



TECHNISCHE
UNIVERSITÄT
WIEN

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

PROF. DR. JOCHEN SCHIECK

und

DR. ROBERT SCHÖFBECK

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

eingereicht an der
TECHNISCHEN UNIVERSITÄT WIEN
FAKULTÄT FÜR PHYSIK
von

ECE AŞILAR

Matrikelnummer: 1428489

Wien, am 14. December 2017



This thesis is dedicated to
the city of Vienna

Acknowledgements

It took three years for this thesis to come out. If it is written as required, the necessary appreciation for these three years will not fit into one page. For this reason only a brief version of my gratitude is shown here.

First of all, I would like to thank Robert Schöffbeck for writing this project and choosing me as his student. At every crucial point of the work, he was there and supporting with very beneficial advice. He always encouraged us to reach our best. I also would like to thank Wolfgang Adam. He was always there to discuss problems and finding solutions together. He was always managing to make us happy even in the very stressful times with his calm and positive attitude. I am also grateful to Prof. Jochen Schieck, my thesis supervisor and director of HEPHY, for providing an inspiring atmosphere within the institute. Furthermore, I want to thank the other jury members, Rosy Nikolaidou and Hubert Kroha for reading this thesis.

Moreover, I want to thank all my colleagues at HEPHY for an efficient and at the same time friendly working atmosphere. I further appreciated support from Daniel Spitzbart, Navid k. Rad, Federico Ambrogy, Matheusz Zarucki, Erica Brondolin, Ilse Kratschmer, Johannes Brandstetter, Wolfgang Waltenberger and Suchita Kulkarni when I had a ski accident at the end of my PhD.

I would like to thank also my ACDVF colleagues, Claudia Seitz, Henning Kirschenmann, Anadi Canepa, Dirk Krücker, Stephan Lammel, Artur Lobanov, Isabell Melzer-Pellmann, Basil Schneider, Anna Stakia, Paris Sphicas and Markus Stoye for their great team work when publishing our analysis.

A thank you should also go to all the CMS groups, the LHC team and all the people working at CERN.

I would like to thank also Eveline Ess, Nathalie Fortin, Zorica Jelovic, Sonja Weiss.

Moreover, thank you to our beloved IT at HEPHY, Dietrich Liko, for his patience and very talented work. He always answered our needs whenever possible including nights and weekends. I would like to thank René Brun for ROOT, Guido van Rossum for Python and Bram Moolenaar for Vim.

A special thank you goes my mom, my brother, his girlfriend Meltem, and my dear friend Aysegül for always providing the support that I need outside the work life.

Final but not least, this thesis was made possible with the FWF project P26771-N20 and Doktoratskolleg Particles and Interactions (DK-PI).

Thanks for all your encouragement!

Abstract

A search for Supersymmetry will be presented to turn one more stone in the quest of what is beyond the Standard Model. The search is performed in events with a single charged lepton, multiple jets, and missing transverse energy. The proton-proton collision data were recorded by the CMS experiment during the 2016 run of the LHC at a center-of-mass energy of 13 TeV. The integrated luminosity of the dataset corresponds to 35.9 fb^{-1} .

Although the search is model independent, a simplified model of gluino pair production with masses in the TeV range is used as an example. In the model, each of the gluinos decays to two light quarks and an intermediate chargino, with the latter decaying to a W boson and a neutralino. The neutralino is considered to be the stable lightest supersymmetric particle which results in substantial missing transverse energy in the final state. Only events with a single charged lepton, which can be an electron or a muon, are considered. This requirement provides a clean event topology as well as suppresses most of the multijet events. No b quark is expected in the final state of the targeted signal model. Therefore, in the event selection, a veto on b-tagged jets is included. The search uses a powerful discriminating variable to distinguish between background and signal events, and 28 exclusive signal regions are defined in different kinematic observables to enhance sensitivity to a range of different mass scenarios. The estimation of the Standard Model background yields in the signal regions is performed using data in the control regions. This methodology is verified using simulated samples and data in validation regions. Finally, systematic uncertainties related to the background prediction and simulated samples are determined.

No significant deviation from the predicted Standard Model background is observed. Therefore, stringent upper limits on the cross section of the considered simplified model are set. As a result, gluino masses below 1.9 TeV are excluded for neutralino masses below 300 GeV with 95% confidence level.

Contents

Introduction	3
1 Supersymmetry: an extension of the Standard Model	4
1.1 Standard Model	4
1.1.1 Particle content	4
1.1.2 Particle interactions	5
1.1.3 From the quantum field theory window	6
1.1.4 Experimental results	10
1.1.5 Shortcomings	11
1.1.5.1 Experimental reasons	11
1.1.5.2 Theoretical reasons	11
1.2 Supersymmetry as a solution	13
1.2.1 Algebra of Supersymmetry	14
1.2.2 Minimal Supersymmetric Standard Model	15
1.2.2.1 Particle decays	17
1.2.3 Simplified models	17
1.2.4 Short history of SUSY searches at colliders	18
2 Experimental setup	22
2.1 The LHC at CERN	22
2.1.1 The CERN accelerator complex	22
2.1.2 The LHC	23
2.2 The CMS detector	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	28
2.3 Event simulation	29
2.3.1 Event generation	29
2.3.2 Detector simulation	29

3	Object reconstruction and identification	30
3.1	Particle-Flow algorithm	30
3.2	Tracks and primary vertices	32
3.3	Jets	33
3.3.1	Identification of b jets	34
3.4	Leptons	35
3.4.1	Muons	35
3.4.2	Electrons	35
3.4.3	Isolation	37
3.5	Missing transverse energy	39
4	Event Selection:	
	Baseline and search regions	43
4.1	SUSY signature	43
4.1.1	Key variables	44
4.1.2	Signal samples	45
4.2	Background processes	46
4.2.1	Scale factors	47
4.3	Baseline selection	49
4.4	Data samples	51
4.4.1	Trigger selection	52
4.5	Event cleaning filters	54
4.6	Control plots	55
4.7	Regions of interest	57
4.7.1	Signal and control region	57
4.7.2	Mainband regions	57
4.7.3	Aggregate regions	60
5	Background estimation: The R_{CS} method	66
5.1	QCD background estimation	67
5.2	Background fraction calculations: b-tag multiplicity fit	69
5.3	R_{CS} method in $t\bar{t}$ + jets events	69
5.4	R_{CS} method in W+jets events	73
5.5	Validation of the background estimation	75
6	Systematic uncertainties	80
6.1	Systematic uncertainties on background estimation	80
6.2	Systematic uncertainties on signal modelling	82
6.3	Common systematic uncertainties for signal and background modelling . . .	83

7 Results and interpretation	88
7.1 Results of the background prediction	88
7.1.1 Result of the validation in sideband regions	88
7.1.2 Result of the background prediction in mainband regions	88
7.1.3 Result of background prediction in aggregated regions	89
7.2 Statistical interpretation	90
7.2.1 Frequentist limit setting procedure	92
7.3 Interpretation on simplified model T5qqqqWW	96
7.4 Comparison to complementary results	97
Conclusion	102
A MC samples and cross sections	103
B Control plots	105
C Statistical tests	112
Bibliography	126

List of Figures

1.1	Fundamental particles of the SM	5
1.2	The experimental tests of the SM	10
1.3	The running couplings of U(1), SU(2), SU(3)	12
1.4	Feynman diagrams for one-loop quantum corrections on the Higgs squared mass parameter	12
1.5	The simplified models diagrams	18
1.6	Feynman diagrams for the gluino pair production	19
1.7	Cross sections of gluino pair production	19
1.8	Cross section limits at a 95% CL for (left) the T1tttt model, and (right) the T5qqqqWW model [68].	20
1.9	Cross section of SM processes as a function of center of energy [69].	21
2.1	CERN accelerator complex [71]	23
2.2	Instantaneous and integrated luminosity recorded by CMS detector	25
2.3	Schematic view of the CMS detector with a cut-out quadrant [86]	26
3.1	Cross section of a slice of the CMS detector	31
3.2	The mean number of interactions per bunch crossing	33
3.3	Isolation efficiency comparisons	39
3.4	PF candidate multiplicity fits for E_x and E_y in different η regions. Different colored lines represent different data taking run ranges. The Y-axis shows the average value of the x- and y-components of missing energy while X is the sum p_T of the corresponding PF candidate.	41
3.5	$E_T \phi$ (left) and p_T (right) distributions before (blue) and after (red) corrections in simulated Drell-Yan events.	42
4.1	Diagram of the simplified model T5qqqqWW.	43
4.2	H_T (left), L_T (middle), n_{jet} (right) distributions for three different signal mass point	46
4.3	Simulated signal mass points for the simplified T5qqqqWW model.	46
4.4	The normalized distributions of mean number of interactions per bunch crossing	49
4.5	Distributions of M_{T2}	50
4.6	The selection criteria flow for the electron (left) and muon (right) channel	51

4.7	The trigger efficiency as a function of L_T	53
4.8	The trigger efficiency as a function of H_T	53
4.9	The trigger efficiency as a function of lepton p_T	54
4.10	Main kinematic distributions after baseline selection	56
4.11	Sketch of the n_{jet} vs. $\Delta\phi$ plane	57
4.12	The $\Delta\phi$, after the baseline selection, MC only	58
4.13	The figure of merit as a function of $\Delta\phi$ is shown for T5qqqqWW (1.9,0.1) after the baseline selection.	60
4.14	Two dimensional distributions of event counts for the main background samples, $t\bar{t}$ + jets + W + jets simulation (top), and the signal T5qqqqWW (1.9,0.1) (left) and T5qqqqWW (1.5,1.0) in the $\Delta\phi$ vs. L_T plane after the baseline selection.	61
4.15	The distributions on top represent T5qqqqWW signal counts (left) and selection efficiency (right) after the baseline requirements in the gluino-neutralino plane. The distributions at the bottom show the efficiency of signal region selection with respect to the baseline selection; $\Delta\phi > 1$ (left) and $\Delta\phi > x$ (right) where x is the threshold according to L_T bin.	62
4.16	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).	63
4.17	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).	64
4.18	r-value calculated for each of the inclusive aggregate bins.	65
5.1	The L_P shape fit result for $3 \leq n_{\text{jet}} \leq 4$ and $n_{b\text{-jet}} = 0$ in the $250 \leq L_T \leq 350$ bin (left). Ratio of selected to anti-selected electron events from QCD multijet events for $3 \leq n_{\text{jet}} \leq 4$ and $n_{b\text{-jet}} = 0$, in bins of L_T in data (right).	68
5.2	The b-tag multiplicity fit examples	70
5.3	R_{CS} as a function of n_{jet} for dileptonic (left), and single leptonic (right) $t\bar{t}$ + jets events.	71
5.4	The simulated R_{CS} for $t\bar{t}$ + jets background	71
5.5	Double Ratio of n_{jet} distributions	72
5.6	Comparing $\kappa_{t\bar{t}}$ and $\kappa_{t\bar{t}}^{DL-Corr}$	73
5.7	Simulated R_{CS} for W + jets events	74
5.8	The top row shows the two simulated R_{CS} values (Eq. 5.13 numerator and denominator). The bottom row shows the resulting κ_w	75
5.9	The factors κ_w^{val} and the factors $\kappa_{t\bar{t}}^{\text{val}}$ are shown.	76

6.1	Relative uncertainty on $\kappa_{t\bar{t}}$ due to the different composition of dileptonic events in sideband and mainband regions. Color filled areas represent the dilepton uncertainties, while the black line shows the uncertainty which is calculated with a fit on the n_{jet} distribution.	81
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.	85
6.3	Distribution of pileup for the low mass gap signal benchmark 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.	85
6.4	Summary of systematic uncertainty for background prediction	86
6.5	Summary of systematic uncertainty for signal simulation.	87
7.1	The results of background estimation method validation.	89
7.2	The results of main background estimation.	91
7.3	The inclusive $\Delta\phi$ distributions	92
7.4	Simplified example of a text file file for one search region. The entries corresponding to rare backgrounds and QCD have been suppressed.	94
7.5	Observed significance for the model T5qqqqWW.	96
7.6	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 neutralino masses.	97
7.7	Cross section limits at a 95% CL for the T5qqqqWW 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	98
7.8	Best exclusion limits on sparticle masses from searches for SUSY using SMS by ATLAS (top) and CMS (bottom) collaborations [136, 137]. 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, results summarizing the 13 TeV results, the orange bars represent the 12.9 fb ⁻¹ results while the blue bars show the 15.9 fb ⁻¹ results.	99
7.9	95% CL 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.	100
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).	106
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).	107

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).	108
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).	109
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).	110
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).	111
C.1	Prefit (grey), s+b postfit (red) and, b-only postfit (blue) values of nuisance parameters included in the fit.	113
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.	114

List of Tables

1.1	Chiral supermultiplets of the MSSM	15
1.2	Gauge supermultiplets of the MSSM	16
3.1	List of selections for the tight and loose muon identification.	35
3.2	List of selections for the muon identification.	36
3.3	List of selections for the tight and loose electron identification.	37
3.4	List of selection criteria for the CMS electron identification.	37
4.1	Weights based on the number of ISR jets as given in Ref. [82]	48
4.2	List of baseline criteria and object requirements.	51
4.3	List of HLT paths. Notation is used in terms of CMS internal shortcuts. . . .	52
4.4	Simulation table of mainband regions, 35.9 fb^{-1}	59
4.5	Simulation table of the aggregate signal regions, 35.9 fb^{-1}	60
5.1	Overview of the definitions of sideband and mainband regions	67
5.2	Simulated and measured R_{CS} of the $t\bar{t}$ + jets events.	77
5.3	Summary table for R_{CS} for $t\bar{t}$ + jets and the corresponding $\kappa_{t\bar{t}}$ value from simulation	78
5.4	The R_{CS} values for positively and negatively charged leptons.	79
5.5	Simulation table of the validation regions, 35.9 fb^{-1}	79
7.1	Summary of the results in the main search regions. The $\Delta\phi$ value that is used to define the CR and SR is provided as well.	90
7.2	The results of background estimation in aggregated regions.	90
7.3	Conversion of R_{CS} regions to corresponding ABCDEF regions	91
A.1	List of simulated $t\bar{t}$ + jets background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.	103
A.2	List of simulated W+jets background samples with a 25 ns bunch crossing processed in CMSSW version 8_0_x.	103
A.3	List of the rest of simulated background samples with a 25ns bunch crossing processed in CMSSW version 8_0_x.	104

Introduction

The Standard Model (SM) of particle physics [1–5] describes fundamental particles and their interactions through the electromagnetic, the weak and the strong force. The discovery of the entire SM particle content was completed [6–10] with the various high-energy experiments, for example, the Tevatron at Fermilab between 1983 and 2011, and the Large Hadron Collider (LHC) at CERN starting its first scientific run in 2010. Up to now, results are consistent with the SM within uncertainties. However, results from the experiments such as Planck [11], successor of COBE [12] and WMAP [13], suggest a large amount of dark matter in the universe which the SM fails to explain. Many extensions of the SM provide solutions to these problems, and Supersymmetry (SUSY) [14–21] is among the most promising candidates.

A generic search for SUSY in events with a single electron or muon is performed using proton-proton collisions at a center-of-mass energy of 13 TeV. The data were recorded by the Compact Muon Solenoid (CMS) [22] experiment during the 2016 Run of the LHC, corresponding to an integrated luminosity of 35.9 fb^{-1} . The signal model is a simplified model (SMS) [23–26] of gluino pair production with masses in the TeV range. Each of the gluinos decays to a three-body final state consisting of a pair of light quarks and an intermediate chargino. The chargino further decays to a neutralino and a W boson. The chargino mass is assumed to be the average of the masses of the gluino and the neutralino. The neutralino is considered to be the stable lightest supersymmetric particle which, in the detector, is reconstructed as substantial missing transverse energy in the final state. No b-tagged jet is expected in the final state of the targeted signal model. Therefore, in the event selection, a veto on b-tagged jets is included.

The SM background is dominated by $W + \text{jets}$ and $t\bar{t} + \text{jets}$ events, where the isolated lepton stems from a leptonic decay of a W boson. Therefore, the lepton is geometrically aligned with the W boson momentum. The neutrino originating from the W boson causes an energy imbalance in the detector. The energy imbalance and the lepton form the reconstructed W boson. In the case of a supersymmetric signal event, the existence of additional energy imbalance from neutralinos randomizes the angle between the lepton and the reconstructed W boson. Because of this distinguishing feature, we can define a separate signal and control region. To enhance the sensitivity to a range of different mass configurations, multiple signal rich search regions are defined based on the number of jets, the scalar sum of all jet transverse momenta, and the scalar sum of the transverse missing momentum and the transverse lepton momentum.

The contribution of SM background events in each search region is estimated using the

data in the corresponding control regions. Additional sideband regions are defined in order to obtain transfer factors from signal- to control regions in data. Prior to measuring the transfer factors, the multijet events contributions are subtracted from these control regions. The prediction of multijet events is performed using measurements of lepton misidentification probabilities from data.

Comprehensive studies of systematic uncertainties are performed. One of the most important systematic uncertainty originates from the different shapes of the angular distribution in dileptonic and single leptonic events. The presence of two neutrinos in $t\bar{t}$ events results in larger angles between the lepton and the reconstructed W boson candidate than in single lepton $t\bar{t}$ events. The systematic uncertainty related to the modeling of this difference, is evaluated in data in control regions with events containing two leptons.

Finally, the results are interpreted in terms of the SMS of the aforementioned model of gluino pair production.

The results of this work is also presented in a smaller number of aggregated search regions in order to facilitate future reinterpretations.

This work is published in

- CMS Collaboration, “Search for supersymmetry in events with one lepton and multiple jets exploiting the angular correlation between the lepton and the missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV”, arXiv:1709.09814, submitted to *Phys. Lett. B*.
- CMS Collaboration, “Search for supersymmetry in events with one lepton and multiple jets in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2016”, CMS-PAS-SUS-16-019, 2016.
- 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), doi:10.1103/PhysRevD.95.012011, arXiv:1609.09386.

This thesis is organized as follows.

- Chapter 1 starts with a brief introduction to the SM of particle physics and its shortcomings. The chapter continues with discussion on supersymmetric models, as one of the most appealing theories Beyond the Standard Model (BSM). A short review of SUSY searches at colliders is also presented.
- Chapter 2 gives an overview of the experimental setup used to collect data for this analysis. Prior to the description of CMS experiment, the LHC is briefly explained. Additionally, complementary simulation tools are reviewed.
- In Ch. 3, the reconstruction and identification of objects, such as jets, electrons, muons, used in the CMS experiment is discussed.
- The event selection is motivated and described in Ch. 4. It is followed by the definition of the search regions.
- The estimation of the SM background using data is explained in Ch. 5.
- Chapter 6 provides a description of systematic uncertainties.
- The final results of the main search regions as well as the aggregated search regions are presented in Ch. 7.

Chapter 1

Supersymmetry: an extension of the Standard Model

This chapter commences with a discussion of the 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 models, as one of the possible theories Beyond the Standard Model (BSM). The Supersymmetry (SUSY) discussion will include the framework of Simplified Model Spectra (SMS): The results presented in this thesis are interpreted with in SMS.

The details of this chapter can be found in various textbooks and reviews on the SM, quantum field theory [27–31] and SUSY [32].

1.1 Standard Model

1.1.1 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 and W 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 (EWK). 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. Fig. 1.1 shows the particle content of the SM¹. All are experimentally confirmed. The range of these forces is inversely propor-

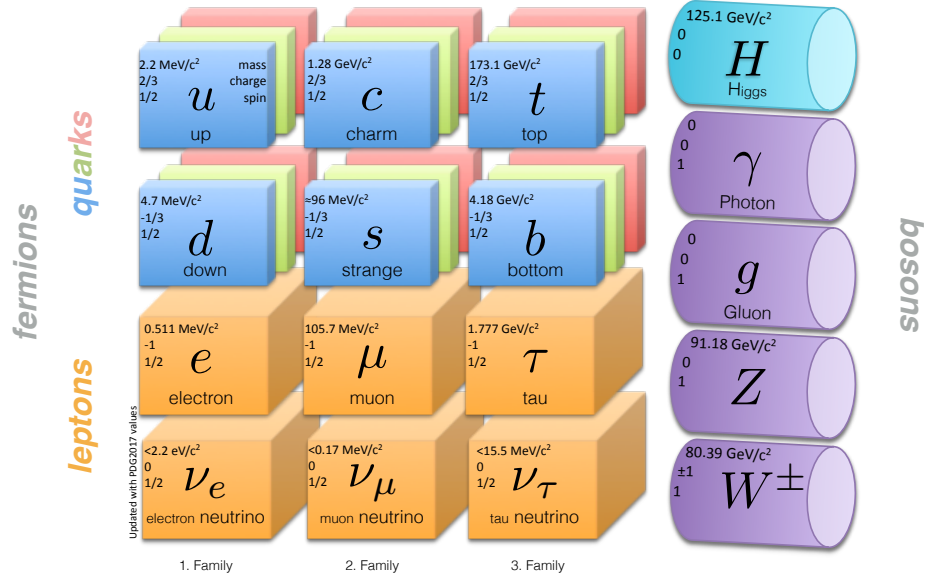


Figure 1.1: All fundamental particles of the SM, the gauge and Higgs bosons are shown in the diagram. The electric charge, spin and mass (or limit for neutrinos) are given in the corners of the boxes [33].

tional 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 Particle interactions

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

After observation of bound states such as $\Delta_{(uuu)}^{++}$ [34, 35], violating Pauli's exclusion principle, it is suggested that the quarks possess three color (red, blue, green) charges [36, 37]. 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 non-abelian 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 reading, [38, 39] can be consulted.

QCD has two important postulates.

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

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. 3.3, 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. This asymptotic freedom helps to build a perturbative description of QCD interactions; in the non-perturbative regime, QCD calculations are extremely difficult.

The interaction of charged particles with energies of the order of the W boson mass can be described by **quantum electrodynamics** (QED). The symmetry group of the electromagnetic theory is a unitary transformation $U(1)_{EM}$ and there is a massless field associated to the photon. The β -decay, where the neutron decays to a proton, electron and corresponding antineutrino through the W boson was explained by Enrico Fermi by an effective theory [40]. The theory is based on an $SU(2)_L$ symmetry group. The subscript “L” stands for the left handed particles. Due to the V-A structure of the coupling to fermions, only left handed fermions contribute to the **weak interaction**. In 1960s, Glashow-Salam-Weinberg proposed a unified theory of the QED and weak interactions [1–3]. Brout, Englert [4] and Higgs [5] solved the problem of gauge symmetry violating mass terms in the non-abelian $SU(2)$ gauge theory by implementing Electroweak Symmetry Breaking (EWSB), called BEH mechanism (or simply Higgs mechanism). This 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 the weak interaction gain their mass through a non-zero vacuum expectation value (VEV) while for fermions direct Yukawa couplings to the Higgs boson provide the mass term. 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. The 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 From the quantum field theory window

In particle mechanics, a Lagrangian³ is a function of the coordinates, and their time derivatives. In field theory 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 on equal footing. The Lagrangian plays an important role in physics because it encodes both the dynamics and the symmetries of the theory. Given a Lagrangian, the equation of motion from the Euler-Lagrange equations can be derived by considering the

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.

least action principle, i.e requiring that the variation of action is zero⁴. Then, the Euler-Lagrange equation can be written as:

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

The SM is a relativistic Quantum Field Theory (QFT). Its Lagrangian is built on a global Poincaré symmetry, which implies symmetry under translations, rotations and Lorentz boosts. According to Noether's theorem [41], each continuous symmetry is accompanied by a conservation law. The Poincaré symmetry implies the conservation of energy, momentum, and 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 Eq. 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 such that the kinetic term stays invariant under gauge transformations. In general, the form of the covariant derivative is $D_\mu = \partial_\mu - igX_\mu$ where with the g coefficient fermion interacts with the X boson. Concretely, the kinetic terms are read as $\bar{\Psi}\gamma^\mu D_\mu \Psi$, and the covariant derivative for QCD is defined as: $D_\mu = \partial_\mu - ig_s T_a G_\mu^a$.

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

⁵ $N_c^2 - 1$, where $N_c = 3$

⁶ f_{abc} are the structure constants

- **Electroweak interaction**

The electroweak Lagrangian, L_{EWK} , can be written as a sum of four contributions:

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

The covariant derivative that leaves this Lagrangian invariant under gauge 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 denote the Pauli matrices. The weak hypercharge is denoted by Y_W and it is defined as $2(Q - I_3)$. The electric charge is denoted by Q and I_3 is the third component of the weak isospin. The first term in Eq. 1.4 is defining the interaction among 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 Eq. 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:

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

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

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

In the case where m_H^2 in Eq. 1.8 is positive, the potential acquires a ground state at origin. Then the theory is in the form of QED with a massless photon and charged scalar field ϕ with a mass m . However, if $m_H^2 < 0$, the potential has then an infinite number of minima. The shape of the potential is generally referred as a ‘‘Mexican hat’’. In this ground state, system has a broken symmetry and the Higgs field can be written as:

$$|\Psi| = \sqrt{\frac{-\mu^2}{2\lambda}}. \quad (1.10)$$

A direction choice of the Higgs field as: $\phi^+ = 0$, $a^0 = 0$ and $\phi^0 = \sqrt{\frac{-\mu^2}{2\lambda}} = v$ ($v \equiv VEV$), results in three massless Goldstone bosons and one massive Higgs boson. Moreover, these

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

Goldstone bosons disappear when gauge invariance is required. The fluctuation around the minimum v can be written as :

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.11)$$

where the scalar field $h(x)$ portrays a physical Higgs boson. In the expanded version of the potential the coefficient of the h^2 term gives the Higgs boson mass:

$$M_H^2 = 2\lambda v^2, M_H = \sqrt{2}|m_H|. \quad (1.12)$$

The final term in Eq. 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 breaking of the SM gauge symmetry (SSB). It can be written in the most general way [31]:

$$L_{Yukawa} = -\epsilon^{ij} \phi_i \ell_{jI} y_{Ij} \bar{e}_j - \epsilon^{ij} \phi_i q_{\alpha j I} y'_{Ij} \bar{d}_j^\alpha - \phi^{\dagger i} q_{\alpha i I} y''_{Ij} \bar{u}_j^\alpha + h.c. \quad (1.13)$$

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 desirable:

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B^0 \\ W^3 \end{pmatrix} \quad (1.14)$$

$$W^\pm = W^1 \pm iW^2. \quad (1.15)$$

The θ_W term is the Weak mixing angle or Weinberg angle. This quantity is measured experimentally as well. Rewriting the Lagrangian in Eq. 1.5 in terms of the physical gauge bosons, the mass terms for massive bosons V arise in the form of $\frac{1}{2} M_V^2 V_\mu^2$. Using the coefficients masses of the charged and neutral bosons can be written as:

$$M_{W^\pm} = \frac{1}{2} v g, \\ M_Z = \frac{1}{2} v \sqrt{(g^2 + g'^2)}. \quad (1.16)$$

Since g and g' are free parameters, the SM makes no absolute predictions for M_W and M_Z . However, it was possible to set limits before the discovery of W , Z bosons. Considering the muon decay, using the relations in Eq. 1.16 and the Fermi constant ($1.167 \cdot 10^{-5} \text{ GeV}^{-2}$), v is found as 246 GeV. Looking back at the mass term for the mass of the Higgs boson in Eq. 1.12, since λ is a free parameter, the mass of the Higgs boson is not predicted in the SM.

1.1.4 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 [6, 7]. 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 [8–10]. Further studies with a much larger data set have provided precise measurements of its mass, production properties and decay rates. Results are consistent with the SM within uncertainties. 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, directly probing the energies much larger than TeV scale is not possible with collider experiments and higher energies must be studied in e.g. astrophysical observations. In addition to cosmological and collider experiments, there are smaller scale table-top experiments where particle properties can be further investigated with high precision at low energies ($\leq$ GeV). In Fig. 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, including benchmark experiments. The colors coarsely indicate the degree of the consistency of the measurements with the SM. 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. The apparent 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 ($\gg$ TeV)
High precision			
Measurement	• Particle content • Electro weak unification	• Anomalous magnetic moments • Electric dipole moments • Mass measurements	• 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 [33]. Additionally further information on g-2 measurements can be found from the paper published by Muon G-2 experiment [42].

1.1.5 Shortcomings

As mentioned in the previous section, no evidence of new physics beyond the SM has been found at the LHC. Moreover, up to now, the SM has successfully explained the world of subatomic physics. However, there are reasons hinting at the incompleteness of the SM.

1.1.5.1 Experimental reasons

Lacking explanation of gravity: Certainly, the observation of gravitational waves is one of the two most exciting discoveries of our decade. The announcement was done by LIGO and Virgo collaborations on 11 Feb 2016 [43]. The observed waveform satisfies the predictions of general relativity [44]. However, the SM does not contain gravitational interactions and a unification with general relativity is not in sight.

No Dark Matter candidate: In 1933, Fritz Zwicky observed a discrepancy between the observed gravitating mass distribution in a galaxy and the velocity distribution of stars by applying the Virial theorem, which relates the gravitational potential energy of a system to its kinetic energy, to the Coma cluster. He introduced this unseen mass as Dark Matter (DM), originally in German called *dunkle Materie* [45]. The evidence for DM became stronger with the observations made by V. Rubin and K. Ford on the velocity distribution of more than twenty spiral galaxies [46]. 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 [12] and WMAP [13] experiments, is in good agreement with the Λ CDM (Lambda cold dark matter) model [11]. A fit to the model returns 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 candidate⁸, in fact the SM fails to explain 95% of content of our universe.

Neutrino masses: The observation of neutrino oscillations [47, 48] implies that at least two of the neutrinos should have non zero mass⁹. Neutrino masses can be included in the SM by means of a Yukawa coupling term. However, it is difficult to explain their smallness. The various theoretical models that alleviate this apparent hierarchy, go beyond the SM.

1.1.5.2 Theoretical reasons

In addition to the observations described above, the SM has also theoretical shortcomings.

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 unification of electroweak and strong interactions. But in fact, the running couplings of the electromagnetic, weak and strong forces, as seen in the Fig. 1.3 (left), do not meet at any energy scale. The renormalization group evolution of the SM couplings does not unify and hence it's not possible to embed the group structure,

⁸In this context, Neutrinos are hot (relativistic) particles

⁹The oscillation frequencies are proportional to the mass difference of the neutrino flavors. Therefore, only upper bounds can be measured.

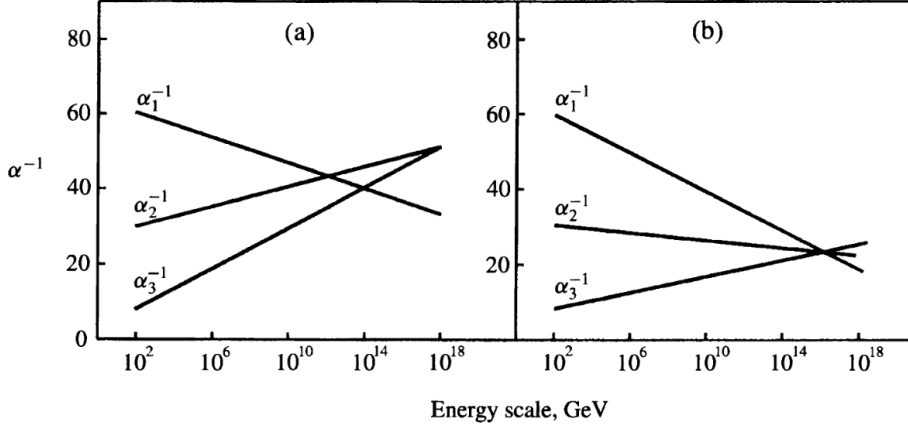


Figure 1.3: The running couplings of U(1), SU(2), SU(3) is denoted by α_1 , α_2 , α_3 respectively. Figure represents the evolution of inverse of couplings with energy scale, for (a) non-supersymmetric SU(5) and (b) supersymmetric SU(5). Figure is taken from [49].



Figure 1.4: Feynman diagrams for 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 [32]

which is described in the previous sections, in a single gauge group.

Hierarchy problem: Loop corrections from heavy fermions in the SM, most prominently the top quark, to the SM Higgs boson mass suggest that the natural value of m_H is at the GUT scale and not at the EWK scale. This subject will be discussed carefully in the upcoming sections.

There are many theories Beyond the Standard Model (BSM) claiming to solve some of the aforementioned SM shortcomings. The candidate BSM theory is expected not to belie the 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 the variety of BSM theories which satisfies these constraints, the Supersymmetric extensions to the SM can be considered as the most promising ones. This thesis focuses on a specific subset of supersymmetric models which will be discussed in the following section.

1.2 Supersymmetry as a solution

The Coleman-Mandula theorem [50] restricts the symmetry group of a wide class of local quantum field theories, including most extensions of the SM, to a direct product of the Poincaré group and an internal symmetry group. It was later found that anti-commutators can eschew this limitation by extending the Poincaré to the Superpoincare group [51]. The new symmetry relates fermionic states to bosonic states. This theoretical concept is known as SUSY. The supersymmetric transformations change the spin of the original SM fields by $\frac{1}{2}$ and thus turns fermionic states into bosonic states, and vice versa. The new fields are called superpartners. If these superpartners exist in TeV scale they can help overcome several problems mentioned in Sec. 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, the 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 Eq. 1.8, receives large quantum corrections. The one-loop radiative correction terms (see Fig. 1.4 left) from a Dirac fermion f is

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \dots \quad (1.17)$$

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 SM breaks down. Because the top quark is the one with the largest mass among the SM fermions¹⁰ ($\lambda_f \approx 0.94$), it gives the largest contribution to m_H^2 . 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 required value of m_H^2 for the Higgs boson with a mass around 125 GeV. In addition to the correction term given in Eq. 1.17, a similar correction is predicted for a massive scalar particle S (see Fig. 1.4 right) as:

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} \Lambda_{UV}^2 + \dots \quad (1.18)$$

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 and couplings conspire such that $\lambda_S = \frac{1}{2}|\lambda_f|^2$. SUSY overcomes the hierarchy problem by imposing superpartners, which possess the same quantum numbers with original fields except for the spin. 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. The discussion will be further expanded in Sec. 1.2.1.

¹⁰Eq. 1.17 should be multiplied by 3 to account for colors.

Dark matter candidate:

The lower limit on the proton lifetime is set to 10^{33} years by experiments carried out in the water Cherenkov detector at Super Kamiokande in Japan [52]. This fact strongly suggests that the baryon number is conserved. Unlike the SM, in SUSY theories the baryon and lepton number conservation can be violated. In SUSY, to prevent proton decay, a new quantum number, called as R -parity is introduced:

$$R = (-1)^{3B+L+2S}, \quad (1.19)$$

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 [53] 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) [54, 55]. In many models, the LSP is electrically neutral and has no color charge. Therefore, it interacts at most weakly. Depending on the model parameters, SUSY theories predict a wide range of LSP masses. Constraints on the LSP mass come from cosmological observation of the DM relic density. In particular we see that, assuming a standard thermal evolution of the universe, the correct observed relic density [11] is achievable if the new physics lies around the EW scale. This represents a good motivation to look for SUSY at the LHC, where the LSP can be directly produced, or result from the decays of heavier sparticles. So-called R -parity violating SUSY theories do not respect this symmetry and therefore do not typically predict a pair of LSPs in the final state. While dedicated collider searches exist, these models 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 Fig. 1.3 a) [56, 57]. In Ref. [56], for unification of gauge coupling constants (see Fig. 1.3 b), 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 [58]. The MSSM and its particle content will be explained in Sec. 1.2.2.

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 chiral¹¹ SM fermions, there is a separate scalar partner and together they form chiral supermultiplets. The supersymmetric transformation operator Q can also be called as supercharge operator, can be read as:

$$\begin{aligned} Q|fermion\rangle &= |boson\rangle \\ Q|boson\rangle &= |fermion\rangle. \end{aligned} \quad (1.20)$$

¹¹left or right eigenstate of the chirality projector

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, the generators Q and $Q^\dagger$ satisfy the commutation and anti-commutation relations:

$$\{Q, Q^\dagger\} = P^\mu, \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0, [P^\mu, Q] = [P^\mu, Q^\dagger] = 0 \quad (1.21)$$

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 Eq. 1.21 implies that the partner and the superpartner states have the same mass. However, no scalar superpartners have been found yet, although, e.g. the selectron with mass m_e would be easy to detect. This indicates that the supersymmetry is broken and the superpartners have much larger mass.

1.2.2 Minimal Supersymmetric Standard Model

The MSSM is the supersymmetric extension of the SM with the minimal particle content. In other words, this is the model with the least number of additional particles and degrees of freedom. The MSSM chiral supermultiplets are shown in Tab. 1.1 while the gauge multiplets are listed in Tab. 1.2. As in the table, all superpartners are represented by a version containing a tilde ($\tilde{}$) 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.

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

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 $Y = -\frac{1}{2}$ gives masses to the down-type quarks and charged fermions. Therefore, the $SU(2)_L$

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

Table 1.2: Gauge supermultiplets of the MSSM

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 the half-integer spin color octet supersymmetric partner of the gluon. The superpartners of the electroweak gauge bosons are called winos and binos.

EWSB in the SUSY sector leaves the gauge quantum numbers intact and therefore the higgsinos and electroweak gauginos mix with each other. The mixing of $\tilde{W}^0$ and $\tilde{B}^0$ are called zino ($\tilde{Z}^0$) and photino($\tilde{\gamma}$). The mixing between neutral higgsinos and the neutral gauginos results in four mass eigenstates called neutralinos and denoted by $\tilde{\chi}_i^0 (i = 1, 2, 3, 4)$. Similarly, the charged higgsinos and winos combine to form two mass eigenstates called charginos denoted by $\tilde{\chi}_i^\pm (i = 1, 2)$.

The soft breaking of the MSSM is introduced by adding a term (L_{soft}^{MSSM}) to the Lagrangian. The Lagrangian, then, can be written as:

$$L = L^{MSSM} + L_{soft}^{MSSM}, \quad (1.22)$$

where the Yukawa and gauge interactions are contained in the first term. The soft terms are chosen such that the correction terms of the Higgs mass contain only logarithmic divergences:

$$\Delta m_H^2 = -m_{soft}^2 \left[\frac{\lambda}{16\pi^2} \ln(\Lambda_{UV}/m_{soft}) + \dots \right]. \quad (1.23)$$

In order to satisfy the naturalness 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 the 19 SM parameters. These 105 additional parameters are categorized as follows: 48 CP-violating phases in the gaugino/higgsino and squark/slepton sectors, 21 squark/slepton masses, and 36 mixing angles, which modify the mixing of the gauge eigenstates into the mass eigenstates.

The large number of parameters of the MSSM make the interpretation of the experimental results challenging. Therefore, using universality assumptions at the GUT scale, the number of free parameters is reduced to five. This model, appropriately, is called the constrained MSSM (cMSSM) and a wide range of its low-mass parameter configurations are excluded by the Higgs mass measurement [60].

1.2.2.1 Particle decays

To stay in the scope of this work, only the R -parity conserving scenarios are 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. For more details see Ch. 9 of [32].

Neutralino and chargino are the mass eigenstates of the higgsinos and gauginos which 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:

$$\begin{aligned}\tilde{\chi}_i^0 &\rightarrow Z\tilde{\chi}_j^0, W\tilde{\chi}_j^\pm, h^0\tilde{\chi}_j^0, l\tilde{l}, \nu\tilde{\nu}, \\ \tilde{\chi}_i^\pm &\rightarrow W\tilde{\chi}_j^0, Z\tilde{\chi}_j^\pm, h^0\tilde{\chi}_j^\pm, l\tilde{\nu}, \nu\tilde{l}.\end{aligned}\tag{1.24}$$

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 the dominating channels if $m_{\tilde{t}}$ and $m_{\tilde{b}}$ are the smallest squark masses. Otherwise the squarks are off-shell and this favors the following decays: $\tilde{g} \rightarrow qq'\tilde{\chi}_i^0$ and $\tilde{g} \rightarrow qq'\tilde{\chi}_i^\pm$.

1.2.3 Simplified models

The SMS [61] signal model covers only a part of the parameter space of the 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, a limited number of decay channels are considered. The branching ratios are often assumed to be 100% or, alternatively, combinations of such models is used for mixed decays. In this way, results can also be reinterpreted within other (non-)SUSY theories.

In a pair production process, primary particles can decay directly or undergo a cascade decay through an intermediate new particle. In each of the models, the particle decay chain ends with a neutral, undetected particle, LSP. In the direct decay case, the parameters include the mass of the primary particle and the LSP. When a decay includes an intermediate particle state, the mass of the intermediate particle is also included in the parameters. To reduce the three-dimensional mass space, intermediate particle mass (m_{int}) is considered to be dependent to primary particle (m_{primary}) and LSP masses (m_{LSP}). The relation is given by: $m_{\text{int}} = m_{\text{LSP}} + x \cdot (m_{\text{primary}} - m_{\text{LSP}})$. The kinematics of the decay strongly depend on the value of x .

A variety of SMS approaches has been investigated by the ATLAS and CMS collaborations at the LHC. An example for the direct decay is a model of gluino pair production, where each gluino decays to top quark pair and a neutralino. This model is called T1tttt¹². An example for a cascade decay can be a model of gluino pair production, where each gluino decays to two light quarks and an intermediate chargino, with the latter decaying to a

¹²this naming is used within the CMS collaboration

gauge boson and a neutralino. When the gauge boson is a W boson, the model is called T5qqqqWW. The diagrams for T1tttt and T5qqqqWW can be seen in Fig. 1.5 left and right respectively. In this thesis, the T5qqqqWW model, with the parameter $x=0.5$, is used to interpret the results which is well motivated in the searches at LHC [26, 62].

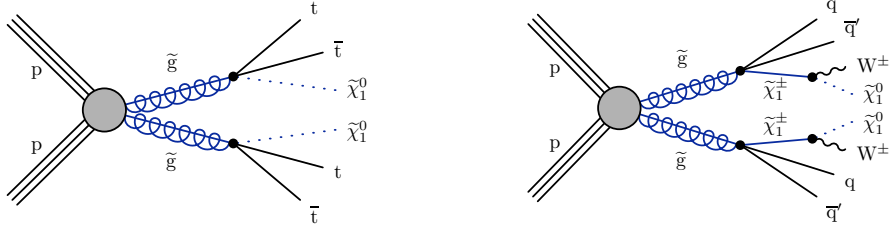


Figure 1.5: Diagrams showing the simplified models T1tttt (left) and T5qqqqWW (right).

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 [63]. 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. The Higgs mass measurement at around 125 GeV excludes a large range of low-mass cMSSM parameters [60, 64]. 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 Fig. 7.8 including the results from the present search. These results have already pushed the sparticle masses to values where sizable fine-tuning is present in the MSSM. 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 [65]. Nevertheless, there is still room for expanding the searches, for example, in the regions with compressed spectra detector acceptance reduces sensitivity.

Expectations for gluino searches at 13 TeV:

Gluino pairs can be produced in LHC, like any other hadron collider, by gluon fusion and quark-anti-quark annihilation (see Fig. 1.6). As it can be seen in Fig. 1.7 (left), the cross sections of gluino pair production are exponentially increased with the center of mass energy. The black line in the figure corresponds to $m_{\tilde{g}} = 1.5 \text{ TeV}$, and the production cross

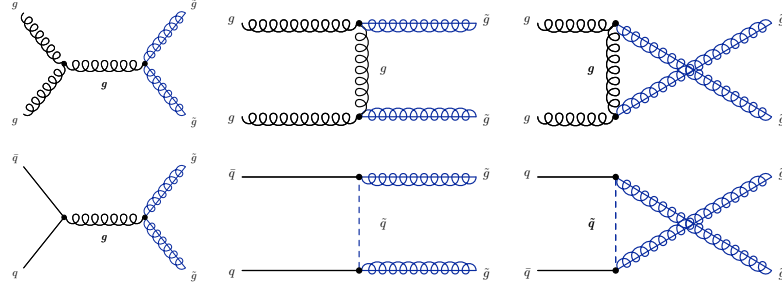


Figure 1.6: Feynman Diagrams showing the gluino pair production. SUSY particles are shown in blue.

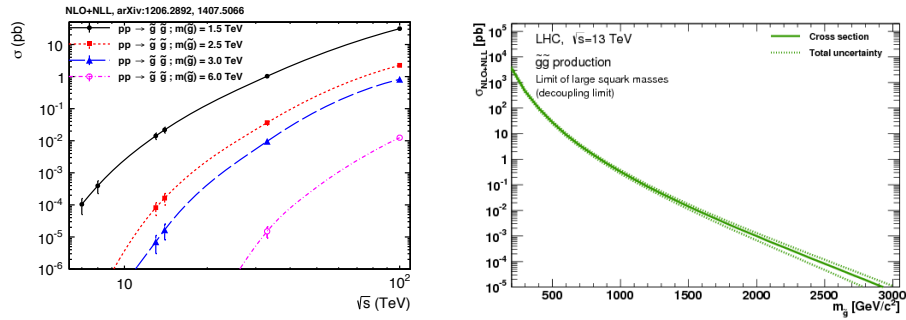


Figure 1.7: left: Cross sections of gluino pair production for four different mass points are given as a function of the center of mass energy. right: Cross section of gluino pair production at 13 TeV center of mass energy is given as a function of gluino mass [66].

section increases from 0.4 fb (8 TeV) to 14 fb (13 TeV). On the other hand, the expected cross sections of SM processes such as the production of b-quarks, $W^\pm$ and Z-bosons, jets for proton-proton scattering at LHC is large (see Fig. 1.9). Given that the total cross section of SM processes is around 10^8 nb, while the gluino production cross section is 14 fb, 1 event including gluino pair production is expected for 10^{13} SM events. Nevertheless, it is still feasible to design a robust analysis strategy. An example of such a search will be presented in this thesis and in Sec. 7.4 a comparison of the latest results from gluino searches can be found.

Earlier results 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 Run 1 with a search for a similar model, which was T1tttt. In this model, the presence of two top-antitop quark pairs requires a search for multiple b-tagged (see Sec. 3.3.1 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. [67]. In Run 2, to increase the discovery potential of the analysis, the channels, which are sensitive to zero b (tagged) final states, are included. These sensitive processes are called T5qqqqWW models. The results had already been presented in the *Moriond 2016* conference in 2016 [68], 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 Fig. 1.8, the exclusion of the masses, for $m_{\tilde{g}}$ and $m_{\tilde{\chi}_1^0}$ is shown for the model T1tttt (left) and T5qqqqWW (right).

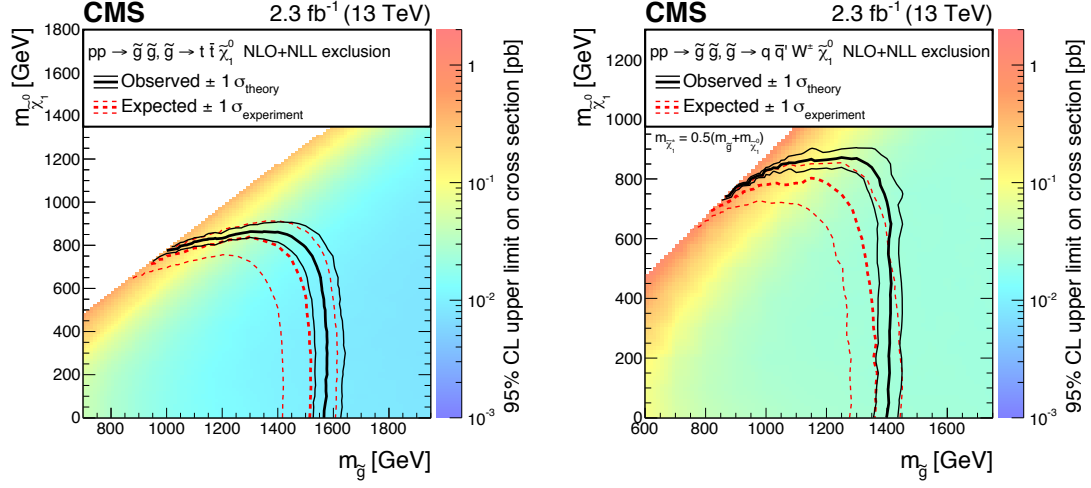


Figure 1.8: Cross section limits at a 95% CL for (left) the T1tttt model, and (right) the T5qqqqWW model [68].

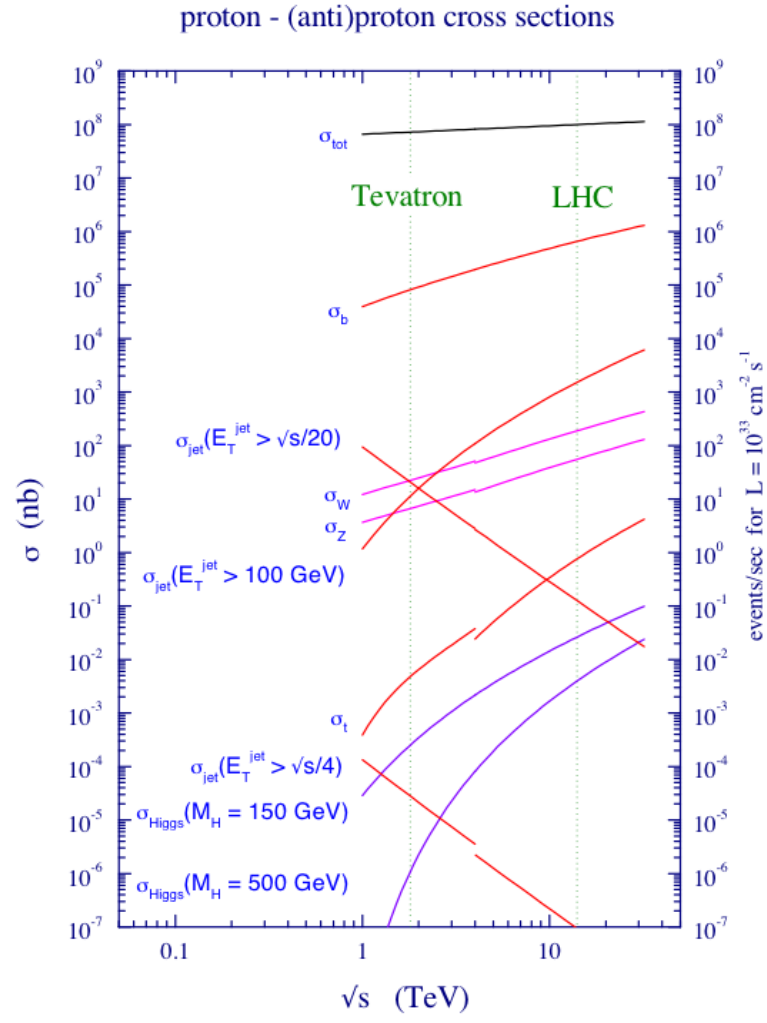


Figure 1.9: Cross section of SM processes as a function of center of energy [69].

Chapter 2

Experimental setup

An overview of the experimental setup used to collect data for this analysis is explained in this chapter. The proton-proton collision data, which is recorded by the CMS experiment at the 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. 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 The LHC at CERN

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 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 [70].

Figure 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 the LHC is used. The LHC is situated about 100 meters below the ground, between the Jura Mountains and Lake Geneva. It 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 many experiments at CERN since 1978 and it will be replaced by newly commissioned LINAC4 in 2020 [72]. The beams provided by the LINAC are then injected to the Proton Synchrotron Booster where the particles are further accelerated to 1.4 GeV [73]. Next, the protons delivered by the Booster are accelerated up to 25 GeV by the Proton Synchrotron (PS) [74]. This accelerator is operating since 1959 and it is the first synchrotron of CERN. The last machine before the LHC is the Super Proton Synchrotron (SPS) [75], where the protons reach the energy of 450 GeV. The SPS was operating as a proton-antiproton collider from 1981 to

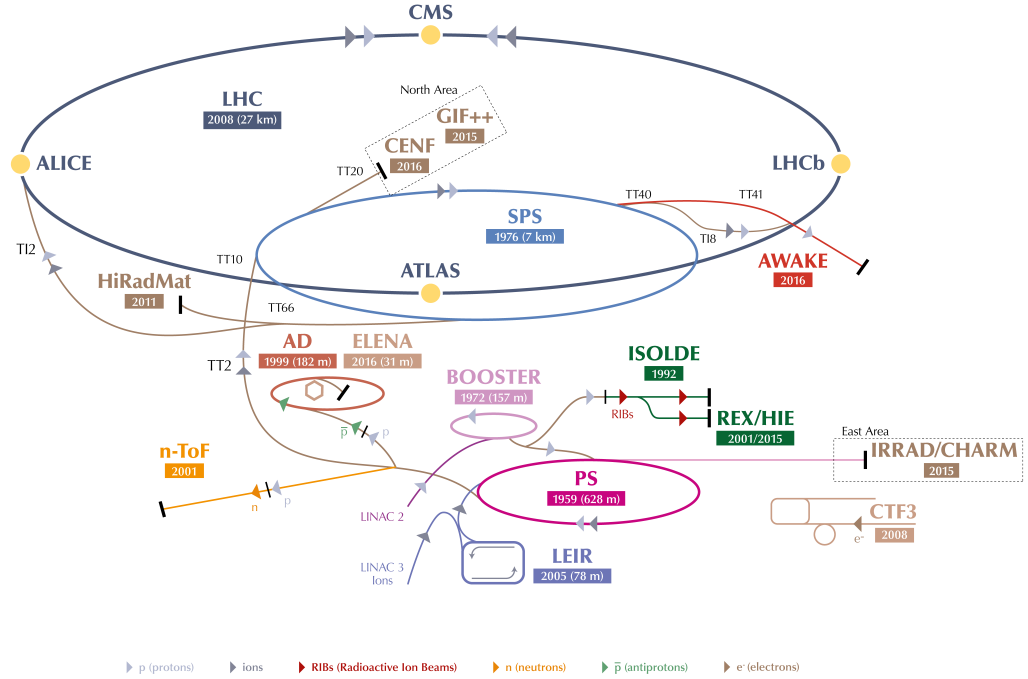


Figure 2.1: CERN accelerator complex [71]

1984. The studies resulted in the discovery of W [76] and Z [77] bosons. This journey of protons to reach the LHC takes 16 minutes in the current injection scheme [78].

2.1.2 The LHC

The LHC is primarily designed to collide protons at a center-of-mass energy of 14 TeV. The LHC tunnel was 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 were 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.

The proton bunches from SPS are injected as counter-rotating beams in the two rings of the LHC, each of them has an energy of 450 GeV. The beam acceleration takes place at a point between the ALICE and the CMS experiments, where 16 superconducting cavities are mounted. Each proton gains an energy 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 [78].

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 $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ would be achieved. However, to reach these conditions, prior studies and commissioning of the machine at lower levels should be performed. Therefore, the first long proton-proton (pp) run of the LHC was performed at a centre-of-mass energy of 7 TeV with a peak luminosity of $\mathcal{L} = 3.6 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. Total integrated luminosity of around 5 fb^{-1} was delivered to the major ATLAS (A Toroidal LHC ApparatuS) and CMS experiments until the end of 2011 [79]. Next, following the 8 TeV run that ended LHC Run 1 in 2012, the first long shutdown, which took place for 1.5 year, allowed protons to accelerate to their current energy of 6.5 TeV. In 2015 the second LHC run, namely Run 2, has begun with a peak luminosity of $\mathcal{L} = 5 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ at a bunch spacing of 25 ns. A full list of the LHC parameters for the 13 TeV runs can be accessed here [80]. The present analysis uses the data collected in 2016 by the CMS detector. The peak instantaneous luminosity (left) and the measured integrated luminosity (right) on a day-by-day basis are shown in Fig. 2.2. Recently, in June 2017, a new luminosity record ($\mathcal{L} = 1.58 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$) is achieved [81]. The LHC has exceeded its design luminosity, with this new record.

The Future of the LHC

According to the LHC plans, the current research program will continue until 2023 including Run 3 with 14 TeV center-of-mass energy and the target integrated luminosity is 500 fb^{-1} [84]. LHC further future plans are constrained by the fixed ring size, thus the limited proton energy. Therefore the LHC proceeds in the direction of the intensity frontier. High-Luminosity (HL-LHC) project is planned to start at the beginning of 2024. HL-LHC could reach luminosities of the order of $\mathcal{L} = 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ which is ten 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.

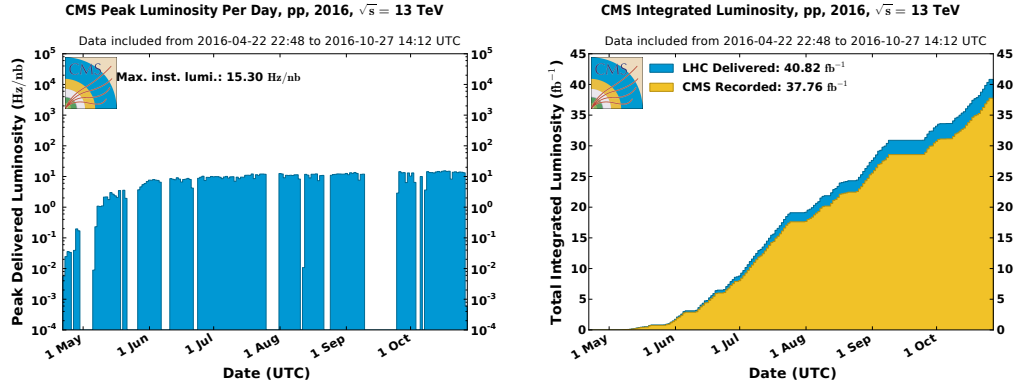


Figure 2.2: Left figure shows the cumulative measured luminosity versus delivery day [83]. Right figure shows peak instantaneous luminosity on a day-by-day basis. The maximum peak luminosity measured is $15.30 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$.

2.2 The CMS detector

The CMS detector is a multi-purpose device designed 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 [85], it was followed by the construction of the individual components soon afterwards.

As it can be seen from the Fig. 2.3, the CMS detector has an 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 the 6m wide solenoid magnet. The outermost layer is the extensive system of muon trackers which in fact makes up more than 80% of the detector volume. In addition to these, forward calorimeters are placed along the beam-pipe immediately inside 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 points vertically upward, and the x-axis points 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, a 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 θ , is also a widely used variable and it characterizes the boost of the particles: $\eta = -\ln[\tan(\frac{\theta}{2})]$. The pseudorapidity is an approximation 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 spatial distance: $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The proton is a composite particle, which includes partons. Therefore, the proton-proton collisions are interactions of the corresponding partons. Furthermore, even though the

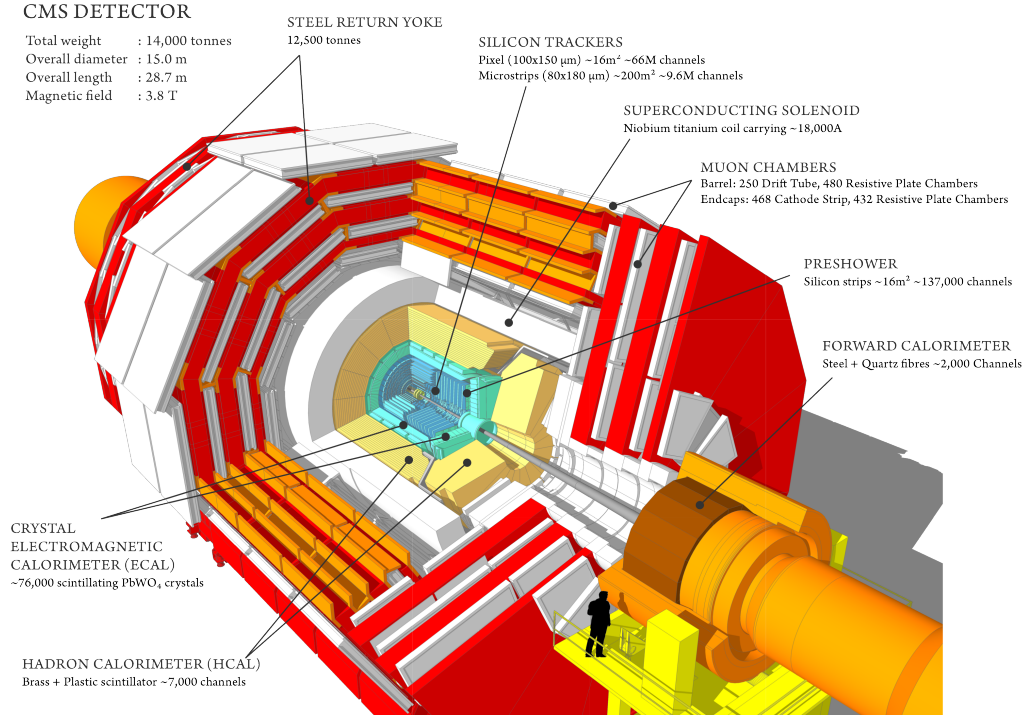


Figure 2.3: Schematic view of the CMS detector with a cut-out quadrant [86]

energy of the incoming proton beams are symmetric, this condition does not necessarily need to hold for individual partons. Therefore, the center-of-mass energy of the parton collision is an unpredictable quantity of an event. However, since the incoming protons, therefore the 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, 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 4 T within its volume. In order to increase its lifetime, it is currently being operated at slightly lower field strength of 3.8 T. Further information on the measurement of the magnetic field inside the barrel yoke can be found here [87].

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 constraints in terms of detection efficiency, spatial 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. Ten layers of silicon strip detectors are located after the pixel detector. The tracker covers the pseudorapidity range of $|\eta| < 2.5$.

2.2.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) comes after the Tracker detector. ECAL is a hermetic homogeneous detector, which is composed of lead tungstate (PbWO_4) crystals. The lead tungstate scintillators lead to a compact calorimeter design, which is important to deal 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 [88]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2, \quad (2.4)$$

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. The HCAL is a sampling calorimeter in other words it finds a particle's position, energy and arrival time using layers of brass and scintillator plates. When the particle passes through the scintillator layers, a rapid light pulse is produced. Then optic fibers feed the light into readout boxes where photodetectors amplify the signal. A particle's 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 that are 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 (HF). The HF is installed outside the CMS endcaps at a distance of $|z| = \pm 11.2$ 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. This 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. The 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 the barrel region the magnetic field is uniform and the muon flux is low. Therefore, drift tubes (DT) are used in this section corresponding to a pseudorapidity range of $|\eta| < 1.2$. The drift tube system serves as tracking detector by measuring the electromagnetic cascade 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 (RPC) are installed to supplement the triggering system in a fast and independent way. In the muon reconstruction, information 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 25 ns time interval between the proton bunches, the collision rate inside the detector reaches up to 40 MHz, which is the LHC design frequency. An event with all the measured information occupies about 1 MB in size. To save all events, a bandwidth of 60 TB/s would be needed. However, achieving such high transmission of data is impossible with today's 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 100 kHz 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 of several 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 Eq. 2.3. The instantaneous luminosity is measured using two dedicated detectors: the Fast Beam Conditions Monitory (BCM1f) [90] 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.

The total data collected by the CMS experiment in 2016 which corresponds to a 35.9 fb^{-1} integrated luminosity with a measured uncertainty 2.5% [91] is used in this thesis.

2.3 Event simulation

Monte Carlo (MC) simulations provide events close-to-reality. Both the experimental and theoretical physics communities benefit from MC simulations to study the dedicated physics processes as well as the detector response. In the following, a short overview of simulation techniques 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. 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. Finally, hadrons are formed using hadronization models. PYTHIA 8.212 [92] with the CUETP8M1 tune [93] is used for parton shower and hadronisation.

MC samples used in this thesis are simulated with the MADGRAPH5_AMC@NLOv2.2.2 or v.2.3.3 [94] and POWHEGv2.0 [95–99] generators. 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 [100].

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 material as well as the readout electronics of CMS. In CMS the main technique is called Full simulation, for which the GEometry ANd Tracking (GEANT4) toolkit is used [101]. However, the Full simulation requires massive computer power; as a result, for specific large-scale MC production, like SUSY parameter phase scans, a more effective technique is necessary. Therefore, CMS established a Fast Simulation (FastSim) software [102, 103]. Several features of the detector geometry and particle-matter interactions are simplified to reduce the computing time per event. Detector responses of FastSim samples are tuned to match to the full simulation.

Chapter 3

Object reconstruction and identification

The format 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 analysis.

In this chapter, the reconstruction and identification of objects, such as jets, electrons, muons, used in the CMS experiment is discussed. These objects and their kinematic characteristics are subsequently used for defining the event selection criteria, which will be discussed in Ch. 4.

Figure 3.1 displays the tracks of the particles as they pass through the sub-detectors of the CMS experiment. Only energy deposits in the most relevant sub-detectors are presented, because a particle can leave trace in multiple detector layers. The charged particle trajectories (tracks) and origins (vertices) can be determined from the signals (hits) in the tracker layers. The charged particle trajectories are bent due to the magnetic field and this allows to measure the electric charge and the momentum. After the tracker, electrons and photons are absorbed in the ECAL and leave their signatures as clusters of energy recorded in adjacent cells. Using this information, energy and direction of the electromagnetically interacting particles can be determined. Charged (ch.) and neutral hadrons are fully absorbed in the HCAL; however, they may initiate a hadronic shower in the ECAL as well. As in the case of electromagnetically interacting particles, energies and directions of hadrons can be estimated using the corresponding clusters. Muons produce hits in muon detectors which are actually additional trackers located outside the calorimeters. The only particles known from the SM and do not leave any trace in the detector are neutrinos.

3.1 Particle-Flow algorithm

The particle flow (PF) algorithm [105], uses the information obtained from all of the detector subsystems in the object reconstruction. The three main object collections that feed into the PF algorithm are the tracks reconstructed from hits in the tracker, energy deposits in the calorimeters, and muon candidates reconstructed from hits in the muon system. In order to achieve object reconstruction with a high efficiency and low fake rate, iterative

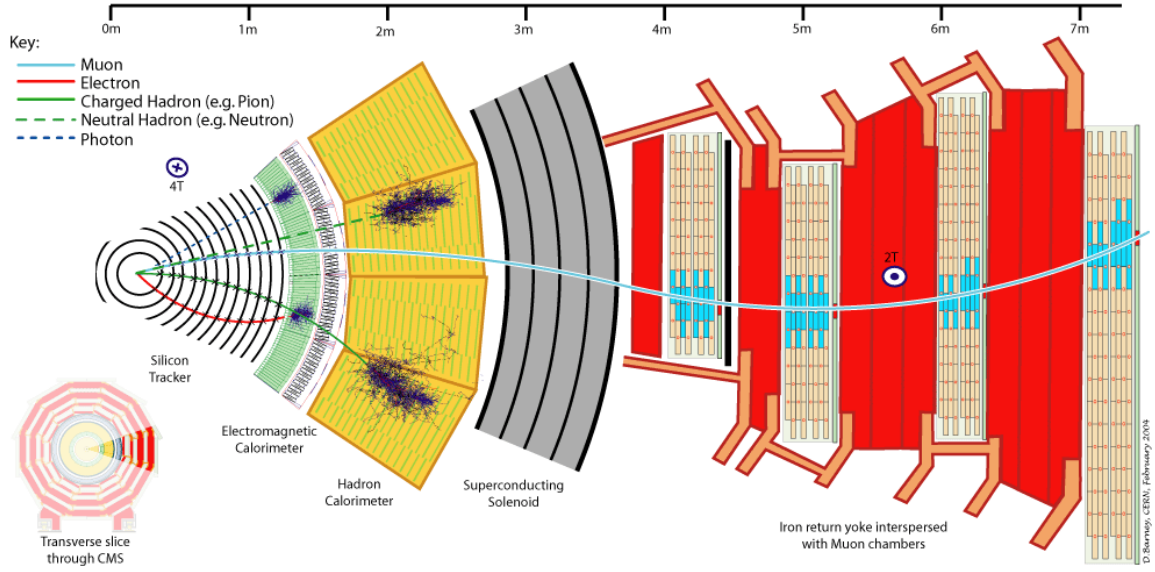


Figure 3.1: Cross section of a slice of the CMS detector in the transverse plane with different particles crossing the subsystems is shown [104]. Only energy deposits in the most relevant subdetectors are presented for simplicity.

tracking and the calorimeter clustering methods are used in the PF algorithm.

The Kalman filter [106] method is adopted to determine single particle tracks (See Sec. 3.2). Calorimeter clustering is responsible for four tasks: the detection and measurement of the energy and the direction of the stable neutral particles, their isolation from charged hadron energy deposits, reconstruction and identification of electrons and all corresponding Bremsstrahlung photons, and help the energy measurement of charged hadrons for which the track parameters were not determined accurately. After determining the single elements, a linking algorithm is used to form a block from separate elements such as tracks, muons, ECAL clusters, HCAL clusters that are possibly related to a single object. After the construction of blocks, the actual particle reconstruction and identification is performed.

The PF algorithm proceeds in steps according to the approximate level of ambiguity for each reconstructed objects; it starts with muons, then electrons follow, finally, it produces photons and neutral particle candidates from the remaining calorimeter clusters to arrive at a global event description. Following to each step, already identified blocks are gradually discarded.

For muons, the tracks in the muon system is re-fit including the (inner) tracker track to form "global" muons if the χ^2 is acceptable [107]. If the momentum of a global muon is compatible with the tracker-only measurement within three standard deviations, the muon hypothesis is kept in PF.

Electrons emit a bremsstrahlung radiation, which necessitates a complicated reconstruction procedure to properly associate the radiated energy with the electron candidate. The Gaussian sum filter algorithm (GSF) [108] allows such recovery and is hence used in the

electron reconstruction. To ensure an optimal energy containment, all ECAL clusters in the PF block, which are linked to the electron GSF track or to the supercluster, are associated with a PF electron candidate if stringent requirements on the compatibility of the track and the cluster are satisfied.

Having removed deposits associated with the reconstructed muons and electrons from the list of PF inputs, the neutral and charged hadrons are reconstructed from the remaining tracks. Naturally, for charged hadrons the matching calorimeter clusters are also used in this stage.

After the PF algorithm, the list of particle candidates (electrons, photons, muons, charged and neutral hadrons) is fed into the higher level algorithms for example jet clustering which are discussed in the following section.

In this work, and from now on, all event observables are based on the list of PF candidates as produced by the PF algorithm.

3.2 Tracks and primary vertices

As discussed in Sec. 2.1.2, the LHC is designed to run with a peak instantaneous luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ with the proton bunches crossing at intervals of 25 ns. In 2016, the peak luminosity measured by the CMS experiment is $15.30 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ (see Fig. 2.2). Figure 3.2 shows the recorded luminosity with respect to the mean number of interactions per bunch crossing in the 2016 pp run at 13 TeV with $\langle n_{\text{vertex}} \rangle = 27$. These multiple interactions are known as pileup and together with the finite time and energy resolution of the silicon detectors, reconstructing tracks and subsequently the primary vertices is a challenging task [109].

The purpose of primary-vertex reconstruction [110] is to determine the location of all proton-proton interaction vertices in each event and the associated uncertainty on them, including the ‘signal’ vertex and any vertices from pileup collisions. It uses the available reconstructed tracks and consists of three steps. 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 tracks are clustered using a deterministic annealing (DA) algorithm [111] which is finding the global minimum for a problem with many degrees of freedom. To overcome outliers that originate from a secondary or mismeasured tracks, the DA algorithm is extended with a rejection term which acts as a cutoff against the outliers. As a result the DA algorithm becomes a one-dimensional robust adaptive multi-vertex fit [112]. After determining candidate vertices with the DA clustering in z , the candidates containing at least two tracks are then re-fitted using an adaptive vertex fitter [113]. In the adaptive vertex fit, each track in the vertex is assigned a weight from 0 to 1. The tracks get a weight close to 1 when they are consistent with the position of the reconstructed vertex. The number of degrees of freedom:

$$n_{\text{dof}} = -3 + 2 \sum_{i=1}^{\# \text{tracks}} w_i, \quad (3.1)$$

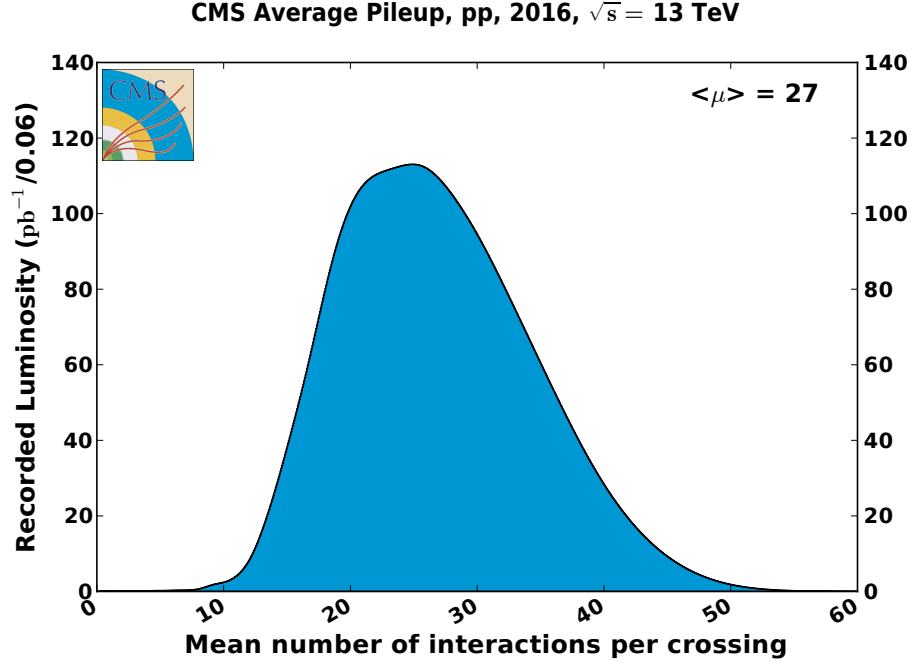


Figure 3.2: The recorded luminosity of data with respect to the mean number of interactions per bunch crossing in the 2016 pp run at 13 TeV [83]. The cross section is taken to be 80 mb.

where w_i is the weight for the i th track which is associated to a vertex. As a result, the value of n_{dof} is correlated with the number of tracks arising from the interaction region. Therefore, this can be also used to select true proton-proton interactions.

Finally, the number of good primary vertices includes vertices consistent with the luminous region (where the collisions happen), which is also known as beam spot, and with a $n_{\text{dof}} \geq 4$, corresponding to a minimum of four tracks.

The primary vertex of interest is chosen to be the vertex with the largest sum of the squared track momenta whereas the other vertices are considered as pileup vertices. The pileup vertex reconstruction and identification efficiency is about 70% [114].

In the present analysis, at least a good primary vertex is required in the event selection.

3.3 Jets

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 lead to collimated sprays of hadrons that are reconstructed as jets. The main purpose is to reproduce the energy of the original parton prior to the shower. Jets are clustered from all PF candidates using the anti- k_T algorithm [116]. The algorithm starts with clustering two of the closest objects together and iteratively reconstructs a closed area of objects until

a truncation criterion is satisfied. The distance used as the termination criterion can be formulated:

$$d_{i,j} = \min(k_{t,i}^{-2}, k_{t,j}^{-2}) \frac{\Delta_{i,j}^2}{R^2}, \quad (3.2)$$

$$d_{i,B} = k_{t,i}^{-2}, \quad (3.3)$$

where $d_{i,j}$ represents the distance between entities (particles, pseudojets) i and j with $R=0.4$ 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.

If the energy fraction of one of the components (e.g. the neutral energy fraction which is dominated by the HCAL subdetector) exceeds 99% then the jet is rejected, because the probability of a spurious jet from detector noise is high in such cases.

The measured jet energy differs from the corresponding parton energy. There are several reasons responsible for this difference such as the non-uniform detector response, non-instrumented regions and contributions from pileup. Four multiplicative correction factors are applied to the reconstructed jet to obtain the calibrated energy [115]: an offset correction, a MC calibration factor, a residual calibration for the relative energy scale, a residual calibration for the absolute energy scale. Furthermore, only jets with the calibrated transverse momentum larger than 30 GeV are selected.

3.3.1 Identification of b jets

Jets that originate from b quarks, are important to identify processes that e.g. involve top quarks. For T5qqqqWW (see Sec. 1.2.3), a veto on b-tagged jets can be used to reduce backgrounds from processes involving top quarks. Inverting the veto, in turn, allows a data driven background estimation of the contributions from $t\bar{t}$ events (see Sec.5.3).

The identification techniques [117, 118] utilize the rather long lifetime, $c\tau \simeq 450\mu\text{m}$, of the b quark which leads to the formation of a displaced vertex.

A variety of algorithms have been developed by CMS to select b jets based on variables such as the impact parameters of charged-particle tracks, the properties of reconstructed decay vertices, and the presence of a lepton in the jet. Each of these algorithms results in a single discriminator value for each jet. In this analysis, the combined secondary vertex (CSV) algorithm, which uses both secondary-vertex and track-based information, is employed. The medium working point (0.8484), which corresponds to a tagging efficiency of about 70% and a misidentification rate of about 1% according the p_T and η of the jet, is chosen.

The differences between the performance of the b-tagging algorithm when measured separately in data and simulation, are compensated by applying scale factors to simulated events, as described in Sec. 4.2.1.

3.4 Leptons

3.4.1 Muons

Muons are reconstructed in the pixel and silicon strip tracker (tracker track) and in the DT, the CSC, and the RPC of the muon system (standalone-muon track) [107]. The muon system provides a high purity while the inner tracker provides a precise momentum measurement. The candidate muons are further categorized into: global muon and tracker muon.

Each standalone-muon track is matched to a track in the inner tracker by requiring compatibility of the parameters of the two tracks. Then the matched hits from inner tracker and from standalone-muon track are combined and fit to form a global-muon track.

The tracker muon is obtained by requiring that each inner track with $p_T > 0.5$ GeV and total momentum p larger than 2.5 GeV is extrapolated to the muon system. The inner track is then called a tracker muon track if at least one muon segment is matched to the extrapolated track.

For tracks inside the geometrical acceptance, the overall reconstruction efficiency is 99%. As briefly mentioned earlier, the Pf algorithm uses the global and the tracker muon properties to identify muons. Isolated global muon are selected including the information from inner tracker and calorimeter energy deposits in a ΔR cone with radius 0.3 in (η, ϕ) plane in the muon direction. It is also required that the sum of the p_T of the tracks and of the E_T of the deposits do not exceed 10% of the muon p_T . This isolation criterion is sufficient to reject hadrons that are misidentified as muons. For the nonisolated global muons, the PF tight muon selection [107] is applied. After PF muons have been selected, they are labeled as tight or loose muon according to the selections that are summarized in Tab.3.1.

Muon Selection	
Loose	Tight
global or tracker muon	muon Id. in Tab.3.2
$p_T \geq 10$ GeV	$p_T \geq 25$ GeV
$ \eta \leq 2.4$	$ \eta \leq 2.4$
	$\text{sip}_{3D} \leq 4.0$

Table 3.1: List of selections for the tight and loose muon identification.

3.4.2 Electrons

Reconstruction of electrons is based on momentum information obtained from the tracker detector and energy measurements of the clusters in the ECAL [119]. The tracker material causes significant bremsstrahlung along the electron trajectory. Due to the CMS magnetic

Muon identification		
Fraction of valid tracker hits $\geq 80\%$	Segment compatibility ≥ 0.303	Global muon Normalized global-track $\chi^2 \leq 3$ χ^2 of the matching between the tracker and Standalone muon position ≤ 12 Tracker kink finder ≤ 20
		or
	Segment compatibility ≥ 0.451	

Table 3.2: List of selections for the muon identification.

field, this bremsstrahlung spreads over a large volume in the azimuthal direction. Consequently, an electron loses on average of 33% of its energy before reaching the ECAL at lower pseudorapidity regions while the energy loss can be up to 86% when it passes through the budget material.

The subdetector geometries in the barrel and in the endcaps are different. Therefore dedicated algorithms are used for the clustering of the electron energy in these regions: the hybrid and multi-5x5 algorithms are used for the barrel and the endcaps respectively. The energy is measured in terms of superclusters, that are clusters of the clusters from bremsstrahlung photons.

The electron track reconstruction starts with seeding. To initiate the building of trajectories in the inner tracker, combined results of two algorithms is used: the ECAL-based seeding and the tracker-based seeding. In the ECAL-based seeding, the supercluster energy and position are employed to extrapolate the electron trajectory to the collision vertex. In the tracker-based seeding, pairs or triplets of hits with the vertices obtained from pixel tracks are combined to form the tracker seeds.

A GSF algorithm is used for electron-track that provides GSF electron candidates. The GSF algorithm is designed to follow the track curvature accounting for the bremsstrahlung loss up to the ECAL surface. GSF uses the hit collection obtained with a KF algorithm. It approximates the Bethe-Heitler distribution with a sum of Gaussian distributions.

The charge of the electron is evaluated from the sign of the GSF track curvature. The charge misidentification rate is 1.5% for reconstructed electrons from Z boson decays. The momentum of the electron is calculated from a weighted combination of the measurements from track parameters and from supercluster parameters. The first one is dominant for low energy candidates and the latter is dominant for high energy candidates.

On the reconstructed electrons, the criteria which are listed in Tab. 3.3 are applied to identify and categorize electrons as either loose or tight. The requirements in Tab.3.4 are slightly changing between the barrel and the endcaps, due to the different granularity of ECAL. Those criteria 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. The loose electrons are tuned to an average of 95% efficiency while for tight electrons it is 70%. Due to the identification inefficiency in the gap region between the ECAL barrel and endcaps, the corresponding pseudorapidity region of $1.44 < |\eta| < 1.56$ is excluded for reconstructed electrons.

Electron Selection	
Loose	Tight
loose Id. in Tab.3.4	tight Id. in Tab.3.4
$p_T \geq 10 \text{ GeV}$	$p_T \geq 25 \text{ GeV}$
$ \eta \leq 2.4$	$ \eta \leq 2.4$
	$\text{sip}_{3D} \leq 4.0$

Table 3.3: List of selections for the tight and loose electron identification.

Selection	Definition	Barrel loose/tight	Endcap loose/tight
$\sigma_{i\eta i\eta}$	ECAL crystal-based shower covariance in the direction of η	0.0114/0.0101	0.0352/0.0279
$\Delta\eta_{in}$	Difference between the supercluster position in the ECAL and the track direction at the innermost tracker position in η	0.0152/0.00926	0.0113/0.00724
$\Delta\phi_{in}$	Difference between the supercluster position in the ECAL and the track direction at the innermost tracker position in ϕ	0.216/0.0336	0.237/0.0918
H/E	Ratio of energy measured in the HCAL over the energy measured in the ECAL	0.181/0.0597	0.116/0.0615
$ 1/E - 1/p $	Absolute difference between the inverse electron energy measured in the ECAL and the inverse momentum measured in the tracker	0.207/0.012	0.174/0.00999
Δ_{xy}	Track-vertex distance in the transverse plane	0.0564/0.0111	0.222/0.0351
Δ_z	Track-vertex distance along the beam axis	0.472/0.0466	0.921/0.417
N_{hits}^{miss}	Number of missing hits in the electron inner layer track	≤ 2	$\leq 3 / \leq 1$

Table 3.4: List of selection criteria for the CMS electron identification.

3.4.3 Isolation

In an event, naturally, electrons and muons are also produced in b- and c-quark decays. These secondary leptons are an important background for searches with leptons that originate from W boson decays. Fortunately, these secondary leptons can be distinguished

with the help of the surrounding hadronic activity. The absence of such an energy flow is qualified as isolation. The absolute isolation is quantified as:

$$\begin{aligned}
I = & \sum_{\Delta R < R} p_T(\text{ch. from PV}) \\
& + \max\left[0, \sum_{\Delta R < R} p_T(\text{photons})\right. \\
& + \sum_{\Delta R < R} p_T(\text{neutral hadrons}) \\
& \left. - \frac{1}{2} \sum_{\Delta R < R} p_T(\text{ch. from PU})\right],
\end{aligned} \tag{3.4}$$

where ΔR defines the angular distance to the lepton trajectory at the interaction vertex. The last term is a correction to the contributions from photons and neutral hadrons for the accompanying pileup energy. The average fraction of charged and neutral pileup contributions has been determined empirically to be one-half. Then, a relative isolation for a given radius R , I_{rel} , can be defined as the ratio of absolute isolation and the lepton p_T . In a TeV scale SUSY scenario or, more generally, in the decay of a highly boosted top quark, the decay products can be highly collimated. In such events, the probability of an overlap of the lepton and a jet increases.

To mitigate this effect, a p_T dependent size of the isolation cone radius was proposed in [120, 121]. The main source of overlap are particles from the hadronization of the b-jet from the top quark. The cone size of the jet can be estimated as:

$$\Delta R_{b\text{-jet}} \approx \frac{2m_{\text{mother}}}{p_{T\text{mother}}} = \frac{2m_b}{p_T^b} \simeq \frac{10\text{ GeV}}{p_T^{\text{lep}}}. \tag{3.5}$$

Using this approximation, the following isolation cone size is defined:

$$R_{\text{iso}} = \begin{cases} 0.2, & p_T^{\text{lep}} \leq 50\text{ GeV} \\ \frac{10\text{ GeV}}{p_T^{\text{lep}}}, & p_T^{\text{lep}} \in (50, 200)\text{ GeV} \\ 0.05, & p_T^{\text{lep}} \geq 200\text{ GeV} \end{cases} \tag{3.6}$$

This varying cone size criteria is large enough for identifying secondary leptons from b-quark decays and sufficiently small for avoiding possible overlap of the lepton and jets. This new isolation is also known as mini-isolation, I_{mini} . For electrons, tight(loose) lepton candidates are required to have I_{mini} below 0.1(0.4) while for tight(loose) muon candidates I_{mini} required to be below 0.2(0.4).

The isolation efficiency is defined as:

$$\epsilon_{\text{Iso}} = \frac{N(\text{passing lepton ID and isolation requirements})}{N(\text{passing lepton ID requirements})}. \tag{3.7}$$

A comparison of the efficiency of I_{rel} and I_{mini} criteria is given in Fig. 3.3 for one of the mass scenarios of T5qqqqWW model and $t\bar{t}$ + jets MC samples. Although mini-isolation

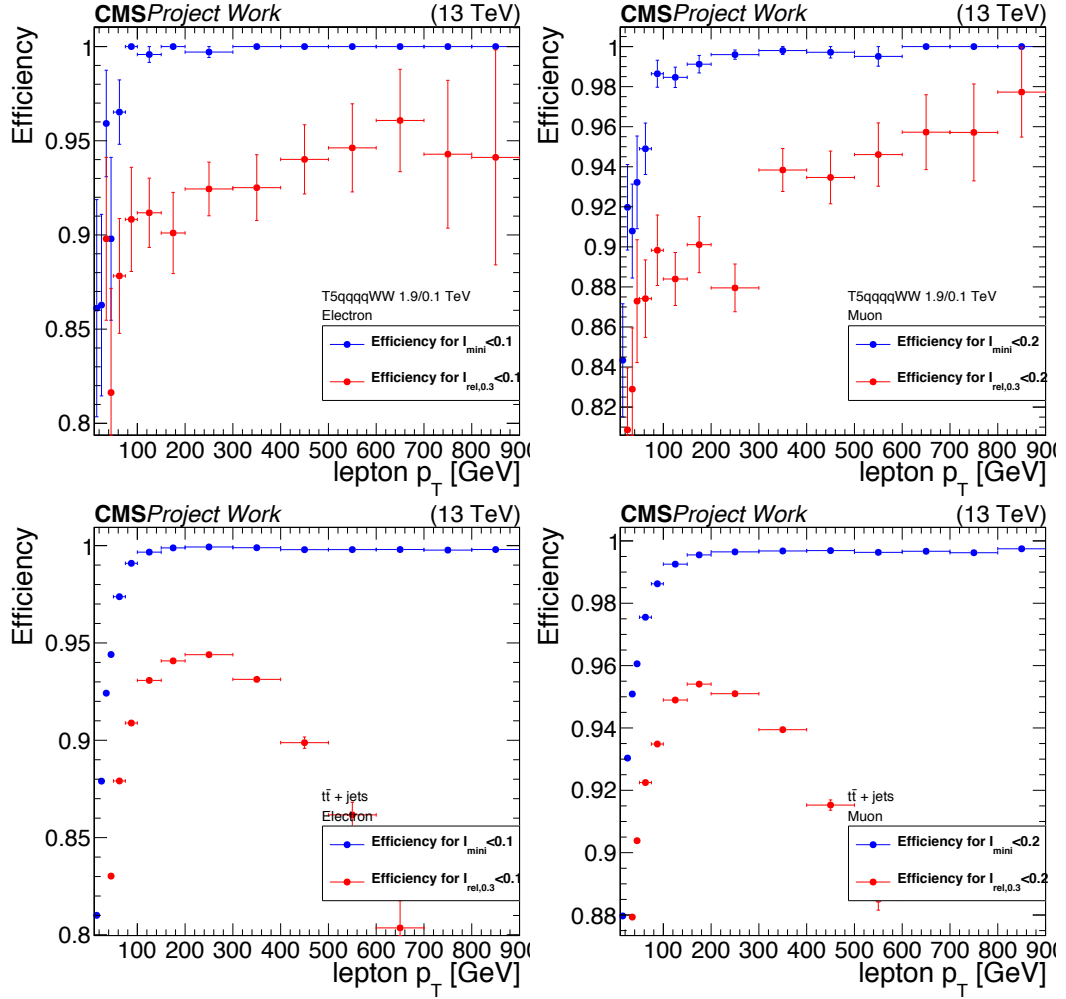


Figure 3.3: Isolation efficiency for simulated events for both of the mini and relative isolation is shown. The upper panel plots show T5qqqqWW (1.9,0.1) model for electrons and muons while the lower panel plots display the efficiency in $t\bar{t}$ + jets events.

is optimized in the context of searches with top quarks, it has a higher efficiency also for 0 b-tag final states as can be seen in Fig. 3.3 (upper). In the lower panel of the figure, isolation efficiency in $t\bar{t}$ + jets events are shown for electrons and muons. Finally, it can be observed that the traditional relative isolation efficiency decreases at large lepton p_T , as the hadronic activity around the lepton increases.

3.5 Missing transverse energy

Weakly interacting massive particles, as predicted by the signal models, leave no trace in the CMS detector. Only the imbalance in transverse momentum can indirectly hint towards their existence. Missing transverse momentum ($\vec{E}_T$) [122] is calculated from the

negative transverse vector sum of all PF candidates,

$$\vec{E}_T = - \sum_i \vec{p}_{T,i}. \quad (3.8)$$

The E_T can be mismeasured for several reasons, such as the non linearity of the calorimeter response for hadronic particles, or the minimum energy thresholds in the calorimeters. The bias in the E_T measurement due to the possible reconstruction problems of PF particles can be reduced by correcting E_T by the vectorial difference in the jet momenta that the jet energy corrections account for. In CMS, particles are produced uniformly in ϕ , thus $\vec{E}_T$ is expected to have a flat distribution in ϕ . However, due to the ϕ -dependence of the detector response, imperfect alignment of different detector subsystems, and a ~ 4 mm shift between the centre of the detector and beamline [123], an asymmetry in ϕ is observed in data and in simulated events.

The observed ϕ -asymmetry can also be viewed as a mean shift in the $\vec{E}_T$ components along the x and y detector axes. The shift shows an increasing trend in shape of a second order polynomial as a function of $\sum p_T$ of the PF candidates. To obtain the corrections, this correlation is used. First, polynomial fits are performed on E_x and E_y distributions as a function of $\sum p_T$ of PF candidates in various η bins. Examples of these fits are shown in Fig. 3.4, where the dependence of $\langle E_x \rangle$ and $\langle E_y \rangle$ on multiplicity of PF candidates can be formulated as:

$$\begin{aligned} \langle E_x \rangle &= c_{x_0} \cdot x + c_{x_s} \cdot x^2, \\ \langle E_y \rangle &= c_{y_0} \cdot x + c_{y_s} \cdot x^2. \end{aligned} \quad (3.9)$$

After the fits are performed, the corrected E_x and E_y can be obtained on an event-by-event basis as:

$$\begin{aligned} E_x^{\text{corr}} &= E_x - \langle E_x \rangle = E_x - (c_{x_0} \cdot x + c_{x_s} \cdot x^2), \\ E_y^{\text{corr}} &= E_y - \langle E_y \rangle = E_y - (c_{y_0} \cdot x + c_{y_s} \cdot x^2). \end{aligned} \quad (3.10)$$

The coefficients c_{x_0} , c_{x_s} , c_{y_0} , and c_{y_s} are determined separately from data and simulated samples. In data, the corrections are obtained in dimuon events with the invariant mass of the dimuon system is inside the Z-boson mass window, $60\text{GeV} < M_{\ell\ell} < 120\text{GeV}$. In simulated samples, the corrections are obtained by using DY events where a Z boson decays to two opposite charged leptons. In such events the expected genuine- E_T contribution is very low, thus they provide a very good environment to measure E_T mismodelings. The validation of the procedure is performed by using $t\bar{t}$ + jets and W+jets simulated samples. Figure 3.5 shows E_T before and after the correction for simulated events. It is evident that there is no significant change to the spectrum while $E_{T\phi}$ is properly corrected.

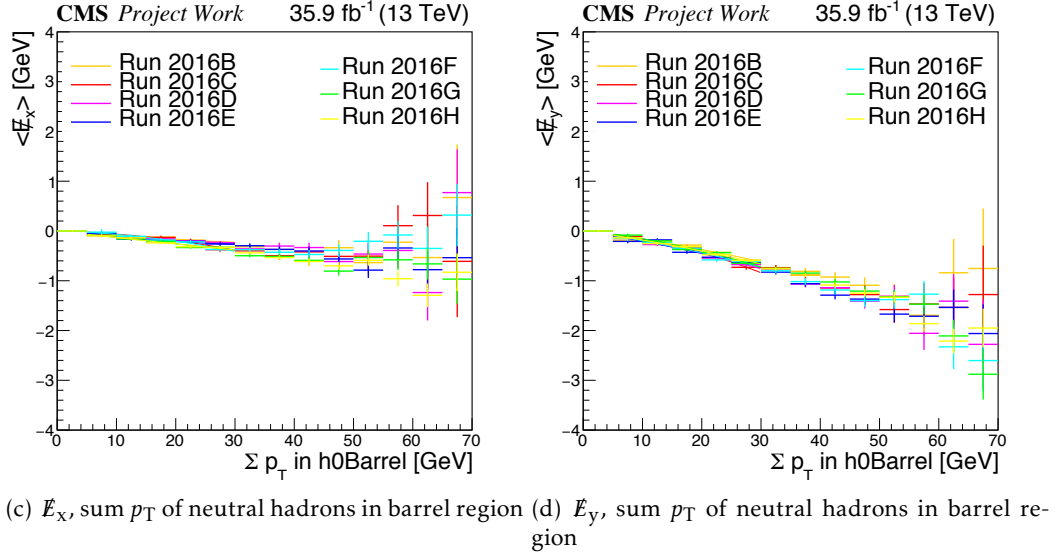
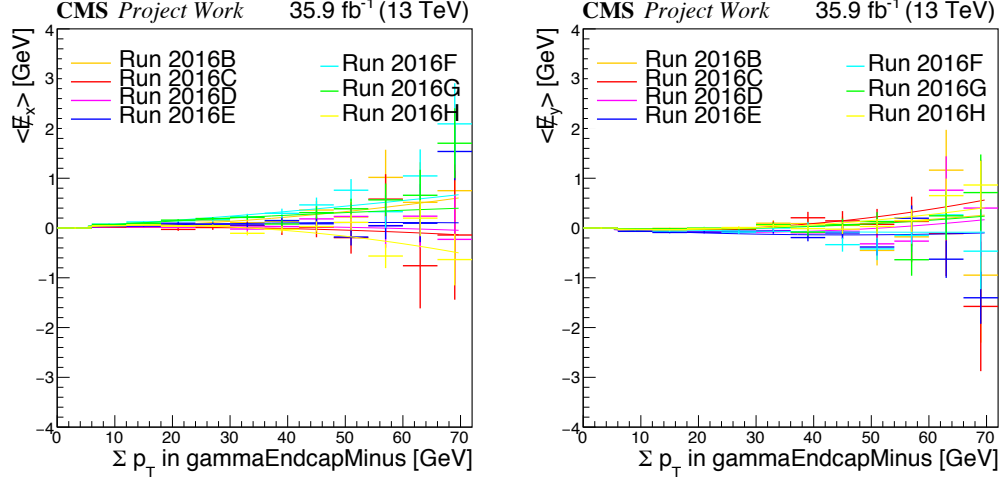


Figure 3.4: PF candidate multiplicity fits for $\langle E_x \rangle$ and $\langle E_y \rangle$ in different η regions. Different colored lines represent different data taking run ranges. The Y-axis shows the average value of the x- and y-components of missing energy while X is the sum p_T of the corresponding PF candidate.

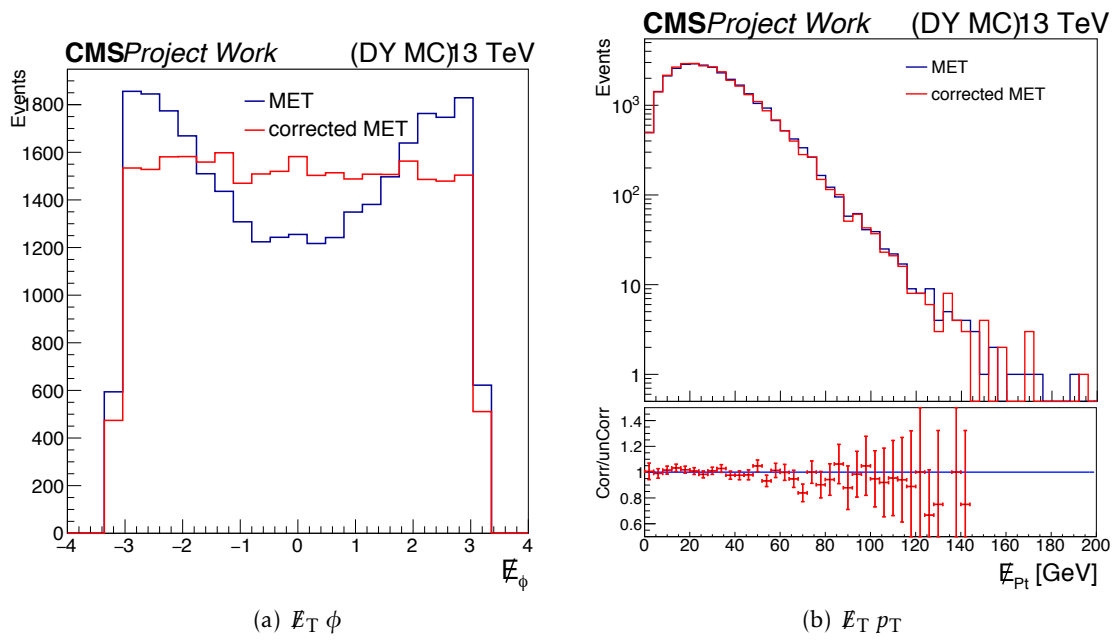


Figure 3.5: $E_T \phi$ (left) and p_T (right) distributions before(blue) and after(red) corrections in simulated Drell-Yan events.

Chapter 4

Event Selection: Baseline and search regions

4.1 SUSY signature

As mentioned in Sec. 1.2.3, a search for SUSY is performed in the context of the simplified model T5qqqqWW (see Fig. 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.

The event selection is designed to have the largest possible signal efficiency hence it is

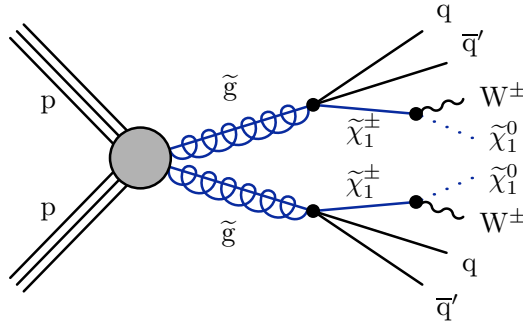


Figure 4.1: Diagram of the simplified model T5qqqqWW.

related to expected final state of the SUSY signal signature.

In the event selection, one of the W bosons 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. The visible final state includes then at least six jets and one lepton (either electron or muon). In order not to loose events where two jets with high energy are produced aligned and reconstructed as one jet, or a jet is out of acceptance, the jet multiplicity is required to be at least five. Moreover, as a distinct feature of this model, it is required that none of the jets are tagged as originating

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

from b quark. Furthermore, in such events, a high E_T is expected due to the presence of the neutralino and the neutrino which leave no trace in the detector.

Merely selecting events with one lepton, five jets, and high E_T is not enough to reveal the existence of SUSY-like events which are rare among the many SM like events with similar final states. The W+jets and $t\bar{t}$ +jets events are the leading SM background processes which can mimic the T5qqqWW signature. Therefore, some advanced kinematic variables need to be defined to eliminate these SM background processes.

4.1.1 Key variables

The list of variables consists of fundamental properties of the reconstructed particles, such as the transverse momentum p_T of these objects, and functions of them. The reconstruction and identification of these particles are introduced in Ch.3.

- n_{jet} is the number of all jets above a p_T threshold in an event.
- $n_{\text{b-tag}}$ is the number of all jets tagged as coming from a b quark above a p_T threshold, as with other jets, in the event.
- H_T , $\sum p_T(\text{jets})$, is the scalar sum of the transverse momenta of all jets above a p_T threshold in the event. For signal events, 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 reflects the "hadronic mass scale" of the event.
- L_T is the scalar sum of the charged lepton p_T and the missing transverse momentum E_T . Given no additional source of E_T , 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 starts by expanding the squared transverse momentum and the transverse mass of the W boson:

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

For events with a single 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.

- $\Delta\phi(W, \ell)$ is the azimuthal angle between the W boson candidate, formed from the E_T and lepton $\vec{p}_T$, and the charged lepton. In W+jets and $t\bar{t}$ +jets background events, the lepton comes from a leptonic decay of the W boson, $W \rightarrow \ell \nu$, which strongly constraints allowed values for $\Delta\phi(W, \ell)$. It is small when the daughter particle is aligned with its mother as in the case of single leptonic W+jets and $t\bar{t}$ +jets events. On the other hand, in SUSY decays, the missing transverse energy comes from two neutralinos and the neutrino, which randomizes $\Delta\phi(W, \ell)$, thus resulting in an almost flat

distribution, making $\Delta\phi(W, \ell)$ a highly discriminating observable. The high values of $\Delta\phi(W, \ell)$ are used as a signal region² while the low values of $\Delta\phi(W, \ell)$ are the control region³.

- M_{T2} [124] 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 an isolated track and the selected lepton respectively. This variable is used only in the baseline selection to reduce contributions with a second, unidentified electron or muon, or the hadronic decay of a tau lepton. It reduces the dileptonic $t\bar{t}$ + 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 Sec. 1.2.3. This model in principle includes three free mass parameters⁴: 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 dimensional mass space the chargino mass is fixed to the arithmetic mean of gluino and neutralino mass. For each mass point, a separate MC sample is produced. The complete mass plane is shown in Fig.4.3 where the z-axis represents the total number of events generated. The MADGRAPH5 event generator is used for signal events modelling. The scan of parameter space includes 657 mass points which are simulated using FastSim (Sec. 2.3.2) in order to reduce CPU consumption. The samples are normalized according to the gluino production cross section [125]. 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 $m_{\tilde{g}} = 1.9$ TeV and $m_{\tilde{\chi}_1^0} = 0.1$ TeV. These signal events have high H_T and L_T .
- **T5qqqqWW(1.5,1.0)** represents a point close to the compressed region with $m_{\tilde{g}} = 1.5$ TeV and $m_{\tilde{\chi}_1^0} = 1.0$ TeV. The compressed region is where the $m_{\tilde{g}} - m_{\tilde{\chi}_1^0} \leq 2 m_W$. These signal events have lower H_T and L_T with respect to high mass gap ones.

In Fig. 4.2, H_T (left), L_T (middle), n_{jet} (right) distributions for three different signal mass points are shown. When the mass gap between gluino and neutralino is low, it results in softer hadronic and leptonic activity and vice versa.

²where signal to background event counts are sufficiently large.

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

⁴there are many more parameters that are ignored

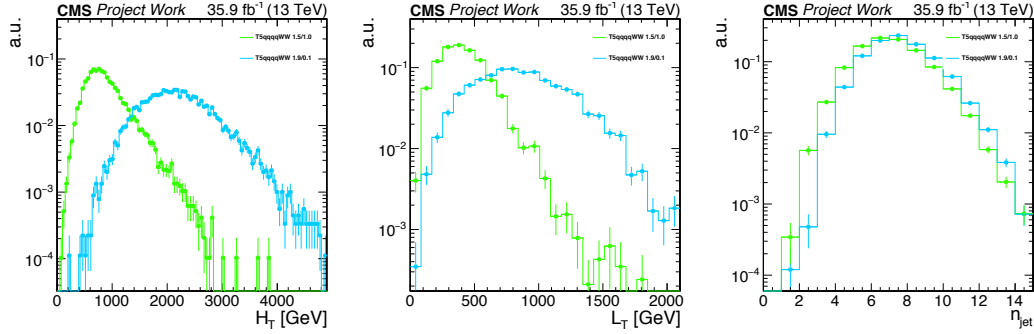


Figure 4.2: H_T (left), L_T (middle), n_{jet} (right) distributions for three different signal mass point. Distributions are normalized to one. The green line represents the low mass gap signal point ($m_{\tilde{g}} = 1.5$ TeV and $m_{\tilde{\chi}_1^0} = 1.0$ TeV), the light blue line shows the high mass gap ($m_{\tilde{g}} = 1.9$ TeV and $m_{\tilde{\chi}_1^0} = 0.1$ TeV) benchmark point.

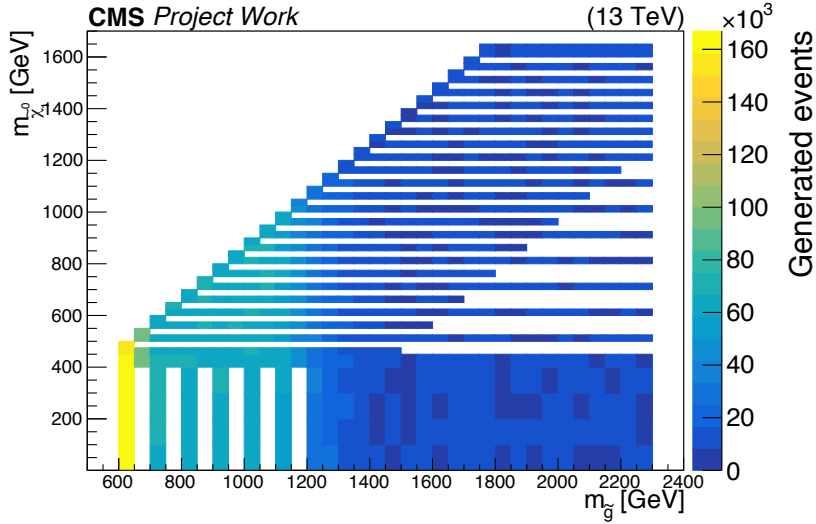


Figure 4.3: Simulated signal mass points for the simplified T5qqqqWW model.

4.2 Background processes

As mentioned earlier in this section, the most important background processes are W +jets and $t\bar{t}$ +jets events. W +jets events becomes the main background, when zero b -tagged jets are required in the event selection. All of the background processes considered in this analysis are listed below.

- **$t\bar{t}$ + jets** is generated with MADGRAPH5_AMC@NLOv2.2.2. To enhance the sample size, the combination of H_T binned samples, dedicated semi- and di-leptonic samples is used.
- **W +jets** is generated with MADGRAPH5_AMC@NLOv2.2.2. H_T binned samples, where W boson decaying leptonically, are used.

- **QCD multijet** events are produced with `MADGRAPH5_AMC@NLOv2.2.2` in bins of H_T .
- $t/\bar{t}$ samples are produced with `POWHEGv2.0` except for the single top quark 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 $m_{\ell\ell} > 50\text{GeV}$ sample is generated in H_T bins.
- **Di-boson** samples are produced with `AMC@NLO` except for WW samples that 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

Studies of data/MC ratios reveal that event-by-event scale factors are needed to remove or mitigate shortcomings of the simulation.

Lepton identification and reconstruction efficiency

MC samples have to be corrected according to the reconstruction and identification efficiencies of leptons. In this work, scale factors are applied as a function of p_T and η of the selected lepton [126].

Efficiency and event weights of b tagging

The b tagging efficiency SFs [126] are measured, ($SF = \epsilon_{\text{DATA}}/\epsilon_{\text{MC}}$), for b and light flavor jets as a function of jet p_T and η . To predict the correct event yield in data, SFs are applied to the selected MC events. For the weighting of simulated events, a probabilistic approach is used. The probability of a jet to be b-tagged, which is denoted by p_i , and is measured, as a function of jet p_T , η and flavor of the original parton in simulation. The index i labels the jets in the event. The p_i is equal to b tagging efficiency, ϵ_i (Eq. 4.3), when the jet comes from a generated b quark.

$$\epsilon_f(p_T, \eta) = \frac{N_f^{\text{b-tagged}}(p_T, \eta)}{N_f^{\text{Total}}(p_T, \eta)}, \quad (4.3)$$

where $N_f^{\text{b-tagged}}$ and N_f^{Total} are the number of b-tagged and all jets, respectively, for flavor f in (p_T, η) . The scale factors from simulation to data, SF_i , are applied on the b-tagging probability, $P_i = SF_i \cdot p_i$. The probability of an event to be reconstructed with exactly n b-tagged jets, becomes:

$$w(n \text{ b-tagged jets}) = \dots \sum_i^n \sum_j^n \left[\dots \overbrace{P_i P_j}^n \prod_k (1 - P_k) \right], \text{ with } j > i \text{ and } k \neq i, j. \quad (4.4)$$

This method allows reusing the events, because each event will contribute to any b-tag multiplicity, irrespective of the parton multiplicities. The advantage of this method is that

it allows to use the full statistical power of the MC events.

ISR reweighting:

Since the 8 TeV Run of LHC, it is known that the p_T spectrum and the number of additional jets 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). Each event is corrected according to number of ISR jets. The weights can be seen in Tab. 4.1 where the factor D is selected in order to preserve the normalization of the inclusive sample.

Table 4.1: Weights based on the number of ISR jets as given in Ref. [82]

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 Sec. 3.2, the number of pileup interactions in simulated samples is sampled from a prior distribution, approximately reflecting 2016 data taking conditions. Thus, the MC pileup distribution may vary from the actual pileup distribution in Data. The simulated distribution is rescaled with the Data/MC ratio in Fig. 4.4, in order to achieve agreement in the true number of interactions.

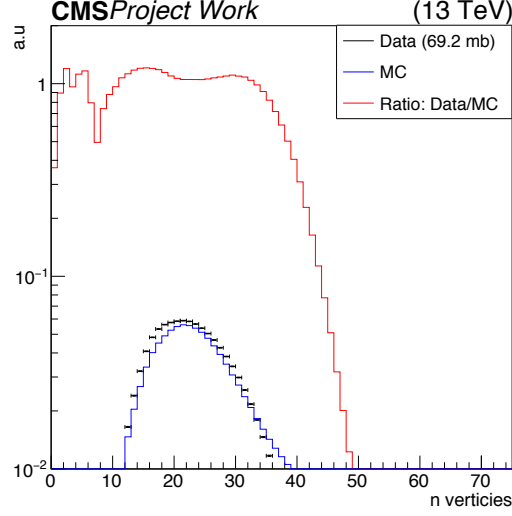


Figure 4.4: The normalized distributions of mean number of interactions per bunch crossing during the analyzed data Run2016 (black) and in the MC samples (blue). The red distribution represents the Data to MC ratio. For the data distribution the latest luminosity calibrations [83] and an inelastic pp cross-section of 69.2 mb are used.

4.3 Baseline selection

Object selections have been discussed in Sec. 3 and are now used to define the event selection.

Selections on leptons

Exactly one lepton (electron or muon) is selected. The selected lepton is required to have a minimum p_T of 25 GeV while at the same time it satisfies the tight lepton criteria introduced in Sec.3.4. Additionally, events are vetoed if there is an extra lepton with $p_T > 10$ that satisfies the loose lepton selection criteria.

Selections on jets

Jets are required to have $p_T > 30$ GeV and $|\eta| < 2.4$. Furthermore, 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 model is expected to have at least five jets. The mass difference of the gluino and the intermediate chargino affects the typical p_T of the jets. Therefore, it is moreover required that two highest p_T jets satisfy $p_T > 80$ GeV. Additionally, in the baseline selection, the number of jets tagged as coming from b quarks is required to be zero in order to suppress the events containing top quarks.

Energy scales thresholds

The signal model favors events with high H_T and L_T as can be seen in Fig. 4.2. The H_T is chosen to be at least 500 GeV while the L_T threshold is 250 GeV.

Isolated track veto

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, events that contain at least one isolated high- p_T charged track are rejected. These tracks originate from $\tau \rightarrow \nu_\tau + \text{hadron}$

decays (hadronic tracks) or electron or muon tracks (leptonic tracks) of poor quality. For this veto, the M_{T2} variable, which is introduced in Sec. 4.1.1, is used with an isolated track t , a lepton ℓ and the missing transverse momentum $\vec{p}_T^{miss}$. In the calculation of M_{T2} , it is assumed that the missing energy comes from the two neutrinos from the dileptonic $t\bar{t}$ decay. The minimization runs over all possible splitting of $\vec{p}_T^{miss}$. Figure 4.5 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 Fig. 4.5 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 seen from Fig. 4.5 (c) that only about 15% of the signal events have an opposite charged isolated track at all compared to 40% of the dileptonic $t\bar{t}$ events.

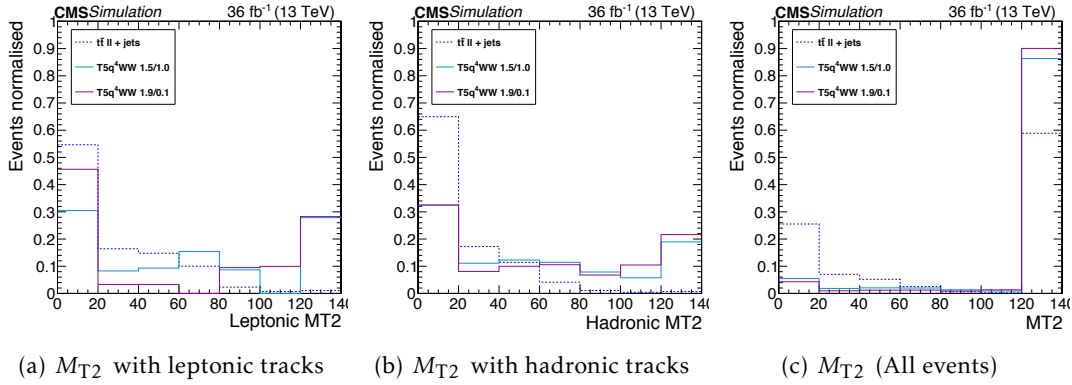


Figure 4.5: Distributions of M_{T2} for events with electron or muon veto tracks (a) and hadronic veto tracks (b), for dileptonic $t\bar{t}$ (blue dotted) and $T5qqqqWW$ (purple and azure) signal samples in Sec. 4.1.2. The highest bin is always an overflow bin. The majority $t\bar{t}$ events have $M_{T2} < 80$ 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 Fig. 4.6 for the different background processes (stacked) and for two signal benchmark points, starting with a selection $H_T > 350$ GeV and $L_T > 150$ GeV. After this selection, the cuts are applied consecutively and the n_{jet} requirement has the biggest impact. It can be also noted that, naturally, the QCD background almost vanishes after the single muon selection. The estimation of important backgrounds will be explained in the Ch. 5.

Table 4.2: List of baseline criteria and object requirements.

Selection	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	$I_{rel} < 0.3$, $\Delta R(\ell, t) < 0.1$, $t_{charge} = -l_{charge}$
$n_{jets} \geq 5$	Jets with $p_T \geq 30$ GeV and $ \eta < 2.4$ and cleaned from close leptons
$p_{T,jets(1,2)} \geq 80$ GeV	
$H_T \geq 500$ GeV	$\sum_{jets} p_T$
$L_T \geq 250$ GeV	$\cancel{E}_T + p_T^{lep}$
$n_{b-tag} = 0$	b-tagged jets with CSVv2 Medium working point (0.8484)

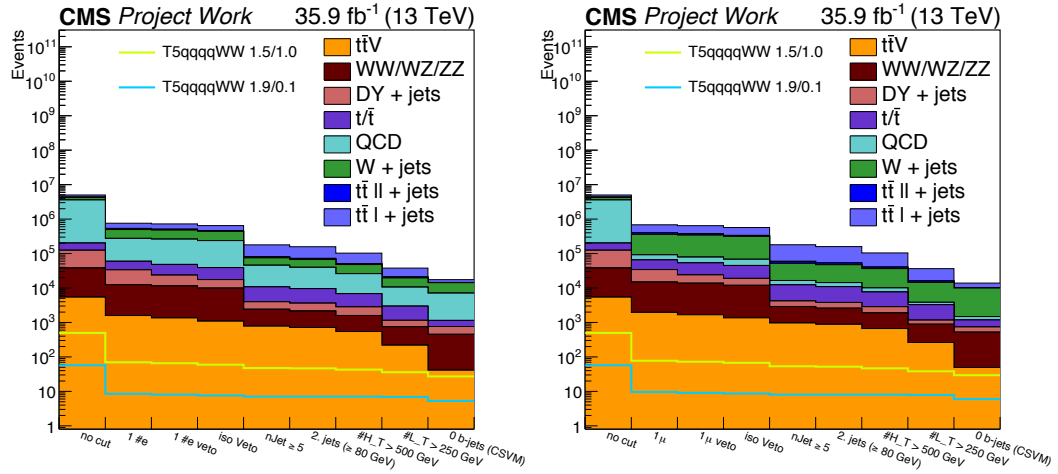


Figure 4.6: The selection criteria flow for the electron (left) and muon (right) channel. Samples are scaled to the luminosity with their cross section and scale factors discussed in Sec. 4.2.1. The color filled histograms represent the background samples while the solid lines represent the two signal models that are shown in Sec. 4.1.2.

4.4 Data samples

In this work, the data recorded by the CMS detector during the 2016 LHC run at 13 TeV is used. As shown in Fig. 2.2 (right), the total integrated luminosity collected by CMS is 37.76 fb^{-1} . The data validated according to the CMS Data Quality Monitoring (CMS-DQM) certification requirements is $L=35.9 \text{ fb}^{-1}$.

4.4.1 Trigger selection

Most events in this work are selected by a combination of triggers requiring an isolated single lepton, electron or muon, with p_T of 15 GeV and H_T of 350 GeV. The complete list of trigger paths can be found in Tab. 4.3. The naming of the trigger paths relates to the conditions introduced on the corresponding object. For example, HLT_Ele105_CaloIdVT_GsfTrkIdT involves the conditions that the electron has p_T larger than 105 GeV and it satisfies the very tight calorimeter identification and the tight GSF identification.

Events recorded with these trigger paths are allocated in three primary datasets (PDs):

Table 4.3: List of HLT paths. Notation is used in terms of CMS internal shortcuts.

Single Electron dataset
HLT_Ele105_CaloIdVT_GsfTrkIdT HLT_Ele115_CaloIdVT_GsfTrkIdT HLT_Ele50_CaloIdVT_GsfTrkIdT_PFJet165 HLT_Ele27_WPTight_Gsf HLT_Ele15_IsoVVVL_PFHT350 HLT_Ele15_IsoVVVL_PFHT400
Single Muon dataset
HLT_Mu50 HLT_IsoMu24 HLT_IsoTkMu24 HLT_Mu15_IsoVVVL_PFHT350 HLT_Mu15_IsoVVVL_PFHT400
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

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 Eq. 4.5. The trigger efficiencies are measured as a function of L_T , H_T and lepton p_T and can be seen in Figs. 4.7, 4.8, 4.9 respectively. The efficiency in the baseline region ($L_T > 250$ GeV; $H_T > 500$ GeV; lepton $p_T > 25$ GeV) is flat and its value is close to 100%(98%) for the muons

(electrons).

$$\epsilon = \frac{N(\text{all events passing probed trigger(s) + preselection + reference trigger})}{N(\text{all events passing preselection + reference trigger})} \quad (4.5)$$

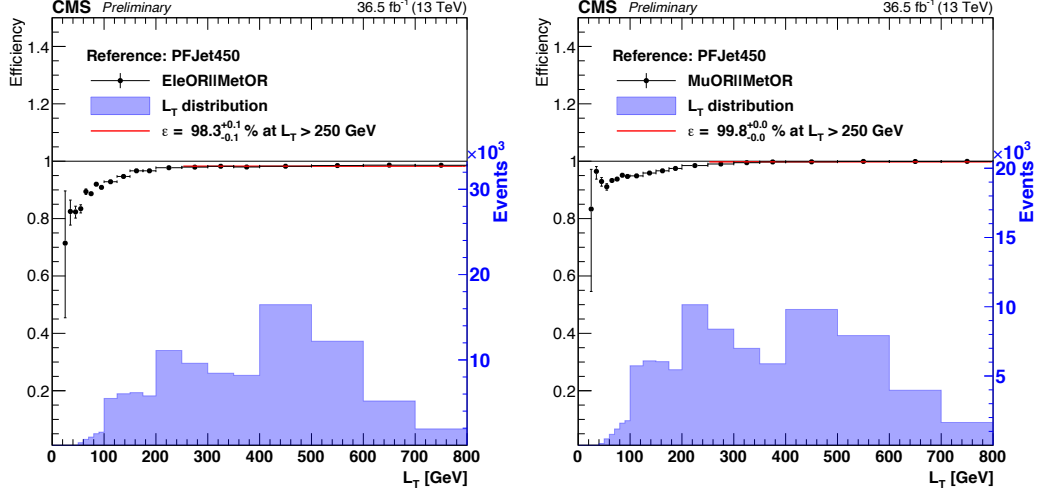


Figure 4.7: Measurement of the trigger efficiency as a function of L_T for the electron trigger selection (see Tab. 4.3) on the right [127].

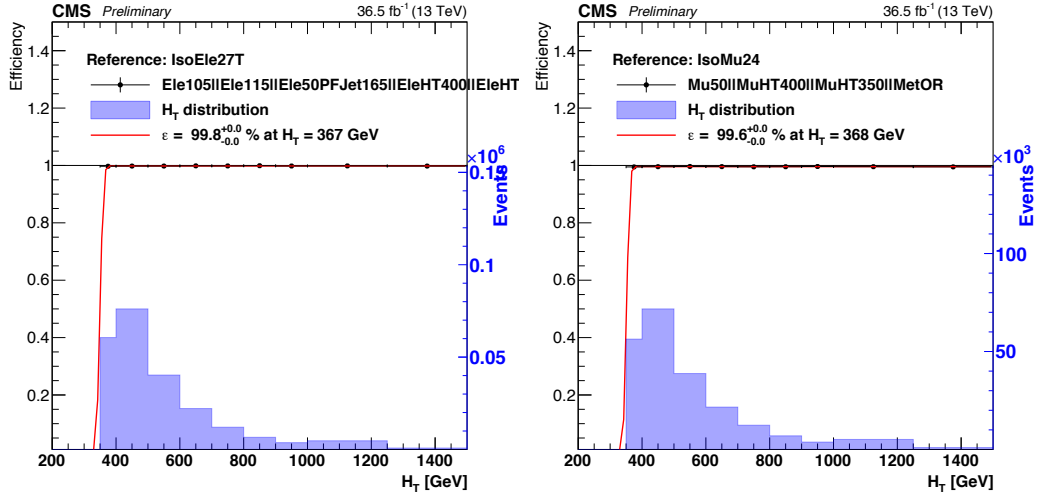


Figure 4.8: Measurement of the trigger efficiency as a function of H_T for the electron trigger selection on the left and the muon trigger selection on the right [127].

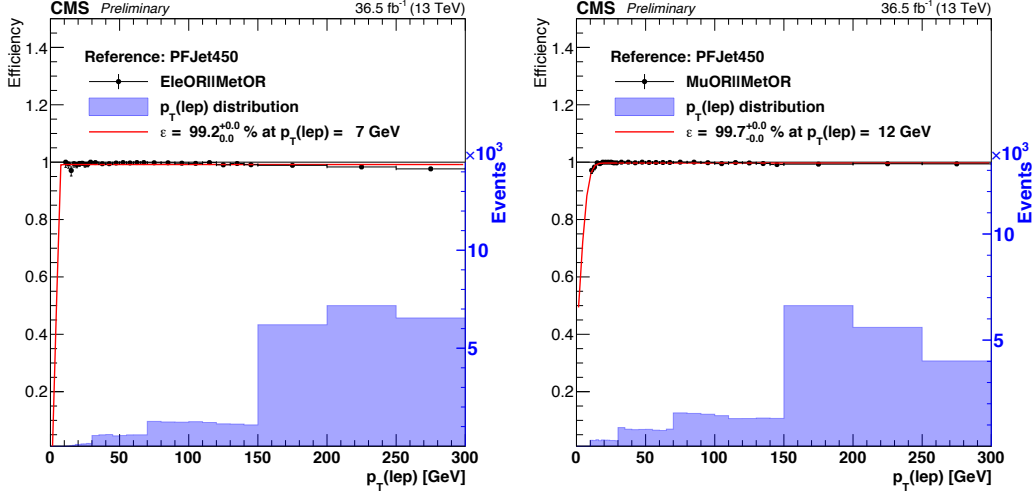


Figure 4.9: Measurement of the trigger efficiency as a function of lepton p_T for the electron trigger selection is shown on the left and the muon trigger selection is shown on the right. The lepton p_T trigger efficiency is measured after an $L_T > 250$ GeV requirement which is the analysis baseline selection [127].

4.5 Event cleaning filters

Spurious detector signals can produce artificially large $\cancel{E}_T$. Sources and dedicated algorithms that are used to identify and remove the events with fake $\cancel{E}_T$ are briefly discussed in the following.

Primary vertex filter

A good primary vertex (see Sec. 3.2) is required.

Beam halo filter

The collisions of the beam with residual gas inside LHC cause showers of secondary particles. Charged particles (muons) 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. To detect events with beam halo it uses the timing information and hit topology in the CSC, ECAL and HCAL subsystems.

HB-HE noise filter

This noise is originated from the Hybrid PhotoDiodes and Readout Boxes of the HCAL subdetector. The timing, pulse shape as well as topological information of the deposits are used to detect the noise.

ECAL dead cell trigger primitive filter

The existence of noisy crystals in ECAL can lead to fake $\cancel{E}_T$. The events in which the noisy cells deposit high energy are filtered.

Bad PF muon filter

This filter fires when there is a PF muon (see Sec. 3.4.1) with too low quality and has large p_T . The quality of the muon is determined using segment compatibility and other detector

related features. If there is a muon with these properties and a $p_T > 100$ GeV, the event is filtered. Unlike the other filters explained above this filter is applied to background simulation samples.

Bad charged hadron filter

The events, where there is a standalone or global muon but it fails the PF muon requirements but it still contributes to $\cancel{E}_T$ calculation as a charged hadron candidate, are filtered, because the misidentification probability is high in such cases. As in the previous case, this muon is required to have $p_T > 100$ GeV in order to trigger the filter. Moreover, it is required that the standard muon and the charged hadron traces share the same tracker track and have similar transverse moment. Similarly to the bad PF muon filter, this is applied to background MC samples as well as the data events.

Duplicate muon filters

At the end of 2016, it was observed that there is an artificially high $\cancel{E}_T$ tail of $Z \rightarrow \mu\mu$ data. It was understood that there were fake duplicate muons reconstructed such events. Therefore, two filters were designed to mitigate the problem. One is to remove events with the ratio of PF $\cancel{E}_T$ to calorimeter $\cancel{E}_T$ is more than 5. The second is to remove events containing jets which have $p_T > 200$ GeV, muon energy fraction > 0.5 and $\Delta\phi(\cancel{E}_T, \text{jet}) > \pi - 4$.

Filter on FastSim

Events in FastSim samples are removed if there is a jet satisfying the following conditions: $p_T > 20$ GeV, $|\eta| < 2.5$, unmatched to a generated jet (using $\Delta R(\text{jet}, \text{gen jet}) > 0.3$), and charged hadron fraction < 0.1 . This requirement removes spurious jets related to a rare mismodeling of the calorimetric showers in FastSim.

After the baseline selection, approximately 2% of the events are removed by the filters.

4.6 Control plots

In this section, the distributions of the main kinematic variables are discussed. 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 shows the observed data points requiring the triggers introduced in Sec. 4.4.1. In all the control plots, the events are cleaned by the filters discussed in Sec. 4.5. The signal and background events are scaled by the luminosity factor and additional weights introduced in Sec. 4.2.1. Figure 4.10 (top left) shows the $n_{b\text{-tag}}$ distribution after the baseline selection. The distribution of simulated signal events peaks at zero, as expected. One can see from this distribution that choosing events with zero b-tagged jets suppresses the $t\bar{t} + \text{jets}$ background significantly. The other distributions, showing the main kinematic variables, exhibit reasonable MC-data agreement. Fig. 4.10 shows the number of jets distribution (top right), H_T (bottom left) and L_T (bottom right).

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 , which is due to the gluino-neutralino mass splitting. In the non-compressed T5qqqqWW (1.9/0.1 [TeV])

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 (1.5,1.0 [TeV]) 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.

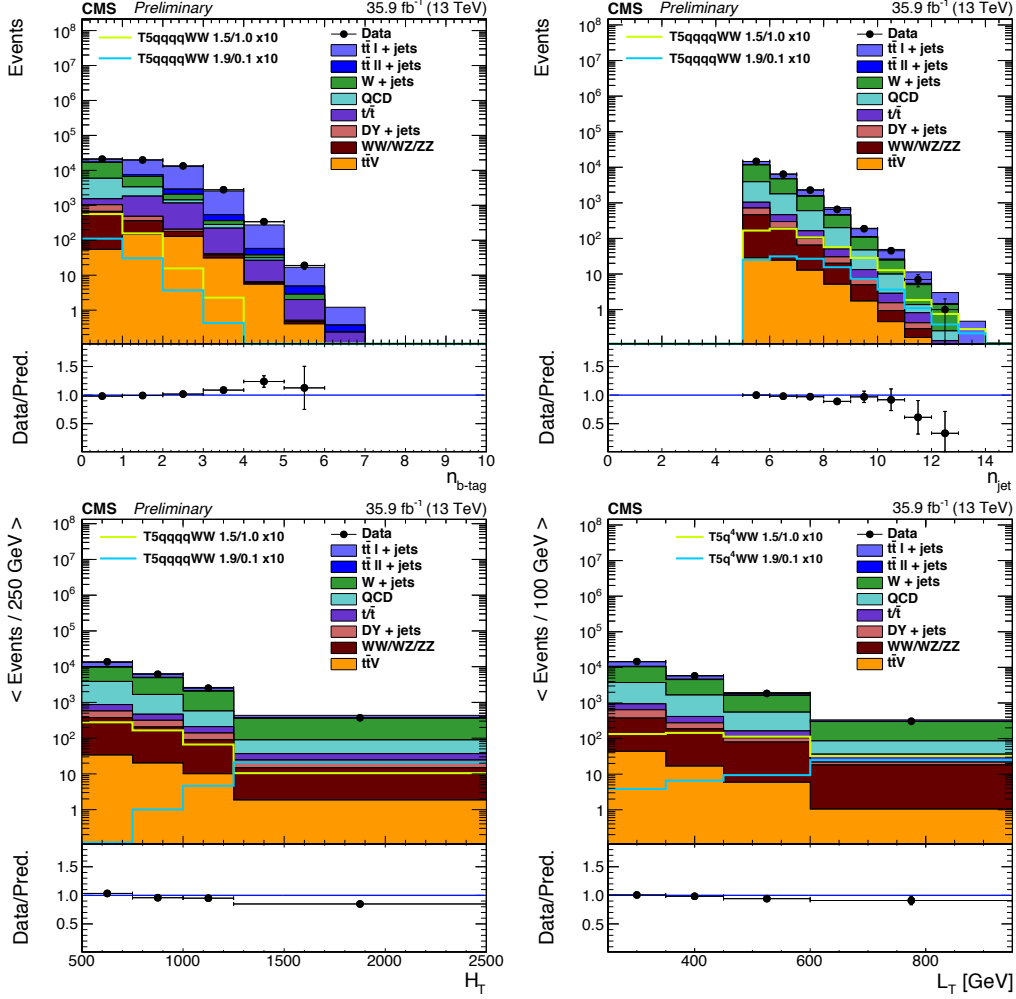


Figure 4.10: Number of b-tagged jets (top left), after the baseline selection, requiring at least five jets, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton with $p_T > 25$ GeV. The other distributions (top right and bottom) shown with an additional b jet veto. Number of jets distribution (top right), H_T (bottom left) and L_T (bottom right) are shown. The benchmark signal models are scaled up by a factor of 10.

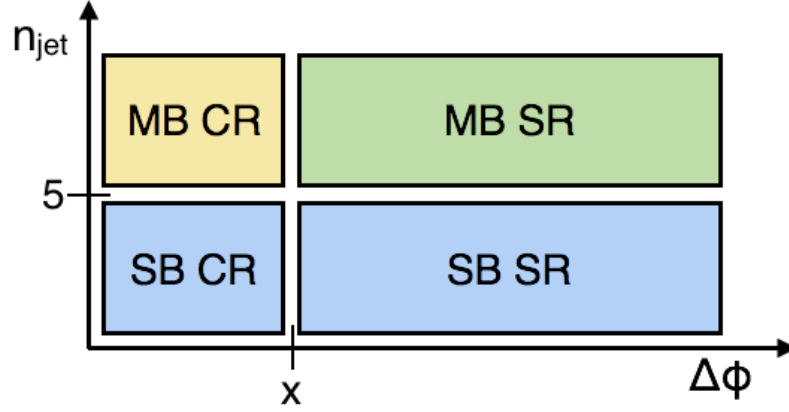


Figure 4.11: The n_{jet} vs. $\Delta\phi$ plane for SR, CR, MB, SB.

4.7 Regions of interest

For a given selection in H_T and L_T , the plane n_{jet} vs. $\Delta\phi$ is divided in 4 kinematic regions. The bins at low and high $\Delta\phi$ are called control region (CR) and signal region (SR), respectively. The low and high n_{jet} bins are called sideband (SB) and mainband (MB), respectively. A sketch of the regions is shown in Fig. 4.11. In this section, SR, CR, MBs are defined. SBs are designed to measure transfer factor from CR to SR for data driven background estimation. SBs are described in Ch. 5 when background estimation is discussed.

4.7.1 Signal and control region

SRs are defined based on the expected deviation from the SM under a SUSY signal hypothesis. Fig. 4.12 shows the main search variable $\Delta\phi$ after the baseline selection. The plot shows that the SM background processes peak at zero while the signal models are almost flat. This shape difference suggests high values of $\Delta\phi$ for SR and the low values as CR, which are used in the background estimation. It can also be seen from the distribution that after the baseline selection, W +jets and $t\bar{t}$ +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.

4.7.2 Mainband regions

Mainband regions are exclusive kinematic regions defined according to n_{jet} , H_T and L_T . These regions are designed to further increase the sensitivity to a range of different gluino and neutralino mass scenarios.

First, the phase space after baseline selection is subdivided into three jet multiplicity bins: $n_{\text{jet}} = 5$, $6 \leq n_{\text{jet}} \leq 7$, $n_{\text{jet}} \geq 8$. As shown in Fig. 4.2, signal model with a lower mass gap has a peak in lower L_T and H_T values than the higher mass gap signal model. Therefore, further binning is introduced in L_T and H_T . To determine the binning, high sensitivity is

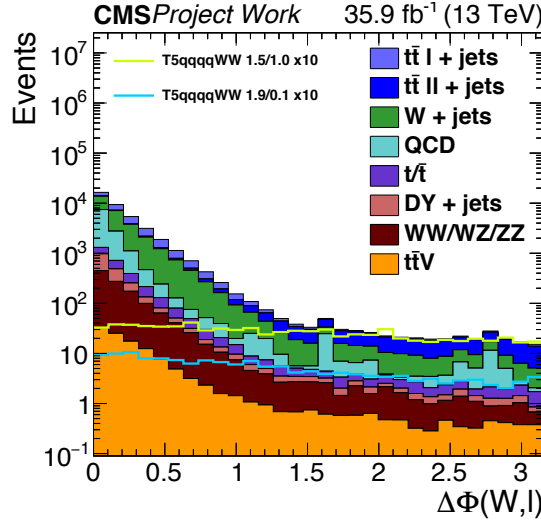


Figure 4.12: The $\Delta\phi$, after the baseline selection, requiring at least five jets and non of b-tagged, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton (electron or muon) with $p_T > 25$ GeV. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked.

considered. A profiled likelihood analysis is performed to measure the sensitivity. Additionally, to have a robust background estimation, the expected yield of background MC for each bin is required to be at least 0.5.

Because $\Delta\phi$ and L_T are both constructed from lepton p_T and E_T , some degree of correlation can be assumed. As shown in Fig. 4.14, for increasing L_T the SM background events are located at 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. Again, the $\Delta\phi$ cut which provides the highest sensitivity is used. Figure 4.13 shows the figure of merit as a function of $\Delta\phi$ for four different L_T bins. The figure of merit is defined based on an assumed systematic uncertainty of 20%:

$$FOM = \frac{N(\text{signal})}{\sqrt{N(\text{background}) + (0.2 \cdot N(\text{background}))^2}}. \quad (4.6)$$

The $\Delta\phi$ cut that defines the signal regions is defined separately for each mainband region. The cut values vary according to L_T of the corresponding bin. For the highest L_T bin (black line in Fig. 4.13), the $\Delta\phi$ cut is found to be 0.5. For the next highest bin (blue), it is 0.75. For the two lowest L_T bins (red, green), $\Delta\phi$ is chosen to be 1. The resulting binning, which is composed of 28 exclusive kinematic regions, is given in Tab. 4.4. The table also shows the simulated background and two benchmark signal yields for signal regions ($\Delta\phi > X$). In Fig. 4.16 and Fig. 4.17, individual simulated SM background yields are presented for single muon and single electron events. The upper plots show the high $\Delta\phi$ regions while the lower plots show the low $\Delta\phi$ regions. Figures manifest that the dileptonic decay of

$t\bar{t}$ + jets events are favoured in SR (high $\Delta\phi$). Furthermore, it can also be seen that the simulated QCD multijet events populate the CR (low $\Delta\phi$) in the electron channel. As mentioned before, the W+jets and $t\bar{t}$ + jets are dominating the SM background. The number of the events in the MB SR is estimated with data sidebands as described in Ch. 5.

Table 4.4: Simulation table of mainband regions, 35.9 fb⁻¹

n_{jet}	L_T [GeV]	$\Delta\phi$ [rad]	H_T [GeV]	Bin name	T5qqqqWW $m_{gl}/m_{\tilde{\chi}^0}$ [TeV]			tot. background	
					(1.5/1.0)	(1.9/0.1)		Simulation	
5	[250, 350]	1.0	[500, 750] ≥ 750	G01	1.82 $\pm$ 0.29	0.0 $\pm$ 0.0		109.14 $\pm$ 2.99	
				G02	0.21 $\pm$ 0.09	0.01 $\pm$ 0.01		81.27 $\pm$ 2.38	
	[350, 450]	1.0	[500, 750] ≥ 750	G03	2.25 $\pm$ 0.32	0.0 $\pm$ 0.0		25.82 $\pm$ 1.55	
				G04	0.29 $\pm$ 0.11	0.04 $\pm$ 0.01		21.62 $\pm$ 1.2	
	[450, 650]	0.75	[500, 750] [750, 1250] ≥ 1250	G05	3.02 $\pm$ 0.37	0.0 $\pm$ 0.0		18.54 $\pm$ 1.5	
				G06	1.4 $\pm$ 0.25	0.04 $\pm$ 0.02		16.23 $\pm$ 1.14	
				G07	0.08 $\pm$ 0.06	0.25 $\pm$ 0.04		4.96 $\pm$ 0.69	
	≥ 650	0.5	[500, 750] [750, 1250] ≥ 1250	G08	0.74 $\pm$ 0.18	0.01 $\pm$ 0.01		3.81 $\pm$ 0.83	
				G09	0.49 $\pm$ 0.15	0.12 $\pm$ 0.03		7.92 $\pm$ 0.79	
				G10	0.14 $\pm$ 0.07	1.15 $\pm$ 0.08		3.51 $\pm$ 0.57	
[6, 7]	[250, 350]	1.0	[500, 1000] ≥ 1000	H01	3.02 $\pm$ 0.36	0.0 $\pm$ 0.0		98.73 $\pm$ 2.35	
				H02	0.31 $\pm$ 0.1	0.09 $\pm$ 0.02		35.97 $\pm$ 1.62	
	[350, 450]	1.0	[500, 1000] ≥ 1000	H03	4.13 $\pm$ 0.41	0.01 $\pm$ 0.01		20.47 $\pm$ 1.12	
				H04	0.52 $\pm$ 0.14	0.14 $\pm$ 0.03		9.64 $\pm$ 0.83	
	[450, 650]	0.75	[500, 750] [750, 1250] ≥ 1250	H05	3.63 $\pm$ 0.39	0.0 $\pm$ 0.0		7.52 $\pm$ 0.95	
				H06	3.79 $\pm$ 0.39	0.03 $\pm$ 0.01		11.41 $\pm$ 0.81	
				H07	0.36 $\pm$ 0.12	0.47 $\pm$ 0.05		4.36 $\pm$ 0.62	
	≥ 650	0.5	[500, 750] [750, 1250] ≥ 1250	H08	0.89 $\pm$ 0.19	0.0 $\pm$ 0.0		1.11 $\pm$ 0.33	
				H09	1.77 $\pm$ 0.26	0.15 $\pm$ 0.03		4.68 $\pm$ 0.57	
				H10	0.83 $\pm$ 0.18	2.83 $\pm$ 0.12		2.67 $\pm$ 0.48	
≥ 8	[250, 350]	1.0	[500, 1000] ≥ 1000	I01	0.88 $\pm$ 0.18	0.0 $\pm$ 0.0		9.04 $\pm$ 0.6	
				I02	0.26 $\pm$ 0.09	0.03 $\pm$ 0.01		8.07 $\pm$ 0.62	
	[350, 450]	1.0	[500, 1000] ≥ 1000	I03	0.55 $\pm$ 0.14	0.0 $\pm$ 0.0		1.88 $\pm$ 0.25	
				I04	0.72 $\pm$ 0.15	0.11 $\pm$ 0.02		2.69 $\pm$ 0.38	
	[450, 650]	0.75	[500, 1250] ≥ 1250	I05	2.07 $\pm$ 0.26	0.01 $\pm$ 0.01		1.02 $\pm$ 0.16	
				I06	0.45 $\pm$ 0.12	0.3 $\pm$ 0.04		1.29 $\pm$ 0.33	
	≥ 650	0.5	[500, 1250] ≥ 1250	I07	0.97 $\pm$ 0.18	0.04 $\pm$ 0.01		0.99 $\pm$ 0.33	
				I08	1.12 $\pm$ 0.18	1.37 $\pm$ 0.08		0.5 $\pm$ 0.08	

Signal acceptance

The yield after baseline selection for the simulated signal events is shown in Fig. 4.15 in the gluino-neutralino mass plane for the T5qqqqWW model. The acceptance of at least one event for 35.9 fb⁻¹ integrated luminosity reaches up to 2250 GeV for gluino mass for the neutralinos masses below 1600 GeV. The selection efficiency, defined as the fraction of simulated signal events after the baseline requirement over the total number of simulated signal events, is presented in the same figure (top right). Figures 4.15 (bottom) show the MB SR efficiencies. The plot on the left displays the efficiency using a constant $\Delta\phi > 1$ cut. The efficiency is ranges from 50% to 70%. On the other hand, the right plot shows the efficiency for the $\Delta\phi$ cut taking values according to corresponding L_T bin and this time the efficiency goes up to 80%.

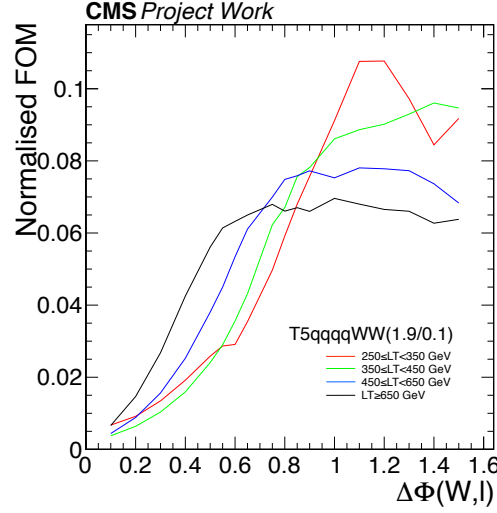


Figure 4.13: The figure of merit as a function of $\Delta\phi$ is shown for T5qqqqWW (1.9,0.1) after the baseline selection.

4.7.3 Aggregate regions

In order to facilitate future reinterpretations, aggregate signal regions are defined in this section. The goal of the aggregate signal regions is, to have a lower number of bins, while keeping sensitivity at an acceptable level.

To find optimal aggregate signal regions, three signal benchmark points have been used. The signal point with gluino mass 1800 GeV and neutralino mass 100 GeV is used to represent the compressed region. For the bulk region two lower gluino mass points (1600, 1300 GeV) has been used only for defining the aggregated signal regions.

First, an attempt is made to merge adjacent bins. Second, the expected exclusion is evaluated individually per bin. The result is shown in Fig. 4.18 as a function of the excluded signal strength modifier r , which multiplies the nominal signal strength. Last, a choice is made. Bins at high n_{jet} , high H_T , and/or high L_T are selected to form the aggregate regions and they are summarized in Tab. 4.5.

Table 4.5: Simulation table of the aggregate signal regions, 35.9 fb^{-1}

n_{jet}	L_T [GeV]	$\Delta\phi$ [rad]	H_T [GeV]	Bin name	T5qqqqWW $m_{\tilde{g}}/m_{\tilde{\chi}^0}$ [TeV]				tot. background Simulation	
					(1.5/1.0)	(1.9/0.1)				
≥ 5	≥ 650	0.5	≥ 750	LT2i, HT1i	6.15 $\pm$ 0.57	6.29 $\pm$ 0.2			20.24 $\pm$ 1.27	
≥ 6	≥ 450	0.75	≥ 500	LT1i, HT0i	16.59 $\pm$ 0.94	5.28 $\pm$ 0.19			31.22 $\pm$ 1.55	
	≥ 650	0.5	≥ 1000	LT2i, HT2i	4.01 $\pm$ 0.46	4.98 $\pm$ 0.18			6.53 $\pm$ 0.71	
≥ 7	≥ 450	0.75	≥ 500	LT1i, HT0i	9.47 $\pm$ 0.71	3.54 $\pm$ 0.15			10.97 $\pm$ 0.86	
	≥ 650	0.5	≥ 500	LT2i, HT0i	4.28 $\pm$ 0.48	3.3 $\pm$ 0.15			4.28 $\pm$ 0.59	
≥ 8	≥ 250	1.0	≥ 1250	LT0i, HT3i	1.82 $\pm$ 0.31	1.71 $\pm$ 0.11			7.85 $\pm$ 0.73	

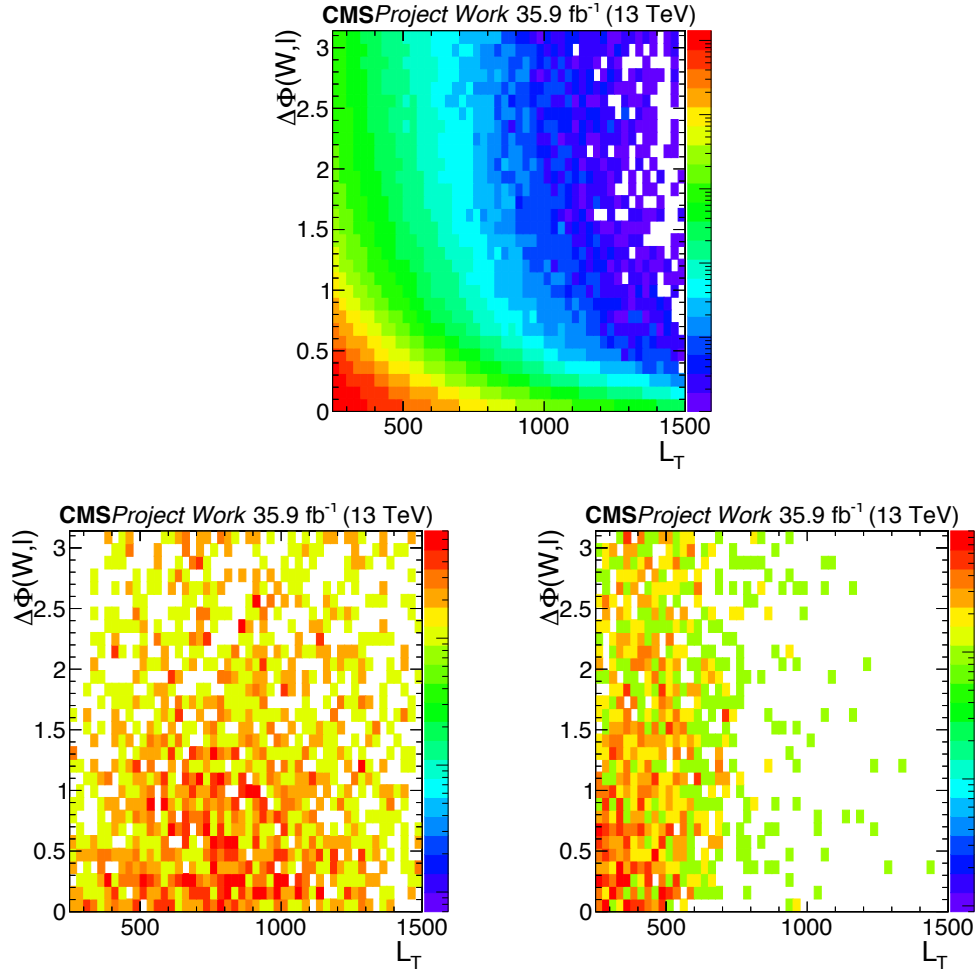


Figure 4.14: Two dimensional distributions of event counts for the main background samples, $t\bar{t}$ + jets + W + jets simulation (top), and the signal T5qqqqWW (1.9,0.1) (left) and T5qqqqWW (1.5,1.0) in the $\Delta\phi$ vs. L_T plane after the baseline selection.

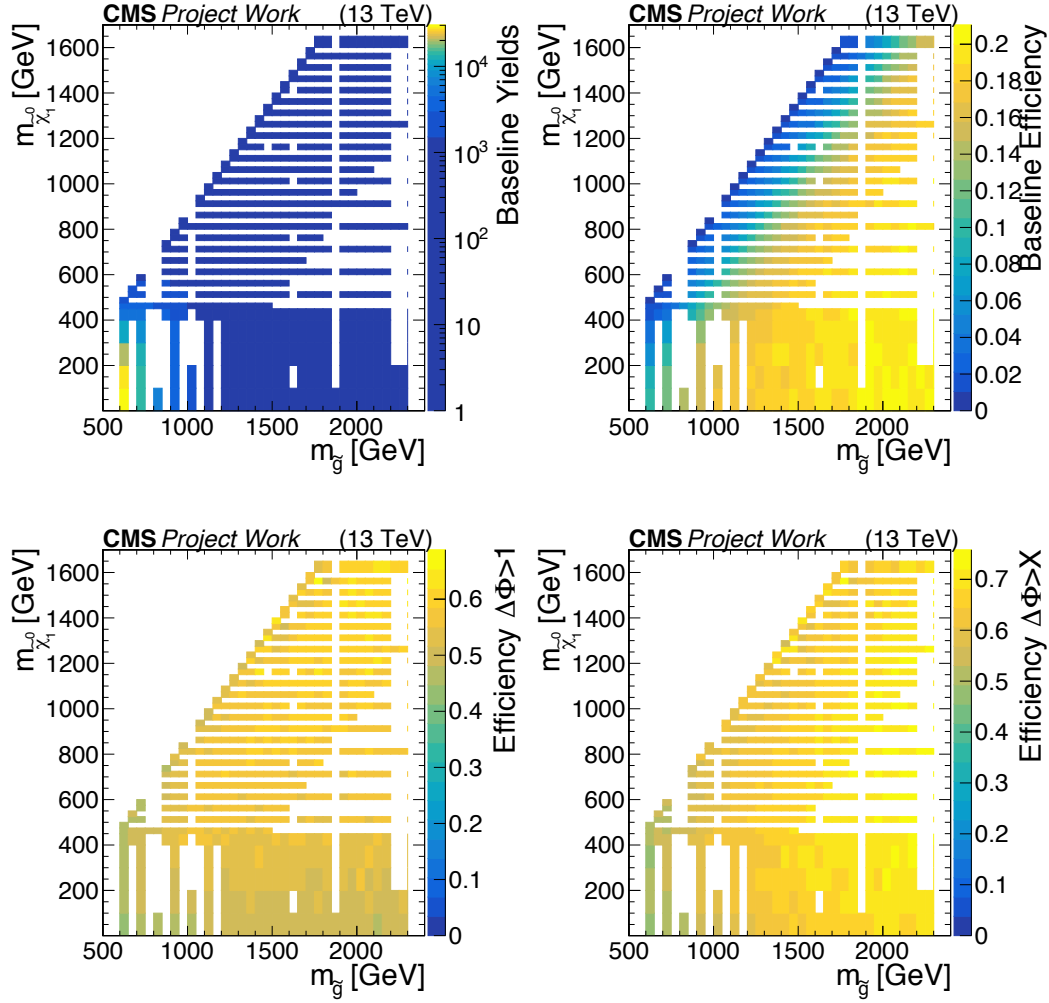


Figure 4.15: The distributions on top represent T5qqqqWW signal counts (left) and selection efficiency (right) after the baseline requirements in the gluino-neutralino plane. The distributions at the bottom show the efficiency of signal region selection with respect to the baseline selection; $\Delta\phi > 1$ (left) and $\Delta\phi > x$ (right) where x is the threshold according to L_T bin.

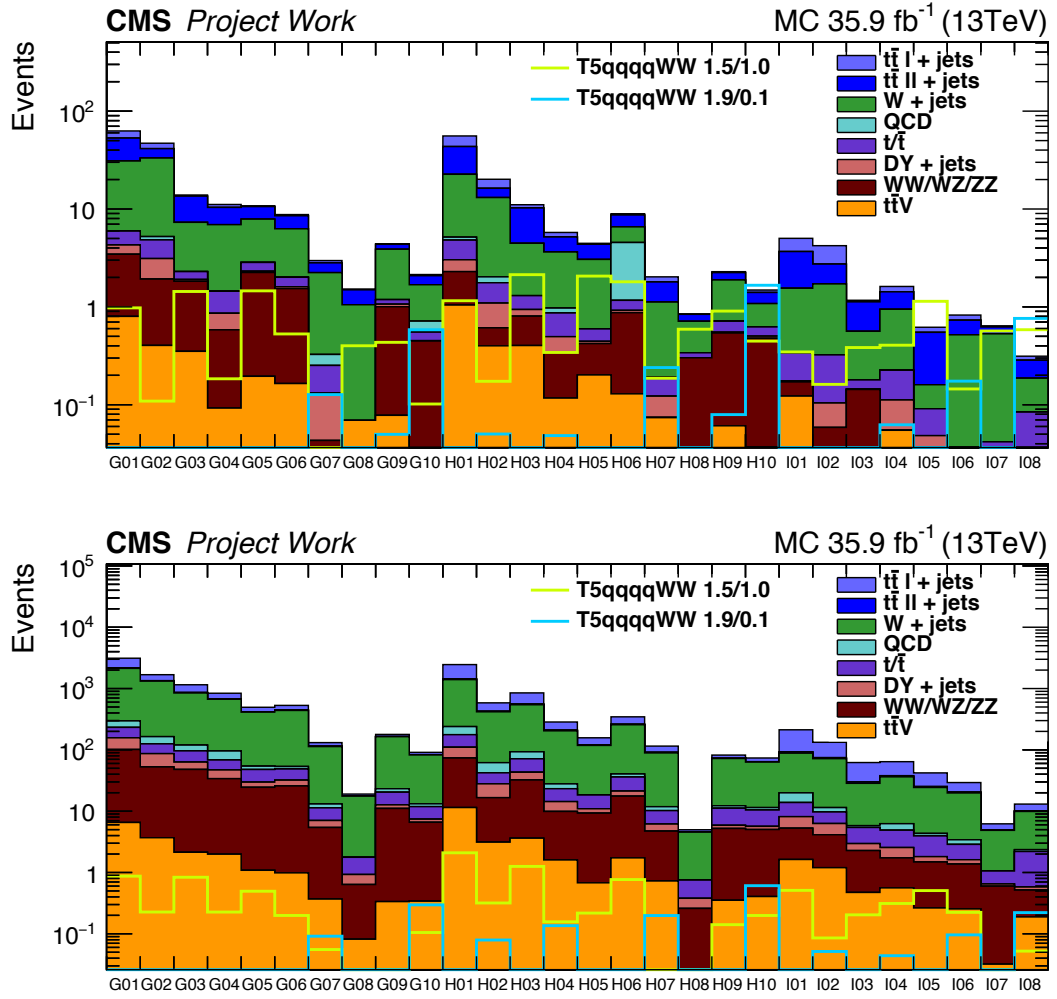


Figure 4.16: Simulated single muon event yields for the background processes are shown as color filled stacked histograms for all 28 search bins. The two signal benchmark models are overlaid and shown by line histograms. The upper plot shows the high $\Delta\phi$ regions (SR) while the lower plot shows the low $\Delta\phi$ regions (CR).

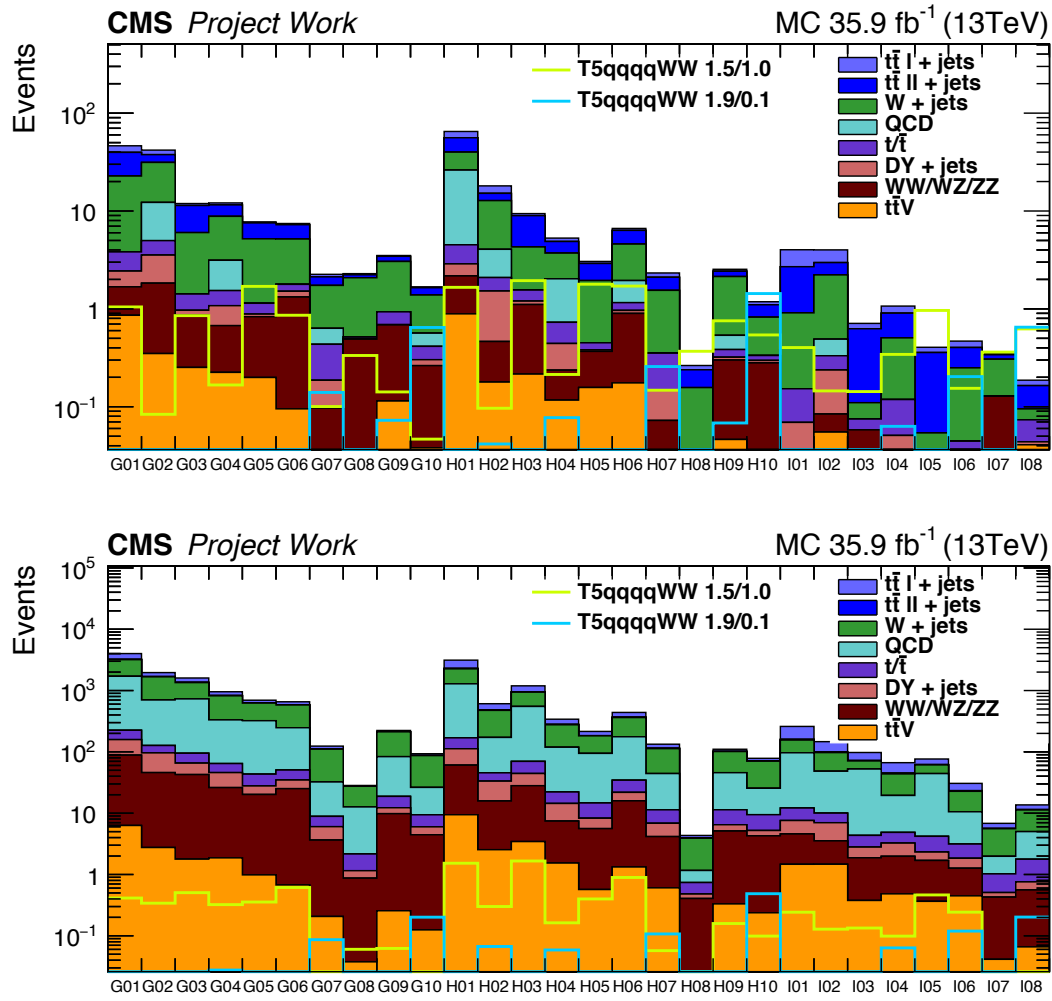


Figure 4.17: Simulated single electron event yields for the background processes are shown as color filled stacked histograms for all 28 search bins. The two signal benchmark models are overlaid and shown by line histograms. The upper plot shows the high $\Delta\phi$ regions (SR) while the lower plot shows the low $\Delta\phi$ regions (CR).

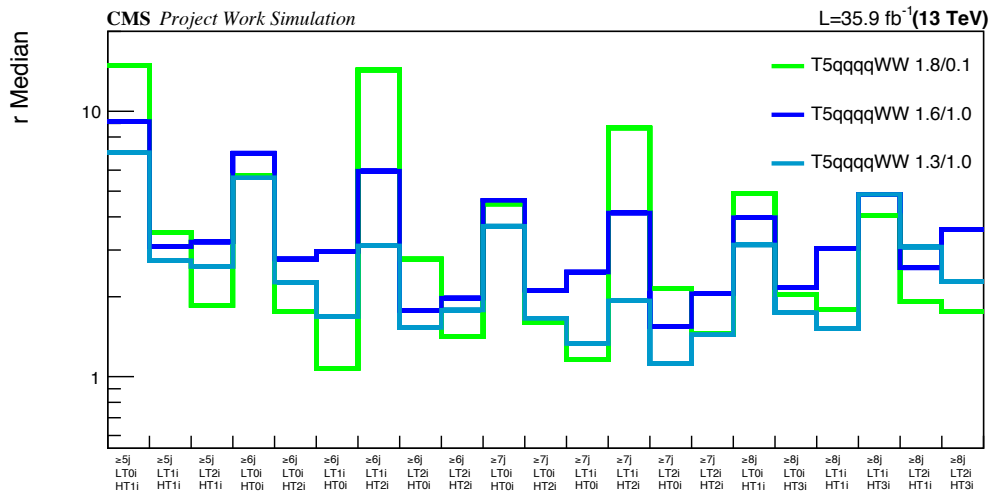


Figure 4.18: r-value calculated for each of the inclusive aggregate bins.

Chapter 5

Background estimation: The R_{CS} method

When searching for new physics, it is important to know precisely the SM background contributions in the MB SRs which are explained in Sec. 4.7.

The modeling of the SM is not trivial due to complicated processes such as QCD, and lack of perfectly realistic detector performance to the simulation tools. As a result, simulated samples have imperfections. However, they mostly provide enough information on the underlying kinematical properties of the simulated samples. Therefore, a data-driven approach is used to estimate the main backgrounds.

To predict the number of events in the MB SRs, the number of events in each of the CRs is multiplied with a transfer factor, which is called R_{CS} . The factor R_{CS} is measured in lower jet multiplicity regions, that are called SBs, in data. The SBs are largely free of hypothetical signal contributions. This procedure constitutes the core of the background estimation method. The factor R_{CS} is defined as the ratio of the number of events in the SR to the number of events in the CR:

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

The formulation of the procedure is:

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

Measuring the R_{CS} in SBs and assuming it takes similar values in MBs require that the R_{CS} is stable as a function of n_{jet} . Vetoing b-jets in the final state makes the W+jets (see Sec. 5.4) and $t\bar{t}$ +jets (see Sec. 5.3) background components approximately equal in the MB SRs. 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 takes into account the differences in R_{CS} values of these two components. In this perspective, the R_{CS} can be defined as a combination of R_{CS} from different source of backgrounds:

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

where i stands for either W +jets or $t\bar{t}$ +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 sideband regions. These regions are chosen, such that they mimic the kinematics of the mainband as closely as possible. For $t\bar{t}$ +jets, the sideband region is $4 \leq n_{jet} \leq 5$ and $n_{b-jet} \geq 1$ while for W +jets the sideband is chosen as $3 \leq n_{jet} \leq 4$ and $n_{b-jet} = 0$.

Despite the fact that QCD multijet events have low contribution in mainband signal regions, their contamination in sidebands as well as the control regions of mainbands has to be estimated and subtracted prior to the application of the R_{CS} method. Again a data based approach is used (see Sec. 5.1). Tab. 5.1 summarizes the mainband and sideband regions. 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 $R_{CS}(W$ +jets) the muon (μ) sample is used. Empty cells are not used in the analysis.

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

5.1 QCD background estimation

Due to the complicated nature of quantum chromodynamics (see Sec. 1.1.2), simulation of QCD multijet events is challenging and often fails to be accurate. In such a case, a data-driven prediction of the QCD multijet contribution 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, their contamination needs to be subtracted from the SBs and CRs of MBs.

Contributions from QCD multijet to signal region occur when a jet is misidentified as a lepton (fake lepton). The QCD multijet contamination is negligible in the after single muon selection, therefore, the prediction is only performed in single electron events. A control sample enriched with fake electrons is obtained by inverting the criteria on lepton identification variables (see Sec. 3.4.2). These anti-selected electrons are required to fail

the tight criteria in Tab. 3.4, but are still required to satisfy the loose electron Id. Additionally, H/E is required to exceed 0.01 and the requirements on the IP parameters d_{xy} and d_z as well as the photon conversion veto are removed. The isolation observable I_{mini} is required to be below 0.4, in contrast to 0.1 for the tight electron selection.

To estimate the fraction of events with fake leptons, that pass the analysis selection, the ratio of selected to anti-selected events (F_{sel}) is used. To ensure the orthogonality to the mainband regions, this ratio is measured in the $3 \leq n_{\text{jet}} \leq 4$ and $n_{b\text{-jet}} = 0$ sideband.

The kinematic properties of the fake leptons helps to distinguish QCD multijet events from the EWK ones (containing prompt electrons) by using a variable which reflects the polarization of the W boson. The variable L_P was introduced in [128], and was used for the first measurement of the W polarization at LHC. The variable L_P is defined:

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

where $p_T(\ell)$ and $p_T(W)$ are the transverse momenta of the charged lepton and the W boson respectively.

As seen in Fig. 5.1, QCD multijet events (dashed cyan colored lines) mostly populate the L_P close to 1 while the EWK events have a falling distribution from 0 to 1. The number of QCD multijet 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 right, as a function of L_T .

A profound description of the method can be found in [129].

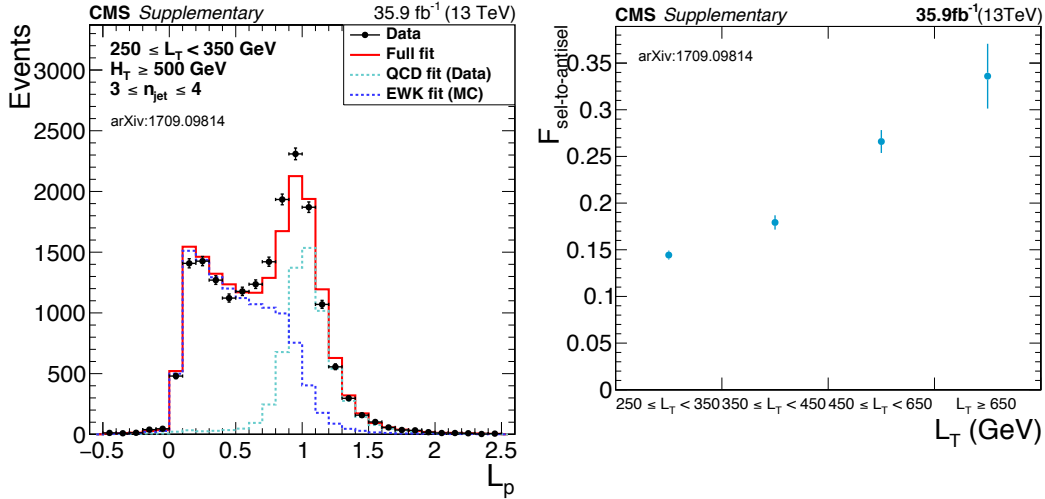


Figure 5.1: The L_P shape fit result for $3 \leq n_{\text{jet}} \leq 4$ and $n_{b\text{-jet}} = 0$ in the $250 \leq L_T \leq 350$ bin (left). Ratio of selected to anti-selected electron events from QCD multijet events for $3 \leq n_{\text{jet}} \leq 4$ and $n_{b\text{-jet}} = 0$, in bins of L_T in data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

5.2 Background fraction calculations: b-tag multiplicity fit

Fractions of the background processes in the control regions (f_i^{CR}) are calculated using a likelihood fit of b-tag multiplicity distribution templates to data. The templates are obtained from simulation, with the exception of QCD multijet events. The latter contribution is taken from the data driven prediction which is described in Sec. 5.1. W+jets events are not produced symmetrically in lepton charge at a pp collider. To account for this, the fits are performed separately for the positive and negative lepton charge. The $t\bar{t}$ +jets and QCD multijet contribution is assumed to be symmetric. Other background templates are also produced separately for positive and negative charged leptons. Fig. 5.2 shows data (black points) and the b tag multiplicity fit (blue dashed line).

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

At the LHC, the top quark pair production mechanism is dominated by gluon-gluon fusion (80%) and quark-quark annihilation(20%). According to the SM, the top quark decays into W and b quark in almost 100% of the cases¹, thus predicting two b quark partons in the $t\bar{t}$ +jets final state. However, the b tagging algorithms are not fully efficient; therefore $t\bar{t}$ events can survive veto on b-tagged jets. The remaining events have a very similar final state as the T5qqqWW signal events.

The $t\bar{t}$ decay can be split in three categories according to the lepton content of the final state; single leptonic, dileptonic and fully hadronic. The branching ratios of these decay modes are 43.8%, 10.5% and 45.7%, respectively. After the baseline selection, the single leptonic decay channel, where one of the two W bosons decays leptonically, populates the low $\Delta\phi$ region while the dileptonic decay channel, where one of the leptons is lost to misidentification or the limited detector acceptance, dominates the high $\Delta\phi$ region. As a result of the induced $\cancel{E}_T$, the dileptonic $t\bar{t}$ contribution can have high R_{CS} values. For illustration, Fig. 5.3 shows the simulated R_{CS} as a function of n_{jet} for dileptonic (left) and single leptonic (right) decays separately. In the primary background estimation procedure, R_{CS} values of $t\bar{t}$ + jets yields are measured in data sideband regions, requiring four to five jets and at least one b-tagged jet. This selection increases the purity of $t\bar{t}$ + jets events and keeps the W+jets contamination low. In addition to this, the yield of predicted QCD multijet events is subtracted from the CR, in summary:

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

Figure 5.4 shows the simulated R_{CS} in low H_T (left), and high H_T (right) values for the first search bin which is $n_{jet} = 5$, $L_T \in [250, 350]$ GeV. The simulated and measured R_{CS} values are shown in Tab. 5.2.

Residual differences between R_{CS} in the sideband and mainband are obtained in simulation as a correction factor κ . In $t\bar{t}$ + jets background estimation, $\kappa = \kappa_b \cdot \kappa_{t\bar{t}}$. The first factor κ_b corrects the residual difference between the R_{CS} values of the b-tagged region and

¹CKM induced flavour violating decays are neglected in this thesis

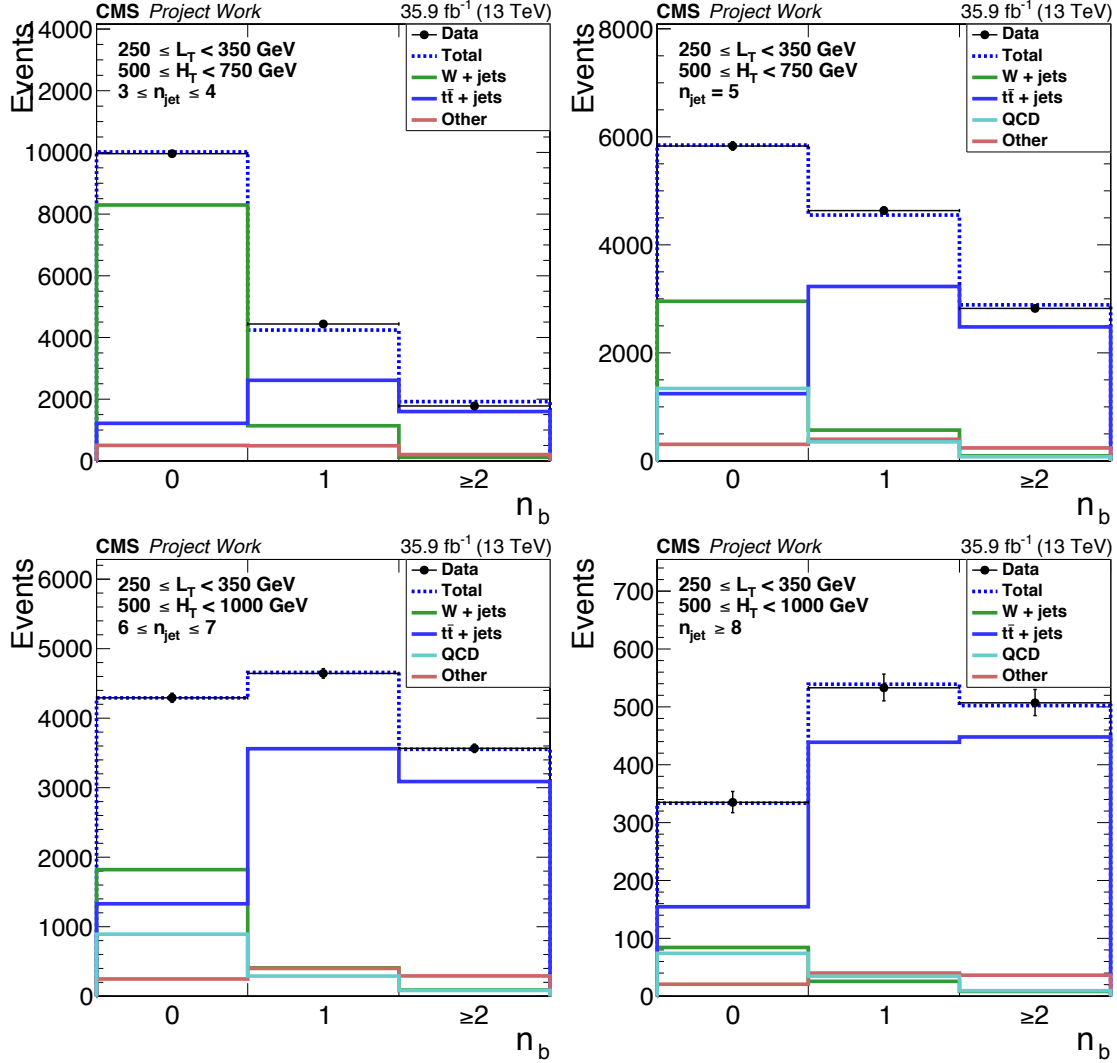


Figure 5.2: b -tag multiplicity fit is performed in control regions. 3-4 jets sideband (top/left), 5 jets mainband (right/top), 6-7 jets mainband (left/bottom) and 8 or more jets mainband fits are shown. Black points represent data, green is for W +jets, blue is for $t\bar{t}$ + jets, orange is for other EWK backgrounds and cyan colored lines show QCD contamination.

no b -tagged region. This correction also accounts for small contributions from processes other than $t\bar{t}$ + jets and QCD multijet production. 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, $\kappa_{t\bar{t}}$, accounts for a residual dependence of R_{CS} on jet multiplicity:

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

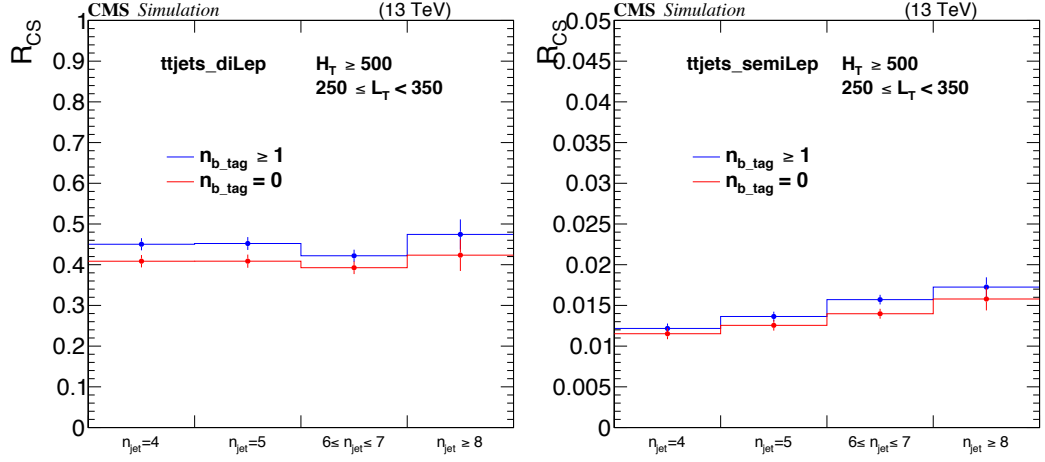


Figure 5.3: R_{CS} as a function of n_{jet} for dileptonic (left), and single leptonic (right) $tt + jets$ events.

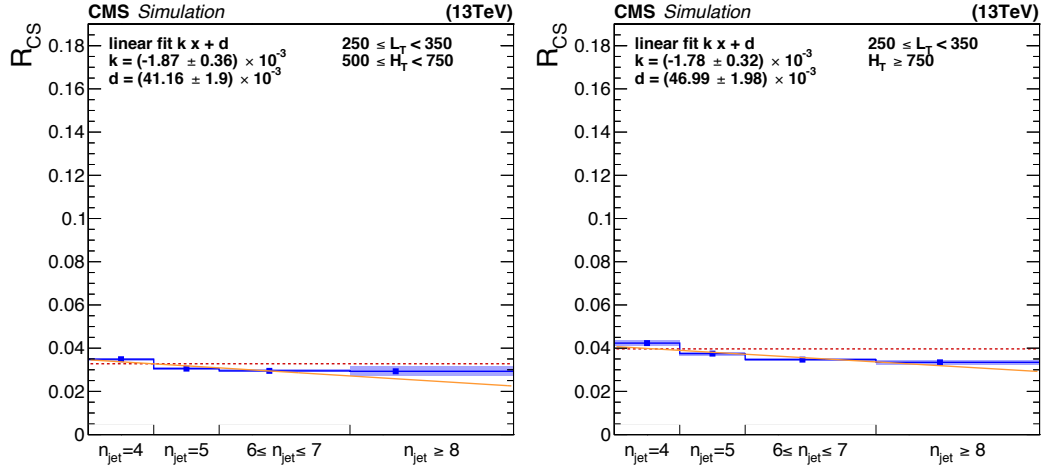


Figure 5.4: The factor R_{CS} as a function of n_{jet} from simulation of $t\bar{t} + jets$ events for low H_T (left), and high H_T (right). The red line shows the R_{CS} value in the SB that is 4-5j, $b \geq 1$. The difference between the linear fit and the red line is used to measure systematic uncertainty.

As mentioned in the beginning of this chapter, the total R_{CS} is based on the fraction of different backgrounds and their individual R_{CS} values. To this end, the difference between R_{CS} values of single leptonic and dileptonic events should be considered.

In order to obtain a high-purity dileptonic $t\bar{t}$ control sample in data, two leptons of opposite charge are required. It is moreover required, that the mass of two same flavor leptons satisfy $|m_{\ell\ell} - m_Z| > 10$ GeV. To mimic the single lepton contribution from dileptonic events,

one of the two leptons is removed. Since the lost leptons are in data primarily coming from $\tau \rightarrow \text{hadrons} + \nu$ decays, the removed lepton is replaced by a jet with $2/3$ of the original lepton's p_T and the remainder is assumed to come from the neutrino and is properly accounted for in $\cancel{E}_T$. The L_T , H_T , n_{jet} and $\Delta\phi$ values of the reconstructed single lepton event are subsequently recalculated. In the dilepton control region, no $\Delta\phi$ requirement is applied, and all events are used twice, separately for each lepton. In Fig. 5.5 (top row), the n_{jet} distributions for the lost lepton background as obtained from dileptonic tt+jet events (left) and the nominal single lepton events (right) in CR. The distributions are obtained after the baseline requirement for L_T and H_T but with a looser jet multiplicity requirement ($n_{\text{jet}} \geq 3$) to include the SBs as well. The correctness of the description of the n_{jet} distribution in simulation is determined from the double ratio, which is the single lepton and dilepton ratio between data and simulation, as shown in Fig. 5.5 (bottom).

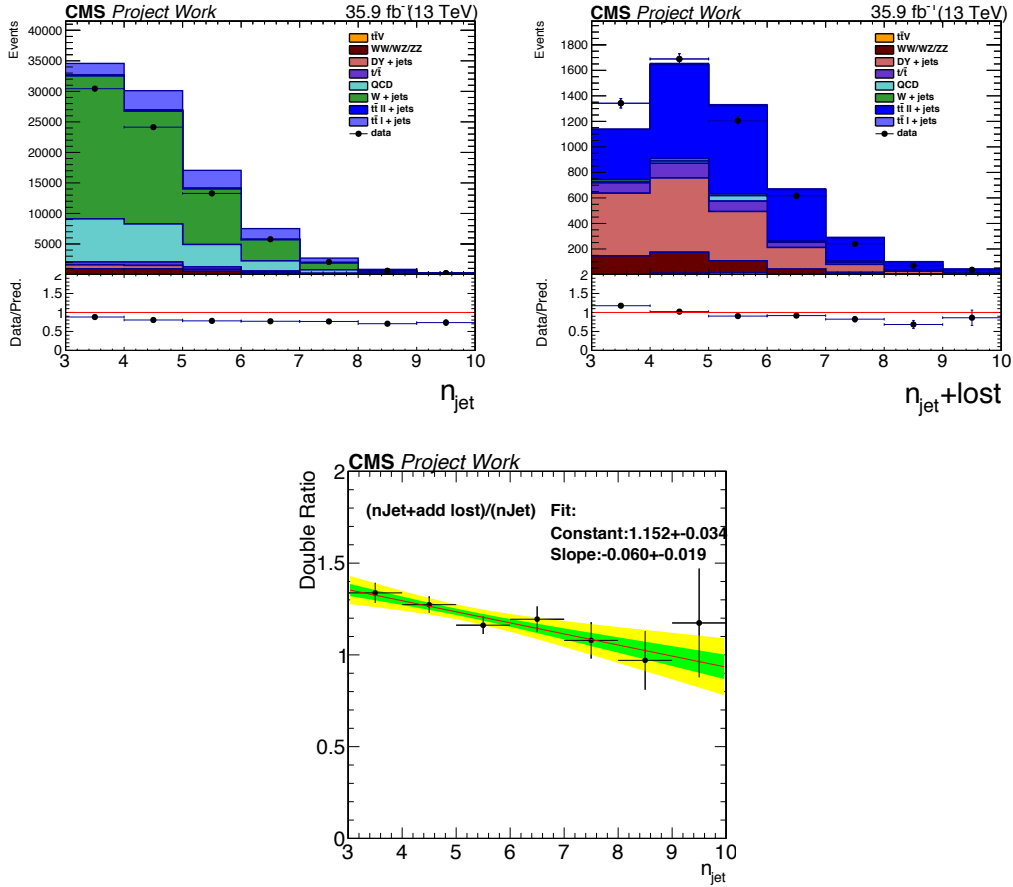


Figure 5.5: Jet multiplicity distribution after the single lepton baseline event selection (in the control region) with $H_T > 500$ GeV, $L_T > 250$ GeV, and $n_{\text{b-tag}} = 0$ (left) and in the dileptonic control region with recalculated H_T and L_T cuts (right). Only nISR reweighting is applied. The bottom figure shows the ratio of the distributions in the lower panels of single leptonic (top left) and dileptonic (top right) events.

If the simulation described the dileptonic contribution perfectly, this double ratio would be flat at unity. However, it is seen in Fig. 5.5 (bottom), that the dileptonic events in simulation should be corrected to obtain a $\kappa_{t\bar{t}}^{DL-corr}$. The double ratio in Fig. 5.5 (bottom) is fitted with a linear parametrization:

$$f(n_{\text{jet}}) = a + b(n_{\text{jet}} - \langle n_{\text{jet}} \rangle), \quad (5.9)$$

where $\langle n_{\text{jet}} \rangle$ is the weighted mean value of n_{jet} in single lepton selection, a determines the offset and b determines the slope. The weight that is used to rescale dileptonic events are then written as:

$$w_{DL} = \text{constant} + \text{slope} \cdot (n_{\text{jet}} - \langle n_{\text{jet}} \rangle), \quad (5.10)$$

where the $\langle n_{\text{jet}} \rangle$ is found to be 5.9, constant and slope values are measured 1.152 and -0.060, respectively. A new $\kappa_{t\bar{t}}$, which is denoted by $\kappa_{t\bar{t}}^{DL-Corr}$, measured as in Eq. 5.8 but this time using the samples in which dileptonic $t\bar{t}$ + jets events are weighted with w_{DL} . The comparison of $\kappa_{t\bar{t}}$ before and after correction can be seen in Fig.5.6.

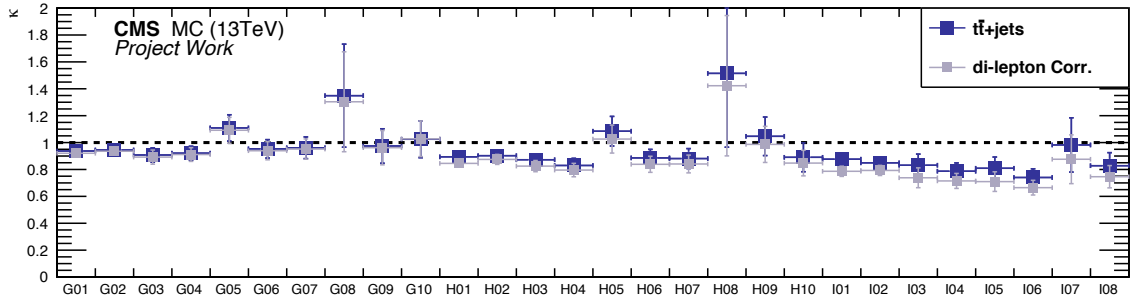


Figure 5.6: Comparing $\kappa_{t\bar{t}}$ and $\kappa_{t\bar{t}}^{DL-Corr}$ where the correction accounts for the impact of dileptonic $t\bar{t}$ events.

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

$$R_{CS}^{t\bar{t}}(0b, MB) = \kappa_b \cdot \kappa_{t\bar{t}}^{DL-Corr} \cdot R_{CS}^{data}(b \geq 1, n_{\text{jet}} \in [4, 5]). \quad (5.11)$$

A summary of κ_b corrected R_{CS} , R_{CS} in MB regions from simulation, and different κ values are shown in Tab. 5.3.

5.4 R_{CS} method in W+jets events

For the W+jets estimation, we first note that we need not separately estimate WV(V= W, Z) if the boson V decays hadronically. The contribution from leptonic V decays is a subleading uncertainty in the main analysis and taken from simulation in the SBs. For the main measurement of R_{CS} values for W+jets background estimation, we define a sideband region with three or four jets. To suppress the $t\bar{t}$ + jets events and to be kinematically as close as possible to mainband regions, we further veto events with b-tagged jets. Furthermore,

in order to suppress QCD multijet contamination, R_{CS} is measured only in the muon channel, where no QCD multijet subtraction scheme is needed. The $t\bar{t}$ +jets contamination in the W+jets SB is subtracted from SRs and CRs. Its fraction $f_{t\bar{t}}$ in the SB CR is taken again from a b tag multiplicity fit, and $R_{CS}^{t\bar{t},MC}$, that is used to obtain the $t\bar{t}$ +jets contamination in the W+jets SB SR, can be taken from simulation. The R_{CS} can then be written as:

$$R_{CS}^{corr,data}(0b, n_{jet} \in [3, 4]) = \frac{N_{data}^{SR} - R_{CS}^{t\bar{t},MC} \cdot f_{t\bar{t}}^{fit} \cdot N_{data}^{CR}}{(1 - f_{t\bar{t}}^{fit}) \cdot N_{data}^{CR}}. \quad (5.12)$$

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+jets events. The simulated R_{CS}^{MC} values and the corrected measured R_{CS} in events with 3 and 4 jets, with the b jet veto applied, are given in Tab. 5.4.

Figure 5.7 shows simulated R_{CS} in low H_T (left), and high H_T (right) values for the first

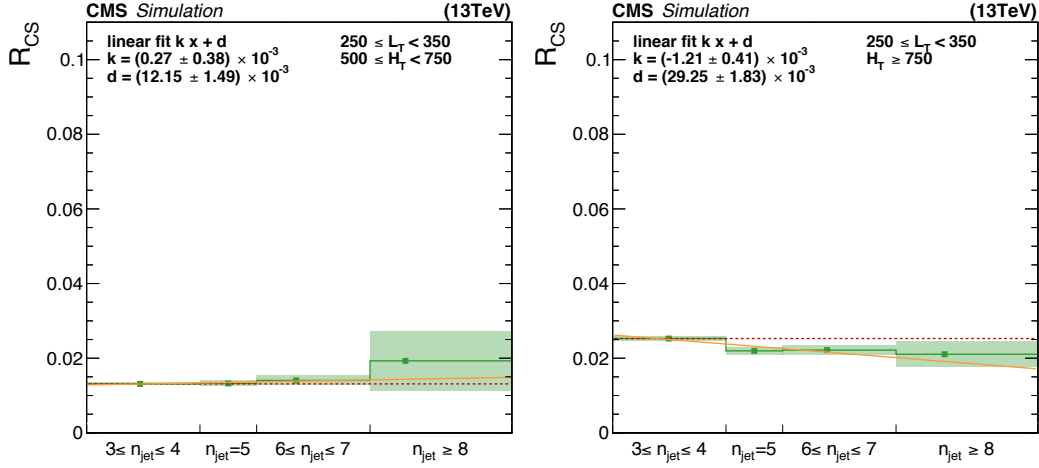


Figure 5.7: R_{CS} as a function of n_{jet} from simulation: low H_T (left), and high H_T (right) values for the first search bin which is $n_{jet} = 5$, $L_T \in [250, 350]$ GeV.

search bin which is defined by $n_{jet} = 5$, $L_T \in [250, 350]$ GeV.

Again, as in the $t\bar{t}$ case, residual differences between the R_{CS} in the sideband and mainband are calculated in simulation as a correction factor κ . This time, only one κ factor is used, due to the fact that the sideband and mainband regions are sharing the same requirement for number of b-tagged jets, $n_{b-tag} = 0$. The factor κ_w accounts for residual dependence of R_{CS} on the jet multiplicity and also covers variations of R_{CS} values in the muon channel and the inclusive single lepton channel. The factor κ_w is calculated as:

$$\kappa_w = \frac{R_{CS}^{MC}(0b, n_{jet} \text{ as in MB, W+jets})}{R_{CS}^{corr,MC}(0b, n_{jet} \in [3, 4], \mu)}. \quad (5.13)$$

Resulting in:

$$R_{CS}^W(0b, MB) = \kappa_w \cdot R_{CS}^{corr,data}(0b, n_{jet} \in [3, 4], \mu). \quad (5.14)$$

Fig. 5.8 shows the κ_w values in the lower panel and the 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_w values are still compatible with 1.

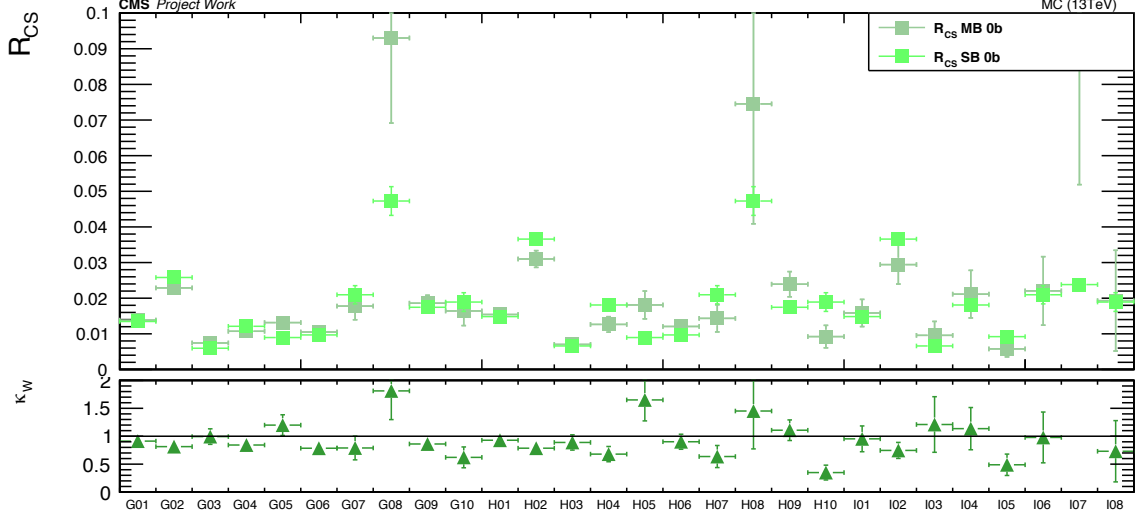


Figure 5.8: The top row shows the two simulated R_{CS} values (Eq. 5.13 numerator and denominator). The bottom row shows the resulting κ_w .

5.5 Validation of the background estimation

A 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 $n_{jet} = 3$ selection is used as the sideband. The $t\bar{t}$ +jets sideband remains unchanged since, it is not used as the validation search region. The simulated signal and background yields, together with the bin naming, can be found in Tab. 5.5. Fig. 5.9 presents the obtained κ_w^{val} and $\kappa_{t\bar{t}}^{val}$ values.

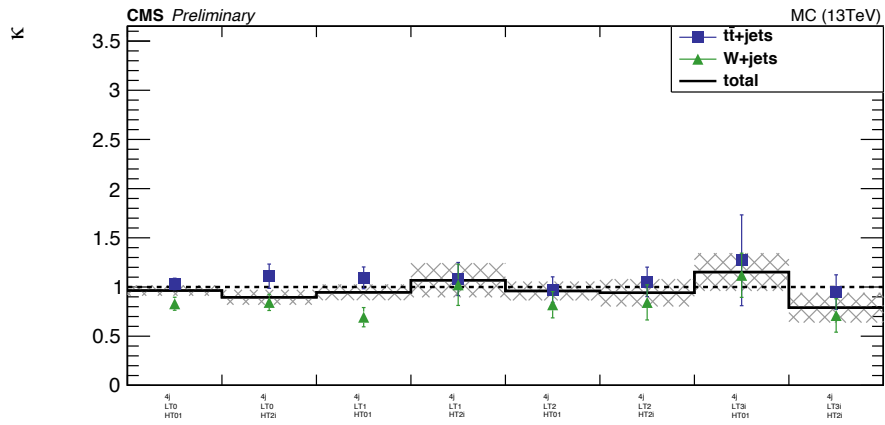


Figure 5.9: Correction factors κ_w^{val} (Green) and κ_{tt}^{val} (Blue) are shown. The black line represents a κ , only for illustration, which is the ratio of prediction to the simulation. The shaded area displays the uncertainty from the statistical precision of the simulated samples.

Table 5.2: Simulated and measured R_{CS} of the $t\bar{t}$ + jets events.

n_{jet}	L_T [GeV]	H_T [GeV]	R_{CS} in 4-5j, $\geq 1b$					
			simulated			measured		
5	[250, 350]	[500, 750] ≥ 750	0.0373	$\pm$	0.0005	0.0413	$\pm$	0.0017
			0.0454	$\pm$	0.0008	0.0549	$\pm$	0.0035
	[350, 450]	[500, 750] ≥ 750	0.0265	$\pm$	0.0007	0.0303	$\pm$	0.0026
			0.0314	$\pm$	0.0009	0.0319	$\pm$	0.0038
	[450, 650]	[500, 750] [750, 1250] ≥ 1250	0.0375	$\pm$	0.0015	0.0359	$\pm$	0.005
			0.0338	$\pm$	0.0012	0.0324	$\pm$	0.0049
	≥ 650	[500, 750] [750, 1250] ≥ 1250	0.054	$\pm$	0.0019	0.0403	$\pm$	0.0137
			0.282	$\pm$	0.037	0.1967	$\pm$	0.0971
[6, 7]	[250, 350]	[500, 1000] ≥ 1000	0.0449	$\pm$	0.0031	0.0476	$\pm$	0.0127
			0.0663	$\pm$	0.0033	0.0641	$\pm$	0.0252
	[350, 450]	[500, 1000] ≥ 1000	0.0382	$\pm$	0.0004	0.0422	$\pm$	0.0016
			0.0536	$\pm$	0.0013	0.0762	$\pm$	0.0076
	[450, 650]	[500, 1000] ≥ 1000	0.027	$\pm$	0.0006	0.0291	$\pm$	0.0022
			0.0379	$\pm$	0.0013	0.0459	$\pm$	0.0082
	≥ 650	[500, 750] [750, 1250] ≥ 1250	0.0375	$\pm$	0.0015	0.0359	$\pm$	0.005
			0.0338	$\pm$	0.0012	0.0324	$\pm$	0.0049
≥ 8	[250, 350]	[500, 1250] ≥ 1250	0.054	$\pm$	0.0019	0.0403	$\pm$	0.0137
			0.282	$\pm$	0.037	0.1967	$\pm$	0.0971
	[350, 450]	[500, 750] [750, 1250] ≥ 1250	0.0449	$\pm$	0.0031	0.0476	$\pm$	0.0127
			0.0663	$\pm$	0.0033	0.0641	$\pm$	0.0252
	[450, 650]	[500, 1000] ≥ 1000	0.0382	$\pm$	0.0004	0.0422	$\pm$	0.0016
			0.0536	$\pm$	0.0013	0.0762	$\pm$	0.0076
	≥ 650	[500, 1000] ≥ 1000	0.027	$\pm$	0.0006	0.0291	$\pm$	0.0022
			0.0379	$\pm$	0.0013	0.0459	$\pm$	0.0082
≥ 8	[450, 650]	[500, 1250] ≥ 1250	0.0357	$\pm$	0.001	0.0343	$\pm$	0.0035
			0.054	$\pm$	0.0019	0.0403	$\pm$	0.0137
	≥ 650	[500, 1250] ≥ 1250	0.0662	$\pm$	0.0042	0.0608	$\pm$	0.0138
			0.0663	$\pm$	0.0033	0.0641	$\pm$	0.0252

Table 5.3: Summary table for R_{CS} for $t\bar{t}$ + jets and the corresponding $\kappa_{t\bar{t}}$ value from simulation

n_{jet}	L_T [GeV]	H_T [GeV]	$R_{CS}(4-5j, b \geq 1) \cdot \kappa_b^{MC}$		$R_{CS}(MB, 0b)$		$\kappa_{t\bar{t}}$ MB/SB		κ_b^{MC} 0b/ $b \geq 1$	
5	[250, 350]	[500, 750] ≥ 750	0.0393	$\pm$ 0.0019	0.0305	$\pm$ 0.0006	0.92	$\pm$ 0.03	0.95	$\pm$ 0.02
			0.0498	$\pm$ 0.0034	0.0375	$\pm$ 0.0009	0.94	$\pm$ 0.03	0.91	$\pm$ 0.02
	[350, 450]	[500, 750] ≥ 750	0.0316	$\pm$ 0.003	0.0222	$\pm$ 0.001	0.91	$\pm$ 0.06	1.04	$\pm$ 0.04
			0.0295	$\pm$ 0.0037	0.0245	$\pm$ 0.0009	0.93	$\pm$ 0.05	0.93	$\pm$ 0.03
	[450, 650]	[500, 750] [750, 1250] ≥ 1250	0.0389	$\pm$ 0.006	0.0354	$\pm$ 0.0024	1.09	$\pm$ 0.1	1.08	$\pm$ 0.06
			0.0352	$\pm$ 0.0057	0.0286	$\pm$ 0.0014	0.95	$\pm$ 0.07	1.09	$\pm$ 0.05
	≥ 650	[500, 750] [750, 1250] ≥ 1250	0.0354	$\pm$ 0.0123	0.0396	$\pm$ 0.0014	0.97	$\pm$ 0.08	0.88	$\pm$ 0.05
			0.3336	$\pm$ 0.1809	0.2974	$\pm$ 0.0656	1.31	$\pm$ 0.39	1.7	$\pm$ 0.3
[6, 7]	[250, 350]	[500, 1000] ≥ 1000	0.0589	$\pm$ 0.0168	0.0386	$\pm$ 0.0032	0.94	$\pm$ 0.12	1.24	$\pm$ 0.11
			0.0867	$\pm$ 0.035	0.0505	$\pm$ 0.0042	1.05	$\pm$ 0.14	1.35	$\pm$ 0.12
	[350, 450]	[500, 1000] ≥ 1000	0.04	$\pm$ 0.0017	0.0301	$\pm$ 0.0005	0.85	$\pm$ 0.02	0.95	$\pm$ 0.02
			0.0652	$\pm$ 0.0069	0.0408	$\pm$ 0.0008	0.87	$\pm$ 0.04	0.86	$\pm$ 0.03
	[450, 650]	[500, 1000] ≥ 1000	0.0297	$\pm$ 0.0025	0.021	$\pm$ 0.0008	0.82	$\pm$ 0.04	1.02	$\pm$ 0.03
			0.0408	$\pm$ 0.0075	0.0255	$\pm$ 0.0008	0.8	$\pm$ 0.05	0.89	$\pm$ 0.04
	≥ 650	[500, 750] [750, 1250] ≥ 1250	0.0389	$\pm$ 0.006	0.0348	$\pm$ 0.0029	1.03	$\pm$ 0.11	1.08	$\pm$ 0.06
			0.0352	$\pm$ 0.0057	0.0266	$\pm$ 0.0013	0.85	$\pm$ 0.06	1.09	$\pm$ 0.05
≥ 8	[250, 350]	[500, 1250] ≥ 1250	0.0354	$\pm$ 0.0123	0.0363	$\pm$ 0.0012	0.85	$\pm$ 0.07	0.88	$\pm$ 0.05
			0.3336	$\pm$ 0.1809	0.2827	$\pm$ 0.0951	1.21	$\pm$ 0.48	1.7	$\pm$ 0.3
	[350, 450]	[500, 750] [750, 1250] ≥ 1250	0.0589	$\pm$ 0.0168	0.0422	$\pm$ 0.0044	0.98	$\pm$ 0.14	1.24	$\pm$ 0.11
			0.0867	$\pm$ 0.035	0.0413	$\pm$ 0.002	0.82	$\pm$ 0.09	1.35	$\pm$ 0.12
	[450, 650]	[500, 1000] ≥ 1000	0.0288	$\pm$ 0.0014	0.0288	$\pm$ 0.0014	0.77	$\pm$ 0.04	0.95	$\pm$ 0.02
			0.0652	$\pm$ 0.0069	0.0388	$\pm$ 0.0012	0.8	$\pm$ 0.04	0.86	$\pm$ 0.03
	≥ 650	[500, 1000] ≥ 1000	0.0297	$\pm$ 0.0025	0.02	$\pm$ 0.0019	0.74	$\pm$ 0.08	1.02	$\pm$ 0.03
			0.0408	$\pm$ 0.0075	0.0243	$\pm$ 0.0014	0.72	$\pm$ 0.06	0.89	$\pm$ 0.04
≥ 8	[450, 650]	[500, 1250] ≥ 1250	0.0372	$\pm$ 0.0043	0.0245	$\pm$ 0.0024	0.7	$\pm$ 0.07	1.08	$\pm$ 0.04
			0.0354	$\pm$ 0.0123	0.0302	$\pm$ 0.0013	0.67	$\pm$ 0.06	0.88	$\pm$ 0.05
	≥ 650	[500, 1250] ≥ 1250	0.081	$\pm$ 0.0203	0.0599	$\pm$ 0.0112	0.89	$\pm$ 0.19	1.33	$\pm$ 0.11
			0.0867	$\pm$ 0.035	0.0395	$\pm$ 0.0019	0.74	$\pm$ 0.09	1.35	$\pm$ 0.12

Table 5.4: The simulated R_{CS} values and the corrected measured R_{CS} (as in Eq. 5.12) in events with 3 and 4 jets, 0 b-tagged jet are shown separately for positively and negatively charged leptons.

n_{jet}	L_T [GeV]	H_T [GeV]	R_{CS} in 3-4j, 0b					
			ℓ^+			ℓ^-		
			simulated	measured		simulated	measured	
5	[250, 350]	[500, 750] ≥ 750	0.0098 $\pm$ 0.0005	0.0142 $\pm$ 0.0023		0.0155 $\pm$ 0.0005	0.0242 $\pm$ 0.0021	
			0.0219 $\pm$ 0.0009	0.0347 $\pm$ 0.0057		0.0275 $\pm$ 0.0008	0.0403 $\pm$ 0.0045	
	[350, 450]	[500, 750] ≥ 750	0.0044 $\pm$ 0.0005	0.0094 $\pm$ 0.0029		0.0067 $\pm$ 0.0005	0.0116 $\pm$ 0.0022	
			0.01 $\pm$ 0.0009	0.0169 $\pm$ 0.0059		0.0129 $\pm$ 0.0008	0.0252 $\pm$ 0.005	
	[450, 650]	[500, 750] [750, 1250] ≥ 1250	0.0076 $\pm$ 0.0008	0.0105 $\pm$ 0.0036		0.0102 $\pm$ 0.0007	0.0136 $\pm$ 0.0029	
			0.0078 $\pm$ 0.0009	0.0189 $\pm$ 0.0068		0.0106 $\pm$ 0.0007	0.0125 $\pm$ 0.004	
	≥ 650	[500, 750] [750, 1250] ≥ 1250	0.0206 $\pm$ 0.0043	0.0059 $\pm$ 0.0168		0.0213 $\pm$ 0.0034	0.0612 $\pm$ 0.0257	
			0.0468 $\pm$ 0.007	0.0143 $\pm$ 0.0144		0.0472 $\pm$ 0.0051	0.1287 $\pm$ 0.0379	
[6, 7]	[250, 350]	[500, 1000] ≥ 1000	0.0112 $\pm$ 0.0004	0.0178 $\pm$ 0.0023		0.0169 $\pm$ 0.0004	0.027 $\pm$ 0.002	
			0.033 $\pm$ 0.0023	0.0381 $\pm$ 0.0107		0.0371 $\pm$ 0.0019	0.0439 $\pm$ 0.0082	
	[350, 450]	[500, 1000] ≥ 1000	0.005 $\pm$ 0.0004	0.0107 $\pm$ 0.0027		0.0073 $\pm$ 0.0004	0.0132 $\pm$ 0.002	
			0.0157 $\pm$ 0.0023	0.0191 $\pm$ 0.0102		0.0187 $\pm$ 0.0019	0.0377 $\pm$ 0.0106	
	[450, 650]	[500, 750] [750, 1250] ≥ 1250	0.0076 $\pm$ 0.0008	0.0105 $\pm$ 0.0036		0.0102 $\pm$ 0.0007	0.0136 $\pm$ 0.0029	
			0.0078 $\pm$ 0.0009	0.0189 $\pm$ 0.0068		0.0106 $\pm$ 0.0007	0.0125 $\pm$ 0.004	
	≥ 650	[500, 750] [750, 1250] ≥ 1250	0.0206 $\pm$ 0.0043	0.0059 $\pm$ 0.0168		0.0213 $\pm$ 0.0034	0.0612 $\pm$ 0.0257	
			0.0468 $\pm$ 0.007	0.0143 $\pm$ 0.0144		0.0472 $\pm$ 0.0051	0.1287 $\pm$ 0.0379	
≥ 8	[250, 350]	[500, 1000] ≥ 1000	0.0112 $\pm$ 0.0004	0.0178 $\pm$ 0.0023		0.0169 $\pm$ 0.0004	0.027 $\pm$ 0.002	
			0.033 $\pm$ 0.0023	0.0381 $\pm$ 0.0107		0.0371 $\pm$ 0.0019	0.0439 $\pm$ 0.0082	
	[350, 450]	[500, 1000] ≥ 1000	0.005 $\pm$ 0.0004	0.0107 $\pm$ 0.0027		0.0073 $\pm$ 0.0004	0.0132 $\pm$ 0.002	
			0.0157 $\pm$ 0.0023	0.0191 $\pm$ 0.0102		0.0187 $\pm$ 0.0019	0.0377 $\pm$ 0.0106	
	[450, 650]	[500, 1250] ≥ 1250	0.0077 $\pm$ 0.0006	0.0132 $\pm$ 0.0033		0.0104 $\pm$ 0.0005	0.0132 $\pm$ 0.0024	
			0.0206 $\pm$ 0.0043	0.0059 $\pm$ 0.0168		0.0213 $\pm$ 0.0034	0.0612 $\pm$ 0.0257	
	≥ 650	[500, 1250] ≥ 1250	0.018 $\pm$ 0.0018	0.013 $\pm$ 0.0075		0.0267 $\pm$ 0.0015	0.052 $\pm$ 0.0105	
			0.0185 $\pm$ 0.0048	0.0349 $\pm$ 0.0289		0.0207 $\pm$ 0.0035	0.0339 $\pm$ 0.0172	

Table 5.5: Simulation table of the validation regions, 35.9 fb $^{-1}$

n_{jet}	L_T [GeV]	H_T [GeV]	Bin name	T5qqqqWW $m_{\tilde{g}}/m_{\tilde{\chi}^0}$ [TeV]		tot. background		
				(1.5/1.0)	(1.9/0.1)	Simulation		
4	[250, 350]	[500, 1000] ≥ 1000	LT0, HT01	0.8 $\pm$ 0.21	0.01 $\pm$ 0.01	250.18 $\pm$ 10.28		
			LT0, HT2i	0.0 $\pm$ 0.0	0.01 $\pm$ 0.01	63.61 $\pm$ 3.64		
	[350, 450]	[500, 1000] ≥ 1000	LT1, HT01	0.96 $\pm$ 0.23	0.01 $\pm$ 0.01	64.88 $\pm$ 4.74		
			LT1, HT2i	0.11 $\pm$ 0.08	0.0 $\pm$ 0.0	17.46 $\pm$ 1.96		
	[450, 650]	[500, 1000] ≥ 1000	LT2, HT01	1.61 $\pm$ 0.29	0.01 $\pm$ 0.01	60.66 $\pm$ 4.83		
			LT2, HT2i	0.05 $\pm$ 0.05	0.07 $\pm$ 0.02	16.65 $\pm$ 1.66		
	≥ 650	[500, 1000] ≥ 1000	LT3i, HT01	0.48 $\pm$ 0.16	0.03 $\pm$ 0.01	26.46 $\pm$ 3.32		
			LT3i, HT2i	0.0 $\pm$ 0.0	0.31 $\pm$ 0.05	13.49 $\pm$ 1.5		

Chapter 6

Systematic uncertainties

The data driven background estimation largely eliminates the systematic uncertainties in modeling the tails of kinematic distributions such as L_T , H_T , $\Delta\phi$. However, there are still several sources of uncertainties that can affect the background prediction as well as the expected signal events counts. These sources and the methods to estimate the uncertainties are discussed in the following.

6.1 Systematic uncertainties on background estimation

The factors κ_w (Eq. 5.13) and $\kappa_{t\bar{t}}$ (Eq. 5.8) depend on simulated quantities. These factors are influenced by the mismodeling of the R_{CS} and the fractions of the SM background processes. Therefore, the systematic uncertainties on the background estimation are determined as:

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

where κ' reflects a systematic variation 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 obtained from a fit which is performed over the n_{jet} range as in Fig. 5.7. The relative difference between the R_{CS}^{MC} value obtained in the sideband and the value derived from the linear fit is taken as a systematic uncertainty on $W + jets$ events. For $t\bar{t} + jets$, the dominant effect is the different composition of single and dileptonic events. The factor $\kappa_{t\bar{t}}$ is recalculated after reweighting events according to the procedure described in Sec. 5.3. As can be read from the Fig. 5.5 (bottom), the bias on the constant (15.2%) and its uncertainty (3.4%) are added in quadrature. The dileptonic events are then rescaled with $w_{DL(Const)}$:

$$w_{DL(Const)} = 1 \pm 15.57\%. \quad (6.2)$$

The bias on the slope (6%) and its uncertainty (1.9%) are added in quadrature as well and then resulting weight, $w_{DL(Slope)}$ is obtained:

$$w_{DL(Slope)} = 1 \pm (n_{jet} - \langle n_{jet} \rangle) \cdot 6.3\%, \quad (6.3)$$

where $\langle n_{\text{jet}} \rangle = 5.9$. The uncertainties derived from the slope and constant variation for each MB SR are shown in Fig. 6.1. In this Figure, color filled areas represent the dilepton uncertainties while the black line shows the uncertainty which is calculated using a linear fit over the n_{jet} distribution. The offset variation almost has no effect on $\kappa_{t\bar{t}}$ while the slope variation has an effect up to 20%. The figure shows that the impact of the different fractions of single and dileptonic $t\bar{t}$ events are well understood.

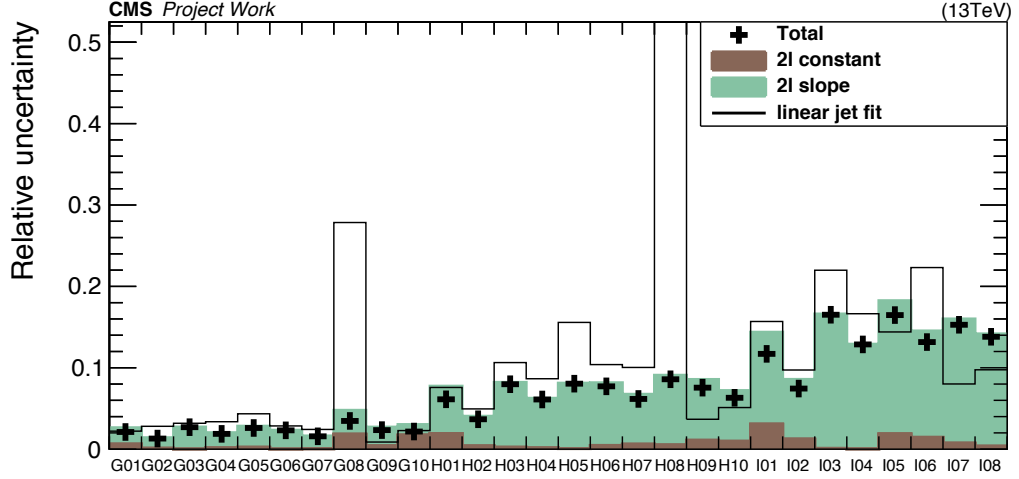


Figure 6.1: Relative uncertainty on $\kappa_{t\bar{t}}$ due to the different composition of dileptonic events in sideband and mainband regions. Color filled areas represent the dilepton uncertainties, while the black line shows the uncertainty which is calculated with a fit on the n_{jet} distribution.

Cross sections

To account for possible biases in the estimation of the background composition in terms of W +jets vs. $t\bar{t}$ +jets events, uncertainties on their cross sections are taken in to account. Although the W boson cross section and $t\bar{t}$ cross section uncertainties are at the order of ten percent level in the inclusive regions, they can be larger in the restricted phase space of the side bands. The W +jets and $t\bar{t}$ +jets cross sections are conservatively varied by $\pm 30\%$, which leads to an average change of 0.3-10% (0.7-13%) in the κ_w ($\kappa_{t\bar{t}}$) values.

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

W polarization

The main search variable $\Delta\phi$ reflects the angular information between the W boson and its decay products. Therefore, the W boson polarization affects the $\Delta\phi$ distribution.

The angular distribution [131] of the leptonic W boson decay products in the W boson rest frame takes values in the interval $0-\pi$, and the distribution is:

$$\frac{1}{\sigma} \frac{d\sigma}{d\cos\theta^*} = \frac{3}{8}(1 \mp \cos\theta^*)^2 f_L + \frac{3}{8}(1 \pm \cos\theta^*)^2 f_R + \frac{3}{4}\sin^2\theta^* f_0, \quad (6.4)$$

where the upper sign is for W^+ boson and the lower sign is for W^- boson, and f_L , f_R and f_0 denote the polarization fractions associated with the W-boson helicity ∓ 1 , helicity ± 1 , and helicity 0, respectively. The normalization are chosen so that

$$\int_{-1}^1 d\cos\theta^* \frac{1}{\sigma} \frac{d\sigma}{d\cos\theta^*} = f_L + f_R + f_0 = 1. \quad (6.5)$$

The effect of the polarization on background estimation of W+jets and $t\bar{t}$ + jets is calculated by reweighting events according to the angle between the lepton and the leptonically decaying W boson in its rest frame within the uncertainties of the measured W boson helicity fractions [130]. The resulting uncertainties are found to be below 3%.

ISR jet multiplicity

In Sec. 4.2.1, the reweighting 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. 4.1. The amount of variation is then estimated using the Eq. 6.1 and found to be ranging between 0.2-11%.

QCD multijet events prediction

QCD multijet prediction is entering in this search via the b-tag multiplicity fit (see Sec.5.2) to measure the background compositions in control regions. Therefore the uncertainty in the QCD multijet prediction has to be propagated. A profound explanation of the QCD prediction and its uncertainty can be found in [129]. The effect of this uncertainty is propagated to the R_{CS} prediction. The resultant uncertainty is found to be approximately 3% on κ_w and $\kappa_{t\bar{t}}$.

R_{CS} in the muon vs. the inclusive lepton selection

As discussed in Sec. 5.4, R_{CS} of W+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 simulation. The discrepancy is assigned as systematic uncertainty. In order to avoid large uncertainties driven by the limited statistics, the estimated systematic uncertainty is constrained to be smaller than the statistical error on R_{CS}^{MC} .

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:

$$\delta = \frac{N'_{\text{events}}}{N_{\text{events}}} - 1, \quad (6.6)$$

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 on the signals are calculated separately for each of the 657 gluino/neutrino mass points.

Initial state radiation

In Run 1, it was observed that the hadronic recoil from initial state radiation (ISR) for

boosted heavy particle pairs such as the $t\bar{t}$ system is not well described by the MADGRAPH event generator. Because gluino pair production is also dominated by gluon fusion, it is expected to be subject to a similar mismodeling. 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 according to Eq. 6.6.

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. The cross section is kept constant and from the total of 8 possible variations, the anti-correlated ones are dropped. Then, an envelope of all variations is computed for each acceptance. The resultant uncertainty on the expected event yields is similar for all the mass points and it ranges between 1-3%.

Reconstruction of E_T

FastSim signals are subject to E_T mismodeling. In order to account for this effect, the yields are obtained based on E_T and E_T^{gen} , where the latter is obtained at generator level. Then, all kinematical observables are recalculated in the two cases, as well as the acceptances. The mean of the acceptances is taken as the central value and half of the difference is taken as systematic uncertainty.

Trigger

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

Luminosity

As shown in Sec. 2.2.7, the pixel cluster counting method [90] is used to calculate luminosity. The uncertainty on this measurement is 2.5% [91].

6.3 Common systematic uncertainties for signal and background modelling

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

Jet Energy Scale

Variations of jet energy corrections within the uncertainties, which affect the jet energy spectrum as well as E_T , are applied to each jet in an event and therefore related kinematic variables, such as $\Delta\phi$, L_T , H_T , n_{jet} , $n_{\text{b-jet}}$ are recalculated.

For the background estimate, κ values are recalculated with up and down scaled jets. The uncertainty, which is obtained using Eq. 6.1, is found to be varying between 0.7-26%. For the signal, yields are recalculated and variation is measured using Eq. 6.6. 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 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

The differences in the identification and reconstruction efficiencies between the simulated and observed data events are evaluated using tag-and-probe methods [119]. For the backgrounds, a flat 5% uncertainty is assigned to account for this difference. For signal, additional uncertainties which originate from the difference between FastSim and full detector simulation are taken into account. The resulting uncertainty on the expected signal yields is found to be 2%.

Pileup

To cover the difference between the simulated pileup and the one in data, the inelastic cross section is varied by 5% up and down and the varied versions of Fig. 4.4 is obtained. For the background, the varied weights are propagated to the κ calculation. For the signal, due to the fact that the signal yields in MB SR do not significantly depend on pileup, FastSim samples are not reproduced for the high pileup environment. Nevertheless, a cross-check on the PU dependence of the two benchmark point is performed. The PU dependence is calculated as follows:

$$\text{PU}_{\text{dependence}} = \frac{\text{Efficiency}_{\text{high}}}{\text{Efficiency}_{\text{low}}}, \quad (6.7)$$

where $\text{Efficiency}_{\text{high}}$ and $\text{Efficiency}_{\text{low}}$ is:

$$\begin{aligned} \text{Efficiency}_{\text{high}} &= \frac{N(n_{\text{vertices}} \geq 20, \text{MB})}{N(n_{\text{vertices}} \geq 20, \text{baseline})}, \\ \text{Efficiency}_{\text{low}} &= \frac{N(n_{\text{vertices}} < 20, \text{MB})}{N(n_{\text{vertices}} < 20, \text{baseline})}. \end{aligned} \quad (6.8)$$

The PU dependence as a function of search regions is shown in Fig. 6.2. The distribution is flat, therefore the uncertainty can be ignored.

However, still an uncertainty has to be assigned to cover the differences between the simulated signal pileup distribution and data, following steps have been employed. 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 (left). Second, the simulated efficiency in MB SR is calculated for low and high pileup region as in Eq. 6.8. Third, these four points, from first and second steps, together compose the two points in Fig. 6.3 (middle). A linear fit is performed to extrapolate for all the pileup range. Finally, the data pileup distribution in Fig. 6.3 is normalized and folded with the fit and the sum is calculated.

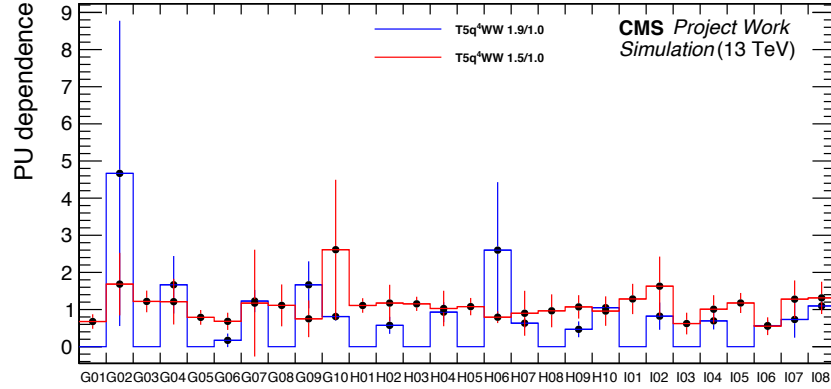


Figure 6.2: Pileup dependence of the two signal benchmark points. The blue line is representing the high mass gap point the red one is for low mass gap region. The histogram points are following a flat line around 1.

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 mainband region. The resulting 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.

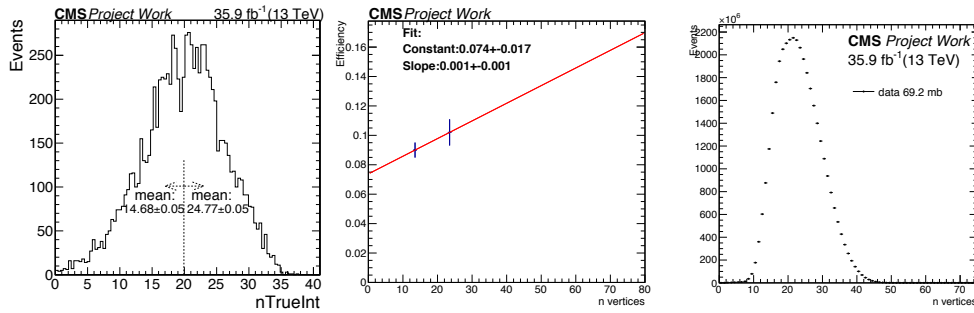


Figure 6.3: Distribution of pileup for the low mass gap signal benchmark sample, divided in to two parts as low and high pileup region (left), example fit performed using the two points explained in the text (middle), data pileup distribution which is folded with the fit.

A summary of different systematic uncertainties on the background prediction for main-band signal regions are shown in Fig. 6.4. Moreover, a summary of the systematic uncertainties on the simulated signal events for the mass point $m_{\tilde{g}} = 1900$ GeV and $m_{\tilde{\chi}_1^0} = 100$ GeV

is shown in Fig. 6.5. 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 and statistical uncertainties are shown in the lower band of the two figures.

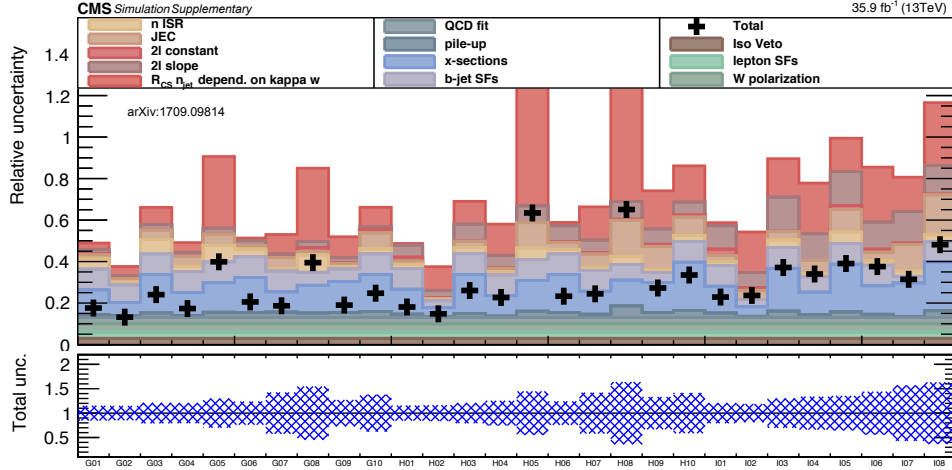


Figure 6.4: Visualization of all systematic uncertainties (upper panel) on the background prediction for mainband regions which are described in Tab. 4.4. The black crosses show an approximate total systematic uncertainty. They show the squared sum of all the different sources which are scaled according to their contribution. The total systematic uncertainty and statistical uncertainties are shown in the lower band.

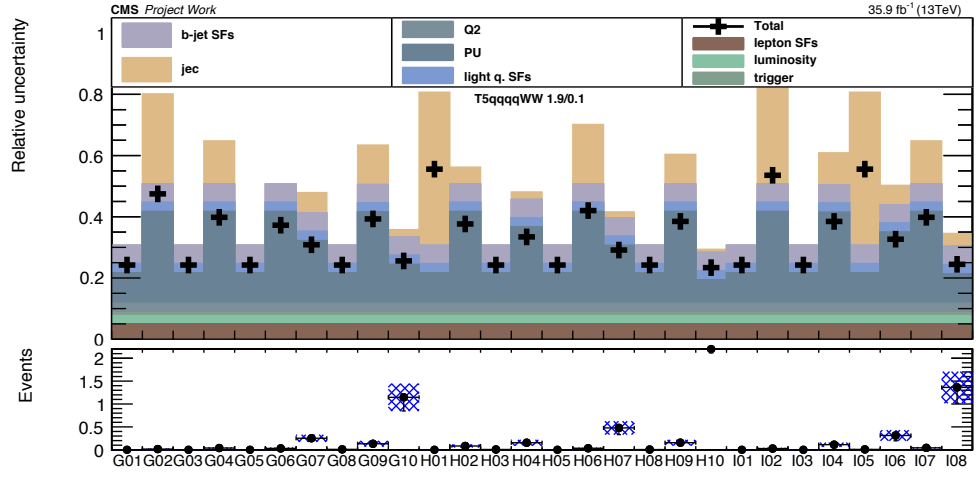


Figure 6.5: Visualization of all systematic uncertainties (upper panel) on the simulated signal yields in main band regions described in Tab. 4.4. Uncertainties are shown only for one mass point which is $m_{\tilde{g}} = 1900$ and $m_{\tilde{\chi}_1^0} = 100$. The lower panel shows the event counts with the total systematic and statistical uncertainty (blue shaded area) on the yields.

Chapter 7

Results and interpretation

In this chapter, the results are presented. The analysis was designed and validated before looking at the data in the SRs. Hence, the chapter commences with the results of the validity test (Sec. 7.1.1) of the data-driven background prediction in a restricted region spared from the W+jets sideband. It proceeds with the comparison between the background prediction and the observed event yields in the mainband signal regions (Sec. 7.1.2) and in the aggregated regions (Sec. 7.1.3). Following a summary of the statistical procedures, the interpretation of results in the framework of simplified SUSY models is presented. The chapter concludes with a discussion on the comparison of the results with complementary analyses.

7.1 Results of the background prediction

7.1.1 Result of the validation in sideband 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 on L_T and H_T are aligned with the mainband regions. Accordingly, the sideband where R_{CS} of W+jets events is measured is defined by $n_{\text{jet}}=3$ in order to avoid overlap. This restriction results in a statistical limitation and a slight reduced performance of the method. The sideband region for measuring $t\bar{t}$ + jets R_{CS} remains unchanged. Results of this validation are shown in Fig. 7.1. The systematic uncertainties shown in this figure are an approximate projection of what is observed in the mainband regions. Figure 7.1 manifests that the data agrees with the SM prediction in the validation regions.

7.1.2 Result of the background prediction in mainband regions

Figure 7.2 shows the observed event yields from data in signal regions for the 28 mainband bins, compared to the data-driven SM background prediction (see Ch. 5). In figure, the stacked colored histograms represent the prediction corresponding to the individual background contributions grouped into $t\bar{t}$ + jets, W+jets and EWK.

No significant deviation from the SM prediction is observed. A summary table of the background prediction and the final observed yields for each 28 bin can be found in Tab. 7.1.

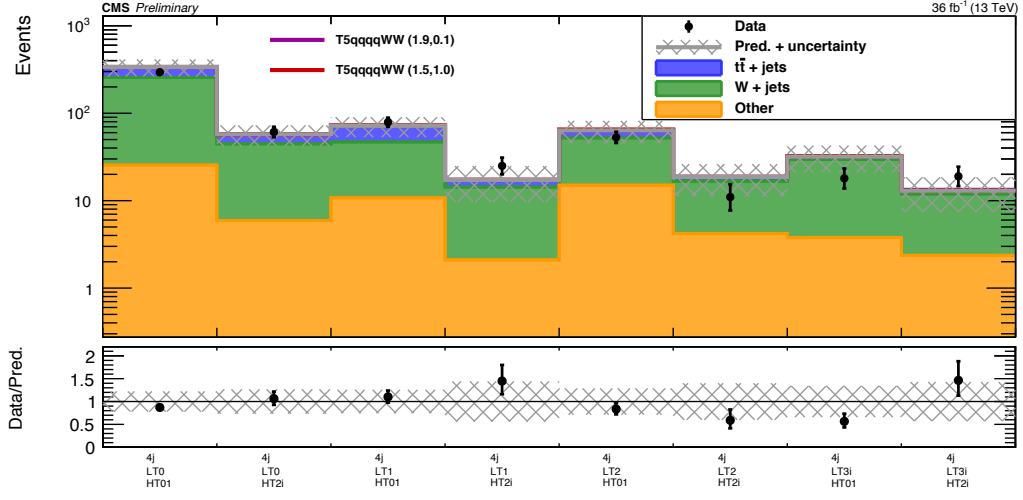


Figure 7.1: Validation of the background estimation method, using mainband regions with 4 jets. The shaded area reflects a rough estimate of the statistical and systematic uncertainties. The colored lines illustrate the expectations for two benchmark points of the T5qqqqWW model, showing the SUSY particle masses $m_{\tilde{g}}/m_{\tilde{\chi}^0}$ in TeV. The lower band shows the ratio between observed and predicted events.

The distributions of $\Delta\phi$ in data and simulation are shown in Fig. 7.3 for the four L_T intervals after the baseline selection with inclusive requirements on $n_{\text{jet}} \geq 5$ and $H_T \geq 500$ GeV. Overall, data and the SM 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 simulation based κ factors.

7.1.3 Result of background prediction in aggregated regions

As mentioned in Sec. 4.7.3, aggregate mainband signal regions have been defined to ease the reinterpretation of the results in terms future models. The comparison of the background predictions and the observed number of events in the SR of the aggregated SRs is presented in Tab. 7.2.

Table 7.1: Summary of the results in the main search regions. The $\Delta\phi$ value that is used to define the CR and SR is provided as well.

n_{jet}	L_T [GeV]	$\Delta\phi$ [rad]	H_T [GeV]	Bin name	Signal T5qqqqWW ($m_{\tilde{g}}, m_{\tilde{\chi}_1^0}$) [TeV]		Predicted background	Observed
5	[250,350]	1.0	[500,750]	G01	1.82 ± 0.29	< 0.01	101.91 ± 47.55	111
			≥ 750	G02	0.21 ± 0.09	0.01 ± 0.01	76.73 ± 16.19	100
	[350,450]	1.0	[500,750]	G03	2.25 ± 0.32	< 0.01	24.43 ± 14.78	25
			≥ 750	G04	0.29 ± 0.11	0.04 ± 0.01	22.78 ± 8.29	22
	[450,650]	0.75	[500,750]	G05	3.02 ± 0.37	< 0.01	14.46 ± 6.5	17
			[750,1250]	G06	1.4 ± 0.25	0.04 ± 0.02	12.13 ± 4.68	10
			≥ 1250	G07	0.08 ± 0.06	0.25 ± 0.04	4.15 ± 1.72	2
	≥ 650	0.5	[500,750]	G08	0.74 ± 0.18	0.01 ± 0.01	2.32 ± 1.49	5
			[750,1250]	G09	0.49 ± 0.15	0.12 ± 0.03	5.79 ± 1.96	6
			≥ 1250	G10	0.14 ± 0.07	1.15 ± 0.08	2.74 ± 1.26	0
[6,7]	[250,350]	1.0	[500,1000]	H01	3.02 ± 0.36	< 0.01	89.32 ± 38.21	85
			≥ 1000	H02	0.31 ± 0.1	0.09 ± 0.02	30.94 ± 5.08	33
	[350,450]	1.0	[500,1000]	H03	4.13 ± 0.41	0.01 ± 0.01	18.91 ± 10.89	31
			≥ 1000	H04	0.52 ± 0.14	0.14 ± 0.03	9.51 ± 2.34	8
	[450,650]	0.75	[500,750]	H05	3.63 ± 0.39	< 0.01	5.71 ± 3.31	13
			[750,1250]	H06	3.79 ± 0.39	0.03 ± 0.01	8.21 ± 3.15	8
			≥ 1250	H07	0.36 ± 0.12	0.47 ± 0.05	3.61 ± 1.78	4
	≥ 650	0.5	[500,750]	H08	0.89 ± 0.19	< 0.01	0.79 ± 0.53	3
			[750,1250]	H09	1.77 ± 0.26	0.15 ± 0.03	3.63 ± 1.37	5
			≥ 1250	H10	0.83 ± 0.18	2.83 ± 0.12	1.83 ± 0.86	1
≥ 8	[250,350]	1.0	[500,1000]	I01	0.88 ± 0.18	< 0.01	6.96 ± 2.83	16
			≥ 1000	I02	0.26 ± 0.09	0.03 ± 0.01	6.32 ± 1.17	4
	[350,450]	1.0	[500,1000]	I03	0.55 ± 0.14	< 0.01	1.67 ± 0.77	3
			≥ 1000	I04	0.72 ± 0.15	0.11 ± 0.02	2.65 ± 0.89	4
	[450,650]	0.75	[500,1250]	I05	2.07 ± 0.26	0.01 ± 0.01	0.63 ± 0.32	0
			≥ 1250	I06	0.45 ± 0.12	0.3 ± 0.04	0.68 ± 0.35	1
	≥ 650	0.5	[500,1250]	I06	0.97 ± 0.18	0.04 ± 0.01	0.27 ± 0.23	1
			≥ 1250	I07	1.12 ± 0.18	1.37 ± 0.08	0.38 ± 0.24	1

Table 7.2: Numbers of expected background events with combined statistical and systematic uncertainty and the observed numbers of events in aggregated SRs. The expected number of signal events for the two corresponding benchmark signals are given as well.

n_{jet}	L_T [GeV]	$\Delta\phi$ [rad]	H_T [GeV]	Pred. background	Observed data	T5qqqqWW (1.5, 1.0)	T5qqqqWW (1.9, 0.1)
≥ 5	≥ 650	0.5	≥ 750	18.4 ± 5.1	14	6.2±0.6	6.3±0.2
≥ 6	≥ 450	0.75	≥ 500	28.8± 6.8	37	16.6±0.9	5.3±0.2
≥ 6	≥ 650	0.5	≥ 1000	5.1± 1.8	4	4.0±0.5	5.0 ±0.2
≥ 7	≥ 450	0.75	≥ 500	9.7± 2.5	11	9.5±0.7	3.5±0.2
≥ 7	≥ 650	0.5	≥ 500	3.8± 1.2	4	4.3±0.5	3.3±0.2
≥ 8	≥ 250	1.0	≥ 1250	7.2± 1.9	8	1.8±0.3	1.7±0.1

7.2 Statistical interpretation

No significant excess is observed, and we proceed to upper limits on signal cross sections in the $m_{\tilde{g}}-m_{\tilde{\chi}^0}$ plane. In this section, the statistical framework used to derive limits is presented. The method includes not only the mainband signal regions (MB SR) but also sideband and control regions. For this purpose, a mechanism based on a statistical ABCD method where each letter is representing MB SR, MB CR, SB SR and SB CR respectively, is

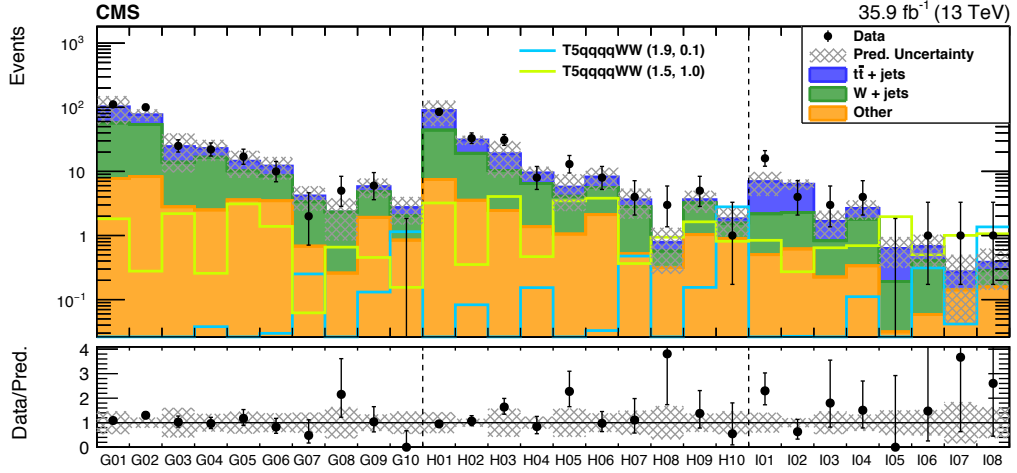


Figure 7.2: Observed data and predicted event yields (from data) in the 28 search regions are shown. The black points with error bars show the number of observed data events and corresponding statistical error. The filled, stacked histograms represent the predictions for $t\bar{t}$ + jets, W+jets, and 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.

used. In this way, potential signal contamination in sideband and control regions is taken into account. The limit procedure makes simultaneous use of two ABCD methods, called thus ABCDEF, since each of the $t\bar{t}$ + jets and W+jets backgrounds is predicted with two parallel but separate R_{CS} methods. A summary of these regions can be found in Tab. 7.3.

Table 7.3: 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 ABCD method is shown in Eq. 7.1, where QCD estimation is subtracted from CRs and possible residual differences between mainbands and sidebands are corrected with κ .

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

The formula representing the ABCDEF method is written as follows:

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

A tool that was developed for the search for the Higgs boson was used. This tool implements the statistical methods described in [139]. The likelihood model has been extended

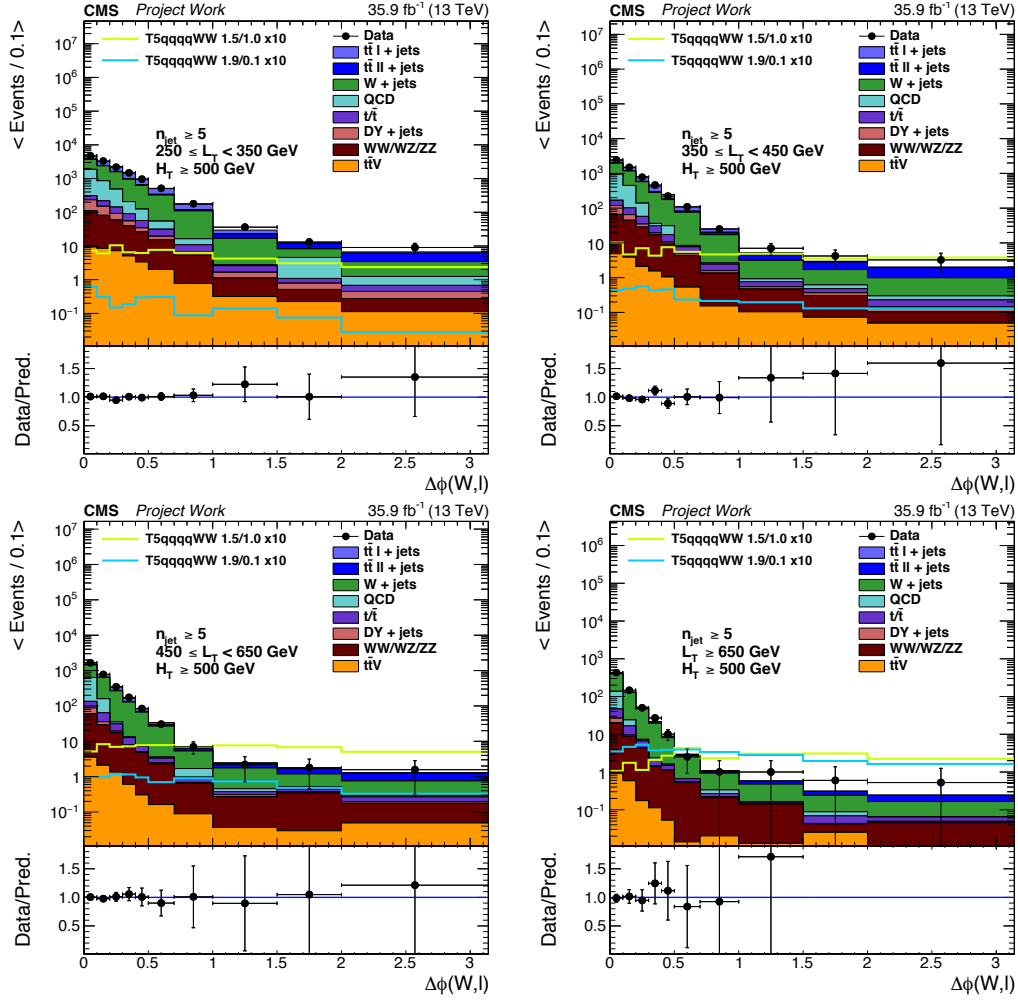


Figure 7.3: $\Delta\phi$ distributions in data which is represented with black points and simulation which is represented with color filled areas. The figure shows the four L_T interval after the baseline selection with inclusive requirements on $n_{\text{jet}} \geq 5$ and $H_T \geq 500$ GeV. The simulated signal events are scaled by 10.

to incorporate the ABCDEF regions and, specifically, the constraint described in Eq. 7.2. 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.

Throughout this section, the expected event yields of SUSY model will be denoted as s_i , the SM background as b_i where index "i" indicates corresponding MB. 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 the statistical treatment, all uncertainty sources are assumed to be either 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 incorporates the ABCDEF regions for each mainband, 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 regions (A)BCDEF involves the Poisson p.d.f.s for the observed number of events given the (estimated) simulated sum of signal and background, and the second factor refers to the nuisance parameters. Both signal and background predictions are subject to multiple uncertainties. Systematic uncertainties which are discussed in Ch. 6 and κ factors are included as constraints on the region A, while the QCD prediction is considered as constraints on the control regions B and F. Additionally, the nuisance parameters include the effect of the statistical uncertainty of all 6 bins.

The entire information is inserted to the limit tool using text files. An example text file is shown in Fig. 7.4 for only one MB and moreover the entries corresponding to rare backgrounds and QCD have not been shown, for simplicity. Example tests on the nuisance parameters can be found in the Appendix C using the text file corresponding to the high mass gap benchmark point.

Test statistic, $\tilde{q}_\mu$

The compatibility of the data with the background-only and signal+background hypotheses, where the signal is allowed to be scaled by a factor μ is measured by constructing the following test statistic based on the profile likelihood ratio [140]:

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

where nominator and denominator are maximized 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 normalization term only. The condition $0 \leq \hat{\mu} \leq \mu$ ensures one-sided confidence intervals for upper limit tests while nonphysical negative signals are avoided.

According to the Neyman-Pearson lemma [141], the ratio of likelihoods provide the most powerful discrimination. However, this statement does not have to be true for a system with multiple free parameters, in fact the existence of a uniformly most powerful test statistic is not certain. On the other hand, the likelihood ratio has proven reliable and is 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 Eq. 7.3 are computed. The case of $\mu = 0$ corresponds to the background-only hypothesis while the non-zero μ values are representing the signal+background hypothesis.

Pdfs $\tilde{q}_\mu$

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 in principle required. This procedure requires to scan a range of μ values and subsequently generate a large number of pseudo-experiments. However, in the present complex case, the pseudo-MC mechanism is prohibitively computationally expensive. Therefore, asymptotic formulae [133] are used to circumvent this computational challenge.

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 Eq. 7.6, directly from the observed test systematic as:

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

This is possible, because the profile likelihood test statistic is distributed as a χ^2 distribution with one degree of freedom. In particle physics, to claim "discovery", the significance must exceed $Z=5$, which is corresponding to a one-sided p-value of the back-ground only hypothesis of $2.87 \cdot 10^{-7}$. The observed significance, which is obtained by combining the main search regions, for the signal model T5qqqqWW is presented in Fig. 7.5. It can be seen that for all the mass points the local significance is less than 3.

Upper limits

Upper limits are calculated at the 95% confidence level(CL) using the asymptotic CL_s criterion [138], 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, $P(q_0 \geq q_{0, \text{obs}}|\mu)$, while CL_b can be constructed using the background only hypothesis, $P(q_0 \geq q_{0, \text{obs}}|\mu = 0)$. The 95% CL exclusion limits are defined as $CL_s(\mu) \leq 5\%$.

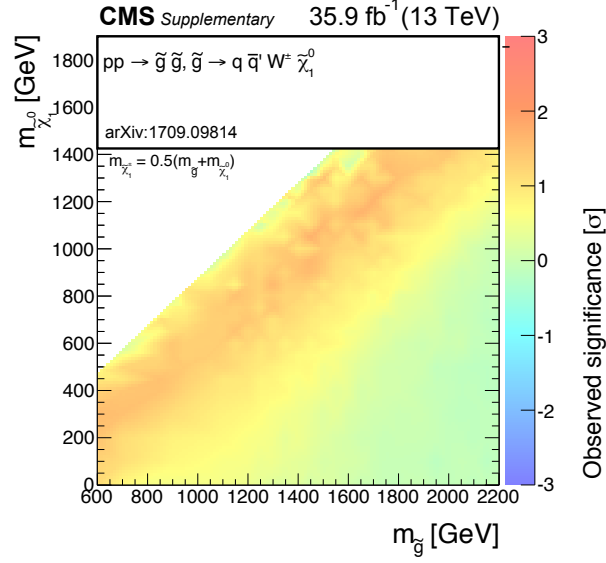


Figure 7.5: Observed significance for the model T5qqqqWW.

Given the fact that the 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 scenario 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 are derived in the mass plane. The procedure introduced in Sec. 7.2 is followed to obtain 95% CL upper limits on the observed and the median expected signal strength. The corresponding results are presented in Fig. 7.6 for the entire mass scan of the T5qqqqWW. The values for the missing mass points, that can be seen in Fig. 4.3, are 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. 7.7, the color map shows the observed upper limit on the theoretical cross section, the black lines show 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 decay products typically have high momenta. Therefore, kinematic regions with high hadronic and leptonic scales provide the largest signal sensitivity. In this region, the 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, where the gluino has very soft decay products resulting in a decrease in acceptance, the upper limit on the cross section takes values up to 1 pb. The signal models with neutralino mass up to 950 GeV are 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 $n_{\text{jet}} \leq 8$, $450 \leq L_T < 650$ GeV, $500 \leq H_T < 1250$ 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 $6 \leq n_{\text{jet}} \leq 7$, $450 \leq L_T < 650$ GeV, $500 \leq H_T < 750$ GeV has 13 events observed while 5.7 ± 3.3 expected.

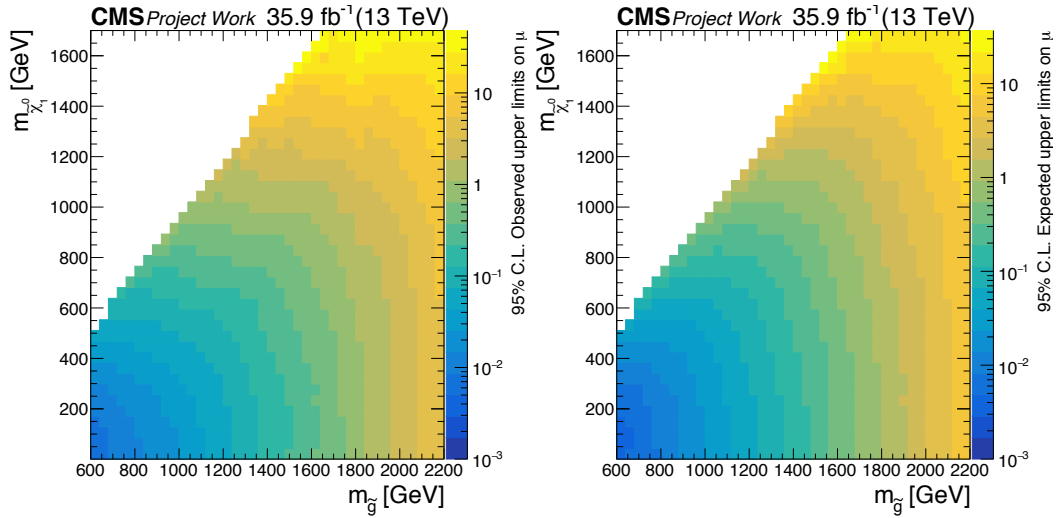


Figure 7.6: 95% CL Observed (Expected) upper limits on the signal strength modifier μ for the T5qqqqWW model is shown at the left (right) side as a function of the gluino and neutralino masses.

7.4 Comparison to complementary results

As discussed 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 were excluded [67]. The searches are repeated at the new center-of-mass energy and the first results were published with integrated luminosity of 2.3 fb^{-1} [68]. In the present analysis, the exclusion on the gluino mass is extended by 500 GeV for the lowest neutralino mass.

This result is compared with complementary results from Run 2 analyses of the CMS collaboration using 35.9 fb^{-1} in Fig. 7.9 (left). The present analysis, which is labeled with SUS-16-042, is represented with the purple curve and published in [134]. The analysis is particularly competitive in the high gluino-neutralino mass region. The fully hadronic analysis SUS-16-033 [135], is similarly competitive, however, as there is no lepton requirement, it selects a different portion of the signal and the background composition and the systematic uncertainties are very different. Covering all the possible final states (without restricting our selves to a specific signal model) provides a robust way to search for what

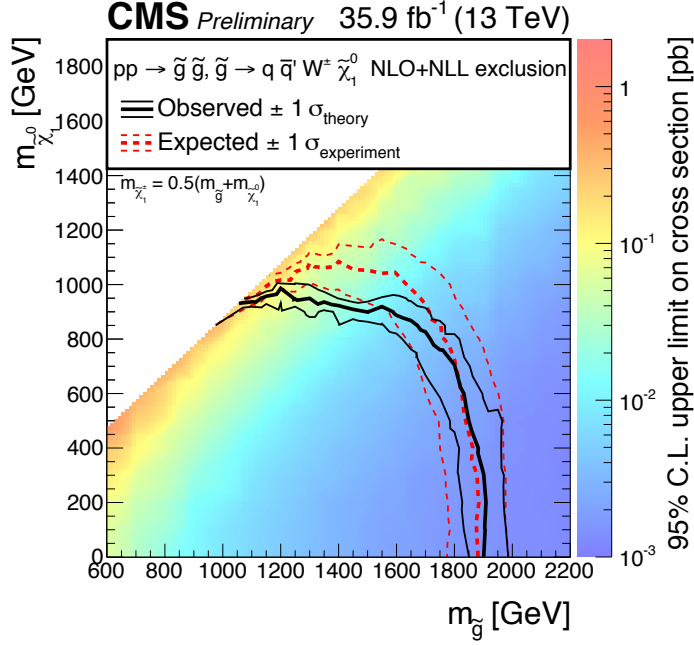


Figure 7.7: Cross section limits at a 95% CL for the T5qqqqWW model, and as a function of the gluino and LSP masses. The solid black (dashed red) lines correspond to the observed (expected) mass limits, with the thicker lines representing the central values and the thinner lines representing the $\pm 1\sigma$ uncertainty bands related to the theoretical (experimental) uncertainties

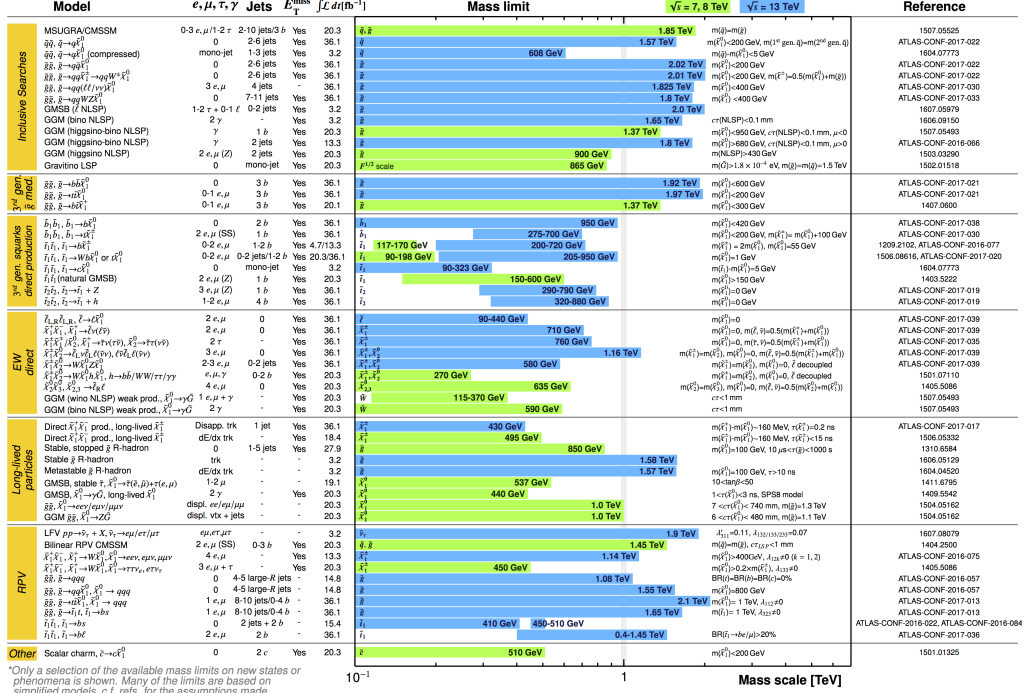
is beyond the SM.

An overall summary of the SUSY results obtained by the CMS collaboration can be found in Fig. 7.8 (bottom). It shows the best exclusion limits on sparticle masses for each search channel. Many analyses, which set limits on gluino mass, reach the TeV scale.

The ATLAS collaboration also performed similar searches for gluino pair production. The results for the T5qqqqWW topology can be seen in Fig. 7.9 (right) where the gray shaded area shows the Run 1 results. Furthermore, a summary of all SUSY models studied in ATLAS is shown in Fig. 7.8 (top).

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

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



Selected CMS SUSY Results* - SMS Interpretation

ICHEP '16 - Moriond '17

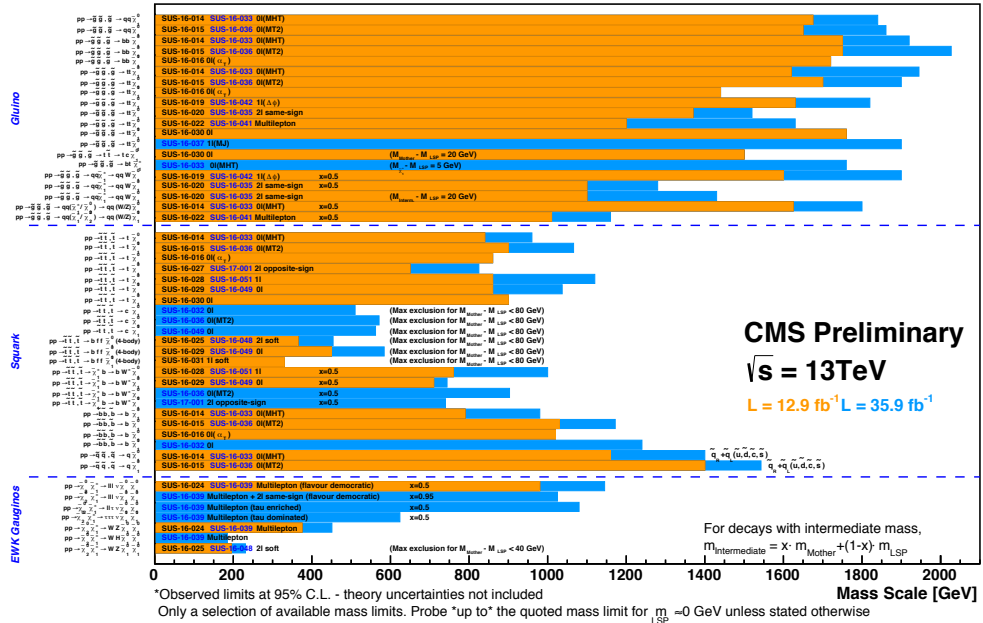


Figure 7.8: Best exclusion limits on sparticle masses from searches for SUSY using SMS by ATLAS (top) and CMS (bottom) collaborations [136, 137]. In the upper plot, the green bars represent the 7-8 TeV results while the blue bars are summarizing the 13 TeV results. In the lower plot, results summarizing the 13 TeV results, the orange bars represent the 12.9 fb^{-1} results while the blue bars show the 15.9 fb^{-1} results.

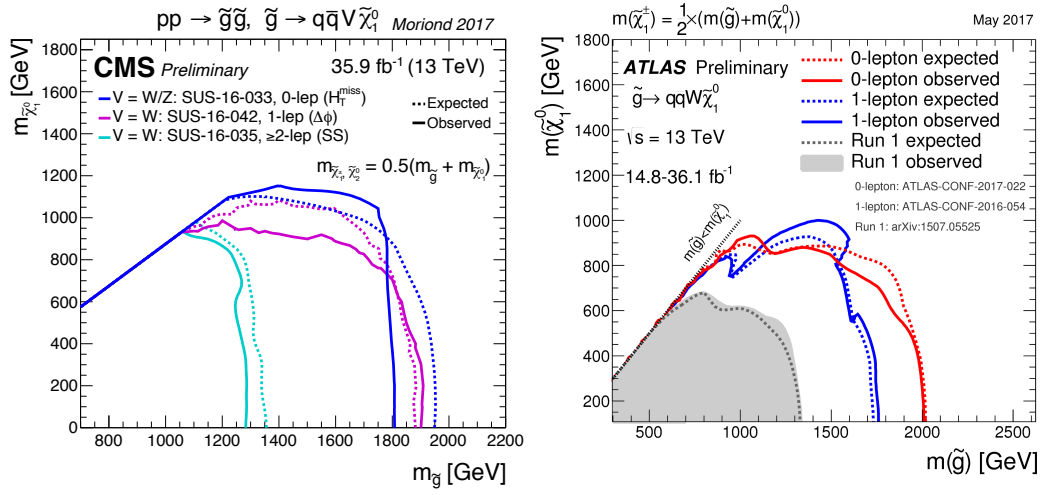


Figure 7.9: 95% CL exclusion limit curves on the simplified SUSY model of gluino pair production with subsequent decay of $\tilde{g} \rightarrow q\bar{q}W\tilde{\chi}_1^0$ obtained by the CMS (left) and ATLAS (right) collaborations. The present analysis is corresponding to the result labeled with SUS-16-042.

Conclusion

The Standard Model of particle physics continues to be a successful theory in explaining the phenomena observed in experiments, such as those at the 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 can the quantum Higgs mass corrections be stabilized?

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

In this thesis, a generic search for SUSY is performed in the single lepton final state with multiple jets and a veto on b-tagged jets. The sensitivity of the search is shown with a simplified model of SUSY, where 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 two neutralinos and the neutrino coming from the leptonic W boson decay combine to a large $\cancel{E}_T$. Moreover, the high multiplicity of predicted final state quarks leads to large hadronic activity. Therefore, the search has high sensitivity to the model in the tails of the kinematical distributions.

The main search variable of the analysis is the azimuthal angle between the lepton and the 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 the LSPs. 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. The method uses control regions to estimate the normalization in signal regions and low jet multiplicity sidebands to obtain the signal to control region transfer factors, R_{CS} . The main SM backgrounds from W+jets and $t\bar{t}$ +jets production are predicted separately, and a further separate estimation method is developed for QCD multijet events. This method uses the polarization information of leptons from W boson decays which manifests itself in the L_p distributions. The predicted QCD multijet event yields 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 estimated. 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 to the result. 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 results of the data-based prediction is compared to data. The data is 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 at 95% CL on the gluino pair production cross section are obtained in the context of the investigated simplified model, where the intermediate chargino mass in the decay chain is taken to be half way between the gluino and the neutralino. As a result, gluino masses below 1.9 TeV are excluded for neutralino masses below 300 GeV. This corresponds to an improvement of about 500 GeV on the gluino mass over the previous result [68]. This is consistent with complementary analyses in e.g. the fully hadronic search channels.

After the analysis of 35.9 fb^{-1} of the 13 TeV run, data still do not favor models with SUSY. While the simplified models used in this work assume 100% branching ratio in the gluino decay, it is clear that a general model as e.g. the MSSM is also tightly constrained. As the limits on the gluino mass approach 2 TeV, natural SUSY as a solution to hierarchy problem is disfavoured. Given the variety of possibilities, excluding SUSY on first principles is, however, difficult. In the future, this analysis will be repeated with the full Run 2 dataset, currently estimated to correspond to 150 fb^{-1} . Meanwhile, SUSY searches can be extended towards long-lived particles with displaced vertex analyses, and there is also still room for natural SUSY with compressed mass configurations.

Although no significant deviation from the Standard Model has been observed, the journey of particle physics is not yet finished. 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.

When viewed from this perspective, the results of the present work may look like a small step. However, even if small, it has increased our knowledge about nature, and, moreover, it is a step that is firmly rooted in data. As such, it will prevail.

Appendix A

MC samples and cross sections

Sample	cross section [pb]
/TTjets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*(3*0.108)*(1-3*0.108)
/TTjets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*(3*0.108)*(1-3*0.108)
/TTjets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76*((3*0.108)**2)
/TTjets_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.610*831.76/502.2
/TTjets_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.663*831.76/502.2
/TTjets_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.12*831.76/502.2
/TTjets_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.001430*831.76/502.2

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

Sample	cross section [pb]
/WjetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1345*1.21
/WjetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	359.7*1.21
/WjetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	48.91*1.21
/WjetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	12.05*1.21
/WjetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.501*1.21
/WjetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.329*1.21
/WjetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.03216*1.21

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

Sample	cross section [pb]
/ST.s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	$(7.20 \pm 4.16) \times 10^{-3}$
/ST.t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	136.02
/ST.t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	80.95
/ST.tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
/ST.tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.55
/DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$147.4^{+1.23}$
/DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$40.99^{+1.23}$
/DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$5.678^{+1.23}$
/DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$1.367^{+1.23}$
/DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$0.6304^{+1.23}$
/DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$0.1514^{+1.23}$
/DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$0.003565^{+1.23}$
/QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	351300
/QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	31630
/QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6802
/QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1206
/QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	120.4
/QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	25.25
/WWTo2L2Nu_13TeV-powheg	10.481
/WWToLNUQQ_13TeV-powheg	43.53
/ZZTo2L2Nu_13TeV-powheg_pythia8	0.564
/ZZTo2L2Q_13TeV_amcatnloFXFX-madspin-pythia8	3.28
/ZZTo2Q2Nu_13TeV_amcatnloFXFX-madspin-pythia8	4.04
/ZZTo2L2Nu_13TeV-powheg_pythia8	0.564
/WZTo1L3Nu_13TeV_amcatnloFXFX-madspin-pythia8	$(47.13) \times (3 \times 0.108) \times (0.2)$
/WZTo1L1Nu2Q_13TeV_amcatnloFXFX-madspin-pythia8	10.71
/WZTo2L2Q_13TeV_amcatnloFXFX-madspin-pythia8	5.60
/WWTo2L2Nu_DoubleScattering_13TeV-pythia8	1.64
/TTWJetsToLNU_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.40620
/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297
/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
/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 Figs. 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 +jets background. Figures B.3 and Figs.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 $t\bar{t}$ +jets background. Distributions in the mainband are shown in Figs. B.5 and Figs. B.6.

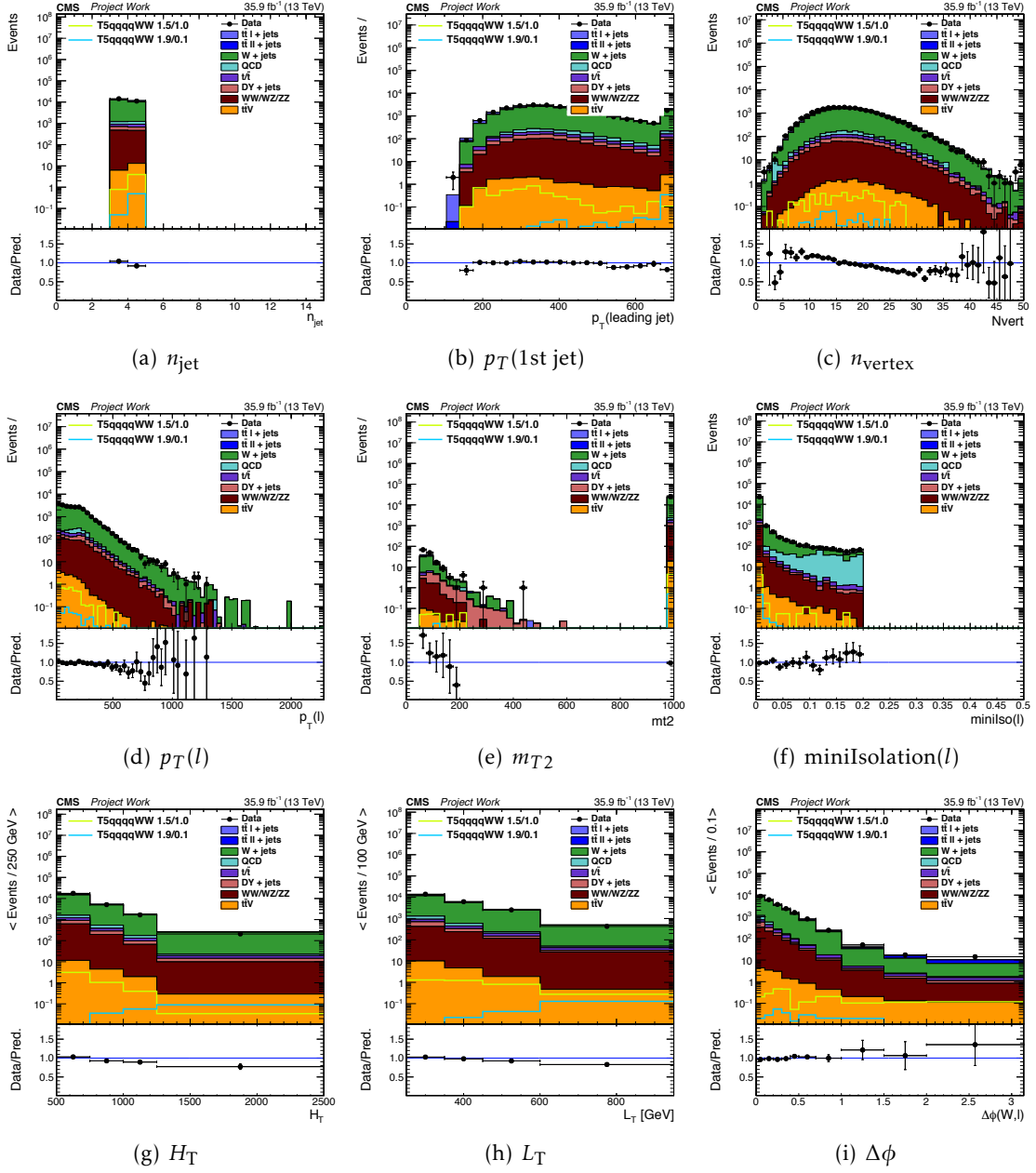


Figure B.1: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $3 \leq \text{jets} \leq 4$ and zero b-tagged jets (1 μ channel).

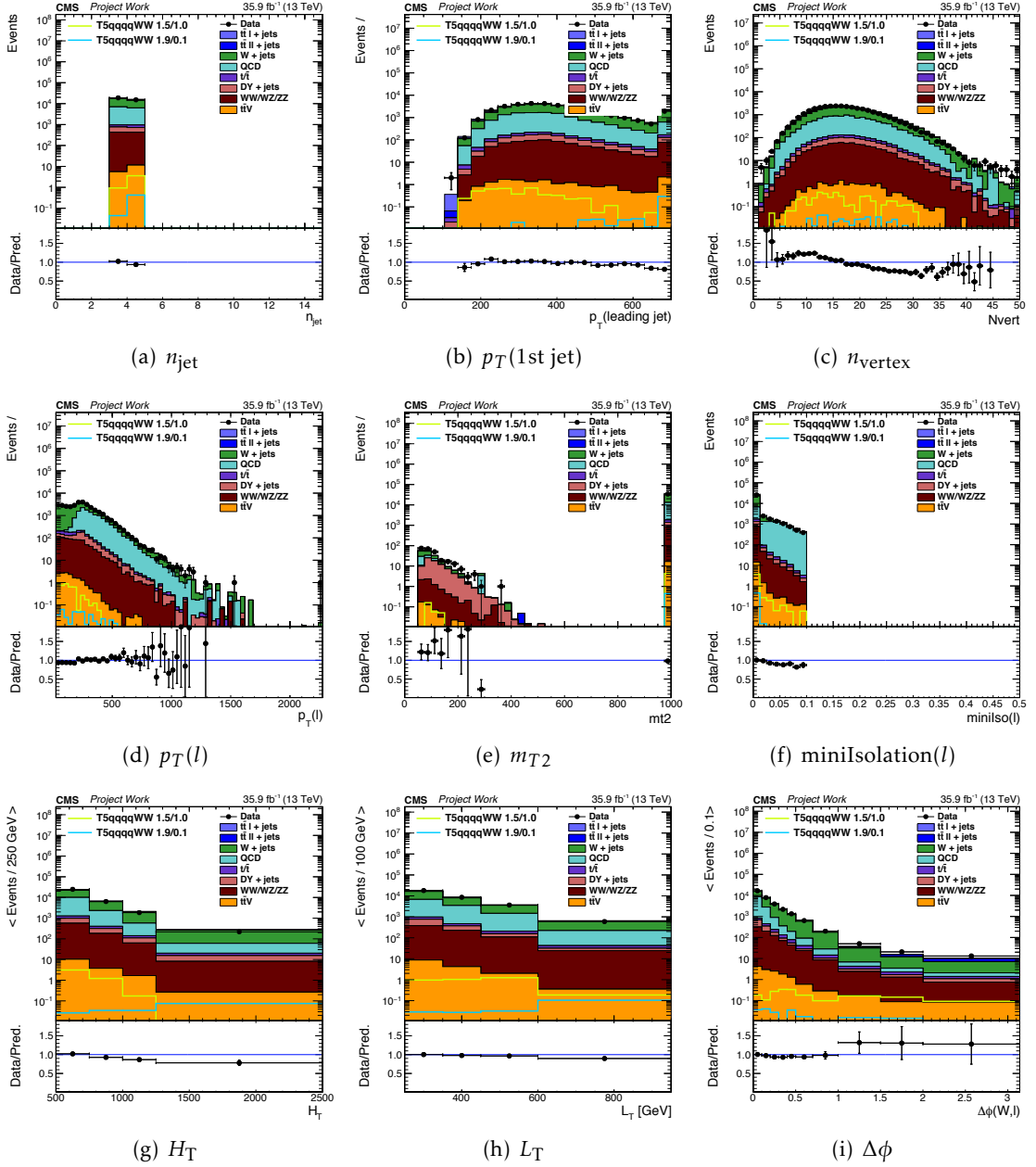


Figure B.2: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $3 \leq \text{jets} \leq 4$ and zero b-tagged jets (1 e channel).

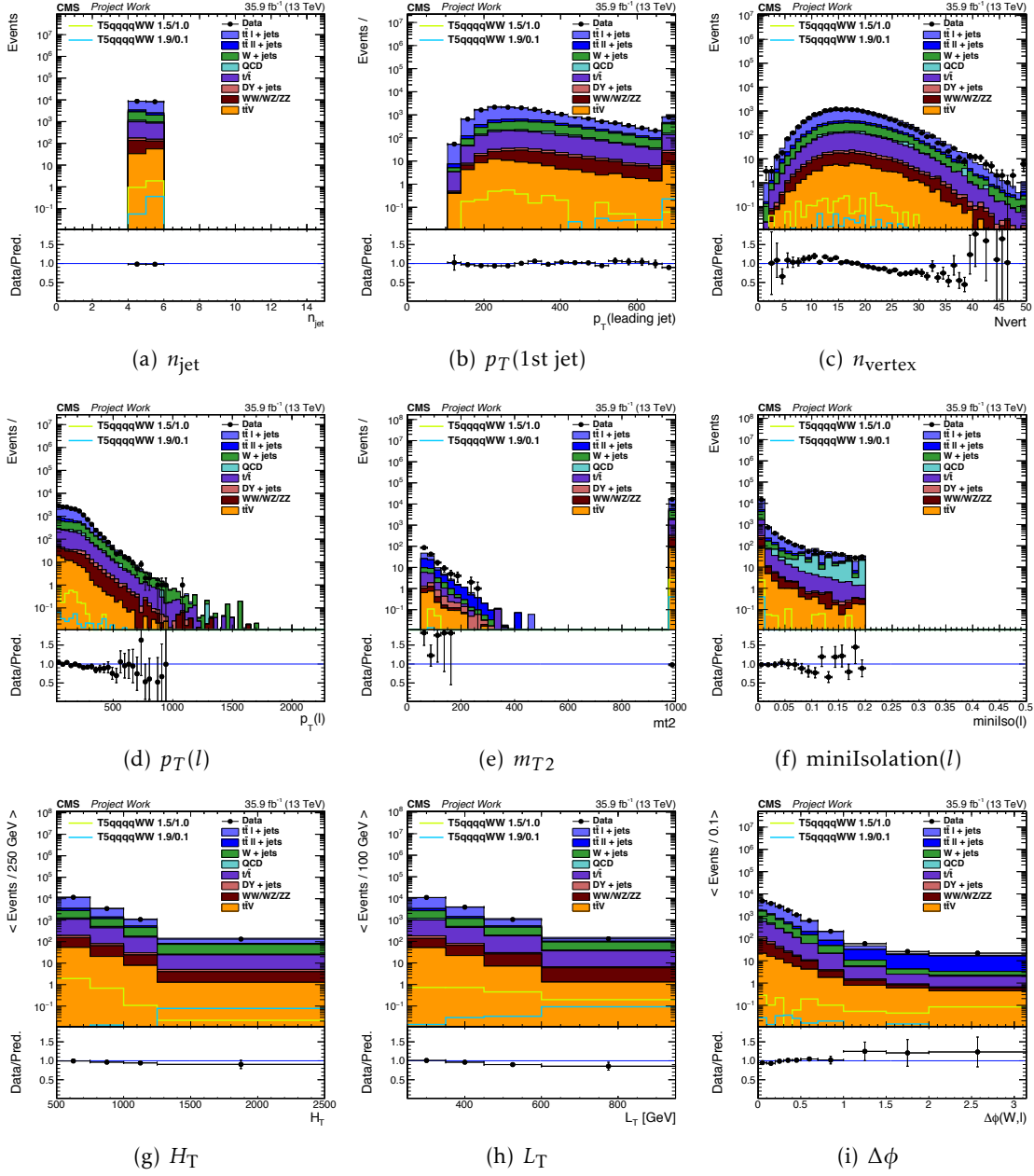


Figure B.3: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $4 \leq \text{jets} \leq 5$ and b-tagged jets (1 μ channel).

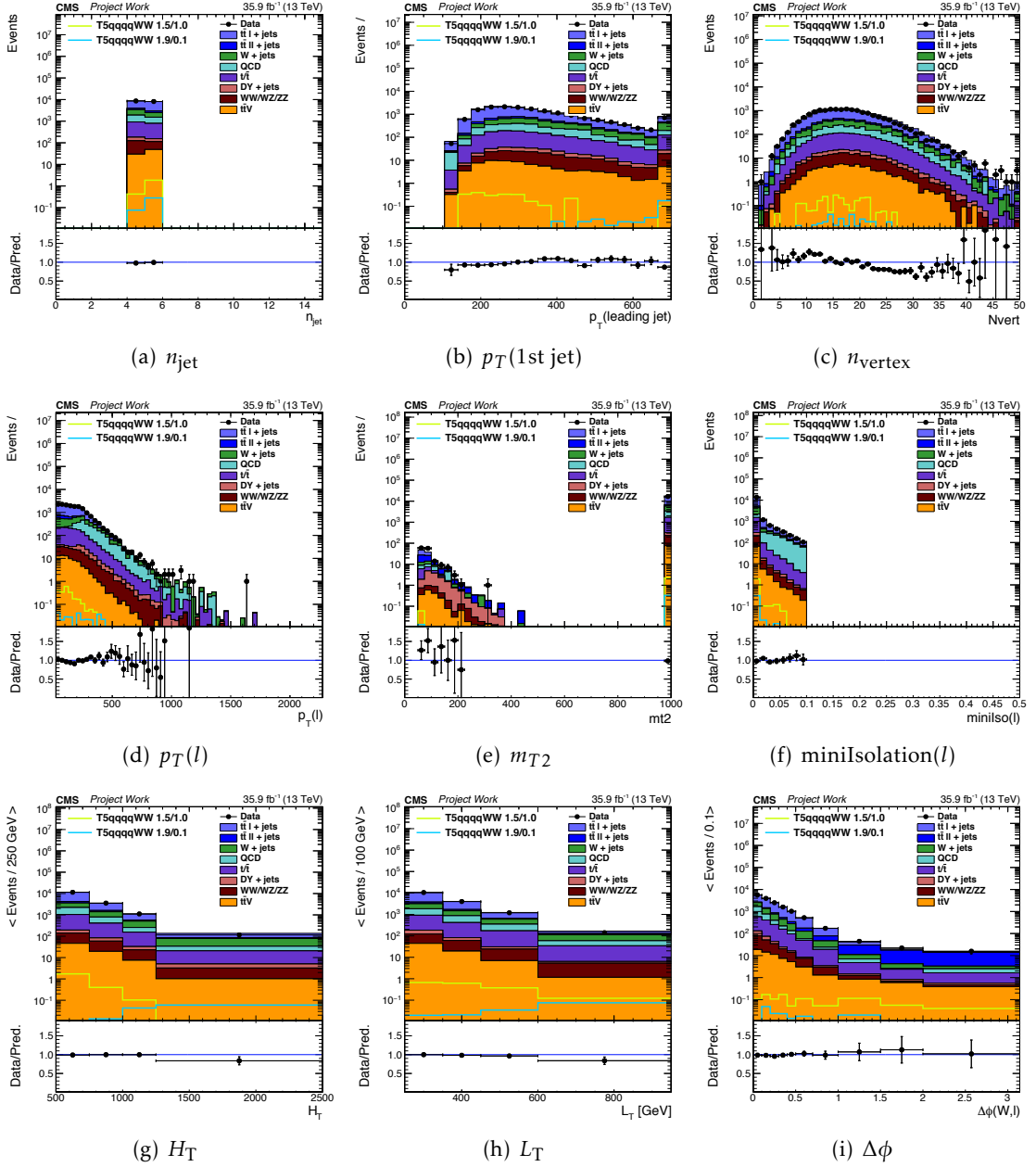


Figure B.4: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, $4 \leq \text{jets} \leq 5$ and b-tagged jets (1 e channel).

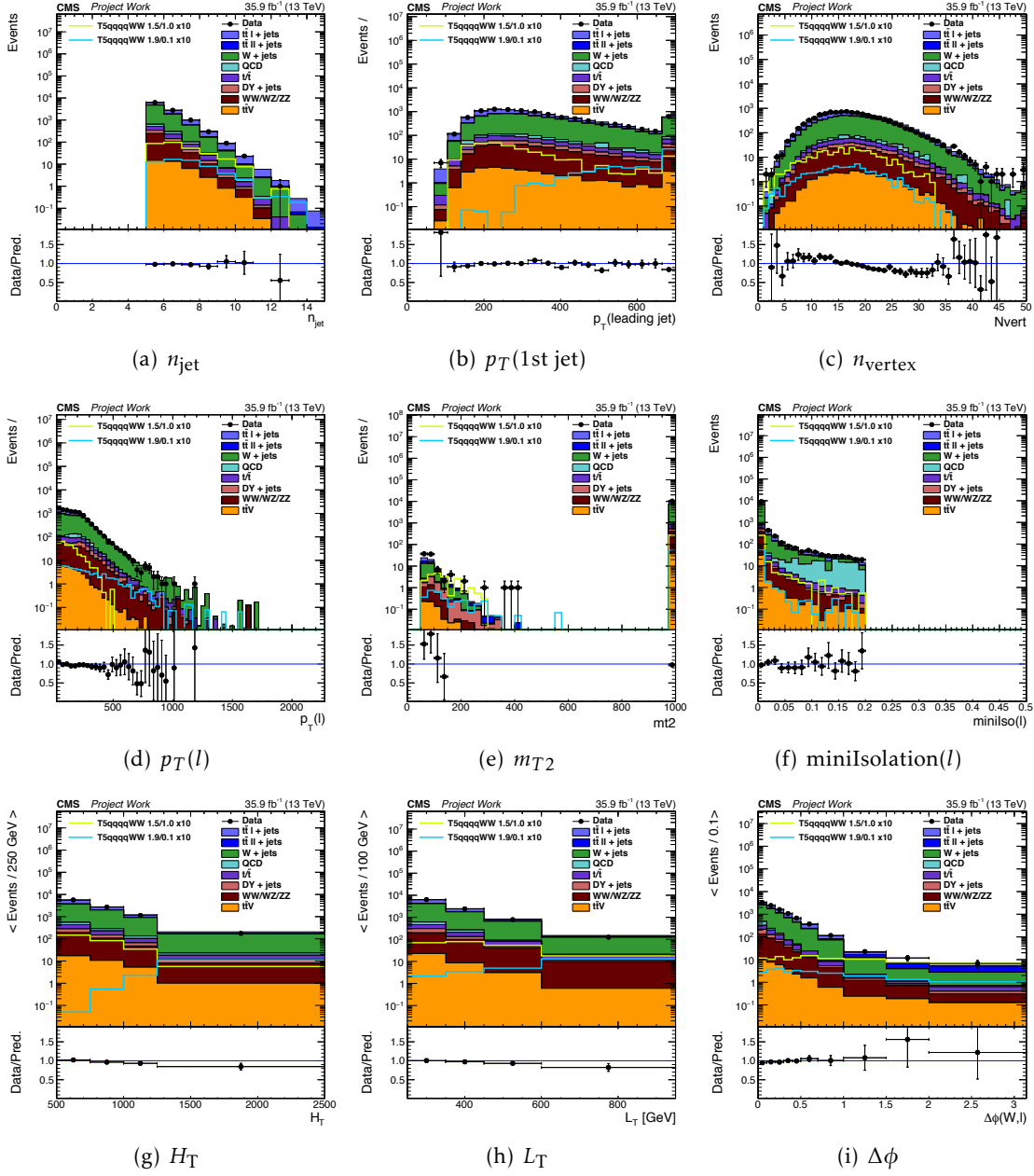


Figure B.5: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, ≥ 5 jets and zero b-tagged jets (1 μ channel).

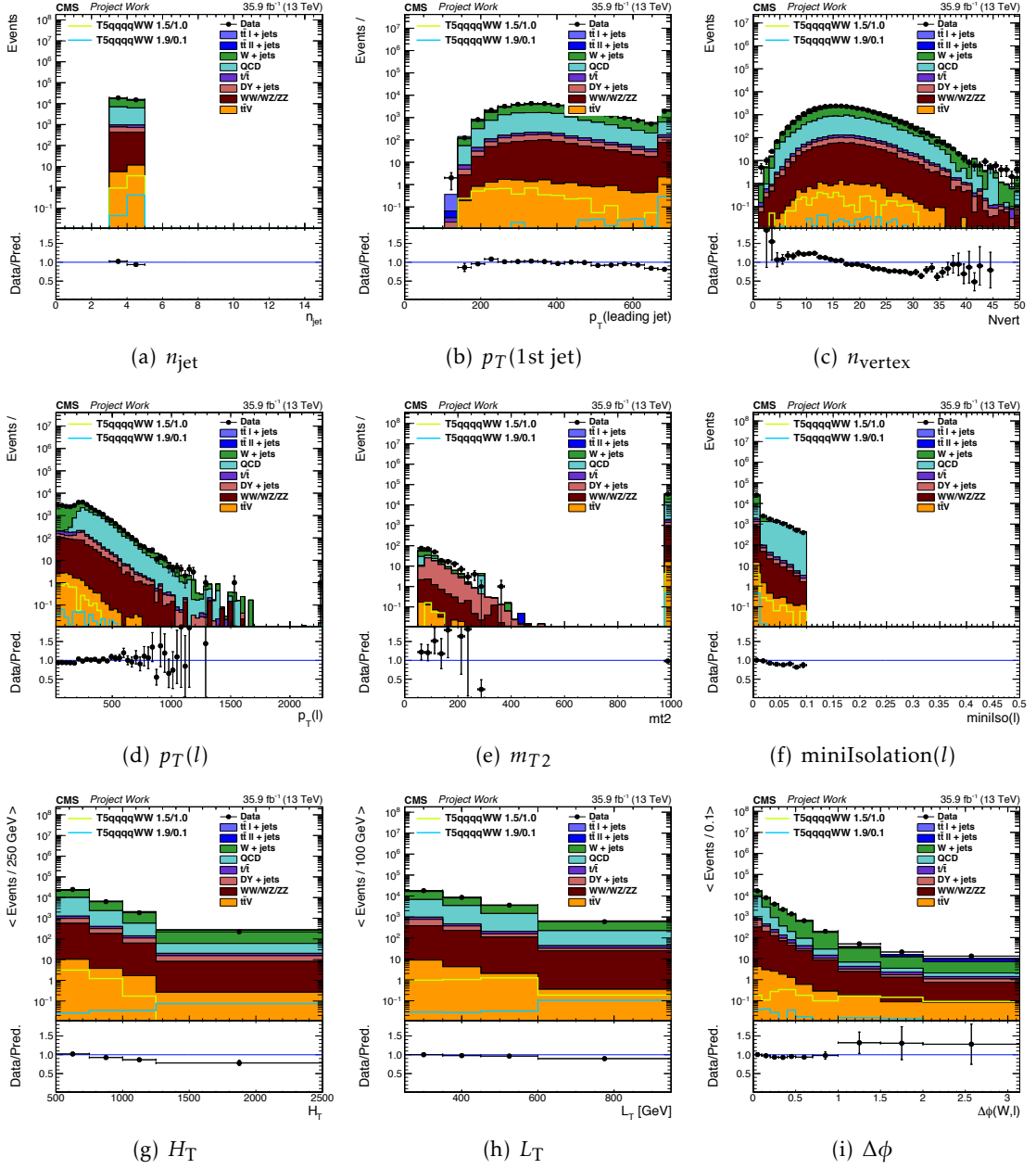
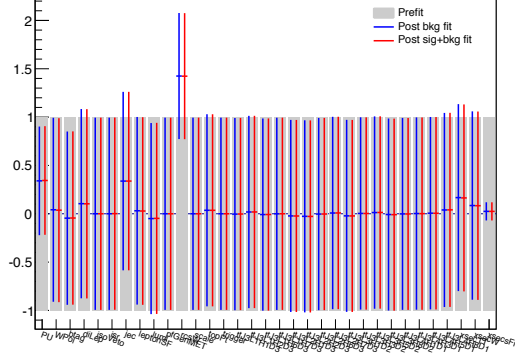


Figure B.6: Distribution of kinematic observables after requiring $H_T > 500$ GeV, $L_T > 250$ GeV, ≥ 5 jets and zero b-tagged jets (1 e channel).

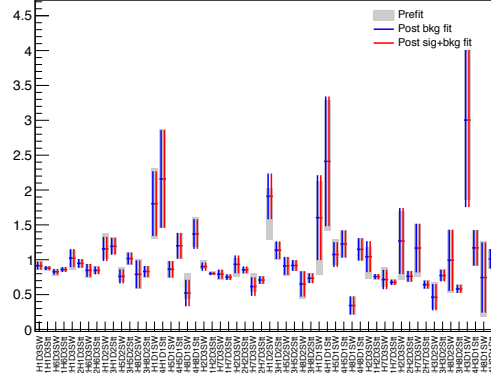
Appendix C

Statistical tests

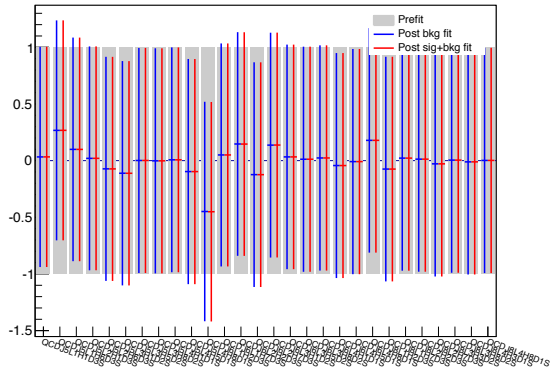
In this appendix, the output of the maximum likelihood fit to the data, in particular the effect of the different nuisance parameters is investigated. The tests are performed with the datacard of the mass point T5qqqqWW 1.9/0.1 [TeV].



(a) lnN constraints on region A

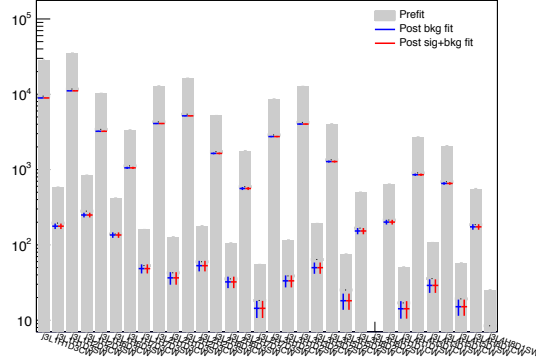


(b) κ parameters

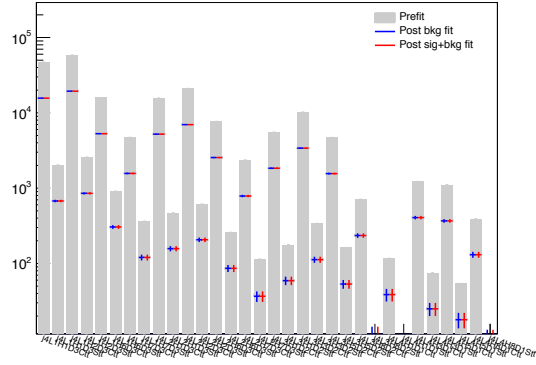


(c) QCD estimate

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.

Bibliography

- [1] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* **22** (1961) 579–588, doi:10.1016/0029-5582(61)90469-2.
- [2] A. Salam, “Weak and Electromagnetic Interactions”, *Conf. Proc. C* **680519**, 367 (1968).
- [3] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.
- [4] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.
- [5] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [6] CDF Collaboration, “Observation of top quark production in $\bar{p}p$ collisions”, *Phys. Rev. Lett.* **74** (1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.
- [7] DØ Collaboration, “Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV”, *Phys. Rev. Lett.* **74** (1995) 2422, doi:10.1103/PhysRevLett.74.2422, arXiv:hep-ex/9411001.
- [8] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [9] 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. B* **716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [10] ATLAS and CMS Collaboration, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments” *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.
- [11] Planck Collaboration, “Planck 2015 results. XIII. Cosmological parameters”, *Astron. Astrophys. A* **13** (2016), doi:10.1051/0004-6361/201525830, arXiv:1502.01589.
- [12] N. W. Boggess et al., “The COBE Mission: Its Design and Performance Two Years after the launch”, *The Astrophysical Journal*. **397** (1992), doi:10.1086/171797.

- [13] 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.
- [14] P. Ramond, "Dual theory for free fermions", *Phys. Rev. D* **3** (1971) 2415, doi:10.1103/PhysRevD.3.2415.
- [15] Yu. A. Golfand and E. P. Likhtman, "Extension of the algebra of Poincaré group generators and violation of P invariance", *JETP Lett.* **13** (1971) 323.
- [16] A. Neveu and J. H. Schwarz, "Factorizable dual model of pions", *Nucl. Phys. B* **31** (1971) 86, doi:10.1016/0550-3213(71)90448-2.
- [17] D. V. Volkov and V. P. Akulov, "Possible universal neutrino interaction", *JETP Lett.* **16** (1972) 438.
- [18] J. Wess and B. Zumino, "A Lagrangian model invariant under supergauge transformations", *Phys. Lett. B* **49** (1974) 52, doi:10.1016/0370-2693(74)90578-4.
- [19] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys. B* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1.
- [20] P. Fayet, "Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino", *Nucl. Phys. B* **90** (1975) 104, doi:10.1016/0550-3213(75)90636-7.
- [21] H. P. Nilles, "Supersymmetry, supergravity and particle physics", *Phys. Rep.* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5.
- [22] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [23] N. Arkani-Hamed et al., "MARMOSSET: The path from LHC data to the new standard model via on-shell effective theories", (2007), arXiv:hep-ph/0703088.
- [24] 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.
- [25] J. Alwall, M.-P. Le, M. Lisanti, and J. G. Wacker, "Model-independent jets plus missing energy searches", *Phys. Rev. D* **79** (2009) 015005, doi:10.1103/PhysRevD.79.015005, arXiv:0809.3264.
- [26] D. Alves et al., "Simplified models for LHC new physics searches", *J. Phys. G* **39** (2012) 105005, doi:10.1088/0954-3899/39/10/105005, arXiv:1105.2838.
- [27] F. Halzen and A. Martin, "Quarks and leptons: an introductory course in modern particle physics", Physics textbook, Wiley, 1984.

- [28] D. Griffiths, “Introduction to Elementary Particles”, Physics textbook, Weinheim, Germany: Wiley-VCH 454 p, 2014.
- [29] B.R. Martin and G. Shaw, “Particle Physics”, 4th Edition, Physics textbook, ISBN: 978-1-118-91190-7
- [30] I. van Vulpen, I. Angelozzi “The Standard Model Higgs Boson”, Lecture notes, October 2013 <https://www.nikhef.nl/~ivov/HiggsLectureNote.pdf> [Accessed on 01/10/2017]
- [31] M. Srednicki, “Quantum Field Theory”, Cambridge: Cambridge University Press, (2016).
- [32] S. P. Martin, “A Supersymmetry primer”, doi:10.1142/9789812839657_0001, arXiv:hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [33] Particle Data Group and C. Patrignani and others, “Review of Particle Physics”, *Chin. Phys. C* **40** (2016), 100001, doi:10.1088/1674-1137/40/10/100001
- [34] H. L. Anderson, E. Fermi, E. A. Long, D. E. Nagle, “Total Cross Sections of Positive Pions in Hydrogen” *Phys. Rev. Band* **85** (1952) S. 936
- [35] Darrah E. Nagle, “The Delta - the first pion nucleon resonance”, *LALP* 84-27 (1983).
- [36] M. Gell-Mann, “The interpretation of the new particles as displaced charge multiplets”, *II Nuovo Cimento* (1956) 4 848, doi:10.1007/BF02748000
- [37] M. Gell-Mann, “A Schematic of Baryons and Mesons”. *Phys. Lett.* **8** (3) 214–215, (1964), doi:10.1016/S0031-9163(64)92001-3.
- [38] F. Br  nner 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.
- [39] 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.
- [40] E. Fermi, “An attempt of a theory of beta radiation”, *Z. Phys.* **88** (1934).
- [41] E. Noether; M. Tavel (translator). “Invariant Variation Problems”, *Transport Theory and Statistical Physics* **1** (1971), doi:10.1080/00411457108231446, arXiv:physics/0503066.
- [42] Muon G-2 Collaboration, “Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL”, *Phys. Rev. D* **7** (2006), doi:10.1103/PhysRevD.73.072003, arxiv:hep-ex/0602035.
- [43] LIGO Scientific and Virgo Collaboration, “Observation of Gravitational Waves from a Binary Black Hole Merger”, *Phys. Rev. Lett.* **116** (2016). doi:10.1103/PhysRevLett.116.061102, arXiv:1602.03837.

- [44] D. Castelvechi, "The black-hole collision that reshaped physics", *Nature* **531** (7595): 428–431, doi:10.1038/531428a.
- [45] F. Zwicky, "On the Masses of Nebulae and of Clusters of Nebulae", *Astrophys. J.* **86** 217–246, 1937, doi:10.1086/143864.
- [46] '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 = 4kpc/ to UGC 2885 /R = 122 kpc/'", *Astrophys. J.*, 238:471, 1980, doi:10.1086/158003.
- [47] Super-Kamiokande Collaboration, "Evidence for oscillation of atmospheric neutrinos", *Phys. Rev. Lett.* **81** (Aug, 1998) 1562–1567, doi:10.1103/PhysRevLett.81.1562.
- [48] OPERA Collaboration, "Observation of a first candidate event in the OPERA experiment in the CNGS beam", *Phys. Lett. B* **691** no. 3, (2010) 138 – 145, doi:10.1016/j.physletb.2010.06.022.
- [49] D. H. Perkins, "Introduction to high energy physics", Cambridge University Press, 2000.
- [50] S. Coleman, J. Mandula, "All Possible Symmetries of the S Matrix, *Phys. Rev.* **159**(5), 1967, pp. 1251–1256.
- [51] R. Haag, J. Lopuszanski, and M. Sohnius, "All Possible Generators of Supersymmetries of the s Matrix", *Nucl. Phys. B* **88**, (1975).
- [52] Super-Kamiokande Collaboration, "Search for proton decay via $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ in 0.31 megaton-years exposure of the Super-Kamiokande water Cherenkov detector", *Phys. Rev. D* **95** (2017) no.1, 012004, doi:10.1103/PhysRevD.95.012004, arXiv:1610.03597.
- [53] G. R. Farrar and P. Fayet, "Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry", *Phys. Lett. B* **76** (1978) 575, doi:10.1016/0370-2693(78)90858-4.
- [54] C. Boehm, A. Djouadi, and M. Drees, "Light scalar top quarks and supersymmetric dark matter", *Phys. Rev. D* **62** (2000) 035012, doi:10.1103/PhysRevD.62.035012, arXiv:hep-ph/9911496.
- [55] C. Balazs, M. Carena, and C. E. M. Wagner, "Dark matter, light stops and electroweak baryogenesis", *Phys. Rev. D* **70** (2004) 015007, doi:10.1103/PhysRevD.70.015007, arXiv:hep-ph/0403224.
- [56] H. B. Kim and J. E. Kim, "Coupling constant unification and LEP data", *J. Korean Phys. Soc.* **27**, 129 (1994), arXiv:hep-ph/9302260.

- [57] ALEPH, DELPHI, L3, OPAL, 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, arXiv:hep-ex/0509008.
- [58] F. Wilczek, S. Raby and S. Dimopoulos, “Unification of couplings”, *Physics Today*, **44**(10), 25-33. (1991, October), doi: <http://dx.doi.org/10.1063/1.881292>
- [59] H. Georgi and S. Dimopoulos, “Softly Broken Supersymmetry and SU(5)”, *Nucl. Phys. B* **193** 150 (1981), doi:10.1016/0550-3213(81)90522-8.
- [60] J. Ellis, F. Luo, K. A. Olive and P. Sandick, “The Higgs Mass beyond the CMSSM”, *Eur. Phys. J. C* **73** (2013) no.4, 2403, doi:10.1140/epjc/s10052-013-2403-0, arXiv:1212.4476.
- [61] 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.
- [62] R. M. Barnett, J. F. Gunion, and H. E. Haber, “Gluino Decay Patterns and Signatures”, *Phys. Rev. D* **37** (1988) , no. 7, 1892–1907.
- [63] Particle Data Group and K. Nakamura and others, “Supersymmetry, Part II (Experiment)”, in: Review of Particle Physics”, *J. Phys. G* **37**, 075021 (2010).
- [64] O. Buchmueller et al., “The CMSSM and NUHM1 in Light of 7 TeV LHC, $B_s \rightarrow \mu^+ \mu^-$ and XENON100 Data”, *Eur. Phys. J. C* **72** (2012) 2243, doi:10.1140/epjc/s10052-012-2243-3, arXiv:1207.7315.
- [65] M. R. Buckley, D. Feld, S. Macaluso, A. Monteux and D. Shih, “Cornering Natural SUSY at LHC Run II and Beyond”, *JHEP* **1708** (2017) 115, doi:10.1007/JHEP08(2017)115, arXiv:1610.08059.
- [66] C. Borschensky, M. Krämer, 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.
- [67] 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.
- [68] 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.
- [69] W.J. Stirling, “proton-(anti)proton cross sections”, private communication, WJS2012.

- [70] CERN, 2017, *CERN Accelerating science*, Available at: <https://home.cern>. [Accessed 9 December 2017].
- [71] CERN, 2016-2017, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, Jul 2016, OPEN-PHO-ACCEL-2016-009, Available at: <http://cds.cern.ch/record/2197559>, [Accessed 9 December 2017].
- [72] CERN, 2012-2017, *Linear accelerator 2*, Available at: <http://cds.cern.ch/record/1997427>, [Accessed 9 December 2017].
- [73] CERN, 2012-2017, *The Proton Synchrotron Booster*, Available at: <http://cds.cern.ch/record/1997372>, [Accessed 9 December 2017].
- [74] CERN, 2012-2017, *The Proton Synchrotron*, Available at: <http://cds.cern.ch/record/1997189>, [Accessed 9 December 2017].
- [75] CERN, 2012-2017, *The Super Proton Synchrotron*, Available at: <http://cds.cern.ch/record/1997188>, [Accessed 9 December 2017].
- [76] 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. B* **122** (1983) 476, doi:10.1016/0370-2693(83)91605-2
- [77] UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c² at the CERN SPS Collider”, *Phys. Lett. B* **126** (1983) 398, doi:10.1016/0370-2693(83)90188-0
- [78] O. S. Brüning et al., “LHC Design Report”, CERN, Geneva, 2004. doi:10.5170/CERN-2004-003-V-1
- [79] 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.
- [80] 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.
- [81] CERN, 2017, *The LHC racks up records*, Available at: <http://cds.cern.ch/record/2272474>, [Accessed 9 December 2017].
- [82] Manuel Franco Sevilla, “Reweight procedure for ttbar based on the number of ISR jets”. Private communication, 2017.
- [83] CMS Collaboration, *CMS Luminosity Public Results - 2016 proton-proton 13 TeV collisions*, Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>, [Accessed 9 December 2017].
- [84] CERN, *The HL-LHC project*, Available at: <http://hilumilhc.web.cern.ch/about/hl-lhc-project> [Accessed 9 December 2017].

- [85] 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.
- [86] 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.
- [87] 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.
- [88] 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.
- [89] CMS Collaboration, “CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger”, CERN-LHCC-2002-026., ISBN: 9290831114.
- [90] A. Bell et al., “Fast beam conditions monitor BCM1F for the CMS experiment”, *Nucl.Instrum.Meth. A* **614** (March, 2010) 433–438, doi:10.1016/j.nima.2009.12.056, arXiv:0911.2480.
- [91] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”, CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [92] T.Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput.Phys.Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [93] W.Beenakker, R.Höcker, M.Spira, and P.M.Zerwas, “Squark and gluino production at hadron colliders”, *Nucl. Phys. B* **492** (1997) 51, doi:10.1016/S0550-3213(97)00084-9, arXiv:hep-ph/9610490.
- [94] 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.
- [95] 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.
- [96] 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.
- [97] 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.

- [98] 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].
- [99] 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.
- [100] Particle Data Group and C. Patrignani and others, “Review of Particle Physics, Monte Carlo event generators”, *Chin. Phys. C* **40**,100001 (2016).
- [101] GEANT4 Collaboration, “GEANT4: A Simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.
- [102] A. Giammanco, “The Fast Simulation of the CMS Experiment”, *J. Phys. Conf. Ser.* **513** (2014) 022012, doi:10.1088/1742-6596/513/2/022012.
- [103] 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.
- [104] CMS Collaboration, *CMS-doc-4172: Static image with tracks*, Available at: <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=4172>, [Accessed 9 December 2017].
- [105] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, arXiv:1706.04965.
- [106] 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.
- [107] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [108] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov, “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”, *J. Phys. G* **31** N9(2005), doi:10.1088/0954-3899/31/9/N01, arXiv:physics/0306087.
- [109] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014) no.10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [110] T. Speer et al., “Vertex Fitting in the CMS Tracker”, CMS Note 2006-032, 2006.
- [111] K. Rose, “Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems”, Proceedings of the IEEE **86** (1998), doi:10.1109/5.726788.

- [112] R. Fruhwirth and A.Strandlie ,“Track fitting with ambiguities and noise: a study of elastic tracking and nonlinear filters”, *Comput. Phys. Commun.* 120 (1999) 197, doi:10.1016/S0010-4655(99)00231-3.
- [113] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, “Adaptive vertex fitting”, *J. Phys. G* 34 (2007) N343, doi:10.1088/0954-3899/34/12/N01.
- [114] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* 12 (2017) no.02, P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [115] CMS Collaboration. “Determination of jet energy calibration and transverse momentum resolution in CMS”, *JINST* 6 P11002, 2011.
- [116] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm”, *JHEP* 4 (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [117] CMS Collaboration, “Identification of b quark jets at the CMS Experiment in the LHC Run 2”, CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, 2016.
- [118] CMS Collaboration, S. Chatrchyan et al., “Identification of b-quark jets with the CMS experiment”, *JINST* 8 (2013) P04013, arXiv:1211.4462.
- [119] CMS Collaboration, “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV”, *JINST* 10 (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [120] K. Rehermann and B. Tweedie, “Efficient Identification of Boosted Semileptonic Top Quarks at the LHC”, *JHEP* 1103 (2011) 059 doi:10.1007/JHEP03(2011)059, arXiv:1007.2221.
- [121] CMS Collaboration, “Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large-radius jets”, *JHEP* 1608 (2016) 122, doi:10.1007/JHEP08(2016)122, arXiv:1605.04608.
- [122] CMS Collaboration, “Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV”, *JINST* 10 (2015), no. 02, P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [123] CMS Collaboration, “Studies of Tracker Material”, CMS-PAS-TRK-10-003.
- [124] C. G. Lester and D. J. Summers, “Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders”, *Phys. Lett. B* 463 (1999) 99, doi:10.1016/S0370-2693(99)00945-4, arXiv:hep-ph/9906349.
- [125] C. Borschensky et al., *LHC SUSY Cross Section Working Group*, Available at: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections>. [Accessed on 12/07/2017].

- [126] CMS Collaboration, *CMS SUSY Results: Objects Efficiency*, Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSMoriond2017ObjectsEfficiency>, [Accessed 10 December 2017].
- [127] Henning Kirschenmann, *Trigger efficiencies*. Private communication, 2017.
- [128] CMS Collaboration, “Measurement of the Polarization of W Bosons with Large Transverse Momenta in W+Jets Events at the LHC”, *Phys. Rev. Lett.* **107** (2011) 021802 doi:10.1103/PhysRevLett.107.021802, arXiv:1104.3829.
- [129] D. Handl, “Background estimation for searches for supersymmetry in the single lepton final state in 13 TeV pp collisions”, Master’s thesis, 2016.
- [130] CMS Collaboration, “Measurement of the Polarization of W Bosons with Large Transverse Momenta in W+Jets Events at the LHC”, *Phys. Rev. Lett.* **107** 021802, (2011). doi: 10.1103/PhysRevLett.107.021802, arXiv:1104.3829.
- [131] Z. Bern et al. “Left-Handed W Bosons at the LHC”, *Phys. Rev. D* **84** 034008, (2011), doi: 10.1103/PhysRevD.84.034008.
- [132] 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.
- [133] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z]. halfChiSq S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”, *Annals Math. Statist.* **9** (1938), no. 1, 60–62, doi:10.1214/aoms/1177732360.
- [134] CMS Collaboration, “Search for supersymmetry in events with one lepton and multiple jets exploiting the angular correlation between the lepton and the missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV”, arXiv:1709.09814.
- [135] CMS Collaboration, “Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV”, *Phys. Rev. D* **96** (2017) no.3, 032003, doi:10.1103/PhysRevD.96.032003 arXiv:1704.07781.
- [136] CERN, 2017, *Summary plots from the ATLAS Supersymmetry physics group*, Available at: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY/index.html> [Accessed 15 December 2017].
- [137] CERN, 2017, *CMS Supersymmetry Physics Results*, Available at: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS> [Accessed 15 December 2017].

- [138] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth. A* **434** (1999) doi:10.1016/S0168-9002(99)00498-2 arXiv:hep-ex/9902006.
- [139] 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.
- [140] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *The Eur. Phys. J. C* **71** (2011).
- [141] J. Neyman and E. S. Pearson, “On the Problem of the Most Efficient Tests of Statistical Hypotheses”, *Phil. Trans. of the Royal Society of London Series A* **231** (1933), doi:10.1098/rsta.1933.0009.