Vrije Universiteit Brussel

Faculteit Wetenschappen
Departement Fysica en Astrofysica

Measurement of the single top tW associated production in the dilepton decay channel in proton collisions at a center of mass energy of 8 TeV

Proefschrift ingediend met het oog op het behalen
van de graad van Master in de Wetenschappen

Isis Van Parijs

Promotor: Prof. Dr. Freya Blekman
Co–promotor: Dr. Rebeca Gonzalez Suarez

Academiejaar 2012-2013

We live, I think, in the century of science and, perhaps, even in the century of physics.

Polykarp Kusch

Contents

List of Figures

List of Tables

Introduction

In this thesis the single top tW channel, with a center of mass energy of 8 TeV is the object of the study. This process is yet to be discovered, since it was not accessible at the Tevatron, and with a cross-section of 22.2 pb at a center of mass energy of 8 TeV[3], it is still a rare process at the Large Hadron Collider (LHC[4]). Studies at the Compact Muon Solenoid (CMS,[1]), and A Toroidal LHC ApparatuS (ATLAS,[5]) at 7 TeV have been able to establish evidence of this single top production mode, but a 5σ observation has not been established up to now. A measurement of the cross-section is made with data collected by CMS in 8 TeV collisions, with as result $24.3^{8.6}_{-8.8}$ pb, and the significance of the tW signal is measured to be 4.0σ. From the cross-section, one can estimate the Cabibbo-Kobayashi-Maskawa (CKM) matrixelement $|V_{tb}|$, this has as result 1.04 ± 0.20 (exp.) ± 0.04 (th.). Under the Standard Model assumption of $0 \leq |V_{tb}| \leq 1$, a value of $|V_{tb}| = 0.999$ is found, with a 90% confidence interval of $[0.763, 1.000]$, and a 68% confidence interval of $[0.795, 1.000]$.

The behavior of all known elementary particles can be described within a single theoretical framework called the Standard Model (SM[6]). This gauge theory deals with quarks and leptons as well as their interactions through the strong force, weak nuclear force and electromagnetism. Though this well-known physics model has been tested experimentally up to a level of high accuracy, some questions remain unanswered. The SM does not give a theoretical solution for the experimentally discovered baryon asymmetry in the universe nor the existence of dark matter [7]. The arbitrariness of the mass spectrum and gauge group[6] forms a first theoretical sign that the current Standard Model does not tell the whole story. Furthermore, the inability to include gravity and the hierarchy problem[8] suggests that at higher energies a more fundamental theory will be necessary to describe nature. In order to find a solution, the Standard Model has to be tested up to high accuracy and thus observation has to be established for all possible production modes in the SM, including the production modes of the top quark sector.

Using the worlds largest and most powerful particle accelerator, the LHC, physicists all over the world try to rediscover the SM and find physics phenomena beyond the SM. This proton collider started up on 10 September 2008, and tries to answer the remaining physics question as well as produce new physics phenomena. It is located in the former Large Electron Positron collider (LEP[9]) tunnel at the European Organization for Nuclear Research (CERN[10]) Laboratory near Geneva, Switzerland, and consists of a 27 kilometer ring of superconducting magnets with a number of acceler-

ating structures to boost the energy of the particles along the way.

The study of the heaviest of all quarks, the top quark, is considered to be a highly sensitive window for new physics. The top quark mass of 173.5 GeV[11] is intriguingly close to the electroweak symmetry breaking scale, and has a Yukawa coupling close to one[6]. Top quarks form a good laboratory for new physics searches since it decays mainly through $t \to bW$, and has a lifetime much smaller than the hadron formation scale[11].
Singly produced top quarks are very rare processes that happen via the weak interaction. Based on the production mode, three channels can be distinguished: the s-channel[12], t-channel[13][14], and tW associated production. The tW associated production has a top quark produced in association with a W boson, and has as main background the top quark pair ($t\bar{t}$) production, since this can give the same signature. These processes provide a way to study the CKM matrix element $|V_{tb}|$, the strength of the Wtb vertex [15], W polarization[16] and charge-parity violation (CP violation, [17]). Further, they are background to many new physics searches and are sensitive to many models of new physics (e.g. Super Symmetry[18][19]).
The tW associated production is the theoretically less understood single top mode because it mixes at next to leading order with top quark pair production. It is the only single top channel that is not affected by particles predicted by physics beyond the SM and flavor changing neutral currents (FCNC [20][21]), and is therefore a key in disentangling the t- and s-channel of single top production. Further, it is an irreducible background for important searches such as Higgs to two W bosons [22][23], and the electroweak production of charginos, neutralinos and sleptons.

Throughout this thesis natural units are being used, which means that $\hbar = 1$ and $c = 1$. In Chapter 1, the theoretical physics needed to perform the analysis is given. The SM is briefly explained, and motivations for going beyond the SM are given. Then, the focus is set on the top quark sector of the SM. The key properties of the top quark and singly produced quarks are given.
In the next chapter, Chapter 2, the experimental set-up used for the analysis is introduced. The design of the Large Hadron Collider and its physics program is provided, and the Compact Muon Solenoid is described.
An introduction to the analysis and determination of signal and background is provided in Chapter 3. Then, in Chapter 4, the reconstruction and simulation of events is explained. Using Monte Carlo simulations, events are simulated in order to compare with data. The samples used for this analysis, their theoretical cross-sections, and the High Level Trigger paths can be found in this chapter, as well as the corrections made to the simulated samples.
The selection of the event topology is discussed in Chapter 5, where the event selection criteria and the use of regions are introduced. In Chapter 6, the systematic uncertainties needed for the statistical analysis are being discussed.
Chapter 7, handles about using the established given tools in the statistical analysis for the estimation of the single top tW associated production cross-section. Conclusions and possible enhancements of the performed measurement are provided.

The study was performed at the Interuniversity Institute for High Energies (IIHE[24]), at the top quark subgroup for the CMS collaboration. The selection of the event topology is based on the recommendations of the single top subgroup at the CMS collaboration[25], which already established a measurement at 7 TeV [1].
With this thesis, a cross-check for the single top tW analysis at the CMS collaboration at 8 TeV is provided. Taking the same reconstruction (Chapter 4) and event selection (Chapter 5), a cut-based analysis with a template fit is made (Chapter 7), taking into account the systematic uncertainties (Chapter 6).

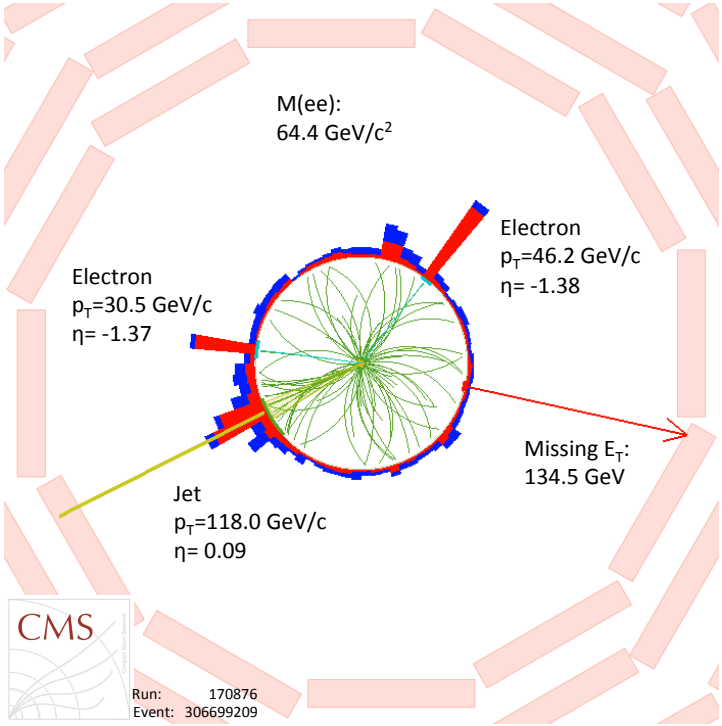


Figure 1: Candidate for a top quark event with a W boson at CMS from the 2010 LHC run. Where the W bosons decay to a b quark and an electron[1].

Chapter 1

Standard Model and Top Quark Sector

The Standard Model of particle physics aims to describe the nature of physics at it most fundamental level. It has been tested up to high accuracy by experimental physicists over the last decades, and is an effective theory only up to the scale of TeV, which means that at higher energy scales new physics phenomena should arise.
The heaviest of all quarks, the top quark, was discovered in 1995 by the collaborations D0 and CDF at the Tevatron collider[26][27]. Its mass is close to the electroweak breaking scale[6], causing the top quark to be sensitive to new physics.

In the first part of this chapter, a brief overview of the Standard Model is given and motivations for physics beyond the Standard Model are stated. Then, a focus is set on the top quark sector. The properties of top quarks are discussed, and the production modes of singly produced top quarks are described.

1.1 Standard Model

Due to the excellent agreement between experimental measurements and theoretical predictions, and the ability to describe all fundamental forces with the exception of gravity, the SM[6] is the general accepted formalism to describe nature at particle level. This gauge theory is developed throughout the mid to late 20th century and uses the fundamental particles and their interactions as building blocks.

All known matter is made of particles with half-integer spin, called fermions f. Fermions are subdivided into two groups, quarks and leptons. Each group is classified into three generations. These generations only differ from another in mass. Thus, all generations have the same quantum numbers (charge, color, spin, ...). According to increasing mass, the generations are labeled 'first', 'second' and 'third'. Each generation contains two quarks and two leptons, and each fermion has a corresponding anti particle $\bar{f}$, making 12 kinds of fermions in total.

The first generation contains the building blocks for stable matter, it consists of the electron e, electron neutrino ν_e, and the up and down quarks. Ever since Chadwick discovered the neutron in 1932[28], the common knowledge is that an atom consists of electrons and a nucleus. This nucleus is built up from protons and neutrons. The electron is known to be an elementary particle (more specifically, a lepton), but the proton and the neutron are not. The proton consists of two up quarks and one down quark, while the neutron consists of two down quarks and one up quark.
The second and third generations contain charged fermions that decay into fermions of a lower generation. In Table 1.1[11], the three generations of fermions, as well as their electric charge Q and mass m are provided.

Table 1.1: The generations in the SM with their measured mass value m and their electric charge Q.

fermions	First	Second	Third	Charge
leptons	electron neutrino ν_e	muon neutrino ν_μ	tau neutrino ν_τ	0
	< 2 eV	< 0.19 MeV	< 18.2 MeV	
	electron e^-	muon μ^-	tau τ^-	-1
	0.51 MeV	105.66 MeV	1776.82 MeV	
quarks	up u	charm c	top t	2/3
	2.3 MeV	1.275 GeV	173.5 GeV	
	down d	strange s	bottom b	-1/3
	4.8 MeV	95 MeV	4.18 GeV	

The fundamental forces or interactions of nature are the strong force, weak force, gravity and electromagnetic force. They are carried by particles of integer spin. These force carrying gauge particles are referred to as bosons. A matter particle (lepton or quark), emits a boson resulting in a change of velocity of the matter particles. The boson in its turn collides with another matter particle and is absorbed, resulting in a change of velocity of the second particle. These mediators of interactions between the fermions are given in Table 1.2[11]. The strong interaction is mediated through eight gluons, while massive W and Z bosons mediate the weak interaction. The well-known photon is responsible for the electromagnetic force between particles with non-zero charge, and since the W bosons are electrically charged, they also couple to the electromagnetic interaction. The range of a force is dependent on the mass of the force carrying particle. When a boson is heavy, it will be difficult to produce and exchange over a large distance. Therefore, a force carried by a high mass boson will have a small range.

A part of the mass of elementary particles is obtained via the interaction with a scalar field (Brout-Englert-Higgs mechanism). This interaction is mediated through a spin-0 particle, called the Higgs boson or the scalar particle. More information about the Standard Model can be found in [6].

Table 1.2: The gauge bosons of the SM and their measured mass value and their electric charge Q.

Boson	Interaction	Mass (GeV)	Charge
photon γ	electromagnetic	0	0
W^+ and W^-	charged current weak	80.398 ± 0.025	+1, -1
Z^0	neutral current weak	91.1876 ± 0.0021	0
gluons g	strong	0	0

1.1.1 Physics Beyond the SM

The SM is tested up to high accuracy and the only significant deviation is the non-zero neutrino masses, discovered in 1988 by the Super-Kamiokande experiment[29]. There is however no issue to incorporate this observation into the theory[6]. Still, there are some remaining questions. The arbitrariness of the parameters and gauge group[1], the presence of only three generations fermions[30], and the electroweak scale ($m_Z \sim 10^2\,\text{GeV}$) being much smaller than the Planck scale ($m_z \sim 10^{19}\,\text{GeV}$, the fundamental scale of gravity), are some of the issues that particle physicists try to address. Another well-known shortcoming is the hierarchy problem, which is elaborately discussed in [8]. The SM does not foresee new physics up to the Planck mass scale. Therefore, a fine tuning of the constants in the theory is required, imposing questions on the naturalness of the theory. Theory would like an unification of the forces at high energy scales, and this seems impossible in the context of the SM. Also, it has proven to be difficult to incorporate gravity in the theory.
Several theories to extend the SM, like Grand Unified Theories[6] or Super Symmetry[19] have been proposed. Some of these extensions predict new physics phenomena at the TeV scale and provide answers to some of the open questions. With the discovery of the Higgs boson ($m_H = 125\,\text{GeV}$) at CERN[31], an answer for physics beyond the SM is within reach.

1.2 Top Quark Sector of the SM

The top quark was first observed in 1995[26][27], by the CDF[32] and D0[33] experiments at the Tevatron collider[34]. With a mass of $173.2 \pm 0.9\,\text{GeV}$[11], the top quark is by far the heaviest of all known quarks. Due to this high mass (almost 40 times the mass of the closest particle, the b quark), the question arises whether or not the top quark plays a special role in the Standard Model, in particular in electroweak symmetry breaking. Due to its mass, the top quark has a large coupling to the Higgs boson, so it can be that the top quark plays a significant role in the mechanism through which particles acquire mass.

[1]The SM is described by the gauge group SU(3)×SU(2)×U(1), which breaks down in SU(3) for the strong force and SU(2)×U(1) for electromagnetism. This group is chosen based on symmetries and can be a part of larger symmetry as is the case for Super Symmetry and Grand Unified theories[6].

Production of top quarks at hadron colliders

At hadron colliders, top quarks are mostly produced via the strong interaction in pairs[35]:

$$\text{gluon fusion: } gg \to t\bar{t} \text{ (85\% at the LHC at 8 TeV)} \tag{1.1}$$

$$\text{quark annihilation: } q\bar{q} \to t\bar{t} \text{ (15\% at the LHC at 8 TeV)} \tag{1.2}$$

But they can also be produced singly via the electroweak interaction, involving the Wtb vertex:

$$\text{W-gluon fusion: } g \to b\bar{b} \text{ and } bu(\bar{d}) \to W^* \to td(\bar{u}) \tag{1.3}$$

$$\text{Drell-Yann type: } du \to W^* \to tb \tag{1.4}$$

$$\text{Associated production: } gb \to b \to tW \tag{1.5}$$

Production of top quarks via the Wtd or Wts vertex are strongly suppressed due to small CKM matrix elements[6]: $0.0048 < |V_{td}| < 0.014$ and $0.037 < |V_{ts}| < 0.043$[36], where V_{ij} represents the decay probability for quark i to quark j. Therefore, their contribution is negligible[37], and only the production via the Wtb vertex is considered.

Because of these two production mechanisms, the top quark production and decay provide important tests of two main forces in the SM.

In Figure 1.1[38], the Feynman diagrams of the top pair production are given, and in Table 1.3 cross-sections are provided.

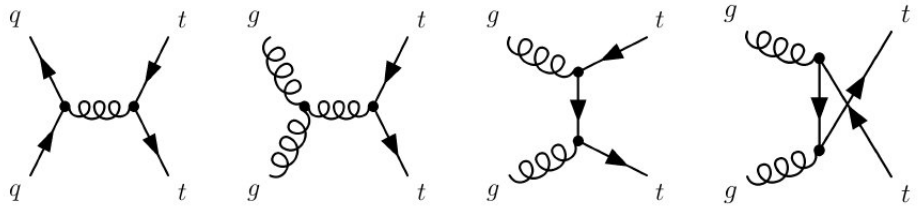


Figure 1.1: Leading order (LO) Feynman diagrams for $t\bar{t}$ production: quark annihilation and gluon fusion

Decay of Top Quarks

Since the top quark is much heavier than the W boson, according to the SM, it can decay in two body final states $t \to Wq$, where the quark $q = b$ has almost a 100% branching fraction ($BR = 99.8\%$ [11]). The decays of the top quark into a W boson and a quark of another doublet ($t \to Ws$ (BR = 0.2 %[11]), $t \to Wd$ (BR = 0.005%[11])) are strongly suppressed in the SM, which can be seen from their corresponding CKM matrix elements (V_{ts}, V_{td}) that are close to zero. More elaborately, one can neglect these decays and calculate the top quark lifetime $\tau_t = 1/\Gamma_t \approx 5 \cdot 10^{-25} s$[11]. This lifetime is about 20 times shorter than the formation time of hadrons, or timescale

for the strong interaction [39]: $\tau \approx 1\,\mathrm{fm/c} \approx 3\,10^{-24}\mathrm{s}$. The fact that $\tau_t << \tau$ implies that the top quark decays before it hadronizes. Since the top quark is the only quark with this property, it is the only quark that can be studied as a bare quark.

1.2.1 Singly Produced Top Quarks

The observation of the single top quark is first reported in 2009 at the Tevatron experiments[40]. At the Large Hadron Collider[4], the study of single top events is much easier. With a total collision energy of 8 TeV, the predicted production rates for the dominant production channel is more or less 30 times more abundant than that from Tevatron (see Table 1.3[3]).

Table 1.3: Cross-sections for single top production (sum of top and anti-top) and top pair production at different center of mass energies at the LHC with a top mass of $173\,GeV$ and Tevatron.

	t-channel	s-channel	tW-channel	$t\bar{t}$
LHC $\sqrt{s} = 7\,TeV$	65.9 pb	4.56 pb	15.6 pb	163 pb
LHC $\sqrt{s} = 8\,TeV$	87.2 pb	5.55 pb	22.2 pb	234 pb
Tevatron	2.08 pb	1.46 pb	0.26 pb	14.16 pb

The production modes via the Wtb vertex are distinguishable by the virtuality[2] (Q^2) of the W boson:

1. In the **t-channel** (see Figure 1.2[1], middle), a space-like ($q^2 < 0$) W boson scatters of a bottom (sea) quark from a proton or a product of gluon splitting ($g \to b\bar{b}$). This mode is also called the W-gluon fusion production and is the most abundant single top production mode at Tevatron and the LHC.

   The t-channel is already being studied since mid 1980, and early 1990s[41][42]. For proton collisions at the LHC, the t-channel events have three quark jets[3] originating from the hard interaction: a b quark jet from top quark decay, a light quark jet, and a $\bar{b}$ quark jet from the initial gluon splitting. The t channel events are most abundant and since they provide a distinct signature, they allowed the first evidence for singly produced top quarks at the LHC.

2. The **s-channel** (see Figure 1.2[1], left) is a Drell-Yann type production and has a time-like ($q^2 \geq (m_{top} + m_b)^2$) W boson that is produced from two quarks belonging to an SU(2) isospin doublet (e.g. $u\bar{d}$), and subsequently decays into $t\bar{b}$.

[2] Q^2 is defined as $-q^2$, where q is the four momentum of the W boson.

[3] Events in hadron collisions are characterized by final state quarks. Due to QCD confinement, color charged particles cannot exist and thus fragment into color neutral bound states. This hadronization produces a cone of hadrons which is defined as a 'jet'. The properties of this jet are used to reconstruct the momentum and energy of the original parton.

At the LHC, this production is very small, and is difficult to distinguish from its background[12].

Due to the fact that the W boson couples only to fermions with left-handed chirality[6], in the t- and s-channel, the restframe of the top quark is a 100% polarized along the direction of the d quark[35]. Therefore, these channels provide a way to study the top quark spin[35]. Further, since the top quark does not hadronize, its decay products contain the information about the top quark polarization.

3. The **tW associated production** (see Figure 1.2[1], right) has a top quark produced associated with a close to real W boson ($q^2 = m_W^2$). This mode is negligible at Tevatron (see Table 1.3), but of relevant size at the LHC.

   The single top tW associated production has yet to be discovered. At Tevatron, this channel was not accessible, and at the LHC it still is a rare process. It is the theoretically less understood single top mode since it mixes at next to leading order with top quark pair production. Also, it shares final states with important searches such as Higgs to two W bosons, and the electroweak production of charginos, neutralinos and sleptons. This single top production is the only channel that is not affected by particles or flavor changing neutral currents. Therefore, it is essential for disentangling the information coming from the t- and s-channel. Another reason for examining this production is because it is sensitive to new physics affecting the Wtb vertex.

The Feynman diagrams for singly produced top quarks can be found in Figure 1.2.

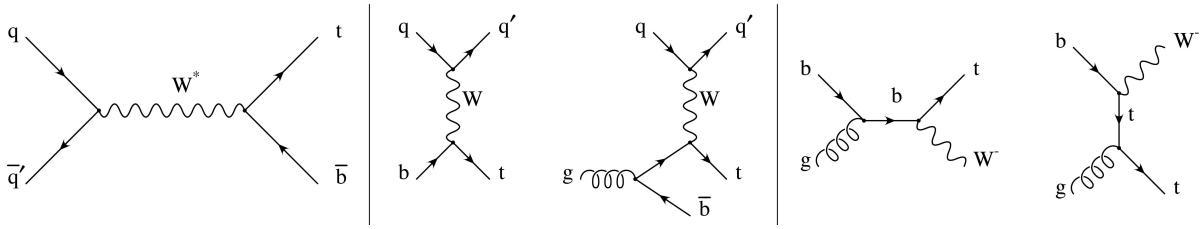


Figure 1.2: Leading order (LO) Feynman diagrams for the three single top production modes: s-channel (left), t-channel (middle) and tW associated production (right)

Next to Leading Order Single Top tW Associated Production

For the tW associated channel, there is interference at higher order with the top quark pair production (see Figure 1.3). This leads to the problem of unambiguously defining

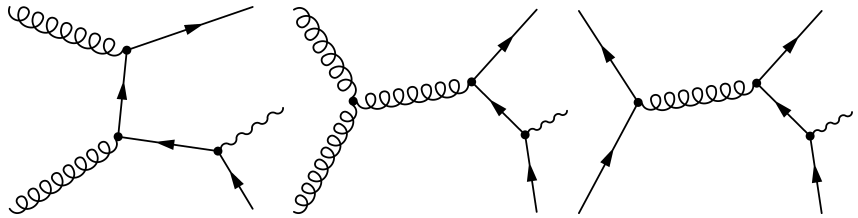


Figure 1.3: Next to leading order (NLO) Feynman diagrams for the single top production in the tW associated channel that mix with $t\bar{t}$ and that are removed from the signal definition in the DR simulation scheme.

the two when considering the Next to leading order (NLO) Lagrangian. To overcome this, two schemes have been proposed to define the tW signal. In the diagram removal (DR) scheme, all signal diagrams which are doubly resonant are removed, while in the diagram subtraction scheme (DS) a gauge invariant term is subtracted which locally cancels the contributions the top quark pair production diagrams. For more information on the DR/DS scheme, see [43].

Chapter 2

Experimental Setup

In this chapter the experimental setup is given. First, in Section 2.1, the design of the Large Hadron Collider is discussed together with its physics program. In section 2.2, the Compact Muon Solenoid experiment is described, where the design of the overall detector concept is introduced. Then, a brief description of the online selection process is given in Section 2.3. More information on these topics can be found in [4] and [44].

2.1 Large Hadron Collider

The Large Hadron Collider is a proton accelerator built at CERN in the existing 27 km long tunnel, between 50 and 175 meters underground, that was used for the LEP[9] collider until 2000. The accelerator complex lies between Switzerland and France, between Geneva Lake and the Jura mountains. Besides being designed for proton collisions, the LHC can also collide heavy ions[45]. Since November 2009, the LHC is the world's highest energy particle accelerator up to date.

2.1.1 Accelerator Complex

The accelerator complex situated at CERN is a succession of machines with increasing energies. As can be seen in Figure 2.1, each machine is injecting the beam into the next one. First, the protons are accelerated to an energy of 50 MeV with a linear accelerator and then fed to the Booster. The Booster produces the first bunches of protons which are injected into the Proton Synchrotron (PS), which accelerates the protons up to an energy of 26 GeV. The Proton Synchrotron delivers bunches of protons which are 25 ns spaced in time to the Super Proton Synchrotron (SPS). There, the protons are accelerated to an energy of 450 GeV and injected into the Large Hadron collider with 2808 bunches at a time. The Large Hadron Collider is the last element in the accelerator chain, and has room for 3654 bunches separated by 25 ns. There is however a 3 μs gap foreseen in the bunch pattern in order to make dumping of the beam in one revolution possible[46]. This 'gap' is in coherence with the time it takes to rise the magnetic field of the beam dumping magnets. The LHC is responsible for the final acceleration of

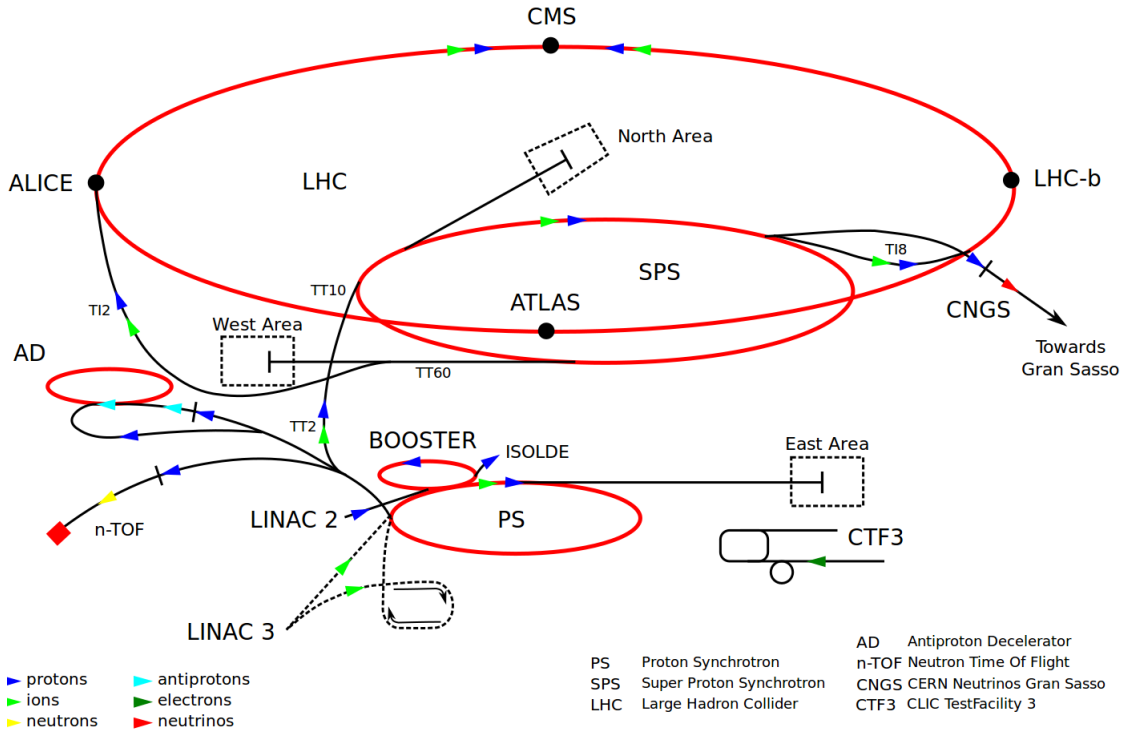


Figure 2.1: The accelerator complex at CERN.

the protons up to the desired energy of (now) 8 TeV. The time needed to fill the LHC beams is about seven minutes and the beam circulates between a few to ten hours, depending on the decrease in luminosity, after which the beam is dumped.

2.1.2 LHC Operation and Design

In particle physics experiments, the number of events N for a physics process is given by the product of the luminosity L, and the production cross-section σ of the process of interest:

$$N = L \cdot \sigma \tag{2.1}$$

Therefore, if one wants to study a process with a low production cross-section, the luminosity of the collider should be as high as possible. For head on collisions this luminosity is given by[47]:

$$L = \frac{f N_p^2 k}{4\pi\sigma_x\sigma_y} \tag{2.2}$$

where f is the revolution frequency, N_p is the number of protons in the colliding bunches, k the numbers of bunches, and σ_i the transverse size of the bunches at the collision point in direction i. At the LHC, the design luminosity is 10^{34} cm^{-2}s^{-1}. In order to obtain such a high luminosity, many particles are needed in each bunch, many bunches should be obtained, and the bunches should have a minimal transverse size at

the interaction point.

For circular accelerators, there is a lot of energy loss due to synchrotron radiation. The particles lose during each revolution an amount of energy ΔE

$$\Delta E \propto \frac{E^4}{Rm^4} \tag{2.3}$$

where E is the energy of the particle, m the particle's mass, and R the radius of the accelerator[48]. In order to reduce this amount of energy loss, when the radius is fixed, one can only go to higher masses.

At the LHC, two proton beams are being used, so one can not use a same magnetic field to keep the particles in their orbit, therefore LHC uses a special design of dipole magnets, more information on these can be found in [49]. In order to obtain the high magnetic fields of 8.33 T to keep particles of such high energies in an orbit, super conducting magnets are necessary. The total length of one dipole magnet is about 15 m and its total mass is about 27.5 ton. At the LHC, there are 1232 dipole magnets through which a current of 15 kA is sent to obtain a 8.33 T magnetic field.

In Figure 2.2[50], the total integrated luminosity throughout the years of data taking is given for CMS. In 2009 and early 2010, the LHC provided initial proton collisions at relatively low center of mass energies. In 2010, the total integrated luminosity accumulated at a center of mass energy of 7 TeV , with up to 368 bunches per beam, corresponds to 44 pb^{-1} . In 2011, this became 6.1 fb^{-1} for a center of mass energy of 7 TeV by raising the number of bunches per beam and optimizing other beam parameters. Since April 2012, the luminosity is even more increased, giving a total accumulated integrated luminosity of 23 fb^{-1}, with a center of mass energy of 8 TeV for CMS.

2.1.3 Experiments at the LHC

The large hadron collider is not a full circle, but built up from eight arcs and eight straight sections. On four of the straight sections the beams cross each other and collisions take place. There, the four main experiments of the LHC are built. This is shown in Figure 2.3[51].

At LHC, the largest contribution to the physics program comes from the two general purpose detectors CMS[44] and ATLAS[52]. Further there are the experiments LHCb, ALICE, TOTEM and LHCf. The LHCb experiment[53], performs precise measurements of CP violations and rare decays. Also heavy ions are being studied at the LHC. In order to study the state of hot nuclear matter (quark-gluon plasma), the ALICE experiment[54] is looking at lead ion collisions. The TOTEM experiment[55], is designed to measure the total proton proton cross-section and is also looking at elastic and diffractive proton collisions, while the LHCf experiment[56] studies the forward production of neutral particles in the proton collisions at low angles.

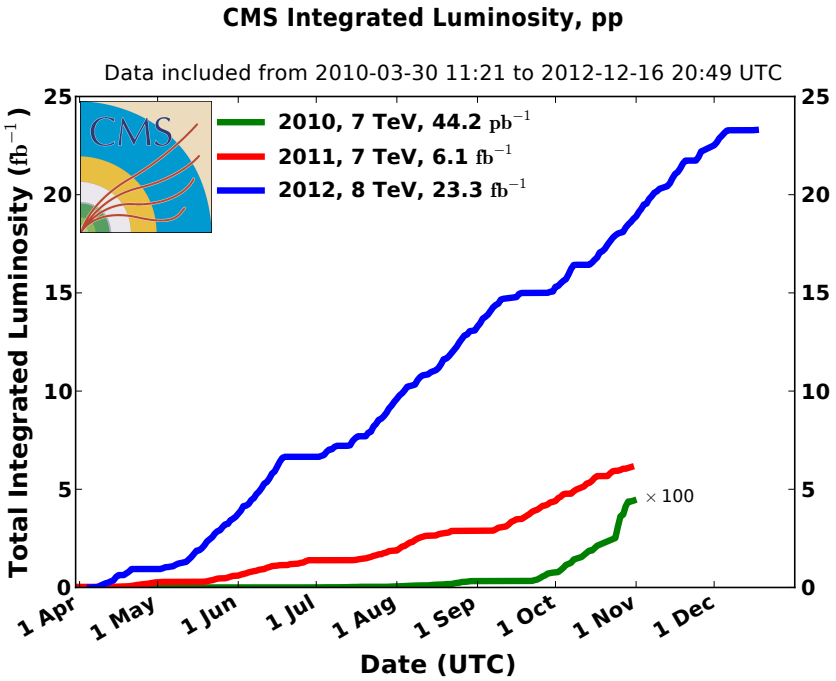


Figure 2.2: Cumulative luminosity versus day delivered to CMS during stable beams and for proton collisions. This is shown for 2010 (green), 2011 (red) and 2012 (blue) data-taking .

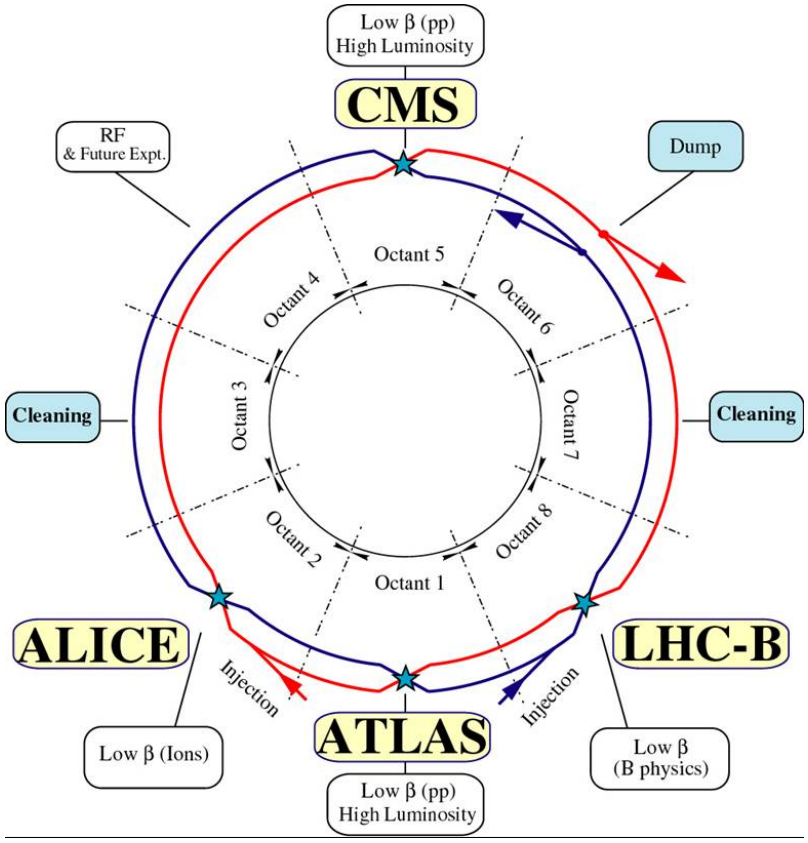


Figure 2.3: The different sections of the LHC.

2.2 Compact Muon Solenoid

The Compact Muon Solenoid is one of the two general purpose detectors at CERN. The CMS collaboration[57] consists of approximately 3 600 people, representing 183 scientific institutes and 38 countries. The experiment is situated in an underground cavern at Cessy in France.

In Figure 2.4[58], a schematic layout of the CMS experiment is shown. This experiment is elaborately discussed in [58]. The overall length of the CMS detector is 21.5

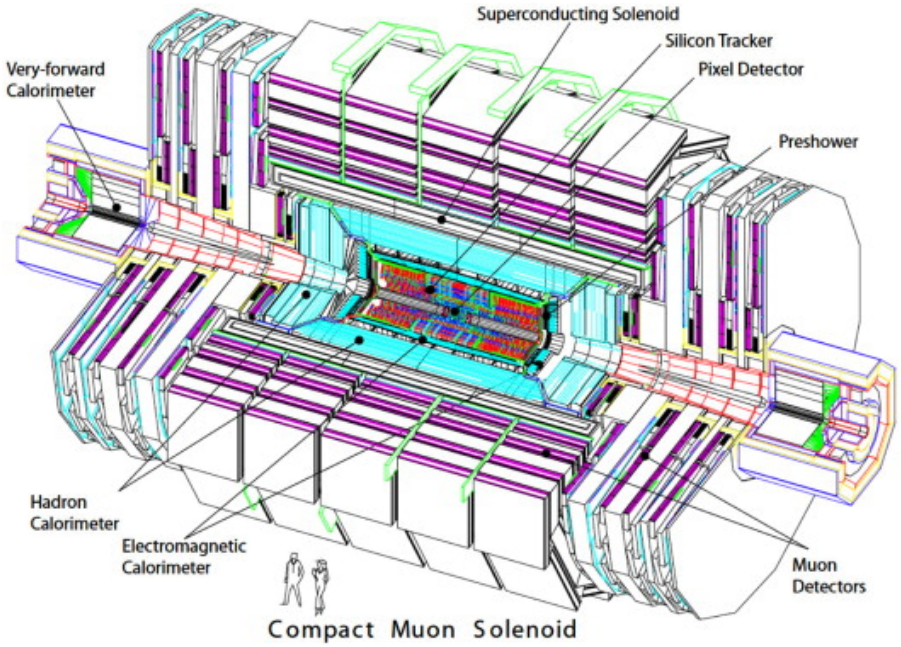


Figure 2.4: Schematic overview of the Compact Muon Solenoid

m, the diameter is 15 m and the weight is 12 500 ton. Since the basic concept of the CMS experiment is to measure the momenta of muons, a large bending power and thus large magnetic field is required. The superconducting solenoid has a length of 12.5 m and a diameter of 6.3 m with an inner diameter of the coil of 5.9 m. It weighs 220 ton and is designed to produce a 4 T field at a current of about 19 kA. The currently used magnetic field is 3.8 T.
The CMS experiment consists of a central barrel part and two end-caps. Within the superconducting solenoid, there is a tracking detector (a silicon strip and pixel tracker), a lead tungstate crystal electromagnetic calorimeter (ECAL), and a scintillating hadron calorimeter (HCAL). The magnetic field lines are closed by the iron return yoke which is placed between the muon detectors. To make the detector more hermetic, a very forward calorimeter is placed in the end-caps along the beam pipe. In Figure 2.5, a transverse slice of the CMS apparatus is given. Each particle can be identified by comparing the hits in each part of the CMS detector.

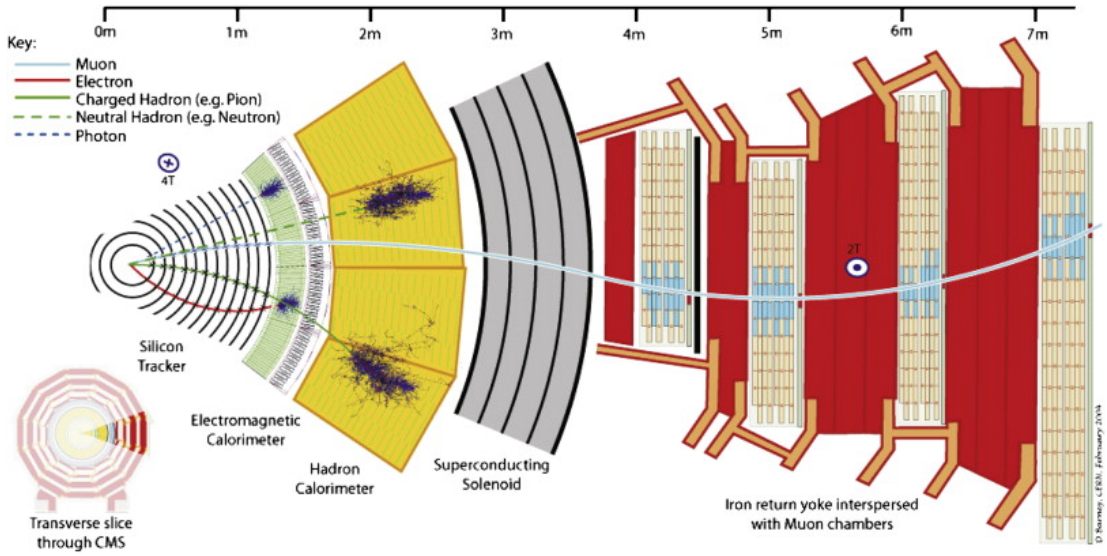


Figure 2.5: Schematic transverse slice of the Compact Muon Solenoid

The CMS experiment uses a right-handed coordinate system, with the origin at the main interaction point, the x-axis pointing to the center of the LHC, the y-axis pointing upwards and the z-axis along counter-clockwise beam direction. The polar angle θ is measured from the positive z-axis and the azimuthal angle ϕ is measured in the xy-plane. Pseudo rapidity is defined as $\eta = \ln[\tan(\theta/2)]$. The transverse momentum p_T ($p_T^2 = p_x^2 + p_y^2$) , the transverse energy E_T ($E_T^2 = m^2 + p_T^2$) and the missing transverse energy E_T^{miss} (see Section 4.3.5) are defined in the xy-plane. The angular distance ΔR is defined as $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$.

Tracks are being reconstructed using the tracking system at CMS, information can be found in [59]. The electromagnetic calorimeter (ECAL) and hadron calorimeter (HCAL) have a barrel and end cap structure. Information about the ECAL and HCAL at CMS can be found respectively in [60] and [61]. Since muons go through the entire detector and deposit a minimum of energy, the muon detection system is placed outside of the solenoid. Information about this system at CMS can be found at [62].

2.3 Online Event Selection Process

With a crossing rate of 40 MHz, LHC has 10^9 proton proton interactions per second. In order to deal with this enormous data rate of 10^9 Hz, an online event selection process or trigger is developed to select the events of interest and reject the other bunch crossings. Since the bunches are 25 ns separated, the decision to keep (or reject) an interesting interaction must be taken very quickly. A problem that arises is that the readout of all the channels on its own takes already more time than the 25 ns. Therefore, a two level triggering system is employed. The Level-1 (L1) trigger uses a customized hardware system, while a High Level Trigger (HLT) is based on software for offline reconstruction (see Figure 2.6). Different trigger streams are defined with as base the reconstruction and identification of the physics objects (see Table 4.2 in

Section 4.1 for the HLT paths used in this manuscript).

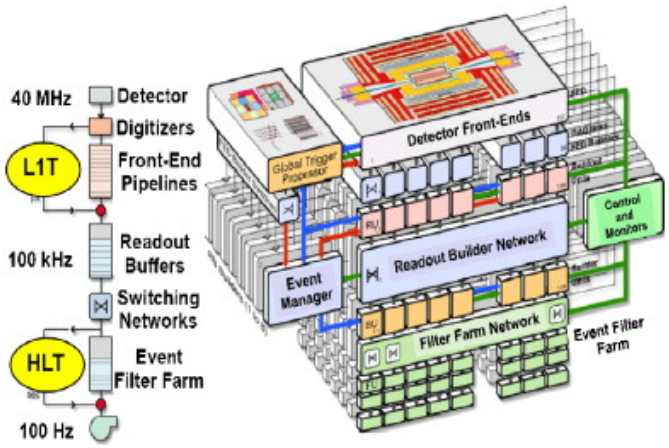


Figure 2.6: The triggering system at the CMS experiment

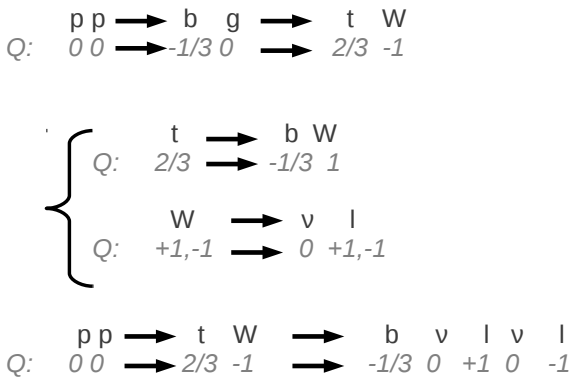
The Level-1 trigger gets a 40 MHz input rate and has to make a decision about the interaction in a frame of 25 ns. As already mentioned, this is impossible and thus a pipeline system is being used. The output rate of this trigger is 100 kHz.
The High Level Trigger reduces the amount of data further. It runs on a farm of mass-market processors[63]. The HLT runs on a single processor for a single event and deals with one event at a time. It has access to the full event information, the full granularity and resolution. The trigger's only limitations are coming from the CPU time usage, limited output rate and the imprecision of the calibration constants available online.
The HLT contains many trigger 'paths', corresponding to dedicated trigger, for example a single electron trigger. A path consists of several steps or software modules, where each module performs a well defined task.

CMSSSW is the overall collection of software at CMS. It is used for the HLT, offline reconstruction, physics analysis and the production of simulated proton collisions. CMSSW is built around a framework, an event data model (EDM) and services needed by the simulation calibration and alignment, and reconstruction modules that process event data. The main goal of the framework and EDM is to facilitate the development and deployment of reconstruction and analysis software. Information on this software framework can be found in [64].

Chapter 3

Introduction to the Analysis

The study is performed in the dilepton channel. Which means that the considered final states are the ones where both W bosons decay leptonically



Only the electron and muon final states are being considered. According to these final states, three underlying sub channels are defined: the ee channel, the eμ channel and the $\mu\mu$ channel, where respectively two electrons, an electron and a muon, and two muons are present.

3.1 Background Processes

The final states of the tW process are characterized by two leptons with opposite charge, a substantial amount of missing transverse energy (E_T^{miss}) due to the presence of neutrinos, and a jet coming from a bottom quark. The challenge for the signal extraction of the tW associated channel lies in the similarities between the signal and backgrounds, and the small cross-section.
All processes that give the same final state are considered background processes[65]. The main background comes from top pair production, with a cross-section of 245 pb (NNLO), this background is 10 times bigger than the signal, which has a cross-section

of 22.2 pb. Other sources of background are Drell-Yann events (or Z/γ^* + jets), W + jets, diboson electroweak production (WW, WZ, ZZ) and other single top processes (t- and s-channel):

- $t\bar{t}$ **events:** in the dileptonic final state, these events are characterized by

$$t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \nu l^+b\nu l^-\bar{b} \tag{3.1}$$

  and tend to give the same signature as the tW signal when one b-jet fails to be identified. This background is irreducible and is the main background for this analysis.

- **Drell-Yann events or Z/γ^* + jets:** This process occurs in high energy hadron hadron scattering. A quark of one hadron and an anti quark of another hadron annihilate, creating a virtual photon or Z boson which then decays into a pair of oppositely charged leptons. The event signature

$$Z + nj \rightarrow l^+l^- + nj \tag{3.2}$$

  where n is a number of jets j, can give rise to

$$Z + nj \rightarrow l^+l^- + \text{ b-jet} \tag{3.3}$$

  which can be mistaken for tW signal.
  The contribution from the Drell-Yann events can easily be identified since it is theoretically well understood. They have two opposite charged, back-to-back leptons that have a large transverse momentum, and through calculation of the invariant mass of the two leptons, these events are easily identified. Though, these events have a cross-section larger than the top quark pair production, and is always of importance in dilepton searches.

- **Diboson electroweak production:** The diboson electroweak production has following relevant decay modes:

$$W^+W^- \rightarrow \nu l^+l^- \tag{3.4}$$
$$W^\pm Z \rightarrow \nu l^\pm l^+l^- \tag{3.5}$$
$$ZZ \rightarrow l^+l^-l^+l^- \tag{3.6}$$

  In combination with the remnants of the colliding protons, these can be mistaken for tW signal events.

- **Other single top processes:** The other single top processes have as signature in the dileptonic

$$\text{s-channel: } t\bar{b} \rightarrow bW^-\bar{b} \rightarrow b\nu l^-\bar{b} \tag{3.7}$$
$$\text{t-channel: } tq \rightarrow bW^-q \rightarrow b\nu l^-q \tag{3.8}$$
$$\text{t-channel: } tq\bar{b} \rightarrow bW^-q\bar{b} \rightarrow b\nu l^-q\bar{b} \tag{3.9}$$

  In combination with the remnants of the colliding protons, these can be mistaken for tW signal events.

- **W + jets:** At CMS, there is a possibility for jets to be misidentified as leptons. The W + jets event signature

$$W^{\pm} + nj \rightarrow \nu l^{\pm} + nj \tag{3.10}$$

  where n is a number of jets j, can be mistaken for

$$W^{\pm} + nj \rightarrow \nu l^{\pm} + l^{\mp} + \text{ b-jet} \tag{3.11}$$

  which is a similar signature as the tW signal.

- **Multi-jet QCD background:** QCD multi-jet background can appear as signal when jets are misidentified as leptons. This multi-jet QCD background becomes negligible after a tight lepton selection is applied.

Chapter 4

Reconstructing and Simulating Events

In this study, Monte Carlo samples corresponding to the Summer12 official CMS production have been used. The samples are generated with pile-up (PU S10 START53 V7A scheme), where pile-up is defined as the interactions that aren't the interaction of interest. In Chapter 2, the several parts of the CMS detector are discussed. The method of reconstruction of the electronic signals registered by the sub-detectors to the high level physical objects are presented in Section 4.3. Monte Carlo generators are being used for the simulation of top quark production, as well as most background processes. This simulation based on the theory presented in Chapter 1, can then be compared with the recorded data. Information about the simulation of collisions can be found in Section 4.2, while in Section 4.4 the corrections applied to the simulated samples is discussed.

4.1 Analysis

A shape analysis is performed, using a discriminant value built from a physical quantity after the application of kinematic selections. A sample collected at a center of mass energy of 8 TeV at the Compact Muon Solenoid, corresponding to an integrated luminosity of 12.1 fb^{-1}, is used. The study has been performed on the dileptonic data streams containing electrons and muons. The used data streams, with their corresponding luminosities, are shown in Table 4.1, while in Table 4.2 the High Level Trigger paths used for this study are given.

Table 4.1: The main data samples used in the analysis.

Final State	Data stream	Run periods	Total luminosity
$\mu\mu$	/DoubleMu/	A,B,C	12.1 fb^{-1}
eμ	/MuEG/	A,B,C	12.1 fb^{-1}
ee	/DoubleElectron/	A,B,C	12.1 fb^{-1}

Table 4.2: The trigger paths used in the analysis (run A,B,C).

Final State	Trigger path
$\mu\mu$	HLT Mu17 Mu8 v*
$\mu\mu$	HLT Mu17 TkMu8 v*
eμ	HLT Mu17 Ele8 CaloIdT CaloIsoVL TrkIdVL TrkIsoVL v*
eμ	HLT Mu8 Ele17 CaloIdT CaloIsoVL TrkIdVL TrkIsoVL v*
ee	HLT Ele17 CaloIdT CaloIsoVL TrkIdVL TrkIsoVL Ele8 CaloIdT CaloIsoVL TrkIdVL TrkIsoVL v*

4.2 Simulating Collision Events

Monte Carlo generators are used for the simulation of the top quark production, as well as the simulation of most important background processes. In order to facilitate comparison with the experimental data, the simulated samples are processed through detailed detector simulations and are subjected to the same reconstruction algorithms and analysis chain as the real data. Signal simulation is used to determine the selection efficiency, and the MC samples are used to determine the differential distributions for signal and background.

For the CMS experiment, simulated events are produced using the CMS Software Framework (CMSSW). This framework will ensure the event generation, the simulation of the full CMS detector and the reconstruction of the particles present in the events. The single top quark events are simulated using the POWHEG event generator version 301[66]. This Monte Carlo generator can describe the full next to leading order (NLO) properties of these processes. MadGraph 5.1.1.[67][68] is used for the top quark pair, as well as the inclusive single boson production (V+X), where $V = W, Z$ and X can indicate light or heavy partons. PYTHIA version 6.4.24[69] is used for simulating the remaining background events, including the diboson production. The CTEQ 6.6M parton distribution function sets[70] are used for all simulated samples. The generated events undergo a full simulation of the detector response according to the CMS implementation of GEANT4[71].
The simulation of the interactions of the particles as they cross the CMS detector, as well as the measured hits of the simulated events are digitized. The latter is done by simulating the electronics response. Hence, the simulation will be influenced by the same detector effects as data (cable interactions, badly working electronics, etc).

The top quark pair production is normalized using next to next to leading order (NNLO) theoretical predictions. The the W+jets and Z/γ^*+jets processes are also normalized to complete NNLO calculations for the inclusive cross-sections, and NLO cross-sections are used for the diboson processes.
In Table 4.3 and 4.4[3][72], the theoretical cross-section of each sample that is being used is stated.

Table 4.3: The cross-section of the used Monte Carlo single top samples at 8 TeV.

Process	Cross-section (pb) at 8 TeV
tW channel: $\bar{t}$ and t DR	22.2 (NNLO)
Single top t channel: $\bar{t}$	30.7 (NNLO)
Single top t channel: t	56.4 (NNLO)
Single top s channel: $\bar{t}$	1.76 (NNLO)
Single top s channel: t	3.79 (NNLO)

Table 4.4: The cross-section of the used Monte Carlo main Summer12 samples at 8 TeV.

Process	Cross-section (pb) $\times$ BR at 8 TeV
$t\bar{t}$	245 (NNLO)
DY $m_{ll} \in [10, 50]$ GEV	860.5 (NNLO)
DY $m_{ll} > 50$ GeV	3532.8 (NNLO)
W + jets	37509 (NNLO)
WW	57.07 (NLO)
WZ	22.44 (NLO)
ZZ	9.03 (NLO)

4.3 Physics Objects

The physics objects are all reconstructed using the CMS particle flow algorithm[73][74]. Particle Flow uses the full list of particles in the final state and the whole detector.

4.3.1 Muons

Muons traverse more matter compared to any other Standard Model particle (neutrino's excepted). They are the only particles that interact with all sub-detectors of CMS and are therefore easily identified and reconstructed.

Muons are reconstructed in CMS by combining the information from the muon chambers and the inner tracking detectors. First, the tracks are reconstructed individually in the inner tracking detector, as well as in the muon system. These tracks are then merged in order to from muon candidates by either using the muon detector or inner tracking detector as seed. The tracks coming from muons have to be of good quality and should be ascending from the reconstructed primary vertex. The muon candidate is required to have a certain minimum amount of hits in the inner tracking detectors, and to have a high quality global fit including a minimum amount of hits in muon detector. The reconstruction efficiency for muons is typically 95%. Information about muon reconstruction at CMS can be found in [75].

For this study the muons are identified and selected by standard quality criteria for

dileptonic searches taken from the single Top working group at CMS[25]. The muon selections can be found in Table 4.5.

Table 4.5: The muon selection criteria.

Particle Flow Muon identification	isPFmuon
Muon Reconstruction Algorithm ID	isGlobalMuon or isTrackerMuon
Transverse momentum	> 20 GeV
$\|\eta\|$	< 2.4
reliso	< 0.20, cone 0.4

The first requirement is that the muon is reconstructed using the particle flow algorithm[73][74]. Further, for the reconstruction, the global muon (tracker and chamber) and tracker muon algorithms are being used[76]. Further the muon is required to have a transverse momentum greater than 20 GeV.
In Figure 4.1[77], the longitudinal view of the CMS detector is given. The entrances of the barrel (muon chambers) can't detect muons very well. For this reason, the absolute pseudo-rapidity should be smaller than 2.4.

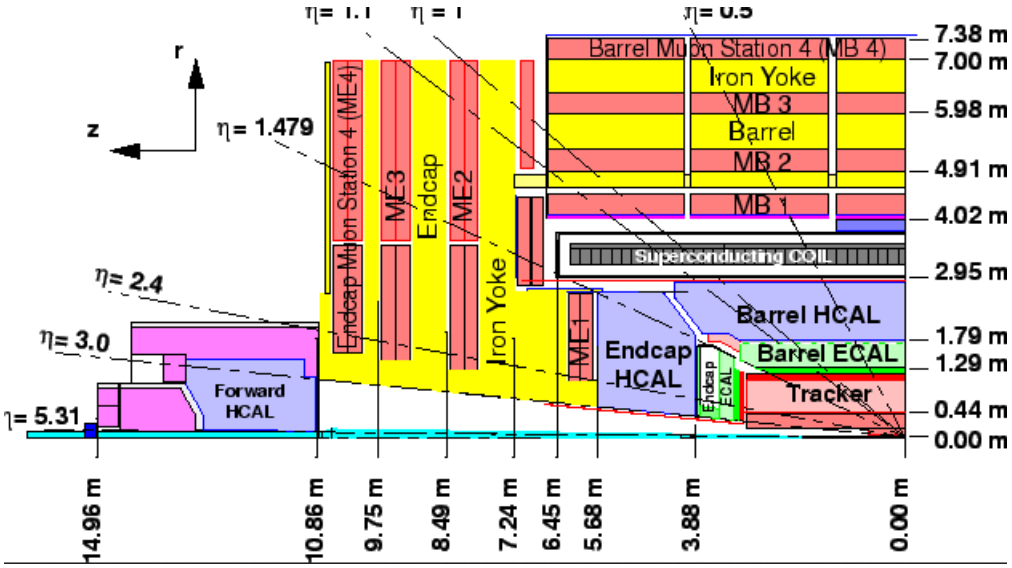


Figure 4.1: Longitudinal view of the CMS detector.

Isolated, prompt leptons coming from W boson decays should be selected. Therefore, isolation requirements are set. A cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ equal to 0.4 is constructed around the muon track. The sum of the momenta inside this track, and the electromagnetic and hadronic calorimeter energy deposits are calculated, excluding the contribution of the muon itself. If this scalar sum exceeds 20% relative to the value of the muons candidate momentum, the candidate is considered to be non-isolated and rejected. The ratio of the scalar sum to the candidate's momentum is called the 'relative isolation' (reliso). Thus, the reliso has to be smaller than 0.20 in order for the muon to be isolated.

Events with extra loose leptons are being vetoed in this analysis. These loose leptons are leptons with relaxed selections. The loose selection criteria for the muons is given in Table 4.6.

Table 4.6: The loose muon selection criteria.

Particle Flow Muon identification	isPFmuon
Muon Reconstruction Algorithm ID	isGlobalMuon or isTrackerMuon
Transverse momentum	> 10 GeV
$\|\eta\|$	< 2.5
reliso	< 0.20, cone 0.4

4.3.2 Electrons

The electrons are reconstructed using the energy deposits or clusters in the electromagnetic calorimeter. The clusters are associated with charged particle tracks coming from the tracking detector. The Gaussian Sum Filter algorithm is being used to take into account Bremsstrahlung. Information on this algorithm can be found in [78][79]. The electrons are identified and selected by some standard criteria given in Table 4.7.

Table 4.7: The electron selection criteria.

identification	GsfElectron
Transverse momentum	> 20 GeV
$\|\eta\|$	< 2.5
Transverse IP of the electron (GSF track)	< 0.04
Conversion rejection	true
MVA	> 0.5
number of Hits	≥ 1
reliso	< 0.15, cone 0.3

The transverse momentum should be higher than 20 GeV and the absolute pseudorapidity should be below 2.5 due to detector limitations. The distance from the reconstructed primary vertex should be smaller than 0.04 cm so that the electron is originating from this vertex. Since a photon can convert into two electrons and give a false signal, the conversion rejection must be set to true. Further, a minimum amount of hits in the inner tracker is set. In order to know if it is a good electron, multi variable triggers (MVA[80]) are being used. Similar as for the muon, the electron should be isolated to ensure that the electron is descending from a W boson decay.
The loose electron selection criteria are given in Table 4.8.

Table 4.8: The loose electron selection criteria.

identification	GsfElectron
Transverse momentum	> 10 GeV
$\|\eta\|$	< 2.5
Transverse IP of the electron (GSF track)	< 0.04
Conversion rejection	true
MVA	> 0.5
reliso	< 0.15, cone 0.3

4.3.3 Jets

The hadronic jets are reconstructed using the infrared and collinear safe anti-k$_T$ algorithm [81] with a parameter size of 0.5.
When two particles are combined during the anti-k$_T$ algorithm, their four momenta are combined according to a recombination scheme. This scheme describes how the constituents of the jet are added together in order to calculate the properties of the jet. In this thesis, the Energy or E-scheme is applied. It simply adds the constituents as four-vectors. The resulting four-vector then represent the kinematic properties of the jet. More information about the E-scheme and about the anti-k$_T$ can be found in [82][83] and [81].

Table 4.9: The jet selection criteria.

Corrected p_T	≥ 30 GeV
$\|\eta\|$	< 2.4
Jet Energy Corrections (JEC)	L1FastJet + L2L3 (+L2L3Residual for data)
JER smearing in MC	applied
jet ID	applied
distance lepton	> 0.3

For the jet selections, again the standard requirements are set. For similar reasons as the leptons, the jets absolute pseudo-rapidity has to be below 2.4 (see Table 4.9). Further, the transverse momentum should be higher than 30 GeV since the transverse momentum is not well reconstructed below this threshold.
Also, corrections on the jet energy scale are applied, information about these corrections can be found in [84][85][86]. It is proven that the jet resolution in data is broader than the one obtained in simulation. For this reason, a smearing is applied on Monte Carlo, for more information on the Jet Energy Resolution (JER) see [87].
The collected data can contain a mixture of real jets originating from the primary interaction, jets from accelerator induced backgrounds, and from non-physical sources. These non-physical sources are, for example, jets due detector noise. In order to reduce these pure fake jets, a set of jet quality criteria, referred to as Jet Identification or Jet ID, are set. More information about Jet ID can be found in [88][89].

In Figure 4.2, the distribution of the number of jets is shown.

Table 4.10: The loose jet selection criteria.

Corrected p_T	≥ 20 GeV
$\lvert\eta\rvert$	< 2.4
Jet Energy Corrections (JEC)	L1FastJet + L2L3 (+L2L3Residual for data)
JER smearing in MC	applied
jet ID	applied

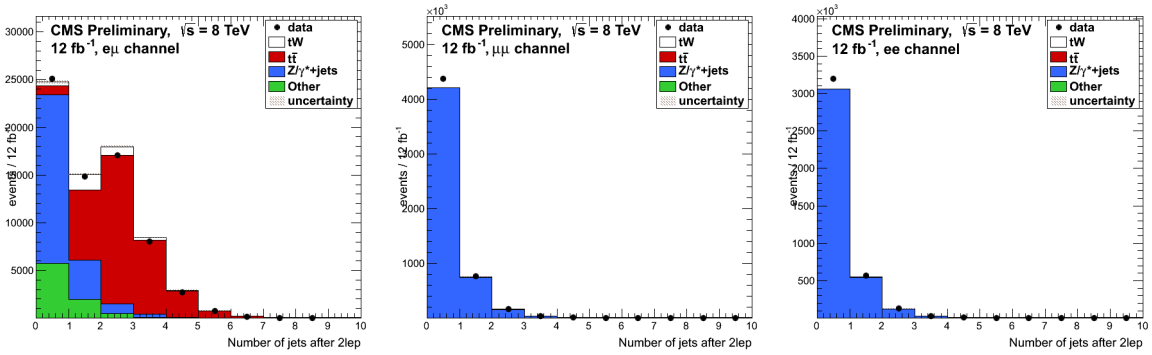


Figure 4.2: The distribution of the number of jets. Left the eμ channel, in the middle the $\mu\mu$ channel, and right the ee channel. After lepton selection (see Chapter 5) with the statistical uncertainty for all samples.

4.3.4 Jets from a b Quark

The CMS collaboration is able to identify jets coming from a b quark. The collaboration has developed a variety of b-tagging algorithms based on the use of impact parameters of charged particle tracks, the properties of reconstructed decay vertices and/or the presence of a lepton.
In order to discriminate between b (or c) and light flavor jets, a variety of reconstructed objects is used. There are algorithms that use only one observable, but there are also b-tagging algorithms that combine observables in order to have a higher discriminating power for each jet. The b-tagging algorithms use a discriminating value or working point. Based on the misidentification probability for light flavor jets the working points are defined as loose (L), medium (M), and tight (T)[90].

B hadrons are produced during fragmentation of b quarks and have a large lifetime (of about 1.5 ps[11]). Their decay length is about 450 μm, resulting in a displaced vertex (see Figure 4.3[91]). This displaced or secondary vertex is observed by the CMS silicon tracker by looking at the intersection point of the tracks. By looking at displaced vertices, one can see if the jet is originating from a b quark. There are two main kinds of b-tagging algorithms. A first category is where the identification is based on the use of track impact parameters, these impact parameters are defined as the distance

between the primary vertex and the linearized track in the point of minimal distance between the track and the jet axis, as shown in Figure 4.3[92].
Another category is the one with b-tagging algorithms based on the use of a secondary vertex. More information about b-tagging can be found in [90].

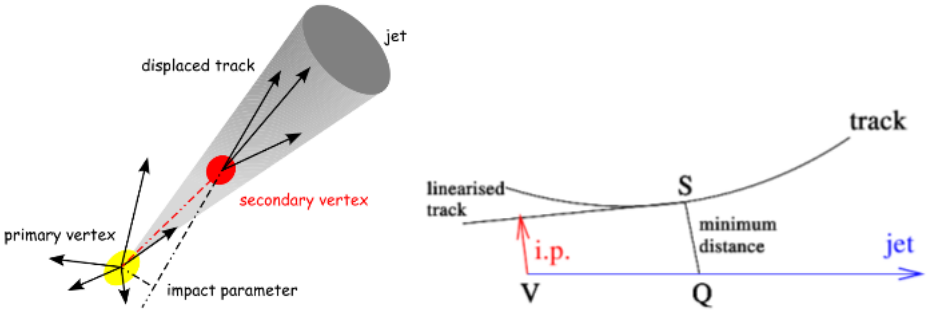


Figure 4.3: Schematic representation of B hadron decaying in a b jet and illustration of the impact parameter (i.p.)

The b-tagging algorithm used in this analysis is the Combined Secondary Vertex (CSV) in the medium operating point, the base of this algorithm is the use of secondary vertices with a track based lifetime information. With a value of the discriminant $\geq$ 0.679, the measured b-tagging efficiency for this point is 62% and the misidentification probability is 0.01%[90]. In Figure 4.4[93], the efficiency of b-tagging for the CSV b-tagging algorithm is given.

In Figure 4.5, the distribution for the number of b-tagged jets in each channel can be found.

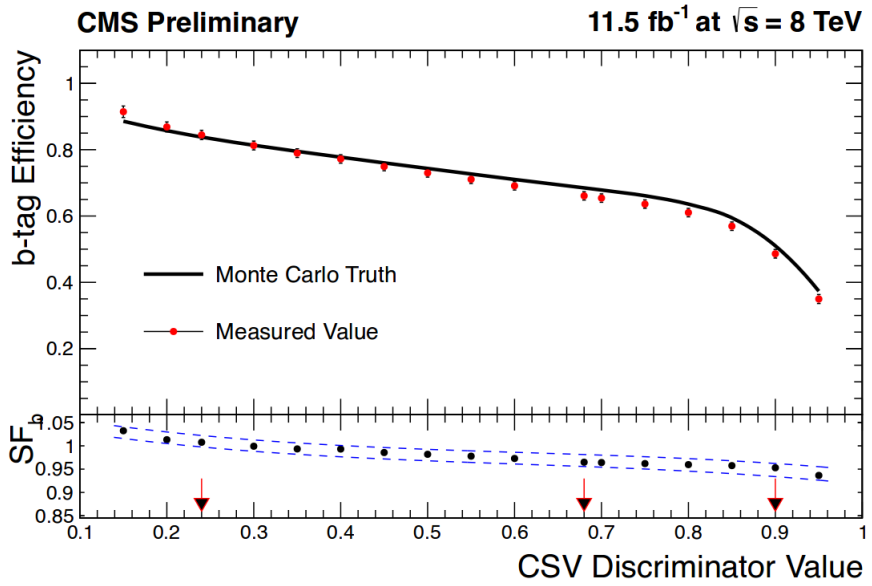


Figure 4.4: Measured b-jet tagging efficiency as a function of the flavor discriminator threshold for the CSV algorithm, measured with the flavor tag consistency (FTC) method. The absolute b-jet tagging efficiency measured from data and predicted from simulation is shown in the upper panel. The scale factors SF for b-tagging is shown in the lower panel, where the blue dashed lines represent the combined statistical and systematic uncertainty. The arrows indicate the standard operating points.

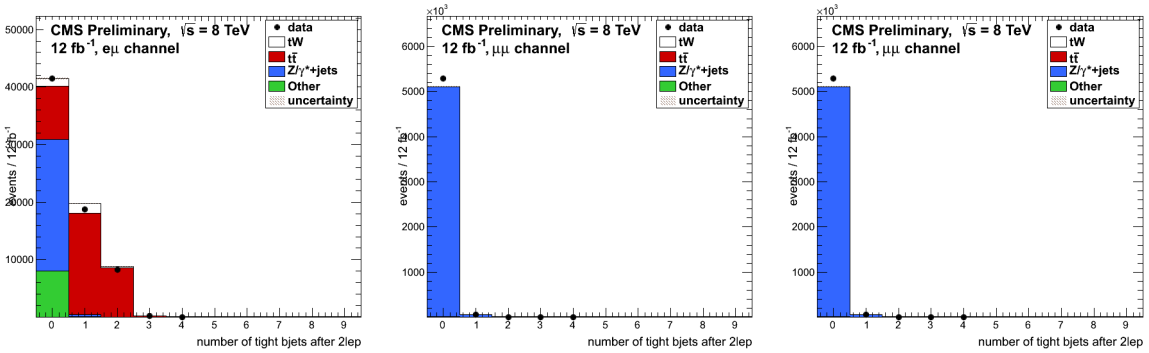


Figure 4.5: The distribution of the number of b-tagged jets. Left, the eμ channel, in the middle the $\mu\mu$ channel, and right the ee channel. After lepton selection (see Chapter 5, with the statistical uncertainty for all samples.

4.3.5 Missing Transverse Energy

Due to the occurrence of neutrinos that pass through the detector without interacting (and mis-measurements), there is missing transverse energy. In general, this is calculated using conservation laws and is thus defined as the negative of the vector sum of the transverse momentum of all final state particles. Then, algorithms are used in order to remove anomalous signals in the calorimeters such as detector noise, as well as beam-halo muons.
A distribution of the missing transverse energy can be found in Figure 4.6.

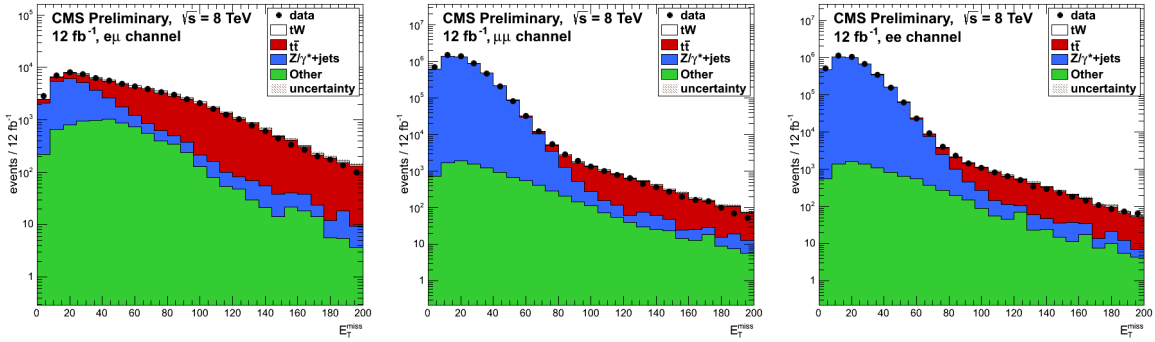


Figure 4.6: The logarithmic distribution of the missing transverse energy. Left the eμ channel, in the middle the $\mu\mu$ channel, and right the ee channel, with the statistical uncertainty for all samples.

4.4 Corrections Applied to the Simulated Events

There are subtle detector related effects that cannot be modeled perfectly. Therefore, studies are made in an independent way for each of these effects and corrections are provided in order to reproduce the data in simulation. The uncertainties introduced by applying the corrections on the simulation, are accounted for in the statistical method (Chapter 7) and are discussed in Chapter 6.

4.4.1 Pile-up Reweighting

For each single bunch crossing at the Large Hadron Collider, several interactions could take place. In time pile-up is defined as all particles descending from the same bunch crossing, but from a different proton proton interaction. While, out of time pile-up, is defined as the left over signal from a previous bunch crossing. The Monte Carlo samples are generated with simulated pile-up meant to roughly cover the conditions for the data taking period. Though,the pile-up is in Monte Carlo is produced before the actual data taking begins, and therefore it is essential to reweight the pile-up.

The pile-up calculation is done centrally by the pile-up studies group at CMS[94]. Information on the pile-up reweighting procedure can be found in [95]
In Figure 4.7, the effect of pile up reweighting is shown. The distribution of the

simulation without pile-up reweighting is broader than the data distribution. In other words, the simulation expects more vertices in the tail of the distribution than reality.

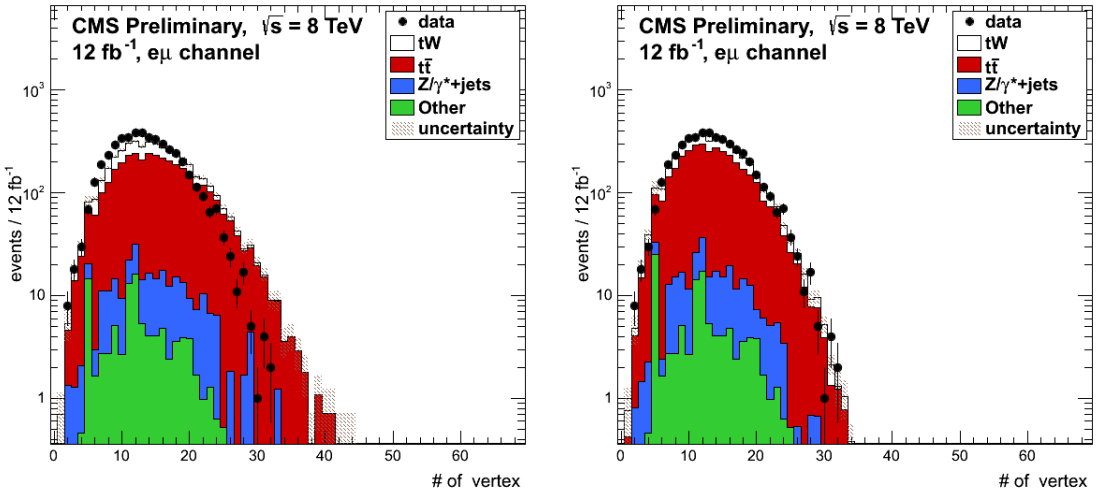


Figure 4.7: The distribution in log scale of the number of vertices in the eμ channel. Left without pile-up reweighting, right with pile-up reweighting, after the 1 jet requirement.

4.4.2 Lepton Isolation, Identification and HLT Reweighting

In order to correct for the High Level trigger modeling, scale factors are applied. Information about the lepton identification and isolation efficiency reweighting calculation can be found in [2]. The resulting scale factors are listed in Table 4.11. These are obtained for each channel of the analysis based on samples reprocessed in CMSSW53X. The trigger scale factors are obtained with data corresponding to 12 fb^{-1}.

Table 4.11: The lepton identification and isolation efficiency scale factors taken from [2].

Channel	SF_{trig}	$SF_{Iso,ID}$	$SF_{Tri,Iso,ID}$
ee	0.975 ± 0.011	0.926 ± 0.019	0.903 ± 0.021
eμ	0.953 ± 0.011	0.960 ± 0.014	0.915 ± 0.017
$\mu\mu$	0.965 ± 0.010	0.998 ± 0.020	0.963 ± 0.022

4.4.3 B-tag Efficiency Reweighting

For each jet in the event, there is an efficiency of b-tagging. The performance of this b-tagging in data and simulation has been studied, and the outcome is that the efficiencies aren't the same. Thus, also here there is a mis modeling due to subtle detector effects. The efficiency of b-tagging in Monte Carlo for each jet is rescaled in order to match the efficiency in data. These scale factors are provided by the b-tagging working group of the CMS collaboration [90], and are function of the transverse momentum and the pseudo rapidity.
The efficiency of the Monte Carlo simulation has to be rescaled so that:

$$\epsilon_{true} = \epsilon_{MC} \times SF \tag{4.1}$$

where SF is the scale factor provided by the b-tagging working group. These scale factors are only provided for jets with a transverse momentum higher than 20 GeV and an absolute pseudo-rapidity below 2.4. For this reason, only jets fulfilling these requirements are considered.

First, the raw efficiency of b-tagging (and fake rates) in simulation is determined. This is done by taking the ratio of the number of b-tagged jets (jets that pass the discriminating threshold) over all jets coming from a b quark (or c, or light quarks):

$$\epsilon_{MC} = \frac{\text{b-tagged jets}}{\text{b-jets}} \tag{4.2}$$

The raw efficiency is a function of the transverse momenta. The efficiencies are calculated for the signal and the main background ($t\bar{t}$), for the other backgrounds a same efficiency as $t\bar{t}$ is assumed. Further, it is assumed that the efficiencies measured in the eμ channel are assumed to the same for all channels. The resulting distributions can be found in Figure 4.8. For the secondary vertex algorithm, the b-tagging efficiency

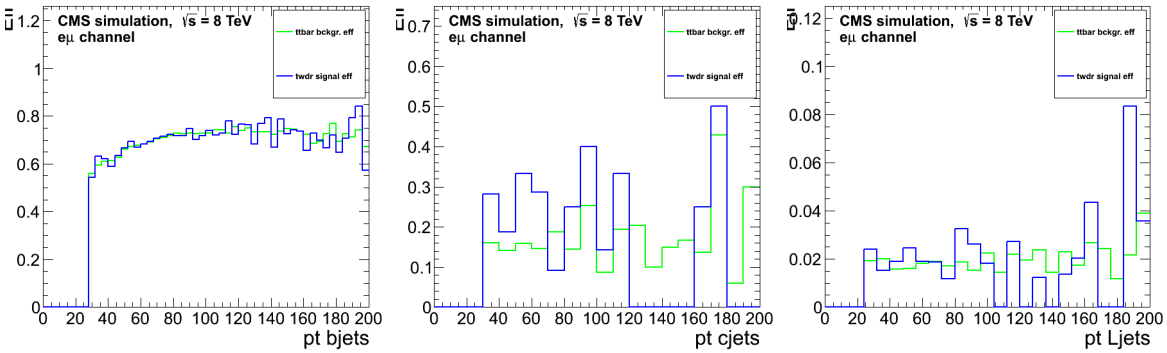


Figure 4.8: The efficiency of b-tagging (left), the fake rate efficiency for c-jets (middle), and fake rate efficiency for light jets (right). All as a function of the transverse momentum of the jet in the eμ channel.

should be 62% with a mis-tag rate of 1.5% for jets with a transverse momentum between 50 and 80 GeV. This trend can easily be seen for the b-jets, but is harder for

the fake rates due to low statistics.

In practice, the true b-tag efficiency is obtained using the algorithm described in [96]. When the b-tag scale factor provided by the b-tagging working group at CMS[90], is greater than one, the efficiency in data is greater than the efficiency in Monte Carlo. To solve this, the simulation efficiency should be adapted such that the it matches the one from data. The percentage of extra jets that have to be tagged as b-jets in simulation is defined as

$$mistag\% = \frac{1 - \mathrm{SF}}{1 - \frac{\mathrm{SF}}{\epsilon_{MC}}} \tag{4.3}$$

Then the non b-tagged jets get randomly tagged according to this percentage.
When the scale factor is smaller than one, less jets get b-tagged in data compared to simulation. B-tagged jets get randomly untagged according to the SF.

4.4.4 Z/γ^* Reweighting

There is a mis modeling of the missing transverse energy, especially for events without genuine neutrinos and thus missing transverse energy.
Therefore, the missing transverse energy is reweighed. Scale factors are calculated in the following way. For the Z/γ^* plus jets sample, the data and the total Monte Carlo minus the Z/γ^* plus jets sample, histograms of the missing transverse energy inside the Z mass window[1] are made. Then the scale factor per bin i is defined as

$$SF_i = \frac{data - (MC_{total} - MC_{Zjets})}{Z/\gamma^* + \mathrm{jets}} \tag{4.4}$$

The scale factors are determined for the ee and $\mu\mu$ final states, after lepton selection and loose lepton veto requirements, and the mean of both final states is taken for the eμ channel. The resulting scale factors can be found in Table 4.12.

Table 4.12: The Z/γ^* missing transverse energy scale factors.

E_T^{miss}	SF $\mu\mu$	SF ee	SF eμ
$< 10\ GeV$	0.8841	0.9215	0.9028
$< 20\ GeV$	0.9386	0.9608	0.9497
$< 30\ GeV$	1.0131	1.0247	1.0189
$< 40\ GeV$	1.1012	1.0964	1.0988
$< 50\ GeV$	1.1850	1.1633	1.17415
$< 60\ GeV$	1.2500	1.2529	1.25145
$> 60\ GeV$	1.3071	1.2194	1.26325

[1]This is where the invariant mass of the two leptons is in $[81, 101]\ GeV$.

Chapter 5

Selection of the Event Topology

After reconstruction of the individual objects, it is important to only select the process of interest among all other events that are recorded. In order to do this, selection criteria are developed based on the topology of the final state of the physics process of interest. Carefully placed selections based on kinematic properties and physics objects are used in order to find a balance between an efficient rejection of the background processes and maintaining as much signal as possible. In Table 5.1, the full analysis chain is given, and in Tables 5.2 and and 5.3, the event rate after the full analysis chain can be found.

Table 5.1: The analysis chain

Criteria	ee	$e\mu$	$\mu\mu$
HLT	yes	yes	yes
primary vertex	yes	yes	yes
Exactly 2 opposite charge leptons, $p_T > 20$ GeV	yes	yes	yes
Loose lepton veto	yes	yes	yes
$m_{ll} > 20$ GeV	yes	yes	yes
m_{ll} outside $[81, 101]$ GeV	yes	no	yes
$E_T^{miss} > 50$ GeV	yes	no	yes
exactly 1 jet, $p_T > 30$ GeV	yes	yes	yes
exactly 1 b-tagged jet, $p_T > 20$ GeV (CSVM)	yes	yes	yes
$H_T > 160$ GeV	no	yes	no

5.1 Event Selection Criteria

Data collected from the dilepton High Level Trigger paths, given in Table 4.2, is taken as a base for the analysis. The event should contain exactly two opposite charge leptons descending from the primary vertex, with a transverse momentum higher than 20 GeV. Further, a loose lepton veto is set. Events where one or both leptons are loose leptons, defined in Table 4.8, are vetoed.

The Drell-Yann (Z/γ^*) events are removed by considering the mass spectrum of the invariant mass of the leptons m_{ll} (see Figure 5.1). In order to reduce background events, the invariant mass m_{ll} for the ee and $\mu\mu$ final states is required to be outside the Z mass window. In order to get rid of low invariant mass Drell-Yann events, for each final state, the invariant mass is required to be larger than 20 GeV. By doing this, also background events from the ZZ and ZW processes are reduced.

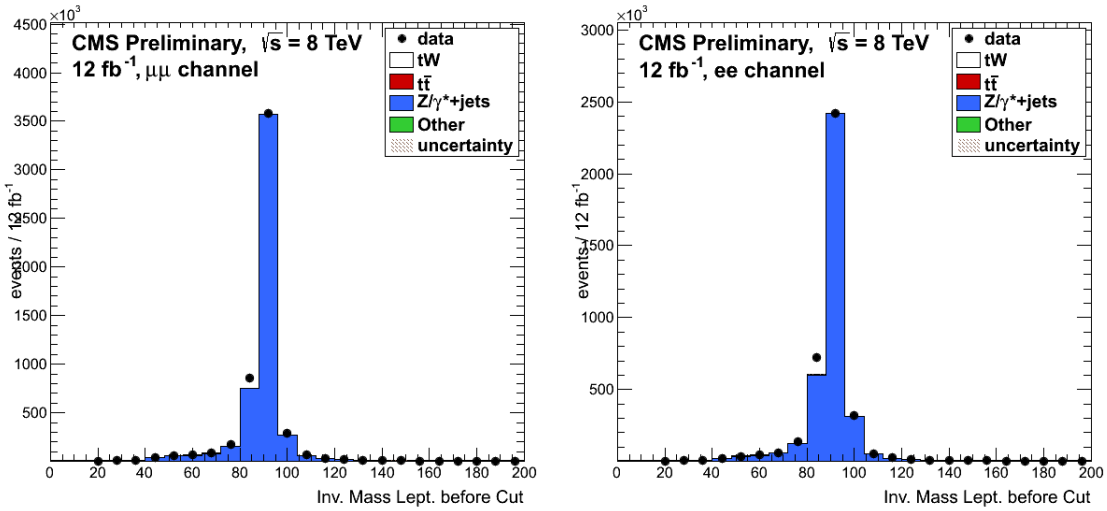


Figure 5.1: The distribution of the invariant mass of two leptons in the $\mu\mu$ (left) and ee (right) final state. After $m_{ll} > 20$ GeV and before $m_{ll} \in [81, 101]$ GeV, with the statistical uncertainty for all samples.

Then, for the ee and $\mu\mu$ final states, a selection is made on the missing transverse energy, in order to reduce the contribution from events without genuine missing transverse energy (mostly Z/γ^* + jets and QCD multi jet production). Since genuine missing transverse energy is expected due to the presence of neutrinos in the final state, the missing transverse energy is required to be larger than 50 GeV. In Figure 5.2 the missing energy distributions are given.

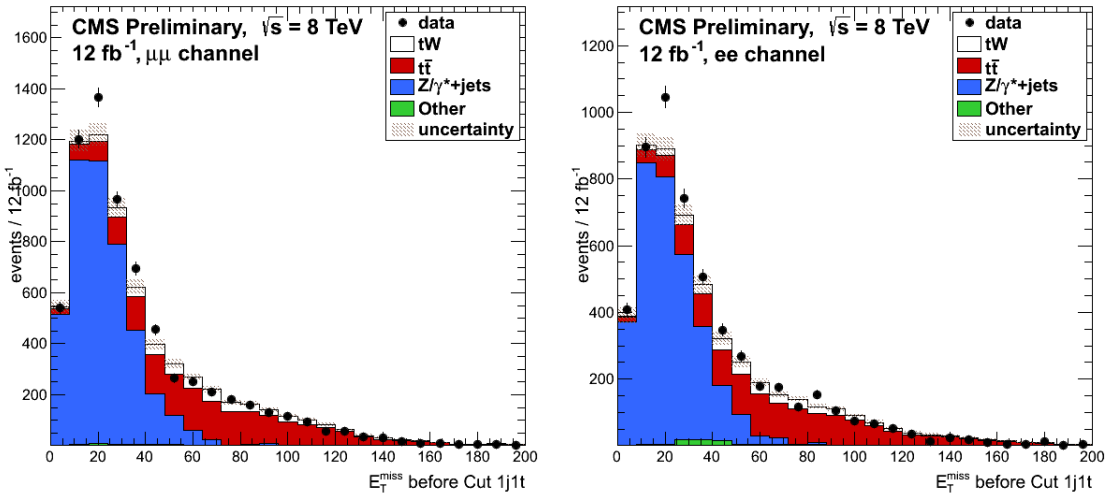


Figure 5.2: The distribution of the missing transverse energy in the $\mu\mu$ (left) and ee (right) final state. After $m_{ll} \in [81, 101]$ GeV and before $E_T^{miss} > 50$ GeV, with the statistical uncertainty for all samples.

Exactly one jet coming from a b quark is required to be present in the event. After this selection, the sample is fully dominated by $t\bar{t}$ background and tW signal. The distribution of the number of b-tagged jets after full event selection is given in Figure 4.5.

In the eμ final state, a selection is made on the scalar sum of all four-momenta (H_T) in the event. The H_T is required to be larger then 160 GeV, such that additional Drell-Yann background is removed. The distribution of H_T is given in Figure 5.3.

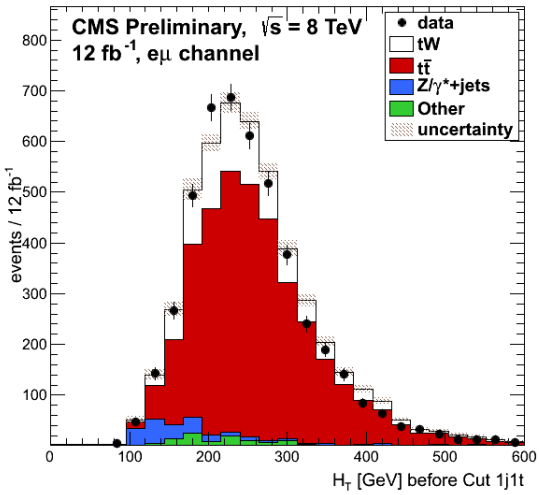


Figure 5.3: The distribution of H_T in the $e\mu$ final state. Before $H_T > 160$ GeV, after 1jet 1 b-tagged jet requirement, with the statistical uncertainty for all samples.

5.2 $t\bar{t}$ Control Regions

After the full signal event selection, the main background is $t\bar{t}$. This background does not only have a large contribution, but is also difficult to separate from the tW signal. As presented in [1], two control regions are established and included in the statistical fit in order to estimate the background in the signal region.
The control regions are defined as closely as possible to the signal region, but with the less amount of signal as possible. This is done by only changing the requirement on the number of jets. The two control regions have the same selections applied as presented in Table 5.1, with the exception that the events have exactly two jets with either one (2j1t) or both jets (2j2t) b-tagged. In Figure 5.4, the distribution of the regions is given and in Table 5.4 the event count can be found.

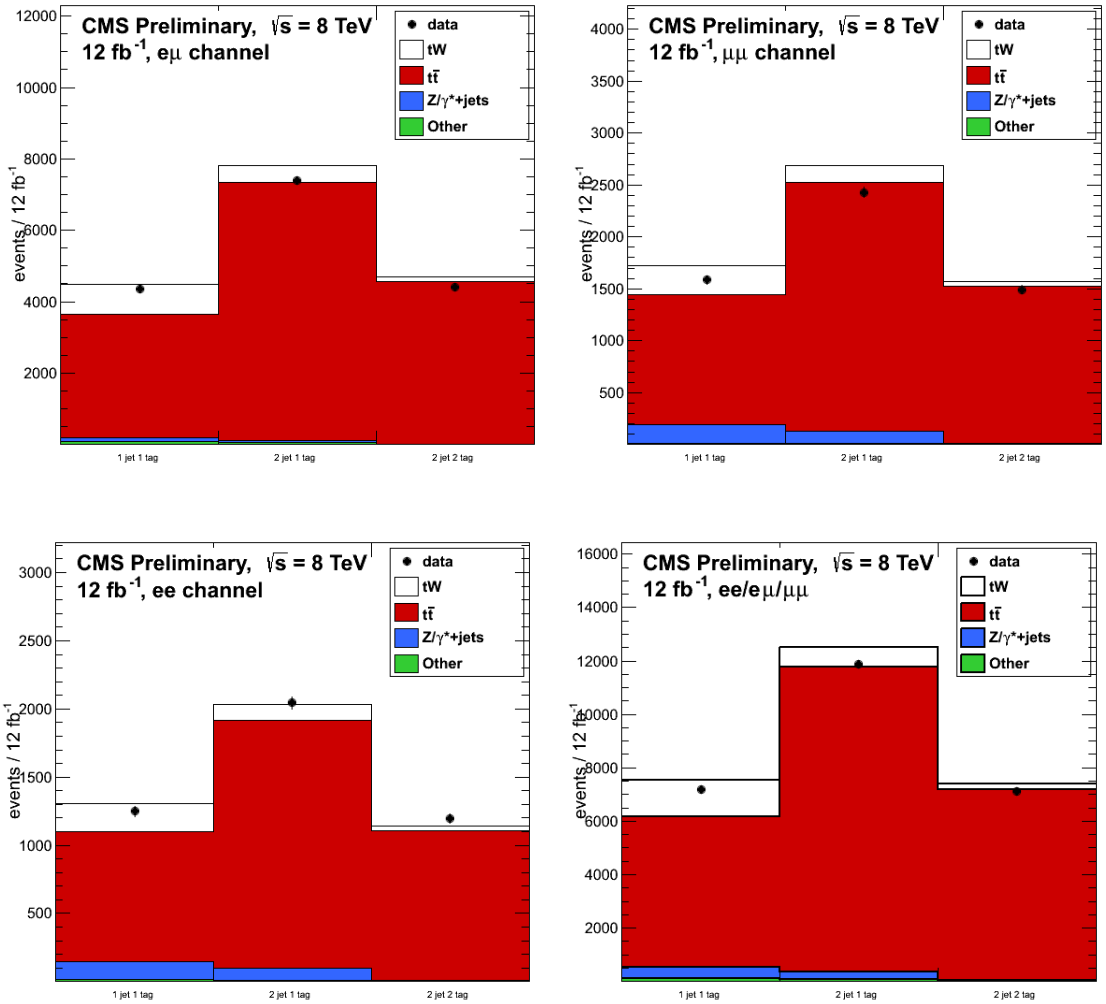


Figure 5.4: The distribution of the regions, in the $e\mu$ (up,left), $\mu\mu$ (up,right), ee (down, left) final state, and all final states together (down,right). After $m_{ll} > 20$ GeV

5.3 Tables with Event Rates

Table 5.2: Rates of events passing the event selection in the 1j1t signal region, with the statistical uncertainty for simulation.

	$e\mu$		$\mu\mu$		ee	
	data	MC	data	MC	data	MC
Lepton Sel.	68692	70212 ±266	5362514	5179534 ±2885	3933949	3771288 ±2416
Inv. Mass	68692	70212 ±266	720881	667115 ±1029	546924	506058 ±883
E_T^{miss}	68692	70212 ±266	29053	29478 ±185	22315	23513 ±171
1 Jet	14865	15078 ±124	6825	6929 ±90	5330	5730 ±86
b-tagging	4691	4846 ±53	1587	1723 ±33	1248	1309 ±28
H_T	4357	4495 ±50	1587	1723 ±33	1248	1309 ±28

Table 5.3: Rates of events passing the event selection in the 1j1t signal region for all final states together, with the statistical uncertainty for simulation and the remaining percentage of events after each selection compared to the original number of events in simulation.

	eμ/$\mu\mu$/ee		%			
	data	simulation	tW	$t\bar{t}$	Z+jets	Other
Lepton Sel.	9365155	17934120 ±5329	start	start	start	start
Inv. Mass	1336497	2397363 ±1939	92	92	13	13
E_T^{miss}	120060	173104 ±468	77	78	0.56	0.69
1 Jet	27020	38304 ±223	38	16	0.12	0.15
b-tagging	7526	8367 ±75	21	8.1	$5.5\ 10^{-3}$	$7.1\ 10^{-3}$
H_T	7192	7917 ±71	20	7.8	$4.4\ 10^{-3}$	$5.8\ 10^{-3}$

Table 5.4: The number of events in each region for data and Monte Carlo after full selection, with the statistical uncertainty.

	data	MC simulation		
	$e\mu$	All samples	tW	$t\bar{t}$
1 jet 1 tag	4357 ±66	4495±50	858 ±16	3451 ±41
2 jets 1 tag	7395 ±86	7803 ±63	458 ±12	7228 ±59
2 jets 2 tags	4416 ±66	4697 ±48	129 ±6	4558±47
	$\mu\mu$	All samples	tW	$t\bar{t}$
1 jet 1 tag	1586 ±40	1723 ±33	275 ±9	1259 ±25
2 jets 1 tag	2423 ±49	2683 ±39	155 ±7	2403 ±35
2 jets 2 tags	1492 ±39	1570 ±28	45 ±4	1515 ±28
	ee data	All samples	tW	$t\bar{t}$
1 jet 1 tag	1248 ±35	1309 ±28	209 ±8	956 ±22
2 jets 1 tag	2044 ±45	2030±33	113 ±6	1816 ±30
2 jets 2 tags	1193 ±35	1140 ±24	37 ±3	1096±23
	$e\mu/\mu\mu/ee$	All samples	tW	$t\bar{t}$
1 jet 1 tag	7191 ±85	7526 ±66	1343 ±20	5666 ±53
2 jets 1 tag	11862 ±109	12517 ±81	726 ±15	11447±75
2 jets 2 tags	7101 ±84	7408 ±60	211 ±8	7169 ±59

Chapter 6

Systematic Uncertainties

In this chapter the main systematic uncertainties and their effect on each final state will be quantified. The used uncertainties are based on the official recommendations of the Single Top group at CMS[1].

In order to perform to study of tW associated channel, a detailed theoretical understanding of the Standard Model and its uncertainties is required. The limited knowledge of the theory as well as the limited statistics will introduce systematic uncertainties. The modifications made to the simulated Monte Carlo samples (introduced in Section 4.4) introduce uncertainties via their scale factors.
The uncertainties that are accounted for in this study are:

- **Luminosity**: The luminosity value used in this analysis is originating from the absolute calibration of the luminosity measured at the CMS experiment. A value of ± 4.4 % is taken as flat rate uncertainty [97].
- **Pile-up reweighing**: The pile-up reweighing scale factors used are adjusted up and down by their uncertainties.
- **Lepton Trigger, Reconstruction, and Identification efficiency**: The efficiencies are scaled up/down by adding/subtracting their uncertainties.
- **The jet energy scale (JES)**: The jet energy corrections are adjusted by adding or subtracting one standard deviation from its central value.
- **The jet energy resolution (JER)**: The scaling of the momentum of the reconstructed jet is adjusted up and down by one standard deviation.
- **Modeling of the missing transverse energy**: The unclustered energy is scaled up/down by adding/subtracting ± 10%.
- **B-tagging**: The b-tagging scale factors used to make the efficiency in simulation match the one for data are adjusted up and down uncertainties.
- **Z/γ^* reweighing**: The Z/γ^* reweighing scale factors used, are doubled or left out in order to make respectively the up and down by their uncertainties.

- $t\bar{t}$ **cross-section**: The cross-section is only known at a limited order of perturbation theory and this will cause a systematic uncertainty on the measurement that is being performed[35]. An uncertainty of $\pm$ 6.7% on the cross-section of the production of top quark pairs is used.

- **PDF**: The limited knowledge of the distribution of the partons within the proton, described by the PDF, will be one of the major systematic uncertainties. A total of 44 new event weights are derived for each event from the 22 eigenvectors of CTEQ66[70]. The largest and smallest of these new weights were used for the plus and minus systematic uncertainties.

- **Factorization or Normalization scale (Q^2)**: The unphysical dependence on the factorization and normalization scales will be source of systematic uncertainties. Dedicated simulated samples with halved and doubled renormalization and factorization scales for the tW signal and $t\bar{t}$ background are produced.

- **ME/PS matching thresholds**: PYTHIA uses PS/ME matching and factorization scales (Q^2) in order to simulate the partonshower[69]. Another choice of these parameters will influence the analysis. Simulated samples for the top quark pair background are used with larger and smaller matching scale.

- **DR/DS scheme**: The choice of diagram removal or diagram subtraction schemes[15] will influence the outcome of the analysis. The difference between the Diagram Removal and Diagram Subtraction scheme for the tW simulated signal samples are accounted for as a systematic uncertainty.

- **Top quark mass**: Samples with top quark masses of 166.5 and 178.5 GeV are used to create uncertainty samples for the top quark mass.

- **Statistics of the simulated data**: The effect due to the size of the sample has been accounted for as a systematic uncertainty.

The systematic samples used in this analysis are given in Table 6.1.

Table 6.1: Summer12 Systematic samples used in the analysis.

Systematic		Sample
Factorization scale: $2m_t$	tW	TToDilepton tW-channel-DR scaleup 8TeV-powheg-tauola
Factorization scale: $2m_t$	$\bar{t}$W	TBarToDilepton tW-channel-DR scaleup 8TeV-powheg-tauola
Factorization scale: $2m_t$	$t\bar{t}$	TTJets scaleup TuneZ2star 8TeV-madgraph-tauola
Factorization scale: $m_t/2$	tW	TToDilepton tW-channel-DR scaledown 8TeV-powheg-tauola
Factorization scale: $m_t/2$	$\bar{t}$W	TBarToDilepton tW-channel-DR scaledown 8TeV-powheg-tauola
Factorization scale: $m_t/2$	$t\bar{t}$	TTJets scaledown TuneZ2star 8TeV-madgraph-tauola
mass: $m_t = 166.5$ GeV	tW	TToDilepton tW-channel-DR mass166 5 8TeV-powheg-tauola
mass: $m_t = 166.5$ GeV	$\bar{t}$W	TBarToDilepton tW-channel-DR mass166 5 8TeV-powheg-tauola
mass: $m_t = 166.5$ GeV	$t\bar{t}$	TTJets mass166 5 TuneZ2star 8TeV-madgraph-tauola
mass: $m_t = 178.5$ GeV	tW	TToDilepton tW-channel-DR mass178 5 8TeV-powheg-tauola
mass: $m_t = 178.5$ GeV	$\bar{t}$W	TBarToDilepton tW-channel-DR mass178 5 8TeV-powheg-tauola
mass: $m_t = 178.5$ GeV	$t\bar{t}$	TTJets mass178 5 TuneZ2star 8TeV-madgraph-tauola
PS/ME matching: up	$t\bar{t}$	TTJets matchingup TuneZ2star 8TeV-madgraph-tauola
PS/ME matching: down	$t\bar{t}$	TTJets matchingdown TuneZ2star 8TeV-madgraph-tauola

6.1 Jet Energy Scale and Jet Energy Resolution

The CMS collaboration uses jet energy calibration in order to relate the energy measured from detected jet with the energy of the true particle jet[87]. Since the calorimeter response is not linear, the translation of the measured jet energy to the true particle energy is not straightforward. Therefore, the CMS collaboration uses jet corrections, a set of tool that allows the mapping of the measured jet energy deposition to the parton's. More on this topic can be found at the Jet Energy Correction group in [86]. The jet energy scale measurement comes with an uncertainty that is given by the Jet Energy Correction group at CMS. Further, it is proven that the jet resolution in data is broader than the one obtained in simulation. For this reason, a smearing is applied on Monte Carlo. This smearing factor as well as the uncertainty on JES is accounted for in the statistical method.

6.2 Modeling of the Missing Transverse Energy

In order to calculate the missing transverse energy, conservation laws are being used.

Energy that is not clustered or is coming from a neutral particle, is called unclustered energy. This energy has the biggest uncertainty since the energy of the muon is very well known (it crosses all detectors), and the energies of the electron and jets can be corrected for[78][87]. Therefore, the uncertainty on the missing transverse energy is almost fully dominated by the uncertainty on the unclustered energy. This uncertainty on the unclustered energy is by default 10% of the energy[98]. Thus, by adding or subtracting 10% of the unclustered energy, the missing transverse energy goes up or down, letting more or less events through the selection. This is accounted for as a systematic uncertainty.

6.3 Contributions from Corrections of the MC Samples

In Section 4.4, the scale factors applied to the Monte Carlo simulation are discussed. The lepton trigger identification and isolation efficiencies, as well as pile-up reweighing scale factors, b-tagging, Z/γ^* reweighing scale factors come from independent studies where data is compared with Monte Carlo samples. Therefore, these scale factors come with an uncertainty[2][95][90].

6.4 Parton Density Function

The hadronic cross-section for the production of a final state X from the collision of two protons, $\sigma(pp \to X)$, can be written as a function of the differential subprocess cross-section of the partons i and j to produce the final state X, weighed with parton distribution functions [99]:

$$\sigma(pp \to X) = \sum_{i,j} \int \int dx_i dx_j \mathrm{PDF}_i(x_1, f_1, Q^2) \mathrm{PDF}_j(x_2, f_2, Q^2) \sigma_{ij \to X} \tag{6.1}$$

where f is the flavor of the parton, x_i is the fraction of the proton's momentum that parton i carries, and Q^2 is the factorization scale. The sum runs over all contributions from the different initial partons (flavor, spin, color, etc), while integrating over the momentum fractions of the partons x_i, x_j. The partonic cross-section $\sigma_{ij \to X}$ can be expanded in the strong coupling constant α_s, making a perturbative calculation of the total partonic cross-section possible.

The parton distribution functions can be obtained by means of global fits to experimental data from deep inelastic scattering, Drell-Yann and jet processes. Different collaborations, such as CTEQ[70], and MRST[100], provide updates when new data and/or theoretical predictions become available. The parton distribution functions used in the production of the Monte Carlo simulated event samples for this thesis is the CTEQ66 set.
For CTEQ PDF's, 40 uncertainty PDF's are provided [101]. These values for the different parton distribution functions provided can be accessed via the LHAPDF (Les Houches Accord Parton Distribution Function) library [102]. The PDF uncertainty determination can then be done by using the reweighing method[103].

6.5 Factorization and Normalization Scale

The partonic cross-section $\sigma_{ij \to X}$ and the parton density functions are functions of the factorization and renormalization scale Q^2. For heavy quark production, this scale is usually set to the order of the heavy quark mass[35], in this case the top quark ($Q = m_{top}$). If the complete perturbation series could be calculated, the result for the cross-section would be independent of Q. However, the calculations are done up

to finite order perturbation theory, and the cross-section in general do depend of Q. This dependence is in general tested by varying the scale between $Q = m_{top}/2$ and $Q = 2m_{top}$. More on this topic can be found in [104].

6.6 ME/PS Matching Thresholds

Pythia[69] uses two complementary methods in order to create the parton showers. The matrix element method (ME) and parton shower method (PS). In order to get a smooth transition between the two, matching is done. This matching is dependent from the normalization scale and this has to be accounted for in the statistical method. More on this subject can be found in [69].

6.7 DR/DS Scheme

As already mentioned in Chapter 1, the tW associated channel has interference at higher order with the top quark pair production (see Figure 1.3)[15]. The problem of unambiguously defining the two can be overcome by defining two schemes, the diagram removal (DR) and diagram subtraction scheme (DS). In the presented single top analysis, the DR approach is used. The differences between the two schemes are accounted in the systematic uncertainties.

Chapter 7

The Statistical Interpretation

The theory from Chapter 1 is tested with a statistical model that gives the probability to have a certain dataset as a function of underlying parameters, such as the cross-section of a process. The theta framework[105] provides a statistical model based on templates. It provides a test statistic distribution for a certain model, large scale Monte Carlo production of pseudo data and efficient calculation of the test statistic.

7.1 Discriminating Variable

The kinematic variable that is chosen as input for the theta framework has to discriminate between the tW signal and the main background ($t\bar{t}$) after full event selection. In order to find such a variable, the kinematic properties of the signal and $t\bar{t}$ events are studied. The considered processes are the following:

$$pp \to t\bar{t} \to b\bar{b}\nu_l\nu_l l^- l^+ \tag{7.1}$$
$$pp \to tW^- \to b\nu_l\nu_l l^- l^+ \tag{7.2}$$
$$pp \to \bar{t}W^+ \to \bar{b}\nu_l\nu_l l^+ l^- \tag{7.3}$$

Considering the kinematic properties of each sample, the transverse momentum of the system p_T^{system} is chosen as discriminating variable. In Figure 7.1, the distribution of this variable for tW signal and $t\bar{t}$ background can be found.

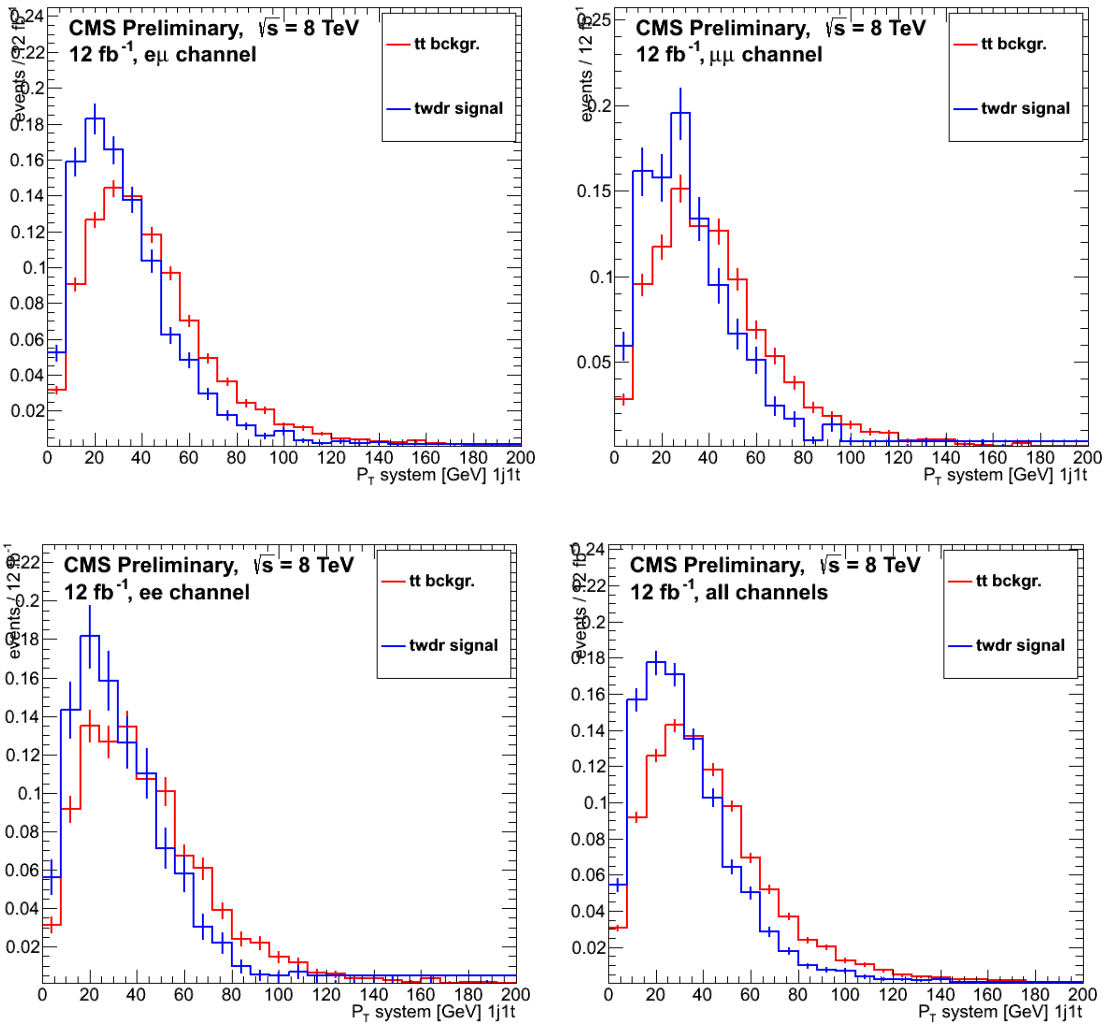


Figure 7.1: The normalized distribution of the transverse momentum of the system for tW signal (blue) and $t\bar{t}$ background (red) for each final state in the signal region, for the $e\mu$ (up,left), $\mu\mu$ (up,right), ee (down,left), and $e\mu/\mu\mu/ee$ (down,right) final states.

The distribution of the transverse momentum of the system p_T^{system}, for all samples and data can be found in Figure 7.2 (see Appendix A.1, for the separate channels).

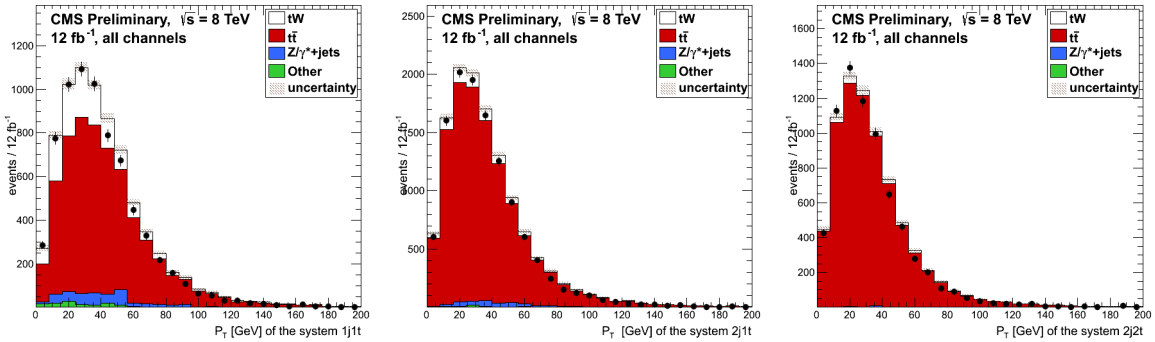


Figure 7.2: The distribution of the transverse momentum of the system for all final states, in the signal region (left), the 2j1t region (middle) and 2j2t region (right).

A study is performed on the effect of the systematic uncertainties on the distribution of this variable in each region with as goal a distinction of shape and rate uncertainties. The outcome from this study is:

- Rate uncertainties: pile-up reweighting, lepton Trigger, reconstruction, and identification efficiency, luminosity, and $t\bar{t}$ cross-section
- Shape uncertainties: JES, JER, b-tagging, Z/γ^* reweighting, modeling of the missing transverse energy, PDF, factorization or normalization scale (Q^2), ME/PS matching thresholds, DR/DS scheme, and top quark mass

The rate effect of each uncertainty (shape and rate) in the signal region can be found in Table 7.1, while the rate effect of each uncertainty in the control regions can be found in Appendix A.2.

Table 7.1: Rate impact of all considered systematic uncertainty sources in the 1j1t signal region, values as a percentage. The estimates are for each of the three channels, unless specified as separate values for each channel ($ee/e\mu/\mu\mu$). If two numbers are listed for a single uncertainty, the upper number is the effect on the rate when the systematic uncertainty source is scaled up and the lower is for when it is scaled down. Entries with a single value indicate that the systematic is symmetric between the scaled up and scaled down effects.

Systematic Uncertainty (ee/$e\mu$/$\mu\mu$)	tW (%)	$t\bar{t}$ (%)	Other (%)
Luminosity	± 4.4	± 4.4	± 4.4
Event pile up	± 0.38/± 0.03/± 0.02	± 0.46/± 0.34/± 0.16	± 1.54/± 1.61/± 2.81
Lepton identification	± 2.33/± 1.86/ ± 2.28	± 2.33/± 1.86/± 2.28	± 2.33/± 1.86/± 2.28
JES	$^{-3.51}_{+1.68}$ / $^{-3.11}_{+1.15}$ / $^{-1.60}_{+2.01}$	$^{-6.50}_{+6.85}$ / $^{-6.71}_{+6.64}$ / $^{-6.77}_{+7.37}$	$^{+12.75}_{+9.80}$ / $^{+3.75}_{+4.11}$ / $^{+13.33}_{+12.08}$
JER	$^{+0.75}_{+1.15}$ / $^{-0.45}_{+0.32}$ / $^{+0.37}_{-0.12}$	$^{-0.14}_{+0.46}$ / $^{-0.35}_{+0.27}$ / $^{-0.33}_{-0.73}$	$^{+8.38}_{-4.55}$ / $^{+3.87}_{+0.01}$ / $^{-0.05}_{-3.01}$
MET modeling	$^{+4.99}_{-6.63}$ / $^{+0.37}_{-0.54}$ / $^{+4.94}_{-5.81}$	$^{+5.86}_{-6.05}$ / $^{+0.56}_{-0.41}$ / $^{+4.98}_{-5.88}$	-
B-tagging data/MC scale factor	$^{+3.61}_{-3.17}$ / $^{+2.61}_{-2.42}$ / $^{+2.55}_{-2.57}$	$^{+2.11}_{-1.78}$ / $^{+1.71}_{-1.69}$ / $^{+1.52}_{-1.69}$	$^{+3.38}_{-3.89}$ / $^{+0.97}_{-0.51}$ / $^{+4.54}_{-0.36}$
Z+jet MET scale factor	-	-	$^{+16.90}_{-16.90}$ / $^{+7.16}_{-7.16}$ / $^{+19.59}_{-19.59}$
$t\bar{t}$ cross-section	-	± 6.71	-
PDF	$^{+6.06}_{-5.94}$ / $^{+5.58}_{-5.61}$ / $^{+5.67}_{-5.66}$	$^{+5.15}_{-4.93}$ / $^{+5.16}_{-4.89}$ / $^{+5.30}_{-5.02}$	$^{+3.56}_{-3.79}$ / $^{+4.01}_{-4.11}$ / $^{+3.95}_{-4.06}$
Q^2 scale	$^{-3.93}_{-8.02}$ / $^{-6.42}_{-8.79}$ / $^{-4.67}_{-6.35}$	$^{+7.39}_{-13.95}$ / $^{+9.56}_{-11.70}$ / $^{+3.38}_{-13.41}$	-
ME/PS matching thresholds	-	$^{-11.37}_{-5.19}$ / $^{-0.52}_{+0.89}$ / $^{-6.50}_{-3.62}$	-
tW DR/DS scheme	$^{-1.40}_{+0.00}$ / $^{-2.94}_{+0.00}$ / $^{-0.42}_{+0.00}$	-	-
Top quark mass	$^{-2.31}_{-7.71}$ / $^{-5.79}_{-10.87}$ / $^{-4.10}_{-9.58}$	$^{-6.11}_{-4.06}$ / $^{-4.65}_{-15.99}$ / $^{-4.20}_{-3.15}$	-

7.2 Statistical Method

The expected event yield for bin i, k_i, is given by the sum over all considered background processes and the signal, scaled with a signal strength modifier μ_S.

$$k_i = \sum_{p=1}^{N_{process,i}} k_{p,i} + \mu_S \times k_{S,i} \tag{7.4}$$

The systematic uncertainties presented in Chapter 6, will affect this event yield k_i. Thus, for each independent source of systematic uncertainty u, a nuisance parameter θ_u is introduced. Rate uncertainties are modeled with a bin independent coefficient for the template β_p, with a log normal prior. Shape uncertainties are modeled by choosing a Gaussian prior for θ_u, and using this parameter to interpolate between the nominal, without systematic uncertainties, template and the shifted templates, which are obtained by applying a plus or minus one sigma systematic shift to the simulated samples and deriving the templates. The uncertainty due to the limited event numbers is accounted for by using the Barlow Beeston lite method (BBL[106][107]). Following

BBL, a Gaussian is used for the distribution of the number of selected events in bin i, for process p after event selection, $N_{i,p,sel}$. Per bin one additional nuisance parameter ν_i is introduced, which denotes the additive shift of the estimated mean k_i to the true mean $k_{i,true}$, which would be the mean when one has an infinite number of events.

$$k_{i,true}(\vec{\theta}, \nu_i) = k_i(\vec{\theta}) + \nu_i \tag{7.5}$$

This additive shift follows a Gaussian prior around zero with width Δ_i.

The complete statistical model can be written as

$$P(X = N|\mu, \vec{\theta}, \vec{\nu}) = \prod_{i=1}^{N_{bin}} \text{Poisson}(X_i = N_i|k_i(\mu, \vec{\theta}) + \nu_i) \tag{7.6}$$

$$k_i(\mu, \vec{\theta}) = \left(\sum_{p=1}^{N_{process}} k_{i,p}(\vec{\theta}) \right) + \mu \times k_{i,S}(\vec{\theta}) \tag{7.7}$$

$$\Pi(\vec{\theta}) = \prod_{u=1}^{N_{syst}} \text{Gauss}(\theta_u|0, 1) \tag{7.8}$$

$$\Pi(\vec{\nu}) = \prod_{i=1}^{N_{bins}} \text{Gauss}(\nu_i|0, \Delta_i) \tag{7.9}$$

$$L(\vec{\theta}, \mu, \vec{\nu}|N) = P(N|\mu, \vec{\theta}) \prod_{u=1}^{N_{syst}} \text{Gauss}(\theta_u|0, 1) \prod_{i=1}^{N_{bin}} \text{Gauss}(\nu_i|0, \Delta_i) \tag{7.10}$$

where Equation 7.6 is the Poisson probability with for each bin the Poisson mean k_i, Equation 7.8 is the prior for nuisance parameters $\vec{\theta}$, Equation 7.9 is the prior for nuisance parameters $\vec{\nu}$, and Equation 7.10 is the binned likelihood.

The significance is estimated by the use of a goodness-of-fit test. For this, a test statistic is defined as

$$q_0 = \frac{\partial}{\partial \mu_S} \left(\max_{\theta_0} L(\mu_S = 0, \vec{\theta_0}|\text{data}) \right) \tag{7.11}$$

where $L(\mu_S = 0, \vec{\theta_0}|\text{data})$ is the likelihood defined in Equation 7.10, and $\vec{\theta_0}$ are the nuisance parameters where the ones corresponding to theory (DR/DS scheme, ME/PS matching thresholds, Q^2 scale, PDF, and top quark mass) are set to the central value of zero during the maximization. The p-value is defined as the probability to have a test statistic q_0 at least as high as the one observed in data, under the background only hypothesis ($\mu_S = 0$). Therefore, pseudo data has to be created for the background only hypothesis, and the nuisance parameters are randomized in each pseudo data set, including those corresponding to theory.
The profile likelihood function L_p carries out the maximization over all nuisance parameters, while the theoretical shape uncertainties are not accounted for:

$$L_p(\mu|N) = \max_{\vec{\theta}} \left[L(\mu_S, \vec{\theta}, 0|N) \right] \tag{7.12}$$

with the likelihood taken from Equation 7.10. The estimated signal strength modifier $\hat{\mu}_S$, is the value of μ_S that maximizes the profile likelihood given in Equation 7.12. The confidence level intervals of 68% (one sigma) are evaluated using this profile likelihood method, without the theoretical shape uncertainties, based on the probability given by the statistical model (Equation 7.6)[108]. The outcome will be some central value for the signal cross-section ($\hat{\mu}$), and some plus/minus uncertainties: $\sigma^{\Delta_{PL,+}}_{\Delta_{PL,-}}$. The uncertainties coming from these theory shape parameters, should be added at the end by hand.

7.3 Results

In Table 7.2, the systematic uncertainties extracted by fixing the sources one at a time and measuring the difference in the cross-section uncertainty is presented.

Table 7.2: Systematic uncertainties extracted by fixing sources one at a time and measuring difference in cross section uncertainty with the theory uncertainties externalized from the fit.

Systematic Uncertainty	$\Delta\sigma$ (pb)	$\frac{\Delta\sigma}{\sigma}$
Luminosity	1.26	0.06
Event pile up	0.11	0.01
Lepton reco/trig/id	0.62	0.03
JES	2.40	0.11
JER	0.56	0.03
MET modeling	1.42	0.06
B-tagging data/MC scale factor	1.09	0.05
Z+jet data/MC scale factor	0.78	0.04
$t\bar{t}$ cross section	0.60	0.03
PDF	1.55	0.07
Q^2 scale	5.49	0.25
ME/PS matching thresholds	4.76	0.21
tW DR/DS scheme	0.63	0.03
Top quark mass	2.49	0.11
Statistical	3.29	0.15

The cross-section of tW is measured to be $24.3^{+8.6}_{-8.8}$ pb . The uncertainty on the result is large due to externalizing the theoretical uncertainties (see Table 7.3), since these uncertainties are handled in a conservative way.

Table 7.3: Results of the fit of for the $e\mu/\mu\mu/ee$ final state .

Theory uncertainties in fit	Observed Significance	Expected Significance	Cross-Section of tW (pb)
Excluded	4.0	$3.2^{+0.4}_{-0.9}$	$24.3^{+8.6}_{-8.8}$
Included	3.9	3.3 ± 0.4	$25.1^{+4.2}_{-4.0}$

There is an excess over the expected background of 4.0σ, with an expected significance of $3.2^{+0.4}_{-0.9}\sigma$ for the tW signal. The effect of the systematic uncertainties which is not only large by itself, but also conservative, is what prevents discovery. This is illustrated in Figure 7.3 and easily proven by calculating the expected significance, taking only the statistical uncertainties into account:

$$\text{sign} = \frac{S}{\sqrt{S + B + (\Delta B)^2}} = \frac{1343}{\sqrt{1343 + 6183 + 69^2}} = 12\sigma \tag{7.13}$$

where S is the number of signal events in simulation for the signal region, B is the number of background events in simulation for the signal region, and ΔB the statistical uncertainty on the number of background events.

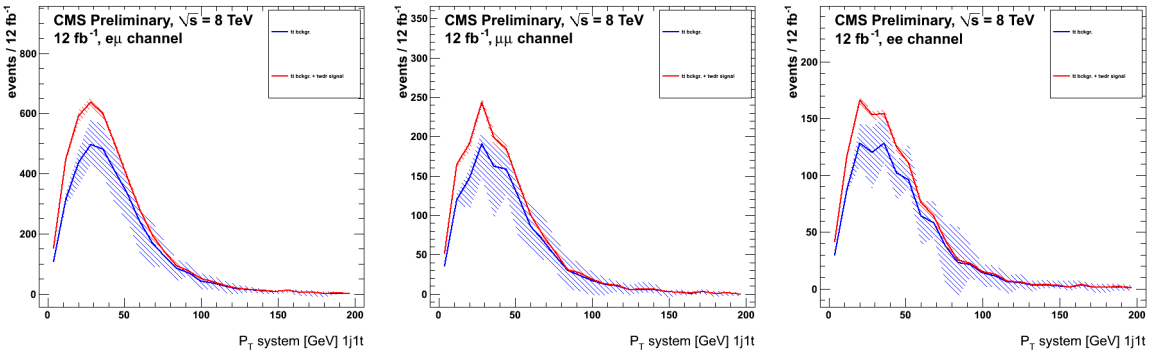


Figure 7.3: The distribution of the transverse momentum of the system in the signal region for all the $e\mu$ (left), $\mu\mu$ (middle) and ee (right) final state. In red the $t\bar{t}$ sample with its total systematic uncertainty is shown, and in blue the $t\bar{t}$ background with tW signal and the total systematic uncertainty on tW is shown.

Based on the assumption that the CKM matrix elements $|V_{ts}|$ and $|V_{td}|$ are much smaller than $|V_{tb}|$, the absolute value of the CKM matrix element $|V_{tb}|$ is calculated using following formula:

$$|V_{tb}| = \sqrt{\frac{\sigma_{tW}}{\sigma_{tW}^{theor}}} = 1.04 \pm 0.20 \text{ (exp.)} \quad \pm 0.04 \text{ (th.)} \tag{7.14}$$

where σ_{tW}^{theor} is the theoretical cross-section of tW assuming that $|V_{tb}|$ is unity.
Under the Standard Model assumption of $0 \leq |V_{tb}| \leq 1$, a value of $|V_{tb}|$ equal to 0.999 is found, with a 90% confidence interval of $[0.763, 1.000]$, and a 68% confidence interval of $[0.795, 1.000]$.

7.4 Conclusions and Outlook

For an excess over the expected background of 4.0σ, with an expected significance of $3.2^{+0.4}_{-0.9}\sigma$, the cross-section for tW is $24.3^{+8.6}_{-8.8}$ pb . There is evidence of the tW associated production, but the effect of the systematic uncertainties which is not only large by itself, but also conservative, is what prevents discovery.

The CKM matrix element $|V_{tb}|$ is measured to be 1.04 ± 0.20 (exp.) ± 0.04 (th.), when asuming that the CKM matrix elements $|V_{ts}|$ and $|V_{td}|$ are much smaller than $|V_{tb}|$. Under the Standard Model assumption of $0 \leq |V_{tb}| \leq 1$, a value of $|V_{tb}|$ equal to 0.999 is found, with a 90% confidence interval of $[0.763, 1.000]$, and a 68% confidence interval of $[0.795, 1.000]$.

The main uncertainties are coming from the factorization and normalization scale and PS/ME matching thresholds. Understanding better the effect of these uncertainties would allow the analysis to consider them in a less conservative way, improving the result. Next in line is the statistics of the sample. Including the full 8 TeV luminosity would improve this uncertainty.

The analysis of the CMS single top group[1] used a cut-based analysis and a shape analysis using a boosted decision tree (BDT) at 7 TeV with a luminosity of 4.9 fb^{-1}. For the cut-based analysis, a cross-section of 15 ± 5 pb is observed with a significance of 3.5σ and an expected significance of $3.2 \pm 0.9\sigma$. The BDT analysis observed a cross-section of 16^{+5}_{-4} pb with a significance of 4σ and an expected significance of $3.6^{0.8}_{-0.9}\sigma$. The main difficulty of this analysis is separating the signal from the very large background contributions (mainly $t\bar{t}$). For this reason a cut-based analysis is not powerful enough. With the proposed shape analysis, the transverse momentum of the system is chosen as discriminating variable for tW signal and the main background $t\bar{t}$. This allows better sensitivity than the cut-based analysis.
In order to achieve a larger significance, all possible information against the background has to be used. Therefore, building a multivariate discriminator would improve the result.

References

[1] The CMS Collaboration, Phys. Rev. Lett. **110.022003** (2010) CMS-TOP-11-022, CERN-PH-EP-2012-266.

[2] The CMS Collaboration, Pardos and J. Kieseler, *Dilepton trigger and lepton identication efciencies for the top quark pair production cross section measurement at 8 TeV in the dilepton decay channel*, CMS AN AN-12-389, 2012.

[3] N. KIDONAKIS, *Differential and total cross-sections for top pair and single top production*, `hep-ph/1205.3453v1`.

[4] *The Large Hadron Collider*, `http://www.lhc.ac.uk/`.

[5] The ATLAS Collaboration, Phys. Lett. **B716** (2012) 142–159. CERN-PH-EP-2012-117.

[6] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. USA: Addison-Wesley, 1995.

[7] D. Perkins, *Particle Astrophysics*, vol. 2nd edition.

[8] T. I. Hisaki Hatanaka and C. Lim, *The Gauge Hierarchy Problem and Higher Dimensional Gauge Theories*,.

[9] *The Large Electron Positron (LEP)collider*, `http://home.web.cern.ch/about/accelerators/large-electron-positron-collider/`.

[10] *European Organization for Nuclear Research*, `http://home.web.cern.ch/`.

[11] *Particle Data Group*, `http://pdg.lbl.gov/`.

[12] The ATLAS Collaboration, *Search for s-Channel Single Top-Quark Production in pp Collisions at (s) = 7 TeV*, Technical Report ATLAS-CONF-2011-118, 2011.

[13] The CMS Collaboration, Phys. Rev. Lett. **107:091802**.

[14] The ATLAS Collaboration, Phys. Lett. **B717** (2012) 330–350. CERN-PH-EP-2012-082.

[15] P. M. B. W. C. D. W. S. Frixione, E. Laenen, JHEP **0807:029** (2008).

[16] The CDF Collaboration, *Combination ofthe W boson polarization measurements from top quark decay using dilepton events with 5.1 fb*$^{-1}$, CDF note 10543.

[17] The ATLAS Collaboration, *Search for CP violation in single top quark events in pp collisions at* $\sqrt{s}$ *= 7 TeV with the ATLAS detector*, ATLAS-CONF-2013-032.

[18] J. W. Q.-H. Cao and C.-P. Yuan, Phys. Lett. **B658** (2007) 50–56.

[19] Martin, *A Super Symmetry primer*, 2011.

[20] The ATLAS Collaboration, *Search for FCNC single top-quark production at sqrt(s) = 7 TeV with the ATLAS detector*,.

[21] The CMS Collaboration, *Search for FCNC in top quark decays in pp collisions at sqrt(s)=7 TeV*,.

[22] The CMS Collaboration, *Search for* $H \to WW \to 2l2\nu$, CMS-PAS HIG-11-014, 2011.

[23] V. Barger, M. McCaskey, and G. Shaughnessy, Phys.Rev. **D81** (2010) 034020.

[24] *Interuniversity Institute for High Energies*, `http://w3.iihe.ac.be/`.

[25] The CMS Collaboration, *Top PAG Reference Selections*, `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiTopRefEventSel`.

[26] The CDF Collaboration, Phys. Rev. Lett. **74** (Apr, 1995) 2626–2631.

[27] The D0 Collaboration, Phys. Rev. Lett. **74** (Apr, 1995) 2632–2637.

[28] Chadwick, Nature **192** (1932) 321.

[29] The Super-Kamiokande Collaboration, Phys. Rev. Lett **81** (1998) 1562–1567.

[30] The CMS Collaboration, Phys. Rev. Letters **D 86** (2012).

[31] The CMS Collaboration, *Observation of a new particle with a mass of 125 GeV*, `http://cms.web.cern.ch/news/observation-new-particle-mass-125-gev`.

[32] *The CDF experiment*, `http://www-cdf.fnal.gov/physics/physics.html`.

[33] *The D0 experiment*, `http://www-d0.fnal.gov/`.

[34] *The Tevatron Collider*, `http://www-bdnew.fnal.gov/tevatron/`.

[35] W. Wagner, *Top quark physics in hadron collisions*,.

[36] T. Teubner, Acta Phys Pol **B30** (1999) 1941–1966.

[37] v. E. B. Anselmo F and B. G, Phys Rev Lett **D45** (1992) 2312–2322.

[38] B. A. S. Heinson A P and Boos, Phys Rev Lett **D 56** (1997) 3114–3128. INP-MSU-96-41/448, UCR/96-25.

[39] A. Quadt, Eur. Phys. J. **C48** (2006) 835–1000.

[40] The CDF Collaboration, Phys. Rev. Lett. (May, 2009).

[41] W. S. S. D and D. D. A, *Production of Heavy Quarks from W Gluon Fusion*, 1986.

[42] D. Sally and W. S. S. D, *Heavy fermion production in the effective W approximation*, 1987.

[43] E. L. F. M. Chris D. White, Stefano Frixione, *Isolating Wt production at the LHC*,.

[44] *The CMS experiment*, `http://cms.web.cern.ch/`.

[45] *Heavy ions and quark gluon plasma at CERN*, `http://home.web.cern.ch/about/physics/heavy-ions-and-quark-gluon-plasma/`.

[46] The LHC Collaboration, *LHC Design Report*, chapter 17.

[47] *Taking a closer look at the LHC*, `http://www.lhc-closer.es/1/4/9/0`.

[48] *Synchrotron radiation*, `http://hyperphysics.phy-astr.gsu.edu/hbase/particles/synchrotron.html`.

[49] J. P. Blewett, *200 GeV intersecting storage accelerators*, Proceedings of the 8^{th} International Conference on High-Energy Accelerators.

[50] *Public CMS Luminosity Information*, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[51] *LHC images*, `http://lhc-machine-outreach.web.cern.ch`.

[52] *The ATLAS experiment*, `http://atlas.ch/`.

[53] *The LHCb experiment*, `http://lhcb.web.cern.ch/lhcb/`.

[54] *The ALICE experiment*, `http://aliceinfo.cern.ch/`.

[55] *The TOTEM experiment*, `http://totem.web.cern.ch/Totem/`.

[56] *The LHCf experiment*, `http://home.web.cern.ch/about/experiments/lhcf`.

[57] *CMS collaboration*, `http://cms.web.cern.ch/content/cms-collaboration`.

[58] M. Della Negra, L. Foà, A. Hervé, and A. Petrilli, *CMS physics Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2006.

[59] The CMS Collaboration, *Track reconstruction in the CMS Tracker*, CMS-NOTE-2006-041.

[60] The CMS Collaboration, *CMS ECAL Wiki Pages*, https://twiki.cern.ch/twiki/bin/viewauth/CMS/ECALWikiHome, 2013.

[61] The CMS Collaboration, *CMS Hadron Calorimeter*, http://cmshcal.web.cern.ch/cmshcal/, 2013.

[62] The CMS Collaboration, *Muon Detectors*, http://cms.web.cern.ch/news/muon-detectors, 2013.

[63] The CMS Collaboration, *CMS trigger system*. 2004.

[64] The CMS Collaboration, *The CMSSW Application Framework*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework, 2006.

[65] E. E. B. A. S. Belyaev and L. V. Dudko, Phys Rev Lett D 59 (1999).

[66] P. N. S. FRIXIONE and C. OLEARI, JHEP **11** (2007) 70.

[67] T. Stelzer and W. F. Long, Comput. Phys. Commun. **81** (1994) 357.

[68] F. Maltoni and T. Stelzer, JHEP **0302** (2003) 027.

[69] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05** (2006) 026.

[70] J. P. et al, JHEP **7** (2007) 12.

[71] The GEANT4 Collaboration, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[72] The CMS Collaboration, *The total top quark pair production cross-section at hadron colliders through O(alpha$_S^4$)*, 1303.6254. CERN-PH-TH/2013-056, TTK-13-08.

[73] The CMS Collaboration, *Particle flow event reconstruction*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookAnalysisOverview, 2005.

[74] The CMS Collaboration, *Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV*, PFT-10-002, 2010.

[75] The CMS Collaboration, JINST. **7** (2012) CMS-MUO-10-004; CERN-PH-EP-2012-173.

[76] The CMS Collaboration, *Muon Local Reconstruction Offline Guide*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonSystemReco.

[77] M. Friedl, *CMS detector layout*, http://www.hephy.at/user/friedl/diss/html/node8.html.

[78] The CMS Collaboration, *Reconstructions of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC*, 2005.

[79] M. Pioppi, *A pre-identification for electron reconstruction in the CMS particle-flow algorithm*, 2008.

[80] A. Hoecker, *TMVA: Toolkit for multivariate data analysis*, arXiv0703039.

[81] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406** (1993) 187.

[82] The CMS Collaboration, *Jet Reconstruction in CMSSW*, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBook130JetAnalysis#JetAlgo`.

[83] G. S. Matteo Cacciari, Gavin P. Salam, *FastJet 2.4.1 user manual*, `http://www.lpthe.jussieu.fr/~salam/fastjet/repo/fastjet-doc-2.4.1.pdf`.

[84] The CMS Collaboration, *Plans for Jet Energy Corrections at CMS*, CMS PAS JME-07/002.

[85] A. Bhatti *et al.*, *Plans for Jet Energy Corrections at CMS*, CMS AN-2007/055.

[86] The CMS Collaboration, *Jet Energy Resolution and Corrections Subgroup*, `https://twiki.cern.ch/twiki/bin/view/CMS/JetEnergyScale`.

[87] The CMS Collaboration, JINST 6 (2011).

[88] The CMS Collaboration, *Calorimeter Jet Quality Criteria for the First CMS Collision Data*, CMS PAPER JME-09-008, 2009.

[89] The CMS Collaboration, *Jet Identification*, `https://twiki.cern.ch/twiki/bin/view/CMS/JetID#Related_Topics`, 2012.

[90] The CMS Collaboration, *b-tag and vertexing physics object group*, `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG`, 2013.

[91] *B-jet production cross section at cdf*, `http://gov.docdat.com/docs/10133/index-10479.html`.

[92] *Quantum Diaries*, `http://www.quantumdiaries.org`.

[93] The CMS Collaboration, *Results on b-tagging identification in 8 TeV pp collisions*, CMS DP -2013/005, 2013.

[94] The CMS Collaboration, *Pile up studies*, `https://twiki.cern.ch/twiki/bin/view/CMS/PileupInformation`, 2013.

[95] M. Hildreth, *Pile up reweighing*, `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData`, 2012.

[96] U. Heinz, *Algorithm to adjust the b-tagging efficiency in simulated events*, `https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFUtil`, 2012.

[97] The CMS Collaboration, *CMS Luminosity Based Pixel Cluster Counting - Summer 2012 Update*, CMS Physics Analysis Summary CMS-PAS-LUM-12-001.

[98] *Official Prescription for calculating uncertainties on Missing Transverse Energy (MET)*,.

[99] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization of Hard Processes in QCD*, `hep-ph/0409313`.

[100] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Phys. Lett. **B 604** (2004) 61.

[101] J. Pumplin *et al.*, Phys. Rev. **D65** (2002) 014013.

[102] D. B. M. R. Whalley and R. C. Group, *The Les Houches Accord PDFs (LHAPDF) and Lhaglue*, `hep-ph/0508110`. `http://projects.hepforge.org/lhapdf/`.

[103] C. H. A. M. H. P. P. Biallass, T. Hebbeker, *Parton Distribution Uncertainty Determination within CMSSW*,.

[104] S. D. E. Collins J C and S. G, Nucl Phys **B 263** (1986) 37–60.

[105] J. W.-K. Thomas Muller, Jochen Ott, *Theta - a framework for template-based statistical modeling and inference*, `http://theta-framework.org/`, 2010.

[106] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples.* 1993. Phys. Commun.

[107] J. S. Conway, *Nuisance Parameters in Likelihoods for Multisource Spectra.*, `1103.0354`.

[108] F. James, *Statistical Methods in experimental physics.* World Scientific, 2nd ed.

Appendix A

Study of p_T^{system}

A.1 Distributions for Each Final State

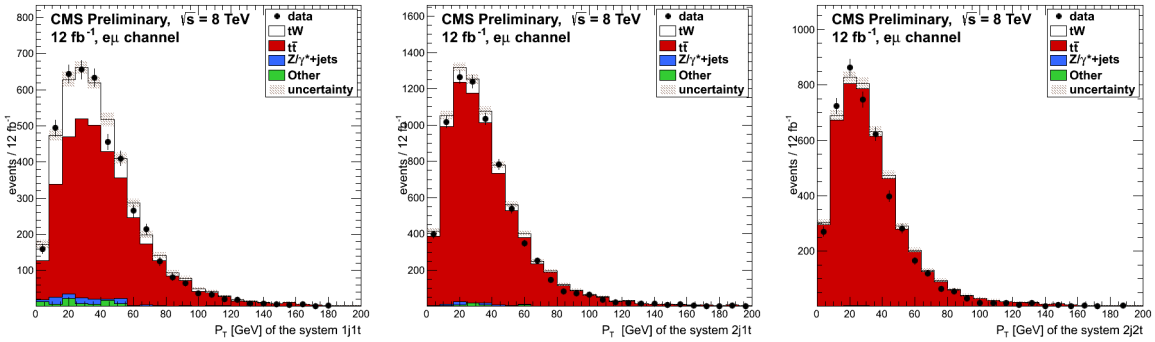


Figure A.1: The distribution of the transverse momentum of the system for the $e\mu$ final state, in the signal region (left), the 2j1t region (middle) and 2j2t region (right).

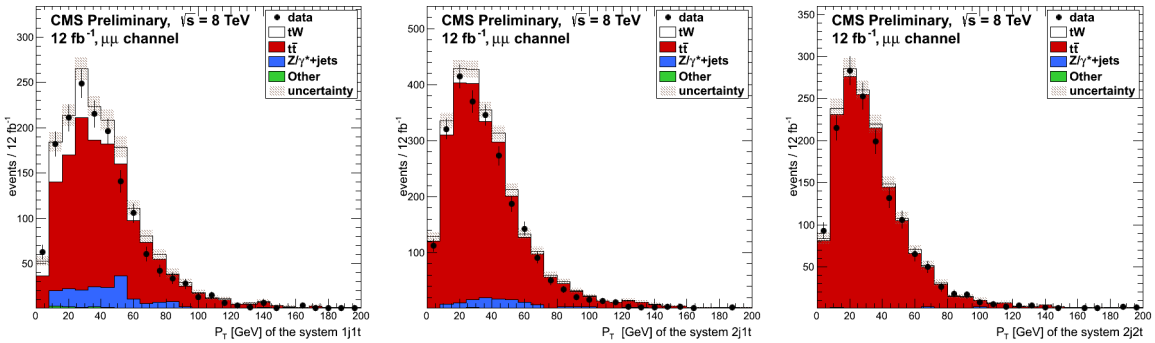


Figure A.2: The distribution of the transverse momentum of the system for $\mu\mu$ final state, in the signal region (left), the 2j1t region (middle) and 2j2t region (right).

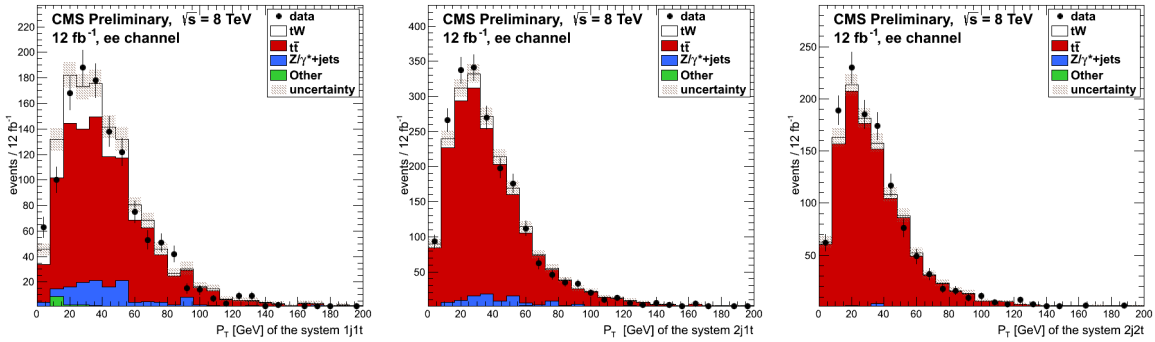


Figure A.3: The distribution of the transverse momentum of the system for *ee* final state, in the signal region (left), the 2j1t region (middle) and 2j2t region (right).

A.2 Rate Impact of the Systematic Uncertainties in the Control Regions

Table A.1: Rate impact of all considered systematic uncertainty sources in the 2j1t control region, values as a percentage. The estimates are for each of the three channels, unless specified as separate values for each channel ($ee/e\mu/\mu\mu$). If two numbers are listed for a single uncertainty, the upper number is the effect on the rate when the systematic uncertainty source is scaled up and the lower is for when it is scaled down. Entries with a single value indicate that the systematic is symmetric between the scaled up and scaled down effects.

Systematic Uncertainty (ee/$e\mu/\mu\mu$)	tW (%)	$t\bar{t}$ (%)	Other (%)
Luminosity	± 4.4	± 4.4	± 4.4
Lepton identification	± 2.33/± 1.86/ ± 2.28	± 2.33/± 1.86/± 2.28	± 2.33/± 1.86/± 2.28
JER	$^{+0.39}_{-0.58}$ / $^{+0.33}_{+0.44}$ / $^{-0.50}_{-0.62}$	$^{+0.11}_{-0.44}$ / $^{-0.10}_{+0.25}$ / $^{+0.26}_{+0.25}$	$^{+7.65}_{-0.02}$ / $^{+0.44}_{-3.63}$ / $^{+13.96}_{-8.72}$
JES	$^{+8.45}_{-1.76}$ / $^{+3.72}_{-3.36}$ / $^{+3.56}_{-2.90}$	$^{+0.06}_{-0.31}$ / $^{-0.96}_{-0.57}$ / $^{-0.25}_{-0.48}$	$^{+19.71}_{+10.67}$ / $^{-5.09}_{-7.14}$ / $^{+19.03}_{+8.91}$
MET modeling	$^{+6.81}_{-3.88}$ / $^{-0.17}_{+0.00}$ / $^{+6.69}_{-7.19}$	$^{+5.72}_{-6.41}$ / $^{-0.07}_{+0.02}$ / $^{+5.48}_{-5.56}$	-
Event pile up	± 0.38/± 0.03/± 0.02	± 0.46/± 0.34/± 0.16	± 1.54/± 1.61/± 2.81
B-tagging data/MC scale factor	$^{+1.46}_{-0.14}$ / $^{+1.18}_{-2.51}$ / $^{+1.05}_{-3.15}$	$^{-0.08}_{+0.68}$ / $^{-0.97}_{+0.85}$ / $^{-0.94}_{+0.47}$	$^{+4.01}_{-2.02}$ / $^{+3.03}_{-3.52}$ / $^{+9.19}_{-1.70}$
Q^2 scale	$^{+3.45}_{+0.12}$ / $^{+1.07}_{-4.21}$ / $^{-0.34}_{-4.27}$	$^{+6.48}_{-6.96}$ / $^{+6.55}_{-5.10}$ / $^{+2.03}_{-5.98}$	-
ME/PS matching thresholds	-	$^{-2.84}_{+1.54}$ / $^{-1.37}_{+0.01}$ / $^{+1.35}_{+1.54}$	-
tW DR/DS scheme	$^{-7.56}_{+0.00}$ / $^{-10.19}_{+0.00}$ / $^{-11.76}_{+0.00}$	-	-
Top quark mass	$^{+6.79}_{-3.36}$ / $^{+2.57}_{-6.54}$ / $^{+0.79}_{-8.38}$	$^{+4.78}_{-4.22}$ / $^{+2.22}_{-18.67}$ / $^{+5.41}_{-2.83}$	-
$t\bar{t}$ cross-section	-	± 6.71	-
PDF	$^{+5.77}_{-5.54}$ / $^{+5.83}_{-5.63}$ / $^{+6.14}_{-5.85}$	$^{+5.26}_{-4.96}$ / $^{+5.01}_{-4.78}$ / $^{+5.16}_{-4.88}$	$^{+4.15}_{-4.14}$ / $^{+3.84}_{-3.61}$ / $^{+3.59}_{-3.66}$
Z+jet MET scale factor	-	-	$^{+17.85}_{-17.85}$ / $^{+6.78}_{-6.78}$ / $^{+20.22}_{-20.22}$

Table A.2: Rate impact of all considered systematic uncertainty sources in the 2j2t control region, values as a percentage. The estimates are for each of the three channels, unless specified as separate values for each channel ($ee/e\mu/\mu\mu$). If two numbers are listed for a single uncertainty, the upper number is the effect on the rate when the systematic uncertainty source is scaled up and the lower is for when it is scaled down. Entries with a single value indicate that the systematic is symmetric between the scaled up and scaled down effects.

Systematic Uncertainty ($ee/e\mu/\mu\mu$)	tW (%)	$t\bar{t}$ (%)	Other (%)
Luminosity	± 4.4	± 4.4	± 4.4
Lepton identification	± 2.33/± 1.86/ ± 2.28	± 2.33/± 1.86/± 2.28	± 2.33/± 1.86/± 2.28
JER	$^{-0.87}_{-1.35}$ / $^{-0.95}_{-1.10}$ / $^{-1.37}_{+1.71}$	$^{-0.43}_{+0.12}$ / $^{-0.52}_{+0.63}$ / $^{-0.02}_{-0.27}$	$^{+3.55}_{-0.93}$ / $^{+0.72}_{+0.28}$ / $^{+1.33}_{+0.44}$
JES	$^{-0.60}_{-2.39}$ / $^{-0.06}_{-3.31}$ / $^{+1.37}_{-2.73}$	$^{-1.86}_{-1.67}$ / $^{-1.73}_{-0.46}$ / $^{-1.15}_{-0.46}$	$^{+3.39}_{-19.81}$ / $^{+0.82}_{+0.50}$ / $^{+2.18}_{+0.72}$
MET modeling	$^{+4.35}_{-14.32}$ / $^{-0.39}_{+0.00}$ / $^{+6.54}_{-7.38}$	$^{+7.41}_{-6.48}$ / $^{-0.09}_{+0.07}$ / $^{+6.07}_{-5.52}$	-
Event pile up	± 0.38/± 0.03/± 0.02	± 0.46/± 0.34/± 0.16	± 1.54/± 1.61/± 2.81
B-tagging data/MC scale factor	$^{+3.10}_{-6.80}$ / $^{+3.88}_{-3.76}$ / $^{+6.97}_{-1.90}$	$^{+4.92}_{-5.11}$ / $^{+5.43}_{-5.19}$ / $^{+5.34}_{-4.76}$	$^{+11.80}_{-0.04}$ / $^{+1.12}_{-0.65}$ / $^{+0.11}_{-0.40}$
Q^2 scale	$^{-16.16}_{-18.55}$ / $^{-1.36}_{-8.00}$ / $^{-3.82}_{-12.60}$	$^{+7.22}_{-10.55}$ / $^{+7.44}_{-11.01}$ / $^{+6.58}_{-12.23}$	-
ME/PS matching thresholds	-	$^{+2.94}_{+2.34}$ / $^{-3.04}_{-0.58}$ / $^{-7.56}_{+1.26}$	-
tW DR/DS scheme	$^{-26.60}_{+0.00}$ / $^{-16.22}_{+0.00}$ / $^{-19.81}_{+0.00}$	-	-
Top quark mass	$^{-10.28}_{-19.77}$ / $^{-1.12}_{-11.03}$ / $^{-3.35}_{-11.98}$	$^{+8.13}_{-8.26}$ / $^{+3.42}_{-20.38}$ / $^{+1.09}_{-11.70}$	-
$t\bar{t}$ cross-section	-	± 6.71	-
PDF	$^{+6.03}_{-5.80}$ / $^{+5.27}_{-5.17}$ / $^{+5.61}_{-5.38}$	$^{+4.67}_{-4.52}$ / $^{+4.68}_{-4.52}$ / $^{+4.60}_{-4.49}$	$^{+4.30}_{-4.87}$ / $^{+3.39}_{-3.29}$ / $^{+3.55}_{-3.62}$
Z+jet MET scale factor	-	-	$^{+16.80}_{-16.80}$ / $^{+8.36}_{-8.36}$ / $^{+19.69}_{-19.69}$

Measurement of the single top tW associated production in the dilepton decay channel in proton collisions at a center of mass energy of 8 TeV

In this manuscript a measurement of the cross-section and significance of the single top tW associated production at a center-of-mass energy of 8 TeV and a luminosity of 12.1 fb^{-1} is presented. This process is yet to be discovered since it was not accessible at the Tevatron, and with a cross-section of 22.2 pb at a center of mass energy of 8 TeV, it is still a rare process at the Large Hadron Collider (LHC).

The difficulty of this process lies in the small signal cross-section compared to its background (almost 10 times smaller than its main background, $t\bar{t}$). After placing kinematic selections in order to get rid of other backgrounds (Drell-Yann, diboson electroweak production, other single top processes, W + jets, QCD multi-jet background) the sample is fully dominated by $t\bar{t}$ and tW signal. Therefore, a variable is chosen based on its discriminating power between tW and $t\bar{t}$, and its shape is used in the statistical fit. This approach of a shape based fit is more powerful than a cut-based fit. However, in order to achieve the largest significance as possible, all possible information against $t\bar{t}$ background should be used. Therefore, building a multivariate discriminant will give a higher significance.

Control regions are introduced in order to get a handle on the enormous amount of $t\bar{t}$ background compared to the signal. These regions are in the ideal case built up with the same selection criteria as the signal region, are completely depleted from signal events, and have as much $t\bar{t}$ events as possible. The mis modeling of the missing transverse energy for events without genuine missing transverse energy is taken into account by reweighting the Drell-Yann events, with scale factors calculated in the Z mass window. A systematic study is performed, considering all systematic uncertainties based on the 7 TeV analysis at CMS.

An excess of events over the expected background is observed with a significance of 4.0 σ, with an expected significance of $3.2^{+0.4}_{-0.9}\sigma$ of the tW signal. The measured cross-section, including statistical and systematic uncertainties, is $24.3^{+8.6}_{-8.8}$ pb, which is in agreement with the Standard Model.
Based on the assumption that the CKM matrix elements $|V_{ts}|$ and $|V_{td}|$ are much smaller than $|V_{tb}|$, the absolute value of the CKM matrix element $|V_{tb}|$ is 1.04 ± 0.20 (exp.) $\pm$ 0.04 (th.). Under the Standard Model assumption of $0 \leq |V_{tb}| \leq 1$, a value of $|V_{tb}|$ equal to 0.999 is found, with a 90% confidence interval of $[0.763, 1.000]$, and a 68% confidence interval of $[0.795, 1.000]$.

Meting van de werkzame doorsnede van de productie van een top in associatie met een W boson in het dileptonische vervalkanaal bij proton botsingen met een massamiddelpuntsenergie van 8 TeV

In dit manuscript is een meting gemaakt van de werkzame doorsnede en de significantie van de enkelvoudige top quark productie in associatie met een W boson met een massamiddelpuntsenergie van 8 TeV en luminositeit van 12.1 fb^{-1}. Dit proces is nog steeds niet ontdekt aangezien het niet bereikbaar was het Tevatron experiment en met een werkzame doorsnede van 22.2 pb bij een massamiddelpuntsenergie van 8 TeV, het nog steeds een zeldzaam proces is bij de Large Hadron Collider (LHC).

De moeilijkheid van deze studie ligt in het feit dat de werkzame doorsnede van het signaal zeer klein is ten opzichte van de achtergrondprocessen (ongeveer 10 maal kleiner dan zijn meest abundante achtergrond, $t\bar{t}$). Om deze reden worden kinematische sneden gemaakt met als doel de meeste achtergrond processen (Drell-Yann, diboson elektrozwakke productie, andere enkelvoudige top modi, W + jets, QCD mult-jet achtergrond) te verwijderen. Na deze selectie blijft er voornamelijk tW signaal en $t\bar{t}$ achtergrond over. Een variabele is gekozen op basis van zijn differentiërende kracht tussen het tW signaal en de top quark paar achtergrond en de vorm van de distributie wordt gebruikt in het statistisch model.
De keuze van een op vorm gebaseerde statistiek is krachtiger dan een statistiek gebaseerd op sneden. Het is namelijk altijd beter alle informatie van de achtergrond te gebruiken en dus zou het gebruiken van een multivariate discriminant een hogere gevoeligheid geven.

Om grip te krijgen op de enorme top quark paar productie, worden controle gebieden geïntroduceerd. In het ideale scenario zijn deze gebieden opgebouwd met dezelfde sneden als het signaal gebied, hebben ze geen tW-data en zijn ze volledig gedomineerd door de top quark paar productie. Verder, om afwijkingen in de modelisatie de ontbrekende transverse energie te corrigeren, wordt de simulatie van de Drell-Yann productie gewogen met factoren berekend in het massavenster van het Z boson. Een volledige systematische studie is uitgevoerd.

Ten opzichte van de verwachte achtergrond is een surplus geobserveerd met een significantie van 4.0σ, met een verwachte significantie van $3.2^{+0.4}_{-0.9}\sigma$ van het tW-signaal. De gemeten werkzame doorsnede, waarbij de statistische en systematische onzekerheden in rekening zijn gebracht, is $24.3^{+8.6}_{-8.8}$ pb en is in overeenkomst met het SM.
Gebaseerd op de veronderstelling dat de CKM matrix elementen $|V_{ts}|$ en $|V_{td}|$ veel kleiner zijn dan $|V_{tb}|$, is de waarde van het CKM matrix element $|V_{tb}|$ gelijk aan 1.04 ± 0.20 (exp.) ± 0.04 (th.). Onder de Standaard Model veronderstelling dat $0 \leq |V_{tb}| \leq 1$, is de waarde van $|V_{tb}|$ gelijk aan 0.999, met een 90% betrouwbaarheidsinterval van $[0.763, 1.000]$, en een 68% betrouwbaarheidsinterval van $[0.795, 1.000]$.

Meting van de werkzame doosnede van de productie van een top quark met een W boson in proton botsingen

Het Standaard Model (SM) beschrijft alle bekende elementaire deeltjes (fermionen) en hun interacties via de fundamentele krachten, met uitzondering van de zwaartekracht. Fermionen zijn opgedeeld in twee groepen van deeltjes, nl. quarks (de bestanddelen van bv. een proton) en leptonen (bv. een elektron). Deze interageren met elkaar door middel van krachtdragende deeltjes, genaamd bosonen (bv. een foton).
In dit manuscript wordt de werkzame doorsnede van de productie van een top quark met een W boson (tW) gemeten. De ontdekking van dit proces is nog steeds niet gebeurd omdat het een zeldzaam proces is. De uitdaging van deze studie is dat de tW gebeurtenissen minder voorkomen ten opzichte van de gebeurtenissen waarin dezelfde deeltjes geproduceerd worden via een andere interactie (achtergrond processen). De productie van twee top quarks (top quark paar) is bijvoorbeeld 10 maal groter dan het beschouwde tW proces (signaal). Om deze reden worden eisen gesteld op de gebeurtenissen met als doel de gebeurtenissen van achtergrond processen te elimineren. Hierna blijven er vnl. signaal en top quark paar gebeurtenissen over. Een variabele wordt gekozen zodat de distributies van de signaal en de top quark paar gebeurtenissen zo verschillend mogelijk zijn. De vorm van deze distributies wordt vervolgens gebruikt in een statistisch model om te schatten hoeveel signaal gebeurtenissen geproduceerd werden. Om de bijdrage van de vele top quark paar gebeurtenissen na de selectie criteria in te schatten, wordt er gebruik gemaakt van controle gebieden. Om deze gebieden te vinden, worden dezelfde eisen als het signaal gebied gebruikt, maar worden de gebieden geconstrueerd zodanig dat er zo goed als geen signaal gebeurtenissen in voorkomen. In deze controle gebieden wordt het aantal top quark paar gebeurtenissen geschat en het resultaat wordt gebruikt in het signaal gebied. Alle systematische onzekerheden worden in rekening gebracht.
De aanwezigheid van tW gebeurtenissen is geobserveerd met een significantie 4.0σ. Er is dus bewijs voor het bestaan van het tW proces, maar nog geen ontdekking (dit is pas bij 5σ). De significantie is een maat voor hoe zeker de meting is. Voor een 5σ betrouwbaarheid, is er 1 op 3.5 miljoen kans dat het resultaat voorkomt door toeval. De werkzame doorsnede voor de productie van het signaal is $24.3^{+8.6}_{-8.8}$ pb, wat overeenkomt met de voorspelling van het SM. Een picobarn (pb) is 10^{-36} cm^2 en wordt in de fysica gebruikt om waarschijnlijkheden uit te drukken. Dit begrip is makkelijk te begrijpen als volgt. Stel dat er genoeg eieren zijn gegooid naar een winkel zodat er ongeveer twee eieren per vierkante meter zijn en stel dat de deur 1.5 m bij 2 m is. Dan is de oppervlakte of "doorsnede" van de deur 3 m^2. Het aantal eieren stukgesmeten op de deur is dan 3 m^2 $\times$ 2 eieren/m^2 of 6 "ei interacties". In de fysica wordt "2 eieren/m^2" uitgedrukt in inverse femtobarn (1 fb = $1/10^{-39}$ cm^2). Voor deze studie is er een luminositeit van 12.1 fb^{-1} en zijn er dus ongeveer 294 030 interacties (24.3 (pb) $\times$ 12.1 (fb^{-1})).
In de veronderstelling dat de top quark enkel kan vervallen naar een b quark, kan er een meting gemaakt worden van het CKM matrix element $|V_{tb}|$, een fundamentele constante van het SM. Dit heeft als resultaat 1.04 ± 0.20 (exp.) ± 0.04 (th.). Volgens het SM kan deze constante niet groter zijn dan één. Het toepassen van deze eis geeft $|V_{tb}|$ gelijk aan 0.999, waarbij er 90% zekerheid is dat als het experiment wordt herhaald, de waarde van $|V_{tb}|$ tussen 0.763 en 1.000 zal liggen.

Acknowledgements

This manuscript would not have been created if I did not have the help of several people. First of all, I would like to thank the IIHE for giving me the opportunity to do my master thesis at their collaboration. The support and patience of my promoter Freya Blekman and co-promoter Rebeca Gonzalez Suarez were invaluable during this study and I value their professional opinions and constructive comments as they made me a better researcher. Further, I would like to express my gratitude to the members of the jury for the evaluation of this thesis.

I want to express my gratitude to Jorgen D'Hondt for helping me throughout my academic career, and giving me feedback along the way. From the side of computing, I want to acknowledge the computing team, with a special thank you to Olivier Devroede and Abdelhakim Boukil for patiently helping me. I am also thankful for the doctoral students of the CMS department at IIHE, team Devil Fish, as they helped me finding a way through the jungle of research techniques at the CMS collaboration. I am specially appreciative for Michael Maes, Petra Van Mulders, James Keaveny and Stijn Blyweert for helping me with the practical aspects of this research. Further, I want to express my gratitude to Daniel Noonan and Jochen Ott, my go-to persons from the Single top group at the CMS collaboration.

Isabelle, Lieselotte, Kevin, Mathias, Bert and Ivo, as we shared astonishments as we made our path into particle physics, it was a pleasure to share these experiences with you. Lieselotte and Kevin, my classmates, roommates and 'colleagues' at IIHE, I think we form a great team and we definitely shared an unforgettable adventure. Further I am grateful for the people who have always been there for me. Sabrina, Sander and Ben, thank you for offering me a shoulder to lean on, and I promise I will make up for the lack of time during the making of this manuscript.
Frank thank you for always being there and put up with my complaints when things were not going my way. I am also indebted to my mother for standing by my side throughout my studies.

I am eager to know what the future might bring and the many opportunities and experiences that lie ahead of me. I will never forget this wonderful period that has come to its end.

List of Abbreviations

Units

b	Barn, a unit of area. It is defined as 10^{-28} m^2 and is approximately the cross-sectional area of a uranium nucleus.
pb	Picobarn, 10^{-12} barn.
fb	Femtobarn, 10^{-15} barn
B	Byte, 8 bit. A bit is a binary digit, the basic unit of information in computing and telecommunications.
MB	Megabyte, 10^6 bit
eV	Electron volt, a unit of energy equal to $\pm 1.6 \cdot 10^{-19}$ Joule. The amount of energy gained by the charge of a single electron moved across a potential of one volt.
MeV	Mega-electron volt, 10^6 eV.
GeV	Giga-electron volt, 10^9 eV.
TeV	Tera-electron volt, 10^{12} eV.
Hz	Hertz, the unit of frequency. It is defined as one cycle per second.
MHz	Megahertz, 10^6 Hz.
kHz	Kilohertz, 10^3 Hz.
m	Meter, a measurement of distance.
km	Kilometer, 10^3 meter.
cm	Centimeter, 10^{-2} meter.
μm	Micrometer, 10^{-6} meter
fm	Femtometer, 10^{-15} meter
s	Second, a measurement of time
μs	Microsecond, 10^{-6} s.
ns	Nanosecond, 10^{-9} s.
T	Tesla, the SI derived unit of magnetic flux density. One tesla is equal to Volt $\cdot$ second $\cdot$ m^{-2}.

Program names/Algorithms

C++	Object-oriented programming language developed by Bjarne Stroustrup in 1979 at Bell Labs.
CMSSW	CMS software framework, a framework to ensure the event generation, the full simulation of the CMS detector, and the reconstruction of the particles present in the event.
CPU	Central processing unit, the hardware within a computer that carries out the instructions of a computer program by performing the basic arithmetical, logical, and input/output operations of the system.
CTEQ	Coordinated Theoretical Experimental Project on QCD. CTEQ is a multi-institutional collaboration devoted to a broad program of research projects and cooperative enterprises in high-energy physics centered on Quantum Chromodynamics (QCD) and its implications in all areas of the Standard Model and beyond.
EDM	Event data model
GEANT	GEometry And Tracking, a detector simulation framework and toolkit. A toolkit for the simulation of the passage of particles through matter, for high energy, nuclear and accelerator physics as well as medical and space science.
LHAPDF	Les Houches Accord Parton Distribution Function library, an interface to modern PDF sets.
MadGraph	Multi-particle matrix-element generator
MC	Monte-Carlo simulation program or technique
MSTW/MRST	Martin-Stirling-Thorne-Watt Parton Distribution Functions
POWHEG	POsitive Weight Hardest Emission Generator, a general computer framework for implementing NLO calculations in shower MC programs according to the POWHEG method. This is also a library that includes ready made processes. It can be interfaced with all modern shower Monte Carlo programs that support the Les Houches Interface for User Generated Processes.
PYTHIA	Monte-Carlo event generator for high-energy physics collisions, thus for the description of collisions at high energies between elementary particles. It contains theory and models for a number of physics aspects, including hard and soft interactions, parton distributions, initial- and final-state parton showers, multiple interactions, fragmentation and decay.
anti-k$_T$	An algorithm for combining jets
CSV	Combined Secondary Vertex algorithm, an algorithm for b-tagging.
GSF	Gaussian Sum Filter, an algorithm that takes into account energy loss due to Bremsstrahlung
FTC	Flavor Tag Consistency method, method to measure b-tag efficiencies

Experiments/Organizations

ALICE	A Large Ion Collider Experiment at the LHC
ATLAS	A Toroidal LHC ApparatuS
CDF	Collider Detector Facility at the Fermi National Accelerator Laboratory Tevatron
CERN	Conseil Européen pour la Recherche Nucléaire, European laboratory for particle physics
CMS	Compact Muon Solenoid
DØ	Experiment at the Fermi National Accelerator Laboratory Tevatron
IIHE	Inter-university Institute for High Energies, Brussels
LEP	Large Electron-Positron collider at CERN
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty experiment
LHCf	Large Hadron Collider forward experiment
PS	Proton Synchrotron
SPS	Super Proton Synchrotron at CERN
S-Kamiokande	Super-Kamioka Neutrino Detection Experiment.
TEVATRON	TeV-range proton–anti-proton accelerator at Fermi National Accelerator Laboratory.
TOTEM	TOTal Elastic and diffractive cross-section Measurement, LHC experiment for the measurement of the total cross section, elastic scattering and diffraction dissociation

Other

ΔR	A distance, the square of the distance is defined as $\Delta\eta^2 + \Delta\Phi^2$
ϵ	Symbol for efficiency
η	Pseudo-rapidity, minus the natural log of the tangent of the center-of-mass scattering angle.
θ	Center-of-mass scattering angle, the polar angle
σ	Symbol for cross-section, the probability that two particles will collide and react in a certain way.
Φ	Azimuthal scattering angle
BDT	Boosted Decision Tree, an analysis technique.
BR	Branching Ratio
CKM	Cabibbo, Kobayashi and Maskawa, used to refer to the quark-mixing matrix
CP	Charge conjugation and Parity
DR	Diagram removal.
DS	Diagrams subtraction.
ECAL	Electromagnetic CALorimeter

E_T^{miss}	The missing transverse energy.
FCNC	Flavor-Changing Neutral Current
HCAL	Hadron Calorimeter
HEP	High-Energy Physics
HLT	High-Level Trigger
H_T	The scalar sum of the four momenta of the objects in the analysis.
ID	Identification
i.p.	Impact parameter, the distance between the primary vertex and the linearized track.
JER	Jet Energy Resolution
JES/JEC	Jet Energy Scale
Jet ID	Jet identification
L1	Level-1 Trigger
LO	Leading Order
m	Symbol for mass
m_{ll}	The invariant mass of two leptons
ME	Matrix Element method
MVA	Multi variate analysis, a statistical analysis technique.
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
PDF	Parton Density Function
PS	Parton Shower or Proton Synchrotron
PL	Profile likelihood method.
p_T	Transverse momentum, the component of the momentum in the transverse plane.
PU	Pile up, the additional interactions in a collision.
PV	Primary Vertex
QCD	Quantum ChromoDynamics
reliso	Relative isolation, the ratio of the scalar sum tot the lepton candidate's momentum. This is used as an isolation requirement.
SF	Scale Factor. A number to reweight efficiencies.
SM	Standard Model
TriDAS	Trigger and Data acquisition system
Q	Electric charge
$\mathbf{Q^2}$	Factorization or normalization scale

I like to say, when asked why I pursue science, that it is to satisfy my curiosity, that I am by nature a searcher trying to understand. If you haven't found something strange during the day, it hasn't been much of a day.

John A Wheeler