5.8 m, with a geometric acceptance of $|\eta| < 2.5$. It has an active area of 200 m^2 and it's made up of

  - **pixel vertex detector**, the closest subsystem to the interaction point, fundamental for the reconstruction of particles with a very short average life, contains 65 million pixels and allows to measure the position with an accuracy of 10 μm for the radial measurements and of 20 μm for the transverse measurements. This detector is formed by cylindrical layers of pixels, placed at 4 cm, 7 cm and 11 cm from the vacuum

tube of the beam, and two discs at each end. As it is so close to the collision, the number of particles passing through is huge (the rate of particles received at the distance of 8 cm from the beam line is around 10 million particles per cm^2/s) and, so, the pixel detector has to disentangle and reconstruct all the tracks they leave behind, and withstand such a harsh radiation over the ten-year duration of the experiment.

- **Silicon strip tracker**, more external, has a resolution between 35-52 μm in the radial direction and 530 μm in the transverse direction. The silicon detectors work in much the same way as the pixels: whether a charged particle crosses the material, it knocks electron from atoms and within the applied electric field these move giving a very small pulse of current, that lasts a few ns. This small amount of charge is, then, amplified by APV25 chips, giving us hits when a particle passes and, because of these hits, allowing us to reconstruct its path. Due to the nature of their job, the tracker and its electronics are pummeled by radiation, but they are designed to withstand it. To minimise disorder in the silicon and avoid it causing damage, this part of the detector is kept at -20 C.

Each one of the detectors is present both in the barrel and in the endcaps. The barrel has three pixel and ten strip layers, while the endcap sections have two pixel and twelve strip layers. The barrel strips and pixels are parallel to the beam axis, while the endcap ones are disposed orthogonally to it, in order to allow measurement of tracks at higher η.

- **Electromagnetic calorimeter (ECAL)**
[58] This is a hermetic homogeneous calorimeter, with radius between 1.2 m and 1.8 m, that measures the energy released by electrons and photons emerging from collisions. They are very useful for finding new elements of physics, as happened in 2012 for the Higgs boson. To find new physics states with the necessary precision in the very extreme conditions of the LHC (such as a high magnetic field, high levels of radiation and only 25 ns between collisions), it was decided to use 61200 lead tungstate crystals ($PbWO_4$), mounted in the central barrel part, and 7342 crystals in each of the two endcaps. The characteristics of the $PbWO_4$ transparent and scintilating crystals, as the high density (8.28 g/cm^3), short radiation length (0.89 cm) and small Molière radius (2.2 cm), lead to a fine granularity and a compact calorimeter that is capable of absorbing electrons and photons in 23 cm of length. The scintilating material covers values of pseudorapidity of $|\eta| < 3.0$ and has a decay time of about 10 ns, which allows to collect 85% of light in 25 ns.

- **Hadronic calorimeter (HCAL)**
[59] It is a hermetic sampling calorimeter that measures energy released, position and arrival time of hadrons, such as protons, neutrons, pions, and kaons. It also provides indirect measurement of non-interacting, uncharged particles, such as neutrinos, as measuring the missing transverse energy. To reduce the effects of inefficiency and not to have areas where a particle can pass through unnoticed, HCAL uses alternating layers of absorber,

made from a non-magnetic material, and fluorescent scintillator materials, that produces a rapid light pulse when the particle passes through. Special optic fibres collect up this light and feed it into readout boxes, where photodetectors amplify the signal. In order to contain and measure the showers of particles produced by the interaction of a hadron with the absorber material, the minimum thickness of material required is about one meter. To achieve the result, HCAL is divided into barrel (HB and HO), endcap (HE) and forward (HF) sections. The barrel and endcap sections are located around the ECAL, respecting the concentric cylindrical geometry, and they over a pseudorapidity range of $|\eta| < 3.0$. They are made up of brass absorber plates interleaved with scintillator tiles that are embedded with wavelenght shifting (WLS) fibers. The brass absorber is commonly known as cartridge brass (C26000) and is composed of 70% Cu and 30% Zn, its density is 8.83 g/cm^3, with radiation lenght of 1.49 cm and nuclear interaction length of 16.42 cm. Into barrel, HB is the last detector placed inside the superconducting coil of the magnet, with a radial distance of 1.8 m < r < 2.9 m, while HO is placed outside the coil, and ensures that there are no losses of energy from the back of the HB.

Lastly, the HF is located 11.2 m away from the interaction point, making possible the pseudorapidity range $|\eta| < 5.0$ to be included, and picks up the particles coming out of the collision region at shallow angles relative to the beam line. This detector receives the bulk of the particle energy contained in the collision and, thus, it has to be very resistant to radiation and use different materials to the other parts of the HCAL. Because of that, it is made up of quartz fibres, embedded within steel absorber, and uses a Cherenkov-based technology.

Now we can briefly make some quantitative considerations about the energy resolution of both the ECAL and the HCAL. Tipically, for a calorimeter this quantity can be parametrized as follows:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{a}{\sqrt{E\ \text{[GeV]}}}\right)^2 + \left(\frac{b}{E\ \text{[GeV]}}\right)^2 + c^2$$

with:

- a stochastic term, dominant at low energies, takes into account intrinsic statistical shower flutuations, sampling fluctuations and signal quantum fluctuations;
- b noise term, negligible at low luminosity, takes into account both apparent energy due to electronic noise, radio-activity and pileup;
- c constant term, dominant at high energies, takes into account fluctuations of longitudinal leakage, inhomogeneities (hardware or calibration), imperfections in calorimeter construction (dimensional variations, etc.), fluctuation of the electromagnetic component in hadronic shower, non-linearity of readout electronics and fluctuations in energy lost in dead material, before or within the calorimeter.

The energy resolution of the electromagnetic calorimeter of CMS is:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{0.0280}{\sqrt{E\ [\text{GeV}]}}\right)^2 + \left(\frac{0.12}{E\ [\text{GeV}]}\right)^2 + 0.0030^2$$

while the energy resolution of the hadronic calorimeter of CMS is:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{0.8470}{\sqrt{E\ [\text{GeV}]}}\right)^2 + 0.0740^2$$

- **Magnet**
  In CMS it is composed of a superconducting coil that produce an almost uniform magnetic field, and is built to reach a magnetic field of 3.8 T when electricity flows through it. The magnet is the instrument around which the whole experiment is designed (and named after), and its purpose is to curve the path of the particles born in the collisions of the LHC beams. Once the trajectory is known, it is possible to measure the impulse of a particle and its electric charge. To prevent border effects and make the lines of force of the external magnetic field to be as uniform as possible, there is an iron yoke outside the coil which surround the magnet. This yoke is composed of alternating layers interspersed with the muon detectors, that induces a 1.8 T constant field in the region outside the magnet, too. The magnet has 2.9 m $< r <$ 3.8 m and $|\eta| < 1.4$.

- **Muon system**
  [60] This part of CMS is a system of muon detectors used for muon identification, momentum measurement and triggering. Because the muons can penetrate several meters of iron and of matter in the tracker and in the calorimeters without decay, they have a clean signature. They can also appear in final states of many new physics processes, and can be reconstructed with high precision thanks to the informations inside and outside the magnet. Their properties make them good candidates for precision measurements. At this stage, CMS uses three different types of gas detectors: Drift Tubes, Cathode Strip Chambers and Resistive Plate Chambers . Let's analyse their properties.

  - **The Drift Tubes (DTs)** are positioned in the barrel. Using a drift chamber as the tracking detector for the barrel muon system turned out to be a good choice because of the low expected rate and the relatively low intensity of the local magnetic field. The CMS barrel muon detector is made of four stations: the first three innermost have 60 drift chambers each, the outermost has 70. Since at momenta larger than 40 GeV the rate of background generated by neutrons and gammas is from 10 to 100 times larger than the one from prompt muons, a redundancy of information is required; this goal is obtained by having several layers of separated drift cells per station.

    The tracking and timing performance of a single chamber was optimized using twelve layers of drift tubes divided into three groups of

four consecutive layers, named Super Layers (SL). Into first three stations, each chamber has the middle SL that measures the coordinate along the direction parallel to the beam, z, and the two outside SLs that measure the perpendicular coordinate (r, φ). The fourth station doesn't contain the z-measuring planes. To have the global resolution in (r, φ) per chamber of 100 μm, the single wire resolution has to be better than 250 μm. The baseline cell has a pitch of 40 mm by 13 mm. At the center there is the anode wire, made out of 50 μm diameter stainless steel type 304L. The cathodes defining the cell width are aluminum I-beams, 1.2 mm thick and 9.6 mm high. A plastic profile, made of 0.5 mm thick extruded poly-carbonate plastic (Lexan), is glued to the upper and lower parts of the I-beams in order to electrically insulate the cathodes from the aluminum plates. Each cell of each chamber is offsetted by a half-cell width with respect to their neighbour, with the goal of eliminating dead spots in the efficiency. The choice of a tube as the basic drift unit was made in order to obtain protection against the damage from a broken wire and, moreover, to partially decouple contiguous cells from the electromagnetic debris that accompany the muon itself.

- **The Chatode Strip Chambers (CSCs)**, located in the endcaps, are able to provide precise space and time information in the presence of uneven magnetic field and high particle rate, thanks to their fast response time, fine segmentation, and radiation resistance. SCs are multiwire proportional chambers with segmented cathode readout and made up of arrays of positively-charged anode wires crossed with negatively-charged copper cathode strips, placed in a gas volume. There are four stations of CSCs in each endcap, with chambers having trapezoidal shape and arranged, perpendicular to the beam line, in a series of concentric rings centered on it. The stations are separated by the iron disks of the flux return yoke, which are thick enough to isolate the electrons in showers. The last station is followed by a 100 mm thick iron disk, whose primary purpose is shielding the station from backsplash backgrounds induced by particles scattered at small angle and interacting with the forward instruments. Each CSC has six layers of wires sandwiched between cathode panels, which provide robust pattern recognition for rejection of non-muon backgrounds and efficient matching of external muon tracks to internal track segments. Therefore, each chamber provides six measurements of the φ coordinate (strips) and six measurements of the r coordinate (wires). Strip width varies from 3 to 16 mm for different chambers, or from about 2 to 5 mrad in φ coordinates. Muon identification is ensured over the angular range corresponding to the range $10 < \theta < 170$.
- **The Resistive Plates Chambers (RPCs)**, positioned both in the barrel that in the endcaps, are gaseous parallel-plate chambers that combine a reasonable level of spatial resolution with impressive time resolution, comparable to that of scintillators, and operate in avalanche mode to ensure good operation at high rates. A resistive plate cham-

ber is constructed of two parallel plates of material made of phenolic resin, with good surface flatness and a high bulk resistivity. Typically, the plate separation is of the order of a few mm. The resin material is coated with a conductive graphite paint to form electrodes and, then, readout is made by means of aluminum strips outside the resin plates, insulated from the electrodes by some plastic material. In the barrel muon system there are six layers of RPCs, two in each of the first two stations, and one in each of the last two stations. In this way it is possible to use the trigger algorithm to work even for low-p_T tracks, that may stop before reaching the outer two stations. In the endcap region, there is a plane of RPCs in each of the four stations to let the trigger to use the coincidences between stations in order to reduce background, to improve the time resolution for bunch crossing identification, and to achieve a good p_T resolution. At the end, a sophisticated alignment system measures the positions of the muon detectors with respect to each other and to the inner tracker, in order to optimize the muon momentum resolution.

These types of gaseous detectors are chosen because both the volume to be occupied is very large, and they are less expensive than other tracking technologies, besides they are reliable and robust to the radiation. The system covers a pseudorapidity range of $|\eta| < 2.4$ and a radial range of $4.0m < r < 7.4m$.

During the actual LS2, CMS is facing the second part of the Upgrade - Phase 1 (with the first one exploited during the LS1), and the CMS Gas Electron Multiplier (GEM) subsytem is going to be installed. It will form an extra layer of detectors, along with the Cathode Strip Chambers (CSC) layers, closest to the beam. The GEMs' ability to cope with a very high particle rate will greatly improve muon measurement in this very difficult region. This will be especially important when LHC will become the High-Luminosity LHC, when the number of collisions will increase by almost a factor of 10.

After this upgrade, 72 GEMs superchambers, called GE1/1, will be installed into the CMS. Together, they will form the GE1/1 subsystem, the first completely new subdetector system to be introduced into CMS since it was built in the late 2000's. This subsystem will be located in the forward muon system of CMS, in the difficult η environment in front of the already-existing cathode strip chambers, as mentioned before. The GE1/1 subsystem adds much-needed redundancy in a high rate and high background environment, allowing for better tracking and measurement of the bending angle at the trigger level. This, in turn, will decrease the number of mis-measured muons by permitting to lower the trigger threshold of soft muons.

- **Trigger and data acquisition (DAQ) system**
  Its function is to select and store events of potential interest for the analysis. Referrig us to the Run 2 data taking phase, at full LHC luminosity twenty inelastic proton-proton collisions occur every 25 ns, but only a small fraction of them are hard scattering interactions containing events with an interesting signature, and the rest are mostly minimum bias (MB) events. Given the

high interaction rates, of about 40 MHz, it is impossible to store and process all produced data with the current technologies, so a drastic rate reduction is achieved to allow the writing on mass memory, and this last operation occurs with a frequency of 100 Hz. At CMS, the rate is reduced by the trigger system in two steps called Level-1 (L1) [15] Trigger and High-Level Trigger (HLT) [21], respectively.

L1 is an extremely fast (its decision has to be made every 25 ns, without deadtime) and wholly automatic process that looks for simple signs of interesting physics, for instance particles with a large amount of energy, and reduces the event rate from 40 MHz to 100 kHz (50 kHz at low luminosity). It takes decisions based on the kinematics of the individual object. The triggered objects are, then, passed to the subsequent DAQ system and HLT for further reconstruction and selection.

HLT reduces the rate to 100 Hz, performing more detailed reconstruction of objects: the system assimilates and synchronises the information from different parts of the detector to recreate the entire event and, after that, sends it to a farm of more than 1000 standard computers. Here, the PCs run very complex physics tests to look for specific signatures, for instance matching tracks to hits in the muon chambers, or spotting photons through their high energy but lack of charge.

The fig. 1.5 shows the data flow in the trigger and DAQ system.

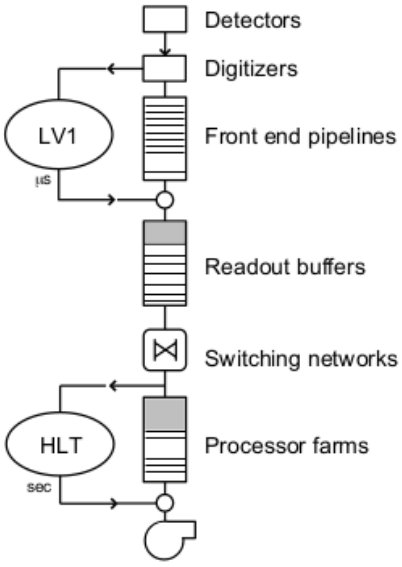


Figure 1.5: Trigger and DAQ system built for CMS

Chapter 2

The Standard Model of Particle Physics

The **Standard Model of Particle Physics** (briefly SM) is the set of quantum field theories that describes the fundamental components of matter and their interactions observed in nature so far at the microscopic scale. Its first formulation in the current form dates back to 1950s and, since then, it has been tested to an ever increasing level of precision. The SM has provided precise predictions of fundamental interactions (except the gravitational one), that have been confirmed by experiments throughout the years [2, 4, 8, 9, 12, 14, 20]. The SM is however not sufficient to provide a full picture of the known universe, as it fails to explain several phenomena, like the evidence for neutrino masses, dark energy and dark matter, and does not include any satisfying interpretation of the gravitational interaction.

2.1 An overview on Standard Model

The SM is capable to describe three of the four known fundamental interactions, as it defines in a coherent model the unification of the electromagnetic and weak interactions [30, 50, 64], the Higgs mechanism [26, 35, 38] for the appearance of the masses, and the strong interaction [33, 47] within the framework of the Quantum Chromodynamics (QCD) [65], including its features and the asymptotic freedom, by introducing a new degree named colour (so the SU$(3)_C$ symmetry) [31, 36].

The SM is a quantum field theory and, in addition to the usual spacetime symmetries, it exhibits invariance under transformations of the three gauge groups. To every gauge symmetry we can relate one of the fundamental interactions. The three gauge group, in fact, are associated to three of the four known fundamental interactions. In a gauge theory, a gauge symmetry group is associated with a set of massless spin-1 (vector, then) fields, of number equals to the number of parameters which features uniquely an element of the group. These fields are called bosons, as they obey the Bose-Einstein statistics, and they are:

- the circle group U$(1)_Y$, whose transformation can be represented by a unitary scalar complex operator multiplied by its quantum number (whose

value depending from the particle under investigation), acting as the equivalent of the electromagnetic charge, called **weak hypercharge** Y; the associated boson is called B_μ;

- the $n = 2$ special unitary group SU(2)$_L$, whose three fundamental transformations can be represented by the 2×2 Pauli matrices σ_i $(i = 1, 2, 3)$ multiplied by a quantum number called **weak isospin** I_3; the three associated vector fields are $W_\mu^{\pm,0}$;
- the $n = 3$ special unitary group SU(3)$_C$, whose eight operations can be represented by the 3×3 Gell-Mann matrices λ_j $(j = 1, \ldots, 8)$ multiplied by a charge called **colour** C, that can takes the value r, g, b. The eight associated vector fields are $G_\mu^{1,\ldots,8}$.

In this way there are 12 vector fields associated with 3 gauge symmetries, all of them can be summarized by the tensor product of groups that defines the gauge symmetry of the Standard Model:

$$\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$$

The subgroup $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ defines the symmetry of the unified electroweak interaction.

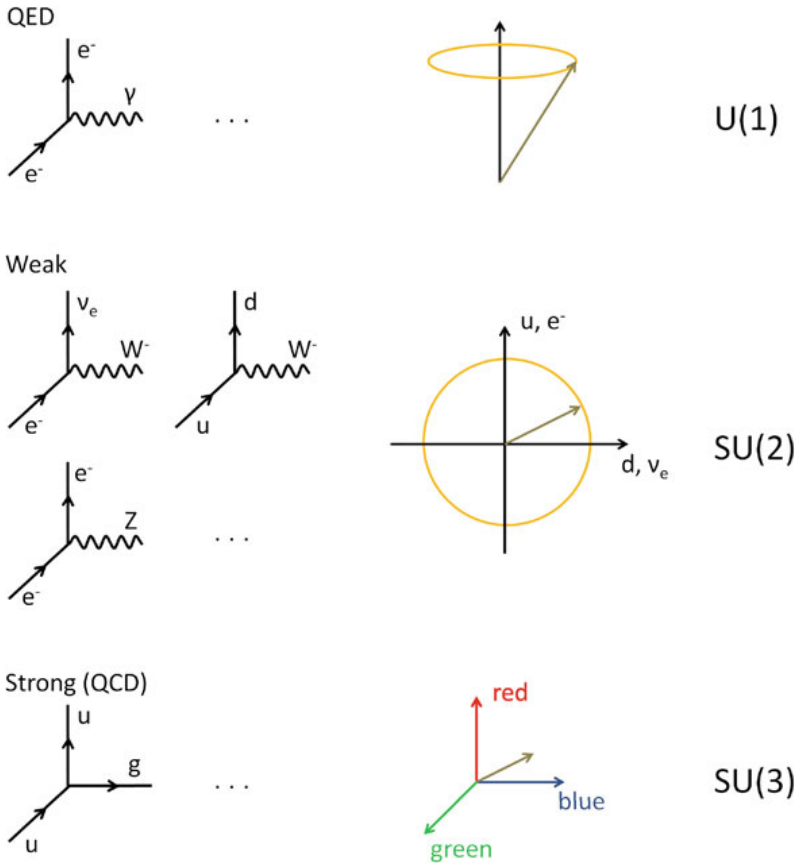


Figure 2.1: A graphic representation of the gauge symmetry groups of the SM and corresponding interactions.

As we will show further, thanks to the Higgs mechanism three of these massless fields acquire mass and become the known physics weak boson: W^+, W^-, Z.

The other nine fields remain massless: the eight gluons g for the strong force and the photon γ for the electromagnetic one. The bosons are the mediators of the known forces and regulate the interactions among other spin-1/2 particles of the SM, called **fermions**. For the sake of clarity, we list the following explanations:

- the electromagnetic force is carried by spin-1 photons γ and acts between electrically charged particles;
- the weak interaction is ruled by the three massive gauge vector bosons, $W^{\pm}$ and Z, discovered in 1983 by the UA1 and UA2 [8, 9, 12, 14] experiments, and it is responsible for phenomena like nuclear β decays;
- the strong interaction, intermediated responsible of holding together nuclei its gauge bosons are the gluons g.

As we pass to the matter sector of the SM, we meet 12 fermion fields with half-integer spin and, because of that, they obey to the Fermi-Dirac statistics; they are classified as lepton or quark fields. Among the lepton fields, fermions are grouped in three generations or families, with each one of them that is an isospin doublet of particles, with specific isospin quantum number, electromagnetic and weak charges. Leptons don't interact via strong force, while quark fields are fermions of spin $1/2$ which carry electromagnetic, weak and colour charges. The latter, in particular, plays for the strong interactions the same role of the charge for the electromagnetic interactions. Quarks are grouped in three families, similarly to whath happens for the fermions. According to Dirac equation, that is the equation of motion of the fundamental fermions, and the CPT conservation theorem, all the fermionic fields have an associated anti-particle one that has the same mass, but opposite quantum numbers.

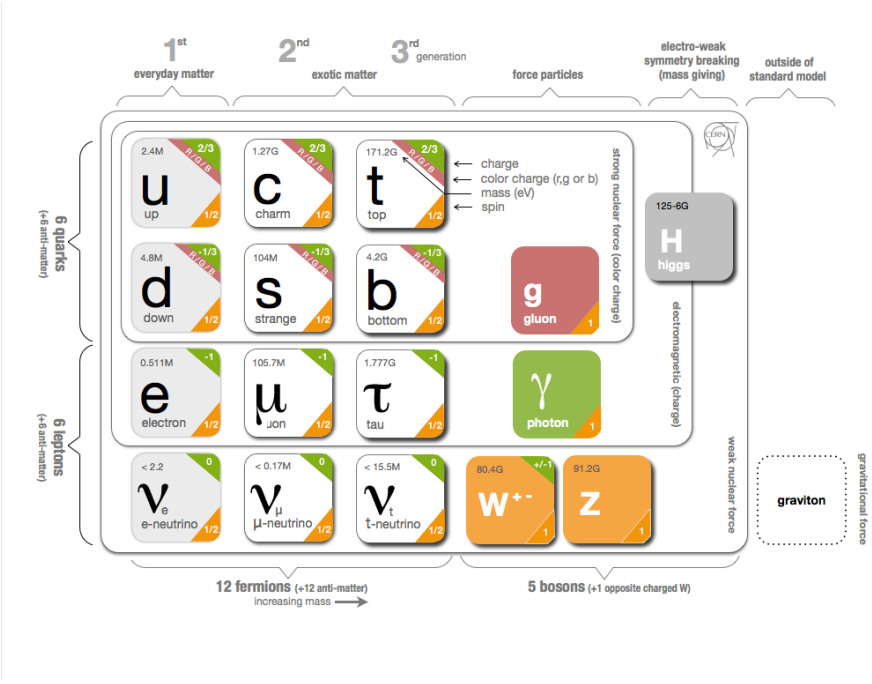


Figure 2.2: A pictorical scheme of the fields included in the Standard Model.

Fermions have also a property called **chirality**, and they can either be left-chiral, or right-chiral. At relativistic velocity regimes, chirality can be considered an observable equivalent to the elicity, that is the projection of the spin of a particle on its own momentum, and, so, the "chiral" tags are sometimes referred to as left- or right-handed. This property is very important, moreover, becauste it is experimentally verified that left-handed fermion fields transform differently under the SU$(2)_L$ gauge symmetry with respect to the behaviour, under the same symmetry transformation, of the right-handed fermion fields. This fact results in an asymmetry in the properties of fermions. Specifically, there are left-handed and right-handed charged leptons, left-handed neutrinos, but no right-handed neutrinos (as, even if they exist, they are undetectable because they would interact with literally nothing, for what we know so far). This does not yet have an explanation from prime principles, and so it is considered as a postulate of the theory. The last field that we have to consider in order to complete the picture of the Standard Model is a complex scalar doublet field φ, named the Higgs field after one of the theorists who predicted its existence in 1964 [38]. In fig. 2.2 the particles predicted by the SM are listed.

In the following, we are going to pay our attention about the three gauge theories of the SM interactions and the Higgs mechanism.

2.2 SM Quantum Field Theories

2.2.1 Quantum Electrodynamics (QED)

As for what happened for the Einstein's relativity, the source of inspiration for the quantum field theories was the electrodynamics. In classical electrodynamics, in fact, we can see Maxwell's equations can be derived by imposing invariance of the lagrangian under a gauge transformation of the involved fields. A similar reasoning led to the formulation of QED in the first place, and later on of the other fundamental interactions. The first relativistic quantum field theory to be developed, in fact, was the **Quantum Electrodynamics (QED)**. It is an Abelian gauge theory, able to describe the dynamics and interactions of fermions with the electromagnetic field. The lagrangian density for the QED can be obtained starting from the free lagrangian density $\mathcal{L}_D$ of the Dirac field ψ with mass m:

$$\mathcal{L}_D = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi \tag{2.2.1}$$

where the first is the kinetic term and the last is the mass term, with γ^μ the Dirac matrices and ψ and $\overline{\psi}$ the 4-components spinor and its overlineoint, respectively. The eq. (2.2.1) describes the kinematics of a free non-interacting fermion and it is under a global gauge transformation of the symmetry group U(1). The QED is associated with this symmetry group, by defining U$(1)_q$ transformations, where the subscript indicates the charge as conserved quantum number. This number assumes the meaning of the charge of the particles that rise as excitations of the fermionic fields. A spinor ψ transforms under a U$(1)_q$ transformation as follows:

$$\psi \longrightarrow \psi' = \psi e^{iq\theta}$$

with θ an arbitrary constant.

In order to embed the interactions of the fermionic fields with electromagnetic one, we have to require the eq. (2.2.1) to satisfy the **local gauge invariance** principle. This condition means that the equation has to be invariant under a local gauge transformation of the fields involved, and not trivially under a global one. If we perform a local gauge transformation of the fields in the following way:

$$\psi \longrightarrow \psi' = \psi e^{iq\theta(x)}$$
$$\overline{\psi} \longrightarrow \overline{\psi'} = \overline{\psi} e^{-iq\theta(x)}$$

we can see the eq. (2.2.1) does not invariantly transform, as the derivative doesn't invariantly transform, too:

$$\partial_\mu \psi \longrightarrow \partial_\mu \psi' = \partial_\mu \phi e^{iq\theta(x)} + iq\psi e^{iq\theta(x)} \partial_\mu \theta(x)$$

In order to restore the lost local invariance, we can introduce the **covariant derivative**, a particular derivative that undergoes the same phase transformation of the field:

$$\mathcal{D}_\mu \psi \longrightarrow \big(\mathcal{D}_\mu \psi\big) e^{iq\theta(x)}$$

and it's defined as follows:

$$\mathcal{D}_\mu \psi \equiv \partial_\mu + iqA_\mu$$

under the condition that the field A_μ transforms in the following way:

$$A_\mu \longrightarrow A'_\mu = A_\mu - \partial_\mu \theta(x)$$

By substituting the partial derivative with the covariant one in eq. (2.2.1), we get the following locally invariant lagrangian density:

$$\mathcal{L}_D = i\overline{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\overline{\psi}\psi$$

If we write explicitly the covariant derivative:

$$\mathcal{L}_D = i\overline{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\overline{\psi}\psi - q\overline{\psi}\gamma^\mu A_\mu \psi = l_0 - J^\mu A_\mu$$

where the latter term contains the interaction between the Dirac particle and the electromagnetic field. The quantity J_μ is interpreted as the charge current, i.e. the probability current of the particle times its charge.

To complete the lagrangian density of the QED, we have to add the following kinetic term for the field A_μ:

$$\mathcal{L}_\gamma = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

where $F_{\mu\nu}$ is the field strength tensor which can be written in terms of 4-vector electromagnetic field A_μ:

$$F^{\mu\nu} = -F^{\nu\mu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

l_γ describes the propagation of free photons and it is invariant for local gauge transformation. Adding l_D and l_γ, we get the complete QED lagrangian density:

$$\mathcal{L}_{QED} = \mathcal{L}_D + \mathcal{L}_\gamma = i\overline{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\overline{\psi}\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{2.2.2}$$

In every gauge theory, to each lagrangian density is possible to associate a Feynman diagram, that is an istrument to write and calculate the amplitude of every process involved in the theory. In this case, the interaction term is used to obtain the amplitude for all electromagnetic processes and in fig. all the fundamental transitions associated to the lagrangian density in eq. (2.2.2) can be drawn using the fundamental vertex shown in fig. 2.3.

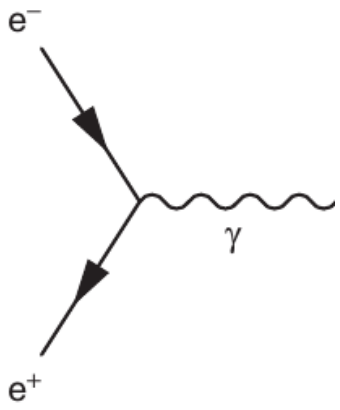


Figure 2.3: The fundamental Feynman vertex of QCD.

2.3 The Electroweak Theory

The Italian physicist Enrico Fermi developed a theory of the β decay that leads to a phenomenological description of the weak interactions. It is determined by the current-current lagrangian density, known as Fermi-like interaction

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}} J^{\dagger\mu}(x) J_\mu(x)$$

where G is the Fermi constant

$$G = 1.16638 \times 10^{-5}\ \text{GeV}^{-2}$$

and J_μ is the weak current, analogue of the electromagnetic one, and is the sum of a weak leptonic current $l_\mu(x)$ and a weak hadronic one $h_\mu(x)$.

These two current have different structure. l^μ, in fact, is a pure vectorial current:

$$l^\mu(x) = \overline{\psi}_l(1-\gamma^5)\gamma^\mu\psi_{\nu_l}$$

where ψ_l is the spinor of the charged lepton of the considered generation, ψ_{ν_l} is the spinor of the corresponding neutrino. h_μ, instead, has the structure:

$$h^\mu(x) = \overline{\psi}_p(G_V - G_A\gamma^5)\gamma^\mu\psi_n$$

where the subscripts p, n are for the members of a hadronic weak isospin doublet. G_V and G_A weights the vectorial and axial parts of the hadronic current, and they depends from the nature of the considered hadrons.

This interpretation, although it manages to describe successfully the short-range interaction approximation, leads to divergences, which manifest themselves in the violation of the unitarity limit of the Fermi-like cross-section, due mainly to the dimensionality of Fermi's constant. The divergent behaviour of the cross section can be avoided by improving the analogy with the electromagnetic interaction, i.e. introducing an adimensional coupling constant and one or more intermediate vector bosons, analogous to the photon, for the weak interaction. In this way, these bosons should be the mediators of the weak interaction. The Electroweak theory, thus, is the natural development of Fermi's theory and has been proposed by S. Glashow, A. Salam and S. Weinberg in 60's and 70's [30, 50, 64]. It is also known as the GWS model of the weak interactions, named after its creators.

2.3.1 The GWS Model

The GWS Model is a quantum field theory based on the symmetry group SU$(2)_L$, where the subscript L means that only the left-handed chiral components of the fields can take part in the weak interactions. Taking account of the analogy with the QED, the lagrangian density of this theory must be locally invariant with respect to transformations of the SU$(2)_L$ group. The generators of this symmetry are the Pauli 2×2 matrices σ_i $(i = 1, 2, 3)$. Rembering that SU$(2)_L$ is also the symmetry group for the rotations of the half-spin particles, we can adopt the corresponding formalism of the angular momentum. In this way, the particles, seen as eigenstates of the weak interaction, are arranged in six doublets of weak isospin. By naming I the weak isospin and I_3 its observed component on an axis of the isospin space, the three doublets for the leptons are:

$$I = \frac{1}{2} \qquad \begin{matrix} I_3 = +1/2 \\ I_3 = -1/2 \end{matrix} \qquad \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$$

and for quarks are:

$$I = \frac{1}{2} \qquad \begin{matrix} I_3 = +1/2 \\ I_3 = -1/2 \end{matrix} \qquad \begin{pmatrix} u \\ d' \end{pmatrix}_L \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L$$

For quark pairs, d', s', b' are the weak interaction eigenstates, that can be obtained as linear combination of the strong interaction eigenstates (that are also mass eigenstates) d, s, b and, thus, the mixing of different flavours is given by:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

where V is the complex unitary matrix named **Cabibbo-Kobayashi-Maskawa (CKM) matrix**:

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

In the next sections is reported the derivation of the CKM matrix, with a focus on its parametrization and its most significant properties. As for the QED development, also in the GSW model a local gauge transformation invariance is required, as mentioned before, and, thus, the action of the SU(2)$_L$ on the weak isospin doublets leads to:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L^{'} = e^{-i\vec{\alpha}(x)\cdot\vec{\tau}} \begin{pmatrix} \nu_e \\ e \end{pmatrix}$$

where $\vec{\tau}$ are the Pauli matrices divided by 2 and $\vec{\alpha}(x)$ is the vector of real parameters of the transformation that, in addiction, depends on the spacetime coordinates.

Although only the left-handed components can take part to the weak charged-current processes, the right-handed ones of charged fermions can take part to the neutral current weak process. The right-handed components are singlet of SU(2)$_L$:

$$I = 0 \qquad e_R^-, \mu_R^-, \tau_R^-, d_R, u_R, s_R, c_R, b_R, t_R$$

The request of local invariance under the SU$(2)_L$ group leads to the introduction of an isospin triplet of Yang-Mills fields: $W_{(i)}^{\mu}$, with $i = (1, 2, 3)$. These gauge fields are not yet the physics boson, known as $W^{\pm}$ and Z, but the latter can be obtained by combination of the $W_{(i)}^{\mu}$ with another gauge field, the isospin singlet B_{μ}, that is the boson field associated to the additional local gauge symmetry with respect to the Abelian group U(1)$_Y$. Thanks to this combination, it's possible incorporate the electromagnetic interaction in the weak one. The quantum number associated to the new group is the weak hypercharge Y, defined by:

$$Q = Y + \frac{I_3}{2} \tag{2.3.1}$$

which gives the electric charge Q, in units of e, of the I_3 member of a weak isospin multiplet, assigned a weak hypercharge Y. The definition (2.3.1) was proposed by Glashow and is an extension of the Gell$-$Mann-Nishijima relation for charges, thought originally for the flavour symmetry group SU$(2)_f$. The symmetry group of transformation, finally, is SU$(2)_L \otimes$ U$(1)_Y$ and leads to four gauge fields; three of them, as said before, coming from the SU$(2)_L$ and the other one coming from the group U$(1)_Y$.

The procedure to obtain the Electroweak lagrangian density is analogous to that followed for the electromagnetic case. The requirement of gauge local invariance under the SU$(2)_L \otimes$ U$(1)_Y$, indeed, leads to the introduction of the covariant derivative:

$$\mathcal{D}_{\mu} = \partial_{\mu} + ig\frac{\vec{\tau}}{2}W_{\mu} + ig'\frac{Y}{2}B_{\mu} \tag{2.3.2}$$

where g and g' are the two coupling constants for the two interactions. Neglecting for now the mass term and introducing the (2.3.2), the electroweak lagrangian density for fermions can be written as follows:

$$\mathcal{L}_{fermions} = \sum_f \overline{\psi}_f \gamma^{\mu} \mathcal{D}_{\mu} \psi_f$$

We can also add the kinetic term for the gauge fields:

$$\mathcal{L}_{gauge} = -\frac{1}{4} W^{\mu\nu}_{(i)} W^{(i)}_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}$$

where

$$\begin{aligned} W^{\mu\nu}_{(i)} =& \partial^\mu W^\nu_{(i)} - \partial^\nu W^\mu_{(i)} \\ B^{\mu\nu} =& \partial^\mu B^\nu - \partial^\nu B^\mu \end{aligned}$$

Finally, we can write down the complete electroweak lagrangian density:

$$\begin{aligned} \mathcal{L}_{EW} =& - i\psi_L \gamma^\mu \left(\partial_\mu + ig \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu + ig' \frac{Y}{2} B_\mu \right) \psi_L + \\ & - i\psi_R \gamma^\mu \left(\partial_\mu + ig' \frac{Y}{2} B_\mu \right) \psi_R + \\ & - \frac{1}{4} W^{\mu\nu}_i W^i_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} + \\ & + \frac{1}{2} g \varepsilon_{ijk} W^{\mu\nu}_i W_{j\mu} W^{k\nu} + \frac{1}{4} g^2 \varepsilon_{ijk} \varepsilon_{imn} W_{j\mu} W_{k\nu} W^\mu_m W^\nu_n \end{aligned} \tag{2.3.3}$$

where ψ_L and ψ_R are the left and right-handed chiral components of the particle, and the term in the last line describes the three and four-point self interactions of the vector bosons, arose because of the non-Abelian nature of the SU$(2)_L$ group. The four gauge fields can be combined to produce the physical vector fields for the $W^\pm, Z$ bosons and the photon:

$$\begin{aligned} W^\pm_\mu =& \frac{1}{\sqrt{2}} (W^1_\mu \mp i W^2_\mu) \\ \begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix} =& \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^\mu \\ B^\mu \end{pmatrix} \end{aligned}$$

with

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \qquad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}$$

The parameter θ_W W has to be determined experimentally, and it's called **Weinberg angle** or **weak mixing angle**. The electromagnetic charge therefore values:

$$q = g' \cos\theta_W = g \sin\theta_W$$

To conclude, the two vector bosons $W^\pm$ are electrically charged and can induce transitions between the members of the weak isospin doublets. The third gauge boson Z of the triplet is electrically neutral and is the source of the neutral currents.

2.3.2 Spontaneous symmetry breaking and Higgs mechanism

We can see the gauge fields so far are all massless, as the introduction of a mass term ad hoc, like:

$$\frac{1}{2}m^2 B_\mu B^\mu$$

in eq. (2.3.3) would break the local gauge symmetries. To make mass terms of the gauge fields invariant with respect to these tranformations, the simplest and most elegant way is the **spontaneous symmetry breaking (SSB)** and the **Higgs mechanism**. The key element of the latter mechanism is the Higgs boson field, that is a doublet of complex scalar fields that can be written as follows:

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

It has, thus, the the form of $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ multiplet, in order to ensure that the lagrangian invariance is conserved. The complex scalar field ϕ^+ destroys positive charged particles and creates negative particles, while ϕ^0 destroys neutral particles and creates neutral antiparticles. The lagrangian for the ϕ field is:

$$\begin{aligned} \mathcal{L}_H =& (\mathcal{D}^\mu \phi)^\dagger \mathcal{D}_\mu \phi - V(\phi) = \\ =& (\mathcal{D}^\mu \phi)^\dagger \mathcal{D}_\mu \phi - \frac{1}{2}\mu^2 \phi^\dagger \phi - \frac{1}{4}\lambda \big(\phi^\dagger \phi\big)^2 \end{aligned} \tag{2.3.4}$$

where $V(\phi)$ is the potential responsible of the symmetry breaking, and the parameter λ is assumed to be positive. In order to get the ground state ϕ_0, we have to minimise the potential. For $\mu^2 > 0$ the potential V assumes a unique minimum at $\phi = 0$ and, consequently, the ground state is symmetric under $\mathrm{SU}(2)_L$. On the other hand, for $\mu^2 < 0$ the shape of the potential is modified, as can be seen in the fig. 2.4, and V assumes a non-trivial minimum:

$$|\phi_0|^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}$$

The vacuum expectation value, defined as the absolute value of the field at the minimum of the potential, is thus non-zero and corresponds to the radius of a circumference in the complex plane $Re(\phi) - Im(\phi)$. Without any loss of generality, a reference minimum can be chosen among all possible ground states:

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

Adding to the lagrangian density of the gauge field sector the term in eq. (2.3.4), when the covariant derivative acts, one has:

$$\mathcal{L}_H = \big(\mathcal{D}^\mu \phi\big)^\dagger \mathcal{D}_\mu \phi - \frac{1}{2}\mu^2 \phi^\dagger \phi - \frac{\lambda}{4}(\phi^\dagger \phi)^2 - \frac{1}{4}F^{\mu\nu} F_{\mu\nu} - \frac{1}{4}B^{\mu\nu} B_{\mu\nu} \tag{2.3.5}$$

with

$$\mathcal{D}^{\mu}\phi = \left(\partial^{\mu} + ig\frac{\vec{\tau}}{2}\cdot\vec{W}^{\mu} + ig'\frac{Y}{2}B^{\mu}\right)\phi$$
$$F^{\mu\nu} = \partial^{\mu}W^{\nu} - \partial^{\nu}W^{\mu} - g\vec{W}^{\mu}\times\vec{W}^{\nu}$$
$$B^{\mu\nu} = \partial^{\mu}B^{\nu} - \partial^{\nu}B^{\mu}$$

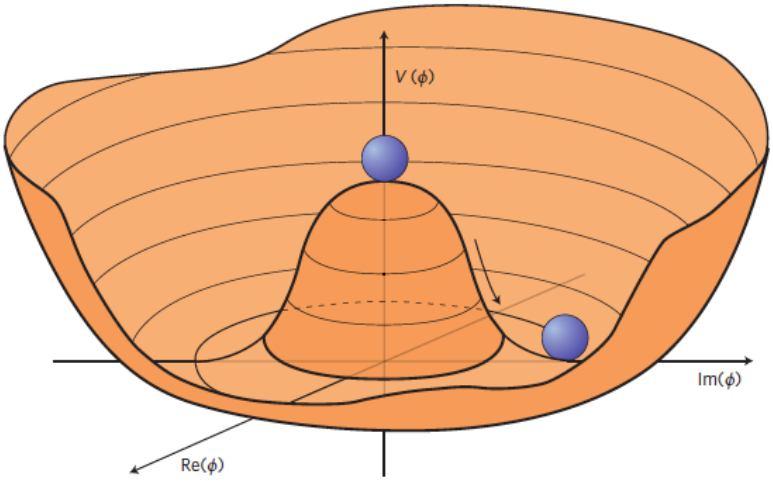


Figure 2.4: The shape of $V(\phi)$ when $\mu^2 < 0$.

Now we can make the ϕ field to fluctuate its vacuum, and we can parametrise the value assumed by the Higgs field in this state in the following way:

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix} \tag{2.3.6}$$

where:

$$v = \sqrt{-\frac{\mu^2}{\lambda}}$$

By substituting (2.3.6) in (2.3.5), we get as follows:

$$\begin{aligned}\mathcal{L}_{G\phi} =& \frac{1}{2}\partial_{\mu}H\partial^{\mu}H - \mu^2H^2 + \\ &- \frac{1}{4}\left(\partial_{\mu}W^1_{\nu} - \partial_{\nu}W^1_{\mu}\right)\left(\partial^{\mu}W^{1\nu} - \partial^{\nu}W^{1\mu}\right) + \frac{1}{8}g^2v^2W^1_{\nu}W^{1\nu} \\ &- \frac{1}{4}\left(\partial_{\mu}W^2_{\nu} - \partial_{\nu}W^2_{\mu}\right)\left(\partial^{\mu}W^{2\nu} - \partial^{\nu}W^{2\mu}\right) + \frac{1}{8}g^2v^2W^2_{\nu}W^{2\nu} \\ &- \frac{1}{4}\left(\partial_{\mu}Z_{\nu} - \partial_{\nu}Z_{\mu}\right)\left(\partial^{\mu}Z^{\nu} - \partial^{\nu}Z^{\mu}\right) + \frac{1}{8}(g^2 + g'^2v^2Z_{\nu}Z^{\nu} \\ &- \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\end{aligned} \tag{2.3.7}$$

The first line of (2.3.7) is the lagrangian of a massive scalar field, the Higgs one, with mass equals to $\sqrt{2\mu}$. The next two lines show that the components W^{μ}_1

and W_2^μ of the triplet $\vec{W}^\mu$ acquire mass:

$$M_1 = M_2 = \frac{1}{2}gv \equiv M_W$$

The fourth line shows that the field Z^μ acquires the following mass:

$$M_Z \equiv \frac{1}{2}v\sqrt{g^2 + g'^2} = \frac{M_W}{\cos\theta_W}$$

and the last line shows that the field A^μ has a null mass:

$$M_A = 0$$

2.3.3 How leptons acquire masses

The spontaneous breaking of the SU(2)$_L \otimes$ U$(1)_Y$ gauge symmetry can also generate the masses of the fermions. Anyway, the procedure is different from the one used for the bosons case, as the fermion mass term $-m\psi\overline{\psi}$ is not invariant under the SU$(2)_L \otimes$ U$(1)_Y$ group, because of the different transformation of the right and left-handed chiral components of the fields. In a theory where the symmetry is spontaneously broken, however, introducing in the lagrangian density an explicit mass term is not necessary, but we can be obtain the fermionic mass term by coupling their fields to a scalar field that acquires a vacuum expectation value. This can be achieved by introducing a Yukawa coupling between the fermion field and the Higgs field:

$$\mathcal{L}_Y = g_f\left(\overline{\psi}_L\phi\psi_R - \overline{\psi}_R\phi^\dagger\psi_L\right) \tag{2.3.8}$$

where g_f is the Yukawa coupling constant. Substituting (2.3.6) in (2.3.8), we obtain:

$$\begin{aligned}\mathcal{L}_Y =& \frac{g_f}{\sqrt{2}}\left[\begin{pmatrix}\nu_l & l\end{pmatrix}\begin{pmatrix}0\\ v+H\end{pmatrix}l_R + l_R\begin{pmatrix}0 & v+H\end{pmatrix}\begin{pmatrix}\nu_l\\ l\end{pmatrix}\right]\\ =& \frac{g_f}{\sqrt{2}}(v+H)(l_Ll_R + l_Rl_L)\end{aligned} \tag{2.3.9}$$

The second line of (2.3.9) is a Dirac-like mass term and, because of that, allows us to identify the constant coefficient of $(l_L + l_R + l_R + l_L)$ with the mass term for leptons:

$$m_f = \frac{v}{\sqrt{2}}g_f$$

This is the least satisfactory part of the model because, even if this kind of Yukawa coupling solves the problem of leptons' masses, it does not arise from a gauge principle. It is, in fact, purely phenomenological and needs a specific coupling constant for each fermion- Higgs interaction. Moreover, the couplings are very different given the wide range of fermion masses experimentally observed.

2.3.4 Masses of quarks

The same mechanism adopted to give mass to leptons can be used also for the quarks:

$$\begin{aligned}\mathcal{L}_Y =& \frac{1}{\sqrt{2}}\left[g^d_{i,j}\begin{pmatrix}u_{i,L} & d_{i,L}\end{pmatrix}\begin{pmatrix}0\\ v+H\end{pmatrix}d_{j,R} + g^u_{i,j}\begin{pmatrix}u_{i,L} & d_{i,L}\end{pmatrix}\begin{pmatrix}-(v+H)^*\\ 0\end{pmatrix}u_{j,R} + h.c.\right]\\ =& \frac{1}{\sqrt{2}}(v+H)\left[g^u_{ij}(u_{i,L}u_{j,R}+u_{j,R}u_{i,L}) + g^d_{ij}(d_{i,L}d_{j,R}+d_{j,R}d_{i,L}+h.c.\right]\end{aligned}$$

with $u_i = (u, c, t)$ and $d_i = (d, s, b)$. The matrix of mass terms is not diagonal:

$$m^u_{ij} = -\frac{v}{\sqrt{2}}g^u_{ij} \qquad m^d_{ij} = -\frac{v}{\sqrt{2}}g^d_{ij}$$

We can, nevertheless, diagonalise these matrices with four different transformations on the family triplets $u_{i,L}, u_{i,R}, d_{i,L}, d_{i,R}$ in the following way:

$$u_{\alpha,L} = \left(\mathcal{U}^u_L\right)_{\alpha i}u_{i,L} \qquad u_{\alpha,R} = \left(\mathcal{U}^u_R\right)_{\alpha i}u_{i,R} \qquad d_{\alpha,L} = \left(\mathcal{U}^d_L\right)_{\alpha i}d_{i,L} \qquad d_{\alpha,R} = \left(\mathcal{U}^d_R\right)_{\alpha i}d_{i,R} \tag{2.3.10}$$

where α is the index in the mass diagonal basis and i is the index in the non-diagonal weak interaction basis. All these matrices are unitary. In the diagonalised basis, we can write:

$$\mathcal{L}_Y = \frac{1}{\sqrt{2}}(v+H)\left[m^u u\overline{u} + m^d d\overline{d} + m^c c\overline{c} + m^s s\overline{s} + m^t t\overline{t} + m^b b\overline{b}\right]$$

We must apply the same transformations to the interacting term, even if it is invariant under the $SU(2)_L \otimes \mathrm{U}(1)_Y$ symmetry, as it still contains the eigenkets of the weak interaction. When this operation is worked out, the term of the coupling with the Z boson, i.e. neutral current coupling term, is diagonal also in the mass basi, because the transformations of eq. (2.3.10) are unitary; the term of the coupling with the W boson, instead, i.e. charged current coupling term, is written in this basis:

$$\begin{aligned}\mathcal{L}_{CC} &= -\frac{g}{\sqrt{2}}\left(\overline{u}_{i,L}\overline{d}_{i,L}\right)\gamma^\mu\tau_+ W^+_\mu\left(u_{L,i}d_{L,i}\right) + h.c.\\ &= -\frac{g}{\sqrt{2}}\overline{u}_{i,L}\gamma^\mu d_{L,i}W^+_\mu + h.c.\\ &= -\frac{g}{\sqrt{2}}\overline{u}_{\alpha,L}\left[\left(\mathcal{U}^u_L\right)_{\alpha i}\left(\mathcal{U}^d_L\right)^\dagger_{\beta i}\right]\gamma^\mu d_{L,\beta}W^+_\mu + h.c.\end{aligned}$$

where the matrix:

$$V_{\alpha\beta} = \left[\mathcal{U}^u_L\mathcal{U}^{d\dagger}_L\right]_{\alpha\beta}$$

is unitary, but not diagonal. V is the **CKM matrix** and it is a 3×3 unitary matrix and display the mismatch between the weak eigenstates and the mass eigenstates. It is also responsible of the transitions between quark generations through flavour changing interactions. The definition of the CKM matrix implies that it is defined by three real parameters and one unavoidable phase factor, with the latter that

leads to CP violation. The charged current lagrangian density can be ultimately written as:

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}\overline{u}_{L,\alpha}\gamma^{\mu}V_{\alpha\beta}d_{L,\beta}W_{\mu}^{+} - \frac{g}{\sqrt{2}}\overline{d}_{L,\alpha}\gamma^{\mu}V_{\alpha\beta}^{\dagger}d_{L,\beta}W_{\mu}^{-}$$

Analogously a mixing matrix can be introduced also for the neutrino sector, namely the PMNS matrix, named after Pontecorvo, Maki, Nakagawa e Sakata.

2.4 Quantum Chromodynamics

The **Quantum Chromodynamics**, QCD, is the gauge theory of strong interactions. It is analogous to the QED, but, in this case, the symmetry group of the theory is symmetry group SU$(3)_C$, where the subscript C stands for the charge associated with this symmetry, named **colour**. In the strong interactions, the colour is identified with the strong charge and, so, as the source of the cromodynamic field. An important feature of SU$(3)_C$ is that it isn't an Abelian group, and the consequence of this fact is that the generators of the symmetry group do not commute. Because of that, we have to introduce in the QCD lagrangian density of interaction terms among the gauge fields, called gluons, as they bring the charge of the group, the colour. In QED, instead, photons do not have electrical charge and therefore a self-interaction terms do not exist in the QED. For what we said so far, the QCD is invariant under local gauge transformations of SU$(3)_C$ group, i.e.:

$$\psi \longrightarrow \psi' = e^{ig_s\vec{\alpha}(x)\cdot\vec{T}}\psi$$

where g_s is the strong coupling constant, $\vec{\alpha}(x)$ are eight functions of the space-time coordinates and $\vec{T} = T^{\alpha}$, with $\alpha = (1, \ldots, 8)$ are generators of the SU$(3)_C$ group. The generators are related to the Gell-Mann matrices:

$$T^{\alpha} = \frac{1}{2}\lambda^{\alpha}$$

and follow the commutation rules:

$$[T_{\alpha}, T_{\beta}] = if_{\alpha\beta\gamma}T_{\gamma}$$

where $f_{\alpha\beta\gamma}$ are the structure constants of the group SU$(3)_C$, with the indices running from 1 to 8. Since the generators of SU$(3)_C$ are represented by 3×3 matrices, a new degree of freedom is needed, the colour, whose values parametrise the group. In this way, the field ψ has three possible states labelled as red, green, and blue. The lagrangian density of free quarks, neglecting the mass terms, is:

$$\mathcal{L} = \sum_{f=1}^{6}\overline{\psi}^{f}i\gamma^{\mu}\partial_{\mu}\psi^{f}$$

The local gauge invariance under SU$(3)_C$ introduces 8 massless fields of gauge, called gluons, and the covariant derivative, $\mathcal{D}_{\mu}$, given by:

$$\mathcal{D}_{\mu} = \partial_{\mu} + ig_sT_{\alpha}G_{\mu}^{\alpha}$$

where G_μ^α are the 8 gluons fields which transforms as follows:

$$G_\mu^\alpha \longrightarrow G'^\alpha_\mu + ig_s f^{\alpha\beta\gamma}\theta_b(x)G_{\gamma,\mu}$$

Adding the contribution of the kinetic energy for each gluon, we obtain the complete lagrangian density for the QCD:

$$\mathcal{L}_{QCD} = \overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi - ig_s\overline{\psi}\gamma^\mu\lambda_a\psi G_\mu^\alpha - \frac{1}{4}G_\alpha^{\mu\nu}G_{\mu\nu}^\alpha \tag{2.4.1}$$

with $G_\alpha^{\mu\nu}$ the strong tensor field, defined as:

$$G_\alpha^{\mu\nu} = \partial^\mu G_\alpha^\nu - \partial^\nu G_\alpha^\mu - g_s f_{\alpha\beta\gamma}G^{\beta,\mu}G^{\gamma,\nu}$$

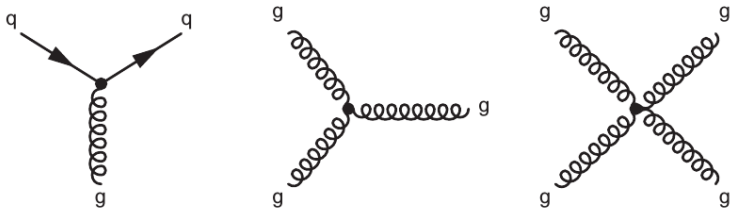


Figure 2.5: The fundamental vertices of the QCD.

As for the other gauge theories, it is possible to associate each lagrangian interaction term with a Feynman diagram. In fig. 2.5 are shown all the fundamental transitions associated to the lagrangian in (2.4.1).

Two relevant properties, which stem from experimental evidence and are described in SM by the strong interactions sector, significantly differentiate the QCD from QED: colour confinement and asymptotic freedom. The asymptotic freedom has to cope with the experimental fact that no coloured hadrons are observed. Hadrons, therefore, are colour singlets, since they are interpreted as bound states of quarks in the QCD parton model. This imposes restrictions on the types of bound quark state configurations that can exist. All this can be summarised by saying that the quark colour degree of freedom must be confined.

The colour confinement can be explained by looking at the running coupling constant form of strong force:

$$\alpha_s(|q^2|) = \frac{\alpha_s(\mu^2)}{\left[1 + \alpha_s(\mu^2)\frac{33-2N_f}{12\pi}\ln\frac{|q^2|}{\mu^2}\right]}$$

where:

- q^2, the transferred 4-momentum;
- μ, scale parameter for the strenght of the coupling;
- N_f, number of fermions capable of strong interactions at the scale considered.

We can see that $\alpha_s(|q^2|)$ decreases as $|q^2|$ increases. For $|q| \sim 200$ MeV the value of α_s is large enough that any perturbative approach cannot be applied. In this region of the energy spectrum, the calculations are carried on with the QCD-lattice approach. For increasing values of $|q^2|$, $\alpha_s(|q^2|)$ decreases and we move towards a regime in which perturbative approach is a good approximation.

2.5 Unsolved questions in the Standard Model

In the last 30 years experiments have tested the Standard Model theory in many ways. Every predicted particle has been found, features of the electromagnetic and strong interactions have been described with accuracy and, in many cases, as for instance for the case of the electron gyromagnetic ratio [44], predictions of the SM were verified with a precision up to 12 orders of magnitude. The SM model, however, fails in giving explanation of other phenomena and observations in nature and, because of that, it isn't a comprehensive and fully satisfactory theory of the known universe. Some of the most important questions that do not find satisfactory answer within the SM are:

- **Electroweak unification**
  The GWS model cannot be considered a real unification theory, as the symmetry group is the product of two different groups, each one with its own constant, g and g', not linked by the theory. The ratio:

  $$\frac{g'}{g} = \tan\theta_W$$

  has to be determined experimentally.

- **Large set of free parameters**
  In the model there are many parameters that are not postulated by theory and, then, must therefore be obtained via measurements:

  - 3 coupling costants, g, g', g_s;
  - fermion masses (or the corresponding Yukawa coupling with the Higgs field);
  - mass of Higgs boson;
  - CKM matrix elements;
  - PMNS matrix elements.

- **Fermion masses**
  Since there is no explanation or prediction of fermion masses, a problem arise in this sector because of the differences of magnitudes, till 5 order of magnitudes between the top quark and the electron, as we can see in fig. 2.6.

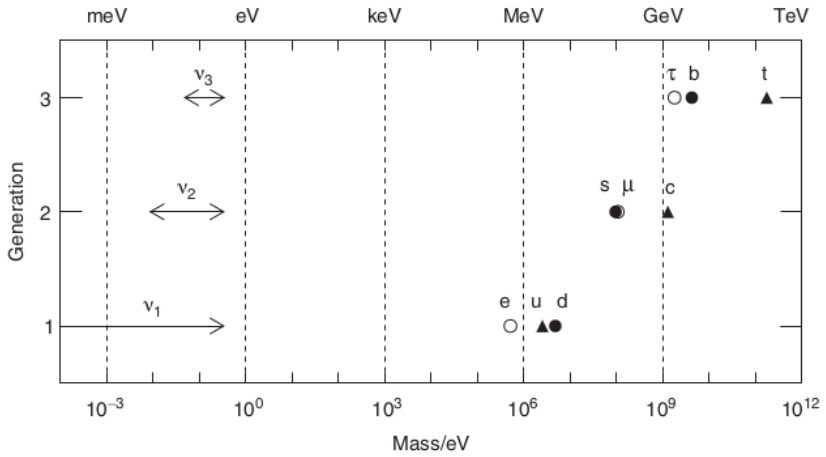


Figure 2.6: The hierarchy of the mass of the fermions.

- **Hierarchy problem**
  Also known as naturalness, this problem is related to the radiative corrections which the Higgs mass receives through boson and fermionic loops, as can be seen in the fig. 2.7. Between these corrections, self-interaction terms arise, such as:

$$m_H^2 \approx M_{H,0}^2 + \frac{\lambda}{4\pi^2}\Lambda^2$$

  where $M_{H,0}$ is the bare mass of the Higgs, λ is the strength of the coupling and Λ^2 is the energy scale of possible new physics (from whose value the latter has to be compulsory considered). Since the Higgs boson has to couple to every massive particle, one can recalculate the one-loop corrections to Higgs mass as:

$$m_H^2 \approx M_{H,0}^2 + \frac{g_f}{4\pi^2}\left(\Lambda^2 + m_f^2\right) + \frac{g_s}{4\pi^2}\left(\Lambda^2 + m_S^2\right)$$

  with g_f and g_S the coupling constant of fermion and scalar particles to the Higgs and m_f and m_S the related mass terms. These corrections are quadratically divergent with the cut-off Λ. Usually, a typical cut-off scale is the Planck scale, where the corrections are $\sim 10^{30}\ m_H^2$, many orders of magnitude larger than the Higgs mass at tree level. A precise tuning between fermions and scalars, known as **fine tuning**, is required in order to reduce or cancel this divergences. This balance should also keep into account the tight constraints on the Higgs mass, as shown in fig. 2.8.

- **Flavour Changing Neutral Currents (FCNC)**
  Another non-predicted observation by the SM is the presence of just three families of quarks and leptons, and each family can be identified by the flavour quantum number. The suppression of flavour changing neutral currents at tree level, as expected from the Glashow, Iliopoulos and Maiani mechanism, GIM, is also entered by hand in the full theory, without explanations from first principles.

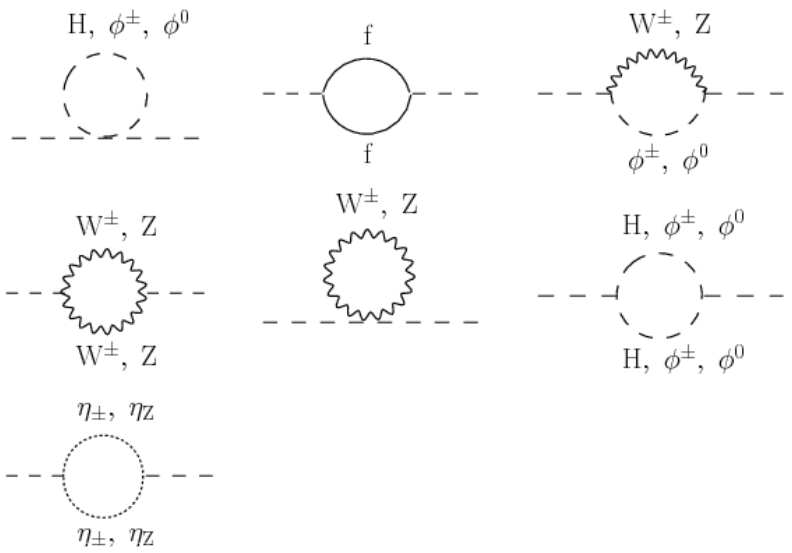


Figure 2.7: One-loop self-energy corrections to the Higgs mass.

- **Symmetries and conservations**
  Some experimental facts have to be inserted ad hoc in the theory, such as electrical charge quantization, the proton stability and the conservation of the baryonic number, the latter does not come from a symmetry. On the other hand, there are other charges, such as the coloured oand the electrical ones, that are related to exact symmetries, as they come from U$(1)_Y$ and SU$(3)_C$ groups.

- **Neutrino sector**
  SM assumes neutrinos to be massless, while observations of flavour oscillations can only be explained if neutrinos have masses, via a mixing of the electroweak eigenstates. The seesaw mechanism, indeed, incorporates neutrino masses into the SM by introducing heavy Majorana neutrinos, whose masses are inversely coupled to the light SM neutrino masses, hence motivating their small values of $O(1$ eV). Despite extensive searches, however, the experimental proof is unluckly still missing.

- **Cosmological problem**
  SM in not able, trough the SM CP-violation in the quark sector, to justify the actual matter-antimatter unbalance measured in the Universe, neither to predict and include some results from astroparticle physics.

- **Gravity**
  One of the most striking shortcoming of the Standard Model is that it lacks a description for gravity. Gravity was the first force to be fully understood over large distances, but it will likely be the last to be fully understood at very short distances. This is because the coupling strength of gravity is very weak if compared to other interactions, with a coupling constant that is 10^{34} times smaller than the electromagnetic coupling α. Gravitational effects would not be observable in particle collisions below a centre of mass energy close to the Planck scale $(10^{19}$ GeV), very larger than the energy limits of current particle colliders. Most theoretical models that could describe the

Standard Model and gravity, sometimes referred to as theories of everything or ultraviolet completion models, manifest new phenomena only above a large energy called Λ_{UV}, roughly in the vicinity of the Planck scale.

- **Dark matter and dark energy**
  Astronomical evidences show the Universe is made up of only for the 5% of ordinary matter, while the rest does not correspond to the known matter. Cosmological observation of galaxy rotation profiles, indeed, provides evidence for a large amount of non SM-matter, that very weakly interact with SM particles. This matter is referred to with the name of **dark matter**, and should represent the 24% of the Universe. The remaining 71% is ascribed to a constant vacuum energy, called **dark energy**. The existence of the dark energy would account for the accelerating expansion of the Universe.

In the next chapter, we are going to discuss about the possible scenarios that we may encounter when we try to push our knowledge about physics beyond the Standard Model of Particles.

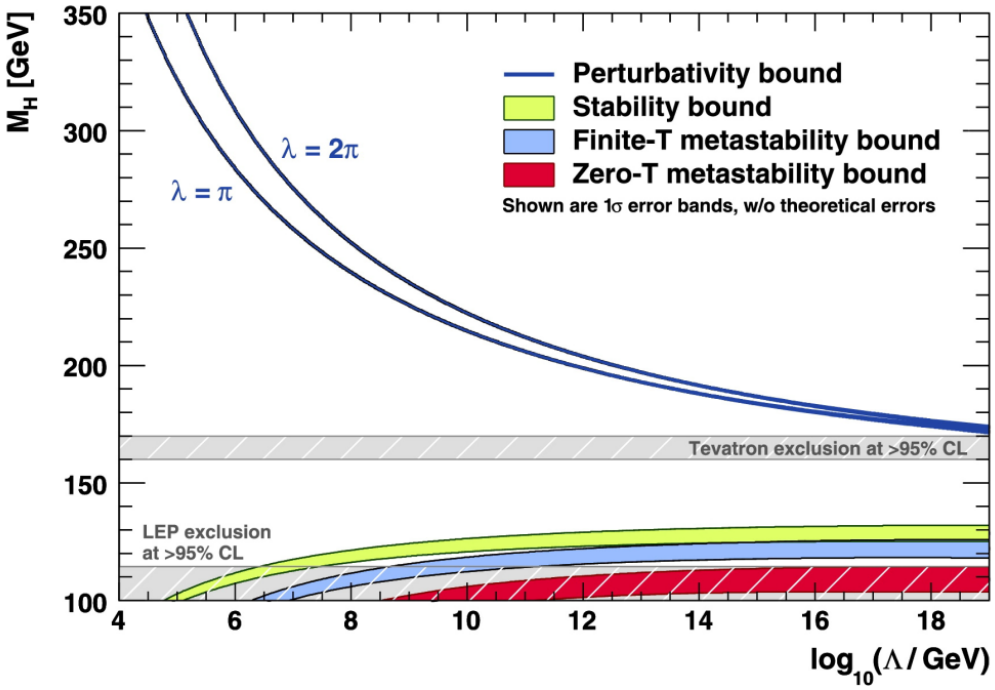


Figure 2.8: The scale Λ at which the two-loop renormalisation-group equations (RGEs) drive the quartic SM Higgs coupling non-perturbative, and the scale Λ at which the RGEs create an instability in the electroweak vacuum $(\lambda < 0)$. The width of the bands indicates the errors induced by the uncertainties. The perturbativity upper bound (sometimes referred to as "triviality" bound) is given for $\lambda = \pi$ (the blue lower bold line) and $\lambda = 2\pi$ (the blue upper bold line). Their difference indicates the size of the theoretical uncertainty in this bound. The absolute vacuum stability bound is displayed by the light shaded in green band, while the less restrictive finite-temperature and zero-temperature metastability bounds are medium, in blue, and dark shaded, the red one, respectively. The grey hatched areas indicate the LEP and Tevatron exclusion domains.

Chapter 3

Beyond SM physics and W' boson

As we have seen before, there are many problems that are not solved in the SM. Considering the experimental results obtained up to now, we can think that SM is only a very good approximation of a more general theory, not known yet, that also incorporates naturally th gravitational interaction. Much thought has been put into how the Standard Model might be modified to understand these puzzles. The vast majority of the possible answers to the questions above fall into one or more of three broad classes of proposals:

1. We can consider the known fundamental fields but introducing new interactions. This road leads to Great Unification, Supersymmetry, String Theory.

2. We can, instead, consider new fundamental fields with new interactions; in this class of theories we can find Compositeness, i.e. that some of the known particles might be composites of still-smaller things, condensed fermion-anti-dermal, Technicolour, Extended Technicolour, Preons.

3. Finally, we can consider the existence of large extra dimensions. This means that there could be more than the standard spacial three dimensions, that is a likely possibility in string theory, and that some of these are large enough to be seen in high-energy accelerators.

Many theories that extend the SM, which we, from now, we briefly refer to as BSM (i.e. Beyond Standard Model), predict additional charged gauge bosons, often referred to as W' bosons. In the following, we are going to discuss about the main theories this hypotetical boson can be considered in and, then, generally about the theoretical features of W' and the experimental searches for it performed so far.

3.1 Theories including W'

3.1.1 Extra Dimensions

First ideas about a spacetime with more than three spatial dimensions date back to the 1920s, mainly through the work of Kaluza and Klein [56]. Although their initial proposal did not find confirmation by experiments, the formalism that they

and others developed is still useful nowadays. Around 1980, within the context string theory it was proposed again to enlarge the number of space dimensions, this time as a requirement for describing a consistent theory of quantum gravity. The extra dimensions were supposed to be compactified at a scale close to the Planck scale, and thus not testable experimentally in the near future.

A different approach was given by Arkani-Hamed, Dimopoulos, and Dvali (briefly ADD). In 1998, they showed that the weakness of gravity could be explained by postulating two or more extra dimensions, in which only gravity could propagate. The size of these extra dimensions should range between roughly a millimeter and $\sim 1/$TeV, leading to possible observable consequences in current and future experiments. A year later, Randall and Sundrum (RS) found a new possibility using a warped geometry, postulating a five-dimensional Anti-de Sitter (AdS) spacetime[1], with a compactification scale of order TeV. The origin of the smallness of the electroweak scale versus the Planck scale was explained by the gravitational redshift factor, present in the warped AdS metric. As in the ADD model, originally only gravity was assumed to propagate in the extra dimensions, although it was soon clear that this was not necessary in warped extra-dimensions and, thus, also the SM gauge fields and SM fermions could propagate in the five-dimensional spacetime.

The physics of warped extra-dimensional models has an alternative interpretation by means of the AdS/CFT correspondence. Models with warped extra dimensions are related to four-dimensional strongly-interacting theories, allowing an understanding of the properties of five-dimensional fields as those of four-dimensional composite states. This approach has opened new directions for tackling outstanding questions in particle physics, such as the flavor problem, grand unification, and the origin of electroweak symmetry breaking or supersymmetry breaking.

Constraints on extra-dimensional models arise from astrophysical and cosmological considerations. In addition, tabletop experiments exploring gravity at sub-mm distances restrict certain models. Collider limits on extra-dimensional models are dominated by LHC results. Most of recent limits are published results based on LHC data collected in 2015-16 at a center-of-mass energy of 13 TeV, and legacy results from 20 fb^{-1} of 8 TeV data collected in Run 1. In addition, there are a few preliminary 13 TeV results, which can be found on the public internet pages of ATLAS and CMS. For most of the models, Run 2 results surpass the sensitivity of Run 1, even in the cases when the integrated luminosity is smaller.

3.1.2 Composite Higgs Model

[56] Within the SM, the Electroweak Symmetry Breaking, that gives is posited but has no dynamical origin. Furthermore, the Higgs boson appears to be unnaturally light. In the context of weakly coupled models of the Electroweak Symmetry Breaking we can also consider multiple Higgs SU(2)$_L$ doublets, as well as additional Higgs singlets, triplets or even more complicated multiplet structures, with or without low energy supersymmetry. In general, for such models one needs

[1] In mathematics and physics, n-dimensional anti-de Sitter space (briefly named as AdS$_n$) is a maximally symmetric Lorentzian manifold with constant negative scalar curvature.

to take into account experimental constraints from precision measurements and flavour changing neutral currents.

A scenario that remedies these two catches is to consider the Higgs boson as a bound state of new dynamics becoming strong around the weak scale. The idea that the Higgs boson itself could be a composite bound state emerging from a new strongly-coupled sector has regained recently some interest. The composite Higgs idea is an interesting incarnation of Electroweak Symmetry Breaking via strong dynamics that smoothly interpolates between the Technicolor approach and the true SM limit. To avoid the usual conflict with electroweak data, it is sufficient, if not necessary, that a mass gap separates the Higgs resonance from the other resonances of the strong sector. Such a mass gap can naturally follow from dynamics if the strongly-interacting sector exhibits a global symmetry, G, broken dynamically to a subgroup H at the energy scale f. These have to be such that the coset G/H contains, in addition to the three Nambu–Goldstone bosons of SO(4)/SO(3) that describe the longitudinal components of the massive W and Z, a fourth Nambu–Goldstone boson that can be identified with the physical Higgs boson. Simple examples of such a coset are SU(3)/SU(2) or SO(5)/SO(4), the latter being favoured, since it is invariant under the custodial symmetry. It is also possible to have non-minimal custodial cosets with extra Goldstone bosons leading to additional Higgs bosons in the spectrum,

The Higgs boson can be made significantly lighter than the other resonances of the strong sector if it appears as a pseudo-Nambu–Goldstone boson. The main prediction of the so-called Composite Higgs Model is, moreover, the existence of new particles with mass around a TeV, that are excitations of the composite Higgs. Within the most compelling scenarios each SM particle has a partner with equal quantum numbers but heavier mass. For example, the γ, W and Z bosons have heavy replicas with mass determined by the compositeness scale, expected around TeV. Composite Higgs Models, indeed, typically require a larger global symmetry of the underlying theory and, hence, additional relatively light scalar particles, extra electroweak vector bosons, e.g. an additional SU$(2)\times$U(1) gauge group), and vector-like partners of the top-quark of charge +2/3 and possibly also +5/3.

3.1.3 Little Higgs Model

The idea behind the Little Higgs models is to identify the Higgs doublet as a (pseudo) Nambu–Goldstone boson while keeping some sizable non-derivative interactions, in particular a largish Higgs quartic interaction. If we switch on some interactions that break explicitly the global symmetry, we can generate masses for the would-be massless Nambu–Goldstone bosons of the order of $g\Lambda_{G/H}/(4\pi)$, where g is the coupling of the symmetry breaking interaction and $\Lambda_{G/H} = 4\pi f_{G/H}$ is the dynamical scale of the global symmetry breaking G/H. In the case of the Higgs boson, the top Yukawa interaction or the gauge interactions themselves will certainly break explicitly (part of) the global symmetry since they act non-linearly on the Higgs boson [56].

In the SM the Higgs mass suffers from an instability under radiative corrections, but in the Little Higgs models the Higgs mass is not UV sensitive, thanks

to the their Electroweak Symmetry Breaking dynamic mechanism. Explicitly, the cancellation of the SM quadratic divergences is achieved by a set of new particles around the Fermi scale, such as gauge bosons, vector-like quarks, and extra massive scalars. These ones are related, by the original global symmetry, to the SM particles with the same spin. Contrary to supersymmetry, the cancellation of the quadratic divergences is achieved by same-spin particles. These new particles, with definite couplings to SM particles as dictated by the global symmetries of the theory, are perfect experimental goals for the LHC.

In addiction to what we've said so far, warped extra dimensions can give rise to scenarios, often called gauge-Higgs unified models, where the Higgs boson appears as the fifth component of a 5D gauge boson, A_5. The Higgs mass is protected by the 5D gauge invariance and can only get a nonzero value from non-local one-loop effects.

The motivation for Little Higgs models is to solve the little hierarchy problem, i.e., to push the need for new physics (responsible for the stability of the weak scale) up to around 10 TeV. Finally, the Little Higgs models are effective theories valid up to their cutoff scale $\Lambda_{G/H}$. Their UV completions could either be weakly or strongly coupled.

3.1.4 Left-Right symmetry

The spatial parity (P) is explicitly broken in the SM by the asymmetry between left- and right-handed multiplets. Because of that, the breaking of P is not at the same footing as the breaking of the gauge group SU$(2)_L\times$U$(1)_Y$, which is spontaneous in SM. This could be a starting points for extensions of the Standard Model itself [32, 52].

A recent proposal suggested that the observed $V-A$ structure of weak interactions may only be a low-energy phenomenon, which ought to disappear when we reach energies of the order of 10^3 GeV. In this contest, all interactions above these energies are supposed to be parity-conserving and, in addiction, are describable by a single gauge coupling constant g. A left-right symmetric model has been devised within the framework of unified gauge theories and, in this picture, it not only exhibits the above properties, but also provides a natural basis for the CP-violating interactions, observed, for example, in K decays. Another relevant feature of this model is the close link between the magnitude of CP violation and the departure from exact left-right symmetry observed in nature. Thanks to that, we may have a model for CP violation where its magnitude isn't entirely arbitrary.

Translating these ideas into a realistic model of weak and electromagnetic interactions, we get a renormalizable gauge model, based on the gauge group SU$(2)_L\times$SU$(2)_R\times$SU$(4')$, where SU$(4')$ unifies the lepton and the quark sectors. The unification of the left and right gauge couplings in this model can be achieved by demanding that the entire Lagrangian, except for Higgs boson mass terms, ha to be invariant under a discrete symmetry, that transforms SU$(2)_L$ to SU$(2)_R$.

The requirement that the Higgs boson mass term is not invariant under the latter transformation guarantees that the right-handed gauge bosons are heavier than the left-handed gauge bosons, as required by the non observation of the right-handed current interactions at energies we experimentally reached so

far. However, this requirement induces a not spontaneous parity breaking, but it is possible to obtain genuine spontaneous breakdown of left-right discrete symmetry in a dynamically way, without imposing the request about the Higgs mass terms from the beginning. All distinctions between left- and right-handed sectors, then, could arise purely as a result of spontaneous breakdown of the local symmetry.

3.1.5 Sequential Standard Model (SSM)

The Sequential Standard Model predicts the existence of a new massive W_{SSM} boson, with couplings to fermions that are identical to those of the SM W boson, i.e. with final states consisting either of a charged lepton and neutrino or a quark pair. It also assumes ν from W_{SSM} decay to be light and stable. This model represents a good benchmark, as the results can be interpreted in the context of other models of new physics, and is useful for comparing the sensitivity of different experiments.

3.1.6 The 331 Model

Some recent experiments of precision flavour physics showed that there are several anomalies in the present data, even if none of them are conclusive. We will discuss about some of these anomalies in the last section of this chapter. It is useful, so, to make up theories that investigate how the pattern of flavour violation comes up, and what results it can induces in the future experiments. Among the simplest extensions of the SM which pursue this goal are the so-called 331 models, based on the gauge group SU$(3)_C\times$SU$(3)_L\times$U$(1)_X$ (331). In this framework, the new sources of flavour and CP violation originate dominantly through the flavour violating interactions of ordinary quarks and leptons with a new heavy Z' gauge boson. Also one-loop contributions involving new charged gauge bosons and new heavy quarks with exotic electric charges can be relevant in certain processes [18].

In a 331 theory, the group SU$(3)_C\times$SU$(3)_L\times$U$(1)_X$ is spontaneously broken to the SM group SU$(2)_L\times$U$(1)_X$ and, subsequently, broken down to the electromagnetic group U$(1)_Q$. Therefore, the 331 model has an extended Higgs sector, with the first symmetry breaking occurring at a scale much larger than the electroweak scale. One of the interesting features of this model is that the requirement of anomaly cancelation, together with that of asymptotic freedom of QCD, constrains the number of generations to be equal to the number of colours, thus providing an explanation for the existence of 3 generations. The requirement of anomaly cancellation has also consequences on the transformation properties of fermions. In fact, as a first consequence of the extension of the SM gauge group SU$(2)_L$ to SU$(3)_L$, one has that left-handed fermions transform as triplets (or antitriplets) under the action of SU$(3)_L$. In order to have a flavour-anomaly-free theory, the number of triplets should be equal to the number of antitriplets. A possible choice is to assume that the three lepton generations transform as antitriplets, so that, taking into account the three colour possibilities for the quarks, the number of quark triplets should be equal to the number of antitriplets minus

one. Two quark generations, hence, should transform as triplets, one as an antitriplet. The choice of having the third generation with different transformation properties might be at the origin of the large mass of the top quark with respect to the other quarks. However, in contrast to the SM, where anomaly is canceled for each generation of fermions, in the 331 model the cancelation is fulfilled only when all the generations are considered.

The extension of the SM gauge group SU$(2)_L$ to SU$(3)_L$ also implies the existence of 5 new gauge bosons. These are a new neutral boson, Z', plus other four. The latter ones might be charged depending on the variant of the model that one chooses. They are generically denoted as $V^{\pm Q_V}$ and $Y^{\pm Q_Y}$. The new bosons $V^{\pm Q_V}$ and $Y_{\pm Q_Y}$ couple two SM leptons, while this is not possible for SM quarks, which can only be coupled to new heavy quarks by means of these gauge bosons.

3.2 W' boson

After we did an overview obout the most relevant Beyond Standard Model theories in which the existence of a new vector boson W' is predicted, we switch to discuss more profoundly about the W' itself.

[56] The W' boson is a massive hypothetical particle of spin 1 and electric charge equals to ± 1, which is a color singlet.

3.2.1 Coupling to the fermions and the bosons

The Lagrangian terms describing couplings of a W'^{+} boson to fermions are the following:

$$\frac{W'^{+}_{\mu}}{\sqrt{2}}\left[\overline{u}_i\Big(C^R_{q_{ij}}P_R + C^L_{q_{ij}P_L}\Big)\gamma^\mu d_j + \overline{\nu}_i\Big(C^R_{l_{ij}}P_R + C^L_{l_{ij}}P_L\Big)\gamma^\mu e_j\right] \tag{3.2.1}$$

and for the W'^{-} are the hermitian conjucate of the eq. (3.2.1). There, u, d, ν, e are the SM fermions in the mass eigenstate basis, $i, j = 1, 2, 3$ label the fermion generation and $P_{R,L}$ the projectors on the chirality eigenstates:

$$P_{R,L} = \frac{1 \pm \gamma_5}{2}$$

The coefficients $C^L_{q_{ij}}, C^R_{q_{ij}}, C^L_{l_{ij}}, C^R_{l_{ij}}$ are complex dimensionless parameters. If $C^R_{l_{ij}} \neq 0$, then the ith generation includes a right-handed neutrino. For the sake of clarity, in this notation the SM W couplings are $C^L_q = gV_{CKM}$, $C^L_l = g \approx 0.63$ and $C^R_q = C^R_l = 0$.

Unitarity considerations imply that the W' boson is associated with a spontaneously broken gauge symmetry. This is true even whether it is a composite particle, as in the cases in which its mass is much smaller than the compositeness scale, or a Kaluza-Klein mode in theories where the W boson propagates in extra dimensions. The simplest extension of the electroweak gauge group that includes a W' boson is SU$(2)_1\times$SU$(2)_2\times$U(1), but we can encounter larger groups

in some theories. A generic property of these gauge theories is that they also include a Z' boson, and the ratio between the Z' and W' masses is often a free parameter of the theory.

In such new theories, a tree-level mass mixing may be induced between all the electrically-charged gauge bosons. If we diagonalise their mass matrix, the $W-Z$ mass ratio and the couplings of the observed W boson are shifted from the SM values. Their actual measurements imply that the mixing angle θ_+ between the gauge eigenstates, both SM's and not, must be smaller than about 10^{-2}. In certain theories the mixing is negligible, even when the W' mass is near the electroweak scale. The W' coupling to W and Z is fixed by Lorentz and gauge invariances and, to leading order in θ_+, it is given by:

$$\frac{g\theta_+ i}{\cos\theta_W}\Big[W'^+_\mu\big(W^-_\nu Z^{\nu\mu} + Z_\nu W^{-\mu\nu} + Z^\nu W^{-\mu} W'^+_{\nu\mu}\big)\Big] + h.c.$$

where $W^{\mu\nu} \equiv \partial^\mu W^\nu - \partial^\nu W^\mu$, and so on for the similar tensors. The W' coupling to the SM Higgs boson h^0 rises thanks to a three-field term in the langrangian density, namely

$$-\xi_h g_{W'} M_W W'^+_\mu W^{\mu-} h^0 + h.c.$$

where $g_{W'}$ is the gauge coupling of the W' boson, and the coefficient ξ_h satisfies the condition $\xi_h \leq 1$ in simple Higgs sectors.

In models based on the left-right symmetric gauge group, as $\mathrm{SU}(2)_L{\times}\mathrm{SU}(2)_R{\times}\mathrm{U}(1)_{B-L}$, the SM fermions that couple to the W boson transform as doublets under $\mathrm{SU}(2)_L$, while the other fermions transform as doublets under $\mathrm{SU}(2)_R$. Consequently, in such a theory the W' boson couples primarily to right-handed fermions, and its coupling to left-handed fermions arises due to the θ_+ mixing, so that C_q^L is proportional to the CKM matrix and its elements are much smaller than the diagonal elements of C_q^R. Generically, C_q^R does not need to be proportional to the entries of the V_{CKM} matrix.

There are many other models based on the $\mathrm{SU}(2)_1{\times}\mathrm{SU}(2)_2{\times}\mathrm{U}(1)$ gauge symmetry. In the so-called alternate left-right model, all the couplings shown in eq. (3.2.1) vanish, but there are some new fermions such that the W' boson couples to pairs involving a SM fermion and a new one. In the ununified SM, the left-handed quarks are doublets under one $\mathrm{SU}(2)$, and the left-handed leptons are doublets under a different $\mathrm{SU}(2)$, leading to a mostly leptophobic W' boson, namely $C_l^L \ll C_{q_{ij}}^L$ and $C_l^R = C_{q_{ij}}^R = 0$.

Fermions of different generations may also transform as doublets under different $\mathrm{SU}(2)$ gauge groups, as it happens in 331 model. In particular, the couplings to third generation quarks may be enhanced.

It is also possible that the W' couplings to SM fermions are highly suppressed. For example, if the quarks and leptons are singlets under the $\mathrm{SU}(2)$ related to the W', then the couplings are proportional to the tiny mixing angle θ_+. Similar suppressions may arise if some vector-like fermions mix with the SM fermions. Gauge groups that embed the electroweak symmetry, such as $\mathrm{SU}(3)_W{\times}\mathrm{U}(1)$ or $\mathrm{SU}(4)_W{\times}\mathrm{U}(1)$, also include one or more W' bosons.

3.2.2 The searches at the colliders

Several searches of W' have been performed at different colliders in the last decades, as both it was clear SM is not enough to explain completely and satisfactorily the Nature, and the development of the experimental instruments allowed us to investigate increasingly energy ranges.

At LEP-II, W' bosons could have been produced in pairs via their photon and Z couplings. The production cross section at this collider is large enough to rule out the range of mass $M_{W'} < \sqrt{s}/2 \approx 105$ GeV for most patterns of decay modes.

At hadron colliders, W' bosons can be detected through resonant pair production of SM fermions or bosons. Assuming that the W' width is much smaller than its mass and the sensitivity of our actual detectors, the contribution of the s-channel W' boson exchange to the total rate for $pp \to f\overline{f}'X$, where f and f' are fermions whose difference of electric charges is ± 1, and X is any final state physically compatible with the considere decay, may be approximated by the branching fraction $B(W' \to f\overline{f}')$ times the following production cross section of W':

$$\sigma(pp \to W'X) \simeq \frac{\pi}{48s} \sum_{i,j} \left[\left(C^L_{q_{ij}} \right)^2 + \left(C^R_{q_{ij}} \right)^2 \right] w_{ij} \left(\frac{M^2_{W'}}{s}, M_{W'} \right)$$

The functions w_{ij} include the information about the proton structure and, to leading order in α_s, are given by

$$w_{ij}(z,\mu) = \int_z^1 \frac{dx}{x} \left[u_i(x,\mu)\overline{d}_j\left(\frac{z}{x},\mu\right) + \overline{u}_i(x,\mu) d_j\left(\frac{z}{x},\mu\right) \right]$$

The most commonly studied W' signal consists of a high-momentum lepton, such as electron or muon, and large missing transverse momentum. The transverse mass distribution of the final states forms a Jacobian peak with its endpoint at $M_{W'}$ (cfr. fig. 3.1). As the branching fractions for $W' \to e\nu$ and $W' \to \mu\nu$ could be very different, the results in these channels are preferentially presented separately. Searches in these decays often implicitly assume that the left-handed couplings vanish, which means there's no interference between W and W', and that the right-handed neutrino is light compared to the W' boson and escapes the detector. An example of parameter values that satisfy these assumptions is $C^R_q = gV_{CKM}$, $C^R_l = g, C^L_q = C^L_l = 0$, which define a model that preserves lepton universality and is essentially equivalent to the Sequential SM used in many W' searches. However, if a W' boson were discovered with the final state fermions left-handed, the effects of $W - W'$ interference could be observed, providing useful information about the W' couplings.

In the $e\nu$ channel, the ATLAS and CMS Collaborations set limits on the W' production cross section times branching fraction and, thus, indirectly on the W' couplings, when $M_{W'}$ is in the $0.15 - 6$ TeV range, based on $2 - 36$ fb^{-1} at $\sqrt{s} = 13$ TeV, as shown in fig. 3.2(a). ATLAS sets the strongest mass lower limit[2] $M_{W'} > 5.2$ TeV in the Sequential SM. The coupling limits are much weaker for $M_{W'} < 150$ GeV, a range last explored with the Tevatron at $\sqrt{s} = 1.8$ TeV.

[2]From now, all the statistical limits are intended at the 95% CL.

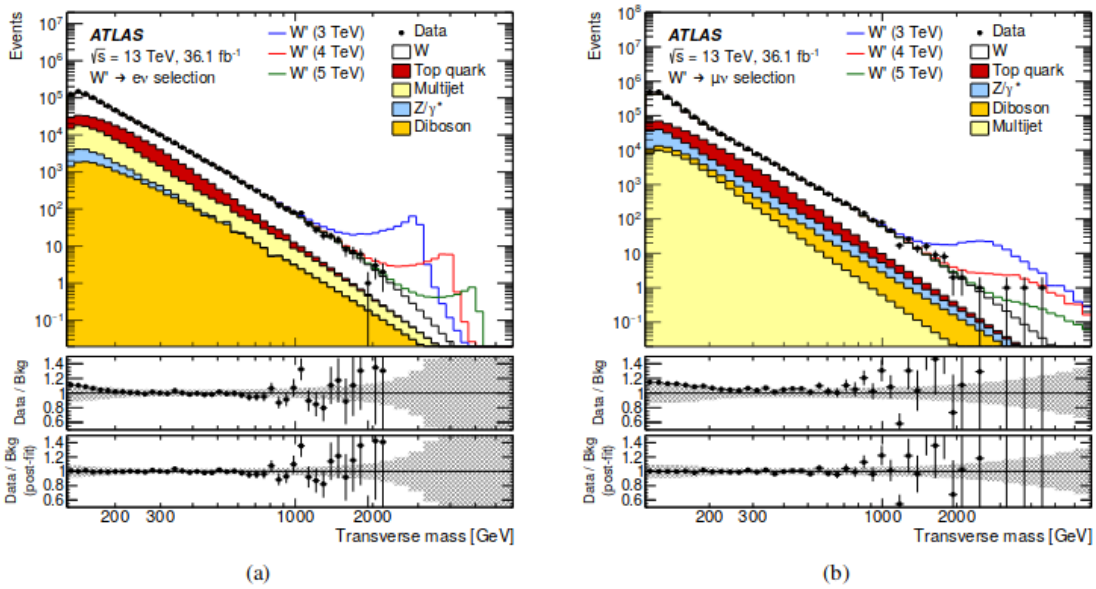


Figure 3.1: The shape of the W' Jacobian peak in the transverse mass spectrum of the final states in a totally leptonic decay. In (a) the considered decay involves the electronic generation, in (b) the muonic one. Courtesy of ATLAS Collaboration.

In the $\mu\nu$ channel, ATLAS and CMS set rate limits for $M_{W'}$ in the $0.15 - 6$ TeV range of energy from the same analyses as mentioned above, with the strongest lower mass limit of 4.5 TeV set by ATLAS using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV data. When combined with the $e\nu$ channel assuming lepton universality, the upper limit on the $\sqrt{s} = 13$ TeV cross section times branching fraction to $l\nu$ varies between 0.2 and 4 fb for $M_{W'}$ between 1 and 5 TeV. Only weak limits on $W' \to \mu\nu$ exist for $M_{W'} < 150$ GeV. Note that masses of the order of the electroweak scale are interesting from a theorethical point of view, while lepton universality does not necessarily apply to a W' boson.

Dedicated searches for $W' \to \tau\nu$ have been performed by CMS at 8 TeV and 13 TeV . Limits are set on $\sigma \cdot B$ for $M_{W'}$ between 0.3 and 4 TeV for the former and between 1.0 and 5.8 TeV for the latter. A lower mass limit of 3.3 TeV is set in the Sequential SM.

The W' decay into a lepton and a right-handed neutrino, ν_R, could also be followed by the ν_R decay through a virtual W' boson into a lepton and two quark jets. The CMS and ATLAS searches in the $eejj$ and $\mu\mu jj$ channels have set limits on the cross section times branching fraction as a function of the ν_R mass or of $M_{W'}$. These searches are typically performed with same-charge lepton pairs, that provide strong background reduction and are motivated by models with a left-right symmetry. However, it is also interesting to search in final states with opposite-charge lepton pairs, as done in the CMS analysis. A related W' search in the $\tau\tau jj$ channel, with τ decaying in jets and ν, was also performed by CMS.

The $t\bar{b}$ channel is particularly important, as a W' boson which couples only to right-handed fermions cannot decay to leptons when the right-handed neutrinos are heavier than the W' boson. Additional motivations for the relevance of this channel are provided by a W' boson with enhanced couplings to the third gen-

eration, and by a leptophobic W'. The usual signature consists of a leptonically-decaying W boson and two b-jets. Recent studies have also incorporated the fully hadronic decay channel for $M_{W'} \gg m_t$, with the use of jet substructure techniques to tag highly boosted top-jets.

Searches for dijet resonances may be used to set limits on $W' \to q\overline{q}'$. CMS and ATLAS provide similar coverage in the $\sim 0.75 - 7.0$ TeV mass range, with data collected at $\sqrt{s} = 8$ and 13 TeV, with the most stringent lower W' mass limit in the Sequential SM set to 3.6 TeV, obtained using 37 fb^{-1} of 13 TeV data. For lower masses, the best limits on W' couplings to quarks have been set by CDF in the $300 - 500$ GeV range, and by CMS in the $500 - 750$ GeV range. Limits for W' masses in the $50 - 300$ GeV range can be derived from the dijet limits on Z' bosons set by CMS.

In some theories the W' couplings to SM fermions are suppressed by discrete symmetries. W' production, then, occurs in pairs through a photon or Z boson. The decay modes are model-dependent and often involve other new particles. The ensuing collider signals arise from cascade decays and typically include missing transverse momentum.

Searches for WZ resonances at the LHC have focussed on the process $pp \to W' \to WZ$, with the production mainly from $u\bar{d}W'$ vertices assuming SM-like couplings to quarks. ATLAS and CMS have set the strongest upper limits on the $W'WZ$ coupling for $M_{W'}$ in the $0.2 - 5.0$ TeV range with a combination of fully leptonic, semi-leptonic and fully hadronic channels at both 8 and 13 TeV [1, 53]. The strongest single lower limit on the mass is set by CMS at 13 TeV, with 35.5 fb^{-1} in the $WZ \to (jj)(\nu\overline{\nu})$ final state, where the parentheses represent a resonance. In the limit is $M_{W'} > 3.2$ TeV in the context of the Heavy Vector Triplet (HVT) weakly-coupled, scenario A.

A fermiophobic W' boson that couples to WZ may be produced at hadron colliders in association with a Z boson, or via WZ fusion. This would give rise to $(WZ)Z$ and $(WZ)jj$ final states.

W' bosons have also been searched for recently in final states with a W boson and a SM Higgs boson in the channels $W \to l\nu$ or $W \to qq$ and $h^0 \to b\bar{b}$ or $h^0 \to WW$ by ATLAS and CMS at $\sqrt{s} = 8$ and 13 TeV. Cross section limits are set for W' masses in the range between 0.4 and 4.5 TeV. The ATLAS and CMS 13 TeV analyses both set the most stringent lower limit on the mass, that is $M_{W'} > 2.4$ TeV for the HVT weakly-coupled, scenario A.

3.2.3 Low-energy constraints

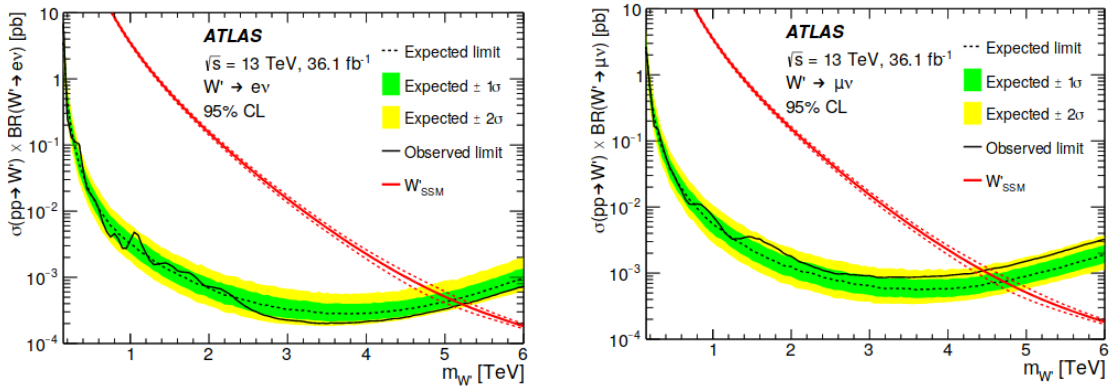
The properties of W' bosons are also constrained by measurements of processes at energies much below $M_{W'}$. The bounds on $W - W'$ mixing are mostly due to the change in W properties, compared to the ones predicted by the SM. Limits on deviations in the ZWW couplings provide a leading constraint for fermiophobic W' bosons.

Constraints arising from low-energy effects of W' exchange are strongly model-dependent. If the W' couplings to quarks aren't suppressed, then box diagrams involving a W and a W' boson contribute to neutral meson-mixing. In the case of W' couplings to right-handed quarks, as in the left-right symmetric model, the

limit from $K_L - K_S$ mixing is severe, namely $M_{W'} > 2.9$ TeV for $C_q^R = gV_{CKM}$. However, if no correlation between the W' and W couplings is assumed, then the limit on $M_{W'}$ may be significantly relaxed.

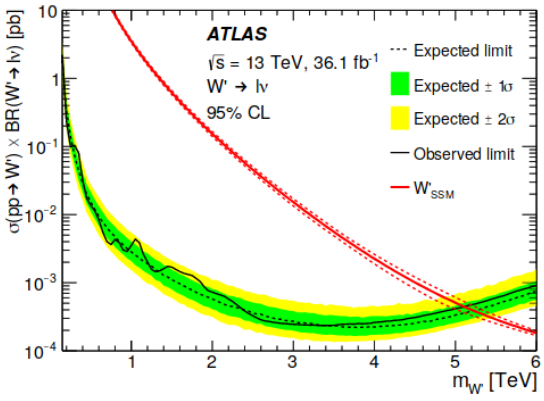
W' exchange also contributes at tree level to various low-energy processes. It would impact, particularly, the measurement of the Fermi constant G_F in muon decay, which, in turn, would change the predictions of many other electroweak processes. A recent test of parity violation in polarized muon decay has set limits of about 600 GeV on M'_W, assuming W' couplings to right-handed leptons as in left-right symmetric models and a light ν_R. There are also W' contributions to the neutron electric dipole moment, β decays, and other processes.

If right-handed neutrinos have Majorana masses, then there are tree-level contributions to neutrinoless double-beta decay, and a limit on M'_W versus the ν_R mass may be derived. For ν_R masses below a few GeV, the W' boson contributes to leptonic and semileptonic B meson decays, so that limits may be placed on various combinations of W' parameters. For ν_R masses below ~ 30 MeV, the most stringent constraints on $M_{W'}$ are due to the limits on ν_R emission from supernovae.



(a) (b)

Figure 3.2: Upper limit on $\sigma(pp \to W'X \cdot B(W' \to l\nu)$ from ATLAS, at 95% CL. In (a) the lepton l is the electron, in (b) is the muon. The red line shows the theoretical prediction in the Sequential SM.



(c)

Figure 3.3: Upper limit on $\sigma(pp \to W'X \cdot B(W' \to l\nu)$ from ATLAS, at 95% CL, obtained combining the data in 3.2. The red line shows the theoretical prediction in the Sequential SM.

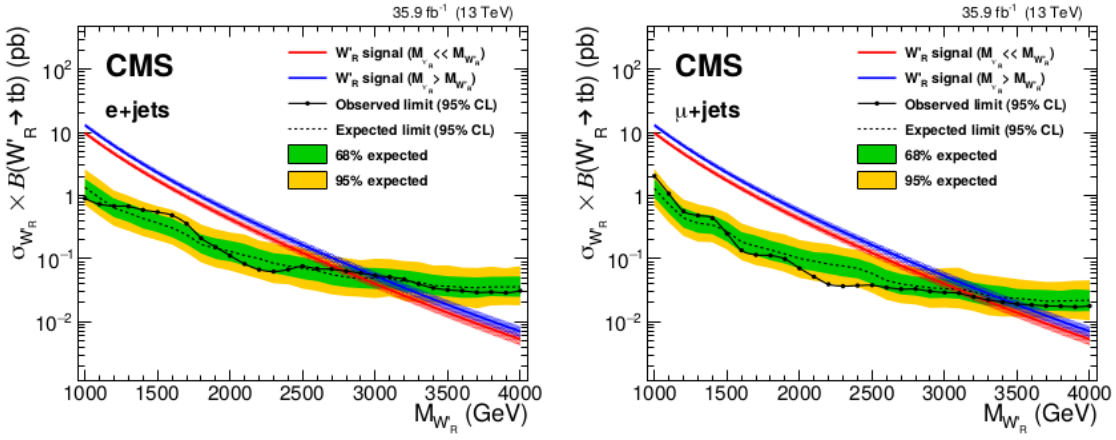


Figure 3.4: Upper limit at 95% CL on the W'_R boson production cross section separately in the electron (left) and muon (right) channels. Signal masses for which the theoretical cross section (in red and blue for $M_{\nu_R} \ll M_{W'_R}$ and $M_{\nu_R} > M_{W'_R}$, respectively) exceeds the observed upper limit (in solid black) are excluded at 95% CL. The green and yellow bands represent the ± 1 and 2 standard deviation uncertainties in the expected limit, respectively.

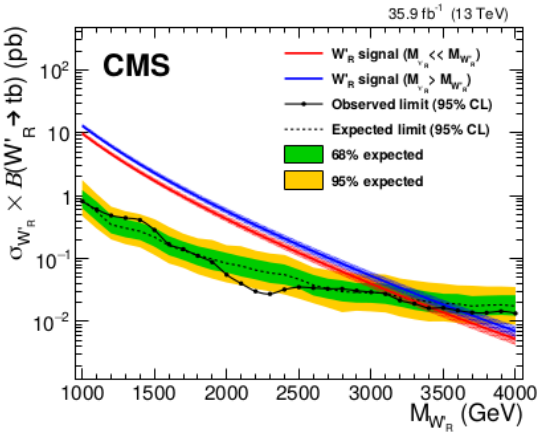


Figure 3.5: Upper limit at 95% CL on the W'_R boson production cross section for the combined electron and muon channels. Signal masses for which the theoretical cross section (in red and blue for $M_{\nu_R} \ll M_{W'_R}$ and $M_{\nu_R} > M_{W'_R}$, respectively) exceeds the observed upper limit (in solid black) are excluded at 95% CL. The green and yellow bands represent the ± 1 and 2 standard deviation uncertainties in the expected limit, respectively.

3.3 W' and the anomalies of the SM

3.3.1 W' and anomalies of the meson sector

[3, 10] Recently, the BaBar, Belle, and LHCb collaborations have measured the semi-leptonic decays of B mesons to D and D^*, and found a sizable discrepancy from the value predicted by the SM. More specifically, the anomaly is manifest in the observables

$$R(D) = \frac{BR(\overline{B} \to D\tau^-\overline{\nu}_\tau)}{BR(\overline{B} \to Dl^-\overline{\nu}_l)}$$
$$R(D^*) = \frac{BR(\overline{B} \to D^*\tau^-\overline{\nu}_\tau)}{BR(\overline{B} \to D^*l^-\overline{\nu}_l)}$$

where $l = e, \mu$. The denominator is averaged over electrons and muons in BaBar and Belle, while for LHCb only muons contribute. The special property of these observables is that, in the SM, the hadronic factors are expected to cancel out, which reduces the uncertainty in our prediction. The measurements are found to disagree with the SM predictions at about 4σ.

The discrepancy suggests that it may be possible there is a lepton flavour universality violation, and that has prompted many works to explain the anomaly in the context of new physics models. The underlying interaction that accommodates the experimental result arises from the charged-current mediated decay $b \to c\tau\nu$, which is CKM-suppressed in the SM. One simple way to obtain a new physics contribution to this charged current is to use a W' gauge boson which couples to the second and third generation fermions. In order to explain the anomaly, the W' does not need to couple to the first generation, and this fact seems to make the model harder to explore at the colliders like LHC.

As mentioned before, the BaBar, Belle and LHCb collaborations have measured $R(D)$ and $R(D^*)$ to very high precision. The combined experimental values of these two observables are:

$$R(D)_{Exp} = 0.407 \pm 0.039 \pm 0.024$$
$$R(D^*)_{Exp} = 0.304 \pm 0.013 \pm 0.007$$

while the theory values are:

$$R(D)_{SM} = 0.298 \pm 0.003$$
$$R(D^*)_{SM} = 0.255 \pm 0.004$$

If we take account of the contribution to these processes of phenomenons of new physics, including the W', boson, the observable above are modified as follows:

$$R(D^{(*)})_{NP} = \left(1 + \frac{g'_q g'_\tau}{m^2_{W'}} \frac{\sqrt{2}}{4G_F V_{cb}}\right) R(D^{(*)})_{SM}$$

where $m_{W'}$ is the W' mass, g'_q, g'_τ are, respectively the W' couplings to the q-flavour quark and to the τ lepton, $G_F = 1.16 \times 10^{-5}$ GeV^{-2} is the Fermi constant,

and $V_{cb} = 0.04$ is the cb entry of the CKM matrix. Taking both new couplings to be positive, the central values of $R(D)$ and $R(D^*)$ require the factor $g'_q g'_\tau / m^2_{W'}$ to be $0.002(100\text{ GeV}/m_{W'})^2$ and $0.001(100\text{ GeV}/m_{W'})^2$, respectively. In fig. 3.6 are reported the limits in the $m_{W'} - g'_q$ plane for several representative values of g'_τ, as they are constrained by the experimental results [10]. In these plots they show, for representative choices of g'_τ, the projected sensitivity of the 13 TeV LHC, using their analysis for 30, 300, and 3000 fb^{-1} of luminosity in the $g'_q - m'_W$ plane, overlaying the 1σ fits to the $R(D)$ and $R(D^*)$ measurements. Points above the black lines have a projected significance equal to or higher than 3σ.

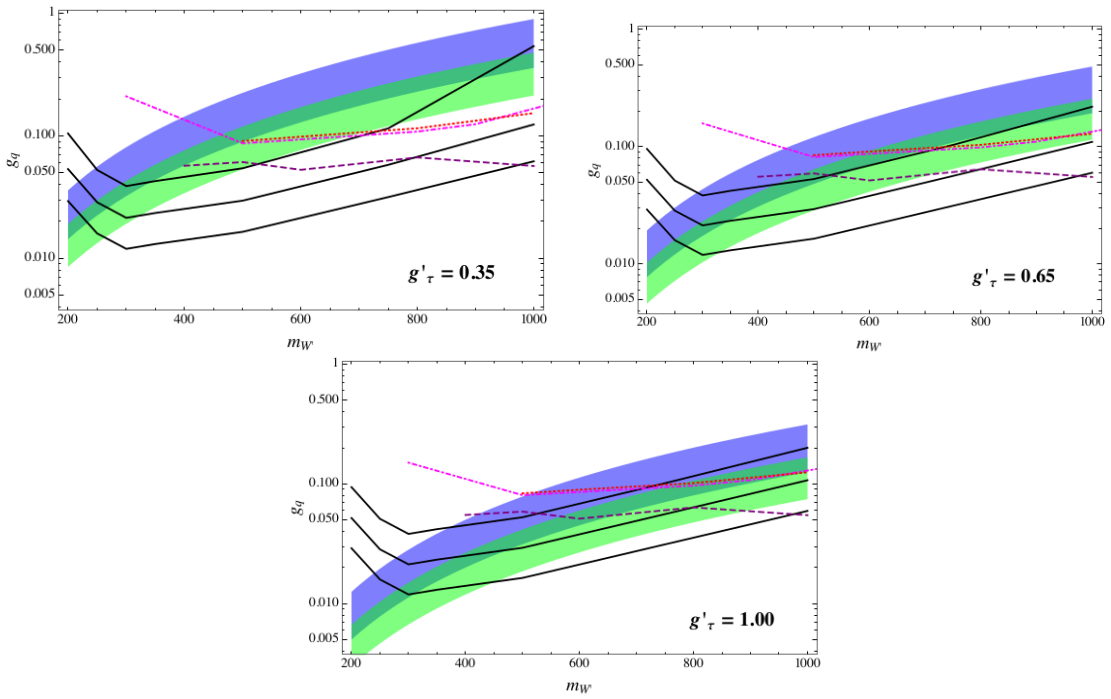


Figure 3.6: Projected 3σ sensitivity on the quark coupling g'_q as a function of $m_{W'}$ for $g'_\tau = 0.35, 0.65, 1$ at the 13 TeV LHC and $L = 30, 300, 3000$ fb^{-1} (solid black curves). Superimposed, the 1σ bands explaining the $R(D)$ (blue) and $R(D^*)$ (green) anomalies. Also shown are the recast inclusive limits from ATLAS at 13 TeV (dotted red curve), CMS at 13 TeV (dashed purple curve) and CMS at 8 TeV (dot-dashed magenta curve).

3.3.2 W' and the anomaly of the baryon asymmetry

It is known that the leptogenesis mechanism offers a very good possibility to explain the baryon asymmetry of the universe. The recent evidence for neutrino masses, indeed, has brought forward leptogenesis as a very attractive mechanism to explain this asymmetry of the Universe. Along this mechanism, the baryon asymmetry of the universe is explained by the same interactions as the ones which can explain the neutrino masses. However, it will be very hard, if not impossible, to test experimentally in a foreseeable future, as most of the mechanism typically takes place at high energy or results from suppressed interactions, without unavoidable low-energy implications. It's possible, however, taking an alternate approach by asking if it can be, at least, falsified. Possible discoveries at current and future colliders, particularly that of right-handed gauge interactions, would indeed forbid at least the leptogenesis mechanisms based on right-handed neutrino decay. For high mass right-handed neutrinos and W_R[3], around 10^{10} GeV or higher, the W_R, in fact, can have suppression effects on leptogenesis through dilution and scattering of the right-handed neutrino (the latter reported as N_R in fig. 3.7). With a low scale W_R, moreover, the suppression effects are dramatically enhanced [28].

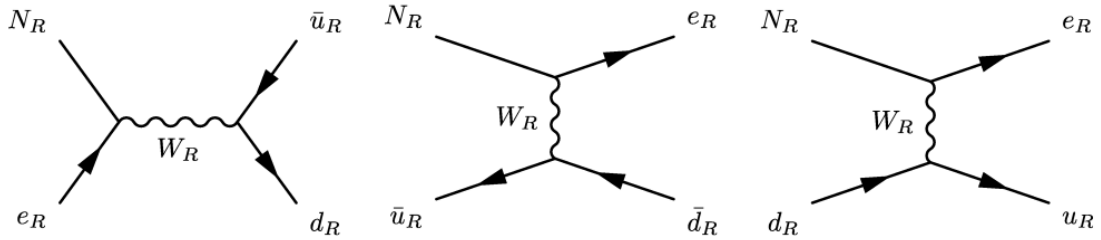


Figure 3.7: Scatterings involving the W_R. N_R is the hypothetical right-handed neutrino, L a not-SM lepton.

[3] W_R is the name for the right-handed form of the hypothetical W' boson.

Chapter 4

Decay channel and physical objects of interest

The goal of this thesis work is the search for the W' boson resonance using the following datasets:

- 41.5 fb^{-1} of proton-proton collision data[1], collected by the CMS experiment in 2017;
- 59.74 fb^{-1} of proton-proton collision data, collected by CMS in 2018.

In both 2017 and 2018 the aforementioned data were delivered by the LHC at a centre of mass energy of $\sqrt{s} = 13$ TeV.

In this chapter we are going to describe the decay channel of the W' boson this thesis focuses on, characterising the searched signal and listing the SM events which have to be considered as not negligible backgrounds for our analysis.

We are going also to briefly display the reconstruction event tools needed to reveal the physical objects produced during the collisions and finalise the analysis we are intended to pursue.

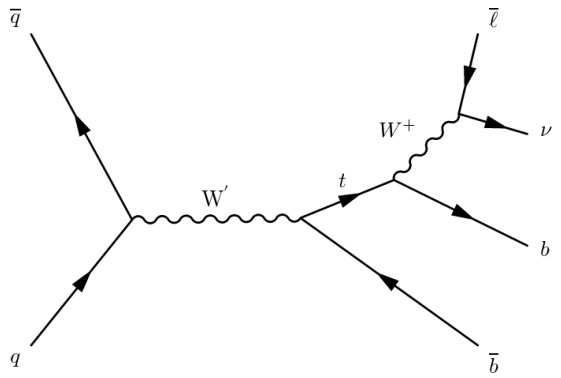


Figure 4.1: Representative Feynman diagram at tree level for the considered W' boson production and decay mode.

[1] Evaluated in integrated luminosity L.

4.1 Signal and background characterisation

We want to perform our search in the semileptonic decay channel of the W' boson, namely:

$$W' \rightarrow t\bar{b} \rightarrow l^{+}\nu_{l}\bar{b} \tag{4.1.1}$$

$$\bar{t}b \rightarrow l^{-}\overline{\nu_{l}}b \tag{4.1.2}$$

In the events recorded by the detectors, thus, we search for a final state characterized by two jets, one charged lepton and missing energy. This topology is not a *unicum*, as it can also be reproduced by some SM phenomenons. A selection on such final states is therefore not sufficient to fully separate signal events from the ones coming from its competing SM processes. We shall henceforth refer to such processes as backgrounds for our search.

The aforementioned SM backgrounds have to be investigate and properly modeled, in order to get an estimate of their contributions to our event selection.

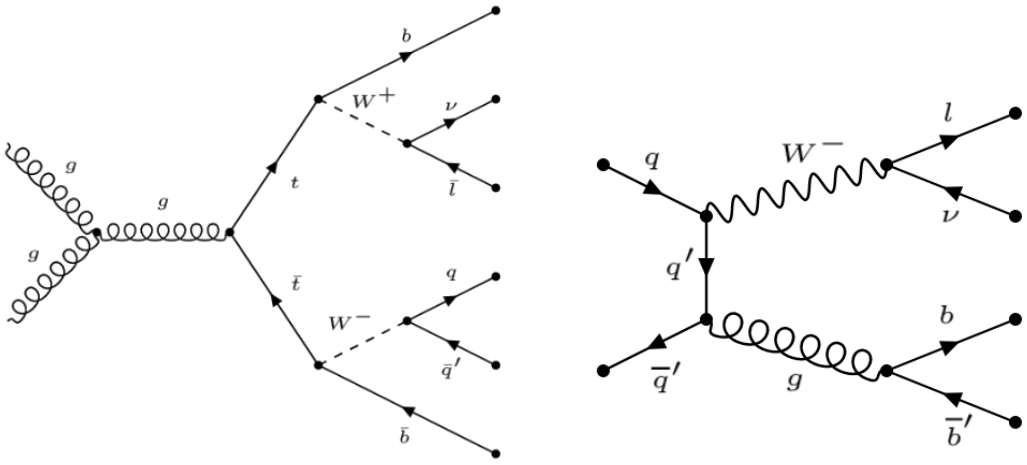


Figure 4.2: Representative Feynman diagrams at tree-level of the background processes considered in this thesis. The charge-conjugated events are possible, too.

In our study we considered the main SM backgrounds contributing to the investigated topology.:

- $t\bar{t}$, processes where a top quark-antiquark pair is produced by strong interaction. In case one t quark decays through a $t \rightarrow Wb \rightarrow l\nu b$ chain, such final state exactly replicates the decay of the t quark from the W' boson. The decay products of the remaining t quark can be misidentified as a b -jet produced promtly by the W' boson.

- $W+$ jets processes, where the coupling between two quarks with different flavours create a W boson and a gluon. If the W does decay leptonically, and the gluon produces two b quarks and, therefore, two jets, this event does replicate the W' signature. .

4.2 Reconstruction and identification of the physics objects

The processes of reconstruction and the identification of all stable particles produced through the proton-proton collisions, such as electrons, muons, photons, charged and neutral hadrons, including their own kinematic features, is performed thanks to the **Particle Flow** (PF) algorithm [23, 45]. The PF exploits a set of informations coming from all CMS sub-detectors towards optimal identification and determination of the particles, as such as their direction, energy and type. The list of particles identified by the PF algorithm is, then, used in order to pursue the following goals:

- build jets which the quark and gluon energies and directions are inferred from;
- determine the missing transverse energy, briefly named **MET**, which gives an estimate of the direction and energy of the neutrinos and other invisible particles;
- quantify charged lepton isolation with respect to other particles;
- tag jets coming from the hadronisation of b or t quarks, and, to a minor extent, c quarks or gluons.

An iterative tracking algorithm is the key to link all the reconstructed elements to each other, and this procedure is performed by identifying blocks of elements that are compatible, each one of them to the others. These blocks are generally named clusters. Starting from the latter, PF candidates are reconstructed, then, in the following order:

- **muons**: a global muon, reconstructed from the hits in the tracker and the track in the muon system, gives rise to a PF muon. If such particle is identified, the corresponding tracks are removed from the block;
- **electrons**: a particle is identified as an electron if the link between a charged particle track and one or more ECAL clusters are compatible. Also these tracks, then, are removed from the list;
- **charged hadrons**: after the previous identifications are made, the remaining tracks are used to reconstruct charged hadrons. The tracks can be linked to ECAL and HCAL clusters, while the momentum is directly taken from the tracker and corrected with the help of the data recorded by the calorimeters;
- **photons and neutral hadrons**: clusters in the ECAL that are not compatible with charged tracks give rise to PF photons. In a similar way, neutral hadrons are identified if there are unmatched clusters in HCAL.

When the list of PF candidates is complete, PF jets can be reconstructed, using dedicated jet reconstruction algorithms, and, finally, the MET of the event can be estimated.

The PF algorithm also pursues the important task of detecting the vertices of the collisions and the decays. They are reconstructed thanks to a best fit procedure to the intersection of tracks reconstructed in the tracking system and, after that, are defined as good if they satisfy the following conditions:

- more than 4 degrees of freedom in the fit;
- less than 2 cm away in the xy plane from the interaction point;
- less than 24 cm away in the z direction from the interaction point.

Those requirements guarantee that the reconstructed vertices are in the luminous region. In our thesis only events where at least on good primary vertex is found are selected. Among the good vertices, it is defined as primary vertex the one corresponding to the highest value of the sum of the squared transverse momentum of all the tracks associated with it.

In the following, we are going to extensively discuss about the physical object the Particle Flow is capable to reconstruct that will be used in the analysis.

4.2.1 Electrons

Electron candidates are reconstructed from a collection of electromagnetic clusters with matched pixel tracks. Their tracks, then, are fitted using a Gaussian Sum Filter (GSF) algorithm along the trajectory, taking into account the likely emission of *bremsstrahlung* photons in the silicon tracker. All electrons are required to have $p_T > 30$ GeV and $|\eta| < 2.5$. Different requirements are applied to electrons reconstructed in the calorimeter barrel and endcaps. Electron candidates are subsequently discriminated according to the *veto* selections defined in CMS, based on identification criteria with high efficiency but low purity. One of the variables used for identification of electrons is the relative isolation I^e_{rel} , defined as follows:

$$I^e_{rel} = \frac{I^{h_c h} + \max\left[(I^\gamma + I^{h_n} - \rho \times A), 0\right]}{p_T}$$

where p_T is the transverse momentum of the electron, ρ is the average energy density non clustered in jets, measured event-by-event, by the cone area A and $I^{h_c h}, I^{h_n}$ and I^γ are the scalar transverse momentum sums of charged hadrons, neutral hadrons and photons, respectively. The sums are computed in a con of $\Delta R = 0.3$ around the electron direction, where ΔR is angular distance in space defined as follows:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

4.2.2 Muons

Muons are, at first, reconstructed in the tracking system using a technique based on Kalman filter algorithm and, then, in the muon chamber. That is performed combining the informations from the drift tubes (DT), the cathode strip chambers (CSC) and resistive plate chambers (RPC). The reconstructed track in the

muon chamber is named **stand-alone muon**. Starting from these objects, two approaches can be used to complete the task:

- the **Global Muon reconstruction**, that, for each stand-alone track, founds a matching track, in order to reject the background of muons coming from hadronisation and from the particles that interact around the beam pipe. This approach is especially useful at large transverse momenta because it improves the momentum resolution.
- the **Tracker Muon reconstruction**, that consider a track in the tracker as a muon candidate with $p_T > 0.5$ GeV and a total momentum $p > 2.5$ GeV. If some hits deposited in the muon chambers can match this track, also considering the energy losses due to the Coulomb scattering, the corresponding track is identified as a Tracker Muon. This method is efficient at low momentum, $p < 5$ GeV, because it requires only a hit in one of the components of the muon chambers.

When a single muon is identified both as Global and Tracker, if the latter ones share the same track in the tracking system they are merged together. The isolation variable I^{μ}_{rel} is defined, in the muon case, as:

$$I^{\mu}_{rel} = \frac{I^{h_{ch}} + \max\left[(I^{\gamma} + I^{h_n} - 0.5 \times I^{PU})\right], 0}{p_T}$$

where $I^{h_{ch}}, I^{\gamma}, I^{h_n}$ and I^{PU} are the scalar transverse momentum sums of the charged hadrons, the photons, the neutral hadrons and the charged hadrons associated with pileup vertices, respectively. The sums are computed in a cone of $\Delta R = 0.4$ around the muon direction.

4.2.3 Jets

In the majority of the LHC events hadronic partons are produced in the final states, but we are not able to directly observe them, as it is forbidden by the strong interaction. Quarks and gluons manifest themselves, then, trough hadronization to stable particles which can be detected in the tracking chambers and calorimeters. The interactions between constituent partons and the showering into stable particles can be modeled thanks to the perturbative regime of the QCD theory and the hadronization model. We report in fig. 4.3 the evolution of a hadronic jet from hard interaction to observable energy deposits. The jet reconstruction algorithm employed at CMS is based on clustering all the PF candidates reconstructed. The particles identified by PF are clustered by applying the **anti-k_T** algorithm [19].

At the beginning of the clustering procedure, a correction named **Charged Hadron Subtraction (CHS)** is applied on the group of hardest particles, in order to remove the energetic contribution coming from the pile-up interactions before build the cluster itself for what concerns the narrow jets [46]. For the fat jets, the same work is exploited by the procedure called **PileUp Per Particle Identification (PUPPI)** [37]. The anti-k_T algorithm is named after the metric used to decide how to cluster the reconstructed physics object. Namely, the distance between

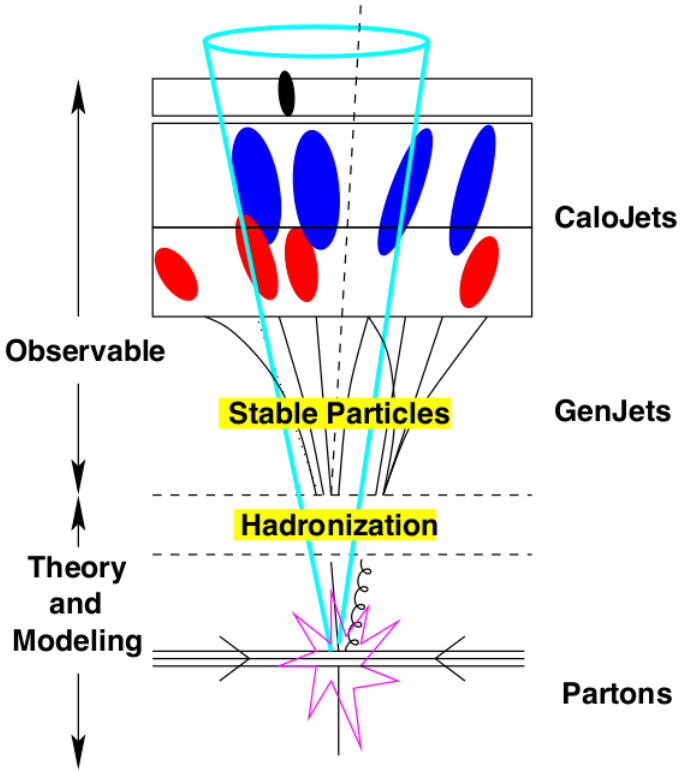


Figure 4.3: Representation of the hadronisation process experienced by the partons, within the framework of a detector such as CMS.

the particles i and j d_{ij}, and the one between the particle i and the beam, d_{iB} are defined as follows:

$$d_{ij} = \min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right)\frac{\left(\Delta R\right)_{ij}^2}{R^2}$$

$$d_{iB} = \frac{1}{p_{T,i}^2}$$

where

$$\left(\Delta R\right)_{ij}^2 = \left(\eta_i - \eta_j\right)^2 + \left(\phi_i - \phi_j\right)^2$$

as we reported before. Jets with cone parameter of $\Delta R = 0.4$ are named **narrow** or **resolved jets** (AK4), while jets clustered with a cone parameter of $\Delta R = 0.8$ are defined as **fat** or **boosted jets** (AK8). Generally, the latter are considered for the selection of boosted t quark candidates or W boosted candidates, while the former are used to select a forward light-flavour jet, as the ones originated by a b quark [11, 41, 48, 57]. In both collections, jets are required to pass a set of identification criteria, defined globally as *loose* working point, according to the guidelines of the dedicated group of CMS that works on Jets and MET.

Several corrections at various levels have to be taken in account when calculating the energy of the jets, and they are known as **Jet Energy Corrections (JEC)**. These corrections have been introduced to reproduce the energy at particle level, and they consist in scaling the jet energy is scaled by a factor that describes the detector response depending on the transverse energy and the

pseudorapidity of the jet [22] (cfr. fig. 4.4). The corrections pass trough the following steps:

- removal of the energy due to the contribution of pileup events;
- correction of the response of jets in order to achieve a flat distribution in η;
- correction of the jet response for p_T.

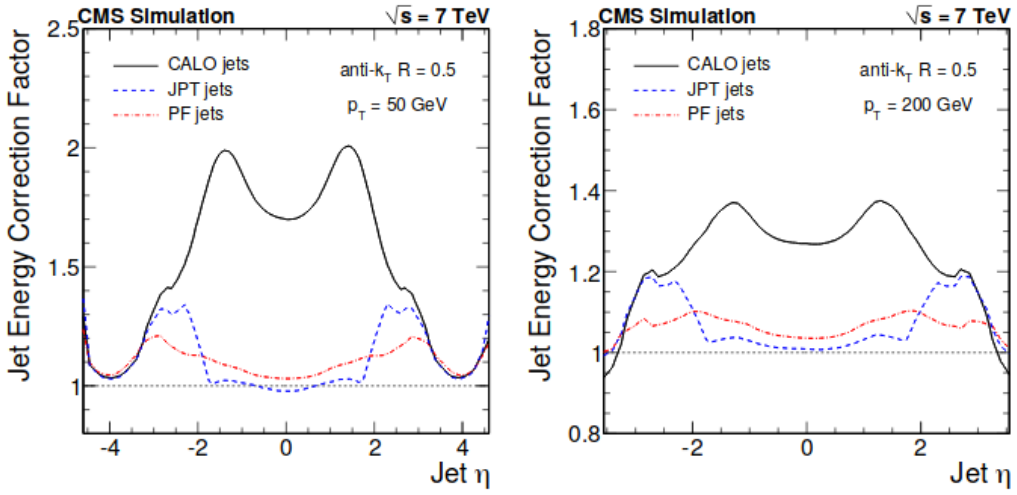


Figure 4.4: Monte Carlo Jet Energy Correction Factors for the different jet types, as a function of jet η. On the left hand, correction factor required to get a corrected jet with $p_T = 50$ GeV; on the right hand, correction factor required to get a corrected one with $p_T = 200$ GeV.

4.2.4 b-tagging procedure: the DeepCSV algorithm

In high energy physics searches the jet flavour classification is a critical element of the analysis strategy. It plays a central role in identifying heavy-flavour signals and reducing the enormous backgrounds from light-flavor processes.

Jets originating from heavy-flavour quarks tend to produce longer-lived particles than those found in jets from light-flavor quarks. These long-lived particles have decays which are displaced from the primary vertex of the considered event. To identify such vertices, the central tracking chamber measures the trajectories of charged particles, which allows for the reconstruction of vertex locations. The variable used to define the distance between the two vertices is the **Impact Parameter (IP)** (cfr. fig. 4.5), that is a Lorentz invariant. That means IP is invariant with respect to changes of the long lived particle kinetic energy, and for the B hadrons this quantity corresponds to $c\tau_B \approx 450\mu$m. In CMS the IP can be measured with precision between 30 μm and hundreds μm. Also the uncertainty can be of the same order of IP magnitude and, then, the IP significance, defined as IP$/\sigma_{\text{IP}}$ is a better observable. It can be positive or negative, depending on the signs of the scalar product of the IP vector and the jet direction. It is expected to

be symmetric around 0 for decays with short life and mostly positive for particles with longer lifetime.

The high mass of the heavy partons leads, in addition, to decay products with a larger transverse momentum relative to the jet axis with respect to the ones typically found in jets from light partons. Finally, heavy hadrons have a sizable branching ratio for semileptonic decays and, hence, the presence of soft leptons in the produced jets provides another tool for heavy jet identification [51].

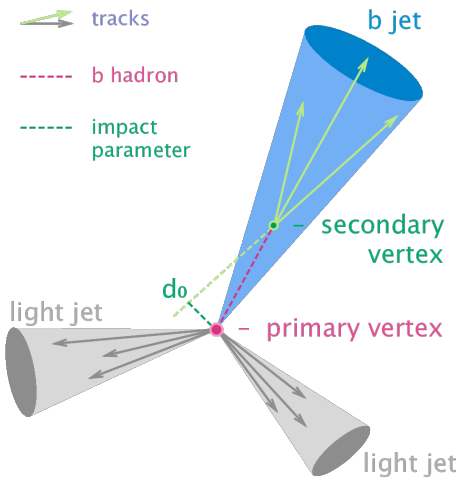


Figure 4.5: Definition of Impact Parameter.

The large and varying number of particles in a jet leads to a difficult classification problem with large and variable dimensionality, without a natural ordering. The first step in typical approaches involves vertex-finding algorithms, which transform the task into one of reduced, but still variable, dimensionality.

Recent applications of deep learning to similar problems in high-energy physics, combined with the lack of a clear analytical theory to provide dimensional reduction without loss of information, suggests that deep learning techniques applied to the lower-level higher-dimensional data can yield improvements in the performance of jet-flavour classification algorithms [34].

The CMS collaboration has implemented these ideas in their flavour-tagging algorithms, developing increasingly efficient taggers [39].

In this thesis we're going to use the DeepCSV tagger, where "Deep" stands for Deep Neural Network and "CSV" for Combined Secondary Vertices.

DeepCSV uses the first six most displaced tracks of the considered jet and seven features that are traditionally used for b tagging, based on secondary vertex and track-based lifetime information. Moreover, it uses the characteristic quantities of the most displaced vertex, as well as ones related to the jet, to build all the requested deep learning features. The latter, that are around 60, are the input to a dense neural network, which has six layers and 100 nodes per layer.

To train properly the DeepCSV tagger, for the first time in CMS a multi-class classification approach is used for the classes. In this case, the light quark class and the gluon class are merged, as usually done in flavour tagging. Another change with respect to previous taggers is that a larger and more diverse sample is used for the training, reducing, in this way, the danger of obtaining a classifier

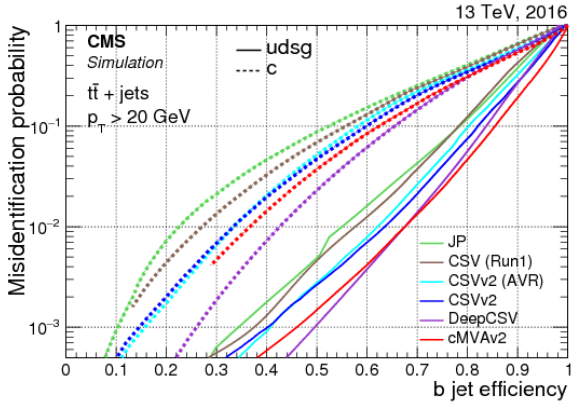


Figure 4.6: True positive rate, reported as b jet efficiency, vs. false positive rate of light jets and c quarks in simulation with 2016 conditions for jets with more than 30 GeV transverse momentum from top quark pair production. The other names in the legend are other taggers developed in CMS collaboration.

too specific to a process. The larger sample sizes, moreover, allow to increase the complexity of the tagger without danger of overfitting [7, 37].

DeepCSV has three working points, named Loose, Medium and Tight, with increasingly tight threshold applied to the rate of the flavour-tagging classifier.

4.2.5 Missing transverse energy (MET)

At the LHC, and in general at any hadron collider, the energy imbalance is computed in the transverse view only. Only the longitudinal momentum of the whole proton, indeed, can be measured, as the one of the individual colliding partons is not known, while, with large precision, we can assume their transverse momentum to be zero. As mentioned before, the PF algorithm reconstructs the missing transverse energy, **MET**, from the vectorial sum of all the candidates transverse momentum, requiring total value of its magnitude is zero. The MET is one of the most important variables for discriminating leptonic decays of W bosons and top quarks from background events which do not contain neutrinos. It is also a crucial variable to search for Beyond Standard Model physics, since many models predicts particles that would not leave any trace in the detectors. The presence of particles that do not interact in the detector creates an energy imbalance in the measurements and, then, can be detected as MET. In this thesis the MET is fundamental to identify events where the t quark decays leptonically. The MET is also one of the most complex variables to reconstruct, since it is very sensitive to detector malfunctions and particles crossing poorly-instrumented regions of the detector. To obtain the best definition of MET the energy corrections applied to the jets (JEC) must be propagated to the MET, as well as corrections to the scale of the particles that are not clustered into jets. Filters to the MET are applied in order to remove pathological events from data, referred in this thesis as **MET filters**.

In the process considered in this thesis MET comes exclusively from the neutrino produced in the t decay (cfr. fig. 4.1). For this reason, we can reconstruct the $4-$momentum of the invisible neutrino in a kinematically way: the magnitude

of the total $4-$momentum of the leptonic couple $l\nu_l$ born in the t decay has to be equal to the mass of the W boson [16]. Given an event with one lepton and MET, following this approach we can directly reconstruct the $4-$momentum of the parent t quark.

4.3 W' candidate reconstruction

The physical objects whose reconstruction has been treated so far are the basis for forging a semileptonically decaying W' candidate.

The W' could hypothetically decay in τ lepton and corresponding neutrino. This lepton, however, cannot be directly reconstructed, since it decays too early to be promptly detected. As from the τ decays the leptons of the other generations are produced, we indirectly include events such as $W' \to \tau\nu_\tau$ in our study.

At first stage, we have to reconstruct the two quarks, t and b, that are the prompt products of a hypothetical W' decay. The t quark, then, decays producing another b quark and a lepton-neutrino pair $l\nu_l$ (we may call it leptonic couple) through the SM W boson. As the neutrino is the only undetectable final state, its kinematic features can be reconstruced considering the only component of the MET.

For what we have just said, we need the following physics objects in order to construct a W' candidate:

- one lepton;
- MET;
- two AK4 jets.

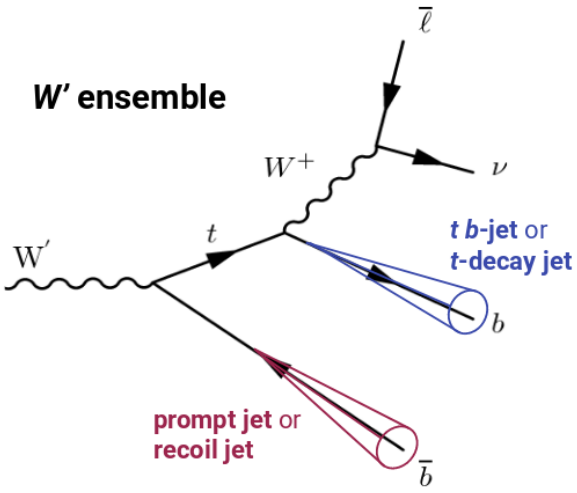


Figure 4.7: Representation of a so-called ensemble.

We call, from now, this set **ensemble**. The first three objects are combined in such a way we are able to reconstruct directly the t quark, as mentioned before [16]. In the same time, we can reconstruct the b quark that completes the

W' vertex with the other AK4 jet, in order to have all the kinematic information about the prompt products of the investigated decay. Finally, we can sum up the $4-$momenta of the reconstructed t e b to get the one associated with the boson candidate. We can, in particular, compute the invariant mass of the latter as follows:

$$M_{W'} = \sqrt{\left(E_t + E_b\right)^2 - \left|\vec{p}_t + \vec{p}_b\right|^2} \tag{4.3.1}$$

Because of that, the spectrum of total invariant mass of the four objects considered to build a W' candidate, namely $M_{l\nu bb}$, is going to be the main character of the analysis strategy. Our purpose is, indeed, to evaluate the contribute to the observed $M_{l\nu bb}$ distribution of a new resonance identifiable as the searched W'.

Given an event with one only lepton and more than one jet, we could build more than one ensemble per event. Coupling the right jet to the leptonic couple to reconstruct the t quark and, subsequently, associate the latter to the right jet to reconstruct the W' vertex is not immediate. A strategy to reconstruct the event in the right way, i.e. by selecting the correct permutation of reconstructed jets and leptons that corresponds to the prompt W' decay products, is necessary. In the following, we are going to show some algorithms and how to evaluate their goodness in discriminating the ensembles.

4.3.1 Reconstructing the event via Monte-Carlo truth

In a Monte-Carlo simulated sample We can obtain all information about the leptons, quarks, and gluons in the hard interaction and, therefore, associate it to the reconstructed jets and leptons, which are the observable physics objects we do reveal in the detector. In such a way, to the simulated reconstructed jet we can associate a unique parton with known flavour. In addition, thanks to that we know which ones are precisely the two jets produced by the b quarks involved in the investigated process. This set of information is called **Monte-Carlo truth**. We can use the latter to get the ideal invariant mass spectrum that we can produce with the physical objects detected and reconstructed by CMS, if we exactly know how they are linked to the primary vertex.

In data this is not possible, since we can reconstruct the events starting only by the *a posteriori* information provided by the detector about the final states. We can, nevertheless, use the invariant mass spectra reconstructed with the MC truth as a benchmark to make up a ensemble-selection strategy. By using exclusively the experimentally available information on jets and leptons, one can define an ensemble selection that best identifies the true ensemble and best reproduces its invariant mass spectrum. In this way, we can arrange the W' best candidate ensemble using only the properties of the detected physical objects.

In order to construct the W' event with the MC truth, we have to do the following considerations. As we explained before, we have to reconstruct the two t and b quarks, directly produced by the W' boson. In order to do that, we search for the b jet and the lepton with the right flavour to be, both of them, products of the t decay. A neutrino is needed to physically complete the t semileptonic decay.

Considering the information provided by the MC truth, we define the MC truth ensemble in the following way:

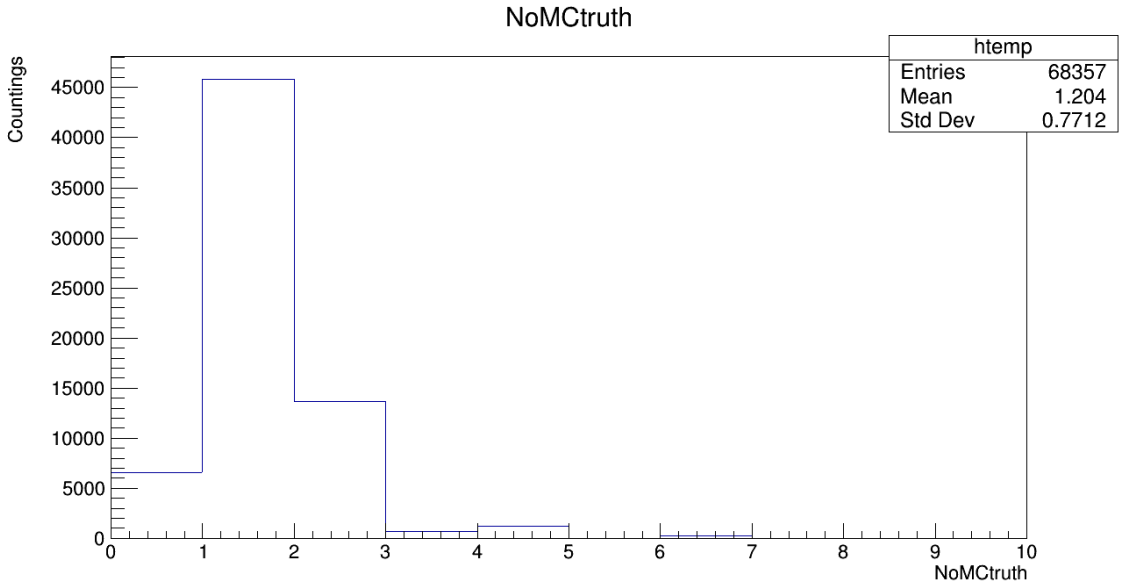


Figure 4.8: The distribution of the number of MC truth ensembles reconstructed per event. Here, we show the results obtained with W' simulated sample for $m_{W'} = 3200$ GeV, right-handed. For more details about the simulated signal samples, see the next chapter.

- one charged lepton, either an electron or a muon;
- MET;
- a AK4 jet generated by a b quark with the charge opposite with respect to the leptonic one, to associate to the lepton itself to reconstruct the t quark, we call it **top b-jet** or **t-decay jet**;
- a AK4 jet generated by a b quark with the same charge with respect to the leptonic one, to associate to the reconstructed t quark, we call it **prompt b-jet** or **recoil jet**.

For what we have said before, we except to find one only MC truth ensemble per event. In fig. 4.8 we can see there are few events with more than one MC truth ensemble. We investigated these events, and we found that, considering two MC truth ensembles in the same event, the recoil jets are very close to each other, and that happens also for the t-decay ones (cfr. fig. 4.9). From this observation we inferred that there are some events in which the reconstruction algorithm produces two jets which are close enough to each other that they are both associated to the same b jet. We will disambiguate those cases by choosing the closest jet.

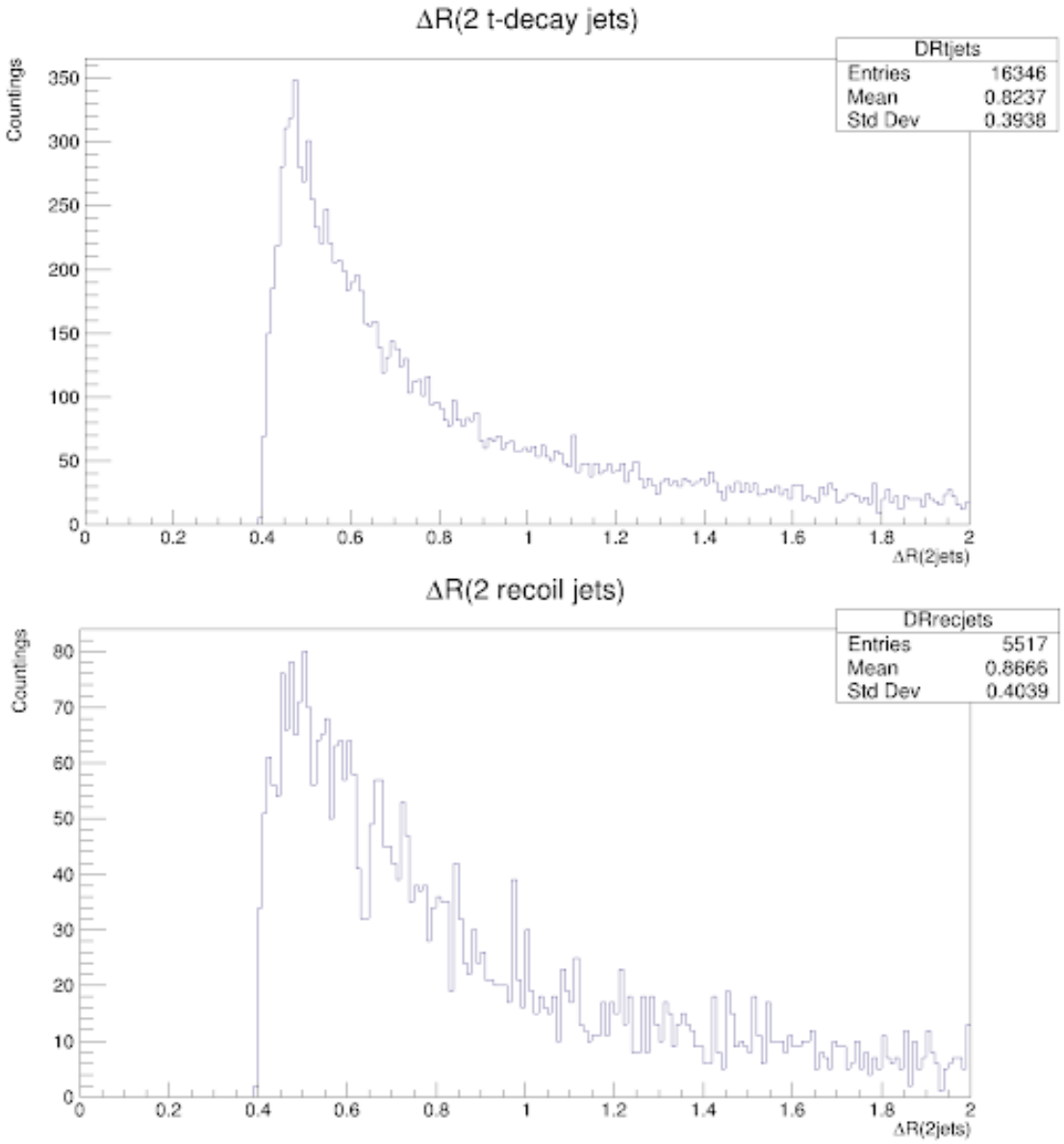


Figure 4.9: The distribution of the angular distance ΔR, calculated the angular distance between the axes of the t-decay jets (top) and the recoil jets (bottom) of the MC truth ensemble, for events with more than one of the latter. The distribution stars from $\Delta R = 0.4$, since is the cone aperture of a single AK4 jet.

Chapter 5

Analysis strategy and fit simulations to data

In this chapter we describe the analysis strategy developed to extract the W' signal. At a first stage, a series of topological and kinematic selection requirements are applied in order to discriminate the signal from the background. We, subsequently, identified one data-enriched region and two background-enriched ones, in order to properly estimate the background contribution when comparing to data. We remember the considered background in this thesis are the $t\bar{t}$ pair production and the $W+$jets events.

An extended maximum likelihood fit to data is performed on the invariant mass of reconstructed W' candidates for different hypotheses on the W' boson mass, into range $1500 - 4000$ GeV, and chirality of its coupling to the fermions, left or right. In absence of signal evidence, upper limits are derived on the W' production cross section times the branching ratio of the investigated channel decay. Comparing the observed limits with the theoretical expectations, we inferred a lower limit to the mass of W', as well as an upper limit to its production cross section as a function of the candidate W' mass.

5.1 Data and simulated samples

The list of data sets employed in this analysis is reported in table 5.1. The label `MET` refers to the set of triggers used to select the data set, based on the properties of jets and hadronic energy reconstructed in the events at trigger-level; the labels `Run2017B` through `F3` and `Run2018A` through `D` refers to the data taking period, the label like `31Mar2017` refers to the date the data has been reprocessed, the final label `MINIAOD` refers to the file format and content, according to CMS standard definitions. The data set used corresponds to two different data-taking periods:

- 2017 data taking, amounting to an integrated luminosity of 41.53 fb^{-1} known with a relative uncertainty of 2.5%;
- 2018 data taking, amounting to an integrated luminosity of 59.74 fb^{-1} known with a relative uncertainty of 2.3%.

Dataset	Luminosity (fb^{-1})	Year
/MET/Run2017B-31Mar2018-v1/MINIAOD	4.82	2017
/MET/Run2017C-31Mar2018-v1/MINIAOD	9.66	2017
/MET/Run2017D-31Mar2018-v1/MINIAOD	4.25	2017
/MET/Run2017E-31Mar2018-v1/MINIAOD	9.28	2017
/MET/Run2017F-31Mar2018-v1/MINIAOD	13.54	2017
/MET/Run2018A-PromptReco-v2/MINIAOD	14.00	2018
/MET/Run2018B-PromptReco-v2/MINIAOD	7.10	2018
/MET/Run2018C-PromptReco-v2/MINIAOD	6.94	2018
/MET/Run2018D-PromptReco-v2/MINIAOD	31.9	2018

Table 5.1: Details about the datasets used in this thesis.

In order to properly develop a strategy analysis and and to model the background distributions used in the fit to data, we make use of simulations to reproduce signal and main background processes. We used these simulated samples to determine the shape of their distributions in every region we are going to define to perform the fit.

For the backgrounds, the simulations have been performed separately for 2017 and 2018, taking into account the different experimental conditions of LHC and CMS in this two years. For the signal, only simulated samples which are gauged to the 2017 conditions are available up to date.

In 2018 data samples we had to face an instrumental problem, since a module of the detector system failed to properly work from the run 319077. We, therefore, vetoed events reconstructed in the spatial region covered by the aforementioned module, consequently reweighting the background and signal simulated sample.

The simulation of events at the LHC are produced with dedicated algorithms, and in the latter the following three main steps can be identified:

- the hard scattering process,
- the parton-shower of quarks that hadronize,
- the interaction of particles with the detector.

The simulation of such levels it is possible thanks to Monte-Carlo (MC) event generators, that manage to:

- generate matrix elements either at leading order corrections (LO) or at next-to- leading order corrections (NLO), with programs like `Madgraph5_aMC@NLO` [6], `Powheg 2.0` [29] or `CompHEP` [17];
- generate and simulate the hadronization of outgoing particles produced in proton-proton collisions, with the goal to reproduce accurately the event properties of a wide range of processes, such as `Pythia` [54, 55] or Herwig [13];
- simulate particle interaction with CMS sub-detectors, such as `GEANT 4` [5].

The signal event samples with CompHEP with different benchmark models based on the hypothesis made on the W' boson mass, ranging from 1500 up to 4000 GeV with nine points chosen, considering the W' boson right-handed; for the left-handed W', we investigated the samples produced for the mass values of 2900 and 3300 GeV. For all the aforementioned signal samples, the width of the resonance mass of 1%. In table 5.2 are reported the mass points and corresponding theoretical cross section of the W' event samples.

CompHEP is used for the leading order (LO) cross section calculation, which is, then, scaled to next-to-leading order (NLO) using a factor of 1.2 [17].

Mass (GeV)	Cross section (pb)	Chirality
1500	0.46602	RH
1800	0.19584	RH
1900	0.14922	RH
2300	0.053847	RH
2400	0.042275	RH
3000	0.010756	RH
3100	0.008669	RH
3200	0.006997	RH
4000	0.001425	RH
2900	2.5551	LH
3300	2.5672	LH

Table 5.2: Details about the W' simulated samples.

The $t\bar{t}$ pair production process is generated with Powheg 2.0 at NLO accuracy, and also its cross section is calculated at the NLO in perturbative QCD. In order to have a reasonable statistics of this background, two simulated samples are generated, the first where the mass of the $t\bar{t}$ pair $m_{t\bar{t}}$ is in the range of 700 to 1000 GeV, the latter where $m_{t\bar{t}} > 1000$ GeV.

The W+jets events are, instead, generated with Madgraph5_aMC@NLO at LO accuracy, but its cross section is calculated at NLO. The samples are produced for seven several range of the jet transverse energy H_T, reported in the table 5.31.

Process	Cross section (pb)
$t\bar{t}$ (700 GeV $< m_{t\bar{t}} < 1000$ GeV)	80.5
$t\bar{t}$ ($m_{t\bar{t}} > 1000$ GeV)	21.3
W+jets (100 GeV $< H_T < 200$ GeV)	1345
W+jets (200 GeV $< H_T < 400$ GeV)	359.7
W+jets (400 GeV $< H_T < 600$ GeV)	48.91
W+jets (600 GeV $< H_T < 800$ GeV)	12.05
W+jets (800 GeV $< H_T < 1200$ GeV)	5.501
W+jets (1200 GeV $< H_T < 2500$ GeV)	1.329
W+jets ($H_T > 2500$ GeV)	0.03216

Table 5.3: Details about the $t\bar{t}$ pair production and W+jets event simulated samples.

For what concerns the background, in the following analysis strategy the contributions from different simulated samples referring to the same physics process are summed together.

5.1.1 Pile-up reweighting

In order to take into account the presence of additional interactions beyond to the primary one, known as **pile-up (PU)**, that create energy deposits in the tracker and in the calorimeters, CMS computes the number of primary vertices from the instantaneous luminosity. The pile-up distribution in simulation, however, does not reproduce the one observed in data. A reweighting is therefore needed, consisting in applying multiplicative factors to MC sample events. We get, in this way, a correction for the discrepancy with the distribution of primary vertices in the data.

5.2 Event selection

In order to discriminate the signal from the background, a series of topological and kinematic selection requirements, that we call from now **cuts**, are applied, taking into account that the final state considered is semileptonic and what we have in section 4.3.

In the following, we are going to show step by step the sequence of selections we imposed and what have motivated us to apply such cuts.

5.2.1 Trigger

Signal events are characterised by the presence of one energetic neutrino, since, for the W' mass values considered this thesis, we except neutrino coming as the result of a decaying t has energy of some hundreds of GeV, Events are, therefore, required to satisfy trigger conditions based on high values of MET, that, as we know, is the basis of the reconstruction of the neutrino itself. The request consists of the logical `OR` between two conditions, namely the online reconstructed MET has to pass the threshold of either 110 GeV or 120 GeV. This combination of trigger conditions is henceforth referred to as the **hadronic trigger**.

5.2.2 Preselection cuts

Remembering what are the final states produced by the investigated decay channel of the W', we imposed the following preselection rules, in order to make the QCD events contribution to decrease significantly:

- one "HighPt" lepton, electron or muon, with $p_T > 100$ GeV;
- two or more jets with $p_T > 30$ GeV and $|\eta| < 2.5$;

In order to discriminate events from t quark-antiquark pairs, an additional selection can be performed studying the collected AK8 jets, aiming at selecting b-jets that are not originating from hadronic top quark decay chains. For every AK4 jet, we do associate the fat one which was the closest to the first between the reconstructed AK8 jets. In this way, it is possible to use some parameters, that compare the kinematic and structure features of the AK8s to the corresponding one of the AK4s, to eventually discriminate the signal from the background.

After we applied the trigger and preselection cuts, we have to reconstruct the W' ensemble starting from the preselect physical object and, then, properly choose which ensemble is the best W' candidate.

The leading jet p_T is a quantity that is a little bit hard to be reproduced in the simulations at high energy range and, then, the simulated spectra of it can have discrepancies with the real ones. This fact explains why, as we can see in fig. 5.1, data and MC simulations do not agree so much.

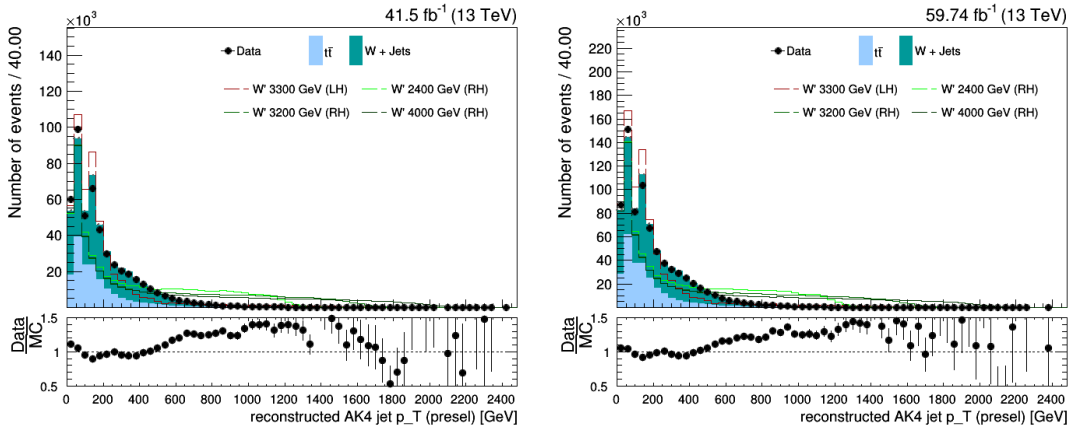


Figure 5.1: AK4 jet p_T distribution, after preselection, for the 2017 (left) and 2018 (right) data taking. Countings are normalised to total data yield.

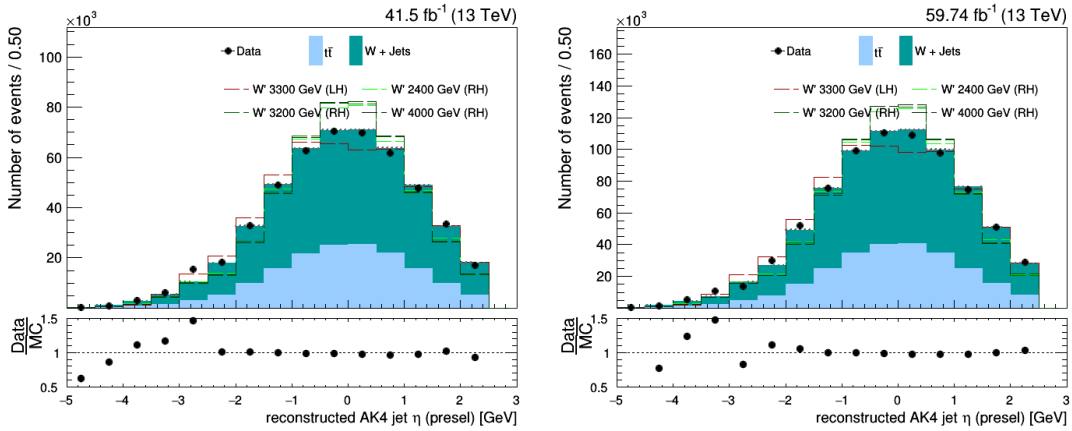


Figure 5.2: AK4 jet η distribution, after preselection, for the 2017 (left) and 2018 (right) data taking. Countings are normalised to total data yield.

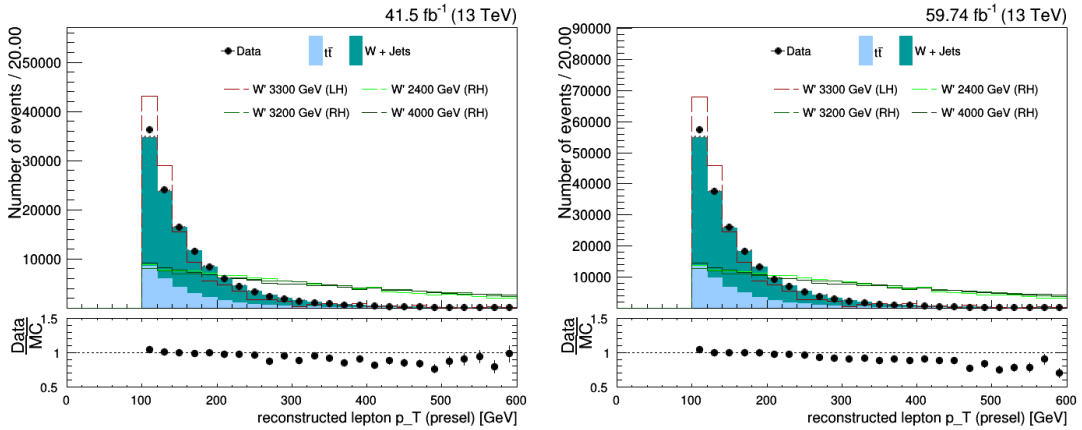


Figure 5.3: Lepton p_T distribution, after preselection, for the 2017 (left) and the 2018 (right) data taking. Countings are normalised to total data yield.

5.3 Reconstruction of the best W' candidate

If we have more than two jets being selected after the aformentioned cuts, more than one W' ensemble can be formed and, as we mentioned in the previous chapter, we have to find an algorithm capable of choosing the best W' candidate reconstructed with the experimentally available physical objects.

In order to achieve such goal, the first step is to properly reconstruct the prompt t quark, as shown in 4.2.5, coupling the right AK4 jet to the leptonic couple $l\nu_l$. We tested two such association methods two ways to choose the latter:

- t **mass** criterion: we choose the resolved jet that, coupled to the leptonic set, reconstructs the t candidate whose invariant mass is the closest to the t mass itself $m_t = 173.1$ GeV; the distance with respect to m_t has been evaluated with the χ^2 metrics. Since the lepton, the neutrino and the b-originated jet are all together products of the t quark decay, the sum of their 4-momenta must have invariant mass close to m_t.

- **closest** ΔR **criterion**: we choose the resolved jet that, coupled to the leptonic set, reconstructs, in the way shown in sec. 4.2.5, the t candidate closest to the MET in the transverse xy plane. Since the t decay is boosted in the energy range of the investigated process, we expect leptons and t-decay jet to be very close to each other.

After performing this reconstruction, we have to associate to the t quark the right AK4 jet, between the remaining ones, to obtain finally the ultimate W' candidate. Also in this case, we made up two possible reconstruction strategies:

- **furthest** ΔR **criterion**: we choose the resolved jet whose angular distance in the xy plane with MET is the closest to π to be originating from b of the W' direct decay. As the collision between protonic partons is longitudinal, we expect the total transverse momentum to be zero. In order to respect the momentum conservation law, and remembering the boosted topology of the t decay process, the t and b quarks are approximately produced back to back at the primary vertex.

- **maximum** p_T criterion: we choose, between the remaining AK4 jets, the one with the highest transverse momentum. Since the searched AK4 jet is produced by the prompt b quark, indeed, it will be the most energetic in the collection of the reconstructed resolved jets.

Combining the previous criteria for the reconstruction of the t and W' vertices, we obtain three methods to globally reconstruct the W', in order to get the best W' candidate ensemble:

- **tMass** algorithm:
  - t mass criterion for the t vertex and p_T criterion for the W' vertex;
- **one**ΔR algorithm:
  - ΔR criterion for the t vertex and p_T criterion for the W' vertex;
- **both**ΔR algorithm:
  - ΔR criterion for both the t vertex and W' vertex.

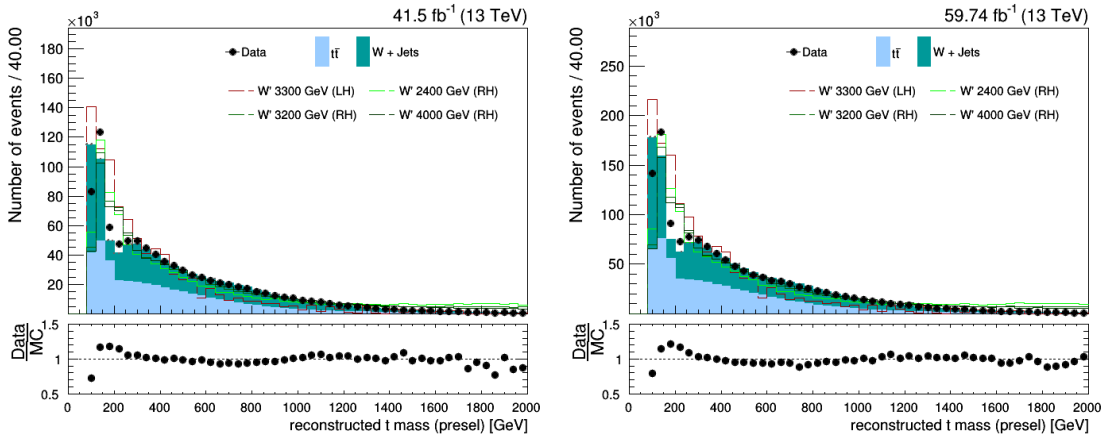


Figure 5.4: Reconstructed t quark mass, after the preselection is performed, for the 2017 (left) and the 2018 (right) data taking. Countings are normalised to the total data yield.

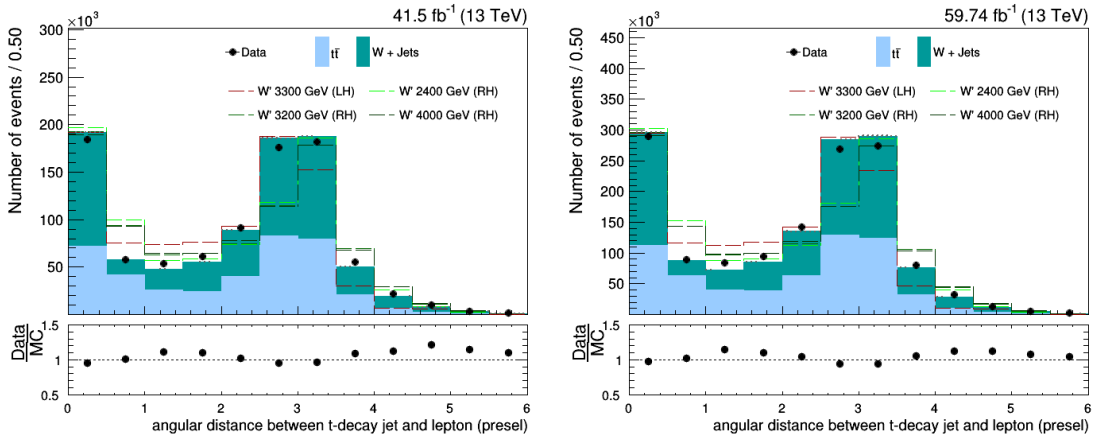


Figure 5.5: Reconstructed angular distance between t-decay AK4 jet and lepton distribution, after the preselection is performed, for the 2017 (left) and the 2018 (right) data taking. The corresponding distribution for the prompt b AK4 jet. Countings are normalised to the total data yield.

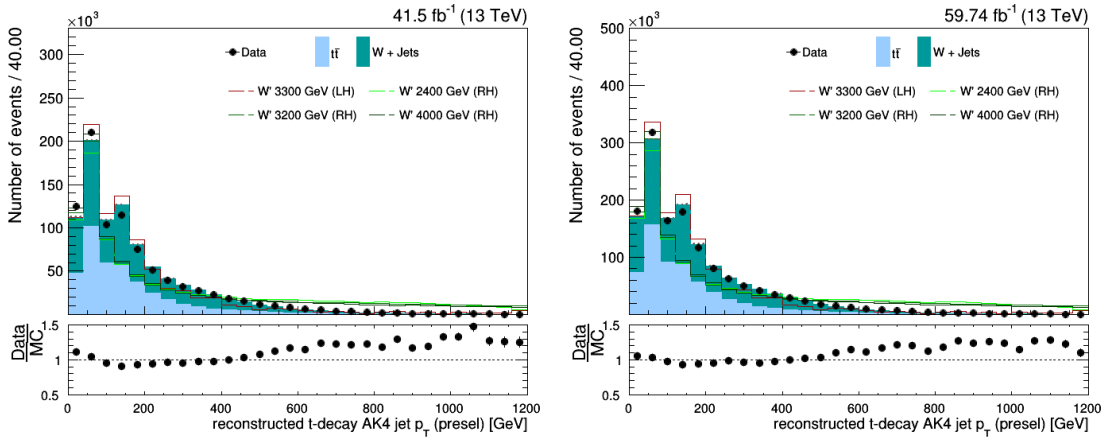


Figure 5.6: Reconstructed t-decay AK4 jet mass distribution, after the preselection is performed, for the 2017 (left) and the 2018 (right) data taking. Countings are normalised to the total data yield.

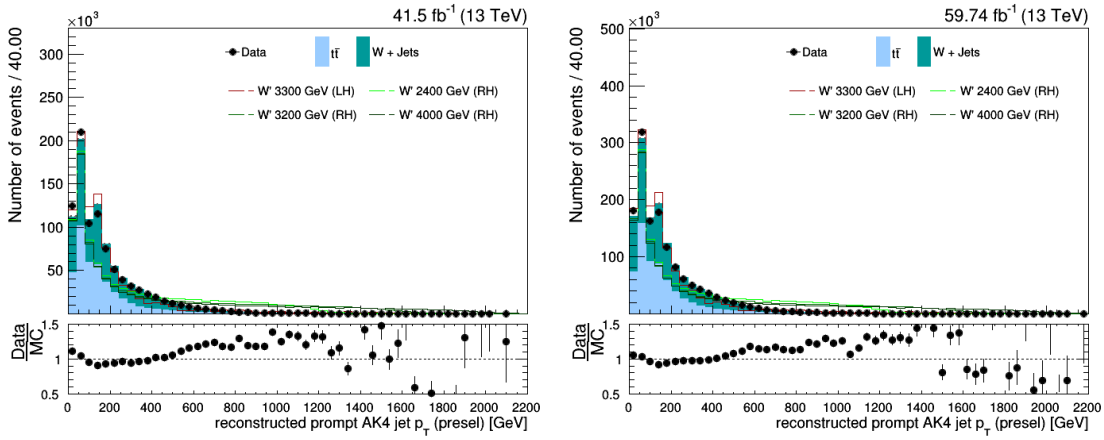


Figure 5.7: Reconstructed prompt b AK4 jet mass distribution, after the preselection is performed, for the 2017 (left) and the 2018 (right) data taking. Countings are normalised to the total data yield.

As we explained in sec. 4.3.1, the MC truth help us to investigate the performance of the ensemble reconstruction algorithms that use only the experimentally available information. We, therefore, compared the invariant mass spectra reconstructed by these three algorithms with the ideal one, reproduced with the MC truth ensembles. We performed this comparison using the simulated signal data, reporting the results for the samples with $m_{W'} = 3200, 4000$ GeV as examples in fig. 5.8.

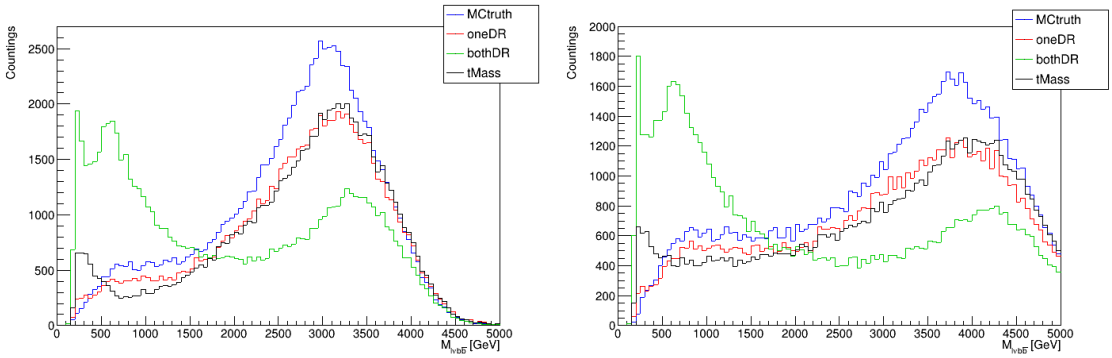


Figure 5.8: W' candidate mass distribution, reconstructed with MCtruth, oneΔR, bothΔR and tMass algorithms. Top, simulated samples with $m_{W'} = 3200$ GeV, bottom, with $m_{W'} = 4000$ GeV. The first is the benchmark for the other three.

In this way, we figured out the oneΔR algorithm is the one which mostly reproduce the ideal invariant mass spectrum of W' candidate. We can observe, indeed, it best reproduces the ideal spectrum of the MCtruth ensemble in all the considered invariant mass range. In consequence of this result, we chose this algorithm to obtain the best W' candidate for all the simulated signal samples. We applied this algorithm, in addition, also to the the simulated background sample and, before the fitting procedure, to the data samples, in order to detect all the possible events compatible with the searched W' signal and properly reconstruct the best candidate.

From now, we call the ensembles chosen the oneΔR algorithm briefly **one ΔR candidate**.

5.3.1 Efficiencies calculated for the cuts performed so far

At this point we can start to compute the efficiencies of the selection rules, applied so far to simulated samples, with respect to the total number of events per sample, nEvents. It is possible to appreciate the difference in efficiency depending on the physics processes, each taken at a characteristic energy regime of interest for this analysis.

From fig. 5.9 we can observe the considered cuts are more efficient in selecting signal-like events than in picking the background-like ones. That means the selections imposed so far properly reduce the contribution of the background in the selected collection of the oneΔR ensembles.

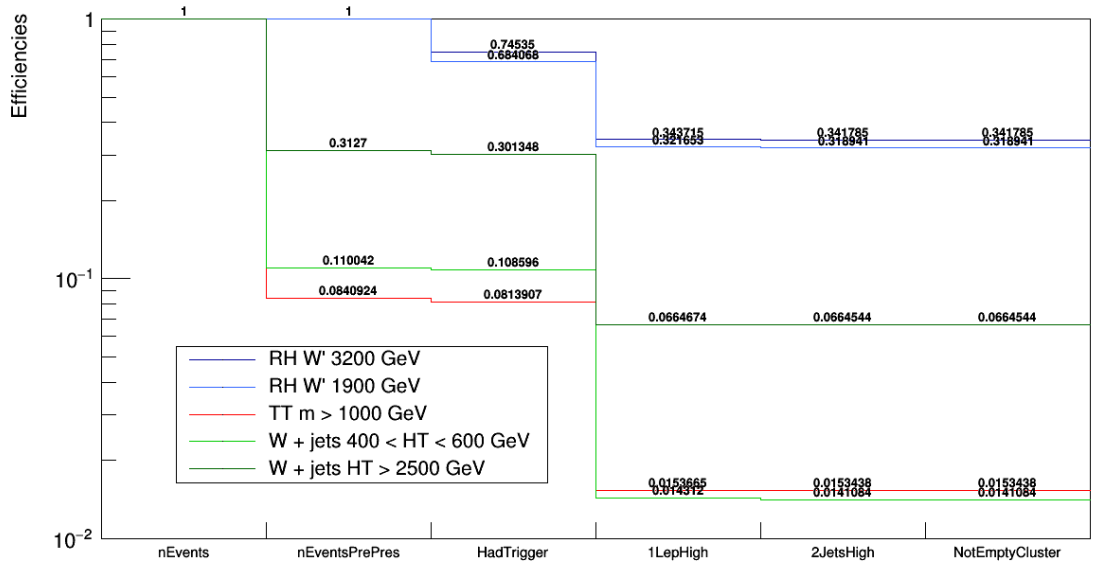


Figure 5.9: Cut efficiencies calculated, from the left to the right, for the quality of the data recorded by the detector (nEventsPrePres), hadronic trigger, selection of events with one only lepton and more than one AK4 jet. NotEmptyCluster is a sanity check performed after the oneΔR ensemble is found in an event.

5.3.2 Selection rules for the oneΔR candidates

For the results obtained so far, the oneΔR algorithm is the best method to reconstruct as much as possible ideal W' candidate ensemble, starting only from the experimentally available information. The selection of the best W' candidate alone, by the way, does not by itself reduce the background contamination, whose efficiencies can be seen in fig. 5.9.

In order to eventually reduce the amount of the background events simulating the signal, we applied other cuts to the kinematic properties of the involved lepton and AK4 jets. For the latter, we also considered the features related to the associated surrounding AK8 jet. We imposed, therefore, the following conditions:

- ΔR between lepton and t-decay jet > 0.4, as we require lepton to not come from a hadronic jet;
- p_T(AK4)/p_T(AK8) > 0.9 for recoil b jet, we call this quantity **RelPt**;
- recoil b jet $p_T > 600$ GeV.

In figg. 5.10 and 5.11 we show the distributions of the latter two quantities just listed for backgrounds and some signal samples. The spectra are normalised to the total yield, in order to better appreciate the distribution shapes.

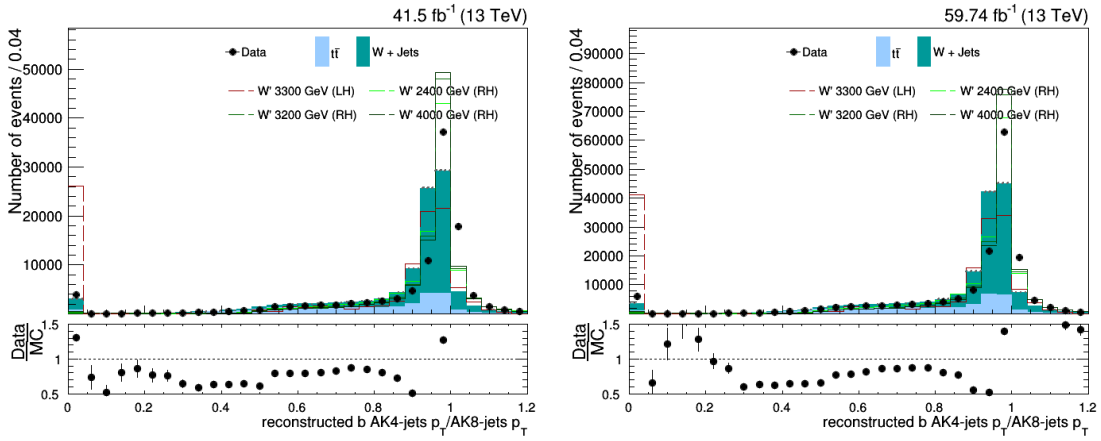


Figure 5.10: Recoil b jet RelPt distribution, reconstructed for the oneΔR ensembles collected the 2017 (left) and the 2018 (right) samples.

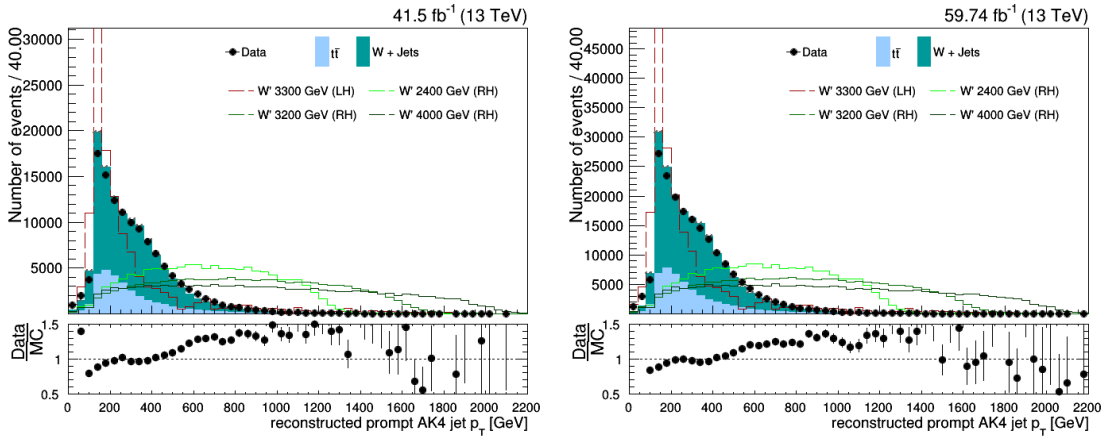


Figure 5.11: Recoil b jet p_T distribution, reconstructed for the oneΔR ensembles collected in the 2017 (left) and the 2018 (right) samples.

5.4 Fitting procedure

5.4.1 Preparing the fitting regions

In our study we define three regions by counting the AK4 jets classified as coming from a B hadron clustering, according to the flavour tagging DeepCSV algorithm (cfr. sec. 4.2.4). Fixing the working point of the latter at the Medium threshold, we have the following fitting regions:

1. **Background Enriched Region (BER)**, defined as containing the oneΔR ensembles with zero b-tagged jets;
2. **Background-Signal Mixed Region (BSMR)**, defined as containing the oneΔR ensembles with one b-tagged jets;
3. **Signal Enriched Region (BER)**, defined as containing the oneΔR ensembles with two b-tagged jets.

The above naming convention can be easily understood if one observes the efficiencies, calculated with respect to the total number of the oneΔR ensembles, of the regions for both backgrounds and signal MC samples, as reported in 5.12.

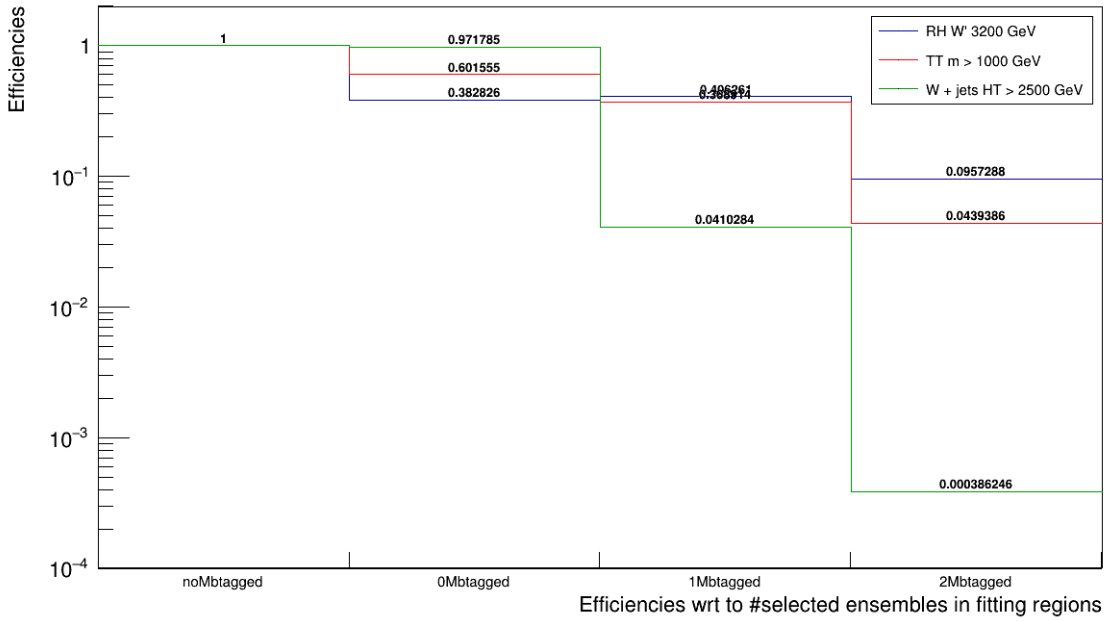


Figure 5.12: Efficiencies for different number of b-tagged jets per oneΔR ensemble, normalised to the total number of the latter. noMBtagged corresponds to total amount of selected oneΔR ensembles, 0Mbtagged to BER, 1Mbtagged to BSMR, 2Mbtagged to SER.

We can see, indeed, the efficiencies of background selection decrease with the number of b-tagged jets more rapidly than the ones of the signal selection. In this way, The SER becomes the one with the best signal-to-background ratio.

As background and signal distributions we do take the binned distributions or “templates”, taken from simulation. No systematic effect is considered on such

shape, but the normalisation of the backgrounds is constrained across the three regions and left floating as free parameter in the fit. This allows us to adjust for any normalisation effect. Residual systematics effects can modify the signal and background shapes, as well as the relative normalisation in between the regions. Such effects are not considered in this analysis, but are expected to be subdominant with respect to the statistical uncertainty in the signal region.

In figg. 5.13, 5.14 and 5.15 are reported the aforementioned mass invariant distributions, properly normalised to the data yield of the corresponding taking period. The contributions for the backgrounds are, for every process, merged together, taking into account how the cross section of the process itself changes with energy, and, then, stacked all together.

We can observe the data distributions almost replicate the total background ones, except in SER for the 2017 data taking period. In the latter we find an overestimate of the background contribution. That might be a peculiar effect of the selection rules imposed so far, and it is worth to be more profoundly investigated in future studies.

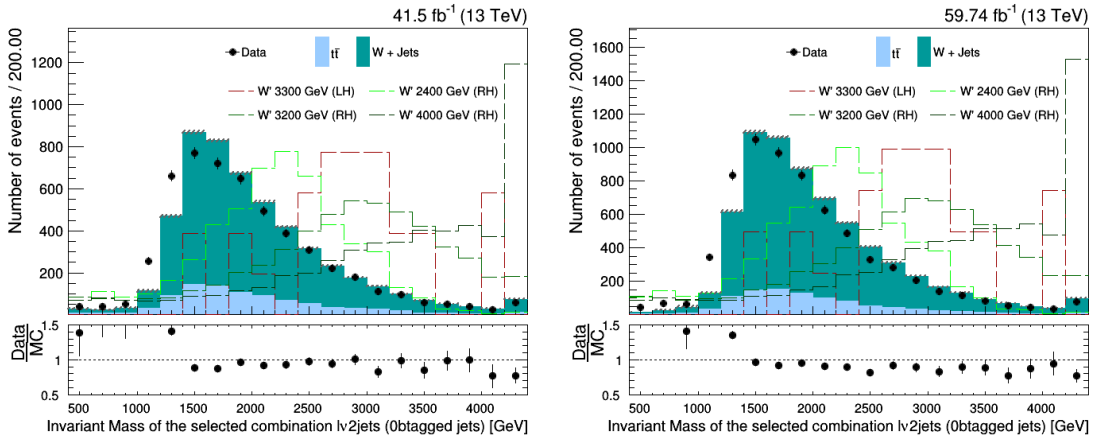


Figure 5.13: Invariant mass of the oneΔR ensembles for backgrounds, some signal mass points and data (left, 2017 and right, 2018) in BER. All the samples are normalised to the corresponding total data yield.

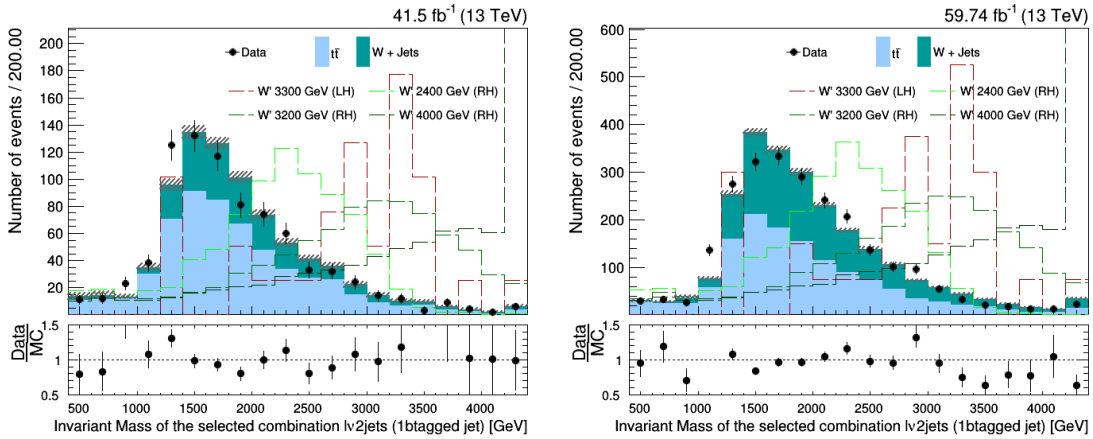


Figure 5.14: Invariant mass of the oneΔR ensembles for backgrounds, some signal mass points and data (left, 2017 and right, 2018) in BSMR. All the samples are normalised to the corresponding total data yield.

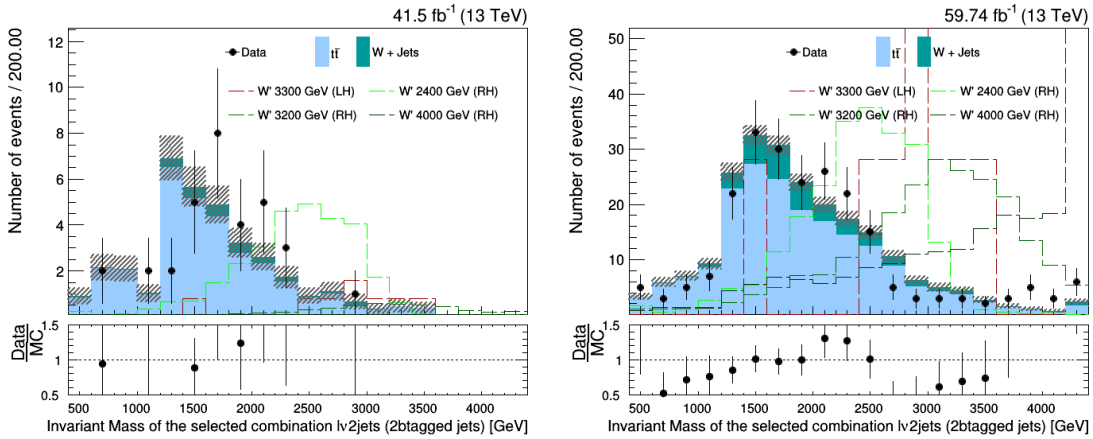


Figure 5.15: Invariant mass of the oneΔR ensembles for backgrounds, some signal mass points and data (left, 2017 and right, 2018) in SER. All the samples are normalised to the corresponding total data yield.

5.4.2 How to perform the fit

After we have optimised the event selection, in order to maximise the contribution of the signal-like events and the minimise that of the background-like ones, we can test the existence of the W' boson via the **Extendend Binned Maximum Likelihood (EBML)** fit method. We are going to perform this procedure simultaneously in three aforementioned fitting regions both for the 2017 and the 2018 data samples.

In order to perform the existence test, we have to discriminate between two hypotheses:

- $\mathcal{H}_0$, absence of new physics, namely the signal is absent or too small to be detected;
- $\mathcal{H}_1$, detection of signals produced by new physics.

For every hypothesis we have a corresponding expected distribution for the $m_{W'}$ spectrum. If we define, then, λ as the expected number of event associated to each one of the expected distribution, it can assume the following values depending which hypothesis is true:

$$\mathcal{H}_0 \Rightarrow \lambda = b$$
$$\mathcal{H}_1 \Rightarrow \lambda = \mu_{sig} s + b$$

where b is the expected total yield of the backgrounds, provided by the MC samples, s the corresponding value for the signal and μ_{sig} is called **signal strenght**. Varying this quantity, we can optimise the signal contribution to the fit of data, changing consequently the signal yield. The value $\mu_{sig} = 1$ corresponds to the case of $\mathcal{H}_1$, $\mu_{sig} = 0$ corresponds to $\mathcal{H}_0$.

Since we also want the background yields to vary for the same purpose, we parameterise b in a way similar to the one applied for the signal yield:

$$b = \beta_{t\bar{t}} + \beta_{Wj} = \mu_{t\bar{t}} b_{t\bar{t}} + \mu_{Wj} b_{Wj}$$

We, therefore, separated the background total yield in its two components $\beta_{t\bar{t}}$ and β_{Wj}, depending by the $t\bar{t}$ and W+jets processes. We also defined the $t\bar{t}$ *strenght* $\mu_{t\bar{t}}$ and W+*jets strenght* μ_{Wj}, with $b_{t\bar{t}}$ and b_{Wj} the corresponding expected number of events calculated by the MC samples.

As we did not consider systematic effects on the distributions obtained via MC simulations, we don't need to introduce any dependence by nuisance parameters for the expected number of events listed before. Between the aforementioned three strengths, μ_{sig} is the one discriminating between the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_0$ and, then, we will report the results of the fit in terms of signal strength.

5.4.3 Extended Binned Maximum Likelihood fit method

The EBML fit method is based on the maximisation of the extended likelihood function $\mathcal{L}(\mathbf{m}|\mu_{sig}, \mu_{t\bar{t}}, \mu_{Wj})$ of the observed data sample $\mathbf{m}$, with cardinality n_{events} [43], defined as follows for our study:

$$\mathcal{L}(\mathbf{m}|\mu_{sig}, \mu_{t\bar{t}}, \mu_{Wj}) = \text{Poisson}\big(n_{\text{obs}}, \lambda\big) \prod_{k=1}^{n_{\text{events}}} f(m_k|\mu_{sig}, \mu_{t\bar{t}}, \mu_{Wj}) \tag{5.4.1}$$

where m_k is the invariant mass $m_{W'}$ value for the oneΔR ensemble selected for the k^{th} event, and

$$f(m_k|\mu_{sig},\mu_{t\bar{t}},\mu_{Wj}) = \frac{\mu_{sig}s}{\lambda} f_{sig}(m_k) + \frac{\mu_{t\bar{t}}b_{t\bar{t}}}{\lambda} f_{t\bar{t}}(m_k) + \frac{\mu_{Wj}b_{Wj}}{\lambda} f_{Wj}(m_k)$$

is the combined probability density function (briefly, PDF) of the invariant mass distribution that takes into account contributions both from the signal and the background, evaluated in m_k and normalised to λ. $f_{sig}, f_{t\bar{t}}$ and f_{Wj} are the single PDFs for signal, $t\bar{t}$ e W+jets, respectively.

The signal distribution, as we said before, depends on the $m_{W'}$ value and, then, the fitting procedure has to be performed for every $m_{W'}$ value considered to produce the simulations.

With this approach, we perform a so-called shape analysis of the observed invariant mass $m_{W'}$ distribution, fitting it with the signal and backgrounds ones obtained thanks to the corresponding MC samples. This fit is performed for every mass point used to produce the signal simulated samples, obtaining, in this way, an estimate of μ_{sig} for every considered W' boson mass in producing the MC samples.

5.4.4 Hypothesis test and limit extractions

In order to verify which hypothesis is true and which is rejectable, we have to perform a quantitative hypothesis test. The latter use an appropriate random variable, called **test statistics**, and its PDF to evaluate how the observed set of quantities $\mathbf{m}$ are consistent with $\mathcal{H}_0$ or $\mathcal{H}_1$ [25, 42]. A very performing test statistic is the one defined by the Neyman-Pearson lemma, built with the help of the likelihood function of the observed data sample as follows:

$$\lambda(\mu_{sig}) = \frac{\mathcal{L}(\mathbf{m}|\mathcal{H}_1)}{\mathcal{L}(\mathbf{m}|\mathcal{H}_0)} = \frac{\mathcal{L}_{s+b}(\mathbf{m})}{\mathcal{L}_b(\mathbf{m})}$$

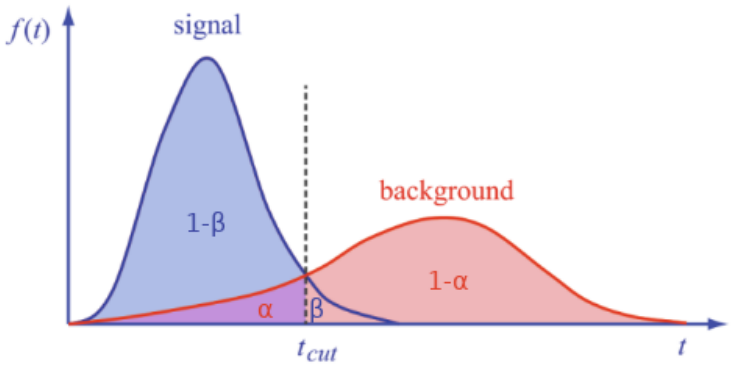


Figure 5.16: The pdf $f(t)$ of a test statistics t, corresponding to the $\lambda(\mu_{sig})$ defined before, in case of signal ($\mathcal{H}_1$ true) in blue and in case of background ($\mathcal{H}_0$ true) in red. t_{cut} is the fixed threshold to discriminate between the hypotheses. I-type error α and II-type error β are also represented for the sake of clarity.

It turns out to be the most powerful discriminator, since, with fixed significance level α, i.e. the probability to reject $\mathcal{H}_0$ if it is true, minimises the type-II error β, i.e. the probability of rejecting $\mathcal{H}_1$ if $\mathcal{H}_1$ is true[1] [42].

When is verified the case in which $\mathcal{H}_0$ is true, $\mathcal{H}_0$ is rejected if the ratio is smaller than a chosen constant value λ_{cut}, depending on the chosen value for α. Taking into account of what we have said so far, we can define as test statistics, equivalent to $\lambda(\mu_{sig})$, the profile likelihood ratio q_μ in the following way:

$$q_{\mu_{sig}} = -2\ln\lambda(\mu_{sig}) = -2\ln\frac{\mathcal{L}_{s+b}(\mathbf{m})}{\mathcal{L}_s(\mathbf{m})}$$

Writing the likelihood function as reported in eq. (5.4.1), this quantity can be written down as follows:

$$q_{\mu_{sig}} = 2\left[\mu_{sig}s - \sum_{i=k}^{n_{\text{events}}}\ln\left(\frac{\mu_{sig}sf_s(m_k)}{\mu_{t\bar{t}}b_{t\bar{t}}f_{t\bar{t}}(m_k) + \mu_{Wj}b_{Wj}f_{Wj}(m_k)} + 1\right)\right]$$

Data samples with $q_{\mu_{sig}} \geq 0$ appear to be consistent with the hypothesis $\mathcal{H}_0$ of background only, while the ones with $q_{\mu_{sig}} \leq 0$ are compatible with the hypothesis $\mathcal{H}_1$ of background plus signal.

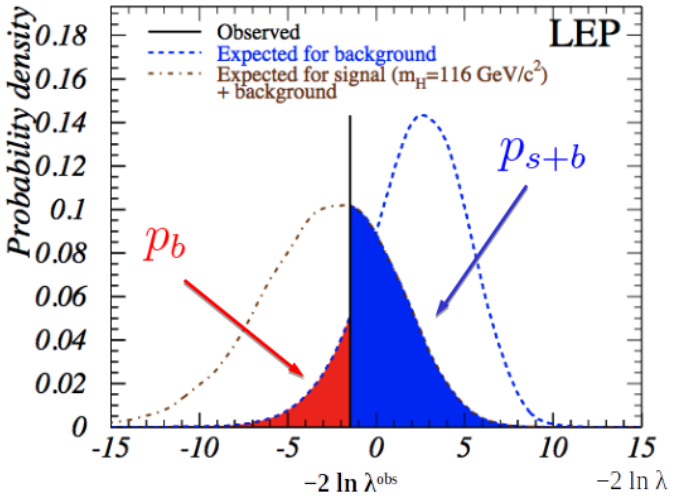


Figure 5.17: Example of evaluation of CL_s from pseudoexperiments performed for the LEP search of the Higgs boson. The distribution of the test statistics $q_{\mu_{sig}}$ is shown in blue assuming the $\mathcal{H}_1$ hypothesis and in red assuming the $\mathcal{H}_0$ one [42].

If a hypothesis test rejects $\mathcal{H}_1$, in any case is possible to get upper limits on the discriminating parameter, in our case μ_{sig}, in order to quantify the level of incompatibility of data with a signal hypothesis. First of all, we obtain the distribution of $q_{\mu_{sig}}$, when μ_{sig} varies, considering $\mathcal{H}_1$ is true and, then, calculate the $p-$value as follows:

- if $\mathcal{H}_1$ is true, it is the probability that the considered random quantity $q_{\mu_{sig}}$ assumes a value greater or equal to the observed one $q^{obs}_{\mu_{sig}}$, we call it $p_{s+b}(\mu_{sig})$;

[1] or, equivalently, of not rejecting $\mathcal{H}_1$ if $\mathcal{H}_0$ is true.

- if $\mathcal{H}_0$ is true, it is the probability that $q_{\mu_{sig}}$ assumes a value smaller or equal to $q^{obs}_{\mu_{sig}}$, we call it $p_b(\mu_{sig})$.

The just defined p_{s+b} and p_b, therefore, are used in the **modified frequentist approach**, also called **CL**$_s$ [49], to calculate the confidence level for the signal as the ratio of $p-$values:

$$\mathrm{CL}_s(\mu_{sig}) = \frac{p_{s+b}(\mu_{sig})}{1 - p_b(\mu_{sig})} \tag{5.4.2}$$

Upper limits at CL $= 1 - \alpha$ are determined excluding the range of parameters of interest for which CL$_s(\mu_{sig})$ is lower than the fixed α. In case of background only, the upper limit is a random variable that depends on the observed data sample and its distribution can be predicted using MC simulation [42]. In our study we follow the common choice in CMS collaboration, and widely used across the LHC experiments, to impose $\alpha = 0.05$.

5.5 Fit results and limits

Fitting the simulated distribution to 2017 and 2018 data samples in the way described in the previous section, no evidence for presence of signal is found. In consequence of that, we have to reject the $\mathcal{H}_1$ hypothesis. We decided, then, to extract upper limits for the signal strength μ_{sig} with the CL$_s$ method for every experimental mass value of W' boson, $m_{W'}$, both for RH and LH chiralities. We also evaluated the signal strength PDF in absence of signal.

In figg. 5.18$-$5.20 we show, by way of example, the results of the fit in all the three fitting regions for $m_{W'} = 3200$ GeV and RH coupling. We can observe the effects of the background overestimate, noted in sec. 5.4.1, in the fitted plots.

If we call s_{obs} the observed signal yield, s_{th} the expected one, obtained with the MC samples, and L the integrated luminosity of the considered data sample, we have:

$$\mu_{sig} = \frac{s_{obs}}{s_{th}} = \frac{L\sigma_{obs}}{L\sigma_{th}} = \frac{\sigma_{obs}}{\sigma_{th}} \tag{5.5.1}$$

where σ_{obs} the observed cross section of the investigated process, σ_{th} the expected one. The upper limits on μ_{sig}, therefore, can also be interpreted as upper limits on the ratio σ_{obs}/σ_{th}. Once fixed the set of free parameters of theory which is the basis for the MC simulation, and $m_{W'}$ is part of it, if $\mu_{sig} < 1$ the aforementioned set can be excluded with the same CL used to extract the μ_{sig} limit. Since in our study $m_{W'}$ is the only parameter free to vary, we can exclude the mass ranges in which $\mu_{sig} < 1$, imposing in this way a limit also for the mass of the W' boson. In table 5.1 we report the observed and expected limits obtained for μ_{sig} at 95% CL. These limits are calculated with a dedicated software, the **Combine tool** [24].

We subsequently performed the fit procedure also on a dataset obtained combining the 2017 and 2018 collections, with integrated lumonisity equals to 101.24 fb^{-1}, in order to exploit the statistical advantages originated by a larger sample. At this stage, we separately introduced background and signal strengths and fitting regions for each year, resulting, in this way, in one set of strengths and on set

of fitting regions per data taking period. We, therefore, performed the fit simultaneously in all the above defined regions, in order to get results which take into account 2017 and 2018 data. The observed and expected upper limits on cross sections obtained with the combined dataset are reported in tables 5.2.

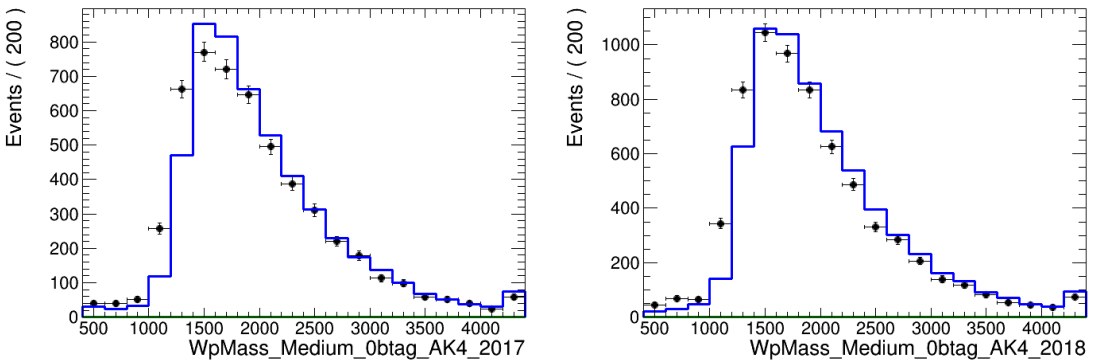


Figure 5.18: Fit results in BER. In blue, the overall distribution fitted to 2017 (left) and 2018 (right) data samples. As no signal (green) evidence is found, the background distribution (red) is superimposed to the overall one.

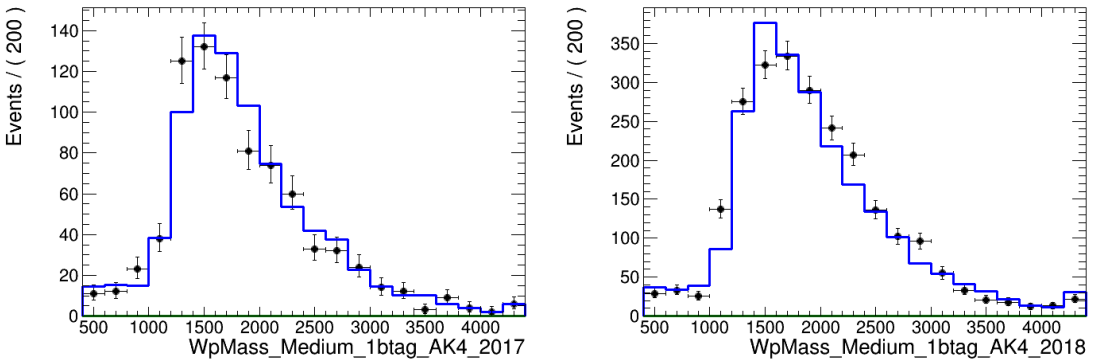


Figure 5.19: Fit results in BSMR. In blue, the overall distribution fitted to 2017 (left) and 2018 (right) data samples. As no signal (green) evidence is found, the background distribution (red) is superimposed to the overall one.

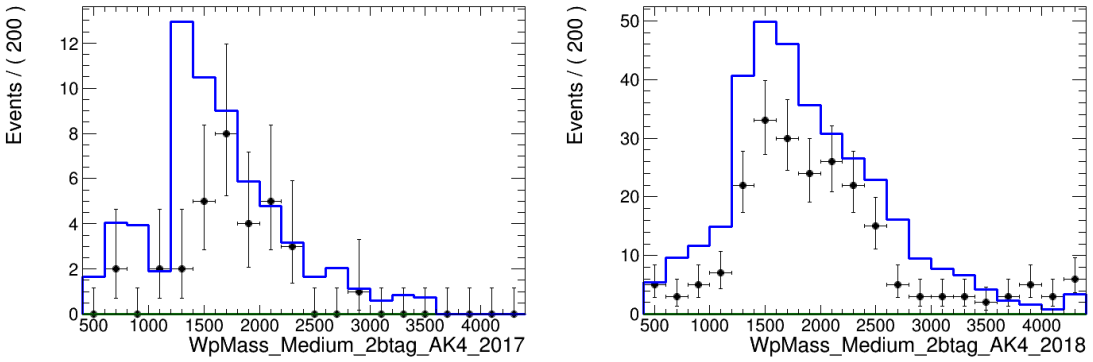


Figure 5.20: Fit results in SER. In blue, the overall distribution fitted to 2017 (left) and 2018 (right) data samples. As no signal (green) evidence is found, the background distribution (red) is superimposed to the overall one.

		2017		2018	
Chirality	Mass (GeV)	Observed	Expected	Observed	Expected
RH	1500	0.0710	0.0962	0.1658	0.1343
RH	1800	0.0737	0.1313	0.2209	0.2207
RH	1900	0.0816	0.1499	0.3003	0.2607
RH	2300	0.1182	0.2568	0.3731	0.4590
RH	2400	0.1343	0.2949	0.3648	0.5059
RH	3000	0.3591	0.6895	0.5639	1.1211
RH	3100	0.4491	0.8086	0.7370	1.3242
RH	3200	0.5027	0.8867	0.8645	1.4805
RH	4000	2.0917	2.8828	3.4245	5.0000
LH	2900	0.3300	0.7012	0.7152	0.8711
LH	3300	0.2852	0.5410	0.3299	0.9805

Table 5.1: Upper limits obtained varying the hypothetical value of $m_{W'}$, both for RH and LH coupling, separately reported for 2017 and 2018 datasets.

		Combined	
Chirality	Mass (GeV)	Observed	Expected
RH	1500	0.0656	0.0776
RH	1800	0.0667	0.1118
RH	1900	0.0788	0.1284
RH	2300	0.1022	0.2217
RH	2400	0.1101	0.2510
RH	3000	0.2407	0.5723
RH	3100	0.3107	0.6699
RH	3200	0.3567	0.7324
RH	4000	1.5269	2.3828
LH	2900	0.2346	0.4473
LH	3300	0.1746	0.5605

Table 5.2: Upper limits obtained varying the hypothetical value of $m_{W'}$, both for RH and LH coupling, reported for the combined dataset.

In figg. 5.21, 5.22 and 5.25 we show graphically the observed and expected upper limits for the ratio σ_{obs}/σ_{th}, obtained as indicated in sec. 5.4.4. We also report the limits directly on the cross section for each considered mass value of the W' in figg. 5.23, 5.24 and 5.26 for every analysed dataset.

Excluding the mass ranges for which the upper limits are less than one, we can extract the lower limits on $m_{W'}$ for the RH coupling, reported in table 5.3.

Year	Observed (GeV)	Expected (GeV)
2017	3450	3245
2018	3242	2882
Combined	3640	3330

Table 5.3: Lower limits obtained for $m_{W'}$, RH coupling, separately reported for 2017, 2018 and combined datasets.

We have to do the following considerations about the reported results:

- The observed σ_{obs}/σ_{th} upper limits are systematically lower than the expected one at high masses for 2018 and the combined dataset, and in all the investigated range for 2017. This likely is an effect of the overestimate of the background in the corresponding MC simulations found in SER, as reported in sec. 5.4.1. Since this region is the one where the fit procedure is more sensitive to the signal, a background overestimate in SER induces an underestimate of the signal yield, as the total, namely background plus signal, yield is fixed by the data and cannot vary. This effect, remembering the eq. (5.5.1), systematically lowers the upper limit on the signal strength μ_{sig}. In order to face this systematic effect, we suggest to properly investigate the cited overestimate, involving a data-driven background estimate and introducing some systematics.

- The 2017 upper limits are lower than the 2018 ones, but we would expect the viceversa to happen, since the 2018 luminosity is higher than the 2017 one. This effect could be induced by the fact we used signal simulated samples, for both the data taking periods, tuned only on the 2017 detector conditions and, additionally, also the physics object definitions and selections requirements were optimised on 2017 data. That obviously negatively affect the performance of the analysis strategy in fitting the 2018 data samples. One of the improvements to this analysis is, therefore, to reperform the analysis for the 2018 datasets as soon as any 2018-suited signal simulations are available.

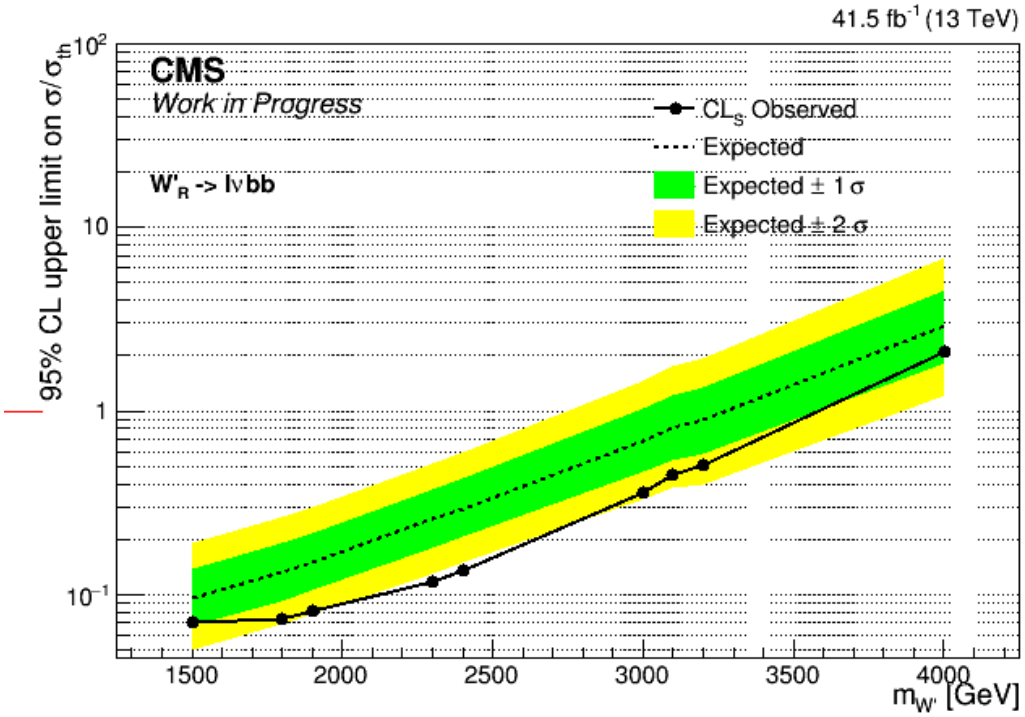


Figure 5.21: Expected and observed upper limits on the ratio $\sigma/\sigma_{th} = \mu_{sig}$ in function of $m_{W'}$, RH coupling, extracted with the 2017 dataset. The green and yellow stripes represent, respectively, the 1 and 2 standard deviations of the statistical fluctuations of the expected upper limit, evaluated with MC simulations.

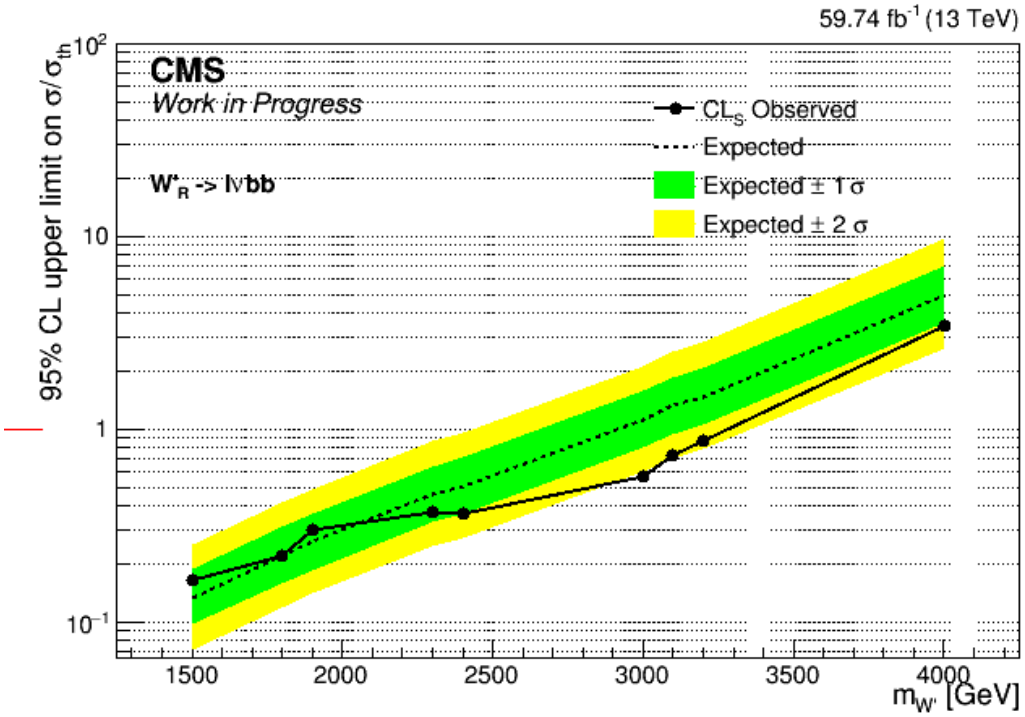


Figure 5.22: Expected and observed upper limits on the ratio $\sigma/\sigma_{th} = \mu_{sig}$ in function of $m_{W'}$, RH coupling, extracted with the 2018 dataset. The green and yellow stripes represent, respectively, the 1 and 2 standard deviations of the statistical fluctuations of the expected upper limit, evaluated with MC simulations.

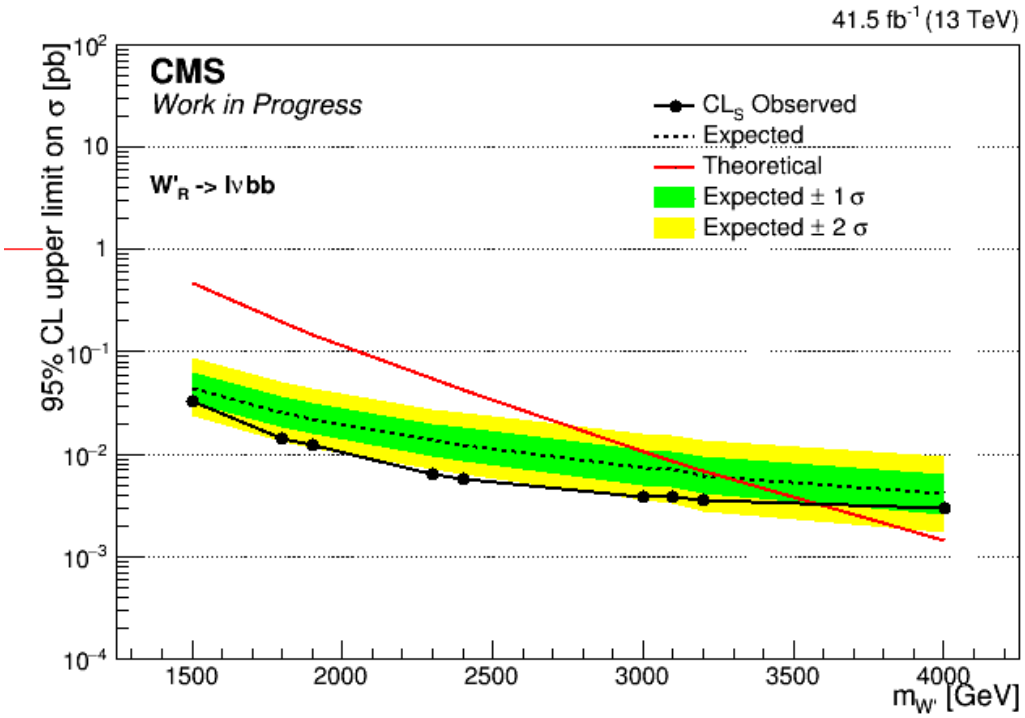


Figure 5.23: Expected and observed upper limits on σ in function of $m_{W'}$, RH coupling, extracted with the 2017 dataset. The green and yellow stripes represent, respectively, the 1 and 2 standard deviations of the statistical fluctuations of the expected upper limit, evaluated with MC simulations. The red line is the cross section predicted by the model in the MC signal samples.

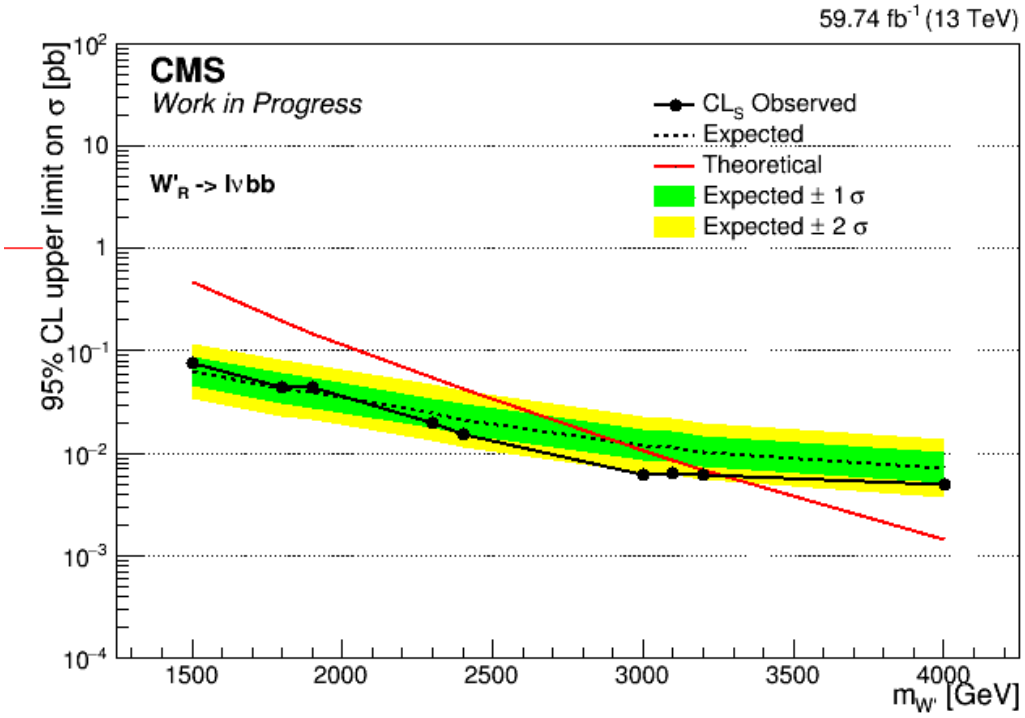


Figure 5.24: Expected and observed upper limits on the ratio $\sigma/\sigma_{th} = \mu_{sig}$ in function of $m_{W'}$, RH coupling, extracted with the 2018 dataset. The green and yellow stripes represent, respectively, the 1 and 2 standard deviations of the statistical fluctuations of the expected upper limit, evaluated with MC simulations. The red line is the cross section predicted by the model in the MC signal samples.

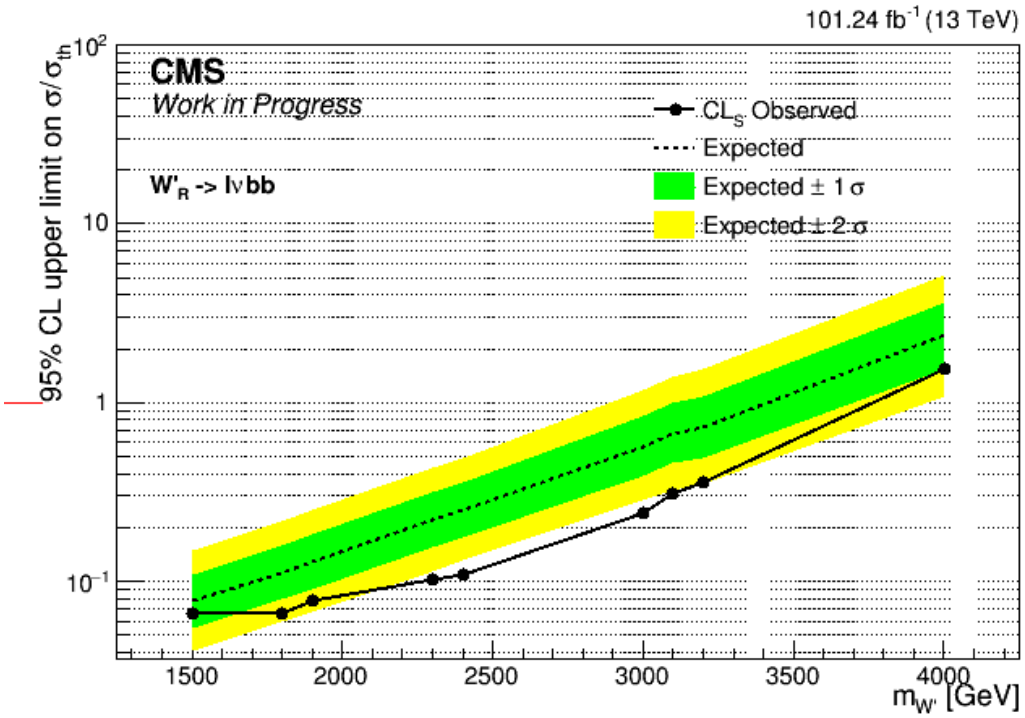


Figure 5.25: Expected and observed upper limits on the ratio $\sigma/\sigma_{th} = \mu_{sig}$ in function of $m_{W'}$, RH coupling, extracted with the combined dataset. The green and yellow stripes represent, respectively, the 1 and 2 standard deviations of the statistical fluctuations of the expected upper limit, evaluated with MC simulations.

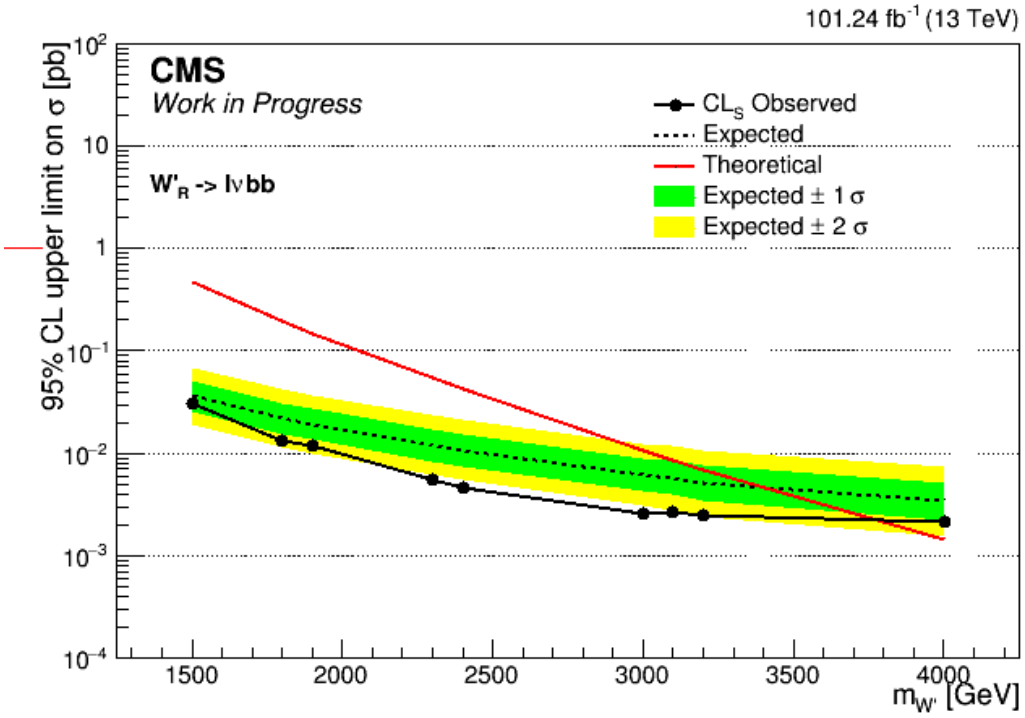


Figure 5.26: Expected and observed upper limits on σ in function of $m_{W'}$, RH coupling, extracted with the combined dataset. The green and yellow stripes represent, respectively, the 1 and 2 standard deviations of the statistical fluctuations of the expected upper limit, evaluated with MC simulations. The red line is the cross section predicted by the model in the MC signal samples.

Conclusions

This thesis work is focussed on the search for a new W' boson, using two different collections of proton-proton collision data delivered by the LHC at a centre of mass energy of $\sqrt{s} = 13$ TeV. They were collected during 2017 and 2018, and correspond to an integrated luminosity of 41.53 fb^{-1} and 59.74 fb^{-1}, respectively.

The search is performed in the semileptonic final states, coming from the direct decay of the hypothetical W' to a tb quark pair, through the chain $W' \to tb \to l\nu_l bb$.

The Standard Model processes that contribute significantly to the background for this decay and, therefore, we considered in this study, are $t\bar{t}$ pair production and $W+$jets production. In order to discriminate the signal from the background, a series of topological and kinematic requirements is applied. The t quark is identified with the $l\nu_l$ pair and a narrow b-jet, while the W'-originated b quark is identified via the corresponding narrow jet, result of its hadronisation. We optimised an event interpretation algorithm in order to reconstruct as consistently as possible the W' primary vertex, naming W' **ensemble** the set of final state objects selected as W' candidate. The mass $m_{W'}$ of the latter is reconstructed from the total 4-momenta of the physical objects included in the ensemble.

We determined the distributions of the contribution to the $m_{W'}$ spectrum of signal and background events by studying the corresponding MC simulated samples.

We defined three fitting regions using the DeepCSV jet flavour tagging algorithm at Medium working point to count the jets identified as coming from b quarks, in order to estimate the background and signal yield via the fit procedure to the data samples on the $m_{W'}$ distribution.

The latter performs an Extended Binned Maximum Likelihood Fit simultaneously in all the fitting regions. In this way we test the hypothesis of the existence of a W' boson. Different benchmark models are tested based on the hypothesis made on the W' boson mass, ranging from 1500 GeV up to 4000 GeV, both in RH coupling scenario and in LH one. No significant deviation from the Standard Model expectations is observed. The results obtained with the presented analysis, therefore, allow to set limits upper limits on the ratio between the observed cross sections and the theoretical one at 95% CL. For the RH coupling scenario from the latter we also inferred lower limit on $m_{W'}$.

This search could be improved by evaluating systematic effects, by introducing control regions to improve the estimate of the background distributions, currently obtained from MC samples, and by using new simulated signal models with 2018 conditions.

Acknowledgments

First of all, I'm very grateful to my supervisor Dr. Alberto Orso Maria Iorio, a great scientific guide and example of professionalism and humanity in the research environment. With his knowledge, kindness, and humorism has made this thesis possible and immersive. I want also to thank Prof. Luca Lista and Dr. Stefano Morisi for their availability, courtesy, advices and time. I'd really like to thank the people of CMS group in Naples, especially Agostino De Iorio, PhD student in Naples and fellow at CERN, and Lorenzo Vigilante, former Master student in Naples and current PhD student in Aachen, for having welcomed and introduced me to the secrets of any sort of the CMS collaboration as kindly as possible. An important thanks, moreover, to the Istituto Nazionale di Fisica Nucleare, including the people working there, for giving me the extraordinary opportunity to realise my dream of visiting and working at CERN, and to CERN itself and its researchers, who honoured me with their precious advices and time.

After offered my thanks to people involved in the scientific community, it's time to thank the non insiders. I dedicate a very heartfelt gratitude to my family, who accompanied, supported and dealt with my person even when it was not so simply, from the primary school to the master degree. I also want to very thanks my real friends, always ready to celebrate any good moments and offer a help when most it was needed.

And last, but not absolutely at least, I am really very grateful to Erika,
who became a very important person in my life while I was working on this thesis, a very smart girl who quickly learned how to deal with a wannabe physicist.

Thank you!

Bibliography

[1] Morad Aaboud et al. “Combination of searches for heavy resonances decaying into bosonic and leptonic final states using 36 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev.* D98.5 (2018), p. 052008. DOI: 10.1103/PhysRevD.98.052008. arXiv: 1808.02380 [hep-ex].

[2] 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”. In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[3] Mohammad Abdullah et al. “Probing a simplified W' model of $R(D^{(*)})$ anomalies using b tags, τ leptons, and missing energy”. In: *Phys. Rev. D* 98 (5 2018), p. 055016. DOI: 10.1103/PhysRevD.98.055016. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.055016.

[4] F. Abe et al. “Evidence for top quark production in pp collisions at $\sqrt{s} = 1.8$ TeV”. In: *Phys. Rev. Lett.* 73 (2 1994), pp. 225–231. DOI: 10.1103/PhysRevLett.73.225. URL: https://link.aps.org/doi/10.1103/PhysRevLett.73.225.

[5] S. Agostinelli, J. Allison, et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[6] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[7] Markus Stoye and. “Deep learning in jet reconstruction at CMS”. In: *Journal of Physics: Conference Series* 1085 (2018), p. 042029. DOI: 10.1088/1742-6596/1085/4/042029. URL: https://doi.org/10.1088%2F1742-6596%2F1085%2F4%2F042029.

[8] G. Arnison et al. “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV”. In: *Phys. Lett.* 122B (1983). [,611(1983)], pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2.

[9] G. Arnison et al. “Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider”. In: *Phys. Lett.* 126B (1983). [,7.55(1983)], pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0.

[10] Pouya Asadi, Matthew R. Buckley, and David Shih. “It’s all right(-handed neutrinos): a new W$'$ model for the R_{D^*} anomaly”. In: *JHEP* 09 (2018), p. 010. DOI: 10.1007/JHEP09(2018)010. arXiv: 1804.04135 [hep-ph].

[11] Ryan Atkin. “Review of jet reconstruction algorithms”. In: *Journal of Physics: Conference Series* 645 (2015), p. 012008. DOI: 10.1088/1742-6596/645/1/012008. URL: https://doi.org/10.1088%2F1742-6596%2F645%2F1%2F012008.

[12] P. Bagnaia et al. “Evidence for $Z^0 \to e^+e^-$ at the CERN pp collider”. In: *Physics Letters B* 129.1 (1983), pp. 130 –140. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90744-X. URL: http://www.sciencedirect.com/science/article/pii/037026938390744X.

[13] M. Bahr et al. “Herwig++ Physics and Manual”. In: *Eur. Phys. J.* C58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[14] M. Banner et al. “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider”. In: *Phys. Lett.* 122B (1983). [,7.45(1983)], pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.

[15] G. L. Bayatyan et al. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS. URL: http://cds.cern.ch/record/706847.

[16] Burton A. Betchart, Regina Demina, and Amnon Harel. “Analytic solutions for neutrino momenta in decay of top quarks”. In: *Nucl. Instrum. Meth.* A736 (2014), pp. 169–178. DOI: 10.1016/j.nima.2013.10.039. arXiv: 1305.1878 [hep-ph].

[17] E. Boos et al. “CompHEP 4.4: Automatic computations from Lagrangians to events”. In: *Nucl. Instrum. Meth.* A534 (2004), pp. 250–259. DOI: 10.1016/j.nima.2004.07.096. arXiv: hep-ph/0403113 [hep-ph].

[18] Andrzej J. Buras et al. “The Anatomy of Quark Flavour Observables in 331 Models in the Flavour Precision Era”. In: *JHEP* 02 (2013), p. 023. DOI: 10.1007/JHEP02(2013)023. arXiv: 1211.1237 [hep-ph].

[19] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti-k_t jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[20] Serguei Chatrchyan et al. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[21] Sergio Cittolin, Attila Rácz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. Geneva: CERN, 2002. URL: http://cds.cern.ch/record/578006.

[22] CMS Collaboration. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *Journal of Instrumentation* 6.11 (2011), p. 11002. DOI: 10.1088/1748-0221/6/11/P11002. arXiv: 1107.4277 [physics.ins-det].

[23] The CMS collaboration. "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector". In: (2010).

[24] The CMS collaboration. *Documentation of the RooStats-based statistics tools for Higgs PAG (rev. 106)*. 2013. URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit?rev=106.

[25] G. Cowan. *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. ISBN: 9780198501558. URL: https://books.google.it/books?id=ff8ZyW0nlJAC.

[26] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[27] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[28] Jean-Marie Frere, Thomas Hambye, and Gilles Vertongen. "Is leptogenesis falsifiable at LHC?" In: *JHEP* 01 (2009), p. 051. DOI: 10.1088/1126-6708/2009/01/051. arXiv: 0806.0841 [hep-ph].

[29] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[30] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579 –588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[31] O. W. Greenberg. "Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons". In: *Phys. Rev. Lett.* 13 (20 1964), pp. 598–602. DOI: 10.1103/PhysRevLett.13.598. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.598.

[32] W. Grimus. "Introduction to left-right symmetric models". In: *Elementary particle physics. Proceedings, 4th Hellenic School, Corfu, Greece, September 2-20, 1992. 1&2*. 1993, pp. 619–632. URL: http://lss.fnal.gov/archive/other/uwthph-1993-10.pdf.

[33] David J. Gross and Frank Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[34] Daniel Guest et al. “Jet Flavor Classification in High-Energy Physics with Deep Neural Networks”. In: *Phys. Rev.* D94.11 (2016), p. 112002. DOI: 10.1103/PhysRevD.94.112002. arXiv: 1607.08633 [hep-ex].

[35] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[36] M. Y. Han and Y. Nambu. “Three-Triplet Model with Double $\mathrm{SU}(3)$ Symmetry”. In: *Phys. Rev.* 139 (4B 1965), B1006–B1010. DOI: 10.1103/PhysRev.139.B1006. URL: https://link.aps.org/doi/10.1103/PhysRev.139.B1006.

[37] “Heavy flavor identification at CMS with deep neural networks”. In: (2017). URL: http://cds.cern.ch/record/2255736.

[38] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[39] *Identification of b quark jets at the CMS Experiment in the LHC Run 2*. Tech. rep. CMS-PAS-BTV-15-001. Geneva: CERN, 2016. URL: http://cds.cern.ch/record/2138504.

[40] V Karimäki et al. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/368412.

[41] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jet Trimming”. In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[42] L. Lista. *Statistical Methods for Data Analysis in Particle Physics*. Lecture Notes in Physics. Springer International Publishing, 2017. ISBN: 9783319628394. DOI: 10.1007/978-3-319-62840-0. URL: https://www.springer.com/gp/book/9783319628394.

[43] Luca Lista. “Practical Statistics for Particle Physicists”. In: *Proceedings, 2016 European School of High-Energy Physics (ESHEP2016): Skeikampen, Norway, June 15-28 2016*. 2017, pp. 213–258. DOI: 10.23730/CYRSP-2017-005.213. arXiv: 1609.04150 [physics.data-an].

[44] B. Odom et al. “New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron”. In: *Phys. Rev. Lett.* 97 (3 2006), p. 030801. DOI: 10.1103/PhysRevLett.97.030801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.030801.

[45] “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”. In: (2009).

[46] *Pileup Removal Algorithms*. Tech. rep. CMS-PAS-JME-14-001. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1751454.

[47] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1346.

[48] Murilo Rangel. "Experimental aspects of jet physics at LHC". In: (2016), pp. 230–234. arXiv: 1608.00057 [hep-ex].

[49] A L Read. "Presentation of search results: theCLstechnique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313. URL: https://doi.org/10.1088%2F0954-3899%2F28%2F10%2F313.

[50] A. Salam and N. Svartholm. "Elementary Particle Theory. Relativistic Groups and Analyticity. Proceedings of the eighth Nobel Symposium, Aspenäsgården, Lerum, Sweden, May 1968. Ed. Interscience (Wiley), New York, and Almqvist and Wiksell, Stockholm, 1969. 400 pp., il..." In: *Science* 168.3936 (1970), pp. 1196–1197. ISSN: 0036-8075. DOI: 10.1126/science.168.3936.1196-a. eprint: https://science.sciencemag.org/content/168/3936/1196.2.full.pdf. URL: https://science.sciencemag.org/content/168/3936/1196.2.

[51] Luca Scodellaro. "b tagging in ATLAS and CMS". In: *5th Large Hadron Collider Physics Conference (LHCP 2017) Shanghai, China, May 15-20, 2017*. 2017. arXiv: 1709.01290 [hep-ex].

[52] G. Senjanovic and Rabindra N. Mohapatra. "Exact Left-Right Symmetry and Spontaneous Violation of Parity". In: *Phys. Rev.* D12 (1975), p. 1502. DOI: 10.1103/PhysRevD.12.1502.

[53] Albert M Sirunyan et al. "Combination of CMS searches for heavy resonances decaying to pairs of bosons or leptons". In: (2019). arXiv: 1906.00057 [hep-ex].

[54] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[55] Torbjörn Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[56] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[57] Jesse Thaler and Lian-Tao Wang. "Strategies to Identify Boosted Tops". In: *JHEP* 07 (2008), p. 092. DOI: 10.1088/1126-6708/2008/07/092. arXiv: 0806.0023 [hep-ph].

[58] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375.

[59] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/357153.

[60] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814.

[61] *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2000. URL: http://cds.cern.ch/record/490194.

[62] CERN Web. URL: https://home.cern.

[63] CMS Web. URL: http://cms.web.cern.ch.

[64] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[65] G Zweig. *An SU_3 model for strong interaction symmetry and its breaking; Version 1*. Tech. rep. CERN-TH-401. Geneva: CERN, 1964. URL: https://cds.cern.ch/record/352337.