



DISSERTATION

Search for supersymmetry in the single lepton final state in 13 TeV pp collisions with the CMS experiment

ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Doktorin der
technischen Wissenschaften unter der Leitung von

UNIV.PROF. DIPL.-ING. DR.TECHN. JOCHEN SCHIECK

und

DR. ROBERT SCHÖFBECK

am Institut für Hochenergiephysik (HEPHY)
der Österreichischen Akademie der Wissenschaften (ÖAW)
und am Atominsttitut (E141)

eingereicht an der
TECHNISCHEN UNIVERSITÄT WIEN
FAKULTÄT FÜR PHYSIK
von

M.S. ECE AŞILAR
Matrikelnummer: 1428489

Wien, am 20. August 2017

This thesis is dedicated to
the city of Vienna

Acknowledgements

Robert Schöfbeck

Wolfgang Adam

Jochen

Office colleagues: Daniel Spitzbart, Navid k. Rad, Federico Ambrogy, Matheusz Zarucki, David Handle..

Other HEPHY people: Erica, Ilse, Johannes, WW, Suchita LHC, and CMS

SUSY group convenors

To all POG

ACDVF People: Clauida ...

Dietrich

Guido van Rossum for python :)

Bram Moolenaar for vim :)

Social Life Friends, mom brother, Gökhan ...

Final but not least, FWF and DKPI

Abstract

In this thesis, an inclusive search for supersymmetry is presented. The search is performed in events containing a single lepton, multiple jets requiring none of them coming from b quarks, and missing transverse energy in the final state. The proton-proton collision data were recorded by the CMS experiment during Run 2 of the LHC at a center-of-mass energy of 13 TeV. The analyzed data corresponds to a total integrated luminosity of 35.9 fb^{-1} . The search uses delta phi, the azimuthal angle between the lepton and four-vector sum of the missing energy and lepton, as a powerful discriminating variable to distinguish between background and signal. Additionally, multiple exclusive search regions are defined in different kinematic observables to enhance sensitivity to a range of different mass scenarios. The latest results in this clean event topology interpreted in the context of simplified models.

Contents

1 Supersymmetry: as an extension of Standard Model	1
1.1 Standard Model	1
1.1.1 Standard Model: Particle Content	2
1.1.2 Standard Model: Particle Interactions	2
1.1.3 Standard Model: from the quantum field theory window	4
1.1.4 Standard Model: Experimental results	7
1.1.5 Shortcomings of the Standard Model	8
1.2 Supersymmetry as a solution	10
1.2.1 Algebra of Supersymmetry	12
1.2.2 Minimal Supersymmetric Standard Model	13
1.2.3 Simplified Models	16
1.2.4 Short History of SUSY searches at colliders	17
2 Experimental Setup	21
2.1 The Large Hadron Collider at CERN	21
2.1.1 The CERN accelerator complex	22
2.1.2 The Large Hadron Collider	22
2.2 The Compact Muon Solenoid experiment at the LHC	25
2.2.1 Superconducting Magnet	26
2.2.2 Tracker	26
2.2.3 Electromagnetic Calorimeter	27
2.2.4 Hadronic Calorimeter	27
2.2.5 Muon System	28
2.2.6 Trigger and Data Acquisition Systems	28
2.2.7 Luminosity measurement	29
2.3 Event simulation	29
2.3.1 Event generation	29
2.3.2 Detector simulation	30

3 Object reconstruction and identification	31
3.1 Particle-Flow algorithm	31
3.2 Physics Object reconstruction and identification	32
3.2.1 Primary vertices	32
3.2.2 Leptons	33
3.2.2.1 Muons	33
3.2.2.2 Electrons	34
3.2.3 Jets	35
3.2.4 b tagged Jets	36
3.2.5 Missing transverse energy	37
4 Event Selection:	
Baseline and Search regions	40
4.1 SUSY signature	40
4.1.1 Key variables	42
4.1.2 Signal samples	43
4.2 Background processes	43
4.2.1 Scale Factors	44
4.3 Baseline selection	46
4.4 Data Samples	48
4.4.1 Triggers	48
4.5 Event cleaning: Filters	49
4.6 Control plots	53
4.7 Search regions	53
4.7.1 Signal and control region	53
4.7.2 Main band regions	54
5 Background Estimation: R_{CS} Method	61
5.1 QCD background estimation	62
5.2 Background fraction calculations: b-tag multiplicity fit	63
5.3 R_{CS} method in $t\bar{t} + \text{jets}$ events	65
5.4 R_{CS} method in $W + \text{jets}$ events	67
5.5 Validation of the background estimation	70

6 Systematic Uncertainties	72
6.1 Systematic uncertainties on background estimation	72
6.2 Systematic uncertainties on signal modelling	75
6.3 Common systematic uncertainties for signal and background modelling . . .	76
7 Results and Interpretation	80
7.1 Results of background prediction	80
7.1.1 Result of validation in side band regions	80
7.1.2 Background prediction in main band regions	81
7.2 Statistical interpretation	82
7.2.1 Frequentist limit setting procedure	84
7.3 Interpretation on simplified model T5qqqqWW	89
7.4 Comparison to other results	89
8 Conclusion	94
A MC Samples and cross sections	97
B Control Plots	99
C Statistical Tests	106
Bibliography	126

Chapter 1

Supersymmetry: as an extension of Standard Model

Contents

1.1	Standard Model	1
1.1.1	Standard Model: Particle Content	2
1.1.2	Standard Model: Particle Interactions	2
1.1.3	Standard Model: from the quantum field theory window	4
1.1.4	Standard Model: Experimental results	7
1.1.5	Shortcomings of the Standard Model	8
1.2	Supersymmetry as a solution	10
1.2.1	Algebra of Supersymmetry	12
1.2.2	Minimal Supersymmetric Standard Model	13
1.2.3	Simplified Models	16
1.2.4	Short History of SUSY searches at colliders	17

This chapter commences with a discussion of the Standard Model (SM) of particle physics. After presenting a brief theoretical overview, its success in explaining the majority of the experimental results will be discussed. We will then ponder on the missing aspects of the SM as a complete theory. The second part of this chapter is dedicated to the supersymmetric (SUSY) models, as one of the possible theories beyond the SM (BSM). The SUSY discussion will include the framework of simplified model spectra (SMS): The results presented in this thesis are interpreted with in SMS.

1.1 Standard Model

The details of this chapter can be found in various textbooks on the standard model and quantum field theory [1–3].

1.1.1 Standard Model: Particle Content

The SM describes fundamental particles and their interactions. In this respect, the SM attempts to explain all physical phenomena with the exception of gravity, which has an insignificant effect on subatomic particles. There are three known fundamental interactions or forces in the SM: the electromagnetic, the weak and the strong force. The interactions are mediated by force carrier particles called bosons. Bosons are quanta of gauge fields describing aforementioned interactions. The electromagnetic interactions are carried by the photon while Z^0 and W^\pm bosons are responsible for the weak interactions. In the SM framework, these two seemingly different forces can be unified through the so-called electroweak theory. For the strong interactions, the force carrier bosons are called gluons and they are massless. Bosons are integer spin particles obeying the Bose-Einstein statistics. The elementary particles that constitute the all known forms of matter are called fermions. Obeying the Fermi-Dirac statistics, fermions are particles with half spin. Fermions incorporate quarks (with color charge) and leptons (without color charge). According to the SM, fermions can be categorized as three families or generations, which are very similar to each other in terms of characteristics of the particles. The first family represents the substance we see around us, the rest can be observed in the colliders or nuclear reactors or in the atmospheric showers. Figure 1.1 shows the particle content of the standard model¹. All are experimentally confirmed. The range of these forces is inversely proportional to their masses. The photon is massless so the range of electromagnetism is infinite while the range of weak force is constrained by the large mass of the corresponding gauge bosons. This mechanism is more complicated in the case of the strong force; it will be discussed in the next paragraphs.

1.1.2 Standard Model: Particle Interactions

The theory that aims to model strong force is called **quantum chromodynamics** (QCD). In Greek, the word $\chi\rho\omega\mu\alpha$ chroma means color.

After observation of bound states such as $\Delta_{(uuu)}^{++}$, violating Pauli's exclusion principle, it is suggested that the quarks possess three color (red, blue, green) charges. The interaction between quarks occurs through gluons. The fact that gluons have color charge make them interact with themselves as well. Accordingly, QCD also admits bound states whose valence constituents are all gluons, the nonabelian gauge bosons of QCD. These additional mesons are one of the most important predictions of the SM and they are known as gluonia or glueballs. However, so far they have not been observed experimentally. For further

¹ In this thesis, the matter and anti-matter particles are not distinguished.

reading, [17, 18] can be consulted.

QCD has two important postulates.

Confinement: In nature we observe only color singlet (colorless) particles. The particles with color charges, i.e. quarks and gluons, immediately coalesce to form colorless bound states (hadrons) given an attempt to separate them. This is also known as hadronization. As a result of this phenomenon, despite the fact that gluons are massless, the range of the strong force is confined. Moreover, in experiments a shower of color-neutral particles, which is called as jet and will be explained in Sec.3.2.3, is observed instead of a single quark.

Asymptotic freedom: The interaction of quarks becomes asymptotically weaker as the energy increases or as the distance decreases. In other words, the strong coupling constant α_s increases with distance, but inside a meson or a baryon, they act as free particles. This asymptotic freedom helps to build a perturbative description of QCD interactions; otherwise QCD calculations are extremely complicated.

The interaction of charged particles with energies of the order of the W^\pm boson mass can be described by **quantum electrodynamics** (QED). The symmetry group of the electromagnetic theory is a unitary transformation $U(1)_{EM}$; there is only a massless field as-

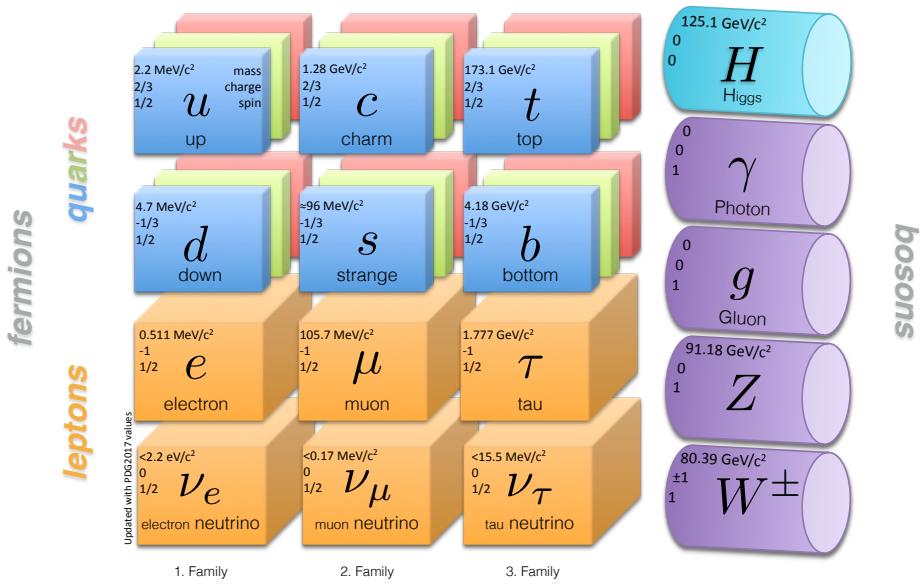


Figure 1.1: All fundamental particles of the standard model, the gauge and Higgs bosons are shown in the diagram. The electric charge, spin and mass (or limit for neutrinos) are given in the corners of the boxes [5].

sociated to photon. The β -decay, where the neutron decays to a proton, electron and corresponding antineutrino through the W boson, explained by Enrico Fermi by an effective theory [7]. The theory was based on $SU(2)_L$ symmetry group. The subscript L stands for the left handed particles. Due to lack of right handed neutrinos in nature, only left handed fermions contribute to the **weak interaction**. In 1960s, Glashow-Salam-Weinberg proposed a unified theory of the QED and weak interactions [8–10]. After unification of electric and magnetic interactions by Maxwell, the unification of electromagnetic and weak interactions is an encouraging step towards the grand unified theories (GUT). In the formulation of the SM, initially all the mass of particles are 0. Brout, Englert [11] and Higgs [12] solved the problem by implementing the electroweak symmetry breaking (EWSB), called BEH or only the Higgs mechanism. This mechanism introduces an additional scalar field, the Higgs field. The SM Higgs boson is a Goldstone boson with mass given by $m_H = \sqrt{2\lambda}\nu^2$. The massive bosons of weak interaction also gain their mass through a non-zero vacuum expectation value (VEV) while for fermions direct Yukawa couplings to the Higgs boson provides their mass. Particles interact with the Higgs field at a different strength. Depending on the coupling between the Higgs field and the particle, its mass can be large or small. Photons, having zero rest mass, do not couple to Higgs field. In the absence of the Higgs field, all the particles would be massless.

1.1.3 Standard Model: from the quantum field theory window

In particle mechanics, a Lagrangian³ is a function of the coordinates, and their time derivatives. In field theory technically a Lagrangian density is used and it is a function of the fields, ϕ_i and their position and time derivatives, $\partial_\mu\phi_i$. In relativistic theory, space and time coordinates are treated in the same way. The Lagrangian plays an important role in physics because the form of the Lagrangian is invariant under some symmetries. Given a Lagrangian the equation of motion from the Euler-Lagrange equations can be derived by considering the least action principle, i.e requiring that the variation of action is zero⁴. Then, the Euler-Lagrange equation in a simple way can be written as follows:

$$\partial_\mu \frac{\partial L}{\partial(\partial_\mu\phi_i)} = \frac{\partial L}{\partial\phi_i}. \quad (1.1)$$

The SM is a relativistic quantum field theory (QFT). Its Lagrangian is built on a global Poincaré symmetry, which implies symmetry under translation, rotation and boost. According to Noether's theorem [30], each continuous symmetry is accompanied by a conservation law. The Poincaré symmetry points to conservation of energy, momentum, and

² $\nu = (\sqrt{2}G_F)^{-\frac{1}{2}} \approx 246\text{GeV}$, G_F is Fermi Coupling

³ $L = T - V$, T is the kinetic energy of the particle in a potential V .

⁴ $\delta S = 0$, where $S = \int L dt$.

angular momentum. The gauge group of the SM, which is a local symmetry, is described as:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad (1.2)$$

where $SU(3)_C$, $SU(2)_L$ and $U(1)_Y$ are representing the gauge groups of the strong, weak and electromagnetic forces respectively. The subscript C refers color, L refers to left-handedness, and Y refers hypercharge. The conserved quantities, which correspond to the symmetry in the Equation 1.2, are color charge, charge, and weak hypercharge.

The rank of the group, i.e. the number of generators of the fields is at the same time the number of mediators, gauge bosons, of the corresponding vector field. For instance, the $SU(3)_C$ group has eight⁵ generators thus it has eight vector fields which are called the gluon fields (G_μ). Following the same argument, the $SU(2)_L$ group has three vector fields, (W_μ^1 , W_μ^2 , and W_μ^3) and the $U(1)_Y$ group has only one vector field, (B_μ).

In addition, the EWSB, or in other words the Higgs mechanism, leads to an additional scalar field, which will be denoted as ϕ in upcoming equations. The SM Lagrangian resides two components: L_{QCD} and L_{EWK} . The first one is for strong interaction while the latter explains electroweak interaction including interaction with the Higgs boson.

- **Strong interaction**

As mentioned in the previous section, the theory that aims to model strong force is QCD. The QCD Lagrangian is given by:

$$L_{QCD} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi + g_s \bar{\Psi} T_a G_\mu^a \gamma^\mu \Psi - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (1.3)$$

where Ψ are the quark fields, γ^μ are the Dirac matrices, g_s is the strong coupling constant, T_a are Gell-Mann matrices and $G_{\mu\nu}^a \equiv \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f_{abc} G_\mu^b G_\nu^c$ ⁶. When constructing the Lagrangian, a covariant derivative is introduced. This way, the terms stay invariant under gauge transformations. In general, the form of the covariant derivative is read as $D_\mu = \partial_\mu - igX_\mu$ where only ∂_μ is for the fermion and with the g coefficient it interacts with the X boson. The $\bar{\Psi}\gamma^\mu D_\mu \Psi$ terms are kinetic terms of the Lagrangian. The covariant derivative for QCD is defined as: $D_\mu = \partial_\mu - ig_s T_a G_\mu^a$.

- **Electroweak interaction**

The electroweak Lagrangian, L_{EWK} , can be written as the sum of 4 pieces:

$$L_{EWK} = L_{Gauge} + L_{Fermion} + L_{Higgs} + L_{Yukawa}. \quad (1.4)$$

⁵ $3^2 - 1$

⁶ f_{abc} are the structure constants

The covariant derivative that leaves this Lagrangian invariant under transformations is described as: $D_\mu = \partial_\mu - igW_\mu^a\tau_a - ig'B_\mu Y_W$ where g and g' are the gauge couplings of the $SU(2)_L$ and $U(1)_Y$ respectively. The τ_a are Pauli matrices. Y_W is the weak hypercharge defined as $2(Q - I_3)$. Q is the electric charge and I_3 is the third component of the weak isospin.

The first term in Equation 1.4 is defining the interaction amongst the gauge bosons, and it can be written as:

$$L_{Gauge} = -\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} \quad (1.5)$$

where $W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon^{abc}W_\mu^b W_\nu^c$, $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$.

The second term in Equation 1.4 stands for the fermion kinetic term and fermion interactions with $SU(2)$ and $U(1)$ bosons:

$$L_{Fermion} = i\bar{\Psi}\gamma^\mu D_\mu \Psi \quad (1.6)$$

The third term in the same equation is the Higgs Lagrangian, describing the Higgs field, its self-interaction and its interaction with the gauge bosons:

$$L_{Higgs} = |D_\mu \phi|^2 - \lambda(|\phi|^2 - \frac{v^2}{2})^2, \quad (1.7)$$

where λ is the Higgs self-coupling strength. According to BEH mechanism, there is a scalar potential, which permeates the whole universe:

$$V(\Phi) = m^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \quad (1.8)$$

, with the Higgs field Φ with weak hypercharge $Y = 1$, and a self-interacting $SU(2)$ complex doublet in Equation 1.9⁷.

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\phi^+ \\ \phi^0 + ia^0 \end{pmatrix} \quad (1.9)$$

If the quadratic term is negative, the neutral element of the scalar doublet gains a non-zero vacuum expectation value (VEV) stimulating the spontaneous breaking of the SM gauge symmetry [5].

The final term in equation 1.4 is the Lagrangian of the Yukawa interaction between the Higgs field and the fermion fields (quarks and leptons). The L_{Yukawa} produces fermion masses through spontaneous symmetry breaking. It can be written in the most general way [31] pg. 559:

$$L_{Yukawa} = -\epsilon^{ij}\phi_i\ell_{jI}y_{Ij}\bar{e}_j - \epsilon^{ij}\phi_iq_{ajI}y'_{Ij}\bar{d}_j^\alpha - \phi^{\dagger i}q_{\alpha iI}y''_{Ij}\bar{u}_j^\alpha + h.c. \quad (1.10)$$

⁷ ϕ^0 :CP-even, a^0 :CP-odd neutral component. ϕ^+ :complex charged component

where y_{Ij} , y'_{Ij} , y''_{Ij} are complex 3×3 matrices, and the generation indices ($I = 1, 2, 3$) are summed. After SSB the neutral and charged current interactions between fermions and gauge bosons can be derived from the L_{EWK} . In order to extract the actual mass terms, a switch between basis with W^a , B fields and a basis with mass eigenstates is required:

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^0 \\ W^3 \end{pmatrix} \quad (1.11)$$

$$W^\pm = W^1 + iW^2. \quad (1.12)$$

The θ_W term is the Weak mixing angle or Weinberg angle. This quantity is measured experimentally as well.

In this section, the formulation of the Lagrangian which explains the SM particle content discussed in section 1.1.1 was briefly introduced. A more comprehensive and pedagogical formulation of the SM can be found in [5, 31]. The next section will continue with experimental status of the SM.

1.1.4 Standard Model: Experimental results

The entire SM particle content was discovered up to now. The fermionic substance was completed with the discovery of the top quark by the CDF and DØ experiments in 1995 [13, 14]. In 2012, a particle, with a mass of approximately 125 GeV, which has similar features to the Higgs boson predicted by SM was discovered by ATLAS and CMS experiments [15, 16]. Further studies with a much larger data set have provided precise measurements on its mass, production and decay rates. Results present consistency with the SM prediction, within errors. With this new discovery, the bosonic content of the SM is also completed.

Testing the SM is not limited to collider experiments. In fact due to the technological constraints, probing the energies much larger than TeV scale is not possible with collider experiments; this much high energy can only be studied by cosmological experiments. In addition to cosmological and collider experiments, there are smaller scale desktop experiments where particle properties can be further investigated with high precision in lower energies (\leq GeV). In the Figure 1.2, an overview of the different types of experiments probing the SM can be seen. In addition to this, an example list of different kinds of measurements is also added to the table. The colors are making an analogy to the degree of the consistency of the measurements with the SM where green indicates that the experimental results are consistent with SM, yellow is showing that the measurements point out a deviation from SM predictions, and red evince that there must be a theory beyond the SM. Taking into account that not all the colors are green, the SM has several shortcomings not

only from the experimental but also from theoretical perspective. These shortcomings are the topic of the next section.

Experiment type	Colliders	Smaller scale	Cosmology
Energy	High/Low energy (≤ 13 TeV)	Low energy (\leq GeV)	Very high energy ($>>$ TeV)
High precision	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Measurement	<ul style="list-style-type: none"> Particle content Electro weak unification 	<ul style="list-style-type: none"> Anomalous magnetic moments Electric dipole moments Mass measurements 	<ul style="list-style-type: none"> Dark matter Dark energy
Consistency with the SM	Up to now, observations are consistent with the SM	g-2 was found to be 3 sigma away from the theoretical value	The SM does not have a dark matter candidate

Figure 1.2: An overview of characteristics of the experimental tests of the SM is shown in the table. The information is mainly collected from “the Electroweak model and constraints on new physics” section in PDG2017 [5]. Additionally further information on g-2 measurements can be found from the paper published by Muon G-2 experiment [19].

1.1.5 Shortcomings of the Standard Model

As mentioned in the previous section, no evidence of the new physics beyond the SM has been found at LHC. Moreover, up to now, the SM has successfully explained the world of subatomic physics. So, why are we searching for what is beyond the SM?

- **Experimental reasons**

Lacking explanation of Gravity: Certainly, the observation of gravitational waves is one of the two most exciting discoveries of our century. The announcement was done by LIGO and Virgo collaborations on 11 Feb 2016 [20]. The observed waveform satisfies the predictions of general relativity [21]. However, unless the gravitational field is quantized, it is impossible for the SM to include gravity.

No Dark Matter candidate: In 1933, Zwicky observed an unseen mass by applying the Virial theorem⁸ to the Coma cluster. He introduced this unseen mass as Dark matter (DM),

⁸relates the gravitational potential energy of a system to its kinetic energy

originally in German called *dunkle Materie* [22]. The evidence of the DM got stronger with the observations made by V. Rubin and K. Ford on the velocity curve of more than twenty spiral galaxies [23]. Moreover, the observations of anisotropies in the cosmic microwave background (CMB) further supported the existence of DM. The data provided by Planck experiment, successor of COBE [25] and WMAP [24] experiments, is in good agreement with the Λ CDM (Lambda cold dark matter) model [26]. This model postulates a dark energy dominated (68%) flat universe, with 5% baryonic matter and 27% dark matter. Therefore, given the SM does not provide a viable cold dark matter candidate⁹, in fact the SM fails to explain 95% of our universe.

Massless neutrinos: The observation of neutrino oscillations [27, 28] implies that at least two of the neutrinos should have non zero mass¹⁰. Neutrino masses can be included as an input to fermion Yukawa couplings, otherwise massless. However, even this input does not explain the observed mass differences between the generations.

- **Theoretical reasons**

Addition to the solid experimental reasons, the SM has also theoretical insufficiencies, which are mostly aesthetic concerns of theorists.

Grand Unification: The unification of forces has started with the integration of electric and magnetic forces into one electromagnetic force. Then it was followed by the unification of electromagnetic and weak forces into the electroweak interaction. At this point, it is inevitable that one expects the merge of electroweak and strong interactions. In fact, the running couplings of electroweak and strong forces, as seen in the Figure 1.3, meet at around 10^{16} GeV. Finally, even the unification of gravity with other fundamental forces is strongly encouraged. **Hierarchy problem:** In the literature of particle physics, there are two kinds of hierarchy problems; the big and the little hierarchy problem. The first one is due to the huge difference between mass scale of weak forces (m_W) and gravity (Planck scale¹¹). In fact, this is not a consistency problem but a naturalness problem. The latter, the little hierarchy problem, is a problem of how large m_H^2 , in Eq.1.8, can get. It receives huge quantum corrections because Higgs boson, including itself, couples to all massive particles in the SM. This subject will be discussed carefully in the upcoming sections.

There are many Beyond the Standard Model (BSM) theories claiming to solve some of the aforementioned SM shortcomings. The candidate BSM theory must not belie the

⁹In this context, Neutrinos are hot (relativistic) particles

¹⁰The oscillation frequencies are proportional to the mass difference of the neutrino flavors. Therefore, only upper bounds can be measured.

¹¹The Planck scale is the scale at which classical gravity theory is no longer valid but quantum gravity dominates. $M_P = 2.4 \times 10^{18}$ GeV

current observations; on the contrary it should predict them. Moreover, in order to be able to test the theory, it should also provide a phenomenological background. In a spectrum of BSM theories which satisfies these constraints, the Supersymmetric extensions of the SM are the most promising ones. This thesis focuses on a subgroup of these models, thus a relatively specified overview will be given in the following section.

1.2 Supersymmetry as a solution

This chapter presents a brief summary of [4] in the context of this thesis search. When searching what is beyond the SM, a familiar concept, symmetry, helped again. A symmetry, which relates fermionic states to bosonic states, had already attracted physicists' attention in the first half of the twentieth century. It was in the 1970s; a theory of four-dimensional relativistic space-time (super)symmetry has emerged. This theory is known

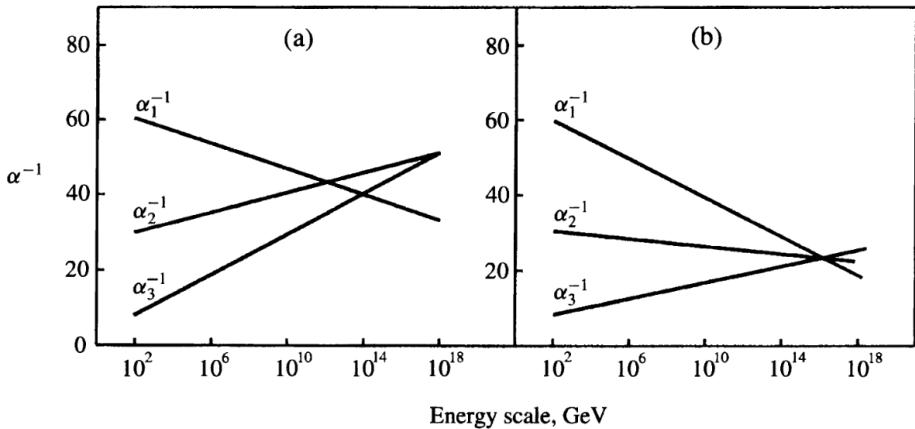


Figure 1.3: The running couplings of U(1), SU(2), SU(3) is denoted by α_1 , α_2 , α_3 respectively. The figure represents the evolution of inverse of couplings with energy scale, for (a) non-supersymmetric SU(5) and (b) supersymmetric SU(5). The figure is taken from [29] pg. 282.



Figure 1.4: One-loop quantum corrections to the Higgs squared mass parameter m_H^2 , due to (left) a Dirac fermion f , and (right) a scalar S [4] p.g. 3

as Supersymmetry (SUSY). The supersymmetric transformation changes the spin angular momentum of the original SM fields by a factor of $\frac{1}{2}$ and this turns fermionic states into bosonic states, and vice versa. These additional partners are called superpartners. If these superpartners exist in TeV scale it can overcome several problems mentioned in section 1.1.5.

The existence of SUSY may lead to spectacular results that can be listed as follows:

Solving the hierarchy problem:

As mentioned in the previous section, Higgs boson couples to each massive SM fermion f through a Yukawa coupling λ_f . Therefore, the Higgs boson bare mass, the m^2 term in Equation 1.8, receives enormous quantum corrections. The one-loop radiative correction terms (see Figure 1.4 left) which are coming from Dirac fermions f are as follows:

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \dots \quad (1.13)$$

where Λ_{UV} is an ultraviolet momentum cutoff, which is used to restrict the loop integral. The cutoff can be interpreted as the energy scale at which the new physics expected. Considering the top quark is the one with the largest mass among the SM fermions, the largest correction comes from it with $\lambda_f \approx 0.94$. If Λ_{UV} is of the order the Planck mass, then this correction to m_H^2 is around 30 orders of magnitude larger than the Higgs bare mass. In addition to the correction term given in Equation 1.13, a similar correction can be assumed from the heavy scalar particle S (see Figure 1.4 right) as:

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} \Lambda_{UV}^2 + \dots \quad (1.14)$$

where λ_S is the coupling strength between the scalar particle and the Higgs field. It should be noted that this time correction carries an opposite sign. In this way, the diverging terms can be eliminated if a scalar particle exist with $\lambda_S = \frac{1}{2}|\lambda_f|^2$.

SUSY overcomes the hierarchy problem by introducing the so-called superpartners, which possess the same quantum numbers with original fields except for the spin. As long as the masses of the superpartners are of the order $O(\text{TeV})$, SUSY can still be considered as natural. In order to achieve this, the SUSY breaking needs to be soft. This means that it remains a valid symmetry of the underlying laws of physics but is broken in the course of the evolution of the state of the universe.

Dark matter candidate:

The lower limit on the proton lifetime is set to 10^{30} years by experiments carried out in

the water Cherenkov detector at Super Kamiokande in Japan [35]. This fact strongly suggests that the baryon number is conserved. Unlike the SM, in SUSY theories the baryon and lepton number conservations can be violated. In SUSY, to prevent proton decay, a new quantum number, called as R-parity is introduced:

$$R = (-1)^{3B+L+2S}, \quad (1.15)$$

where B represents the baryon number, L is the lepton number, and S denotes the spin. All the SM particles have positive R-parity, while the superpartners have negative R-parity. Conservation of R-parity entails that the SUSY particles can only be produced in pairs. Moreover, in SUSY, this conservation leads to a stable the lightest supersymmetric particle (LSP). LSP is neutral in terms of both electric and color charge. Thus it is one of the most viable cold dark matter candidates and it has great importance in this thesis. If it exists, it would appear as a high missing energy in the events observed by experiments. There are also R-parity violating SUSY theories, but these are beyond the scope of this work.

Unification of gauge couplings:

Precise measurements of the running weak, strong and electromagnetic couplings performed at the Large Electron-Positron Collider (LEP) indicates that these couplings fail to unify at high energy (see Figure 1.3 a) [32, 33]. In [32], it is shown that the minimal supersymmetric standard model (MSSM) is the only possibility, without an intermediate mass scale. The MSSM can raise the scale of the unification by introducing a particle called the gluino, the spin half partner of the gluon. The gluino partially cancels the asymptotic freedom effect of the gluon itself [34]. The MSSM and its particle content will be explained in section 1.2.2.

Including gravity:

Even if it is not the subject of this thesis, it is worth mentioning that there is a SUSY model including gravity, namely supergravity. This model introduces a spin-2 massless gauge field, the graviton, which mediates gravity with its superpartner, the gravitino. Supergravity, like any other QFT of gravity, is also not renormalizable.

1.2.1 Algebra of Supersymmetry

As it is stated in the beginning of this chapter, a SUSY transformation turns a bosonic state into a fermionic state, and vice versa. SM particles and their superpartners constitute supermultiplets. For each of the chiral¹² SM fermions, there is a separate scalar partner and

¹²left- and right- handed

together they form chiral supermultiplets. The supersymmetric transformation operator Q can also be called as supercharge operator, can be read as:

$$Q|fermion\rangle = |boson\rangle \quad (1.16)$$

$$Q|boson\rangle = |fermion\rangle. \quad (1.17)$$

In SUSY, the number of supercharges characterizes the theory. If there is only one supercharge then it is called a N=1 supersymmetry. If there are two supercharges then there is a N=2 supersymmetry and so on. For chiral fermions, generators Q and Q^\dagger satisfy the commutation and anti-commutation relations:

$$\{Q, Q^\dagger\} = P^\mu, \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0, [P^\mu, Q] = [P^\mu, Q^\dagger] = 0 \quad (1.18)$$

where $P^\mu = (H, \vec{p})$ stands for the spacetime momentum operator in which H is the Hamiltonian and \vec{p} is the three-momentum operator. Similar to chiral supermultiplets, the SM gauge bosons together with their fermionic supersymmetric partners form gauge supermultiplets. The algebra in Equation 1.18 indicates that the partner and the superpartner states have the same mass. However, no scalar superpartners have been found yet, although it would be expected to detect low mass partners, like selectron, experimentally long ago. This indicates that the supersymmetry has to be broken and the superpartners have much larger mass.

1.2.2 Minimal Supersymmetric Standard Model

The minimal supersymmetric extension of the standard model (MSSM) is the only irreducible representation of the SUSY algebra (the aforementioned N=1 supersymmetry case). In other words, this is the model with the least number of additional particles and degrees of freedom. All other representations can be reduced to combinations of chiral and gauge supermultiplets. The MSSM chiral supermultiplets are shown in the Table 1.1 while the gauge multiplets are listed in the Table 1.2. As in the table, all superpartners are represented by a version containing a tilde (\sim) of the original SM particle symbols. The scalar counterparts, are indicated by adding an “s” as initials to the SM fermion names (e.g. selectron) while the fermionic gauge superpartners are represented by appending “ino” (e.g gluino). The subscripts L and R of the sleptons show only the helicity of their SM partners.

In MSSM, there are two Higgs chiral supermultiplets. They are the cure of the gauge anomaly, in other words, they leave the action invariant under the supersymmetry transformation. Moreover, by the construction of SUSY, the Higgs doublet with hypercharge $Y = \frac{1}{2}$ is only able to give masses to the up-type quarks while the one with hypercharge

Names		Spin 0	Spin $\frac{1}{2}$	$SU(3)_C, SU(2)_L, U(1)_Y$
squarks, quarks (x3 families)	Q	$(\tilde{u}_L \ \tilde{d}_L)$	$(u_L \ d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	\tilde{u}	\tilde{u}_R^*	u_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	\tilde{d}	\tilde{d}_R^*	d_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons (x3 families)	L	$(\tilde{\nu} \ \tilde{e}_L)$	$(\nu \ e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	\tilde{e}	\tilde{e}_R^*	e_R^\dagger	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	H_u	$(H_u^+ \ H_u^0)$	$(\tilde{H}_u^+ \ \tilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	H_d	$(H_d^0 \ H_d^-)$	$(\tilde{H}_d^0 \ \tilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 1.1: Chiral supermultiplets of the MSSM

Names	Spin $\frac{1}{2}$	Spin 1	$SU(3)_C, SU(2)_L, U(1)_Y$
gluino, gluon	\tilde{g}	g	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\tilde{W}^\pm \ \tilde{W}^0$	$W^\pm \ W^0$	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B boson	\tilde{B}^0	B^0	$(\mathbf{1}, \mathbf{1}, 0)$

Table 1.2: Gauge supermultiplets of the MSSM

$Y = -\frac{1}{2}$ gives masses to the down-type quarks and charged fermions. Therefore, the $SU(2)_L$ complex scalar fields with hypercharge $Y = \frac{1}{2}$ and $Y = -\frac{1}{2}$ can be denoted by H_u and H_d respectively. The half spin superpartners are called higgsinos and represented as \tilde{H}_u and \tilde{H}_d .

The gluino is a half spin color octet supersymmetric partner of the gluon. The superpartners of the electroweak gauge bosons are called winos and binos. The mixing of \tilde{W}^0 and \tilde{B}^0 are called zino (\tilde{Z}^0) and photino($\tilde{\gamma}$).

Due to the EWSB the higgsinos and electroweak gauginos mix with each other. The mixing between neutral higgsinos and the neutral gauginos results in four mass eigenstates called neutralinos and denoted by $\tilde{\chi}_i^0$ ($i = 1, 2, 3, 4$). Whereas, the charged higgsinos and winos combine to form two mass eigenstates called charginos denoted by $\tilde{\chi}_i^\pm$ ($i = 1, 2$).

As mentioned when discussing the hierarchy problem, the breaking of SUSY needs to be soft. In the MSSM, SUSY breaking is introduced by including a new component (L_{soft}^{MSSM}) to the Lagrangian. The Lagrangian, then, can be read as:

$$L = L^{MSSM} + L_{soft}^{MSSM}. \quad (1.19)$$

The Yukawa and gauge interactions are contained in the first term, L^{MSSM} , while the L_{soft}^{MSSM} is breaking the symmetry. In order not to cause ultraviolet divergences, L_{soft}^{MSSM} includes only the terms with positive mass dimension. In this way, the correction terms of the Higgs mass contain only logarithmic divergences:

$$\Delta m_H^2 = -m_{soft}^2 \left[\frac{\lambda}{16\pi^2} \ln(\Lambda_{UV}/m_{soft}) + \dots \right]. \quad (1.20)$$

In order to satisfy the natural requirement, $|\Delta m_H^2| < m_H^2|_{measured}$, m_{soft} needs to be of the order of TeV scale. This indicates that the new particles introduced by MSSM, if they exist, should be reachable at the LHC.

The L_{soft}^{MSSM} adds 105 new parameters to 19 parameters coming from pure SM. These 105 additional parameters are categorized as follows: 48 CP-violating phases in the gaugino/higgsino and squark/slepton sectors, and 21 squark/slepton masses with 36 mixing angles to define their mass eigenstates.

The large number of parameters of the MSSM make the interpretation of the experimental results challenging, even almost impossible. Therefore, using universality assumptions at the Grand Unified Theory (GUT) scale, the number of free parameters is reduced to five. This model, appropriately, is called as the constrained MSSM (cMSSM). Another approach would be to focus on a limited set of SUSY particle production and decay modes and allow other parameters to change freely. This effective theory approach is called simplified models spectra (SMS) [36] and in this thesis, results are interpreted in the context of SMS. Before introducing the simplified models, it is beneficial to give an overview of the MSSM particle decays. To stay in the scope of this thesis, only the R-parity conserving scenarios will be discussed. In these models, the cascade decays of SUSY particles always end with an LSP. In this search, the lightest neutralino, $\tilde{\chi}_1^0$, is the LSP. A deepened version of the topic is presented in chapter 9 of [4].

Neutralino and chargino are mixtures of the higgsinos and gauginos that are the spin half partners of the Higgs and gauge bosons. A neutralino or chargino can decay into lepton+slepton or quark+squark pairs, given the condition that slepton and squark are light enough. A neutralino or chargino may also decay into a Higgs scalar or an electroweak gauge boson via lighter neutralino or chargino. Therefore, the possible two-body decay modes for neutralinos and charginos in the MSSM are:

$$\tilde{\chi}_i^0 \rightarrow Z \tilde{\chi}_j^0, W \tilde{\chi}_j^\pm, h^0 \tilde{\chi}_j^0, l \tilde{l}, \nu \tilde{\nu}, \quad (1.21)$$

$$\tilde{\chi}_i^\pm \rightarrow W \tilde{\chi}_j^0, Z \tilde{\chi}_j^\pm, h^0 \tilde{\chi}_j^\pm, l \tilde{\nu}, \nu \tilde{l}. \quad (1.22)$$

Gluino decays via a squark exclusively, which can be either on- or off-shell. If the on-shell decays are allowed, then $\tilde{g} \rightarrow t \tilde{t}$ and $\tilde{g} \rightarrow b \tilde{b}$ are probably the dominating channels.

Otherwise the squarks are off-shell and this favors the following decays: $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_i^0$ and $\tilde{g} \rightarrow q\bar{q}'\tilde{\chi}_i^\pm$.

1.2.3 Simplified Models

The simplified model spectra (SMS) involves only a part of the parameter space of MSSM. This simplification leaves a few phenomenologically relevant parameters to understand the SUSY models; the cross section, branching ratios, and masses. In this simplified framework, normally a limited number of decay channels are considered. In other words, the branching ratios are set to 100% or a linear combination of such models is used for mixed decays. In this way, results can also be reinterpreted within other (non-)SUSY theories.

An example model ($\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0$):

In this simplified model, gluino decays to two light quarks and a neutralino. Generally, the neutralino is considered to be the stable lightest supersymmetric particle or LSP. Hence, It is a strong candidate of Dark Matter. In this model concerning direct decay, the parameters include the mass of the gluino and neutralino ($m_{\tilde{g}}, m_{\tilde{\chi}_1^0}$) and the production cross section of the gluino $\sigma(pp \rightarrow \tilde{g}\tilde{g} + X)$. The case where gluino decays to two light quarks and an intermediate chargino, with the latter decaying to a gauge boson and a neutralino can be read as:

$$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^\pm \rightarrow q\bar{q}(W^\pm\tilde{\chi}_1^0) \quad (1.23)$$

$$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_2^0 \rightarrow q\bar{q}(Z^0\tilde{\chi}_1^0). \quad (1.24)$$

In this case, the mass of the intermediate particle (chargino $\tilde{\chi}_1^\pm$ or a heavier neutralino $\tilde{\chi}_2^0$) is also included in the parameters. In order to reduce the three-dimensional mass space, chargino mass is considered to be dependent to gluino and neutralino masses. The relation is given by: $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_1^0} + r(m_{\tilde{g}} - m_{\tilde{\chi}_1^0})$. The kinematics of the decay strongly related to the value of r . To further decrease the number of parameters, the branching ratio of these decays is set to 100%.

In simplified models of gluino pair production, the case where each gluino decays as shown in the Equation 1.23 with a $r=0.5$, which indicates the mass of the chargino is at the midway, is called T5qqqqWW while the direct decay of gluino to top quark pair and a neutralino is called as T1ttt¹³. The diagrams for T1ttt and T5qqqqWW can be seen in Figure 1.5 left and right respectively. In this thesis, the T5qqqqWW model is used to interpret the results which is well motivated in the searches at LHC [43, 44].

¹³this naming is used within the CMS collaboration

A variety of SMS approaches has been investigated by the ATLAS and CMS collaborations at the LHC. In the following, a short history of SUSY searches at colliders will be discussed.

1.2.4 Short History of SUSY searches at colliders

Before the LHC, the first constraints on SUSY had already been set by UA1 experiment and the UA2 experiment at Super Proton Synchrotron. It was until 2000, SUSY searches continued up to 209 GeV at the electron-positron collider LEP. Until 2011, The CDF and DØ experiments at the Tevatron extended the limits in the context of cMSSM. In parallel between 1992 and 2007, the H1 and ZEUS experiments at HERA (The electron-proton collider) searched for the R-parity violating production of single SUSY particles. The LHC has started proton-proton collision at a center of mass energy of 7 TeV in 2010. Since then, ATLAS and CMS experiments have performed robust searches and provided strong limits in the context of SMS. The 2011 and 2012 runs of the LHC are called Run 1. During Run 1, 20 fb^{-1} data was collected at a center of mass energy 7 and 8 TeV. In 2012, with the discovery of Higgs boson, the SUSY models have been constrained further. Later, in the first stage of Run 2, approximately 36 fb^{-1} data has been collected at a center of mass energy 13 TeV. A summary of the both Run 1 and Run 2 results by ATLAS and CMS experiments is presented in Figure 7.9 including the results from present search. Within the MSSM these results have already pushed the sparticle mass boundaries required to solve hierarchy/naturalness/fine-tunning problems. For instance, It was foreseen that the gluinos cannot be much heavier than 1 TeV and the higgsino mass should be around 200 GeV. Obviously, the latest results challenge these requirements. Nevertheless, there is still room for expanding the searches, for example, in the regions with compressed spectra where the analysis is not so trivial. **Expectations for gluino searches at 13 TeV:**

LHC is a proton-proton collider. Consequently, high sensitivity is favored in the production of color charged particles: the gluino and squarks.

The gluino pair production diagrams are shown in Figure 1.6. The cross sections as a function of the center of mass energy and gluino mass are shown respectively in Figure 1.7 left

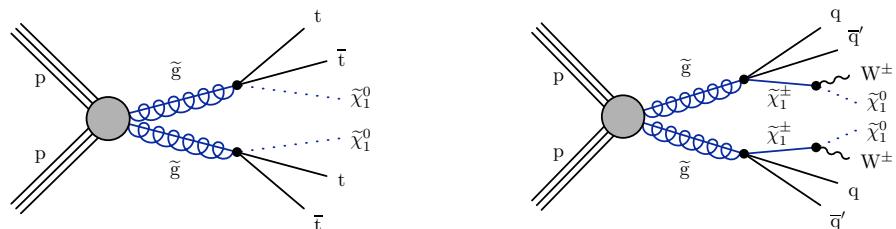


Figure 1.5: Diagrams showing the simplified models T1tttt (left) and T5qqqqWW (right).

and right. The cross section of gluino pair production increased significantly from 8 TeV to 13 TeV. However, the cross section of SM processes increased as well (see Figure 1.9). These figures indicate that at 13 TeV for 10^{10} SM events, 1 event including gluino pair production is expected. This sounds even harder than looking for a 4-leaf clover. Nevertheless, it is still achievable by designing an analysis strategy, with a robust background estimation and assigning the uncertainties carefully. An example of such a search will be presented in this thesis and in the section 7.4 a comparison of the latest results from gluino searches can be found.

History of the present analysis:

As mentioned in the previous section, in the present thesis, a search for SUSY is performed in the context of the simplified model T5qqqqWW. The analysis journey started in Run1 with a search for a similar model, which was T1tttt. In this model, the presence of 2 top-antitop quark pairs requires a search for multiple b-tagged (see Sec. 3.2.4 for the performance of tagging) quarks in the final state. No significant deviation from the predicted SM background is observed. The related Run 1 results from CMS can be found in Ref. [42] In Run2, to increase the discovery potential of the analysis, the channels, which are sen-

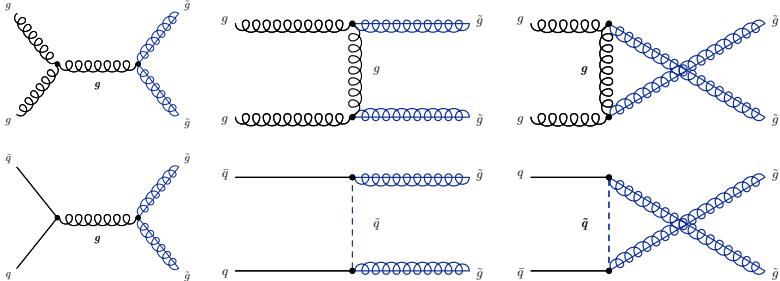


Figure 1.6: Feynman Diagrams showing the gluino pair production. SUSY particles are shown in blue.

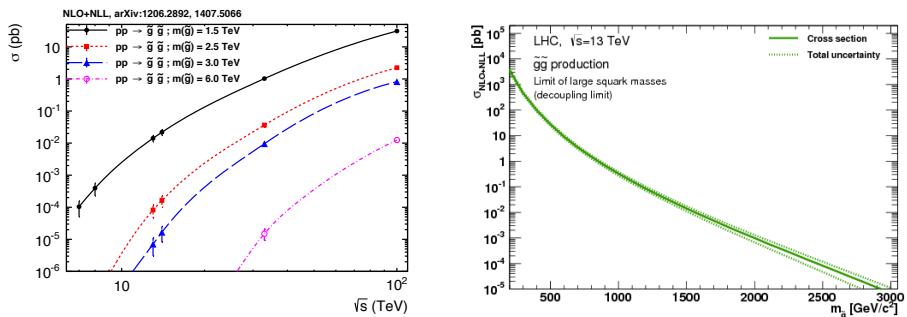


Figure 1.7: left: Cross sections of gluino pair production for four different mass points are given as a function of the center of mass energy. right: Cross section of gluino pair production at 13 TeV center of mass energy is given as a function of gluino mass. [37]

sitive to zero b (tagged) final states, are included. These sensitive processes are called T5qqqqWW models. The results had already been presented in the *Moriond* conference in 2016, with 2.3 fb^{-1} integrated luminosity. Unfortunately, again no significant deviation from the predicted SM background is observed in both of the channels.

In Figure 1.8, the exclusion of the masses, which are presented in *Moriond 2016* conference, for $m_{\tilde{g}}$ and $m_{\tilde{\chi}_1^0}$ is shown for the model T1tttt (left) and T5qqqqWW (right).

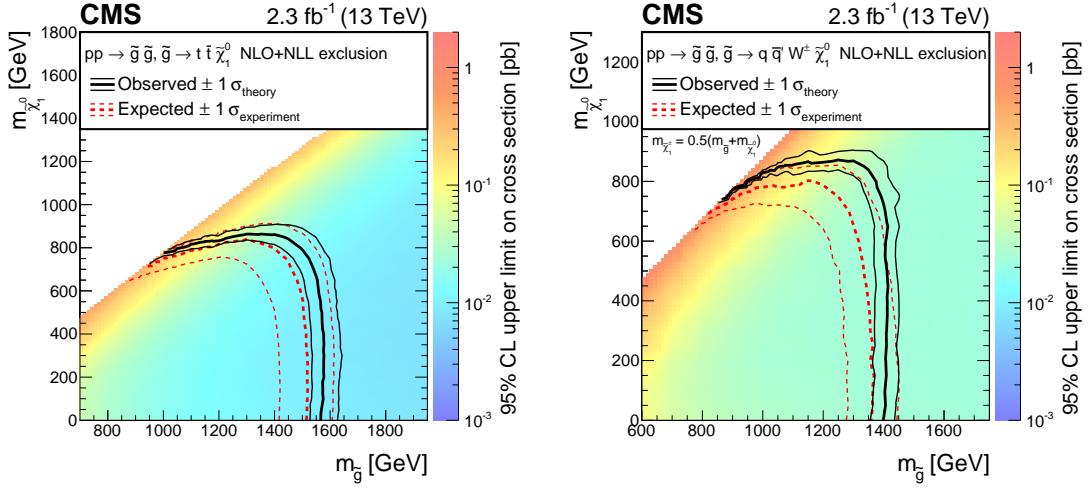


Figure 1.8: Cross section limits at a 95% CL for (left) the T1tttt model, and (right) the T5qqqqWW model [60].

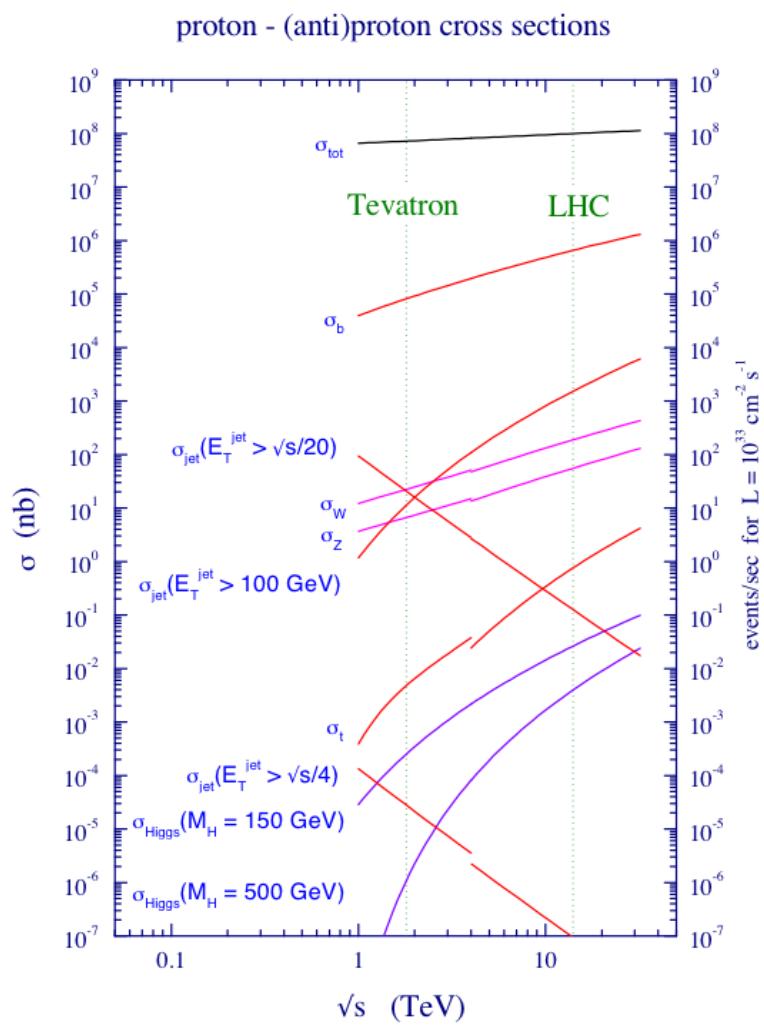


Figure 1.9: Cross section of SM processes as a function of center of energy. [38]

Chapter 2

Experimental Setup

Contents

2.1	The Large Hadron Collider at CERN	21
2.1.1	The CERN accelerator complex	22
2.1.2	The Large Hadron Collider	22
2.2	The Compact Muon Solenoid experiment at the LHC	25
2.2.1	Superconducting Magnet	26
2.2.2	Tracker	26
2.2.3	Electromagnetic Calorimeter	27
2.2.4	Hadronic Calorimeter	27
2.2.5	Muon System	28
2.2.6	Trigger and Data Acquisition Systems	28
2.2.7	Luminosity measurement	29
2.3	Event simulation	29
2.3.1	Event generation	29
2.3.2	Detector simulation	30

2.1 The Large Hadron Collider at CERN

An overview of the experimental setup used to collect data for this analysis is explained in the chapter. In this search, the proton-proton collision data recorded by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) at a center-of-mass energy of 13 TeV is used. The chapter begins with a brief explanation of the LHC including the pre-accelerator complex. The different components and subsystems of the CMS detector will be discussed in Sec. 2.2. The chapter will end with a description of the event simulation tools in Sec. 2.3.

2.1.1 The CERN accelerator complex

A particle accelerator is a machine that moves charged particles by using electromagnetic fields. Nowadays, accelerator machines are using changing electromagnetic fields to propel particles to nearly the speed of light. The large hadron collider (LHC) is the world's largest particle accelerator and it is the last ring of the accelerator chain at CERN. CERN is an abbreviation for the European Organization for Nuclear Research - Conseil Européen pour la Recherche Nucléaire. The complex is situated about 100 meters below the ground, between the Jura Mountains and Lake Geneva.

Fig.2.1 shows the entire CERN accelerating complex including the storage rings, beam transfer lines and four major LHC experiments. In this thesis, the data collected from the proton-proton collisions in LHC is used. The LHC is the final stage of a successive acceleration procedure of protons. The protons begin their journey in a hydrogen bottle which is currently at the beginning of the linear accelerator (LINAC2) which propels the particles to 50 MeV. This linear accelerator has served to many experiments at CERN since 1978 and it will be replaced by newly commissioned LINAC4 in 2020[62]. The beams provided by the LINAC are then injected to the Proton Synchrotron Booster where the particles are further accelerated to 1.4 GeV [63]. Next, the protons delivered by the Booster are accelerated up to 25 GeV by the Proton Synchrotron (PS) [64]. This accelerator is operating since 1959 and it is the first accelerator of CERN. The last machine before the LHC is the Super Proton Synchrotron (SPS) [65], where the protons reach the energy of 450 GeV. The SPS was operating as a proton-antiproton collider from 1981 to 1984, which leads to the discovery of W [66] and Z [67] bosons. This journey of protons to reach the LHC takes 16 minutes in the current injection scheme [68].

2.1.2 The Large Hadron Collider

The LHC is primarily designed to collide protons at a center-of-mass energy of 14 TeV. The LHC tunnel is originally build for the LEP with a total circumference of 27 km. After shutdown of the LEP the construction of LHC began in 2000 and the first beams were circulated in September, 2008. Due to an electrical malfunction in the magnet system the first collisions are delayed until November, 2009.

The LHC incorporates different methods already probed at different synchrotrons worldwide at the same time pushing the limits even further. In total more than 9000 magnets are used in the LHC. 1232 of them are the main superconducting dipoles responsible for the bending of the beam. To focus the beam, superconducting quadrupole magnets are used, while the remaining sextupole and higher order magnets are used to correct the beam chromaticity.

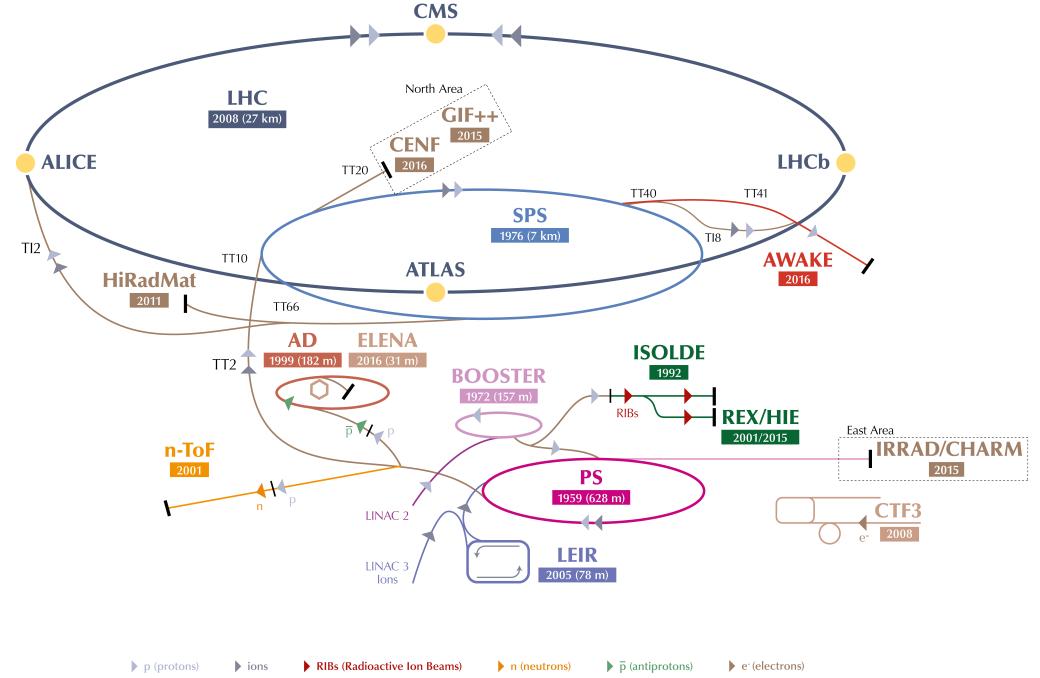


Figure 2.1: CERN accelerator complex [61]:

The proton bunches from SPS are injected to the two counter-rotating rings of the LHC with a 450 GeV. The beam acceleration take place at a point between ALICE and CMS experiments, where 16 superconducting cavities are mounted. Each beam gains an acceleration of 485 keV per turn, which provides the necessary acceleration to reach the 6.5 TeV in about 20 minutes at the moment.

The two beams collide in four intersection points where the detectors of the experiments are located.

For the interested reader, the LHC design details are introduced here [68].

LHC parameters

One important parameter of LHC relevant to the physics analyses is the particle density delivered by LHC which can be quantified with its *luminosity*. The instantaneous luminosity, which is denoted by \mathcal{L} , is estimated using a variety of beam parameters:

$$\mathcal{L} = \frac{N_p^2 N_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F, \quad (2.1)$$

where N_p is the number of particles per bunch, N_b is the number of bunches per beam, f_{rev} is revolution frequency, γ_r is relativistic gamma factor, ϵ_n is the normalized transverse beam emittance, and β^* is the beta function at the interaction point. The parameter F is the reduction factor because of the beam crossing angle.

The total integrated luminosity over time can be calculated:

$$L = \int \mathcal{L} dt. \quad (2.2)$$

Using the instantaneous luminosity the expected event rate of a process can be computed:

$$\frac{dN}{dt} = \sigma \cdot \mathcal{L}, \quad (2.3)$$

where σ is the cross section of a specific process.

At the design energy of 14 TeV and with a total bunch crossing rate of 40 MHz, a peak luminosity of $\mathcal{L} = 10^{34} cm^{-2}s^{-1}$ would be achieved. However, to reach these conditions, prior studies and commissioning of the machine at lower levels should be performed. Therefore, the first proton-proton (pp) run of the LHC was performed at a centre-of-mass energy of 7 TeV with a peak luminosity of $\mathcal{L} = 3.6 \times 10^{33} cm^{-2}s^{-1}$. Total integrated luminosity of around $5 fb^{-1}$ was delivered to the major ATLAS (A Toroidal LHC AparatuS) and CMS experiments until the end of 2011 [70]. Next, following the 8 TeV run (*Run 1*) in 2012 the first long shutdown which took place for 1.5 year, allowed to accelerate protons to their current energy of 6.5 TeV. In 2015 the second LHC run, namely *Run 2*, has begun with a peak luminosity of $\mathcal{L} = 5 \times 10^{33} cm^{-2}s^{-1}$ at a bunch spacing of 25 ns. A full list of the LHC parameters for the 13 TeV runs can be accessed here [69]. The present analysis uses the data collected in 2016 with an integrated luminosity of $35.9 fb^{-1}$ by CMS experiment.

The Future of the LHC

According to the LHC plans, the current research programme will continue until 2023 including *Run 3* with 14 TeV center-of-energy and the target integrated luminosity is $500 fb^{-1}$ [71]. LHC further future plans are constraint by the fixed ring size, thus the limited proton energy. Therefore the LHC proceeds in the direction of the intensity frontline. High-Luminosity (HL-LHC) project is planned to start at the beginning of 2024. HL-LHC could reach luminosities of the order of $\mathcal{L} = 10^{35} cm^{-2}s^{-1}$ which is hundred times more than the LHC design. The main challenge in this environment would be the development of new magnet system to handle with such an increase in number of protons per beam.

2.2 The Compact Muon Solenoid experiment at the LHC

The Compact Muon Solenoid (CMS) [73] detector is a multi-purpose device to operate at the LHC. CMS is installed about 100 metres underground close to the French village of Cessy. The first Letter of Intent for the CMS detector was published in 1992 [74], it was followed by the construction of the individual components soon afterwards.

As it can be seen from the Fig. 2.2, CMS detector is in onion shape, meaning it incorporates successive layers. These layers which are corresponding to various purpose sub-detectors can be listed as follows: the inner tracking system is surrounding the interaction point; afterwards the electromagnetic and hadron calorimeters are located between the tracker and 6m wide solenoid magnet. Outermost layer is the extensive system of muon trackers which infact makes up more than 80% of the detector volume. In addition to these, forward calorimeters are placed along the beam-pipe immediately after the muon detectors.

Coordinate system and the relevant kinematic variables

The origin is centered at the nominal collision point inside the experiment. The y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC. Therefore, the z-axis points along the beam direction following the direction of the counter-clockwise rotating beam. Due to the almost symmetric cylindrical shape of the detector, generally, cylindrical coordinate system is used. In other words, the same z-axis with the 2-D polar coordinates, which involve the radial distance r and azimuthal angle ϕ , is used. The polar angle θ is measured from the z-axis.

The pseudorapidity η , derived from the θ , is also a widely used variable and it characterizes the boost of the particles:

$$\eta = -\ln[\tan(\frac{\theta}{2})]. \quad (2.4)$$

The pseudorapidity is a version of rapidity, which is Lorentz invariant under longitudinal along the beam axis - boosts. In the case where particle masses are negligible, rapidity converges to η .

Another variable used commonly is the spacial distance:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (2.5)$$

The proton is a composite particle, which includes partons. Therefore, the proton-proton collisions are interactions of the corresponding partons. Furthermore, even though the energy of the incoming proton beams are symmetric, this condition does not necessarily need to be hold for individual partons. Therefore, the total center-of-mass energy is an unpredictable quantity of an event. However, since the incoming protons, therefore the

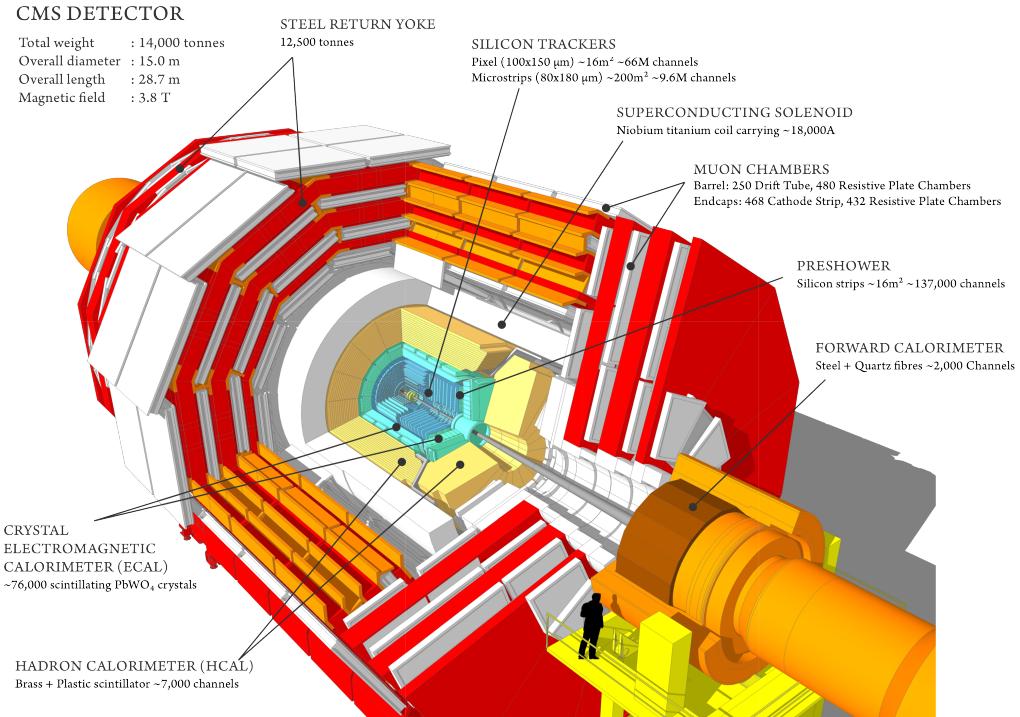


Figure 2.2: Schematic view of the CMS detector with a cut-out quadrant [72]

partons, are aligned along the beam axis, the total momentum in the transverse plane is close to be zero, which is followed also by the momentum conservation of the collision remnants. As a result of this, the transverse momentum of the particles, $\sqrt{p_x^2 + p_y^2}$, are extremely useful for the physics analyses.

2.2.1 Superconducting Magnet

As it can be understood from the name of the CMS experiment, the solenoid magnet is a crucial component of the detector. It has an important role in precise measurements of charged particles. It is designed to host the tracker and calorimeter systems at the same time providing a uniform axial magnetic field of 4T within its volume. Due to its unique design and not yet well-understood aging, it is currently being operated at slightly lower field strength of 3.8T. Further information on the measurement of the magnetic field inside the barrel yoke can be found here [75].

2.2.2 Tracker

The inner tracking system comprises the innermost part of the CMS experiment. The main purpose of this component is the identification and measurement of tracks ascending from

charged particles coming from the interaction vertices. This detector system has several restraint requirements in terms of detection efficiency, spacial resolution and radiation safety due to the large collision rates and high particle multiplicities.

To overcome these challenges, the tracker system consists of silicon pixel and strip sensors. The Silicon Pixel detector is located just around the interaction region. It covers a pseudo-rapidity range of $|\eta| < 2.5$.

2.2.3 Electromagnetic Calorimeter

The succeeding sub-detector to the inner tracker is the electromagnetic calorimeter (ECAL). ECAL is a hermetic homogeneous detector, which is composed of lead tungstate (PbWO_4) crystals. The lead tungstate scintillators leads to a compact calorimeter design, which is important to cope with the space restrictions inside the magnet. The ECAL consists of endcap components on each side and a barrel part. The barrel covers a range of $|\eta| < 1.479$ and the endcap extends the range until $|\eta| = 3.0$. Additionally, a preshower detector is installed in front of the endcap. The main purpose of the ECAL is to measure the energy of electrons and photons coming from the tracking system. The preshower prevents the misidentification of photons in ECAL originating from neutral pion decays. To detect the showers produced in the crystals in the barrel and endcap, photodiodes and phototriodes are used respectively. The energy resolution of ECAL barrel is measured to be [76]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2, \quad (2.6)$$

where the three contributions correspond to the stochastic, noise, and constant terms. The constant terms cover the non-uniformities, intercalibration errors and energy leakage from the back of the crystal. According to this formula, a 100 GeV electron can be measured with a resolution of 0.4%.

2.2.4 Hadronic Calorimeter

The Hadron Calorimeter (HCAL) is installed directly after ECAL without any dead material in between. The principle purpose of the HCAL is to measure the energy of hadrons, which are particles consist of quarks and gluons. The HCAL is a sampling calorimeter in other words it finds a particles position, energy and arrival time using layers of brass and scintillator plates. When the particle passes through from these layers, a rapid light pulse is produced. Then optic fibers feed the light into readout boxes where photodetectors amplify the signal. A particles energy is then measured by summing over the amount of light in a given region of many layers of strips in depth, which is called HCAL tower. HCAL has two components inside the solenoid: the barrel (HB) part covers a pseudorapidity range of

$|\eta| < 1.3$ and the endcap (HE) is placed in the range of $1.3 < |\eta| < 3$. There is an outer hadron calorimeter (HO) just after the magnet to stop the particles surviving after the HB; therefore this component is also known as a tail catcher. HO utilizes the material of the magnet as an absorber. In addition to the central part, the coverage of HCAL is extended by a very forward calorimeter (HCAL). The HF is installed outside the CMS endcaps at a distance of $|z| = \pm 11.2\text{m}$ from the collision vertex. The pseudorapidity range of HF is $2.85 < |\eta| < 5.19$, thus it is expected that the HF is exposed to extreme particle flux. Aforementioned leads to a choice of radiation-hard technology: steel is used as absorber material while quartz-fibers are used as active material.

2.2.5 Muon System

The importance of the muon system has been stressed out in the name of the CMS experiment. Muon system has three major roles: identification, measurement of momentum and triggering of muons. A combination of three different particle detectors is utilized to achieve these goals. In barrel region the magnetic field is uniform and the muon flux is low. Therefore, drift tubes are used in this section corresponding to a pseudorapidity range of $|\eta| < 1.2$. The drift tube system serves as tracking detectors by measuring the electromagnetic cascades accompanying the muons. Outside of the barrel yoke, towards the endcaps, the magnetic field becomes stronger and non-uniform. Moreover, particle flux increases in this region. Therefore, the detectors in the endcaps have to be fast and radiation hard. To overcome these challenges finely segmented cathode strip chambers (CSC), which are located in the pseudorapidity range $0.9 < |\eta| < 2.4$, are used. CSCs are used to measure the radial position of muons. In addition to drift tubes and cathode strip chambers, the resistive plate chambers are installed to supplement the triggering system in a fast and independent way. In the muon reconstruction, informations from the tracker and muon systems are used together to attain the best momentum resolution.

2.2.6 Trigger and Data Acquisition Systems

Given the 25ns time interval between the proton bunches, the collision rate inside the detector reaches up to 45MHz. An event with all the measured information occupies about 1 MB in size. To save all events, a bandwidth of 60TB/s would be needed. However, achieving such high transmission of data is impossible with todays technology. Therefore, a trigger system is necessary to select the most interesting events. The CMS trigger system is composed of two levels. The first level (L1) utilizes the information from the calorimeters and muon detectors. It consists of custom hardware processors. The accepted events are then sent with an output rate of maximum 100kHz to the second level, which is also

known as High Level Trigger (HLT). The main purpose of HLT is to reconstruct physics object and to define events with interesting features. In the HLT stage, events are discarded as soon as the available information is enough to take the decision before being fully reconstructed. The HLT further decreases the event rate to an order of magnitude 100 Hz before storage. The HLT requires an immense parallelism to handle the event rate coming from L1 trigger.

2.2.7 Luminosity measurement

Luminosity information is essential for both the beam parameters and physics analysis. For physics analysis, the integrated luminosity is related to the number of events of a certain process through Eqn. 2.3. The instantaneous luminosity is measured using two dedicated detectors: the Fast Beam Conditions Monitor (BCM1f) [78] and the Pixel Luminosity Telescope (PLT). In addition to dedicated detectors, the pixel tracker, drift tubes and hadron forward calorimeter systems are used as well.

In this thesis, the total data collected by the CMS experiment corresponds to a 35.9 fb^{-1} integrated luminosity and uncertainty on it is measured to be 2.6% [50].

2.3 Event simulation

Monte Carlo (MC) simulations provide events close-to-reality. Both the experimental and theoretical physics searches benefit from MC simulations to study the dedicated physics processes as well as the detector methods. In the following, a short overview of simulation technics used in this analysis is presented.

2.3.1 Event generation

The simulation of proton-proton collisions is an extremely difficult task due to the composite structure of protons. This complicated task can be divided into stages with decreasing energy scale.

First, the scattering amplitude is computed using the matrix element (ME) where the proton parton distribution functions (PDFs) are used to describe the initial state momenta of partons. Naturally, being the first step, the hardest particles of the collision are produced in this stage. Second, the initial-state radiation (ISR) and final-state radiation (FSR) are modeled in the parton shower stage. In parton shower, relative transverse momenta evolve from a high scale to lower values. Here an infrared cutoff on the parton showers is needed due to the perturbation theory. Finally, hadrons are formed using hadronization models.

MC samples used in this thesis are simulated with the `MADGRAPH5_AMC@NLOv2.2.2` or `v.2.3.3` [80] and `POWHEGv2.0` [81–85]. Further details are given in Sec. 4.2 when separate background processes are discussed. A more profound introduction to MC generators used in LHC physics can be found in [79].

2.3.2 Detector simulation

The final state particles, which are emerging after the generation, showering and hadronization states are required to be propagated to detector simulation. This stage involves the modeling of interactions between those particles and the detector hardware material as well as the readout electronics of CMS. In CMS the main technic is called Full simulation, for which the GEometry ANd Tracking (GEANT4) toolkit is used [87]. However, the Full simulation requires massive computer power; as a result, for large-scale MC production, like SUSY parameter phase scans, a more effective technic is necessary. Therefore, CMS established a Fast Simulation (FastSim) software [88]. Several features of the detector geometry and particle-matter interactions are simplified to increase the computational power. Detector responses of FastSim samples are tuned to match to the full simulation.

Chapter 3

Object reconstruction and identification

Contents

3.1	Particle-Flow algorithm	31
3.2	Physics Object reconstruction and identification	32
3.2.1	Primary vertices	32
3.2.2	Leptons	33
3.2.3	Jets	35
3.2.4	b tagged Jets	36
3.2.5	Missing transverse energy	37

The form of the raw data collected by the CMS detector is in detector hits and deposited charge information. Therefore, it is not suitable to be used directly in the physics analyses. In this chapter the reconstruction and identification of physics objects used in the CMS experiment is discoursed. These objects and their kinematic characteristics are used in building the event selection criteria which will be discussed in Chapter 4.

3.1 Particle-Flow algorithm

Figure 3.1 displays the trails of the particles passing trough the sub-detectors of CMS experiment. Only energy deposits in the most relevant sub-detectors are presented otherwise a particle can leave trace in multiple detector layers. The particle-flow (PF) algorithm [90, 91], uses the information obtained from the all detector subsystems in the object reconstruction. Three main features of the CMS PF algorithm are tracks reconstructed from hits in the tracker, energy deposits in the calorimeters and tracks reconstructed from hits in the muon system. In order to archive object reconstruction with a high efficiency and low fake rate, iterative tracking and the calorimeter clustering methods are used in the PF algorithm. The Kalman filter [92] method, which ensures a proper track finding even for

objects with low p_T , is adopted to identify single particle tracks. Calorimeter clustering is responsible for several tasks such as the detection and measurement of stable neutral particles, their isolation from charged objects, reconstruction and identification of electrons and corresponding Bremsstrahlung photons. After determining the single elements, a linking algorithm is used to form a block from separate elements that are possibly related to a single object. After the construction of blocks, the actual particle reconstruction and identification is performed. The objects are reconstructed in a sequence of highest reconstruction performance to lowest, thus already identified blocks are gradually discarded for the more ambiguous object reconstructions.

For muons, the tracks in the muon system is fit to the inner tracks by requiring an acceptable χ^2 . These muons are also called global muons. If the momentum of a global muon is compatible with the tracker-only measurement within three standard deviations, a particle-flow muon is formed.

Generally, an electron goes into a bremsstrahlung radiation, which results in a more complicated reconstruction procedure than the muons. Gaussian sum filter [93] algorithm which also accounts for the possible energy loses is used to perform the track fitting. Moreover, a second algorithm based on ECAL superclusters determines the electromagnetic energy deposits that are aligned to electron candidate tracks. A particle-flow electron is formed, if the track-cluster compatibility is satisfied.

At this point, after removing the identified muons and electrons tracks, the neutral and charged hadrons are reconstructed from the remaining tracks and the matching calorimeter deposits. The reconstruction for photons and τ -leptons will not be described since they are not used in this analysis. Detailed discription of these objects can be found here [94]. In this analysis, the kinematic variables, which are introduced in Chapter 4, are based on the reconstructed objects satisfying certain identification criteria.

3.2 Physics Object reconstruction and identification

3.2.1 Primary vertices

To describe an event, the primary vertex (PV) is the main quantity. The primary vertex is where the interesting parton collision takes place.

The very high instantaneous luminosities of the LHC results in a quite high probability of multiple proton-proton interactions per bunch crossing, pile-up. Therefore, a precise reconstruction of PV is a challenging task.

The reconstruction commences with selecting a set of tracks based on some quality criteria such as the number of hits in the tracker and the impact parameter. Next, the selected

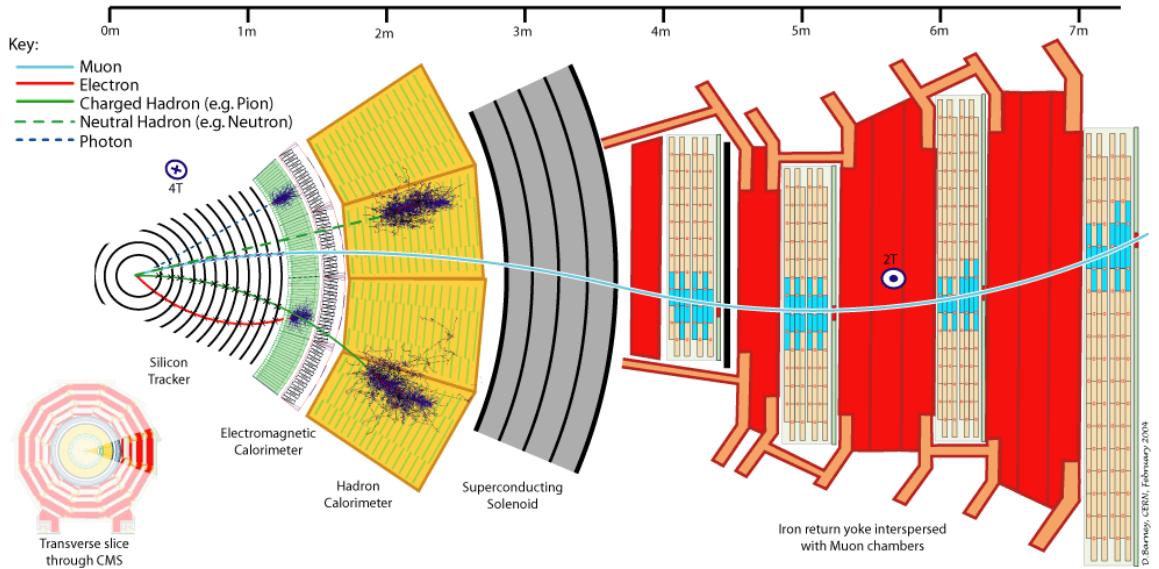


Figure 3.1: Cross section of a slice of the CMS detector in the transverse plane with different particles crossing the subsystems is shown [89]. Only energy deposits in the most relevant subdetectors are presented for simplicity.

tracks are clustered according to their z-coordinate. A vertex fit [95] is then applied on the clusters that more than 1cm apart and contain at least two tracks. Finally, the primary vertex of interest is chosen to be the cluster with the largest sum of the squared track momenta.

A vertex is assigned as good if the number of degrees of freedom of the fit result is larger than four. In the present analysis, at least a good primary vertex is required in the event selection.

3.2.2 Leptons

In the context of this thesis, only the events with single lepton are considered; therefore lepton reconstruction and identification has a great importance. Additional requirements on the particle flow leptons are applied to ensure choosing a good quality lepton or rejecting the fake ones.

In this analysis, only the leptons in the acceptance of $|\eta| < 2.4$ with a transverse momentum of $p_T > 10$ GeV are considered.

3.2.2.1 Muons

First, the muons are defined as good if they satisfy several CMS-wide identification criteria. A candidate is a good muon if the fraction of valid tracker hits above eighty percent.

Furthermore, several conditions are imposed according the segment compatibility evaluation. The segment compatibility is used to quantify the compatibility of a tracker-muon object with the muon hypothesis [96]. If the segment compatibility computation returns a value larger than 0.451 then the muon is selected, otherwise if it is below 0.303 the muon is rejected. For the candidates with intermediate segment compatibility further requirements are levied: the muon is a global muon, the global fit gives a χ^2 less than three, the χ^2 of the matching between the tracker and muon is below twelve and the value returned by the kink finder is below twenty.

In addition to CMS-wide identification selections, several criteria specific to this analysis are also applied. Two types of muons are defined as tight and loose. The loose muons are required to have a mini isolation [97], for which an explanation can be found following to this paragraph, less than 0.4. The candidates with $p_T > 25$ GeV, the mini isolation less than 0.2 and have the three-dimensional impact parameter significance less than four are defined as tight muons.

In this search, only events with one tight muon and no other loose muons are considered if the lepton is not electron. The isolation of a lepton is a measure of the relative additional activity in a predefined lepton cone. For a more analysis specific selection, a varying cone radius is selected according to the p_T of the lepton. For a lepton coming from a boosted W boson, which then will have a high p_T , a narrower cone size is chosen. More precisely, cone radius is set to 0.05 for the leptons with $p_T \geq 200$ GeV, while for the leptons have p_T less than 50 GeV a larger cone radius of 0.2 is considered. For the leptons with intermediate p_T , con size is given by the formula: $\frac{10\text{GeV}}{p_{T,\text{lep}}}$. The mini isolation is also used for the electron selection for which the criteria will be given in the following.

3.2.2.2 Electrons

As mentioned earlier the particle flow reconstruction of good quality electrons [98] is more elaborate than the muons case. Similarly to the muons case, a list of CMS-wide identification requirements are imposed on the particle flow electrons identified by using Gaussian sum filter. The list of criteria with short descriptions is given in Tab. 3.1. The requirements are slightly changing between the Barrel and the Endcap regions, due to the different granularity of ECAL. Those criteria [99] in the list involve the shower shape quality and the cluster energy and track momentum compatibility. In addition to the list no associated photon conversion vertex is required.

On top of the CMS-wide identifications also analysis specific selection criteria is also applied. In the electron case, leptons with $p_T > 25$ GeV and with mini isolation smaller than

0.1 are defined as tight. For electrons, due to the identification inefficiency in the overlapping region between the ECAL barrel and endcaps, the corresponding pseudorapidity region $1.44 < |\eta| < 1.56$ is excluded. The number of leptons fail to satify these criteria is set to be zero if the lepton is not muon.

Selection	Definition	Barrel	Endcap
$\sigma_{i\eta i\eta}$	ECAL crystal-based shower covariance in the direction of η	0.0101	0.0283
$\Delta\eta_{in}$	Difference between the supercluster position in the ECAL and the track direction at the innermost tracker position in η	0.00926	0.00724
$\Delta\phi_{in}$	Difference between the supercluster position in the ECAL and the track direction at the innermost tracker position in ϕ	0.0336	0.0918
H/E	Ratio of energy measured in the HCAL over the energy measured in the ECAL	0.0597	0.0615
$ 1/E - 1/p $	Absolute difference between the inverse electron energy measured in the ECAL and the inverse momentum measured in the tracker	0.012	0.00999
Δ_{xy}	Track-vertex distance in the transverse plane	0.0111	0.0351
Δ_z	Track-vertex distance along the beam axis	0.0466	0.417
N_{hits}^{miss}	Number of missing hits in the electron inner layer track	≤ 2	≤ 1

Table 3.1: List of selection criteria for the CMS electron identification.

3.2.3 Jets

In the context of this thesis, in the decays of supersymmetric gluinos, an abundance of jets is expected in the final state.

As already briefly mentioned in Sec.1.1.2, cascade decays of quarks and gluons leads to a collimated spray of hadrons. In experiments, this signature is also known as jet. As jets of particles propagate through the CMS detector, according to their particle content, they leave traces in all of the sub-detectors. To form a reconstructed jet, all these detector responses are combined using various jet algorithms. In CMS there are two types of jet clustering algorithms in use: Cone algorithms and sequential algorithms. The first one uses a geometrical shape assumption, while the latter starts with clustering two of the

closest objects together and iteratively reconstruct a closed area of objects until a truncation criterion is satisfied [104].

The distance used as the termination criterion can be formulated in generalised form as:

$$d_{i,j} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta_{i,j}^2}{R^2}, \quad (3.1)$$

$$d_{i,B} = k_{t,i}^{2p}, \quad (3.2)$$

where $d_{i,j}$ represents the distance between entities (particles, pseudojets) i and j and $d_{i,B}$ is for the distance between the entity i and the beam (B). In the formula, $\Delta_{i,j}^2$ is the spatial distance in $y - \phi$ plane and $k_{t,i}$ is the transverse momentum. The $d_{i,j}$ is calculated iteratively until it is $d_{i,B}$ and then i is called a jet and it is removed from the list of entities. According to different values of p , which is governing the relative power of the energy versus geometrical ($\Delta_{i,j}$) scales, three main algorithms can be defined.

The case of $p = 0$ corresponds to the Cambridge/Aachen algorithm where the termination condition is described by purely geometrical distance. The $p = 1$ case is called the k_t algorithm which involves the transverse momenta of the entities as well as the geometrical distance parameter.

In this analysis, the anti- k_t algorithm, which is the case of $p = -1$, is used to cluster the particle flow candidates into jets. For the curious readers, the anti- k_t algorithm and its performance in comparison to other algorithms are explained here [103].

In the present search, particle flow candidates are clustered into $R = 0.4$ jets. Jets are required to be within the pseudo rapidity range of $|\eta| < 2.4$ which corresponds to the tracker acceptance. Moreover, if the energy fraction of one of the components such as the neutral, charged electromagnetic or hadronic exceeds 99% then the jet is rejected.

The main purpose of the algorithm is to reproduce the energy of the original parton prior to the shower. However, the measured jet energy is not the same as the corresponding parton. There are several reasons such as the non-uniform detector response, electronic noise responsible for this difference. In CMS, these effects mostly covered with Jet Energy Calibrations (JEC) [105].

In this search, after all corrections are applied to a reconstructed jet, the transverse momentum is required to be larger than 30 GeV.

3.2.4 b tagged Jets

For many SM and new physics processes, jets, which are originating from b quarks, are important since a number of these processes involve top quark decays. However, in the present analysis, the targeted SUSY model T5qqqqWW (see Sec.1.2.3), involves no top

quark; therefore in the final state no jets coming from b quarks is expected. In the analysis, the events including b quark tagged jets are used to perform the $t\bar{t}$ + jets background estimation (see Sec.5.3).

Thus, the identification of jets coming from b-quark hadronization, which is also referred as b-tagging, is an important task. Generally, the identification technics [101] utilize the rather long lifetime and high mass of the b quark which leads to a displaced vertex formation.

A variety of algorithms have been developed by CMS [102] to select b-tagged jets based on the variables such as the impact parameters of charged-particle tracks, the assets of reconstructed decay vertices, and the presence or absence of a lepton or combination of some of them. Each of these algorithms results in a single discriminator value for each jet. In this analysis, the combined secondary vertex (CSV) algorithm is employed with medium working point (0.8484), which corresponds to a tagging efficiency of about 70% and a misidentification rate of about 1% accordingly to the p_T and η of the jet.

The possible differences between the performance of the b-tagging algorithm measured in data and in simulation are compensated by applying scale factors to simulated events, the basic principles are mentioned in Sec.4.2.1.

3.2.5 Missing transverse energy

One of the main purposes of the present search is to investigate the potential discovery of SUSY models, which involve dark matter candidates such as neutralinos in the form of the lightest stable supersymmetric particle. However, these particles as well as the weakly interacting SM neutrinos are invisible in the CMS detector in other words they dont leave any trace in the subsystem. Only imbalance in energy and momentum can be a hint for their existence. The momentum imbalance in the transfers plane is called missing transverse energy and its vector form is denoted as \vec{E}_T while the magnitude is E_T . Therefore, \vec{E}_T can be formulated as the negative vector sum of the transverse momentum p_T of all observed particles:

$$\vec{E}_T = - \sum_i \vec{p}_{T,i}, \quad (3.3)$$

where the index i iterates over all particle flow candidates, which are associated with the primary vertex, to calculate the particle flow E_T , $PF - E_T$. Otherwise, the case where i stands for the calorimeter deposits is a another E_T type used in CMS which is known as $Calo - E_T$. A third type of E_T is an improved version of $Calo - E_T$ by corrections obtained from the tracking detector and it is called $TC - E_T$. Naturally, as it can be understood from the name, $PF - E_T$ is reconstructed by using the particle flow algorithm [106].

In the CMS detector, a momentum imbalance can be also originated through the various subdetector malfunctions, reconstruction effects.

The mis-calculation of \cancel{E}_T due to the possible reconstruction problems of PF particles can be reduced by propagating the aforementioned jet energy corrections to the \cancel{E}_T calculation. In CMS, particles generally are produced uniformly in ϕ , thus $\vec{\cancel{E}}_T$ is expected to be independent of ϕ . However, due to the ϕ -dependence of the detector response, imperfect alignment of different detector subsystems, or a ~ 4 mm shift between the centre of the detector and beamline [107], an asymmetry in ϕ is observed in data and in simulated events. The observed ϕ -asymmetry can also be explained by a shift in the $\vec{\cancel{E}}_T$ components along the x and y detector axes, which are denoted as \cancel{E}_x and \cancel{E}_y respectively. The shift shows a linearly increasing trend as a function of multiplicity of the PF candidates. To obtain the corrections, this correlation is used. First, linear fits are performed on \cancel{E}_x and \cancel{E}_y distributions as a function of number of PF candidates in various η bins. Examples of these fits are shown in Fig. 3.2, where the linear dependence of $\langle \cancel{E}_x \rangle$ and $\langle \cancel{E}_y \rangle$ on multiplicity of PF candidates can be formulated as:

$$\begin{aligned}\langle \cancel{E}_x \rangle &= c_{x_0} \cdot x + c_{x_s} \cdot x^2, \\ \langle \cancel{E}_y \rangle &= c_{y_0} \cdot x + c_{y_s} \cdot x^2.\end{aligned}\tag{3.4}$$

After the fits are performed, the corrected \cancel{E}_x and \cancel{E}_y can be obtained on an event-by-event basis as:

$$\begin{aligned}\cancel{E}_x^{\text{corr}} &= \cancel{E}_x - \langle \cancel{E}_x \rangle = \cancel{E}_x - (c_{x_0} \cdot x + c_{x_s} \cdot x^2), \\ \cancel{E}_y^{\text{corr}} &= \cancel{E}_y - \langle \cancel{E}_y \rangle = \cancel{E}_y - (c_{y_0} \cdot x + c_{y_s} \cdot x^2).\end{aligned}\tag{3.5}$$

The coefficients c_{x_0} , c_{x_s} , c_{y_0} , and c_{y_s} are determined separately from data and simulated samples. In simulated samples, the corrections are obtained by using DY events where a Z boson decays to two opposite charged leptons. In these kind of events the expected *real* – \cancel{E}_T contribution is very low, thus they provide a very good environment to measure \cancel{E}_T mismodelings. The validation of the procedure is performed by using $t\bar{t}$ + jets and W + jets simulated samples. In data, the corrections are obtained in events with the invariant mass of the dimuon or dielectron system is inside the Z-boson mass window, $60\text{GeV} < M_{\ell\ell} < 120\text{GeV}$.

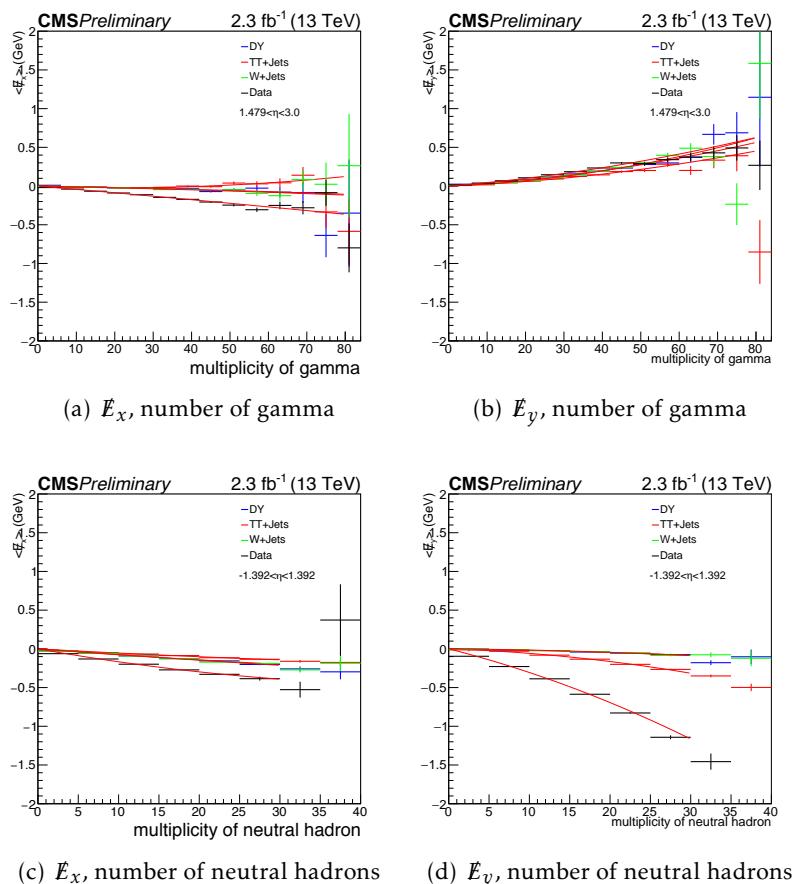


Figure 3.2: PF candidate multiplicity fits for \not{E}_x and \not{E}_y in different η regions.

Chapter 4

Event Selection: Baseline and Search regions

Contents

4.1	SUSY signature	40
4.1.1	Key variables	42
4.1.2	Signal samples	43
4.2	Background processes	43
4.2.1	Scale Factors	44
4.3	Baseline selection	46
4.4	Data Samples	48
4.4.1	Triggers	48
4.5	Event cleaning: Filters	49
4.6	Control plots	53
4.7	Search regions	53
4.7.1	Signal and control region	53
4.7.2	Main band regions	54

4.1 SUSY signature

In this thesis, as mentioned in section 1.2.3, a search for SUSY is performed in the context of the simplified model T5qqqqWW (see Figure 4.1)¹. In this model, each gluino decays to two light quarks (c,s,u,d) and an intermediate chargino, with the latter decaying to a W boson and a neutralino.

In the event selection, one of the W 's is chosen to be decaying leptonically. The choice of single lepton final states provides a cleaner event topology than the full hadronic final states while keeping the signal efficiency sufficiently high. Therefore, the visible final

¹Throughout this thesis, the T5qqqqWW model is sometimes referred as the signal model.

state includes at least six jets and one lepton (can be either electron or muon). Moreover, it is required that none of the jets are tagged as b quark to increase the probability of selecting the events with topology T5qqqqWW. Furthermore, in such events, there is a momentum in balance originated by the neutralino and the neutrino which leave no trace in the detector. Therefore, a high missing transverse energy (MET) is also required. These selections can be further enhanced with the characteristics of individual objects such as; the transverse momentum of lepton $p_{T,\ell}$ or jets $p_{T,jets}$. However, only selecting events with one lepton, many jets, and high MET or using object characteristics is not enough to reveal the existence of SUSY-like events which are rare among the many SM like events with the similar final state. The $W +$ jets and $t\bar{t} +$ jets events are the leading SM processes which can mimic the T5qqqqWW signature. Hence, some advanced kinematic variables need to be invented to eliminate these SM background processes. The table 4.1 shows the list of variables which are used in this analysis. This list does not include the variables that are used in the identification of particle flow objects (see Sec. 3.1). The physics variables are categorised according to their mathematical complexity.

Simple variables	Advanced variables
L_T, H_T	$\Delta\Phi(W, \ell)$
$p_{T,\ell}, \eta_\ell$	$M_{T2}(\vec{p}_T^\ell, \vec{p}_T^t, \vec{p}_T^{miss})$
$p_{T,jets(1,2)}$	
n_{jets}	
n_{b-tag}	

Table 4.1: List of kinematic variables

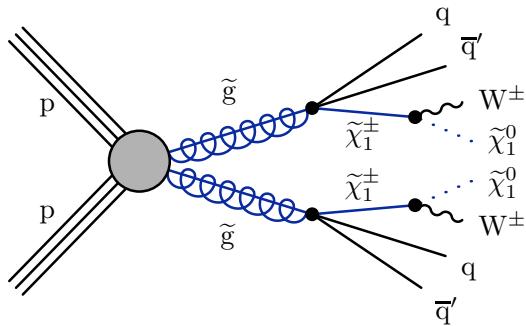


Figure 4.1: Diagram showing the simplified model T5qqqqWW.

4.1.1 Key variables

The list of simple variables consists of fundamental properties of the reconstructed physics objects, such as the transverse momentum p_T of these objects, and their linear combinations.

L_T is the scalar sum of the charged lepton p_T and the missing transverse momentum E_T^{miss} . Given no additional source of E_T^{miss} , except the SM neutrino, the variable L_T can be also written as: $\sqrt{p_{T(W)}^2 + M_{T(W)}^2}$. The derivation of this relation is as follows:

$$\begin{aligned} p_{T(W)}^2 &= (p_{T(\ell)} \cdot \cos\phi_\ell + E_T \cdot \cos\phi_{E_T})^2 + (p_{T(\ell)} \cdot \sin\phi_\ell + E_T \cdot \sin\phi_{E_T})^2, \\ M_{T(W)}^2 &= 2 \cdot p_{T(\ell)} \cdot E_T (1 - (\cos\phi_{E_T} \cdot \cos\phi_\ell + \sin\phi_{E_T} \cdot \sin\phi_\ell)), \\ p_{T(W)}^2 + M_{T(W)}^2 &= p_{T(\ell)}^2 + E_T^2 + 2 \cdot p_{T(\ell)} \cdot E_T (\cos\phi_{E_T} \cdot \cos\phi_\ell + \sin\phi_{E_T} \cdot \sin\phi_\ell) + 2 \cdot p_{T(\ell)} \cdot E_T, \\ (p_{T(\ell)} + E_T)^2 &= L_T^2 \end{aligned} \quad (4.1)$$

For events with a single highly boosted W boson ($p_{T(W)} \gg M_{T(W)}$), $L_T \sim p_{T(W)}$. This variable is also known as “leptonic mass scale” of the event.

H_T is the scalar sum of the transverse momenta of all jets above a p_T threshold in the event. The major contribution to this sum is coming from transverse energy of the jets originated from the gluino decay, and it is related to the mass gap between the gluino and the chargino. This variable is reflecting the “hadronic mass scale” of the event.

n_{jets} is the number of all jets above a p_T threshold, same as in the H_T calculation, in the event.

n_{b-tag} is the number of all jets tagged as coming from a b quark above a p_T threshold, same as with other jets, in the event.

The other variables in the list such as p_T, ℓ or η_ℓ was discussed in Sec. 3.1 regarding particle flow objects.

There are two advanced variables used in this analysis.

$\Delta\Phi(W, \ell)$ is the azimuthal angle between the W boson and the charged lepton. In the $W + \text{jets}$ and $t\bar{t} + \text{jets}$ events, the lepton comes from a leptonic decay of W boson, $W \rightarrow \ell\nu$. Therefore, $\Delta\Phi(W, \ell)$ strongly related to the mass of the W boson and its momentum. In SM events with high missing energy indicates that the W bosons yielding the lepton and the neutrino are boosted, hence resulting in a narrow distribution in $\Delta\Phi(W, \ell)$. On the other side, in SUSY decays, the missing transverse energy comes from two neutralinos and the neutrino, which randomizes the $\Delta\Phi(W, \ell)$, thus resulting in an even distribution in $\Delta\Phi(W, \ell)$. As a result of its features, the variable $\Delta\Phi(W, \ell)$, is the most significant variable in this analysis. The high values of $\Delta\Phi(W, \ell)$ is used as a signal region² while the low val-

²where signal to background event counts are sufficiently large.

ues of $\Delta\Phi(W, \ell)$ is the control region³.

M_{T2} is defined as:

$$M_{T2}(\vec{p}_T^\ell, \vec{p}_T^t, \vec{p}_T^{miss}) = \min_{\vec{p}_T^{(1)} + \vec{p}_T^{(2)} = \vec{p}_T^{miss}} \left\{ \max \left[M_T(\vec{p}_T^\ell, \vec{p}_T^{(1)}), M_T(\vec{p}_T^t, \vec{p}_T^{(2)}) \right] \right\}, \quad (4.2)$$

where M_T is the transverse mass and the indices t, ℓ represent isolated track and the selected lepton respectively. In this analysis, this variable is used only in the baseline selection to reduce the dileptonic $t\bar{t} + \text{jets}$ contribution which is more important in the signal region (high $\Delta\Phi(W, \ell)$).

4.1.2 Signal samples

The simplified model T5qqqqWW is discussed in section 1.2.3. This model originally includes three free parameters: the masses of the gluino $m_{\tilde{g}}$, the intermediate chargino $m_{\tilde{\chi}_1^\pm}$ and the neutralino $m_{\tilde{\chi}_1^0}$. To reduce the three dimension mass space the chargino mass is fixed to midway between gluino and neutralino mass. Hence, this model is investigated in the $m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ mass plane. For each mass point, a separate Monte Carlo sample is produced. The mass plane is shown in Fig. 4.2 where the z-axis represents the total number of events generated. This scan of parameter space includes 657 mass points and it requires more computational power than usual SM process production therefore fastsim is used as explained in section ???. The MADGRAPHv5 event generator is used for signal events modelling. The samples are normalized to the cross section presented by the LHC SUSY Cross Section Working Group [39]. Throughout present thesis, two mass points corresponding to different gluino and neutralino masses are used as benchmarks to study the kinematic properties of the signal.

T5qqqqWW(1.9,0.1) represents a point in the high mass gap region with $m_{\tilde{g}} = 1900$ GeV and $m_{\tilde{\chi}_1^0} = 100$ GeV. These signal events have high H_T and L_T .

T5qqqqWW(1.5,1.0) represents a point close to the compressed region with $m_{\tilde{g}} = 1500$ GeV and $m_{\tilde{\chi}_1^0} = 1000$ GeV. The compressed region is where the $m_{\tilde{g}} - m_{\tilde{\chi}_1^0} \leq 2m_W$. These signal events have lower H_T and L_T with respect to high mass gap ones.

4.2 Background processes

As mentioned earlier in this section the most important background processes are $W + \text{jets}$ and $t\bar{t} + \text{jets}$ events. Requiring zero b-tagged jets in the event selection, $W + \text{jets}$ events becomes the main background. All of the background processes considered in this analysis are listed below. PYTHIA 8 is used for parton shower and hadronisation.

³where signal event counts are negligible with respect to background event counts.

$t\bar{t}$ + jets is generated with `MADGRAPH5_AMC@NLOv2.2.2`. To enhance the statistics the combination of H_T binned samples, dedicated semi- and dilepton samples is used.

W + jets is generated with `MADGRAPH5_AMC@NLOv2.2.2`. H_T binned samples, where W decaying leptonically, are used.

QCD multijets events are produced with `MADGRAPH5_AMC@NLOv2.2.2` in bins of H_T .

t/\bar{t} samples are produced with `POWHEGv2.0` except the t decaying leptonically with s-channel is produced with NLO `MADGRAPH5_AMC@NLOv2.2.2`.

Drell-Yan events are produced with `MADGRAPH5_AMC@NLOv2.2.2` and $m_{\ell\ell} > 50\text{GeV}$ samples are used in H_T bins.

Di-boson samples are produced with `AMC@NLO` except WW samples are produced with `powheg`.

$t\bar{t}V$ samples are produced with `AMC@NLO`.

A list of all background MC samples with cross sections can be found in Appendix A.

4.2.1 Scale Factors

The modeling of physics processes and the simulation of the detector responses are not perfect. Studies of the data-MC discrepancies indicate that MC needs to be corrected by some event-by-event weights, which are called scale factors. The scale factors used in this analysis are as follows:

Lepton identification and reconstruction efficiency:

MC samples have to be corrected according to the reconstruction and identification effi-

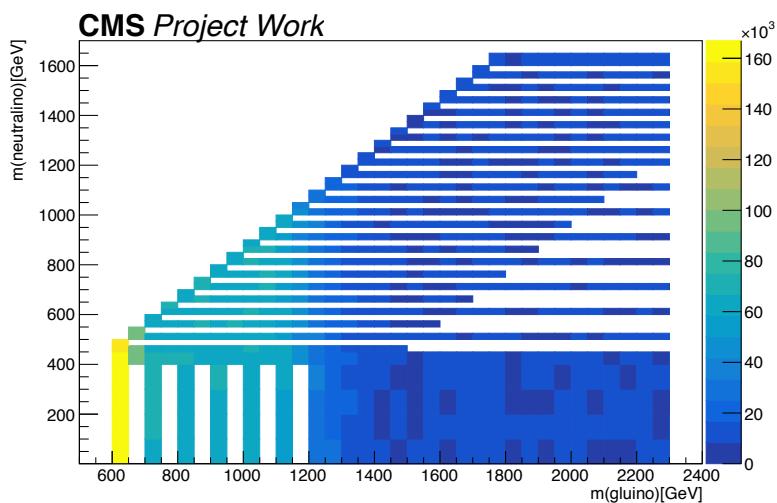


Figure 4.2: Produced signal mass points for the simplified T5qqqqWW model.

ciencies of leptons. In this work, scale factors are applied as a function of p_T and η of the selected lepton. They are calculated for electrons and muons separately by the e-gamma and muon physics object groups respectively[45].

B-tagging efficiency:

The simulated b-tagging efficiency is corrected with respect to the one in data. For each simulated event, weights corresponding to the different number of b-tagged jets values are calculated and this results in multiple weights for each event[45]. This method allows reusing the events. The advantage of this method is that it provides to use full statistical power of the MC events.

ISR reweighting:

Since the 8 TeV Run of LHC, it is known that the p_T spectrum of $t\bar{t}$ events are not well modeled. Therefore, an event-by-event correction is obtained using the jets from initial state radiation (ISR). In the present analysis, each event is corrected according to number of ISR jets. The weights can be seen in table 4.2 where the D factor is for keeping the normalization of the overall sample invariant. This D factor is calculated using an inclusive sample i.e. without the specific selections of the analysis.

Table 4.2: Weights based on the number of ISR jets as given in Ref. [47]

nISR jet	Normalisation weight $D_{t\bar{t}} = 1.071$
0	
1	$D \times (0.920 \pm 0.005 \pm 0.040)$
2	$D \times (0.821 \pm 0.006 \pm 0.090)$
3	$D \times (0.715 \pm 0.009 \pm 0.143)$
4	$D \times (0.662 \pm 0.016 \pm 0.169)$
5	$D \times (0.561 \pm 0.027 \pm 0.219)$
≥ 6	$D \times (0.511 \pm 0.041 \pm 0.244)$

Pileup:

As discussed earlier in Section ??, the number of pile-up interactions in simulated samples is generated with a prior distribution. Thus, the MC pile-up distribution may vary from the actual pile-up distribution in Data. The simulated distribution is rescaled with the Data/MC ratio in Figure 4.3.

4.3 Baseline selection

Earlier in this section 4.1, the SUSY signature considered in this work is introduced. Moreover, a suitable primary event selection for removing the SM background as much as possible while keeping the signal efficiency high was discussed. In the following, details of this event selection are explained.

Selection of leptons:

The selected lepton, which can be an electron or a muon, is required to have a minimum p_T of 25 GeV while at the same time it satisfies the good lepton criteria introduced in Sec. 3.1. Additionally, all the other electrons and muons with p_T greater than 10 GeV are vetoed. These veto leptons satisfy the loose lepton selection with $I_{rel} < 0.4$.

Selection of jets:

The reconstruction of jets is already introduced in Sec. 3.1. In the present analysis, jets are required to have $p_T > 30$ GeV and $|\eta| < 2.4$. In order to avoid double counting of objects, jets that are close ($\Delta R < 0.4$) to either a veto or selected lepton are removed. As discussed earlier in this chapter the T5qqqqWW is expected to have at least five jets. Moreover, the mass difference of the gluino and the intermediate chargino affects the p_T of the jets. Therefore, it is required that two highest p_T jets satisfy the $p_T > 80$ GeV condition. Addi-

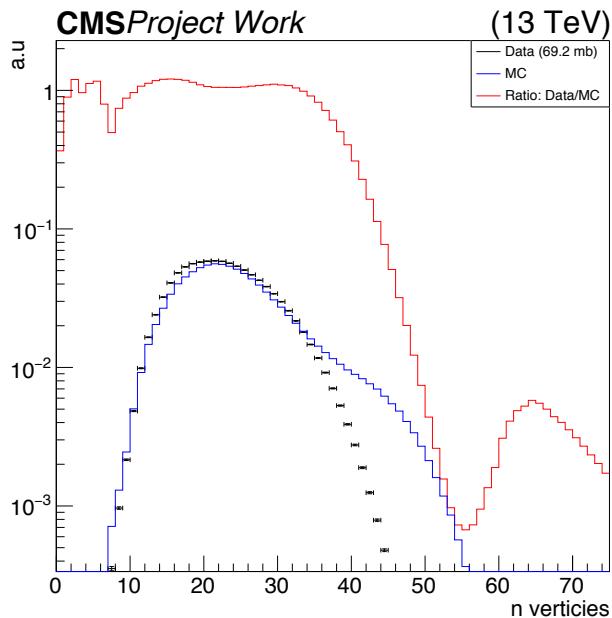


Figure 4.3: The normalised distributions of mean number of interactions per bunch crossing during the analysed data Run2016 (black) and in the MC samples (blue). And the red distribution represents the Data to MC ratio. For the data distribution the latest luminosity calibrations [48] and an inelastic pp cross-section of 69.2 mb are used.

tionally, in the baseline selection of the analysis, the number of jets tagged as coming from b quarks is zero to suppress the events containing top quarks.

Energy scales thresholds:

The signal model considered in this work favors events with high hadronic and leptonic scale. The hadronic scale is chosen to be at least 500 GeV while the leptonic scale threshold is 250 GeV.

Isolated track veto:

The isolated track veto is designed to suppress $t\bar{t}$ events in which both W bosons decay leptonically and one lepton does not satisfy the selection criteria for veto leptons. In the mechanism of this veto, the M_{T2} variable, which is introduced in section 4.1.1, is used with an isolated track \vec{t} , a lepton $\vec{\ell}$ and the missing transverse momentum \vec{p}_T^{miss} . In the calculation of M_{T2} , it is assumed that the missing energy source is the two neutrinos from the dileptonic $t\bar{t}$ decay. The minimization runs over all possible splitting of \vec{p}_T^{miss} . The figure 4.4 shows the M_{T2} distribution separate for hadronic and leptonic tracks after the baseline selection with $H_T > 500$ GeV, $L_T > 250$ GeV, $n_{jet} \geq 5$ and without b-tag requirement. The distribution of M_{T2} is slightly different for the two cases and a lower M_{T2} cut of 60 GeV and 80 GeV is applied for hadronic and leptonic tracks respectively. The rightmost plot in the figure 4.4 shows the M_{T2} distribution for all events i.e. including events that do not have any isolated track at all. In this case events without any isolated track are added to the overflow bin. It can also be derived from the plot that only about 20% of the signal events have an opposite charged isolated track at all compared to 40% of the dileptonic $t\bar{t}$ events.

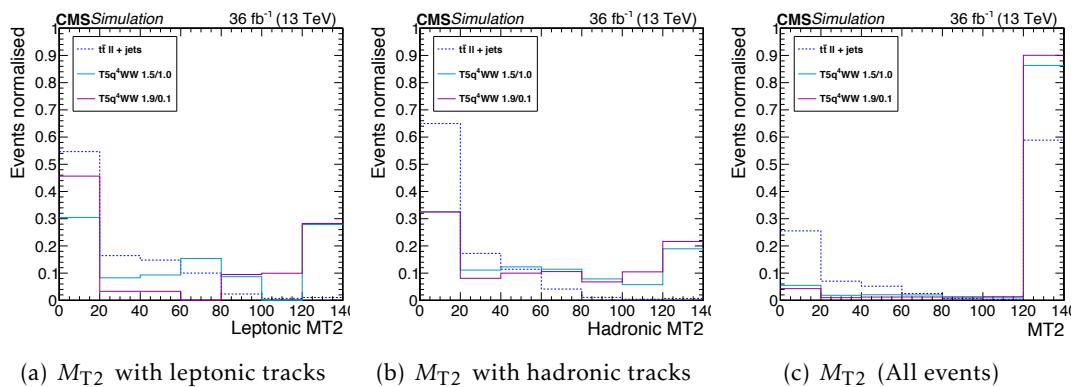


Figure 4.4: Distributions of M_{T2} for events with electron or muon veto tracks (a) and hadronic veto tracks (b), for dileptonic $t\bar{t}$ (blue dotted) and T5qqqqWW (purple and azure) signal samples. The highest bin is always an overflow bin. The majority $t\bar{t}$ events have $M_{T2} < 80$ GeV, while the signal events have longer tails in M_{T2} . The rightmost plot (c) shows the distribution for all events.

The effect of each baseline requirement is demonstrated in Figure 4.5 for the different background processes (stacked) and for two signal benchmark points. It should be noted that the background samples include an initial skim of $H_T > 350$ GeV and $L_T > 150$ GeV to shorten the computation time. It can be observed that the design of baseline selections works such that they reduced the background events significantly while only a small fraction of signal events are eliminated by the selections. It can be also noted that, naturally, the QCD background is almost vanished after the single muon selection. The estimation of important backgrounds will be explained in the next chapter.

Selection	Object definitions
Single lepton	Tight leptons, $p_T \geq 25$ GeV and $ \eta < 2.4$ and $I_{mini} < 0.1(0.2)$ for electrons(muons)
Lepton veto	Loose leptons, $p_T \geq 10$ GeV and $ \eta < 2.4$ and $I_{mini} < 0.4$
Isolated track veto	$I_{rel} < 0.3$, $\Delta R(\ell, track) < 0.1$, $\text{charge}_{track} = -\text{charge}_{lepton}$
$n_{jets} \geq 5$	Good jets with $p_T \geq 30$ GeV and $ \eta < 2.4$
$p_{T,jets(1,2)} \geq 80$ GeV	and cleaned from close leptons
$H_T \geq 500$ GeV	$\sum_{jets} p_T$
$L_T \geq 250$ GeV	$E_T + p_T^{lep}$
$n_{b-tag} = 0$	b-tagged good jets with CSVv2 Medium working point (0.8484)

Table 4.3: List of event selection criteria and object requirements.

4.4 Data Samples

The data used for this analysis is recorded by the CMS detector during the 2016 LHC run at 13 TeV. As in Figure ??, The total integrated luminosity collected by CMS is 37.76 fb^{-1} . In the present analysis, the data validated by the CMS Data Quality Monitoring (CMS-DQM) certification team, which correspond to an integrated luminosity of $L=35.9 \text{ fb}^{-1}$, is used.

4.4.1 Triggers

In this analysis, the main trigger set is a combination of triggers containing an isolated single lepton, electron or muon, with p_T of 15 GeV and HT of 350/400 GeV. To recover inefficiencies due to an update on the online electron ID and the saturated L1-jets, the

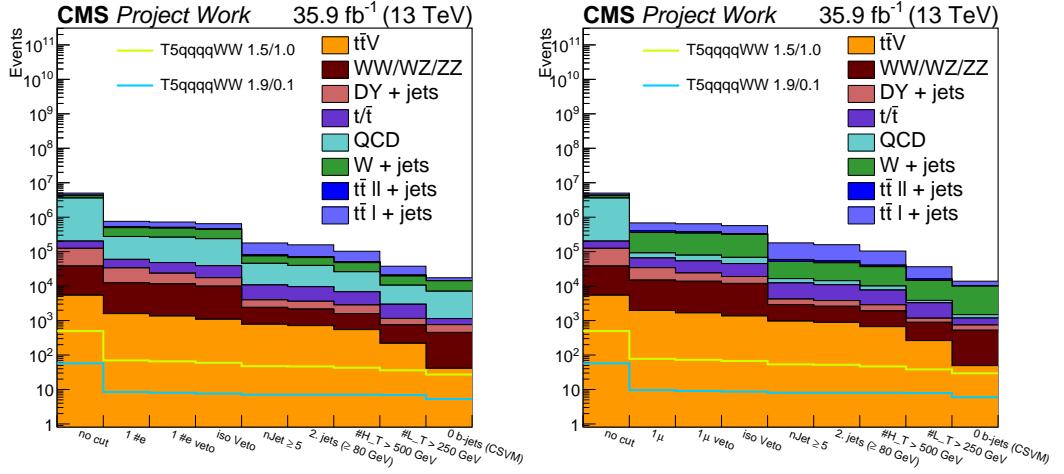


Figure 4.5: The number of events for each sample for the electron (left) and muon (right) channel. Samples are scaled to the luminosity with their cross section and scale factors discussed in section 4.2.1.

trigger strategy had to be extended: The list of trigger paths can be seen in Table 4.4. Events recorded with these trigger paths are allocated in three primary datasets (PDs): the **SingleElectron**, **SingleMuon** and **MET** datasets. Therefore, when successively adding PDs, triggers contained in previous PDs are vetoed to avoid double counting of events contained in more than one of the PDs. The trigger efficiency calculation is shown in the Equation 4.3. The trigger efficiencies are measured as a function of L_T , H_T and lepton p_T and can be seen in Figures 4.6, 4.7, 4.8 respectively. The efficiency distribution in the baseline region ($L_T > 250$ GeV; $H_T > 500$ GeV; lepton $p_T > 25$ GeV) is flat and its value close to 100%(98%) for the muons (electrons).

$$\epsilon = \frac{N(\text{all events passing probed trigger(s) + preselection + reference trigger})}{N(\text{all events passing preselection + reference trigger})} \quad (4.3)$$

4.5 Event cleaning: Filters

The event reconstruction can fail due to the noisy detector cells and other kinds of detector problems. These can result in incorrectly reconstructed physics objects such as: muons, jets and hence MET. In CMS, the physics object groups related to the problem release event filters or cures for the falsely reconstructed objects.

Primary vertex filter:

A good primary vertex is required which is introduced in section 3.2.1.

Beam halo filter:

Single Electron Dataset
HLT_Ele105_CaloIdVT_GsfTrkIdT_v
HLT_Ele115_CaloIdVT_GsfTrkIdT_v
HLT_Ele50_CaloIdVT_GsfTrkIdT_PFJet165_v
HLT_Ele27_WPTight_Gsf_v
HLT_Ele15_IsoVVVL_PFHT350_v
HLT_Ele15_IsoVVVL_PFHT400_v
Single Muon dataset
HLT_Mu50_v
HLT_IsoMu24_v
HLT_IsoTkMu24_v
HLT_Mu15_IsoVVVL_PFHT350_v
HLT_Mu15_IsoVVVL_PFHT400_v
MET dataset
HLT_PFMET100_PFMHT100_IDTight_ OR HLT_PFMETNoMu100_PFMHTNoMu100_IDTight_
HLT_PFMET110_PFMHT110_IDTight_ OR HLT_PFMETNoMu110_PFMHTNoMu110_IDTight_
HLT_PFMET120_PFMHT120_IDTight_ OR HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_

Table 4.4: List of HLT paths

The collisions of the beam with residual gas inside LHC cause showers of secondary particles. Additionally to these scattering effects, charged particles are deflected by the magnetic field of the beam optics. These particles are called beam halo particles and one of the main sources of the beam background of the LHC. In CMS, the beam halo algorithm considers the particles produced outside the CMS cavern and to detect events with beam halo it uses the timing information and hit topology in CSC, ECAL and HCAL subsystems.

HB-HE noise filter:

This noise is originated from the Hybrid PhotoDiodes and Readout Boxes of the HCAL. The timing, pulse shape as well as the other readout errors are used to detect the noise.

ECAL dead cell trigger primitive filter:

The existence of noisy crystals in ECAL can lead to fake MET. The events, in which the noisy cells deposit high energy, are filtered.

Bad PF muon filter:

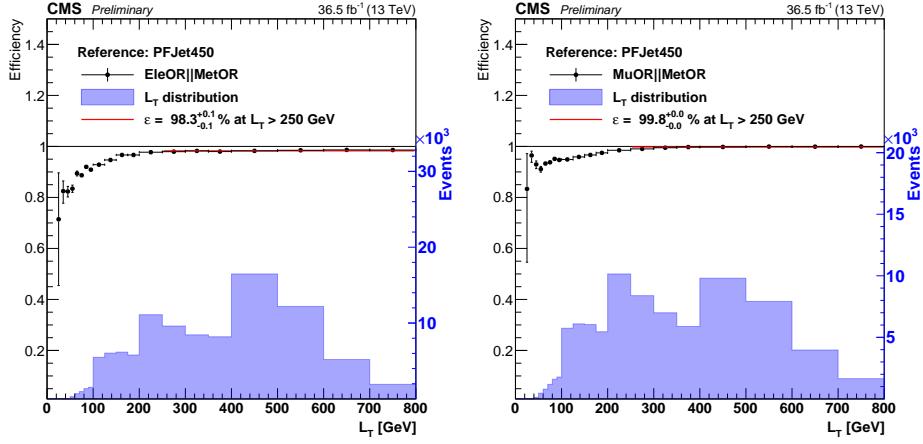


Figure 4.6: Measurement of the trigger efficiency as a function of L_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) on the right [46].

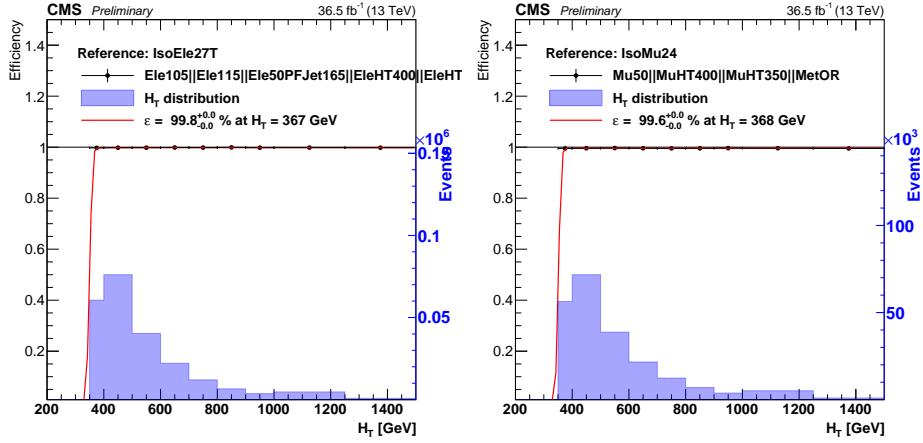


Figure 4.7: Measurement of the trigger efficiency as a function of H_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) on the right [46].

This filter fires when there is a PF muon with too low quality and has large p_T . The quality of the muon is determined according to its tracking uncertainty, segment compatibility and other detector related features. This bad muon is required to have $p_T > 100$ GeV. Unlike the other filters explained above this filter is applied to background simulation samples as recommended.

Bad charged hadron filter:

The events, where there is a muon but it fails to be a PF muon and it still contributes to

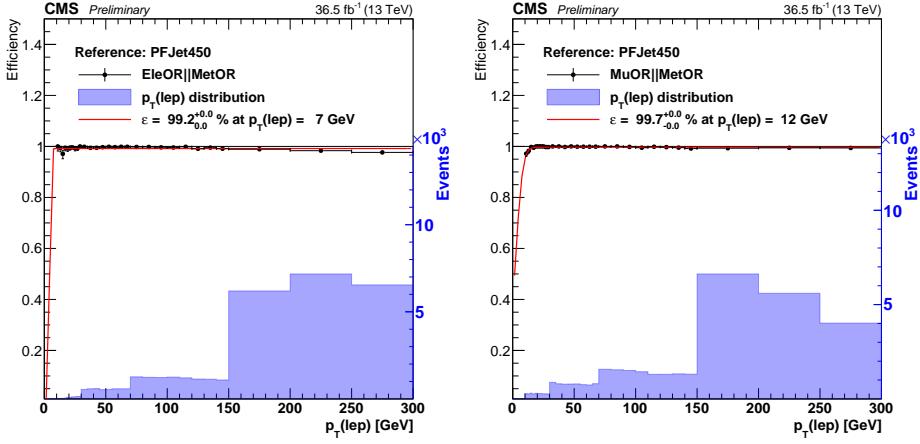


Figure 4.8: Measurement of the trigger efficiency as a function of lepton p_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) is shown on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) is shown on the right. The lepton p_T trigger efficiency is measured after an $L_T > 250$ GeV cut which is the analysis baseline selection [46].

PF MET calculation as a charged hadron candidate, are filtered. Here, as in the previous case, this muon is required to have $p_T > 100$ GeV. Moreover, it is required that the muon and the charged hadron traces are almost overlapping ($\Delta R(\mu, \text{charged hadron}) < 0.00001$). In addition to this, their p_T s are needed to be very close to each other. Similarly to the bad PF muon filter, this is applied to background MC samples as well as the data events.

Duplicate muon filters:

In the Re-reco data it is observed that there is an increase in events in the MET tail of $Z \rightarrow \mu\mu$ data. It is understood by the muon POG that there are duplicate muons in the events due to reconstruction failures. Additionally, two filters are recommended by the SUSY group: One is to remove events with the ratio of PF MET to calorimeter MET is more than 5. The second is to remove events containing bad jets which have $p_T > 200$ GeV, muon energy fraction > 0.5 and $\Delta\phi(\text{met}, \text{jet}) > \pi - 4$.

Filter on fastsim:

This filter is to clean up bad jets from the fastsim events. To remember, in this work, fastsim is used for the signal MC production. The event is removed if there is a jet satisfying the following conditions: $p_T > 20$ GeV, $|\eta| < 2.5$, unmatched to a generated jet ($\Delta R(\text{jet}, \text{gen jet}) > 0.3$), and charged hadron fraction < 0.1 .

The efficiency of all these filters after the analysis baseline selection is approximately 98%.

4.6 Control plots

The distributions of main kinematic variables, also called control plots, are shown in this section. In all the control plots, the colored lines represent the signal models and the color filled stacked histograms display the background processes. Additionally, the black dotted distribution exhibits the observed data points requiring the triggers introduced in section 4.4.1. In all the control plots, the events are cleaned by the filters discussed in section 4.5. The signal and background events are scaled by the luminosity factor and additional weights introduced in section 4.2.1. In the Figure 4.9 top left plot shows the n-btag distribution after the baseline selection (see section 4.3). The distribution of simulated signal events peaks at zero as expected. Clearly, one can see from this distribution that choosing events with zero b-tagged jets suppresses the $t\bar{t}$ background significantly. The other distributions, displaying main kinematic variables, exhibit reasonable MC-data agreement although it is not necessarily expected. The plot shows the number of jet distribution (top/right), H_T (bottom/left) and L_T (bottom/right) are shown.

The multiplicities of jets and b-tagged jets display no difference between the signal scenarios, since the mass splitting has no effect on the decay topology. The two selected signal benchmark models show differences in the distributions of H_T and L_T , this is due to the gluino-neutralino mass splitting. In the non-compressed T5qqqqWW (1900,100) model, the quarks coming from the gluino decay have a large boost, resulting in high leptonic and hadronic energy scales. For the compressed region with T5qqqqWW (1500,1000) no such effect is observed, and the shape of the distributions look similar to the SM processes.

Additional control plots are shown in Appendix B.

4.7 Search regions

4.7.1 Signal and control region

Signal region is where it is expected to observe a deviation from SM background in the existence of SUSY scenario. In order to locate this kinematic region a powerful discriminative variable is used. Figure 4.10 shows the $\Delta\phi$ distribution after the baseline selection. The plot exhibits that the SM background processes peak at zero while the signal models stay flat. Naturally, this shape difference leads to the high values of $\Delta\phi$ as the signal region (SR) and the low values as the control region (CR). The control region is used in the background estimation, which will be discussed in the next chapter. It can also be seen from the distribution that after the baseline selection $W + \text{jets}$ and $t\bar{t} + \text{jets}$ are the main background components. The $t\bar{t}$ composition in the CR is dominated by the single leptonic decays while in the SR the di-leptonic $t\bar{t}$ decays are dominating.

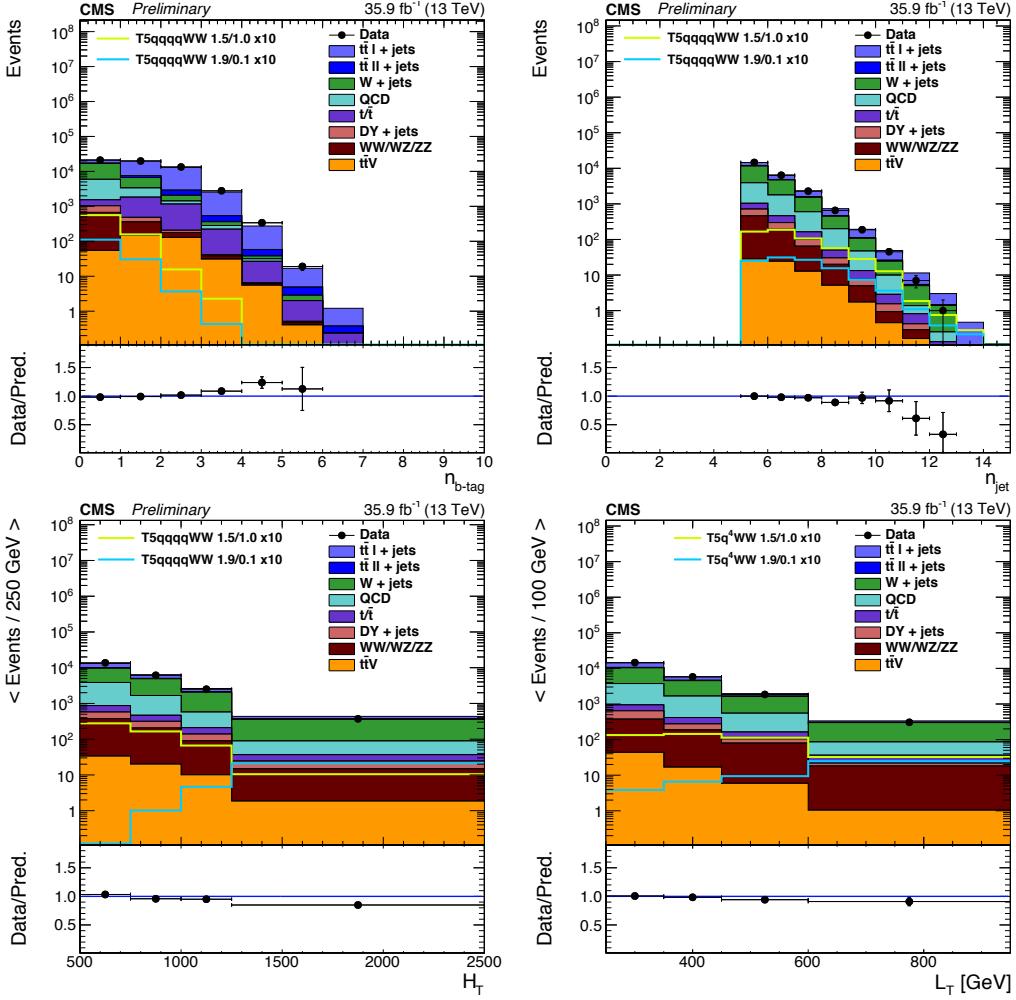


Figure 4.9: The top left distribution shows the number of b-tagged jets, after the baseline selection, requiring at least five jets, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton with $p_T > 25$ GeV. The rest of the distributions are plotted after the same baseline selection additionally requiring no b-tagged jets. In the top row the number of jets (right) while in the bottom row H_T (left) and L_T (right) distributions are shown. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked. The model T5qqqqWW (1.5,1.0) (T5qqqqWW (1.9,0.1)) corresponds to a gluino mass of 1.5 TeV (1.9 TeV) and neutralino mass of 1.1 TeV (0.1 TeV), respectively. The intermediate chargino mass is fixed at 1.25 TeV (1.0 TeV). The two benchmark signal models are scaled up by a factor of 10.

4.7.2 Main band regions

To further enhance the sensitivity, the phase space is subdivided in bins of n_{jet} , H_T and L_T . These sub bins are called main band regions (MB).

The binning is designed for an integrated luminosity of 40 fb^{-1} scenario. In the optimiza-

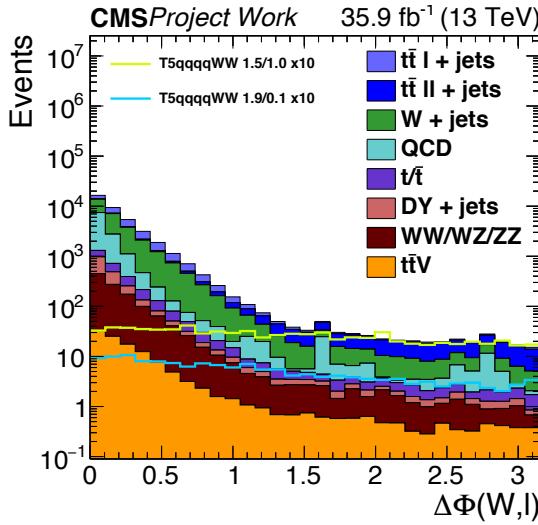


Figure 4.10: The $\Delta\phi$, after the baseline selection, requiring at least five jets and none of b-tagged, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton (electron or muon) with $p_T > 25$ GeV. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked. The model T5qqqqWW (1.5,1.0) (T5qqqqWW (1.9,0.1)) corresponds to a gluino mass of 1.5 TeV (1.9 TeV) and neutralino mass of 1.1 TeV (0.1 TeV), respectively. The intermediate chargino mass is fixed at 1.25 TeV (1.0 TeV).

tion, the bins with the highest figure of merit are chosen among a spectrum of bins. Then, a fully profiled likelihood is performed for the several combinations; The bin combination with the highest sensitivity is chosen. At the same time, for each bin, the yield of background MC is required to be at least one, to have robust background estimation. In addition to these, throughout the bins, boundary cuts are chosen to be symmetric if possible and sensible. For instance, the binning in L_T is same for different n_{jet} bins.

The $\Delta\phi$ and L_T variables are constructed by using the same objects; p_T^{lep} and E_T . Therefore, a correlation can be assumed. As in Figure 4.11, for increasing L_T the SM events are located towards the low values of $\Delta\phi$. On the other hand, the signal plots do not show the same trend. Therefore, a varying $\Delta\phi$ cut, depending on the L_T binning, is used to determine the signal region for each MB.

The resultant binning is given in table 4.5 and the corresponding simulated background and signal yields can be found in Figures 4.13,4.14. Figures manifest that the di-leptonic decay of $t\bar{t}+jets$ events are favoured in SR (high $\Delta\phi$). Furthermore, It can also be extracted that simulated QCD multijet events are populated in the CR (low $\Delta\phi$) in electron channel. As mentioned, the $W+jets$ and $t\bar{t}+jets$ are dominating the SM background. In this analysis, the number of the events in the MB SR is estimated with observed data. In the next chap-

ter, the procedure for the estimation of these background components will be explained.

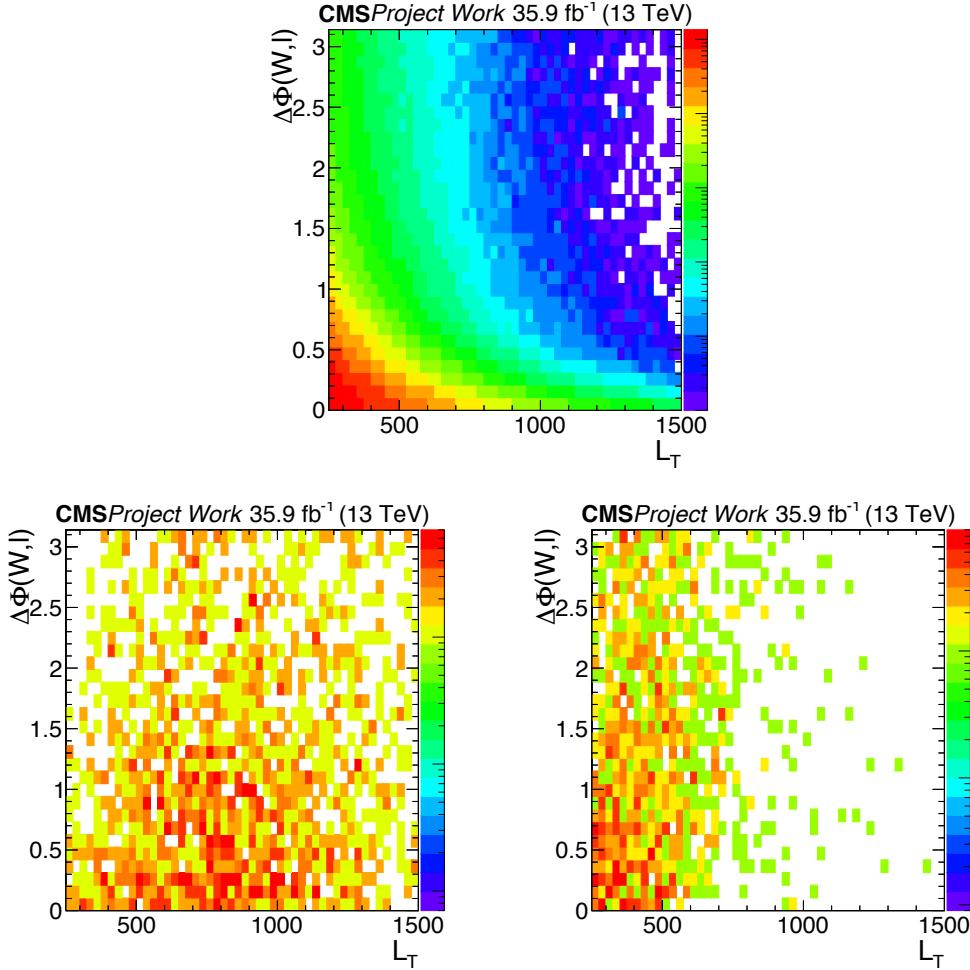


Figure 4.11: 2D distributions of event counts for the main background samples ($t\bar{t} + \text{jets} + W + \text{jets}$) (top), and the signal $T5\text{qqqWW}(1900,100)$ (left) and $T5\text{qqqWW}(1500,1000)$ in the the $\Delta\phi$ vs. L_T plane after the preselection.

Signal acceptance:

The yield after baseline selection for the simulated signal events is shown in Figure 4.12 in the gluino-neutralino mass plane for the $T5\text{qqqWW}$ model. The acceptance of at least one event for 35.9 fb^{-1} integrated luminosity is up to 2250 GeV for gluino mass for the neutralinos masses below 1600 GeV. The equivalent selection efficiency, which is the fraction of baseline events over the total number of simulated events, is presented in the same Figure (top/right). In the bottom part of Figure, the SR efficiencies are shown. The plot on the left displays the efficiency for the constant $\Delta\phi > 1$ cut and it is changing between

50-70%. On the other hand, the left plot shows the efficiency for the variable $\Delta\phi$ cut with respect to L_T bin and this time the efficiency goes up to 80%.

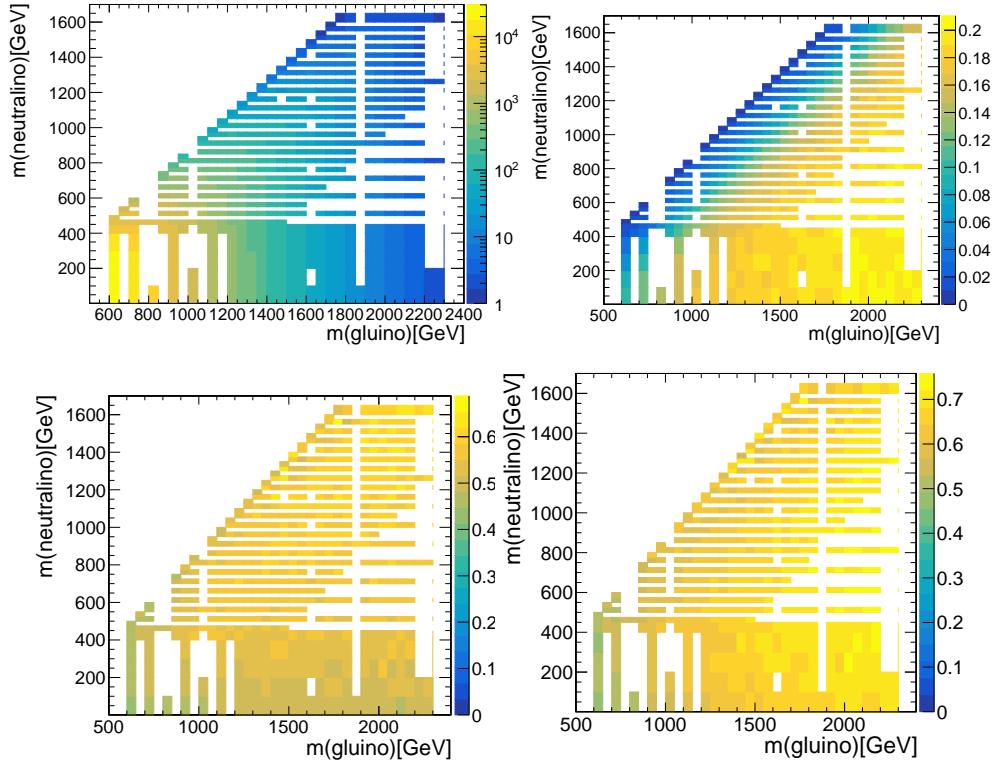


Figure 4.12: The distributions on top represent T5qqqqWW signal counts (left) and selection efficiency (right) after the baseline requirements in the gluino-lsp plane. The distributions at the bottom show the efficiency of signal region selection with respect to the baseline; $\Delta\phi > 1$ (left) and $\Delta\phi > x$ (right) where x stands for the cut value changing according to L_T bin.

Table 4.5: Signal regions.

5 jets, 0 b-tagged jets		
L_T [GeV]	H_T [GeV]	$\Delta\phi$
[250, 350]	[500, 750]	1.0
	≥ 750	1.0
[350, 450]	[500, 750]	1.0
	≥ 750	1.0
[450, 650]	[500, 750]	0.75
	[750, 1250]	0.75
	≥ 1250	0.75
≥ 650	[500, 750]	0.50
	[750, 1250]	0.50
	≥ 1250	0.50
[6,7] jets, 0 b-tagged jets		
L_T [GeV]	H_T [GeV]	$\Delta\phi$
[250, 350]	[500, 1000]	1.0
	≥ 1000	1.0
[350, 450]	[500, 1000]	1.0
	≥ 1000	1.0
[450, 650]	[500, 750]	0.75
	[750, 1250]	0.75
	≥ 1250	0.75
≥ 650	[500, 750]	0.50
	[750, 1250]	0.50
	≥ 1250	0.50
≥ 8 jets, 0 b-tagged jets		
L_T [GeV]	H_T [GeV]	$\Delta\phi$
[250, 350]	[500, 1000]	1.0
	≥ 1000	1.0
[350, 450]	[500, 1000]	1.0
	≥ 1000	1.0
[450, 650]	[500, 1250]	0.75
	≥ 1250	0.75
≥ 650	[500, 1250]	0.50
	≥ 1250	0.50

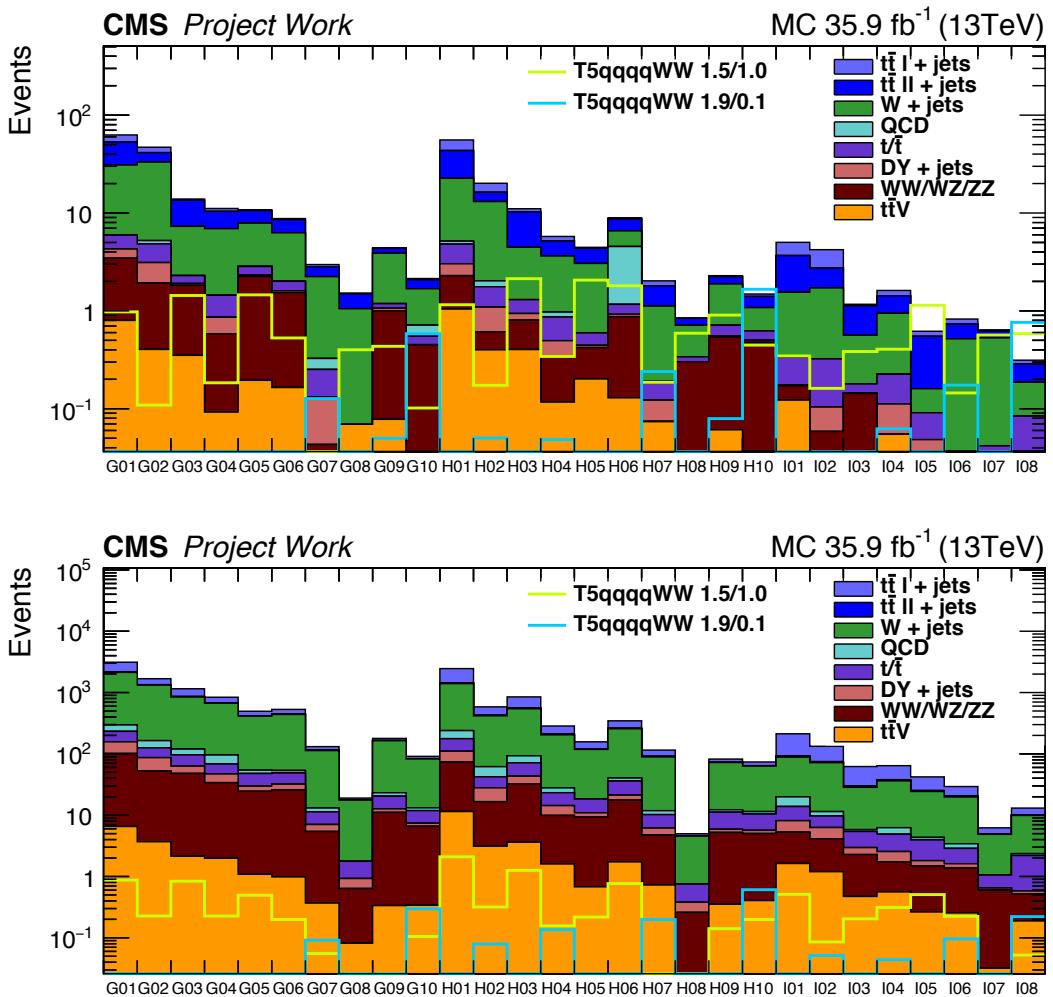


Figure 4.13: Simulated single muon event yields for the background processes are shown as color filled stacked histograms for all 28 search bins. The two signal benchmark models are overlayed and shown by line histograms. The upper plot shows the high $\Delta\phi$ regions (SR) while the lower plot shows the low $\Delta\phi$ regions (CR).

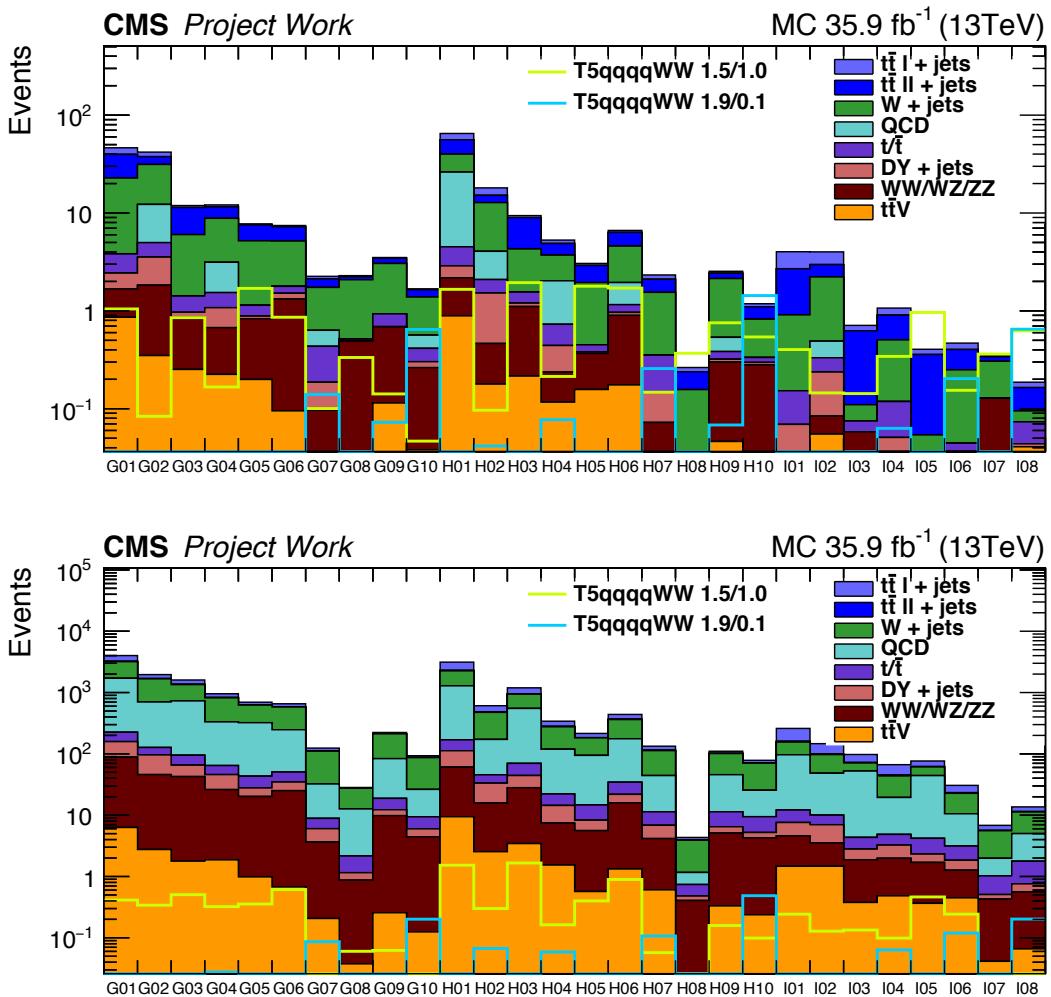


Figure 4.14: Simulated single electron event yields for the background processes are shown as color filled stacked histograms for all 28 search bins. The two signal benchmark models are overlayed and shown by line histograms. The upper plot shows the high $\Delta\phi$ regions (SR) while the lower plot shows the low $\Delta\phi$ regions (CR).

Chapter 5

Background Estimation: R_{CS} Method

Contents

5.1	QCD background estimation	62
5.2	Background fraction calculations: b-tag multiplicity fit	63
5.3	R_{CS} method in $t\bar{t} + \text{jets}$ events	65
5.4	R_{CS} method in $W + \text{jets}$ events	67
5.5	Validation of the background estimation	70

When searching for new physics, it is important to know precisely the standard model background contributions in the search regions (SRs) explained in the previous chapter. The modeling of the SM is not trivial due to complicated processes such as QCD and lack of well-studied detector performance. Therefore, simulated samples can be beneficial to have an initial idea on the kinematics. However, they can not be entirely trustful. In this analysis, a data-driven approach is used to estimate the main backgrounds. To predict the number of events in search regions, the control regions are used with transfer factors, which is called R_{CS} , measured in lower jet multiplicity regions in data. This is the core background estimation method of the analysis. R_{CS} is defined as the ratio of number of events in SR to the number of events in CR:

$$R_{CS} = \frac{N(\Delta\Phi > x)}{N(\Delta\Phi < x)} = \frac{N^{SR}}{N^{CR}}. \quad (5.1)$$

R_{CS} is measured in lower n_{jet} regions, which are not overlapping with the search regions. This requires that the R_{CS} values are stable as a function of n_{jet} . Throughout this thesis, these regions are called as side band regions (SB). The formulation of this procedure is:

$$N_{MB}^{SR} = R_{CS}^{MB} \cdot N_{MB}^{CR}, \quad R_{CS}^{MB} \sim R_{CS}^{SB}, \quad N_{MB}^{SR} = R_{CS}^{SB} \cdot N_{MB}^{CR}. \quad (5.2)$$

Requiring no b-tagged jets in the final state leads to both $W + \text{jets}$ (see Sec. 5.4) and $t\bar{t} + \text{jets}$ (see Sec. 5.3) background components to be equally important in the MB SR.

Other small background contributions are less than 10% and directly taken from the MC. Consequently, the predicted number of events in main band signal regions can be written as the sum of its components:

$$N_{Total}^{SR} = N_{W+jets}^{SR} + N_{t\bar{t}+jets}^{SR} + N_{other}^{SR(MC)}. \quad (5.3)$$

The R_{CS} strategy developed for this analysis takes into account the differences in R_{CS} values of these two components. Furthermore, the R_{CS} can be defined as a combination of R_{CS} from different source of backgrounds:

$$R_{CS} = \frac{N^{SR}}{N^{CR}} = \frac{\sum N_i^{SR}}{\sum N_i^{CR}} = \sum \frac{N_i^{SR}}{N_i^{CR}} \cdot \frac{N_i^{CR}}{N_i^{CR}} = f_i^{CR} \cdot R_{CS}^i, \quad (5.4)$$

where i stands for either $W + \text{jets}$ or $t\bar{t} + \text{jets}$ and f_i^{CR} is the relative yield of the i th background component. The relative compositions in low $\Delta\phi$ control regions are determined from fit to data (see Sec. 5.2). The measurement of R_{CS} is performed in two separate side band regions. These regions are chosen, such that they mimic the kinematics of the main band as closely as possible. For $t\bar{t} + \text{jets}$, the side band region is $4 \leq n_{\text{jet}} \leq 5$ and $n_{b-\text{jet}} \geq 1$ while for $W + \text{jets}$ the side band is chosen as $3 \leq n_{\text{jet}} \leq 4$ and $n_{b-\text{jet}} = 0$. Despite the fact that QCD multijet events have low contribution in main band signal regions, their contamination in side bands as well as the control regions of main bands has to be estimated and subtracted prior to application of the R_{CS} method. Again a data based approach is used (see Sec. 5.1). Tab. 5.1 summarises the main band and side band regions used in this analysis. The background prediction mechanisms are validated with data as described in Sec. 5.5.

Table 5.1: Overview of the definitions of sideband and mainband regions. For the multijet (QCD) fit the electron (e) sample is used, while for the determination (det.) of $R_{CS}(W+\text{jets})$ the muon (μ) sample is used. Empty cells are not used in the analysis.

$n_{b-\text{tag}}$	$n_{b-\text{tag}} = 0$	$n_{b-\text{tag}} \geq 1$
$n_{\text{jet}} = 3$	$R_{CS}(W+\text{jets})$ det. (μ sample)	
$n_{\text{jet}} = 4$	QCD bkg. fit (e sample)	
$n_{\text{jet}} \geq 5$	Main band regions	$R_{CS}(t\bar{t} + \text{jets})$ det.

5.1 QCD background estimation

Due to the complicated nature of quantum chromodynamics (see Sec. 1.1.2), simulation of QCD multi-jet events is challenging and fails to be accurate. Naturally in this case, a data-driven prediction of the QCD events in search regions is necessary. According to simulation, the majority of the QCD multijet events are located in CRs and side band (low n_{jet})

regions. Although, QCD multijet events are not one of the main backgrounds, in order to have a more accurate calculation of R_{CS} , this contamination needs to be subtracted from SBs and as well as the CRs of MBs. Therefore, the prediction of QCD multi-jet events has to be performed prior to the R_{CS} method.

Involvement of QCD events in signal regions occurs through fake leptons, which are misidentified jets or photons. Thus, a fake-lepton enriched data control sample, which is obtained by loosening and inverting the criteria on lepton identification variables, is designed for the estimation. To estimate the fraction of events with fake leptons, that pass the analysis selection, ratio of selected to anti-selected events (F_{sel}) is used. To ensure the orthogonality to the main band regions, this ratio is measured in $3 \leq n_{\text{jet}} \leq 4$ and $n_{\text{b-jet}} = 0$ side band. The feature of containing fake leptons helps also to distinguish QCD events from the EWK ones (containing real leptons) by using a variable which reflects the polarisation of the W boson. A variable called L_P was introduced in [?], it was for the first measurement of the W polarization at LHC. The variable L_P is defined as follows:

$$L_P = \frac{\vec{p}_T(\ell) \cdot \vec{p}_T(W)}{|\vec{p}_T(W)|^2} = \frac{p_T(\ell)}{p_T(W)} \cos(\Delta\phi(W, \ell)), \quad (5.5)$$

where $\vec{p}_T(\ell)$ and $\vec{p}_T(W)$ are the transverse momenta of the charged lepton and the W boson respectively. As seen in Fig. 5.1, QCD events (dashed cyan colored lines) mostly populate the L_P around 1 while the EWK events has a falling distribution between 0 to 1. The number of QCD events in the control regions for selected leptons is then obtained by multiplying the yield of the anti-selected events, which is obtained from the example template fit shown in Fig. 5.1 (left), with F_{sel} which is shown in the same figure as a function of L_T bins.

A profound description of the method can be found in [49].

5.2 Background fraction calculations: b-tag multiplicity fit

Fractions of the background processes in the control regions (f_i^{CR}) are calculated using likelihood fit of b-tag multiplicity distributions templates to data. The templates are obtained from simulation except the QCD events. The QCD multijet contribution in b-tag multiplicity bins is taken from the prediction, which is introduced in the previous section. $W + \text{jets}$ events display a charge asymmetry in R_{CS} due to polarization effects. To account for this, the fits are performed separately for each charge: positive and negative. $t\bar{t} + \text{jets}$ and QCD events are considered as homogenous in charge. Other background templates are also produced separately for positive and negative charged leptons. Fig. 5.2 shows good agreement between data and the b-tag multiplicity fit.

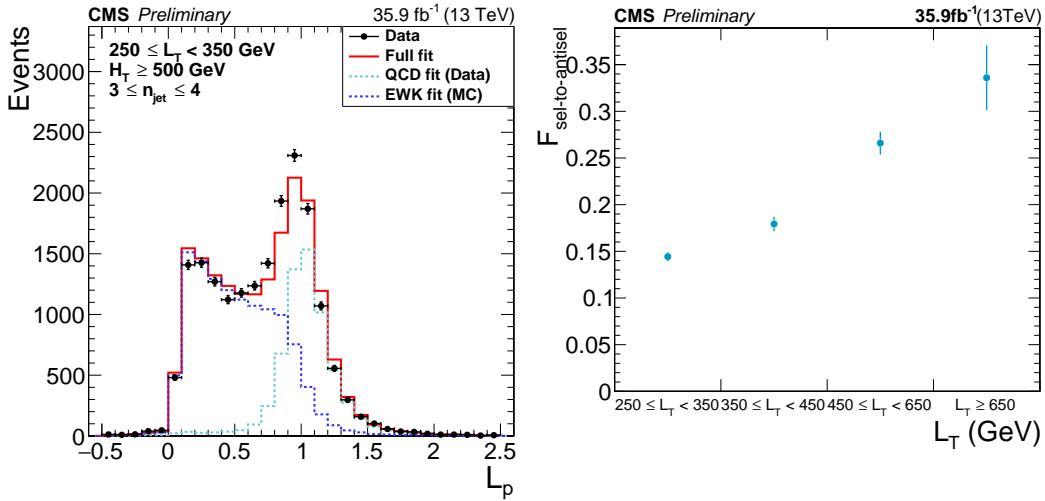


Figure 5.1: L_p shape fit result for $3 \leq n_{\text{jet}} \leq 4$ and $n_{\text{b-jet}} = 0$ in $250 \leq L_T \leq 350$ bin (left). Ratio of selected to anti-selected electron events from QCD for $3 \leq n_{\text{jet}} \leq 4$ and $n_{\text{b-jet}} = 0$, in bins of LT in data (right).

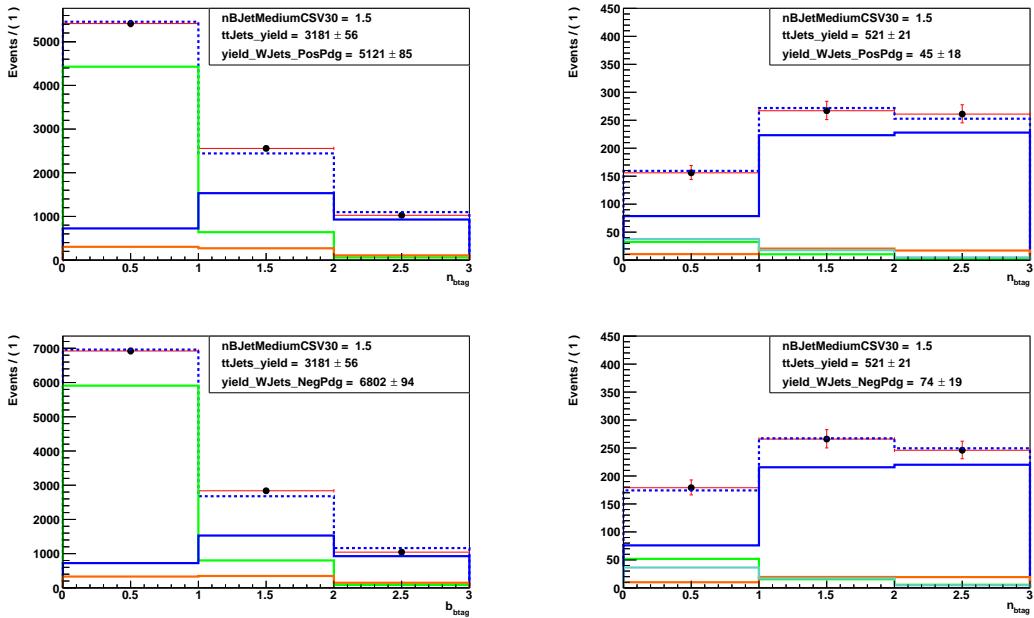


Figure 5.2: b-tag multiplicity fit is performed in control regions. 3-4 jets sideband (left), 8 jets main band (right). Upper plots are showing the negative charged lepton selection while lower plots show the positive charged selection. Black points represent data, green is for $W + \text{jets}$, blue is for $t\bar{t} + \text{jets}$, orange is for other EWK backgrounds and cyan colored lines show QCD contamination.

5.3 R_{CS} method in $t\bar{t} + \text{jets}$ events

In LHC, the top quark pair production mechanism is dominated by gluon-gluon fusion and quark-quark annihilation. According to the SM, the top quark decays into Wb in almost 100% of the cases. In other words, nearly all of the $t\bar{t} + \text{jets}$ events involve two b-tagged jets. However, the b-tagging algorithms are not fully efficient; therefore $t\bar{t}$ events can survive after the b-tagged jet veto requirement. The surviving events have a very similar final state with the T5qqqqWW signal events. Due to its high cross section, $t\bar{t} + \text{jets}$ is one of the two main backgrounds of this search.

The $t\bar{t}$ decay can be categorized into three with respect to lepton content of the final state; single leptonic, di leptonic and fully hadronic. The branching ratios of these decay modes are approximately 45%, 9% and 44%, respectively. After the baseline selection of the analysis, the single leptonic decay channel, where one of the two W bosons decays leptonically, populates the low $\Delta\phi$ region while the di leptonic decay channel, where one of the leptons are lost due to misidentification or limited detector acceptance dominates the high $\Delta\phi$ region. As a result of this, di leptonic $t\bar{t}$ events have higher R_{CS} values. Fig. 5.3 shows the R_{CS} as a function of n_{jet} for di leptonic (left) and single leptonic (right) decays separately. It should be noted that this figure shows only simulated values and it is only for illustra-

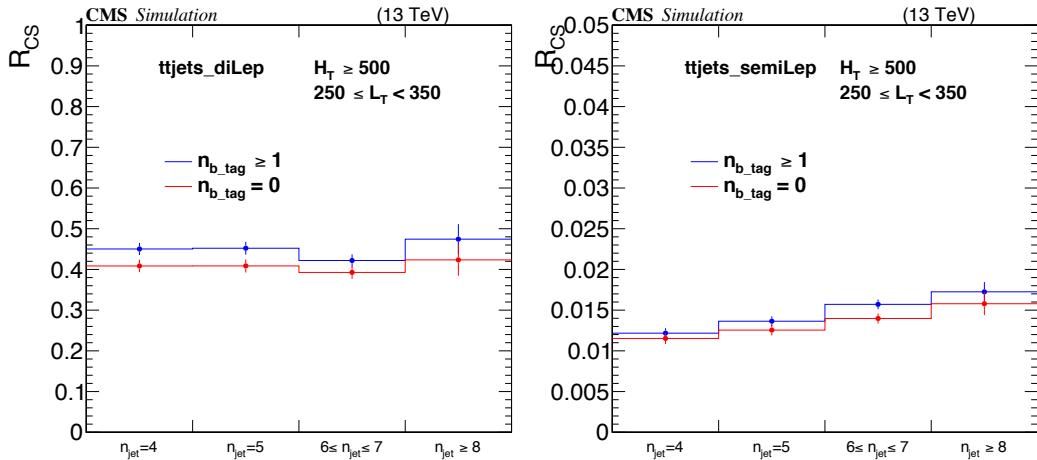


Figure 5.3: R_{CS} as a function of n_{jet} for dileptonic (left), and signle leptonic (right) $tt + \text{jets}$ events.

tion. In the primary background estimation procedure, R_{CS} values of $t\bar{t} + \text{jets}$ events are measured in sideband regions requiring four to five jets and at least one b-tagged jet from data. This selection increases the purity of $t\bar{t} + \text{jets}$ events while at the same time keeps $w\text{Jets}$ events contamination low. In addition to this, the yield of predicted QCD multi-jet

events is subtracted from the CR. The formulation of R_{CS} :

$$R_{CS}^{data}(b \geq 1, n_{jet} \in [4, 5]) = \frac{N_{SR}^{data}}{N_{CR}^{data} - N_{CR}^{QCD(pred)}}. \quad (5.6)$$

Fig. 5.4 shows simulated (left) and measured(right) values for the first search bin which is $n_{jet} = 5$, $L_T \in [250, 350]$ GeV and $H_T \in [500, 1000]$ GeV. Residual differences between the R_{CS} in the side band and main band is calculated in simulation as a correction factor κ . In $t\bar{t}$ background estimation two kind of kinds of κ corrections are used. The first one is

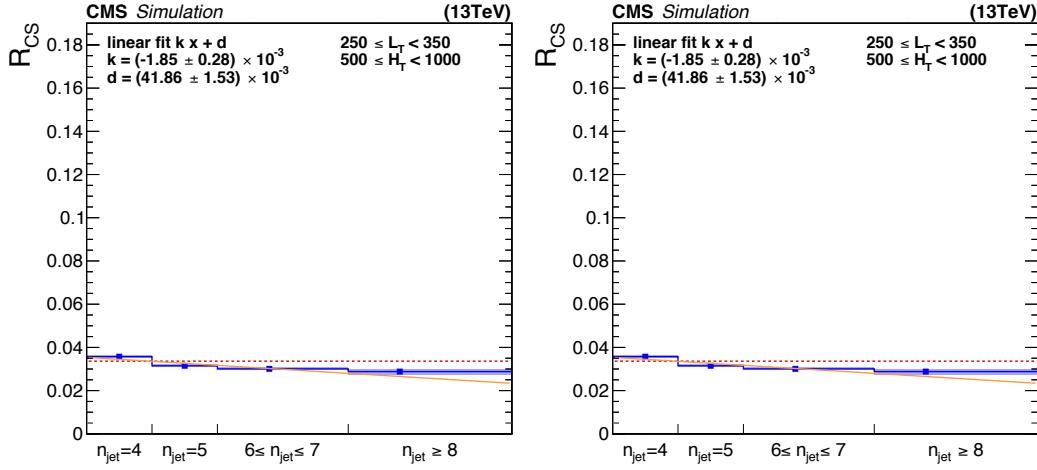


Figure 5.4: R_{CS} as a function of n_{jet} from simulation of $t\bar{t} +$ jets events (left), and data (right) in $t\bar{t} +$ jets sideband.

called κ_b . As the name implies, it covers the possible difference between the R_{CS} values of the b-tagged region and no b-tagged region. This correction also accounts for small contributions from processes other than $t\bar{t} +$ jets and QCD multi-jet. The correction factor κ_b can be written as follows:

$$\kappa_b = \frac{R_{CS}^{MC}(0b, n_{jet} \in [4, 5], t\bar{t})}{R_{CS}^{MC}(b \geq 1, n_{jet} \in [4, 5], EWK)} \quad (5.7)$$

The second correction factor accounts for a residual dependence of R_{CS} on jet multiplicity:

$$\kappa_{t\bar{t}} = \frac{R_{CS}^{MC}(0b, n_{jet} : MB, t\bar{t})}{R_{CS}^{MC}(0b, n_{jet} \in [4, 5], t\bar{t})} \quad (5.8)$$

As mentioned in the beginning of this chapter, the total R_{CS} is based on the fraction of different channels and their corresponding R_{CS} values. In this case, the difference between R_{CS} values of single leptonic and di leptonic events should be considered. In this point,

another control region is designed to understand the effect of di lepton events that pass the single lepton selection. To obtain a high-purity di lepton $t\bar{t}$ control sample in data, two leptons of opposite charge are required. It is also required that the mass of two same flavor leptons is more than 10 GeV away from the Z mass peak. To mimic the single lepton events, one of the two leptons is removed. These so-called lost leptons are mostly coming from $\tau \rightarrow \text{hadrons} + \nu$ decays. Therefore, the removed lepton is replaced by a jet with 2/3 of the original lepton's p_T to account for the missing energy due to the neutrino from the τ decay. The L_T , H_T and $\Delta\phi$ values of the reconstructed single-lepton event are recalculated. No $\Delta\phi$ requirement is applied, and all events are used twice, with each reconstructed lepton being considered as the lost lepton to increase the statistics in this di-lepton control region. In Fig. 5.5, the top row displays the n_{jet} distributions for the lost lepton events obtained from di lepton ones (left) and usual single lepton events (right) in control region (low $\Delta\phi$). The distributions are obtained after baseline requirement for L_T and H_T with a lower jet multiplicity requirement ($n_{\text{jet}} \geq 3$) to account for side band regions. The correctness of the description of the n_{jet} distribution in simulation is determined from the double ratio, which is the single-lepton and di lepton ratio between data and simulation, as shown in the bottom of the same figure. In the case simulation describes the dileptonic events in data entirely, the double ratio would be flat and located around 1. However, double ratio manifests that the di leptonic events in simulation should be corrected to calculate a more accurate $\kappa_{t\bar{t}}$. The weight for dileptonic events are extracted from double ratio :

$$weight_{DL} = 1.22 + (-0.044) * (n_{\text{jet}} - 5.9), \quad (5.9)$$

where the value 5.9 is coming from the average of the n_{jet} of the single lepton distribution. The $\kappa_{t\bar{t}}$ described in Eqn. 5.8 is replaced with a κ calculated after this di lepton correction. The comparison of $\kappa_{t\bar{t}}$ before and after correction can be seen in Fig. 5.6.

The resultant R_{CS} for $t\bar{t} + \text{jets}$ estimation is then written as:

$$R_{CS}^{t\bar{t}}(0b, MB) = \kappa_b \cdot \kappa_{t\bar{t}}^{DL-Corr} \cdot R_{CS}^{data}(b \geq 1, n_{\text{jet}} \in [4, 5]). \quad (5.10)$$

5.4 R_{CS} method in $W + \text{jets}$ events

In this search, when performing the $W + \text{jets}$ estimation, the WV di-boson events, where V stands for W or Z bosons, are included in the prediction mechanism. In the di-boson events, which are considered as a part of $W + \text{jets}$ estimation, the W boson decays leptonically and the second boson, denoted by V , decays hadronically. The similarity of the kinematics of the events, hence the R_{CS} values, makes this addition possible. All other di-boson events are treated as part of the rare EWK backgrounds and taken from simulation.

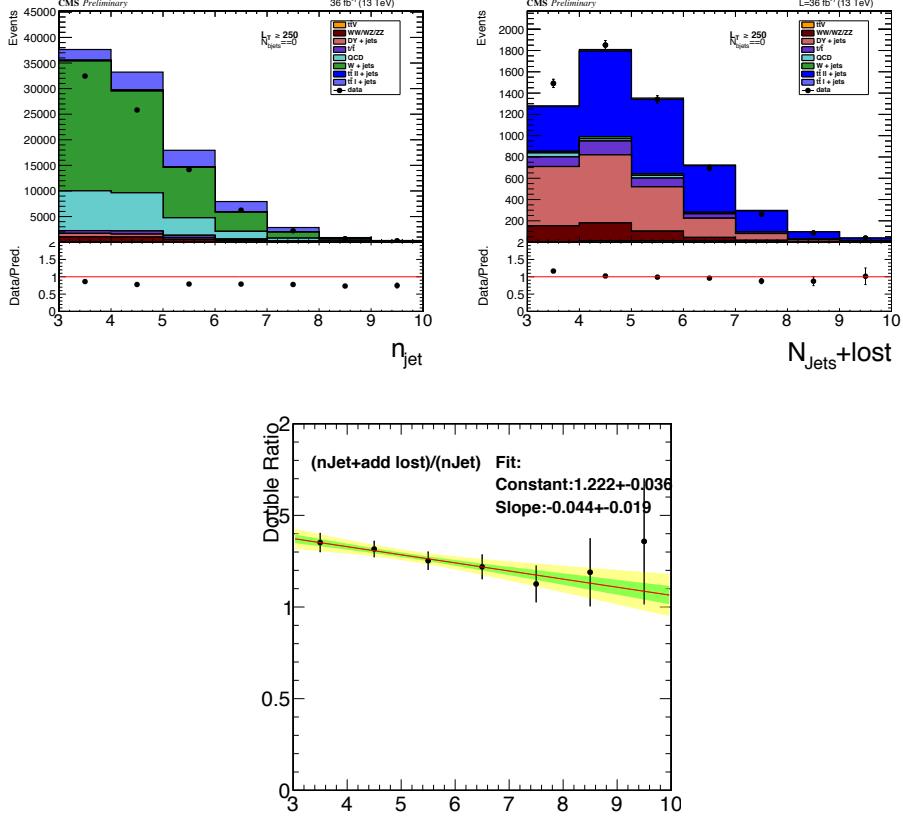


Figure 5.5: Jet multiplicity distribution after the single lepton baseline event selection (in the control region) with $H_T > 500$ GeV, $L_T > 250$ GeV, and $n_{b\text{-tag}} = 0$ (left) and in the dileptonic control region with recalculated H_T and L_T cuts (right). Only nISR reweighting is applied as scale factors.

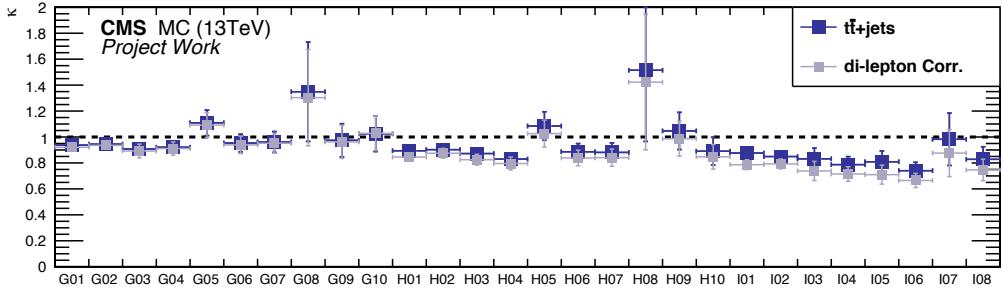


Figure 5.6: Changes to $\kappa_{t\bar{t}}$ due to applying the double-ratio as given by Eqn. 5.9 and reevaluating $\kappa_{t\bar{t}}$. The corrected κ , which is shown with a lighter color, is used as $\kappa_{t\bar{t}}$ in the full background prediction.

R_{CS} values for $W + \text{jets}$ background estimation is measured in a sideband region with three to four jets. To suppress the $t\bar{t} + \text{jets}$ events and to be kinematically as close as possible to main band regions, zero b-tagged region is used. Furthermore, in order to suppress QCD contamination R_{CS} is measured only in muon channel. Therefore, it is no longer needed to subtract predicted yield of QCD multi-jet events. However, remaining $t\bar{t} + \text{jets}$ events contamination is subtracted both from signal and control regions. In the determination of the amount of $t\bar{t} + \text{jets}$ events contamination, fractions $f_{t\bar{t}}$ in the sideband control region is taken from the b-tag multiplicity fit, and R_{CS} is taken from MC directly. The R_{CS} can then be written as:

$$R_{CS}^{Corr.\text{data}}(0b, n_{\text{jet}} \in [3, 4]) = \frac{N_{\text{data}}^{SR} - R_{CS}^{t\bar{t},MC} \cdot f_{t\bar{t}}^{fit} \cdot N_{\text{data}}^{CR}}{(1 - f_{t\bar{t}}^{fit}) \cdot N_{\text{data}}^{CR}}. \quad (5.11)$$

As already discussed in Sec. 5.2, R_{CS} is separately measured for positive and negative charged leptons to account for the charged asymmetry of the $W + \text{jets}$ events. Fig. 5.7

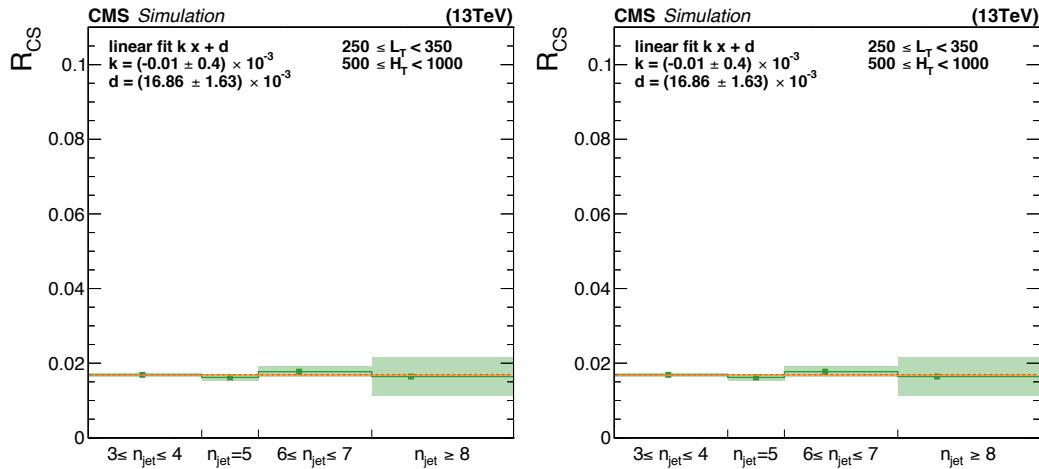


Figure 5.7: R_{CS} as a function of n_{jet} from simulation of $W + \text{jets}$ events (left), and data (right) in $W + \text{jets}$ sideband.

shows simulated (left) and measured (right) values for the first search bin which is $n_{\text{jet}} = 5$, $L_T \in [250, 350] \text{ GeV}$ and $H_T \in [500, 1000] \text{ GeV}$. Again as in the $t\bar{t}$ case, residual differences between the R_{CS} in the side band and main band is calculated in simulation as a correction factor κ . This time only one κ factor is used, due to the fact that the side band and main band regions are sharing the same requirement for number of b-tagged jets, which is zero. The κ_w accounts for residual dependence of R_{CS} on the jet multiplicity and also covers variances between R_{CS} values in muon only channel to combined lepton final state. The

κ_w is calculated as:

$$\kappa_w = \frac{R_{CS}^{MC}(0b, n_{jet} : MB, W + jets)}{R_{CS}^{Corr, MC}(0b, n_{jet} \in [3, 4], \mu)}. \quad (5.12)$$

Then the resultant R_{CS} is described as:

$$R_{CS}^W(0b, MB) = \kappa_w \cdot R_{CS}^{Corr, data}(0b, n_{jet} \in [3, 4], \mu). \quad (5.13)$$

Fig. 5.8 shows the κ_w values in the lower panel and R_{CS} values went into this calculation in the upper panel. In two kinematically extreme bins of high L_T and low H_T the R_{CS} values are different from the bulk but the SB follows the MB, therefore, κ values are still compatible with 1.

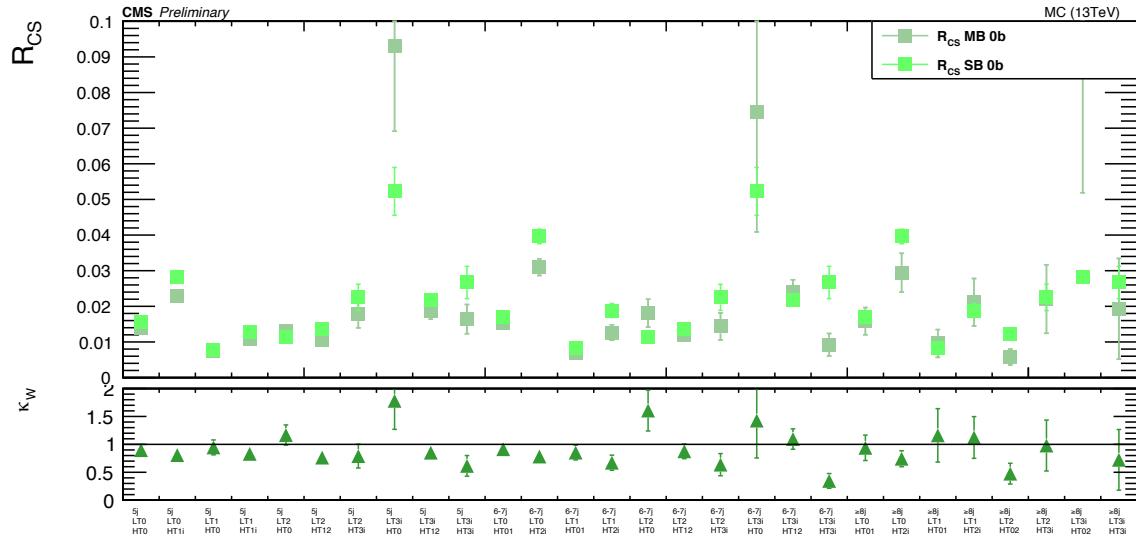


Figure 5.8: R_{CS} values went into κ_w (lower panel) calculation is shown.

5.5 Validation of the background estimation

Validation of the background estimation method is performed in events where there are four jets and zero b-tagged jets. This region is a part of the $W + jets$ sideband of the main estimation and it is dominated by the SM background events. Consequently, in the validation, to perform the $W + jets$ prediction, only three jets region is used as the side band. The $t\bar{t} + jets$ sideband remains unchanged since it is not used as the validation search region. Fig. 5.9 presents the kappa values calculated for the validation. The results of the validation as well as the main prediction will be presented in Sec. 7.1.1 after the discussion of the systematic uncertainties.

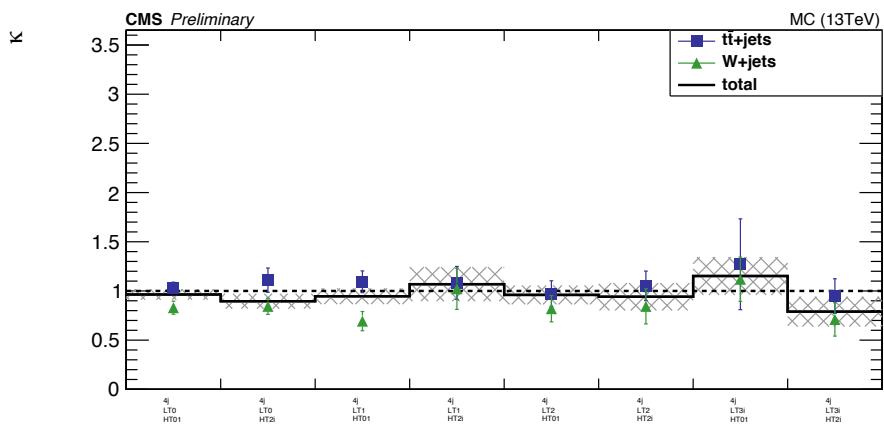


Figure 5.9: κ_w and $\kappa_{t\bar{t}}$ correction factors are shown. The total κ , which is represented with black line, shows the ratio of prediction to the simulation. The shaded area displays the uncertainty coming from simulation statistics.

Chapter 6

Systematic Uncertainties

Contents

6.1 Systematic uncertainties on background estimation	72
6.2 Systematic uncertainties on signal modelling	75
6.3 Common systematic uncertainties for signal and background modelling	76

The data driven background estimation and using the ratio of the number of signal region events to control region events eliminates the systematic uncertainties related to potential mismodeling of the tails of kinematic distributions such as L_T , H_T , $\Delta\phi$. However, there are still several sources that can affect the background prediction and expected signal events counts. These sources and the methods to calculate the uncertainties are discussed in the following.

6.1 Systematic uncertainties on background estimation

These sources and the methods to calculate the uncertainties are discussed in the following. Because the κ factors of the analysis are taken from simulation, these factors are reflecting the influence of mismodeling of the R_{CS} and hence the fractions of background processes on the analysis. Therefore, the systematic uncertainties on the background estimation are calculated through $\kappa_{t\bar{t}}$ and κ_w . The amount of uncertainty is then determined as:

$$\delta = \frac{\kappa'_x}{\kappa_x} - 1, \quad (6.1)$$

where κ' reflects the recalculated κ with varied weights and x stands for $t\bar{t}$ or w .

n_{jet} extrapolation for $t\bar{t} +$ jets:

One of the major systematic uncertainties on the background estimate results from the extrapolation of R_{CS} from the measurement region, low n_{jet} , to the application region, high

n_{jet} . This uncertainty can be calculated with a fit which is performed over the n_{jet} range as in Fig. 5.7 (left). The relative difference between the R_{CS} value obtained in the side band in simulated samples and the value derived from the fit is taken as a systematic uncertainty on $W + \text{jets}$ events.

The same procedure would also be applied for $t\bar{t} + \text{jets}$ events as well. However, in this method limit statistics of some bins affects significantly the systematic uncertainty calculations. Therefore, another method is developed based on the impact of different fractions of dileptonic and single leptonic $t\bar{t}$ events. To account for this effect, as discussed in Sec. 5.3, an event by event weight for correction factor $\kappa_{t\bar{t}}$, which is shown in Eq. 5.9, is derived. In addition to this, two weights are obtained to calculate the systematic uncertainties originated from the procedure:

$$W(DL_{\text{Const}}) = 1 \pm 23\%, \quad (6.2)$$

$$W(DL_{\text{Slope}}) = 1 \pm (n_{\text{jet}} - 5.9) \cdot 5\%. \quad (6.3)$$

The variations are extracted as the quadratic sum of the deviation from 1 for the constant (offset) parameter and its uncertainty or as the quadratic sum of the deviation from 0 for the slope parameter and its uncertainty. The uncertainties derived from the slope and constant variation as a function of the $\kappa_{t\bar{t}}$ is shown in Fig. 6.1. In this Figure, color filled areas represent the dilepton uncertainties while the black line shows the uncertainty calculated with fit along the n_{jet} , which is not used in the analysis. The offset variation almost has no effect on $\kappa_{t\bar{t}}$ while the slope variation has an effect up to 20%. The figure manifests that the two different approach result in the similar values of uncertainties while the uncertainty calculated with n_{jet} fit fluctuates to higher values the statistical effects.

Cross sections:

To account for possible biases in the estimation of the background composition in terms of $W + \text{jets}$ vs. $t\bar{t} + \text{jets}$ events, uncertainties on their cross sections are taken in to account. Although in the inclusive regions the W boson cross section [52] and $t\bar{t}$ cross section [53] uncertainties are at the order of ten percent level, in the analysis regions it can be larger. The $W + \text{jets}$ and $t\bar{t} + \text{jets}$ cross sections are conservatively varied by 30%, which lead to a change 0.3-10% (0.7-13%) in the κ_w ($\kappa_{t\bar{t}}$) values.

For other small EWK background processes, which are taken directly from simulation, again a conservative 50% variation is applied. Because, the fraction of these events are small, the effect of this variation on the κ values is calculated to be changing between 0.1-3.8%.

W polarization:

The main search variable $\Delta\phi$ reflects the angular information between the W boson and

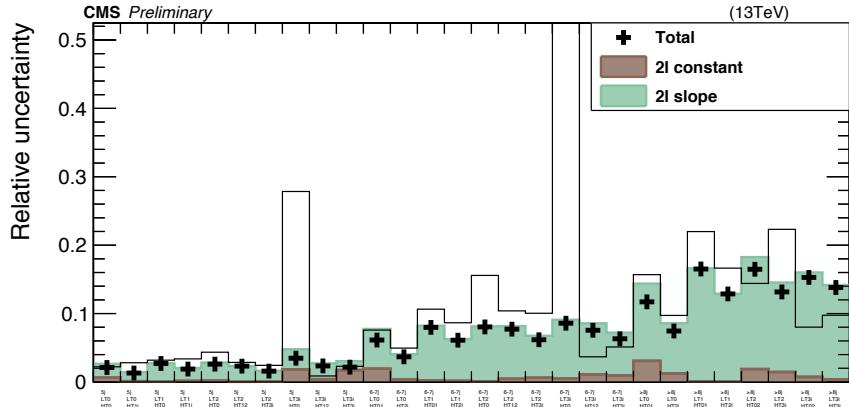


Figure 6.1: Relative uncertainty on $\kappa_{t\bar{t}}$ due to the different composition of dileptonic events in sideband and main band regions. Color filled areas represent the dilepton uncertainties while the black line shows the uncertainty calculated with fit along the n_{jet} .

its decay products. Therefore, the W boson polarization affects the $\Delta\phi$ distribution. The effect of the polarization on background estimation of $W + \text{jets}$ and $t\bar{t} + \text{jets}$ is calculated by reweighting these events by:

$$w = (1 \pm \alpha \cdot (1 - \cos(\theta^*))) \cdot C \quad (6.4)$$

where a is 5% for $t\bar{t} + \text{jets}$ and 10% for $W + \text{jets}$ events, C is a factor to keep the normalization after baseline constant and θ^* is the angle between the charged lepton and the W boson in the W rest frame. This procedure results in an uncertainty below 3% in all the signal regions.

number of ISR jets reweighting:

In Sec. 4.2.1, the reweightenning of events according to number of ISR jets is already explained. κ values are recalculated by varying the ISR weights with the systematic uncertainties listed in Tab. 4.2. The amount of variation is then calculated using the Eqn. 6.1 and found to be varying between 0.2-11%.

QCD multijet events prediction:

QCD multijet events prediction is entering in this search through the b-tag multiplicity fit (see Sec.5.2) to measure the background compositions in control regions and also through calculation of the R_{CS} values. Therefore the uncertainty originated from QCD prediction has to be considered. A profound explanation of the QCD prediction and its uncertainty can be found in [49]. The effect of this uncertainty is propagated to the R_{CS} prediction. The resultant uncertainty is found to be around 3%.

R_{CS} difference muon/(electron or muon):

As discussed in Sec. 5.4, R_{CS} of $W + \text{jets}$ events is calculated only in the muon channel. To account for the inconsistency between the R_{CS} with only muon channel and the true R_{CS} , the discrepancy between these values are taken from the simulation. This discrepancy is assigned as systematic uncertainty. In order to avoid large uncertainties driven by the limited statistics, the systematic uncertainty is restricted to be smaller than the statistical error of the true R_{CS} value.

6.2 Systematic uncertainties on signal modelling

The uncertainties considered in this section are applied only to the simulated signal events. The amount of discrepancy is calculated as:

$$\delta = \frac{N'_{\text{events}}}{N_{\text{events}}} - 1, \quad (6.5)$$

where N'_{events} represents the recalculated number of simulated signal events with varied weights and N_{events} is the true number of events. The systematic uncertainties are calculated for each search bin separately as in the background case, this time uncertainties are also calculated for each 657 gluino/neutrino mass points.

Initial state radiation:

In Run 1 it was observed that the hadronic recoil from ISR for boosted heavy particle pairs such as the $t\bar{t}$ system is not well described by the `MADGRAPH` event generator [54]. As a very similar system, it is expected that the gluino pair lead discrepancies. An uncertainty based on the p_T of the gluino-gluino system is applied:

- p_T (gluino-gluino) less than 400 GeV: no uncertainty
- p_T (gluino-gluino) between 400 GeV and 600 GeV: 15% uncertainty
- p_T (gluino-gluino) above 600 GeV: 30% uncertainty.

The uncertainty on the number of expected signal events is calculated by varying each event with propagating the variations listed above through Eqn. 6.5.

Factorization/renormalization scale:

To account for the impact of renormalization and factorization scales on the signal acceptance, the scales are varied by a factor of 0.5 and 2, respectively. As discussed in Sec. 2.3, there are several weights to be applied for these scale factors. Therefore, to calculate the uncertainty an envelope of all variations is computed. Furthermore, to keep the cross section unchanged, a normalization factor is applied. The resultant uncertainty on the expected event yields is similar for all the mass points and it is changing between 1-3%.

Reconstruction of MET:

On the recommendation of corridor studies group, the analysis is performed separately for reconstructed and generated MET, that is to recalculate all the kinematic variables, which include MET. The average of the two acceptances is taken as the central value of the expected signal yields for each main band region. Furthermore, one-half of the difference between the two acceptances is taken as an uncertainty.

Trigger:

As introduced in Sec. 4.4.1, the uncertainty on the trigger selection efficiency is measured to be around 2%, and also considered as the uncertainty on signal simulation.

Luminosity:

As shown in Sec. 2.2.7, the pixel cluster counting method [50] is used to calculate luminosity. The uncertainty on this measurement is predicted to be 2.5% [51].

6.3 Common systematic uncertainties for signal and background modelling

Systematic uncertainties affecting both background and the signal processes are related to the mismeasurement misidentification of particular objects in the events. These uncertainties have impact on the kinematic variables of the analysis: $\Delta\phi$, L_T , HT , n_{jet} , n_{b-jet} . Therefore, for the signal events they are affecting the acceptance and the selection efficiency and for the background events they may vary the R_{CS} values.

Jet Energy Scale:

Calculating uncertainty on the jet energy scales is another dedicated study performed by JET-MET physics object group at CMS. Within the analysis, these variations which are affecting the jet energy spectrum as well as E_T , are applied to the jets in individual events and therefore related kinematic variables are reconstructed. For the background estimate, κ values are recalculated with up and down scaled jets. The uncertainty, which is obtained using Eqn. 6.1, is found to be changing between 0.7-26%. For the signal, yields are recalculated and variation is measured using Eqn. 6.5. In this case, depending on the mass point and signal region, uncertainty can take values up to 40%.

Tagging of b-jets:

B-tagging uncertainties, which are related to difference in b-tagging efficiency between simulated and observed events, have an influence on this analysis through acceptance and b-tag multiplicity fit. Uncertainties are calculated to be less than 3% for the background, and between 1-6% for the signal.

Lepton identification and reconstruction:

Lepton identification and reconstruction efficiencies are different for simulated and observed events. For the backgrounds, a flat 5% uncertainty is assigned to account for this difference. For signal, also variations originated from the difference between fast simulation [55] and full simulation are taken into account. The resultant uncertainty on the expected signal yields is found to be 2%.

Pileup:

To cover the difference between the simulated pileup and data, the inelastic cross section is varied by 5% up and down and the varied versions of Fig. 4.3 is obtained. For the background the weights for up and down variations are propagated to the κ calculation. For the signal, due to the fact that the signal yields in main band signal region do not significantly depend on pileup, simulation samples are not reproduced for the high pileup environment. In Fig. 6.2 y-axis is calculated as follows:

$$PU_{\text{dependence}} = \frac{\text{Efficiency}_{\text{high}}}{\text{Efficiency}_{\text{low}}}, \quad (6.6)$$

where $\text{Efficiency}_{\text{high}}$ and $\text{Efficiency}_{\text{low}}$ is defined as in the Eqn. 6.7.

$$\text{Efficiency}_{\text{high}} = \frac{N(n_{\text{vertices}} \geq 20, MB)}{N(n_{\text{vertices}} \geq 20, \text{baseline})}, \quad (6.7)$$

$$\text{Efficiency}_{\text{low}} = \frac{N(n_{\text{vertices}} < 20, MB)}{N(n_{\text{vertices}} < 20, \text{baseline})}. \quad (6.8)$$

However, still an uncertainty has to be assigned to cover possible dependence. To drive

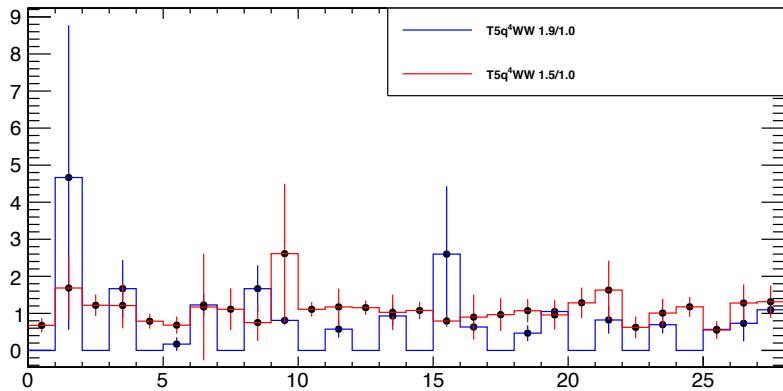


Figure 6.2: Pileup dependence of the two signal benchmark points. The blue line is representing the high mass gap point the red one is for low mass gap region. The histogram points are following a flat line around 1.

an uncertainty, which is covering the differences between the simulated signal pileup distribution in main band regions and data, following steps have been followed. First, the

signal sample is divided into a low and a high PU part according to mean value and then the mean values for each part is calculated as in Fig. 6.3 (middle). Second, the simulated efficiency in main band signal region is calculated for low and high pileup region as in Eqn. 6.7. Third, these four points, from first and second steps, together compose the two points in Fig. 6.3 (right). A linear fit is performed to extrapolate for all the pileup range. Finally, normalized data pileup distribution is folded with the fit and the sum is calculated. This procedure is repeated by varying the two values from second step within their statistical uncertainties independently. The relative difference to the central value is taken as uncertainty. Then, the uncertainty is separately obtained for low and high pileup region and combined in to one by taking the squared sum of the two. The entire procedure is repeated for each mass point in the gluino-neutralino mass plane and for each main band region. The resultant uncertainty is found to be around 5-40% depending on the statistics of the region for the corresponding mass point. In order not to double count the statistical errors, a 10% flat uncertainty is applied to all bins for all mass points.

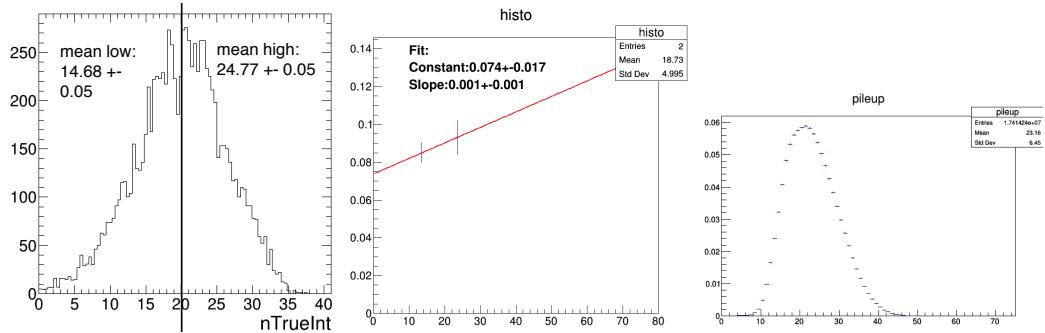


Figure 6.3: Distribution of pileup for the high-mass gap signal sample, divided in to two parts as low and high pileup region (left), example fit performed using the two points explained in the text (middle), data pileup distribution which is folded with the fit.

Fig. 6.4 displays the different systematic uncertainties on the background prediction for main band signal regions while Fig. 6.5 shows the systematic uncertainties on the simulated signal events for the mass point $m_{\tilde{g}} = 1900$ and $m_{\tilde{\chi}_1^0} = 100$. The total systematic uncertainty is calculated as the squared sum of all the different sources are shown with the black crosses and they are just for illustration. The total systematic uncertainty as well as statistical uncertainties is shown in the lower band of the two figures.

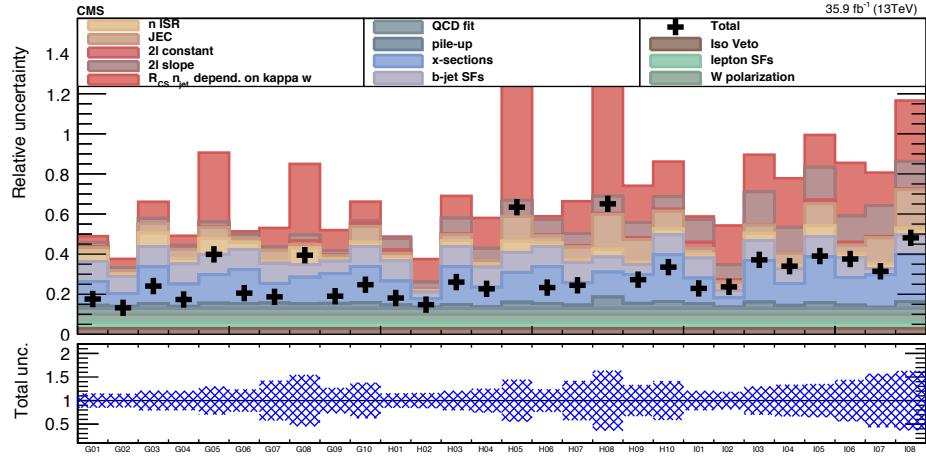


Figure 6.4: Visualization of all systematic uncertainties on the background prediction for main band regions described in Tab. 4.5.

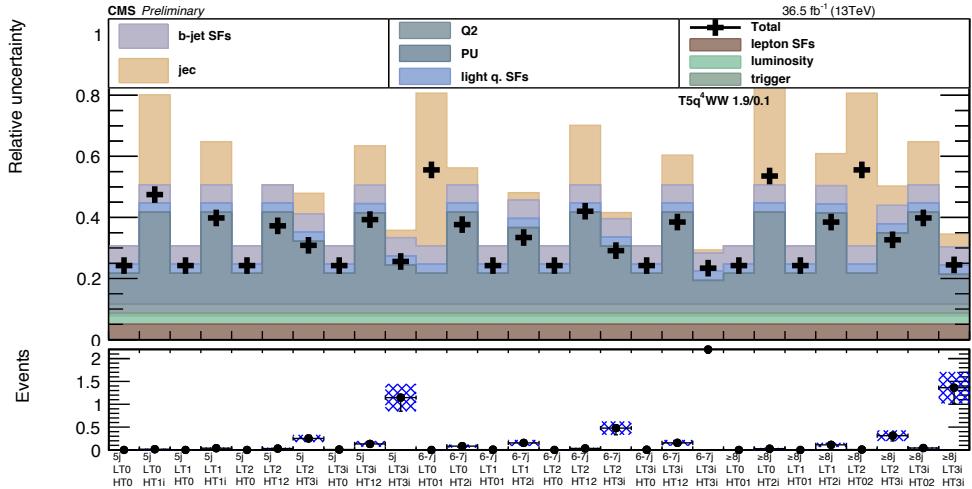


Figure 6.5: Visualization of all systematic uncertainties on the simulated signal yields in main band regions described in Tab. 4.5. Uncertainties are shown only for one mass point which is $m_{\tilde{g}} = 1900$ and $m_{\tilde{\chi}_1^0} = 100$.

Chapter 7

Results and Interpretation

Contents

7.1	Results of background prediction	80
7.1.1	Result of validation in side band regions	80
7.1.2	Background prediction in main band regions	81
7.2	Statistical interpretation	82
7.2.1	Frequentist limit setting procedure	84
7.3	Interpretation on simplified model T5qqqqWW	89
7.4	Comparison to other results	89

In this chapter the results of the analysis are presented. The chapter commence with the validity test (see Sec. 7.1.1) of the data-driven background prediction in a restricted region spared from $W + \text{jets}$ sideband. It proceeds with the comparison between the background prediction and the observed event yields in the mainband signal regions (see Sec. 7.1.2). Following an introduction to statistical procedure, the interpretation of results in the framework of simplified SUSY models are presented. The chapter then concludes with a discussion on the comparison of the results with other analysis.

7.1 Results of background prediction

7.1.1 Result of validation in side band regions

As mentioned in Sec. 5.5, the background estimation method is validated in regions where the n_{jet} selection is restricted to 4 jets and kinematic requirements in L_T and H_T is synched to the mainband regions. Accordinlgy, the sideband where R_{CS} of $W + \text{jets}$ events is measured is chosen to be 3 jets only region, which is necessary to avoid an overlap of sideband and mainband regions of validation. This restriction results in having a statistically limited sideband region, thus a reduced performance of the method. Nevertheless, it is necessary to validate the method before looking into data in mainband regions of the analysis.

On the other hand, the sideband region for measuring $t\bar{t} + \text{jets}$ R_{CS} remains unchanged. Examples of btag multiplicity fit, which are performed for the validation, can be seen in Fig. 7.1. Fig. 5.9 shows the κ values for the validation prediction. Results of this validation are shown in Fig. 7.2. The systematic uncertainties shown in this figure are an approximate projection of what have been observed in the mainband regions. Fig. 7.2 manifests that the data agrees with standard model prediction in the validation region.

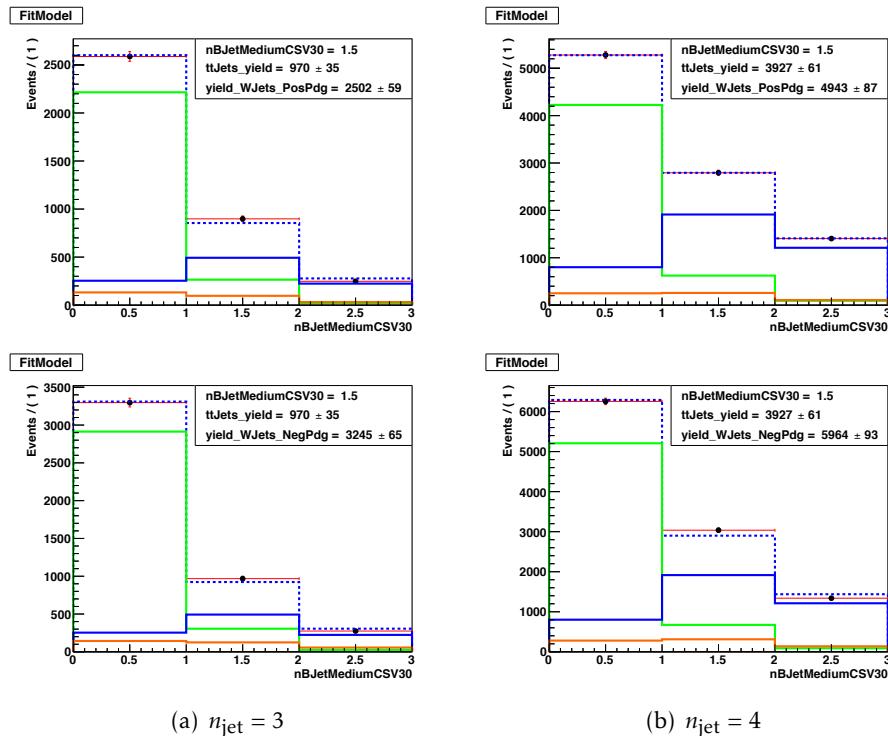


Figure 7.1: Fits to the $n_{\text{b-tag}}$ multiplicity for control regions in (a) $n_{\text{jet}} = 3$ ($250 \leq L_{\text{T}} < 350 \text{ GeV}$, $H_{\text{T}} \geq 500 \text{ GeV}$, $\Delta\phi < 1$, muon channel) and (b) $n_{\text{jet}} = 4$ ($250 \leq L_{\text{T}} < 350 \text{ GeV}$, $H_{\text{T}} \geq 500 \text{ GeV}$, $\Delta\phi < 1$) in data. The solid lines represent the templates scaled according to the fit result (blue for $t\bar{t} + \text{jets}$, green for $W + \text{jets}$, turquoise for QCD, and red for the remaining backgrounds), the dashed line shows the sum after fit, and the points with error bars represent data.

7.1.2 Background prediction in main band regions

Fig. 7.3 presents the observed event yields from data for the 28 main band bins, compared to the data-driven standard model background prediction (see Chapter 5). In the figure, the stacked colored histograms represent the prediction corresponding to the individual background contributions grouped into $t\bar{t} + \text{jets}$, $W + \text{jets}$ and EWK.

No significant deviation from the standard model prediction is observed.

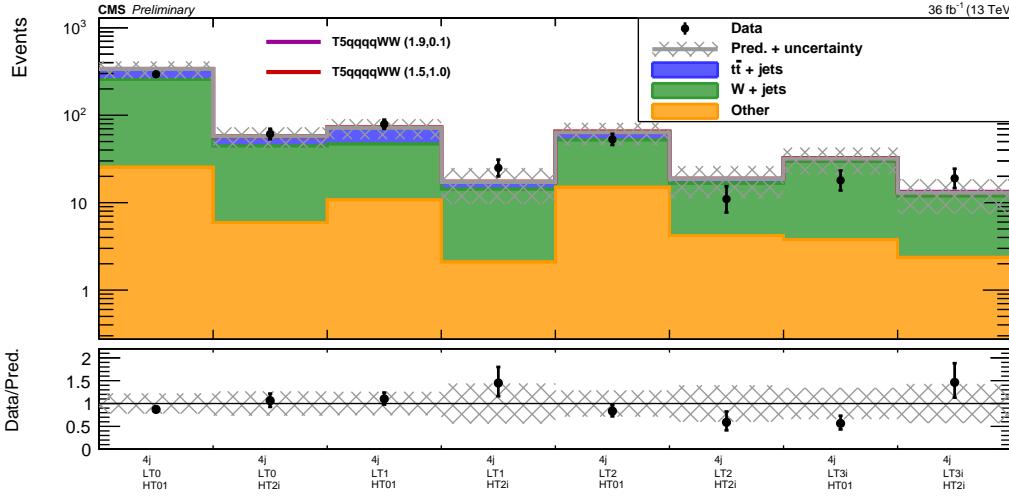


Figure 7.2: Validation of the background estimation method, using mainband regions with 4 jets. The shaded area reflects a rough estimate of the statistical and systematic uncertainties. The colored lines illustrate the expectations for two benchmark points of the T5qqqqWW model, showing the SUSY particle masses $m_{\tilde{g}}/m_{\tilde{\chi}^0}$ in TeV. The lower band shows the ratio between observed and predicted events.

A summary table of the background prediction and the final observed yields for each 28 bin can be found in Tab. ???. Additionally, all the components of the background prediction such as the data yields in the sideband SR and CR, predicted QCD background contamination and various κ factors are listed.

When performing the analysis, for the sake of blindness, number of observed data events has never been evinced in SRs (high $\Delta\phi$ region). In this point, the $\Delta\phi$ distributions in data and simulation are shown in Fig. 7.4 for the four L_T interval after the baseline selection with inclusive requirements on $n_{jet} \geq 5$ and $H_T \geq 500$ GeV. Overall, data and the standard model estimation from simulation is also in good agreement in this inclusive regions. Although, this agreement is not a necessity for the analysis flow, it confirms the reliability of the κ factors which are taken from simulation.

7.2 Statistical interpretation

Given the fact that no significant excess is observed, upper limits on signal cross sections in the $m_{\tilde{g}}-m_{\tilde{\chi}^0}$ plane are calculated. In this section the statistical framework used to derive limits to interpret the results which are summarized in Tab. ???. The method is designed such that it includes not only the mainband signal regions (MB SR) but also sideband and control regions. For this purpose, a mechanism based on ABCD method where each letter is representing MB SR, MB CR, SB SR and SB CR respectively, is used. In this way, ad-

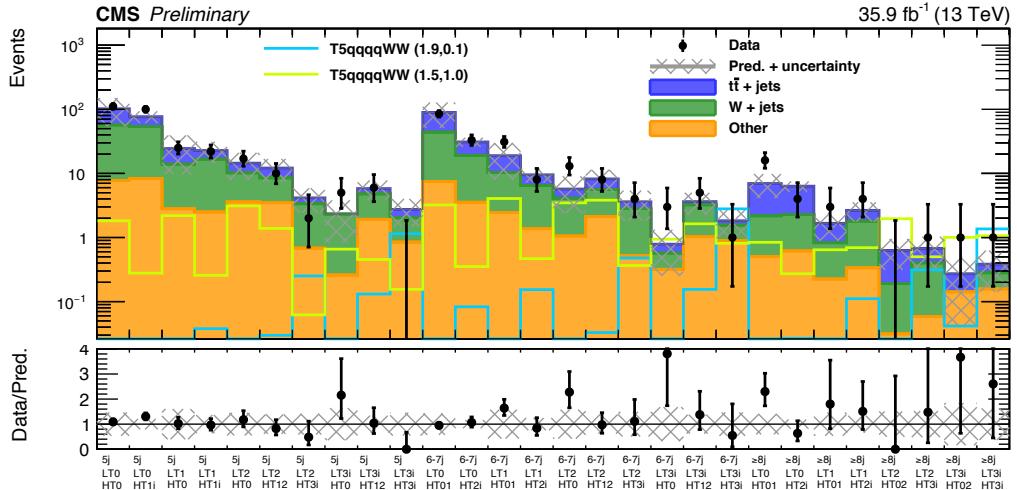


Figure 7.3: Observed data and predicted event yields (from data) in the 28 search regions are shown. The black points with error bars show the number of observed data events and corresponding statistical error. The filled, stacked histograms represent the predictions for $t\bar{t}+j$, $W+j$, and the remaining rare backgrounds. Lower panel shows the ratio of data to prediction and the shaded area reflects the total (stat. and syst.) relative uncertainty of the predicted background.

ditionally, potential signal contamination in sideband and control regions are taken into account. The limit procedure makes simultaneous use of two ABCD methods, called thus ABCDEF, since each of the $t\bar{t}+j$ and $W+j$ backgrounds is predicted with two parallel but separate R_{CS} methods. A summary of these regions can be found in Tab. 7.1.

Table 7.1: Conversion of R_{CS} regions to corresponding ABCDEF regions

	SR	CR
MB	A	B
SB_W	C	D
$SB_{t\bar{t}}$	E	F

The basic formulation of an one dimensional ABCD method is shown in Eqn.7.1, where QCD estimation is subtracted from CRs and possible residual differences between mainbands and sidebands are corrected with κ .

$$A = (B - B_{QCD}) \cdot \frac{C}{(D - D_{QCD})} \cdot \kappa. \quad (7.1)$$

On the other hand, in the present search, the formula representing the ABCDEF method is written as follows:

$$A = (B - B_{QCD}) \cdot f_W \cdot \frac{C - C_{t\bar{t}}}{(D - D_{t\bar{t}})} \cdot \kappa_W + (B - B_{QCD}) \cdot f_{t\bar{t}} \cdot \frac{E}{(F - F_{QCD})} \cdot \kappa_{t\bar{t}}^{DL-corr} \cdot \kappa_b. \quad (7.2)$$

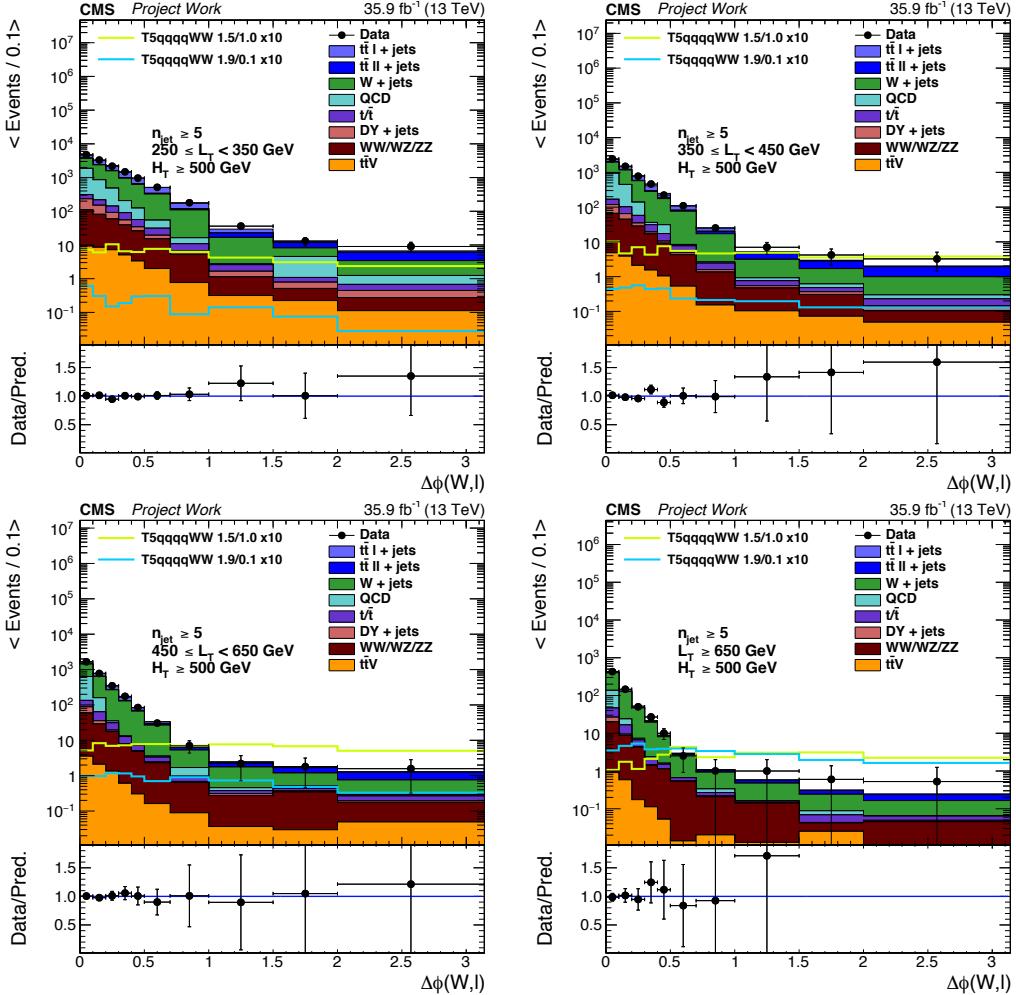


Figure 7.4: $\Delta\phi$ distributions in data which is represented with black points and simulation which is represented with color filled areas. Figure shows the four L_T interval after the baseline selection with inclusive requirements on $n_{jet} \geq 5$ and $H_T \geq 500$ GeV.

For deriving the limits, the Higgs combination tool which is extended with a new functionality to handle non-linear relations between different regions, is used, as needed for the ABCD approach [57].

Although taking a decision of setting limits is a Bayesian approach, the limit procedure for calculating exclusion curves and upper limits in the $m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ plane is based on the modified frequentist method.

7.2.1 Frequentist limit setting procedure

In the frequentist approach, the goal is to find the probability to observe certain data under a given signal hypothesis. It is not to find the probability of that the signal model is true

under the given observation.

Throughout this section, the expected event yields of SUSY model will be denoted as s_i , standard model background as b_i where index "i" indicates corresponding search bin. The systematic uncertainties are included as nuisance parameters θ , such that signal and background predictions can be written as functions of the nuisances: $s_i(\theta)$ and $b_i(\theta)$. In this search, all uncertainty sources are assumed to be independent or 100% correlated. The observed data events will be denoted as *data*. Finally, signal strength modifier is denoted as μ .

Likelihood function

The likelihood function, which is incorporating the ABCDEF regions for each search bin, is given by:

$$L = \prod_i^{ABCDEF} \text{Poisson}(data_i | b_i(\theta) + \mu \cdot s_i(\theta)) \times \prod_j^{\text{nuisances}} \text{Constraints}(\theta_j, \hat{\theta}_j), \quad (7.3)$$

where the six regions are involved with Poisson p.d.f.s for the observed number of events given the estimated signal+background, and the term Constraints refers to the nuisance parameters. Both signal and background predictions are subject to multiple uncertainties. Systematic uncertainties which are discussed in Sec. ?? and κ factors are included as constraints on the region A, while the QCD prediction is considered as constraints on control regions B and F. Additionally, the nuisance parameters include the effect of the statistical uncertainty of all 6 bins.

In the analysis, the entire information is inserted to the Limit Tool using text files, which are called datacards. An example datacard is shown in Fig. 7.5 for only one search region and the entries corresponding to rare backgrounds and QCD have not been shown for the simplicity. Examples tests on the nuisance parameters can be found in the Appendix C using the datacard corresponding to the high mass gap benchmark point.

Test statistic, \tilde{q}_μ

The compatibility of the data with the background-only and signal+background hypotheses, where the signal is allowed to be scaled by some factor μ is measured by constructing the following test statistic based on the profile likelihood ratio [58]:

$$\tilde{q}_\mu = -2 \ln \frac{L(data|\mu, \hat{\theta}_\mu)}{L(data|\hat{\mu}, \hat{\theta})}, \text{ for } 0 \leq \hat{\mu} \leq \mu, \quad (7.4)$$

where nominator and denominator are maximised separately. $\hat{\theta}_\mu$ is the conditional maximum given the signal strength modifier value μ . $\hat{\mu}$ and $\hat{\theta}$ are corresponding to the global maximum of the likelihood. $\hat{\mu}$ and $\hat{\theta}$ are free parameters so the denominator is independent of μ and serves as a normalisation term only. The condition $0 \leq \hat{\mu} \leq \mu$ ensures

Figure 7.5: Simplified example of a datacard file for one search region. The entries corresponding to rare backgrounds and QCD have been suppressed.

one-sided confidence intervals for upper limit tests while unphysical negative signals are avoided.

According to the Neyman-Pearson lemma [59], the ratio of likelihoods are providing the most powerful discrimination. However, this does not have to be true for a system with multiple free parameters, infact the existance of a uniformly most powerful test statistic is not certain. On the other hand, the likelihood ratio is probably one of the closest scenarios to the optimal and easy to use. Therefore, it is commonly used in experimental particle physics.

Observed $\tilde{q}_{\mu,\text{obs}}$ and $\hat{\theta}_{\mu,\text{obs}}$

The observed value of the test statistic ($\tilde{q}_{\mu,\text{obs}}$) and the nuisance parameter ($\hat{\theta}_{\mu,\text{obs}}$) for a given signal strength μ that maximize the likelihood function given in Eqn.7.3 are computed. The $\mu = 0$ case is the background only hypothesis while the non-zero μ values are representing the signal+background hypothesis. If one finds good agreement with data with a large μ , it would probably mean new physics.

Pdfs \tilde{q}_μ

To compute the test statistic pdfs $p_\mu(\tilde{q}_\mu|\mu, \hat{\theta}_\mu)$, generation of toy pseudo-MC samples with a signal strength of μ and $\mu = 0$ is required. This procedure requires to scan a range of μ values and is used in the statistical interpretation of Tevatron data. However, in case a complex system, the pseudo-MC mechanism can be computationally intensive. Therefore, generally at LHC the so-called asymptotic approximation is used to circumvent this computational challenge. The procedure involves usage of analytical formulas instead of toy MC.

Observed p-value for hypothesis μ

The observed p-value for hypothesis μ can be calculated as:

$$P(\mu) = \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} p_\mu(\tilde{q}_\mu|\mu, \hat{\theta}_\mu) d_{\tilde{q}_\mu}. \quad (7.5)$$

The probability of finding a value of test under the background only hypothesis at least as large as the one observed in data, $P(q_0 \geq q_{0,\text{obs}})$, can be used in quantifying an excess.

The p-value can be converted into significance by through one-sided normal distribution tail probability:

$$P = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx. \quad (7.6)$$

The asymptotic approximation mentioned above allows to estimate the significance, Z in Eqn. 7.6, directly from the observed test systematic as:

$$Z = \sqrt{q_{0,\text{obs}}}. \quad (7.7)$$

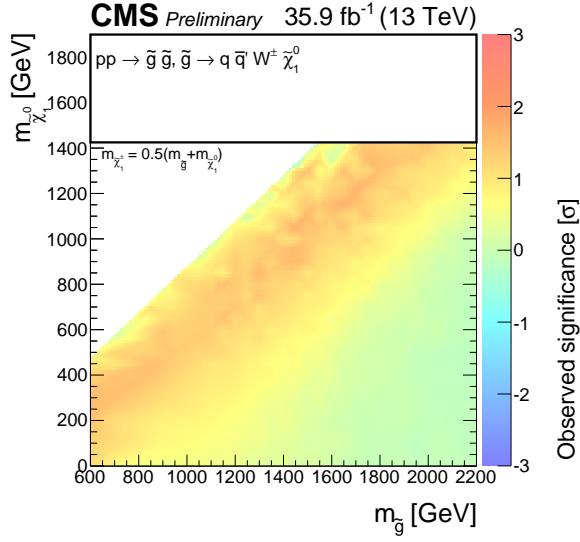


Figure 7.6: Observed significance for the model T5qqqqWW.

This is allowed due to the fact that the profile likelihood test statistic has the property of being distributed as a half χ^2 for one degree of freedom. In particle physics, to claim an observed signal as a "discovery", the significance should exceed $Z=5$, which is corresponding to a one-sided p-value of the back-ground only hypothesis of $2.9 \cdot 10^{-17}$. The observed significance, which is obtained by combining the main search regions, for the signal model T5qqqqWW is presented in Fig. 7.6. It can be seen that for all the mass points the significance is less than 3.

Upper limits

Upper limits are calculated at the 95% confidence level (C.L.) using the asymptotic CL_s criterion [?], where CL_s is defined as:

$$CL_s = \frac{CL_{s+b}}{CL_b}. \quad (7.8)$$

CL_{s+b} can be constructed using the probability of finding a value of q_μ at least as larger as the observed one under the signal+background hypothesis, $P(q_0 \geq q_{0,obs}|\mu)$, while CL_b can be costructed using the background only hypothesis, $P(q_0 \geq q_{0,obs}|\mu = 0)$. The 95% CL exclusion limits are defined as $CL_s(\mu) \leq 5\%$. Given the fact that ratio is also sensitive to downwards fluctuations of the background, this condition is more conservative than only considering signal+background hypothesis, i.e $CL_{s+b}(\mu) \leq 5$. Finally, μ is varied iteratively until $CL_s = 5\%$ and if μ is smaller than 1, a signal mass scenerio is excluded.

7.3 Interpretation on simplified model T5qqqqWW

As discussed earlier, the supersymmetric simplified models (see Sec.1.2.3) are used to interpret the results of the search. Because of the absence of any significant deviation from the data-based SM prediction, upper limits on the cross section of the T5qqqqWW model scan are derived. The procedure introduced in Sec.7.2 is followed to obtain 95% C.L. upper limits on the observed and median expected signal strength. The corresponding results are presented in Fig. 7.7 for the entire mass scan of the T5qqqqWW. The missing mass points, that can be seen in Fig.4.2, is compensated by interpolating between the neighboring μ values. The upper limit on the production cross section of the T5qqqqWW model can be obtained by multiplying the upper limit obtained on signal strength μ with the theoretical cross section. In Fig.7.8, color map shows the observed upper limit on the theoretical cross section, the black lines shows the corresponding exclusion curve and $\pm 1\sigma$ contours due to theoretical uncertainty on cross section. In the same figure, the red lines are the expected exclusion limit contours.

In the region where the mass difference between gluino and neutralino is high, the gluino and decay products have high energies. Therefore, kinematic regions with high hadronic and leptonic scales provide the largest signal sensitivity. In this region, observed limit follows the expected median and the T5qqqqWW simplified model scenarios with gluino mass up to 1.9 TeV for neutralino masses below 300 GeV are excluded.

In the region close to compressed mass scenarios, the upper limit on the cross section takes values up to 1 pb, where the gluino has very soft decay products resulting in a decrease in acceptance. The signal models with neutralino mass up to 950 GeV is excluded. To investigate the reason why the observed limit curve is significantly below the expected one a benchmark point with gluino mass 1.4 TeV and neutralino mass 1 TeV is chosen. It is found to be that the most sensitive bin which is $n_{\text{jet}} \leq 8$, $L_T:[450,650]$ GeV, $H_T:[500,1250]$ GeV does not show an excess. However, there are 6 bins with a similar sensitivity out of which three show light excesses. Especially, the bin with $n_{\text{jet}}:[6,7]$, $L_T:[450,650]$ GeV, $H_T:[500,750]$ GeV has 13 events observed while 5.7+-3.3 expected.

7.4 Comparison to other results

As discussed in the very first chapter in Sec. 1.2.4, the simplified SUSY model T5qqqqWW had already been studied in Run1 of LHC and the models with gluino mass up to 900 GeV was excluded [42]. The searches are repeated at the new center-of-mass energy and the first results are published with integrated luminosity of 2.3 fb^{-1} [60]. In the present analysis, the excursion on the gluino mass is extended 500 GeV for the lowest neutralino mass.

The results of new Run 2 analyses of CMS with 35.9 fb^{-1} are shown in the left plot of Fig. 7.10. In this thesis, the present analysis, which is labeled with SUS-16-042 ($\Delta\Phi$), is represented with the purple curve and published in Ref. [?]. The analysis covers the largest area in the high gluino-neutralino mass gap while another analysis, which is labeled with SUS-16-033, covers the largest area in the region close to compressed scenarios. This second analysis [?] performs a search in the full-hadronic final state, where both of the W

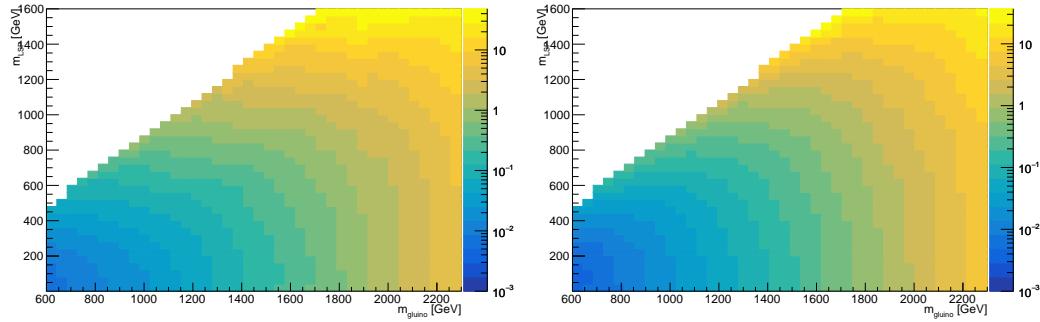


Figure 7.7: 95% CL Observed (Expected) upper limits on the signal strength modifier μ for the T5qqqqWW model is shown at the left (right) side as a function of the gluino and LSP masses.

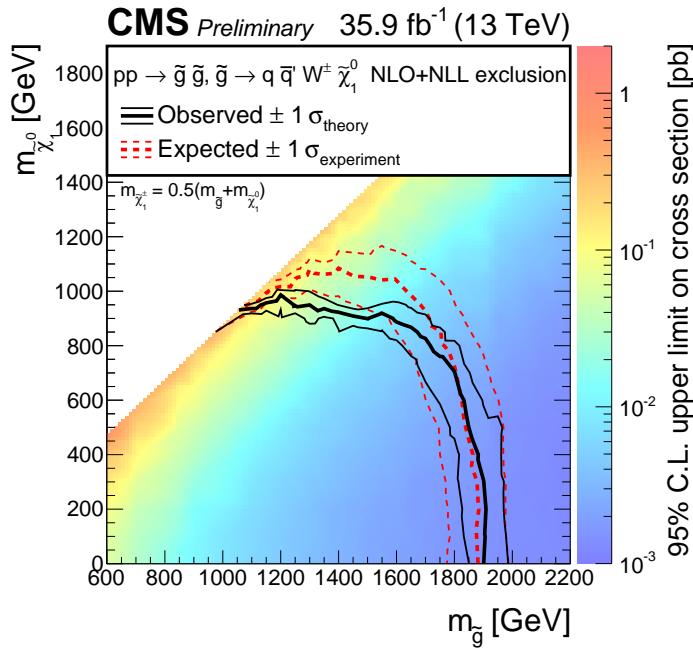


Figure 7.8: Cross section limits at a 95% C.L. for the T5qqWWWW model, and as a function of the gluino and LSP masses. The solid black (dashed red) lines correspond to the observed (expected) mass limits, with the thicker lines representing the central values and the thinner lines representing the $\pm 1\sigma$ uncertainty bands related to the theoretical (experimental) uncertainties

bosons decay hadronically, using the missing H_T variable. The missing H_T is calculated very similar to the \cancel{E}_T but using only the transverse momentum of jets.

An overall summary of the SUSY results analysed by CMS can be found at the bottom plot of the Fig.7.9. The figure shows the best exclusion limits on sparticle masses. Almost all the analyses, which are setting limits on gluino mass, have very competitive results reaching up to 2 TeV.

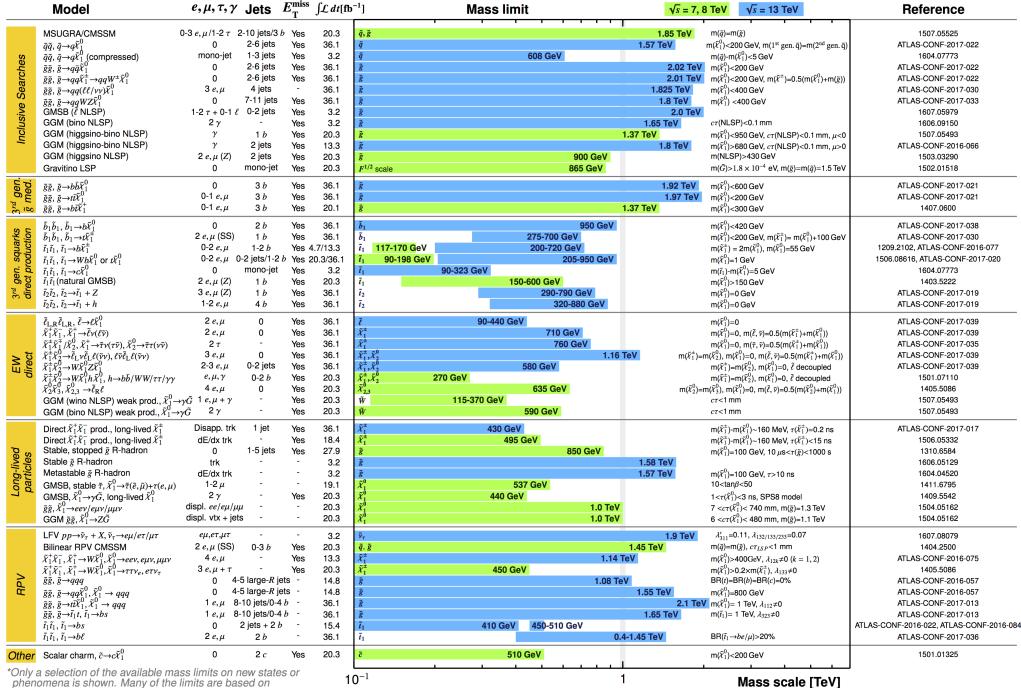
The Atlas collaboration performed also several searches for gluino pair production. The results for the topologies similar to the T5qqqqWW of CMS can be seen in the right plot of Fig. 7.10 where the gray shaded area shows the Run 1 results. Furthermore, a summary of all SUSY models studied in ATLAS is shown at the top in Fig. 7.9.

ATLAS SUSY Searches* - 95% CL Lower Limits May 2017

ATLAS Preliminary

$\sqrt{s} = 7, 8, 13 \text{ TeV}$

Reference



*Only a selection of the available mass limits on new states or phenomena is shown. Many of the limits are based on simplified models. c.f. refs. for the assumptions made.

Selected CMS SUSY Results* - SMS Interpretation

ICHEP '16 - Moriond '17

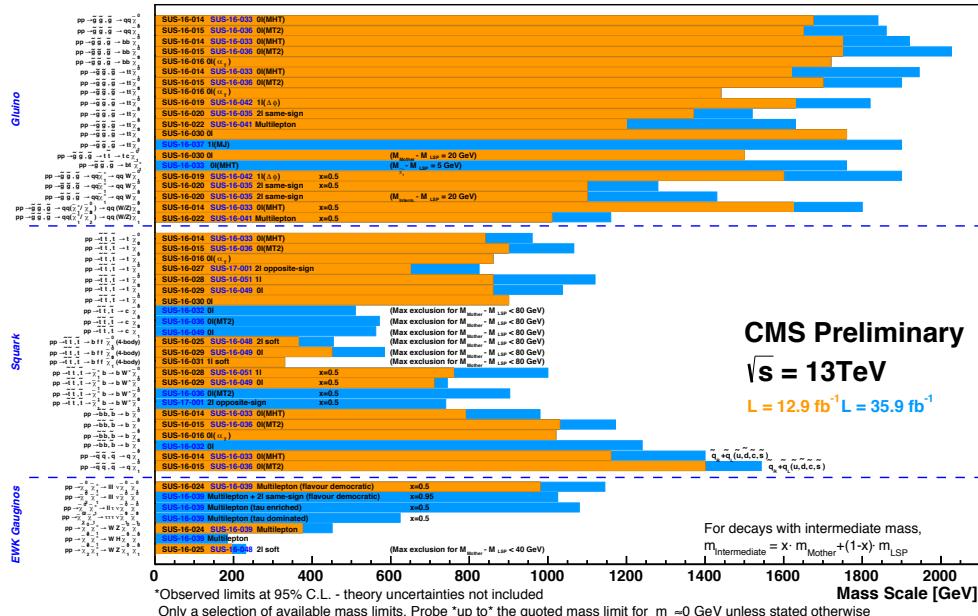


Figure 7.9: Best exclusion limits on sparticle masses from searches for SUSY using SMS by ATLAS (top) and CMS (bottom) collaborations [40, 41]. In the upper plot, the green bars represent the 7-8 TeV results while the blue bars are summarizing the 13 TeV results. In the lower plot, the orange bars represent the 7-8 TeV results while the blue bars are summarizing the 13 TeV results.

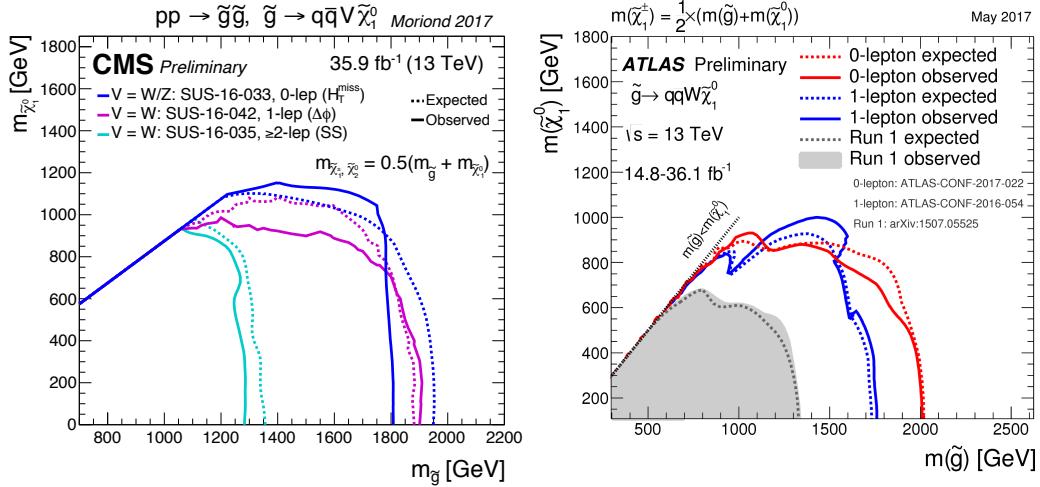


Figure 7.10: 95% C.L. exclusion limit curves on the simplified SUSY model of gluino pair production with subsequent decay of $\tilde{g} \rightarrow q\bar{q}WW\tilde{\chi}_1^0$ obtained by the CMS (left) and ATLAS (right) collaborations. The present analysis is corresponding to the result labeled with SUS-16-042.

Chapter 8

Conclusion

The Standard Model (SM) of the particle physics continues to be a successful theory in explaining the physical phenomenon observed in LHC. However, the SM fails to elucidate several experimental facts that give rise to questions regarding the origin of our universe such as: What is Dark Matter and how gravity can be included?

The supersymmetric extension of the SM (SUSY) is one of the most appealing beyond the SM (BSM) theories that provide solutions to some of these questions. SUSY relates the fermions to bosonic superpartners and bosons to fermionic superpartners. This characteristic allows SUSY to reduce the divergent corrections on the Higgs mass. Furthermore, the R-parity conserved SUSY models provide a cold-dark matter candidate in the form of lightest SUSY particles.

In this thesis, a search for a simplified SUSY model is performed in events with single lepton final states, multiple jets and none of them zero b-tagged. In the targeted model, each of the pair produced gluinos decays to neutralino, which is considered to be stable lightest supersymmetric particle, together with several light quark jets and one electron or muon. The intermediate chargino mass in this decay chain is taken to be half way between gluino and neutralino. The two neutralinos and a neutrino coming from leptonic W boson decay combine a large \cancel{E}_T . Moreover, the high multiplicity of jets leads to large hadronic activity. Therefore the model has high sensitivity in the tails of the kinematic distributions.

The main search variable of the analysis is the azimuthal angle between the lepton and reconstructed W boson, $\Delta\phi(W, \ell)$. The leading background processes tend towards low values of the angle while the expected signal events show a flat distribution, due to the large missing transverse energy contribution from LSP. Thus, the region with high (low) values of this angle is chosen to be signal (control) region. To further increase the sensitivity several signal rich search regions are defined, based on the number of jets, the hadronic scale (H_T), and the leptonic scale (L_T).

The Standard Model background is predicted with a robust data-driven approach, so-called R_{CS} . The method uses the control regions to estimate the normalization in signal regions and low jet multiplicity sidebands to obtain signal to control region transfer factor. The main SM backgrounds $W + \text{jets}$ and $t\bar{t} + \text{jets}$ are predicted separately. A separate estimation method is developed for QCD multijets events. This method uses the polarization information of leptons from W boson decays which manifests itself in the L_p distributions. The predicted QCD multijets event counts are subtracted from the control regions. All of the methods are validated in simulation and data.

Various sources of systematic uncertainties on the background prediction are computed. The largest contributions are from possible residual dependencies of R_{CS} on the jet multiplicity and potential mismodeling of the dileptonic fraction of $t\bar{t}$ events. Additionally, the systematic uncertainties coming from the simulated signal events are studied and their effects are propagated. In this case, the largest uncertainties arise from jet energy correction uncertainties and potential mismodeling of pile up. Other uncertainties are at the 5% level or lower.

Finally, the prediction is compared to data recorded by the CMS experiment during the 2016 run of proton-proton collisions of LHC at 13 TeV center of mass energy. The dataset corresponds to an integrated luminosity of 35.9 fb^{-1} . No deviation from the SM background is observed. Therefore, upper limits on gluino production cross section are obtained in the context of the investigated simplified model. 95% C.L. are calculated with the asymptotic formulae using the CL_s criterion. As a result, gluino masses below 1.9 TeV are excluded for neutralino masses below 300 GeV.

After the latest with 35.9 inverse femtobarns of 13 TeV data still results do not favor the SUSY models. The minimal supersymmetric models become more and more constraint. As the limits on the gluino mass approach to two TeV scales, the natural SUSY become more restricted. Given the variety of possibilities, completely excluding SUSY is difficult. The SUSY searches can be expended towards the direction of long-lived particles with displaced vertex analyses, while there has been still a place for natural SUSY in the compressed regions.

Although no significant deviation from the Standard Model has been observed, the journey of particle physics has not completed yet. The unanswered questions from the observation of astrophysical phenomena remains as they are as long as an evidence of a more universal model is not observed.

Over the last decades, the experimental searches have been performed under the guidance of theoretical physics methods. The success of this scientific path is proven with the Higgs

boson discovery and the gravitational waves. It is also possible that a sudden unexpected discovery can illuminate the theoretical physics path.

Appendix A

MC Samples and cross sections

Sample name	Full name	cross section [pb]
TTJets.SingleLeptonFromTbar	/TTJets.SingleLeptFromTbar.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*(3*0.108)*(1-3*0.108)
TTJets.SingleLeptonFromTbar.ext	/TTJets.SingleLeptFromTbar.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*(3*0.108)*(1-3*0.108)
TTJets.SingleLeptonFromT	/TTJets.SingleLeptFromT.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*(3*0.108)*(1-3*0.108)
TTJets.SingleLeptonFromT.ext	/TTJets.SingleLeptFromT.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*(3*0.108)*(1-3*0.108)
TTJets.DiLepton	/TTJets.DiLept.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*((3*0.108)**2)
TTJets.DiLepton.ext	/TTJets.DiLept.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*((3*0.108)**2)
TTJets.LO_HT600to800	/TTJets.HT-600to800.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.610*831.76/502.2
TTJets.LO_HT600to800.ext	/TTJets.HT-600to800.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.610*831.76/502.2
TTJets.LO_HT800to1200	/TTJets.HT-800to1200.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.663*831.76/502.2
TTJets.LO_HT800to1200.ext	/TTJets.HT-800to1200.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.663*831.76/502.2
TTJets.LO_HT1200to2500	/TTJets.HT-1200to2500.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.12*831.76/502.2
TTJets.LO_HT1200to2500.ext	/TTJets.HT-1200to2500.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.12*831.76/502.2
TTJets.LO_HT2500toInf	/TTJets.HT-2500toInf.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.001430*831.76/502.2
TTJets.LO_HT2500toInf.ext	/TTJets.HT-2500toInf.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.001430*831.76/502.2

Table A.1: List of simulated $t\bar{t} + \text{jets}$ background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.

Sample name	Full name	cross section [pb]
WJetsToLNu_HT100to200	/WJetsToLNu_HT-100To200.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1345*1.21
WJetsToLNu_HT100to200.ext	/WJetsToLNu_HT-100To200.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1345*1.21
WJetsToLNu_HT200to400	/WJetsToLNu_HT-200To400.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	359.7*1.21
WJetsToLNu_HT200to400.ext2	/WJetsToLNu_HT-200To400.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	359.7*1.21
WJetsToLNu_HT200to400.ext	/WJetsToLNu_HT-200To400.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	359.7*1.21
WJetsToLNu_HT400to600	/WJetsToLNu_HT-400To600.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	48.91*1.21
WJetsToLNu_HT400to600.ext	/WJetsToLNu_HT-400To600.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	48.91*1.21
WJetsToLNu_HT600to800	/WJetsToLNu_HT-600To800.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	12.05*1.21
WJetsToLNu_HT600to800.ext	/WJetsToLNu_HT-600To800.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	12.05*1.21
WJetsToLNu_HT800to1200	/WJetsToLNu_HT-800To1200.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.501*1.21
WJetsToLNu_HT800to1200.ext	/WJetsToLNu_HT-800To1200.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.501*1.21
WJetsToLNu_HT1200to2500	/WJetsToLNu_HT-1200To2500.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.329*1.21
WJetsToLNu_HT1200to2500.ext	/WJetsToLNu_HT-1200To2500.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.329*1.21
WJetsToLNu_HT2500toInf	/WJetsToLNu_HT-2500toInf.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.03216*1.21
WJetsToLNu_HT2500toInf.ext	/WJetsToLNu_HT-2500toInf.TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.03216*1.21

Table A.2: List of simulated $W + \text{jets}$ background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.

Sample name	Full name	cross section [pb]
TToLeptons.sch	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo_pythia8_TuneCUETP8M1	(7.20+4.16)*0.108*
T_tch_powheg	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin_pythia8_TuneCUETP8M1	136.02
TBar_tch_powheg	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin_pythia8_TuneCUETP8M1	80.95
TBar_tWch_ext1	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
TBar_tWch_ext2	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
TBar_tWch	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
T_tWch_ext1	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
T_tWch_ext2	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
T_tWch	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
DYJetsToLL_M50_HT100to200	/DYJetsToLL_M_50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	147.4*1.23
DYJetsToLL_M50_HT100to200_ext	/DYJetsToLL_M_50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	147.4*1.23
DYJetsToLL_M50_HT200to400	/DYJetsToLL_M_50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	40.99*1.23
DYJetsToLL_M50_HT200to400_ext	/DYJetsToLL_M_50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	40.99*1.23
DYJetsToLL_M50_HT400to600	/DYJetsToLL_M_50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	5.678*1.23
DYJetsToLL_M50_HT400to600_ext	/DYJetsToLL_M_50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	5.678*1.23
DYJetsToLL_M50_HT600to800	/DYJetsToLL_M_50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	1.367*1.23
DYJetsToLL_M50_HT800to1200	/DYJetsToLL_M_50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	0.6304*1.23
DYJetsToLL_M50_HT1200to2500	/DYJetsToLL_M_50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	0.1514*1.23
DYJetsToLL_M50_HT2500toInf	/DYJetsToLL_M_50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	0.003565*1.23
QCD_HT300to500	/QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	351300
QCD_HT300to500_ext	/QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	351300
QCD_HT500to700	/QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	31630
QCD_HT500to700_ext	/QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	31630
QCD_HT700to1000	/QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	6802
QCD_HT700to1000_ext	/QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	6802
QCD_HT1000to1500	/QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	1206
QCD_HT1000to1500_ext	/QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	1206
QCD_HT1500to2000	/QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	120.4
QCD_HT1500to2000_ext	/QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	120.4
QCD_HT2000toInf	/QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	25.25
QCD_HT2000toInf_ext	/QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	25.25
WWTo2L2Nu	/WWTo2L2Nu_13TeV-powheg	10.481
WWToLNuQQ	/WWToLNuQQ_13TeV-powheg	43.53
WWToLNuQQ_ext	/WWToLNuQQ_13TeV-powheg	43.53
ZZTo2L2Nu	/ZZTo2L2Nu_13TeV_powheg_pythia8	0.564
ZZTo2L2Q	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.28
ZZTo2Q2Nu	/ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8	4.04
ZZTo2L2Nu	/ZZTo2L2Nu_13TeV_powheg_pythia8	0.564
WZTo1L3Nu	/WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	(47.13)*(3*0.108)*(0.2)
WZTo1L1Nu2Q	/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
WZTo2L2Q	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.60
WWDouble	/WWTo2L2Nu_DoubleScattering_13TeV_pythia8	1.64
TTWToLNu_ext	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin_pythia8	0.2043
TTWToLNu_ext2	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin_pythia8	0.2043
TTWToQQ	/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin_pythia8	0.40620
TTZToQQ	/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo_pythia8	0.5297
TTZToLNuNu	/TTZToLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo_pythia8	0.2529
TTZToLLNuNu_m1to10	/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM_pythia8	0.0493

Table A.3: List of the rest of simulated background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.

Appendix B

Control Plots

Figures B.1 to B.6 show distributions of observables in sidebands at low n_{jet} , separately for electrons and muons. Figures B.1 and B.2 present the sideband defined by $L_T > 250\text{GeV}$, $H_T > 500\text{GeV}$, and $3 \leq n_{\text{jet}} \leq 4$, used in the estimation of the $W + \text{jets}$ background. Figures B.3 and B.4 show the sideband defined by $L_T > 250\text{GeV}$, $H_T > 500\text{GeV}$, $n_{\text{b-jet}} \geq 1$, and $4 \leq n_{\text{jet}} \leq 5$, used in the estimation of the $t\bar{t} + \text{jets}$ background. Distributions in the mainband are shown in Figures B.5 and B.6.

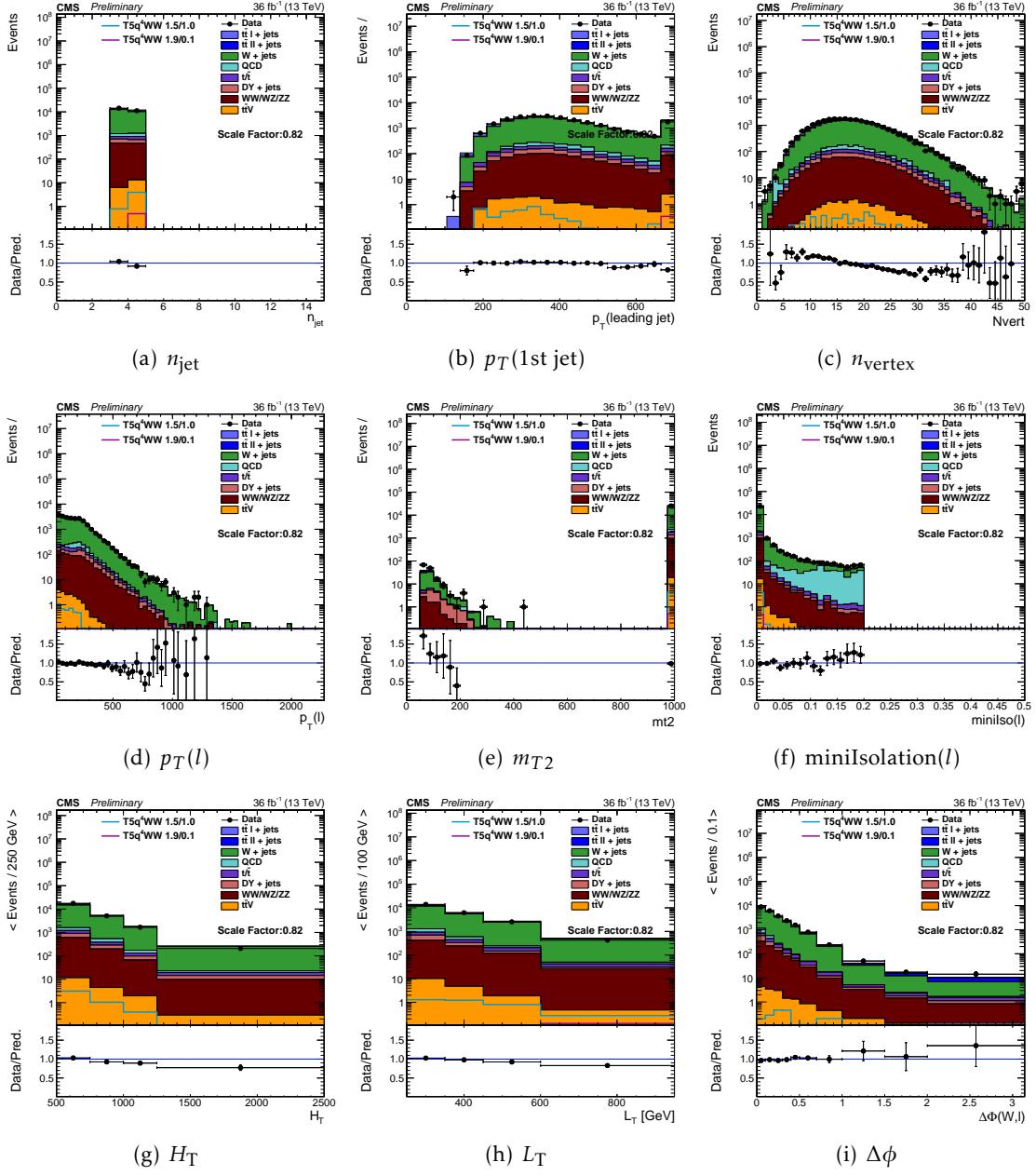


Figure B.1: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, 3 ≤ jets ≤ 4 and zero b-tagged jets (1 μ channel).

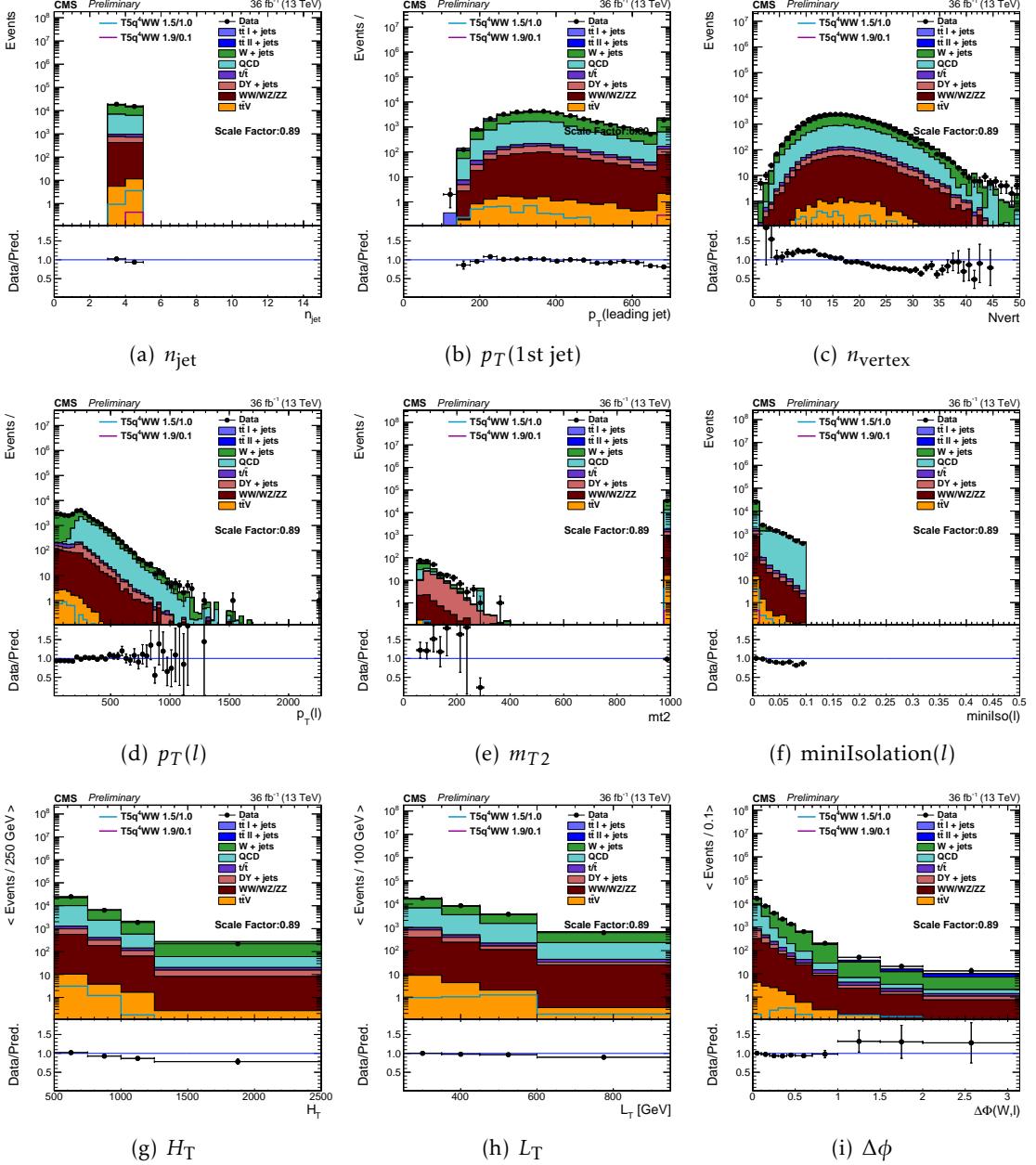


Figure B.2: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $3 \leq \text{jets} \leq 4$ and zero b-tagged jets (1 e channel).

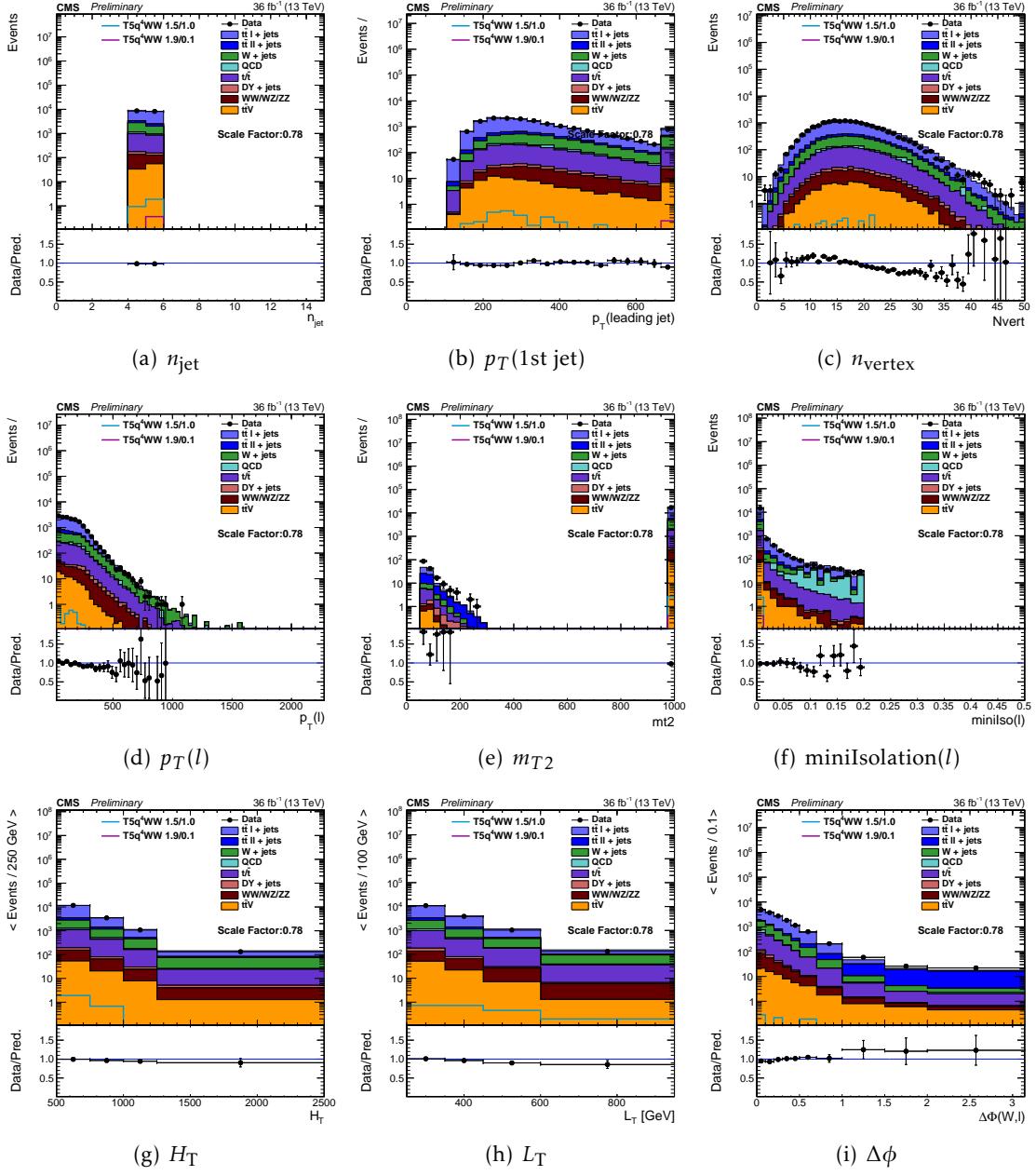


Figure B.3: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $4 \leq \text{jets} \leq 5$ and b-tagged jets (1 μ channel).

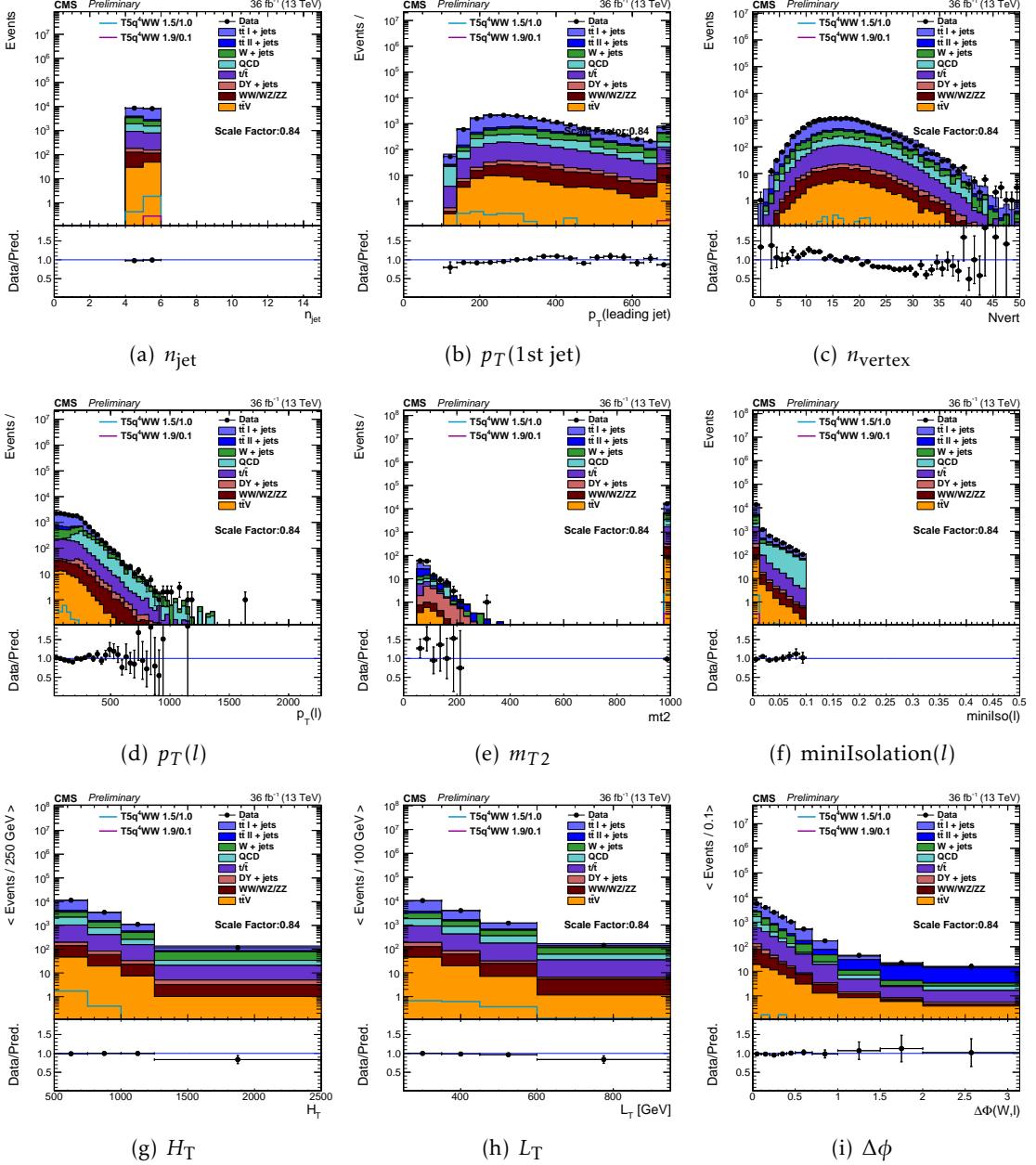


Figure B.4: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $4 \leq \text{jets} \leq 5$ and b-tagged jets (1 e channel).

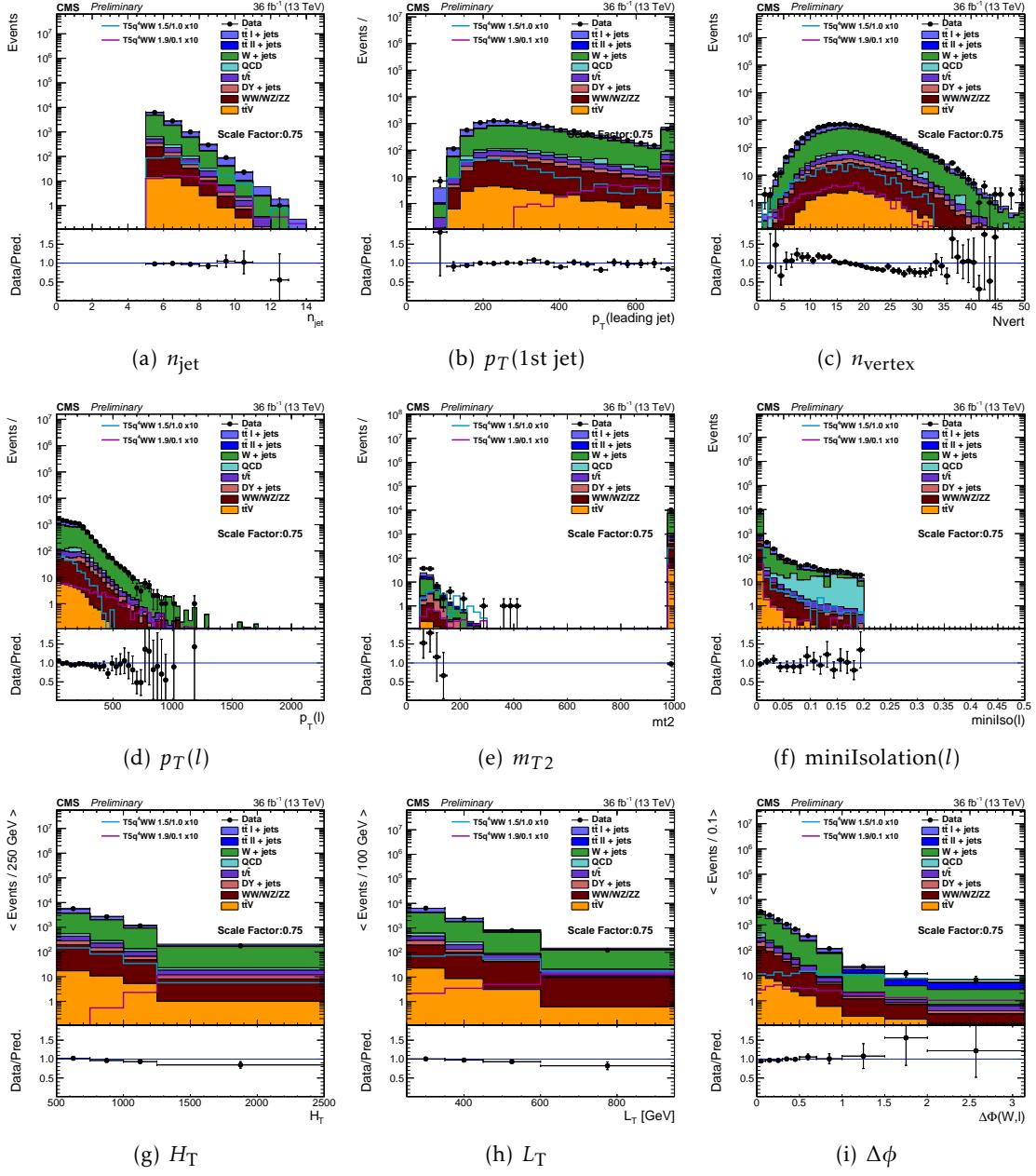


Figure B.5: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, ≥ 5 jets and zero b-tagged jets (1 μ channel).

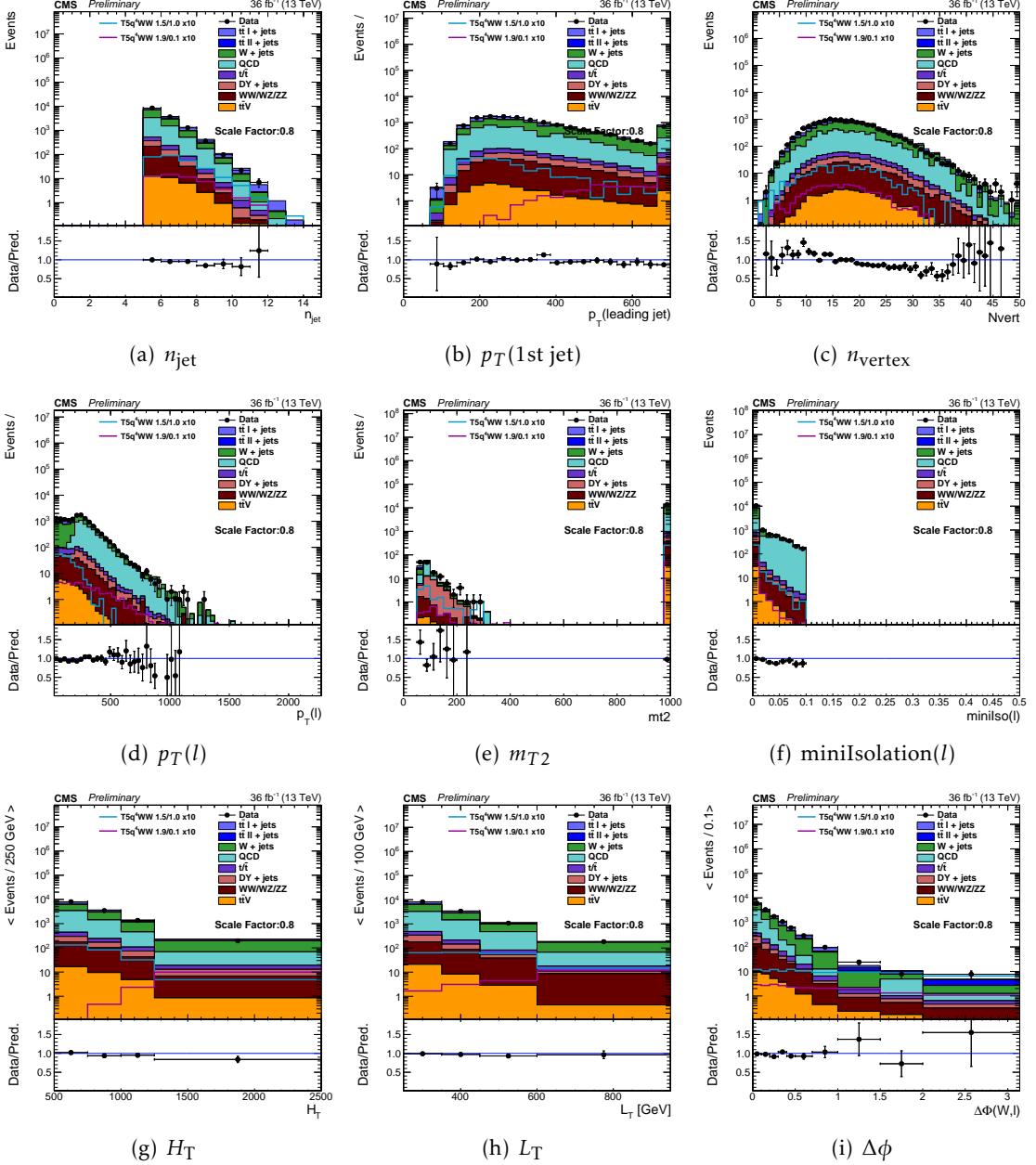
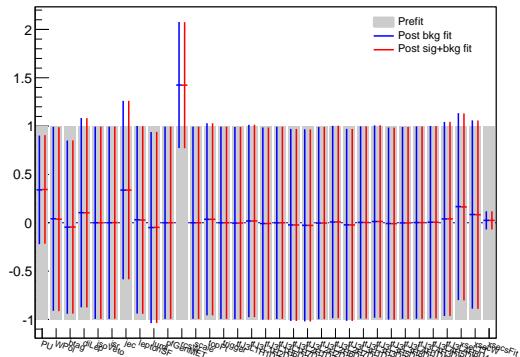


Figure B.6: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, ≥ 5 jets and zero b-tagged jets (1 e channel).

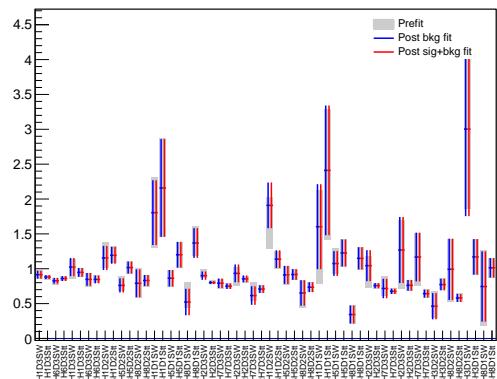
Appendix C

Statistical Tests

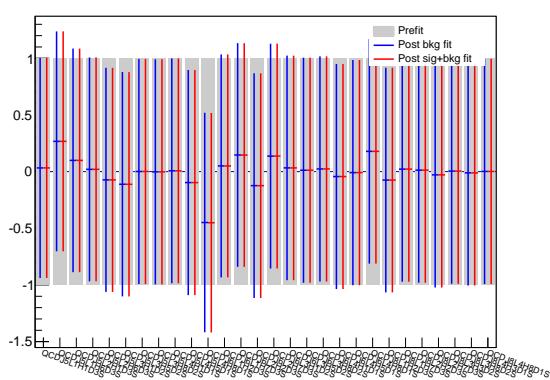
In this appendix, the output of the maximum likelihood fit to the data, in particular the effect of the different nuisance parameters is investigated. The tests are performed with the datacard of the mass point T5qqqqWW 1900/100.



(a) $\ln N$ constraints on region A

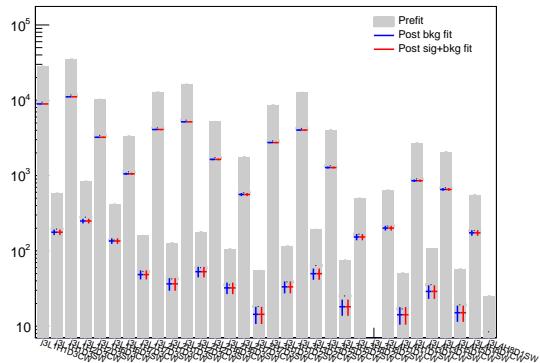


(b) κ parameters

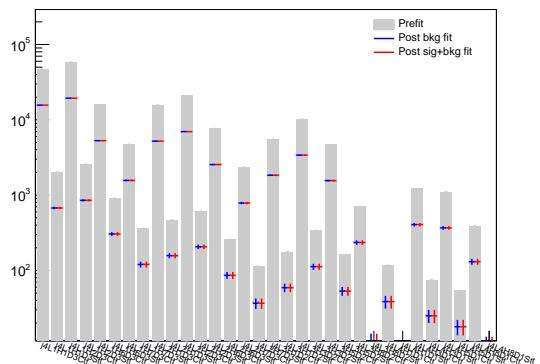


(c) QCD estimate

Figure C.1: Prefit (grey), s+b postfit (red) and, b-only postfit (blue) values of nuisance parameters included in the fit.



(a) nuisances in W Side Band



(b) nuisances in $t\bar{t}$ Side Band

Figure C.2: Prefit (grey), s+b postfit (red) and, b-only postfit (blue) values of nuisance parameters in side band regions included in the fit.

List of Figures

1.1	All fundamental particles of the standard model, the gauge and Higgs bosons are shown in the diagram. The electric charge, spin and mass (or limit for neutrinos) are given in the corners of the boxes [5].	3
1.2	An overview of characteristics of the experimental tests of the SM is shown in the table. The information is mainly collected from “the Electroweak model and constraints on new physics” section in PDG2017 [5]. Additionally further information on g-2 measurements can be found from the paper published by Muon G-2 experiment [19].	8
1.3	The running couplings of U(1), SU(2), SU(3) is denoted by $\alpha_1, \alpha_2, \alpha_3$ respectively. The figure represents the evolution of inverse of couplings with energy scale, for (a) non-supersymmetric SU(5) and (b) supersymmetric SU(5). The figure is taken from [29] pg. 282.	10
1.4	One-loop quantum corrections to the Higgs squared mass parameter m_H^2 , due to (left) a Dirac fermion f, and (right) a scalar S [4] p.g. 3	10
1.5	Diagrams showing the simplified models T1tttt (left) and T5qqqqWW (right).	17
1.6	Feynman Diagrams showing the gluino pair production. SUSY particles are shown in blue.	18
1.7	left: Cross sections of gluino pair production for four different mass points are given as a function of the center of mass energy. right: Cross section of gluino pair production at 13 TeV center of mass energy is given as a function of gluino mass. [37]	18
1.8	Cross section limits at a 95% CL for (left) the T1tttt model, and (right) the T5qqqqWW model [60].	19
1.9	Cross section of SM processes as a function of center of energy. [38]	20
2.1	CERN accelerator complex [61]:	23
2.2	Schematic view of the CMS detector with a cut-out quadrant [72]	26

3.1	Cross section of a slice of the CMS detector in the transverse plane with different particles crossing the subsystems is shown [89]. Only energy deposits in the most relevant subdetectors are presented for simplicity.	33
3.2	PF candidate multiplicity fits for \mathbb{E}_x and \mathbb{E}_y in different η regions.	39
4.1	Diagram showing the simplified model T5qqqqWW.	41
4.2	Produced signal mass points for the simplified T5qqqqWW model.	44
4.3	The normalised distributions of mean number of interactions per bunch crossing during the analysed data Run2016 (black) and in the MC samples (blue). And the red distribution represents the Data to MC ratio. For the data distribution the latest luminosity calibrations [48] and an inelastic pp cross-section of 69.2 mb are used.	46
4.4	Distributions of M_{T2} for events with electron or muon veto tracks (a) and hadronic veto tracks (b), for dileptonic $t\bar{t}$ (blue dotted) and T5qqqqWW (purple and azure) signal samples. The highest bin is always an overflow bin. The majority $t\bar{t}$ events have $M_{T2} < 80$ GeV, while the signal events have longer tails in M_{T2} . The rightmost plot (c) shows the distribution for all events.	47
4.5	The number of events for each sample for the electron (left) and muon (right) channel. Samples are scaled to the luminosity with their cross section and scale factors discussed in section 4.2.1.	49
4.6	Measurement of the trigger efficiency as a function of L_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) on the right [46].	51
4.7	Measurement of the trigger efficiency as a function of H_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) on the right [46].	51

4.8	Measurement of the trigger efficiency as a function of lepton p_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) is shown on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) is shown on the right. The lepton p_T trigger efficiency is measured after an $L_T > 250$ GeV cut which is the analysis baseline selection [46].	52
4.9	The top left distribution shows the number of b-tagged jets, after the baseline selection, requiring at least five jets, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton with $p_T > 25$ GeV. The rest of the distributions are plotted after the same baseline selection additionally requiring no b-tagged jets. In the top row the number of jets (right) while in the bottom row H_T (left) and L_T (right) distributions are shown. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked. The model T5qqqqWW (1.5,1.0) (T5qqqqWW (1.9,0.1)) corresponds to a gluino mass of 1.5 TeV (1.9 TeV) and neutralino mass of 1.1 TeV (0.1 TeV), respectively. The intermediate chargino mass is fixed at 1.25 TeV (1.0 TeV). The two benchmark signal models are scaled up by a factor of 10.	54
4.10	The $\Delta\phi$, after the baseline selection, requiring at least five jets and none of b-tagged, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton (electron or muon) with $p_T > 25$ GeV. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked. The model T5qqqqWW (1.5,1.0) (T5qqqqWW (1.9,0.1)) corresponds to a gluino mass of 1.5 TeV (1.9 TeV) and neutralino mass of 1.1 TeV (0.1 TeV), respectively. The intermediate chargino mass is fixed at 1.25 TeV (1.0 TeV).	55
4.11	2D distributions of event counts for the main background samples ($t\bar{t}$ + jets + W + jets) (top), and the signal T5qqqWW(1900,100) (left) and T5qqqWW(1500,1000) in the the $\Delta\phi$ vs. L_T plane after the preselection.	56
4.12	The distributions on top represent T5qqqqWW signal counts (left) and selection efficiency (right) after the baseline requirements in the gluino-lsp plane. The distributions at the bottom show the efficiency of signal region selection with respect to the baseline; $\Delta\phi > 1$ (left) and $\Delta\phi > x$ (right) where x stands for the cut value changing according to L_T bin.	57

4.13 Simulated single muon event yields for the background processes are shown as color filled stacked histograms for all 28 search bins. The two signal benchmark models are overlayed and shown by line histograms. The upper plot shows the high $\Delta\phi$ regions (SR) while the lower plot shows the low $\Delta\phi$ regions (CR).	59
4.14 Simulated single electron event yields for the background processes are shown as color filled stacked histograms for all 28 search bins. The two signal benchmark models are overlayed and shown by line histograms. The upper plot shows the high $\Delta\phi$ regions (SR) while the lower plot shows the low $\Delta\phi$ regions (CR).	60
5.1 L_P shape fit result for $3 \leq n_{\text{jet}} \leq 4$ and $n_{\text{b-jet}} = 0$ in $250 \leq L_T \leq 350$ bin (left). Ratio of selected to anti-selected electron events from QCD for $3 \leq n_{\text{jet}} \leq 4$ and $n_{\text{b-jet}} = 0$, in bins of LT in data (right).	64
5.2 b-tag multiplicity fit is performed in control regions. 3-4 jets sideband (left), 8 jets main band (right). Upper plots are showing the negative charged lepton selection while lower plots show the positive charged selection. Black points represent data, green is for $W + \text{jets}$, blue is for $t\bar{t} + \text{jets}$, orange is for other EWK backgrounds and cyan colored lines show QCD contamination. . .	64
5.3 RCS as a function of njet for dileptonic (left), and signle leptonic (right) $tt + \text{jets}$ events.	65
5.4 R_{CS} as a function of njet from simulation of $t\bar{t} + \text{jets}$ events (left), and data (right) in $t\bar{t} + \text{jets}$ sideband.	66
5.5 Jet multiplicity distribution after the single lepton baseline event selection (in the control region) with $H_T > 500$ GeV, $L_T > 250$ GeV, and $n_{\text{b-tag}} = 0$ (left) and in the dileptonic control region with recalculated H_T and L_T cuts (right). Only nISR reweighting is applied as scale factors.	68
5.6 Changes to $\kappa_{t\bar{t}}$ due to applying the double-ratio as given by Eqn. 5.9 and reevaluating $\kappa_{t\bar{t}}$. The corrected κ , which is shown with a lighter color, is used as $\kappa_{t\bar{t}}$ in the full background prediction.	68
5.7 R_{CS} as a function of njet from simulation of $W + \text{jets}$ events (left), and data (right) in $W + \text{jets}$ sideband.	69
5.8 R_{CS} values went into κ_w (lower panel) calculation is shown.	70
5.9 κ_w and $\kappa_{t\bar{t}}$ correction factors are shown. The total κ , which is represented with black line, shows the ratio of prediction to the simulation. The shaded area displays the uncertainty coming from simulation statistics.	71

6.1	Relative uncertainty on $\kappa_{t\bar{t}}$ due to the different composition of dileptonic events in sideband and main band regions. Color filled areas represent the dilepton uncertainties while the black line shows the uncertainty calculated with fit along the n_{jet}	74
6.2	Pileup dependence of the two signal benchmark points. The blue line is representing the high mass gap point the red one is for low mass gap region. The histogram points are following a flat line around 1.	77
6.3	Distribution of pileup for the high-mass gap signal sample, divided in to two parts as low and high pileup region (left), example fit performed using the two points explained in the text (middle), data pileup distribution which is folded with the fit.	78
6.4	Visualization of all systematic uncertainties on the background prediction for main band regions described in Tab. 4.5.	79
6.5	Visualization of all systematic uncertainties on the simulated signal yields in main band regions described in Tab. 4.5. Uncertainties are shown only for one mass point which is $m_{\tilde{g}} = 1900$ and $m_{\tilde{\chi}_1^0} = 100$	79
7.1	Fits to the $n_{\text{b-tag}}$ multiplicity for control regions in (a) $n_{\text{jet}} = 3$ ($250 \leq L_T < 350 \text{ GeV}$, $H_T \geq 500 \text{ GeV}$, $\Delta\phi < 1$, muon channel) and (b) $n_{\text{jet}} = 4$ ($250 \leq L_T < 350 \text{ GeV}$, $H_T \geq 500 \text{ GeV}$, $\Delta\phi < 1$) in data. The solid lines represent the templates scaled according to the fit result (blue for $t\bar{t} + \text{jets}$, green for $W + \text{jets}$, turquoise for QCD, and red for the remaining backgrounds), the dashed line shows the sum after fit, and the points with error bars represent data.	81
7.2	Validation of the background estimation method, using mainband regions with 4 jets. The shaded area reflects a rough estimate of the statistical and systematic uncertainties. The colored lines illustrate the expectations for two benchmark points of the T5qqqqWW model, showing the SUSY particle masses $m_{\tilde{g}}/m_{\tilde{\chi}_1^0}$ in TeV. The lower band shows the ratio between observed and predicted events.	82
7.3	Observed data and predicted event yields (from data) in the 28 search regions are shown. The black points with error bars show the number of observed data events and corresponding statistical error. The filled, stacked histograms represent the predictions for $t\bar{t} + \text{jets}$, $W + \text{jets}$, and the remaining rare backgrounds. Lower panel shows the ratio of data to prediction and the shaded area reflects the total (stat. and syst.) relative uncertainty of the predicted background.	83

7.4	$\Delta\phi$ distributions in data which is represented with black points and simulation which is represented with color filled areas. Figure shows the four L_T interval after the baseline selection with inclusive requirements on $n_{jet} \geq 5$ and $H_T \geq 500$ GeV.	84
7.5	Simplified example of a datacard file for one search region. The entries corresponding to rare backgrounds and QCD have been suppressed.	86
7.6	Observed significance for the model T5qqqqWW.	88
7.7	95% CL Observed (Expected) upper limits on the signal strength modifier μ for the T5qqqqWW model is shown at the left (right) side as a function of the gluino and LSP masses.	90
7.8	Cross section limits at a 95% C.L. for the T5qqWWWW model, and as a function of the gluino and LSP masses. The solid black (dashed red) lines correspond to the observed (expected) mass limits, with the thicker lines representing the central values and the thinner lines representing the $\pm 1\sigma$ uncertainty bands related to the theoretical (experimental) uncertainties . .	90
7.9	Best exclusion limits on sparticle masses from searches for SUSY using SMS by ATLAS (top) and CMS (bottom) collaborations [40, 41]. In the upper plot, the green bars represent the 7-8 TeV results while the blue bars are summarizing the 13 TeV results. In the lower plot, the orange bars represent the 7-8 TeV results while the blue bars are summarizing the 13 TeV results.	92
7.10	95% C.L. exclusion limit curves on the simplified SUSY model of gluino pair production with subsequent decay of $\tilde{g} \rightarrow q\bar{q}WW\tilde{\chi}^0$ obtained by the CMS (left) and ATLAS (right) collaborations. The present analysis is corresponding to the result labeled with SUS-16-042.	93
B.1	Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $3 \leq \text{jets} \leq 4$ and zero b-tagged jets (1 μ channel).	100
B.2	Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $3 \leq \text{jets} \leq 4$ and zero b-tagged jets (1 e channel).	101
B.3	Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $4 \leq \text{jets} \leq 5$ and b-tagged jets (1 μ channel).	102
B.4	Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $4 \leq \text{jets} \leq 5$ and b-tagged jets (1 e channel).	103
B.5	Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, ≥ 5 jets and zero b-tagged jets (1 μ channel).	104

B.6	Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, ≥ 5 jets and zero b-tagged jets (1 e channel).	105
C.1	Prefit (grey), s+b postfit (red) and, b-only postfit (blue) values of nuisance parameters included in the fit.	107
C.2	Prefit (grey), s+b postfit (red) and, b-only postfit (blue) values of nuisance parameters in side band regions included in the fit.	108

List of Tables

1.1	Chiral supermultiplets of the MSSM	14
1.2	Gauge supermultiplets of the MSSM	14
3.1	List of selection criteria for the CMS electron identification.	35
4.1	List of kinematic variables	41
4.2	Weights based on the number of ISR jets as given in Ref. [47]	45
4.3	List of event selection criteria and object requirements.	48
4.4	List of HLT paths	50
4.5	Signal regions.	58
5.1	Overview of the definitions of sideband and mainband regions. For the multijet (QCD) fit the electron (e) sample is used, while for the determination (det.) of $R_{\text{CS}}(W + \text{jets})$ the muon (μ) sample is used. Empty cells are not used in the analysis.	62
7.1	Conversion of R_{CS} regions to corresponding ABCDEF regions	83
A.1	List of simulated $t\bar{t} + \text{jets}$ background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.	97
A.2	List of simulated $W + \text{jets}$ background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.	97
A.3	List of the rest of simulated background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.	98

Abbreviations

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC Apparatus
BSM	Beyond the Standard Model
CERN	European Organization for Nuclear Research
CM	Center of Mass
CMS	Compact Muon Solenoid experiment
CMSSW	CMS SoftWare framework
DAQ	Data Acquisition
ECAL	Electromagnetic Calorimeter
EWK	Electroweak Theory
GUT	Grand Unified Theory
HCAL	Hadron Calorimeter
HF	Hadron Calorimeter (Forward)
LHC	Large Hadron Collider
LHCb	the Large Hadron Collider Beauty Experiment
LINAC	Linear particle Accelerator
PDG	Particle Data Group
QED	Quantum Electrodynamics
QFT	Quantum Field Theory
SM	Standard Model
SMS	Simplified Model Spectra
SUSY	Supersymmetry

Bibliography

- [1] F. Halzen and A. Martin, “Quarks and leptons: an introductory course in modern particle physics”. Physics textbook. Wiley, 1984.
- [2] D. Griffiths, “Introduction to Elementary Particles”. Physics textbook. Weinheim, Germany: Wiley-VCH 454 p, 2014.
- [3] B.R. Martin and G. Shaw, “Particle Physics”. 4th Edition. ISBN: 978-1-118-91190-7
- [4] S. P. Martin, A Supersymmetry primer, doi:10.1142/9789812839657_0001, arXiv:hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [5] C. Patrignani *et al.* [Particle Data Group], “Review of Particle Physics,” Chin. Phys. C **40** (2016) no.10, 100001. doi:10.1088/1674-1137/40/10/100001
- [6] Gell-Mann, M. Nuovo Cim (1956) 4(Suppl 2): 848. doi:10.1007/BF02748000
- [7] E. Fermi, “An attempt of a theory of beta radiation. 1.”, Z. Phys. **88** (1934) 161177.
- [8] S. L. Glashow, Partial Symmetries of Weak Interactions, Nucl. Phys. **22** (1961) 579588, doi:10.1016/0029-5582(61)90469-2.
- [9] A. Salam, Weak and Electromagnetic Interactions, Conf. Proc. C680519 (1968) 367377.
- [10] S. Weinberg, A Model of Leptons, Phys. Rev. Lett. **19** (1967) 12641266, doi:10.1103/PhysRevLett.19.1264.
- [11] F. Englert and R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, Phys. Rev. Lett. **13**(1964)321323,doi:10.1103/PhysRevLett.13.321.
- [12] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Phys. Rev. Lett. **13** (1964) 508509, doi:10.1103/PhysRevLett.13.508.
- [13] CDF Collaboration, Observation of top quark production in p p collisions, Phys. Rev. Lett. **74**(1995) 26262631, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.

- [14] D0 Collaboration, Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV, Phys. Rev. Lett. **74** (1995) 24222426, doi:10.1103/PhysRevLett.74.2422, arXiv:hep-ex/9411001.
- [15] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. **B716** (2012) 3061, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [16] ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. **B716** (2012) 129, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [17] F. Brnner and A. Rebhan, “Nonchiral enhancement of scalar glueball decay in the Witten-Sakai-Sugimoto model,” Phys. Rev. Lett. **115**, no. 13, 131601 (2015) doi:10.1103/PhysRevLett.115.131601 [arXiv:1504.05815 [hep-ph]].
- [18] Y. K. Hsiao and C. Q. Geng, “Identifying Glueball at 3.02 GeV in Baryonic B Decays,” Phys. Lett. B **727**, 168 (2013) doi:10.1016/j.physletb.2013.10.008 [arXiv:1302.3331 [hep-ph]].
- [19] Bennett G et al. (Muon G-2 Collaboration) 2006 Phys.Rev. **D73** 072003 (Preprint hep-ex/0602035)
- [20] Abbott, Benjamin P.; et al. (LIGO Scientific Collaboration and Virgo Collaboration) (2016). ”Observation of Gravitational Waves from a Binary Black Hole Merger”. Phys. Rev. Lett. **116** (6): 061102. Bibcode:2016PhRvL.116f1102A. PMID 26918975. arXiv:1602.03837 . doi:10.1103/PhysRevLett.116.061102.
- [21] Castelvecchi, Davide (23 March 2016). ”The black-hole collision that reshaped physics”. Nature. **531** (7595): 428431. doi:10.1038/531428a.
- [22] F. Zwicky. ”On the Masses of Nebulae and of Clusters of Nebulae”. Astrophys. J., 86:217 246, 1937. doi:10.1086/143864.
- [23] ‘V. C. Rubin, N. Thonnard, and W. K. Ford, Jr. ”Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 / $R = 4\text{kpc}$ / to UGC 2885 / $R = 122\text{kpc}$ ”’. Astrophys. J., 238:471, 1980. doi:10.1086/158003.
- [24] G. Hinshaw *et al.* [WMAP Collaboration], ”Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results,” Astrophys. J. Suppl. **180**, 225 (2009) doi:10.1088/0067-0049/180/2/225 [arXiv:0803.0732 [astro-ph]].

- [25] Boggess, N. W.; et al. (1992). "The COBE Mission: Its Design and Performance Two Years after the launch". *The Astrophysical Journal*. **397**: 420. Bibcode:1992ApJ...397..420B. doi:10.1086/171797.
- [26] P. A. R. Ade *et al.* [Planck Collaboration], "Planck 2015 results. XIII. Cosmological parameters," *Astron. Astrophys.* **594**, A13 (2016) doi:10.1051/0004-6361/201525830 [arXiv:1502.01589 [astro-ph.CO]].
- [27] **Super-Kamiokande** Collaboration, Evidence for oscillation of atmospheric neutrinos, *Phys. Rev. Lett.* **81** (Aug, 1998) 15621567. doi: <https://doi.org/10.1103/PhysRevLett.81.1562>
- [28] **OPERA** Collaboration, Observation of a first candidate event in the OPERA experiment in the CNGS beam, *Physics Letters B* **691** no. 3, (2010) 138–145. doi: <https://doi.org/10.1016/j.physletb.2010.06.022>
- [29] D. H. Perkins, "Introduction to high energy physics". Cambridge University Press, 2000.
- [30] Emmy Noether; Mort Tavel (translator) (1971). "Invariant Variation Problems". *Transport Theory and Statistical Physics*. **1** (3): 186207. arXiv:physics/0503066. doi:10.1080/00411457108231446.
- [31] Give reference
- [32] H. B. Kim and J. E. Kim, "Coupling constant unification and LEP data," *J. Korean Phys. Soc.* **27**, 129 (1994) [hep-ph/9302260].
- [33] S. Schael *et al.* [ALEPH and DELPHI and L3 and OPAL and SLD Collaborations and LEP Electroweak Working Group and SLD Electroweak Group and SLD Heavy Flavour Group], "Precision electroweak measurements on the Z resonance," *Phys. Rept.* **427**, 257 (2006) doi:10.1016/j.physrep.2005.12.006 [hep-ex/0509008].
- [34] Wilczek, F., Raby, S., Dimopoulos, S. (1991, October). Unification of couplings. *PHYSICS TODAY*, 44(10), 25-33. doi: <http://dx.doi.org/10.1063/1.881292>
- [35] B. Bajc, J. Hisano, T. Kuwahara and Y. Omura, "Threshold corrections to dimension-six proton decay operators in non-minimal SUSY SU (5) GUTs," *Nucl. Phys. B* **910**, 1 (2016) doi:10.1016/j.nuclphysb.2016.06.017 [arXiv:1603.03568 [hep-ph]].

- [36] S. Chatrchyan *et al.* [CMS Collaboration], “Interpretation of Searches for Supersymmetry with simplified Models,” Phys. Rev. D **88**, no. 5, 052017 (2013) doi:10.1103/PhysRevD.88.052017 [arXiv:1301.2175 [hep-ex]].
- [37] C. Borschensky, M. Krmer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn and X. Portell, “Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV,” Eur. Phys. J. C **74**, no. 12, 3174 (2014) doi:10.1140/epjc/s10052-014-3174-y [arXiv:1407.5066 [hep-ph]].
- [38] WJS2012
- [39] C. Borschensky et al., LHC SUSY Cross Section Working Group. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections>. (Accessed on 12/07/2017).
- [40] <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY/index.html>
- [41] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS>
- [42] V. Khachatryan *et al.* [CMS Collaboration], “Searches for supersymmetry based on events with b jets and four W bosons in pp collisions at 8 TeV,” Phys. Lett. B **745** (2015) 5 doi:10.1016/j.physletb.2015.04.002 [arXiv:1412.4109 [hep-ex]].
- [43] R. M. Barnett, J. F. Gunion, and H. E. Haber, Gluino Decay Patterns and Signatures, Phys. Rev. **D37** (1988) , no. 7, 18921907.
- [44] D. Alves *et al.* [LHC New Physics Working Group], “Simplified Models for LHC New Physics Searches,” J. Phys. G **39**, 105005 (2012) doi:10.1088/0954-3899/39/10/105005 [arXiv:1105.2838 [hep-ph]].
- [45] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSMoriond2017ObjectsEfficiency> . (Accessed on 20/07/2017).
- [46] Henning Kirschenmann, Trigger efficiencies. Private communication, 2017.
- [47] Manuel Franco Sevilla, Reweighting procedure for ttbar based on the number of ISR jets”.Private communication, 2017.
- [48] CMS Collaboration, CMS Luminosity Measurements for the 2016 Data Taking Period, Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [49] D. Handl, Background estimation for searches for supersymmetry in the singlelepton final state in 13 TeV pp collisions, Masters thesis, 2016.

- [50] CMS Collaboration, CMS Luminosity Measurements for the 2016 Data Taking Period, Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [51] Give luminosity ref
- [52] Give Wcross ref
- [53] Give ttcross ref
- [54] Give ISR ref.
- [55] CMS Collaboration, Event generator tunes obtained from underlying event and multiparton scattering measurements, Eur. Phys. J. C 76 (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [56] T. Junk, “Confidence level computation for combining searches with small statistics,” Nucl. Instrum. Meth. A 434, 435 (1999) doi:10.1016/S0168-9002(99)00498-2 [hep-ex/9902006].
- [57] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, Procedure for the LHC Higgs boson search combination in Summer 2011, Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.
- [58] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, The European Physical Journal C 71 (Feb., 2011) 119.
- [59] J. Neyman and E. S. Pearson, On the Problem of the Most Efficient Tests of Statistical Hypotheses, Philosophical Transactions of the Royal Society of London Series A 231 (1933) 289337, doi:10.1098/rsta.1933.0009.
- [60] V. Khachatryan *et al.* [CMS Collaboration], “Search for supersymmetry in events with one lepton and multiple jets in proton-proton collisions at $\sqrt{s} = 13$ TeV,” Phys. Rev. D 95 (2017) no.1, 012011 doi:10.1103/PhysRevD.95.012011 [arXiv:1609.09386 [hep-ex]].
- [61] Mobs, Esma, ”The CERN accelerator complex. Complexe des accélérateurs du CERN” Jul, 2016, OPEN-PHO-ACCEL-2016-009, url:<http://cds.cern.ch/record/2197559>
- [62] ”Linear accelerator 2”, Sep. 2012, url :<http://cds.cern.ch/record/1997427>”
- [63] ”The Proton Synchrotron Booster”, July 2012, url:
<http://cds.cern.ch/record/1997372>”

- [64] "The Proton Synchrotron", Jan. 2012, url: "<http://cds.cern.ch/record/1997189>"
- [65] "The Super Proton Synchrotron", Jan. 2012, url: "<http://cds.cern.ch/record/1997188>"
- [66] M. Banner *et al.* [UA2 Collaboration], "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider," *Phys. Lett.* **122B** (1983) 476. doi:10.1016/0370-2693(83)91605-2
- [67] G. Arnison *et al.* [UA1 Collaboration], "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider," *Phys. Lett.* **126B** (1983) 398. doi:10.1016/0370-2693(83)90188-0
- [68] O. S. Bruning et al., LHC Design Report. CERN, Geneva, 2004. doi: 10.5170/CERN-2004-003-V-1
- [69] R. Bruce *et al.*, "Baseline LHC machine parameters and configuration of the 2015 proton run," doi:10.5170/CERN-2015-002.100 arXiv:1410.5990 [physics.acc-ph].
- [70] D. Fournier, "Performance of the LHC, ATLAS and CMS in 2011," *EPJ Web Conf.* **28** (2012) 01003 doi:10.1051/epjconf/20122801003 [arXiv:1201.4681 [hep-ex]].
- [71] The HL-LHC project — The HL-LHC Project <http://hilumilhc.web.cern.ch/about/hl-lhc-project> (Accessed on 15/09/2017).
- [72] T. Sakuma and T. McCauley, "Detector and Event Visualization with SketchUp at the CMS Experiment," *J. Phys. Conf. Ser.* **513** (2014) 022032 doi:10.1088/1742-6596/513/2/022032 [arXiv:1311.4942 [physics.ins-det]].
- [73] CMS Collaboration, The CMS experiment at the CERN LHC, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [74] CMS Collaboration, Letter of intent: by the CMS for a general purpose detector at LHC, Technical Report CERN-LHCC-92-003. CERN-LHCC-92-3. LHCC-I-1, CERN, Geneva, 1992. Open presentation to the LHCC 5 November 1992, M. Della Negra/CERN, CMS Spokesman.
- [75] S. Chatrchyan *et al.* [CMS Collaboration], "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays," *JINST* **5** (2010) T03021 doi:10.1088/1748-0221/5/03/T03021 [arXiv:0910.5530 [physics.ins-det]].
- [76] P. Adzic et al., Energy resolution of the barrel of the CMS electromagnetic calorimeter, *JINST* **2** (2007) P04004, doi:10.1088/1748-0221/2/04/P04004.

- [77] P. Sphicas [CMS Collaboration], “CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger,” CERN-LHCC-2002-026.
- [78] A. Bell et al., Fast beam conditions monitor BCM1F for the CMS experiment, Nuclear Instruments and Methods in Physics Research **A** **614** (March, 2010) 433438, doi:10.1016/j.nima.2009.12.056, arXiv:0911.2480.
- [79] P. Nason & P.Z. Skands, “Monte Carlo event generators”, in Review of Particle Physics, C. Patrignani et al. (Particle Data Group), Chin. Phys. C40,100001 (2016).
- [80] M. Cacciari, G. P. Salam, and G. Soyez, FastJet User Manual, Eur. Phys. J. C **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [81] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, JHEP **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [82] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method, JHEP **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [83] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, JHEP **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [84] S. Alioli, P. Nason, C. Oleari, and E. Re, NLO single-top production matched with shower in POWHEG: s- and t-channel contributions, JHEP **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].
- [85] E. Re, Single-top Wt-channel production matched with parton showers using the POWHEG method, Eur. Phys. J. C **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [86] J. Alwall et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, JHEP **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [87] GEANT4 Collaboration, GEANT4: A Simulation toolkit, Nucl. Instrum. Meth. **A506** (2003) 250303, doi:10.1016/S0168-9002(03)01368-8.

- [88] A. Giammanco, The Fast Simulation of the CMS Experiment, *J. Phys. Conf. Ser.* **513** (2014) 022012, doi:10.1088/1742-6596/513/2/022012.
- [89] CMS-doc-4172: Static image with tracks. <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB>ShowDocument?docid=4172>. (Accessed on 24/09/2017).
- [90] CMS Collaboration, Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector, 2010.
- [91] J. Alwall, P. C. Schuster, and N. Toro, Simplified models for a first characterization of new physics at the LHC, *Phys. Rev. D* **79** (2009) 075020, doi:10.1103/PhysRevD.79.075020, arXiv:0810.3921.
- [92] W. Adam, B. Mangano, T. Speer, and T. Todorov, Track Reconstruction in the CMS tracker, Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec, 2006.
- [93] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov, Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC, *eConf C0303241* (2003) TULT009, doi:10.1088/0954-3899/31/9/N01, arXiv:physics/0306087. [*J. Phys.G31,N9(2005)*].
- [94] CMS Collaboration, Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET, technical report, 2009.
- [95] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, Adaptive vertex fitting, *J. Phys. G* **34** (2007) N343, doi:10.1088/0954-3899/34/12/N01.
- [96] CMS Collaboration, Muon Identification in CMS, 2008. Unpublished. Available at
- [97] CMS Collaboration, Search for supersymmetry in pp collisions at \sqrt{s} (s) = 13 TeV in the single-lepton final state using the sum of masses of large radius jets, Technical Report CMS-PAS-SUS-15-007, CERN, Geneva, 2015.
- [98] CMS Collaboration, Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at p_T = 8 TeV, *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [99] CMS Collaboration, Electron and Photon performance using data collected by CMS at \sqrt{s} = 13 TeV and 25ns.
- [100] CMS Collaboration, Jet Energy Calibration in the 8 TeV pp data, Technical Report CMS-PAS-JME-13-004, CERN, Geneva, 2014. To be published.

- [101] CMS Collaboration, Identification of b quark jets at the CMS Experiment in the LHC Run 2, Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.
- [102] CMS Collaboration, S. Chatrchyan et al., Identification of b-quark jets with the CMS experiment, JINST **8** (2013) P04013, arXiv:1211.4462 [hep-ex].
- [103] M. Cacciari, G. P. Salam, and G. Soyez, The Anti- $k(t)$ jet clustering algorithm, JHEP **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [104] Salam, Gavin P., Towards jetography, Eur. Phys. J. C **67** no. 3, (2010) .
<http://dx.doi.org/10.1140/epjc/s10052-010-1314-6>.
- [105] CMS Collaboration, Jet Energy Calibration in the 8 TeV pp data, Technical Report CMS-PAS-JME-13-004, CERN, Geneva, 2014. To be published.
- [106] CMS Collaboration, Performance of the CMS missing transverse momentum reconstruction in pp data at $\text{ps} = 8 \text{ TeV}$, JINST **10** (2015), no. 02, P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [107] [CMS Collaboration], “Studies of Tracker Material,” CMS-PAS-TRK-10-003.