

SEARCH FOR SUPERSYMMETRY, LEPTOQUARKS AND
STUDY OF ADVANCED PILEUP MITIGATION
TECHNIQUES

A Thesis

Submitted to the
Tata Institute of Fundamental Research, Mumbai
for the Degree of Doctor of Philosophy
in Physics

by
Bibhuprasad Mahakud

Department of High Energy Physics
Tata Institute of Fundamental Research
Mumbai
April, 2017

Final Version Submitted in April, 2017

DEDICATION

To all the physics lovers and my family

DECLARATION

This thesis is presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledge of collaborative research and discussions.

The work has been done under the guidance of Prof. Gagan Mohanty, at Tata Institute of Fundamental Research, Mumbai.

Bibhuprasad Mahakud

In the capacity as the supervisor of the candidates's thesis, I certify that the above statements are true to the best of my knowledge.

Gagan B. Mohanty

Date:

Acknowledgements

Firstly, I express my sincere gratitude to my advisor Prof. Gagan Mohanty for the continuous support of my Ph.D study and related research, for his patience and motivation. His guidance helped me in all the time of research and writing of this thesis. I am happy to have chosen him as my advisor and mentor for my Ph.D study. I also thank him for his continuous moral support that has helped me to remain cool throughout of my PhD.

Besides my advisor, I thank the rest of my thesis committee: Prof. Tariq Aziz, Prof. Sudeshna Banerjee, and Prof. Shashikant Dugad, Prof. Indranil Mazumdar and Dr. Rishi Sharma for their insightful comments and encouragement as well as for the hard questions which inceted me to widen my research from various perspectives.

My sincere thanks also go to Dr. Andrew Whitbeck, Dr. Jim Hirschauer and Dr. Nhan Viet Tran who provided me an opportunity to join their team at Fermilab, and gave access to the laboratory and research facilities. I am also thankful for useful discussions and guidance that they provided which have been very crucial to me. I thank Bibhuti Parida, my senior, for starting a new collaboration with Fermilab, USA that gave me an opportunity to work in the group.

I thank my friends Jacky, Varghese, Bajarang, Muzamil, Soureek, Nairit, Gouranga, Saurabh, Saranya, Soham, Suman, Pallabi, Arvind, Deepanwita, Sandhya, Adiba, Debashis for useful discussions that has helped my understanding throughout my PhD.

I also thank all the professors of DHEP, TIFR for stimulating discussions that has helped me a lot during my PhD.

Last but not the least, I thank my family: my parents, my wife Dr. Rasmila and my brother for supporting me morally throughout writing this thesis and my life in general.

Contents

Title page	i
Declaration	iii
Acknowledgements	iv
Contents	v
Synopsis	1
0.1 Introduction	7
0.2 CMS Detector	9
0.3 Event Reconstruction	10
0.4 Search for Supersymmetry with 13 TeV pp Collision Data	11
0.4.1 Study on 2.3 fb^{-1} of 2015 Data	12
0.4.2 Study on 12.9 fb^{-1} of 2016 Data	20
0.5 Search for Pair Production of First Generation Leptoquarks	24
0.5.1 Background Estimation	25
0.5.2 Results	25
0.6 Advanced Pileup Mitigation Techniques	25

Contents

0.6.1	Jet Grooming	26
0.6.2	Results	27
0.7	Summary	29
List of Symbols		31
List of Tables		31
List of Figures		33
Chapter 1 Introduction		1
Chapter 2 The Standard Model of Particle Physics		4
2.1	Introduction	4
2.2	Gauge Theory Formulation	7
2.3	Electroweak Theory and Higgs Mechanism	9
2.3.1	Higgs Mechanism	12
2.4	Experimental Verification of SM	14
2.5	Drawbacks of the Standard Model	15
Chapter 3 Beyond the Standard Model		17
3.1	Supersymmetry	17
3.2	Leptoquark models	21
Chapter 4 The Compact Muon Solenoid		22
4.1	CMS detector	22
4.1.1	Tracker	23
4.1.2	Electromagnetic Calorimeter	25
4.1.3	Hadron Calorimeter	28

Contents

4.1.4	Muon System	30
4.2	Object Reconstruction	34
4.2.1	Primary Vertex Reconstruction	34
4.2.2	Electron Reconstruction	34
4.2.3	Muon Reconstruction	35
4.2.4	Jet Reconstruction	35
4.2.5	Missing Transverse Energy	36
Chapter 5	Statistical Procedures for Search	37
5.1	Observed Limits	38
5.2	Expected Limits	40
5.3	Significance	41
Chapter 6	Search for SUSY with 13 TeV pp Collision Data	42
6.1	Study with 2.3 fb^{-1} of the 2015 Data	44
6.1.1	Data-MC Samples and Trigger	44
6.1.2	Event Selection Criteria and Exclusive Search Intervals	47
6.1.3	Backgrounds	51
6.1.4	Results and Statistical Interpretation	64
6.2	Study with 12.9 fb^{-1} of 2016 data	66
6.2.1	Search Binning	66
6.2.2	Z($\rightarrow \nu\bar{\nu}$) + jets Background Estimation	68
6.2.3	Results and Statistical Interpretation	69
Chapter 7	Search for Pair Production of First Generation Leptoquarks	72
7.1	Event Selection	73
7.2	Backgrounds	76

Contents

7.2.1	QCD multijet:	77
7.3	Results and Statistical Interpretation	77
Chapter 8	Advanced Pileup Mitigation Techniques	81
8.0.1	Simulation Results at High Pileup	83
Chapter 9	Summary and Conclusions	88
References		91

Contents

Synopsis

Contents

Search for Supersymmetry, Leptoquarks and Study of Advanced Pileup Mitigation Techniques

Bibhuprasad Mahakud

Department of High Energy Physics
Tata Institute of Fundamental Research,
Homi Bhabha Road, Mumbai 400005, India

We search for supersymmetry in all hadronic final states with proton-proton collision data recorded by the CMS detector at a center-of-mass energy of 13 TeV. We specifically target for pair produced gluinos in the 2015 data that correspond to an integrated luminosity of 2.3fb^{-1} , and using more than five times data collected during 2016 where we also add stop and other squark pairs to our search. The data are examined in search regions of jet multiplicity, tagged bottom quark jet multiplicity, missing transverse momentum, and the scalar sum of jet transverse momenta. The observed numbers of events in all search regions are found to be consistent with the expectations from standard model processes. Exclusion limits are calculated for simplified supersymmetric models of gluino pair production. Depending on the assumed gluino decay mechanism, and for a massless, weakly interacting, lightest neutralino, lower limits on the gluino mass from 1440 to 1700 GeV are obtained, significantly extending previous limits.

Another search for the pair production of first generation scalar leptoquarks decaying to leptons and jets is presented. The data corresponding to 2.6 fb^{-1} were collected from

Contents

proton-proton collisions at a center-of-mass energy of 13 TeV. No significant deviation is found with respect to the standard model predictions. Upper limits are set on cross section times branching fraction.

The increase of instantaneous luminosity for Run II LHC results in a large number of additional proton-proton collisions in a given event (pileup) leading to contamination of jets. We study advanced pileup removal techniques like trimming, pruning and soft-drop etc. The focus was on preparation for Run II for which we expected up to 40 additional pileup events in comparison to Run I LHC data that had typically half the pileup events on average.

Advisor: Prof. Gagan B. Mohanty

List of Publications

Journals

- Search for supersymmetry in the multijet and missing transverse momentum final state in pp collisions at 13 TeV, [Phys. Lett. B 758 \(2016\) 152.](#)

CMS Physics Analysis Summary (Public Results)

- Search for supersymmetry in events with jets and missing transverse momentum in proton-proton collisions at 13 TeV, [CMS-PAS-SUS-16-014.](#)
- Search for supersymmetry in the multijet and missing transverse momentum channel in pp collisions at 13 TeV, [CMS-PAS-SUS-15-002.](#)
- Search for pair-production of first generation scalar leptoquarks in pp collisions at $\sqrt{s}=13$ TeV with 2.6 fb^{-1} , [CMS-PAS-EXO-16-043.](#)
- Study of Pileup Removal Algorithms for Jets, [CMS-PAS-JME-14-001.](#)

Conference Proceedings

- Search for SUSY in jets+MET final state, [PoS\(DIS2016\)094.](#)

CMS Internal Notes

Contents

- Search for supersymmetry in multijet final states in proton-proton collisions at 13 TeV, [CMS-AN-16-188](#).
- Search for supersymmetry in the multijet and missing transverse momentum channel in pp collisions at 13 TeV, [CMS-AN-15-003](#).
- Search for Pair-production of Scalar First Generation Leptoquarks in pp Collisions at $\sqrt{s} = 13$ TeV, [CMS-AN-15-294](#).
- Study of Pileup Removal Algorithms for Jets, [CMS-AN-14-175](#).

0.1 Introduction

The standard model (SM) of particle physics successfully describes a wide range of phenomena. However, it is unable to explain the stability of the Higgs boson mass in the face of higher-order corrections, suggesting that the model is incomplete. Many extensions to the SM have been proposed to provide a more fundamental theory. Supersymmetry (SUSY) [1], one such extension, postulates that each SM particle has a SUSY partner from which it differs in spin by half a unit. As examples, squarks and gluinos are the SUSY partners of quarks and gluons, respectively, while neutralinos arise from a mixture of the SUSY partners of neutral (charged) Higgs and electroweak gauge bosons. Radiative corrections involving SUSY particles can compensate the contributions from SM particles and thereby stabilize the Higgs boson mass. For this cancellation to be natural, the top and bottom squark, and gluino must have masses on the order of a few TeVs, possibly allowing them to be produced at the CERN LHC.

Amongst SUSY processes, the gluino pair production, typically yielding four or more hadronic jets in the final state, has the largest possible cross section, making it an apt channel at the recently started LHC Run 2. In R-parity [34] conserving SUSY models, as are considered here, the lightest SUSY particle (LSP) is stable and assumed to be weakly interacting, leading to potentially large undetected, or missing, transverse momentum. Supersymmetry events at the LHC might thus be characterized by significant missing transverse momentum, numerous jets, and in the context of natural SUSY [69], jets initiated by top and bottom quarks.

In this search, we consider SUSY scenarios in the context of simplified models [37] of new particle production. Diagrams for the three models are shown in Fig. 1. Simplified models contain the minimal particle content to represent a topological configuration. As for SUSY production scenarios, the four simplified models can be interpreted as follows.

Contents

In the first one, shown in Fig. 1 (left), the gluino pair production is followed by the decay of each gluino to a bottom quark and an off-shell bottom squark. The off-shell bottom squark decays to a bottom quark and the LSP, where the LSP is assumed to be the lightest neutralino that escapes detection, leading to significant missing transverse energy. The second scenario, shown in Fig. 1 (middle), is the same as the first one except with top quarks and off-shell top squarks appearing in place of the bottom quarks and squarks. The third scenario, shown in Fig. 1 (right), is the corresponding situation with the gluino decay to a light-flavored quark and off-shell squark up, down, strange, and charm with equal probability, for each of them separately separately.

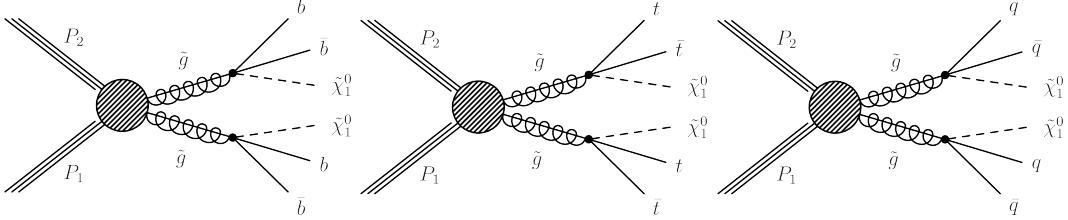


Figure 0.1: Event diagrams for the main new-physics scenarios considered in this study.

The SM quark and lepton sectors are strikingly similar in terms of the number of particles and generations. This hints that the two sectors could be related by a fundamental symmetry. Many beyond the SM theories, include such a symmetry [4; 5; 3; 70], which gives rise to new class of bosons called leptoquarks (LQs). LQs couple to both leptons and quarks, carrying lepton and baryon numbers as well as fractional electric charge. They can be either scalar or vector particles (carrying zero or one unit of spin) and are color triplets. Current experimental searches for rare processes mediated by lepton number violation and flavor-changing neutral currents suggest that LQs come in three generations which do not mix. Their pair-production at the LHC proceeds mainly via gluon-gluon (gg) fusion and quark-antiquark ($q\bar{q}$) annihilation. Due to the gg dominance, along with the fact that only one q-Feynman diagram contains the LQ-

Contents

quark-lepton vertex, the scalar LQ pair-production has a negligible dependence on the LQ-quark-lepton Yukawa coupling, usually denoted as λ . LQ searches are essentially independent of λ . In our search we consider pair produced LQs with each of them decaying to an electron and a quark. Therefore we look for signature in two electron and two jet final states in the data.

In the two physics topics discussed above as well as many other important studies at the LHC, the performance of jets is extremely important. One of the key challenges of the LHC run is the increase of instantaneous luminosity, which results in a large number of additional proton-proton collisions (pileup) in each event. In such high pileup environment, an accurate reconstruction of jet properties and shapes has become more and more demanding. The most common observables for jets at the LHC are primarily the jet p_T , pseudorapidity(η), and ϕ . In recent years, it is becoming increasingly popular to consider the internal structure of the jet. The applications include the discrimination of quark- and gluon-initiated jets as well as the identification of highly boosted hadronically decaying heavy resonances such as W/Z or Higgs boson that are contained in a single jet. In all such cases, contamination from pileup degrades our ability to effectively reconstruct the jet observables. Motivated by this, we study advanced techniques for pileup mitigation for jets in view of high pileup scenarios in Run II of the LHC.

0.2 CMS Detector

The CMS detector is built around a superconducting solenoid of 6-m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and microstrip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass-scintillator sandwich hadron calorimeter (HCAL). The tracking detectors cover a range of $|\eta| < 2.5$. The ECAL and HCAL, each composed of a barrel and two endcap

Contents

sections, extend over $|\eta| < 3.0$. Forward calorimeters on either side of the interaction point encompass $3.0 < |\eta| < 5.0$. Muons are measured within $|\eta| < 2.4$ by gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. The detector is nearly hermetic, permitting accurate measurements of missing transverse energy. A more detailed description of the detector, together with a definition of the coordinate system and relevant kinematic variables, is given in Ref. [71]. The studies presented in the thesis use information from all parts of the CMS detector.

0.3 Event Reconstruction

Physics objects used in our studies are defined using the so-called particle flow (PF) algorithm [72], which reconstructs and identifies individual particles through an optimized combination of information from different detector components. The PF candidates are classified as photons, charged and neutral hadrons, electrons, or muons. Additional quality criteria are imposed on electron, muon and photon candidates. For example, more restrictive conditions are placed on the shower shape and on the ratio of energies deposited in the HCAL and ECAL for electron and photon candidates, and similarly on the matching of track segments between the silicon tracker and muon detector for muon candidates. Photons being neutral particles do not produce tracks in the tracker, while electrons do. The event primary vertex is taken to be the one reconstructed with the largest sum of charged-track p_T^2 values and is required to lie within 24 cm (2 cm) of the center of the detector in the direction along (perpendicular to) the beam axis. Tracks from extraneous pp interactions within the same or a nearby bunch crossing (pileup) are removed . The PF objects serve as inputs for jet reconstruction, based on the anti- k_t algorithm [73] with a distance parameter of 0.4. Jet quality criteria are applied to eliminate, for example, spurious events caused by calorimetric noise. Contributions to an

Contents

individual jets p_T from pileup interactions are subtracted, and corrections are applied as a function of jet p_T and η to account for residual effects of nonuniform detector response.

0.4 Search for Supersymmetry with 13 TeV pp

Collision Data

Because of the large mass scale and their all-hadronic nature, the targeted SUSY events are expected to exhibit large values of H_T , where H_T is the scalar sum of the jet p_T . As a measure of missing transverse momentum, we use the variable H_T^{miss} , which is the magnitude of the vector sum of the jet p_T . We present a general search for gluino pair production leading to final states with large H_T , large H_T^{miss} as well as large jet multiplicity. The data are examined in exclusive bins of N_{jet} , $N_{b-\text{jet}}$, H_T , and H_T^{miss} , where N_{jet} is the number of jets and $N_{b-\text{jet}}$, the number of tagged bottom quark jets.

The principal sources of background arise from the SM production of top quarks, a W or Z boson in association with jets (W+jets or Z+jets), and multiple jets through the strong interaction. We refer to the last class of background as quantum chromodynamics (QCD) multijet events. Although events with top quarks mostly come from top quark-antiquark ($t\bar{t}$) production, a modest contribution is also from single top quark processes. The W and Z bosons in W+jets and Z+jets events can be either on- or off-shell. For top quark and W+jets events, significant H_T^{miss} can arise if the W boson decays leptonically, producing a neutrino and an undetected charged lepton, while Z+jets events can exhibit significant H_T^{miss} if the Z boson decays to two neutrinos. For QCD multijet events, significant H_T^{miss} can arise if the event contains a charm or bottom quark that undergoes a semileptonic decay; however the principal source is the mismeasurement of jet p_T . The signal vs. background composition plots in search variables are shown in Fig. 0.2 and

Contents

0.3.

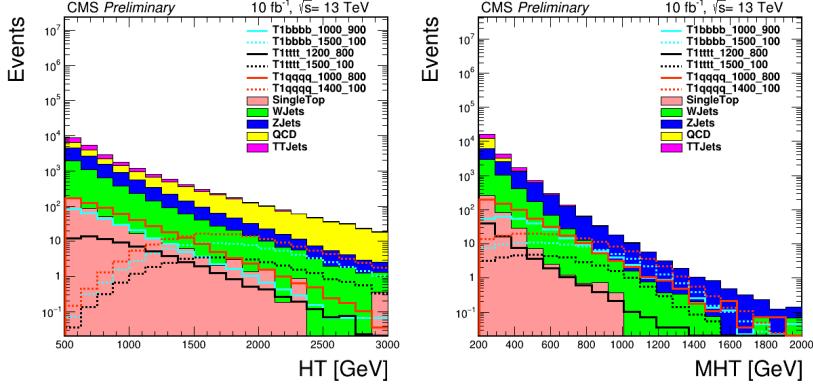


Figure 0.2: Signal vs. stacked backgrounds in H_T (HT, left) and H_T^{miss} (MHT, right)

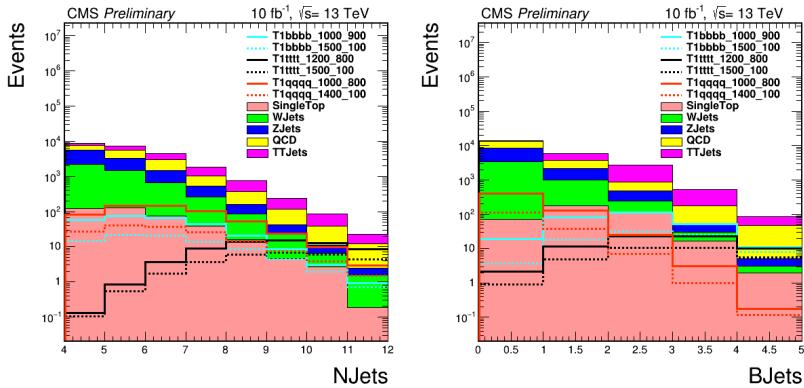


Figure 0.3: Signal vs. stacked backgrounds in N_{jet} (NJets, left) and $N_{b-\text{jet}}$ (MHT, right)

0.4.1 Study on 2.3 fb⁻¹ of 2015 Data

In this section, we discuss the analysis strategy and results of the search conducted using 2015 data. With as little as 2.3 fb⁻¹ data we search for gluino pair production based on the above four variables. To get maximum signal and minimum background we select events with the criteria defined below.

Contents

0.4.1.1 Event Selection and Search Region

The following requirements define the signal event candidates:

- $N_{\text{jet}} \geq 4$, where the jets must lie within $|\eta| < 2.4$; we require at least four jets because of our focus on gluino pair production;
- $H_T > 500 \text{ GeV}$, where H_T is the scalar p_T sum of jets with $|\eta| < 2.4$;
- $H_T^{\text{miss}} > 200 \text{ GeV}$, where H_T^{miss} is the magnitude of \vec{H}_T^{miss} , the negative of the vector p_T sum of jets with $|\eta| < 5$; the η range is extended in this case so that \vec{H}_T^{miss} better represents the total missing transverse momentum in a given event;
- No identified, isolated electron (muon) candidate with $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$ (< 2.4);
- No isolated charged-particle track with $|\eta| < 2.4$, $m_T < 100 \text{ GeV}$, and $p_T > 10 \text{ GeV}$ ($p_T > 5 \text{ GeV}$ if the track is identified as an electron or muon candidate) where m_T is the transverse mass formed from the \vec{p}_T^{miss} and isolated-track p_T vector, with \vec{H}_T^{miss} the negative of the vector p_T sum of all PF objects;
- $\Delta\phi_{\vec{H}_T^{\text{miss}}, j_i} > 0.5$ (> 0.3) for the two highest p_T jets j_1 and j_2 (the next two highest p_T jets j_3 and j_4), with the azimuthal angle between H_T^{miss} and the p_T vector of jet j_i .

The search is performed in the following 72 ($= 3 \times 4 \times 3 \times 3$) exclusive intervals of the four search variables:

- **3** N_{jet} bins: 4-6, 7-8 and ≥ 9 ;
- **4** $N_{\text{b-jet}}$ bins: 0, 1, 2 and ≥ 3 ;
- **3** H_T bins: 500-800, 800-1200 and $> 1200 \text{ GeV}$;

Contents

- **3** H_T^{miss} bins : 200-500, 500-750 and >750 GeV.

0.4.1.2 Background Estimation

In this section, we describe the evaluation of the background from SM processes. The evaluation relies on data control regions (CRs) selected using similar criteria to the search regions. The backgrounds are divided into four different types, namely Z to neutrinos, lost lepton, QCD and hadronic tau.

Z to neutrinos: This is the most important backgrounds being an irreducible one. We are talking about events with a Z boson produced in association with jets when the Z decays to two neutrinos. The most straightforward way to measure this background is to exploit the decays $Z(\ell^+\ell^-)$ +jets in which the Z boson can be reconstructed from the observed pair of muons or electrons. The efficiency-corrected yields from these decays can be translated directly into the $Z(\nu\bar{\nu})$ +jets background yield by the known branching ratios. The limitation of this approach arises from the rather small branching ratio between the charged and neutral leptons, so that the transfer factor from the control sample measurement to the predicted background is larger than one (ignoring efficiencies, the branching ratio itself is approximately 3 when both muon and electron pairs are used).

The alternative approach is to exploit the similarity to Z boson radiation of the more copious radiation of photons. Here the challenge is to obtain validation in data of the MC predictions connecting the two processes.

Our baseline strategy is to use the γ +jets sample to determine the yields in the 18 bins corresponding to $N_{b-\text{jet}}=0$. These are compared with the $Z(\ell^+\ell^-)$ +jets yields in the low- N_{jet} bin to establish the systematic uncertainty of the physics modeling of γ +jets, and the normalization corrected if necessary. The extrapolation to bins with $N_{b-\text{jet}} > 0$

Contents

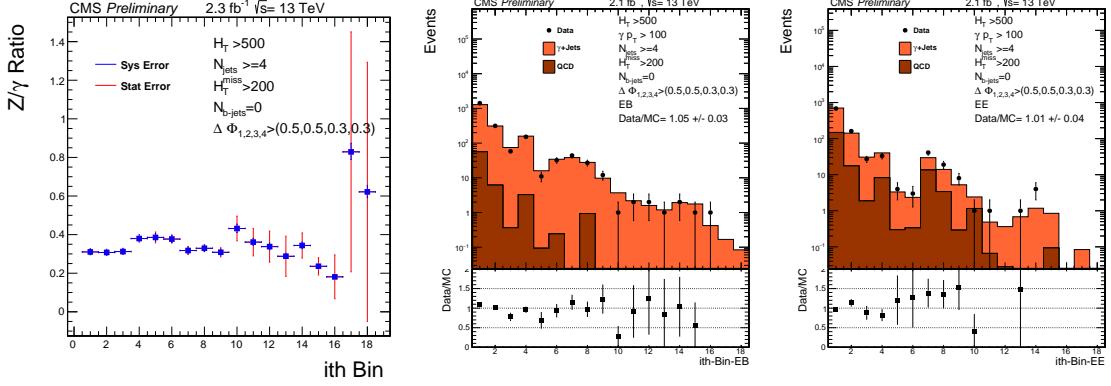


Figure 0.4: Physics ratio $R_{Z(\nu\bar{\nu})}/\gamma$ (left), numbers of observed events in the photon control samples in barrel(EB, middle) and endcap(EE, right) compared to simulation.

is performed to the extent possible with the $Z(\ell^+\ell^-)$ +jets data sample, supplemented with MC information where necessary. We use simulation to correct for the resulting distribution to the higher N_{jet} bins.

As shown in Eq.(0.1), we predict the number of $Z \rightarrow \nu\bar{\nu}$ +jets events contributing to each of the 18 0-btag analysis bins ($N_{Z(\rightarrow\nu\bar{\nu})}^{\text{pred}}$) from the number of events in the corresponding bin of the γ +jets control sample (N_{γ}^{obs}), the purity of the control sample (β), and the ratio of the numbers of $Z \rightarrow \nu\bar{\nu}$ +jets events and γ +jets events obtained from leading order MADGRAPH+PYTHIA ($\mathcal{R}_{Z(\nu\bar{\nu})}/\gamma \cdot \mathcal{F}_{\text{dir}}$). Here \mathcal{F}_{dir} is the fraction of prompt photons that are direct.

$$N_{Z \rightarrow \nu\bar{\nu}}^{\text{pred}} = DR \cdot \mathcal{R}_{Z(\nu\bar{\nu})}/\gamma \cdot \mathcal{F}_{\text{dir}} \cdot \beta \cdot N_{\gamma}^{\text{obs}} \quad (0.1)$$

DR , called the double ratio, is a correction factor to the physics ratio $R_{Z(\nu\bar{\nu})}/\gamma$. This is calculated by measuring the $R_{Z(\ell^+\ell^-)/\gamma}$ both in data and simulation and then taking the ratio.

Prompt photons can be distinguished from non-prompt photons by the shapes of

Contents

their showers in the ECAL, as described by the well-known quantity $\sigma_{inj\eta}$. The purity is determined with a two-component fit to the $\sigma_{inj\eta}$ distribution in the photon control sample. The PDF for the prompt component is fitted directly in data using a gaussian function.

Fig 0.4 shows the data vs MC simulation for 18 kinematic bins (0-btag) where a photon+ jet method is employed.

Lost lepton: SM events (mostly $t\bar{t}$ and $W+jets$) with muons or electrons can satisfy the event selection and enter the signal sample as lost-lepton background if the requirements for any of the following analysis steps are not satisfied

- Kinematic acceptance,
- Reconstruction, or
- Isolation.

The basic idea behind our data-driven method to evaluate the lost-lepton background is to select single-lepton control samples (CS) in the data, through inversion of the lepton vetoes, and to weight each CS event by a factor that represents the probability for a lost-lepton event to appear with the corresponding search-variable values: H_T , H_T^{miss} , N_{jet} , and $N_{\text{jet}}^{\text{miss}}$. The weights are determined through evaluation of the efficiencies for each analysis step. The weighted distributions of the search variables, summed over the events in the CS, define the predicted lost-lepton background in the respective search regions.

QCD: The H_T^{miss} in QCD multijet events is almost always due to a mismeasured jet in the event, thus the H_T^{miss} direction is usually close to the jet. The $\Delta\phi$ variable is the minimum ϕ difference between H_T^{miss} and one of the four highest p_T jets.

Contents

The low $\Delta\phi$ region is significantly enriched in QCD events. The sample of events with the $\Delta\phi$ requirement inverted (i.e., $\Delta\phi_1 < 0.5$ or $\Delta\phi_2 < 0.5$ or $\Delta\phi_3 < 0.3$ or $\Delta\phi_4 < 0.3$) serves as the QCD control sample. The background at high $\Delta\phi$, is estimated from the QCD yield at low $\Delta\phi$ and a high/low ratio R^{QCD} for the QCD component. The $\Delta\phi$ distribution shows that the high/low ratio has some dependence on the search variables H_T , H_T^{miss} , and N_{jet} . We model this dependence by assuming that it factorizes. That is, we assume the H_T dependence does not depend on H_T^{miss} or N_{jet} and similarly for H_T^{miss} and N_{jet} .

Hadronic tau: To evaluate the $t\bar{t}$, single-top and $W+\text{jets}$ backgrounds that arises when a W boson decays to a neutrino and a hadronically decaying τ lepton (τ^h), we employ a tau-template method. In this approach, the τ^h background is estimated from a control sample (CS) of $\mu+\text{jets}$ events, which we select by requiring exactly one muon with $p_T > 20$ GeV and $|\eta| < 2.1$. This single-muon CS is mainly composed of $t\bar{t}(\rightarrow \mu\nu)$ and $W(\rightarrow \mu\nu)+\text{jets}$ events. Since both $\mu+\text{jets}$ and $\tau^h+\text{jets}$ events arise from the same underlying process, the hadronic components of the two event classes are expected to be the same, aside from the response of the detector to a muon or a τ^h jet. The basic idea behind the method is to smear the muon p_T in the CS events, using MC-derived response functions (the “templates”), in order to emulate the τ^h jet response. Global hadronic variables such as N_{jet} , H_T , and H_T^{miss} are then recomputed, and the full analysis procedure is subsequently applied.

0.4.1.3 Uncertainties

Various kinds of systematic and statistical uncertainties that are considered in the analysis. The uncertainties that result from the background estimations including data control region statistics, purity of control region , CR trigger efficiency etc. are discussed in de-

Contents

tails in Ref. [52]. Below we briefly describe various systematic uncertainties that affect the expected signal yield.

- **Luminosity:** A flat 4.6 % uncertainty on luminosity is propagated to the signal yield.
- **b-tag efficiency:** The b-tagging and mistagging scale factors are functions of the jet p_T and η . The scale factors are varied by their uncertainties and these variations are propagated as migrations between the different signal bins.
- **MC statistics:** The MC statistical uncertainties are propagated to the signal yield
- **Trigger efficiency:** The trigger efficiencies are measured in the data. The effect of the statistical and systematic uncertainties is at most 1.1 % at low H_T^{miss} .
- **Pileup reweighting:** The uncertainties in the pileup reweighting correction are derived from the uncertainties in the minimum bias cross section and the difference between the expected and observed number of interactions in the data. The minimum bias cross section in the 13 TeV is estimated to be 69 mb with an uncertainty of $\pm 5\%$. The correction is varied according to these uncertainties, with a maximum effect of 0.5 %.
- **Scale:** The uncertainty is calculated from using the envelope of weights from varying the renormalization and factorization scales. The effect on the yield of non-compressed samples is less than 0.1 % and on compressed samples ranges from 1 % to 3 %.
- **ISR:** The effect on the yield of non-compressed samples is less than 0.1 % and on compressed samples ranges from 3 % to 11 %.

Contents

- **Jet Energy Corrections:** The jet energy corrections (JECs) are varied using the p_T and η dependent jet energy scale uncertainties from the official database, with a separate set of corrections for the fast simulation samples. The overall effect ranges from 0.5 % to 4%.
- **PDF:** The LHC4PDF prescription for the uncertainty on the total cross section is included as $\pm 1 \sigma$ bands in the results.

0.4.1.4 Results

The data in the signal regions are found to be in generally good agreement with the predicted backgrounds. Therefore we do not see any evidence for new physics. For the 72 search bins, the observed data and the pre-fit predictions for each background component are shown in Fig. 0.5. The 95% confidence-level (CL) upper limit is calculated on the production cross section taking all 72 bins. The upper limits on the signal cross section and the exclusion curves are shown in Fig. 0.6. For calculating the upper limits we use a test statistic $q_\mu = -2\ln(L_\mu/L_{max})$, where L_{max} is the maximum likelihood determined by allowing all parameters including the SUSY signal strength μ to vary and L_μ is the maximum likelihood for a fixed signal strength . The details of the statistical procedure can be found in Ref. [51]. For an explanation of the treatment of uncertainties, we refer to Ref. [52]. As can be seen from the plots in Fig. 0.6 the observed exclusion limits for low LSP masses lie around 1600 GeV both for four top and four b-quark final state, for four light quark final state, it is around 1450 GeV of gluino mass. The small disagreement of observed exclusion curves with the expected ones can be ascribed to the small insignificant excesses of events we see in various bins.

Contents

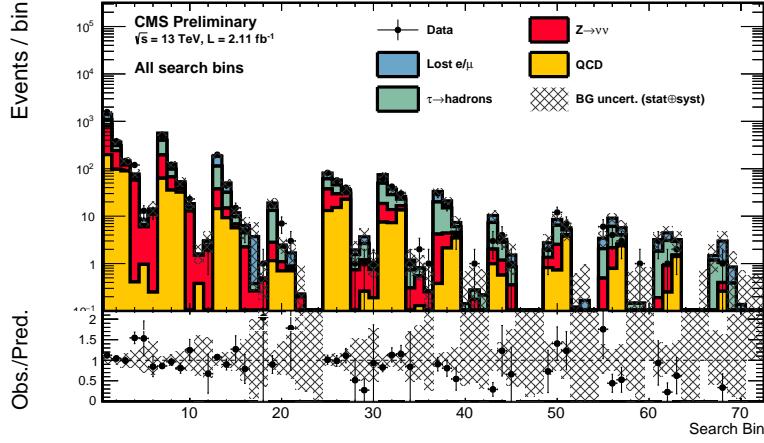


Figure 0.5: Data vs. the SM background before fit

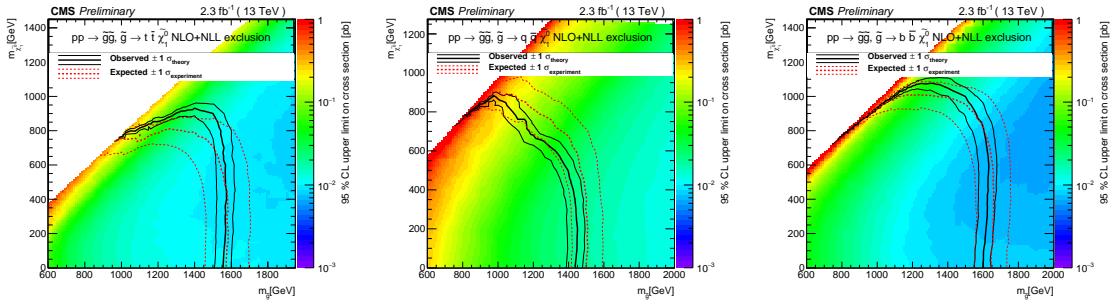


Figure 0.6: The 95% CL upper limits on the production cross sections for four top quark(left),four light quark(middle) and four b-quark(right) in the final state.

0.4.2 Study on 12.9 fb^{-1} of 2016 Data

In 2016 we lowered the H_T and N_{jet} thresholds to enhance the search sensitivity to some stop and squark production models. Also the number of bins is changed from 72 to 160 for an increased sensitivity. The important event selection and search region definitions that are different with respect to the 2015 analysis are described below.

0.4.2.1 Search Region and Event Selection

The following requirements define the selection criteria for signal event candidates:

Contents

- $N_{\text{jet}} \geq 3$, where the jets must satisfy $|\eta| < 2.4$; we change the jet multiplicity threshold to 3 because of our change in focus to direct squark-pair production in addition to gluino-pair production;
- $H_T > 300 \text{ GeV}$;
- $H_T^{\text{miss}} > 300 \text{ GeV}$; All other criteria remain almost the same as for the 2015 analysis (see Sec. 4.1.1).

The search is performed in the following 160 ($= 4 \times 4 \times 10$) exclusive intervals of the four search variables:

- **3** N_{jet} bins: 3-4, 5-6, 7-8, ≥ 9 ;
- **4** $N_{\text{b-jet}}$ bins: 0, 1, 2, ≥ 3 ;
- **10** bins in H_T and H_T^{miss} : defined below

Bin	H_T range [GeV]	H_T^{miss} range [GeV]
1:	300–500	300–350
2:	500–1000	300–350
3:	>1000	300–350
4:	350–500	350–500
5:	500–1000	350–500
6:	>1000	350–500
7:	500–1000	500–750
8:	>1000	500–750
9:	750–1500	>750
10:	>1500	>750

Contents

0.4.2.2 Background Estimation

Similar methods to estimate various SM backgrounds are employed in this analysis even though their relative composition has changed from 2015 owing to the change of baselines. As the phase space has changed, these methods are optimized accordingly. Below we describe the changes for the $Z(\rightarrow \nu\bar{\nu}) + \text{jets}$ background. Details of other background estimations can be found in Ref. [55].

$Z(\rightarrow \nu\bar{\nu}) + \text{jets}$:

- A new high-statistics MC sample is used to calculate the transfer factor $R_{Z/\gamma}$ helping to reduce the related systematic uncertainty.
- For the calculation of photon purity, a charged isolation template is used instead of $\sigma_{i\eta i\eta}$ as we find it to have a better performance.
- Trigger for the photon control region is also changed.

Note that due to the change in H_T^{miss} and N_{jet} threshold the contribution of Z background is increased in many bins.

0.4.2.3 Uncertainties

The uncertainties are treated in the same way as it is done with 2015 data. The % uncertainties are also similar to the previous analysis.

0.4.2.4 Results

Fig. 0.7 shows the predicted background vs. the observed data. The background is found to be statistically compatible with data for all 160 regions. Thus, we do not see any evidence for new physics. The 95% CL upper limits on the cross section are shown in Fig. 0.8. These plots show significant improvements in terms of expected limits

Contents

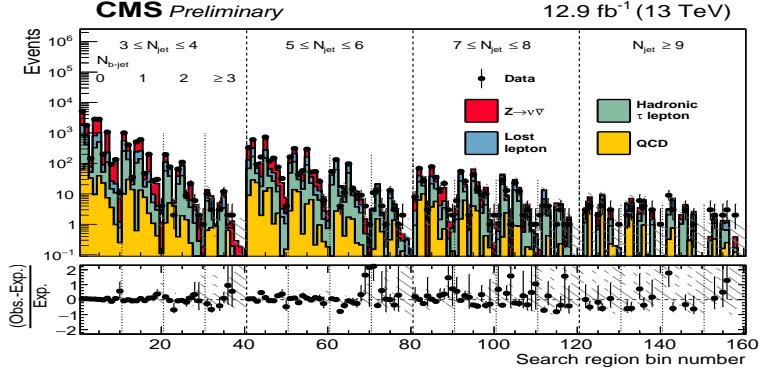


Figure 0.7: Data vs. the SM background before fit.

when compared to the previous results with 2.3 fb^{-1} of data (Fig. 6.12). The mismatch between the observed and expected curves owing to the deficit of events in some of the bins.

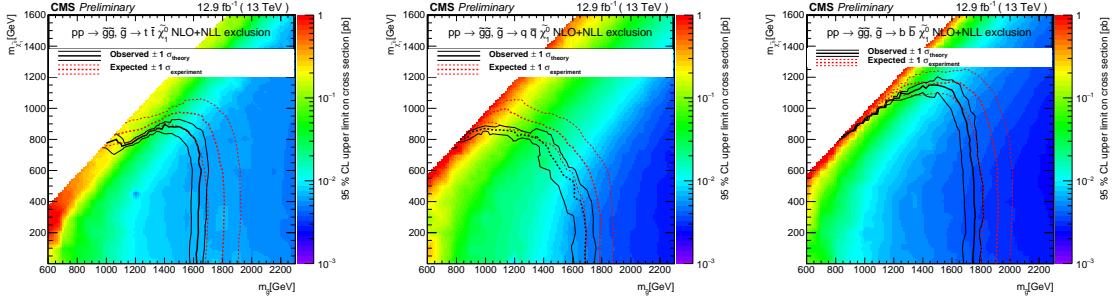


Figure 0.8: The 95% CL upper limits on the production cross sections for four top quark (left), four light quark (middle) and four b-quark (right) in the final state.

Contents

0.5 Search for Pair Production of First Generation Leptoquarks

The eejj final state is the end product of pair-produced leptoquarks with each of them decaying to an electron and a jet. Events containing two electrons and at least two jets are selected, where the two leading p_T electrons and jets are used in the analysis.

A set of cuts on three search variables are optimized for an improved signal sensitivity. The variables are:

- S_T is the scalar sum of p_T of two electrons and the leading two jets;
- M_{ee} is the invariant mass of two leading electrons; and
- M _{ℓ,j} ^{min} is the smaller lepton-jet invariant mass for the assignment of jets and leptons to leptoquarks that minimizes the LQ- $\overline{\text{LQ}}$ invariant mass difference.

The cuts on the above variables are calculated for different leptoquark mass points.

The details of selections and optimization procedures can be found in Ref. [74].

The data vs. backgrounds plots for two variables are shown in the Fig. 0.9.

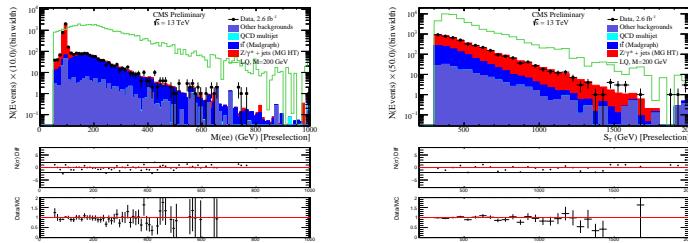


Figure 0.9: Data vs. background plots for M_{ee} (left) and S_T (right).

0.5.1 Background Estimation

The major backgrounds from SM processes are $Z+jets$ and $t\bar{t}$, where single top, $W+jets$, diboson, and $\gamma+jets$ contribute at a lower level. There is also an instrumental background from QCD events with jets faking electrons. Below we describe how these backgrounds are determined in our analysis.

- The $Z+jets$ and $t\bar{t}$ background shapes are taken from MC simulations, and normalized to data using the eejj preselection. More details are given in Ref. [74].
- Single top, $W+jets$, diboson, and $\gamma+jets$ backgrounds are derived completely from MC samples that are scaled to the cross sections.
- QCD background is determined using a data-driven fake rate method, as described in Ref. [74].

0.5.2 Results

Similar to the two SUSY analyses, we observe no significant excess of events as compared to the SM backgrounds. The broad excess of the events that was seen in the 8 TeV analysis [56] has disappeared with the 2015 data. We set the upper limits on the cross section times branching fraction using the same tool as before. Fig. 0.10 shows both the observed and expected limits for different leptoquark masses. We exclude leptoquark masses up to 1130 GeV from this study.

0.6 Advanced Pileup Mitigation Techniques

Contents

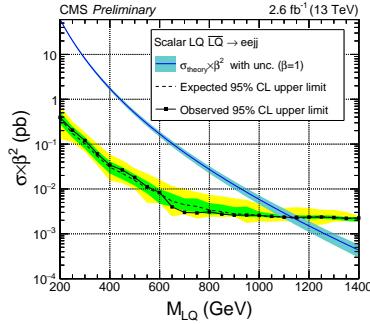


Figure 0.10: The 95% CL upper limits on the production cross sections as a function of leptoquark mass.

0.6.1 Jet Grooming

Grooming, introduced in Ref. [75], is intended to remove soft and wide-angle radiation from a jet. It is typically used to reduce the overall jet mass of QCD (quark- and gluon-initiated) jets while retaining the larger jet mass for jets originating from heavy particles such as the top quark and W/Z/H boson. Additionally, it can also help minimize the pileup dependence on jet mass. In general, grooming alters the soft structure of the jet while other observables may rely on this soft structure. Here we explore three grooming methods to mitigate the effects of pileup on large-R jets ($R=0.8$).

1. Pruning [60] reclusters the constituents of the jet through the Cambridge-Aachen (CA) algorithm [62] using the same distance parameter. At each step in the clustering algorithm, the softer of the two particles i and j to be merged is removed when the following conditions are met:

$$z_{ij} = \frac{\min(p_{T_i}, p_{T_j})}{p_{T_i} + p_{T_j}} < z_{\text{cut}} \text{ and } \Delta R_{ij} > \frac{2 \times r_{\text{cut}} \times m_J}{p_T},$$

where m_J and p_T are the mass and transverse momentum of the originally-clustered jet, and z_{cut} and r_{cut} are parameters of the algorithm.

2. Trimming [61] ignores particles within a jet that fall below a dynamic threshold in p_T . It reclusters the constituents of the jet using the k_t algorithm [73] with a radius

Contents

parameter r_{sub} , accepting only the subjets that have $p_{\text{T,sub}} > p_{\text{T,frac}} \lambda_{\text{hard}}$, where $p_{\text{T,frac}}$ is a dimensionless cutoff parameter, and λ_{hard} is some hard QCD scale chosen to be equal the p_{T} of the original jet.

3. Soft-drop [63] declusters the jet recursively to remove soft and wide-angle radiation from the jet. The jet is reclustered using the CA algorithm. Then the jet is declustered and at each step, subjets j_1 and j_2 are used to define the following condition:

$$\frac{\min(p_{\text{T},j_1}, p_{\text{T},j_2})}{p_{\text{T},j_1} + p_{\text{T},j_2}} > z_{\text{cut}} \times \left(\frac{\Delta R_{12}}{R_0} \right)^\beta$$

where the algorithm parameters are z_{cut} and β . If the condition is met, the declustering continues, otherwise only the leading p_{T} subjet is kept. In the case when $\beta = 0$, soft drop can be considered a generalization of the modified mass drop tagger (MMDT) [62].

The (groomed) masses are corrected for pileup using a four-vector safe subtraction. In the cases of soft drop and trimming, the four-vector subtraction corrects the jet p_{T} and mass at each step in the algorithm. For pruning, however, the correction is applied to the final product using the pruned jet area. The parameters for the grooming algorithms explored in this study can be found in Ref. [59]

0.6.2 Results

The groomed mass distributions of charged hadron subtracted PF (PFCHS) [72; 59] jets are shown in Fig. 0.11. One sees that after the application of grooming due to the reduction of pileup the peaks in the distributions shift to left whereas in the ungroomed case they have an unphysical looking like peak in the higher side. Even though these mass distributions do not tell everything but they say because of the application of grooming the mass distributions change aggressively.

Average jet mass and jet resolutions are taken as quality parameters for determining

Contents

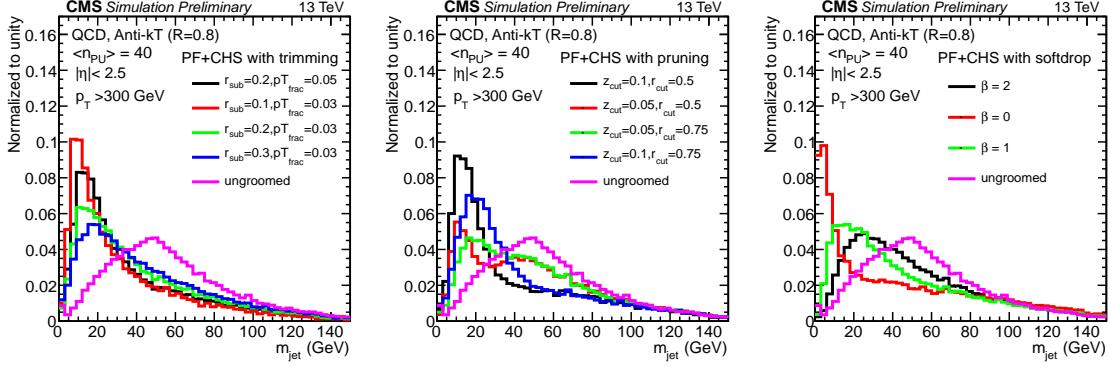


Figure 0.11: The mass distributions of PFCHS jets after application of trimming (left), pruning (middle) and soft-drop (right) for different set of parameters. The ungroomed mass distribution is shown to compare the aggressiveness of grooming methods.

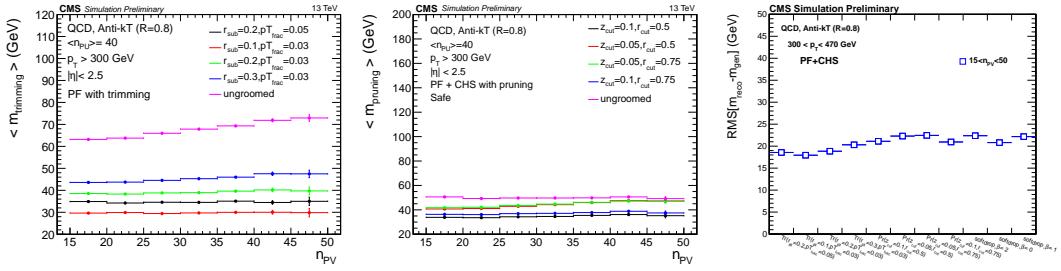


Figure 0.12: Variation of average jet mass vs. number of primary vertices for trimming (left), pruning (middle) and RMS in various grooming techniques (right).

the robustness of pileup mitigation algorithms. The plots of average jet mass vs. the number of primary vertices are shown in Fig. 0.12. The left plot shows the average jet mass is stable for trimmed jets but not for ungroomed jets. However, once the charged hadron subtraction is applied jet mass becomes stable considerably that can be seen from the middle plot. The right plot summarizes the results in terms of jet mass resolution from three different grooming techniques considered here.

Contents

0.7 Summary

We have presented the searches for SUSY and leptoquarks with 13 TeV pp collision data recorded with CMS. The SUSY search targets for direct gluino, stop and squark pair production in all hadronic final state. Two separate searches with 2.3 fb^{-1} of 2015 and 12.9 fb^{-1} of 2016 data yield null results of new physics. Similarly no signatures of any signal are found in the search of first generation scalar leptoquarks performed with 2.6 fb^{-1} of the 2015 data. Upper limits on the signal cross section are obtained in all three cases significantly extending the previous 8TeV limits.

Lastly, a study on advanced pileup mitigation techniques is presented. Performance of various jet grooming methods like trimming, pruning and soft-drop are compared. Optimized quality parameters of these techniques are recommended for future use.

Contents

List of Symbols

Symbols	Definition
LHC	Large Hadron Collider
CMS	Compact Muon Solenoid
PV	Primary vertex
\sqrt{s}	Center-of-mass-energy
MET	Missing transverse energy
Γ	Decay width
fb^{-1}	Inverse femtobarn (Unit of integrated luminosity)
GeV	Giga electron volt
TeV	Tera electron volt

Symbols	Definition
p_T	Momentum in the transverse plane
η	Pseudorapidity
ϕ	Azimuthal angle
j	Jet
J	Fat jet
PF jet	Particle-flow jet

In some parts of the thesis units of p_T and mass are GeV/c and GeV/c² respectively while in some other parts we used natural units by taking $\hbar = c = 1$.

List of Tables

2.1	Fermions in the SM with different generation.	5
2.2	Different interactions and their mediators in nature.	7
3.1	Field content of MSSM in terms of superfields(1st column), sparticles and SM particles(2nd and 3rd column) while in the 4th column the quantum numbers for various fields are shown.	19
3.2	Quantum numbers of scalar leptoquarks and possible renormalizable interactions containing them. We take the table from Ref. [39].	21
6.1	MC samples for signal SMS (simplified SUSY) model points.	44
6.2	SM $t\bar{t}$ MC samples used in the analysis. The cross sections are calculated to NNLO (next-to-next-to leading order).	44
6.3	SM QCD MC samples used in the analysis. All cross sections are calculated to LO.	45
6.4	SM $Z \rightarrow \nu\nu + \text{jets}$ MC samples used in the analysis. The cross sections are calculated to NNLO.	45
6.5	SM $W \rightarrow l\nu + \text{jets}$ MC samples used in the analysis. The cross sections are calculated to NNLO.	45

List of Tables

6.6 SM diboson and other rare process MC samples used in the analysis. The cross sections are calculated to NNLO.	46
6.7 SM DY+jets MC samples used in the analysis. The cross sections are calculated to NNLO.	46
6.8 SM γ +jets MC samples used in the analysis. The cross sections are calculated to LO.	46
6.9 Identification and isolation requirements for photons in the barrel and endcap. EA stands for effective area.	54
6.10 Effective areas used in pileup correction as a function of photon pseudorapidity. Three columns shows different values used for calculating $\text{Iso}_{\text{pfCh.}}$, $\text{Iso}_{\text{pfNu.}}$ and $\text{Iso}_{\text{pfGa.}}$, respectively, in various η bins.	54
6.11 $Z \rightarrow \nu\nu$ +jets prediction ($N_{Z \rightarrow \nu\nu}^{\text{pred}}$) in the 40 0-btag analysis bins as determined with the γ +jets method. We show the inputs to the determination of $N_{Z \rightarrow \nu\nu}^{\text{pred}}$ as that vary for the 40 bins, including N_{γ}^{obs} in EB and EE, $R_{Z/\gamma}$, and double ratio. The first uncertainty on $R_{Z/\gamma}$ comes from the statistical uncertainty on the simulation and the second from variation of the data-MC scale factors for photon reconstruction, ID, and isolation. .	71
7.1 HEEP electron identification criteria used for the analysis.	76
7.2 Optimized final selection thresholds for the $eejj$ analysis.	80
8.1 Summary of grooming parameters.	84

List of Figures

0.1	Event diagrams for the main new-physics scenarios considered in this study.	8
0.2	Signal vs. stacked backgrounds in H_T (HT, left) and H_T^{miss} (MHT, right)	12
0.3	Signal vs. stacked backgrounds in $N_{\text{jet}}(N\text{Jets}$, left) and $N_{b-\text{jet}}$ (MHT, right)	12
0.4	Physics ratio $R_{Z(\nu\bar{\nu})/\gamma}$ (left), numbers of observed events in the photon control samples in barrel(EB, middle) and endcap(EE, right) compared to simulation.	15
0.5	Data vs. the SM background before fit	20
0.6	The 95% CL upper limits on the production cross sections for four top quark(left),four light quark(middle) and four b-quark(right) in the final state.	20
0.7	Data vs. the SM background before fit.	23
0.8	The 95% CL upper limits on the production cross sections for four top quark (left),four light quark (middle) and four b-quark (right) in the final state.	23
0.9	Data vs. background plots for M_{ee} (left) and S_T (right).	24
0.10	The 95% CL upper limits on the production cross sections as a function of leptoquark mass.	26

List of Figures

0.11 The mass distributions of PFCHS jets after application of trimming (left), pruning (middle) and soft-drop (right) for different set of parameters. The ungroomed mass distribution is shown to compare the aggressiveness of grooming methods.	28
0.12 Variation of average jet mass vs. number of primary vertices for trimming (left), pruning (middle) and RMS in various grooming techniques (right).	28
 1.1 Looking deeper and deeper into matter, from atom to quark.	2
 2.1 Summary of the interactions among the SM particles. The plot is taken from Ref. [14].	6
2.2 Masses and other properties of SM particles. The plot is taken from Ref. [14]	7
2.3 The potential $V(\phi)$ for a complex scalar field $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ where $\mu^2 < 0$ and $\lambda > 0$ with a minimum at $v^2 = -\mu^2/\lambda$	14
 3.1 SM and their SUSY partners.	20
 4.1 An inside view of the CMS detector, showing various subdetectors that are placed around the beam pipe and form a series of cylindrical layers of the experiment. This image is taken from Ref. [44].	23
4.2 Various detector components that contribute to event reconstruction are (from left to right) tracker, electromagnetic calorimeter, hadron calorimeter, superconducting solenoid and muon system. The paths of different particles such as photons, muons, electrons, neutral and charged hadrons passing through the detector indicated by different solid or dashed color lines.	24
4.3 A schematic diagram of the CMS Tracker. The plot shows one quadrant of the tracking detector along the rz plane.	25
4.4 A schematic diagram of the CMS ECAL showing its different components.	27

List of Figures

4.5	Amount of material (in units of radiation lengths) upstream of the ECAL (left). ECAL energy resolution (right) as a function of pseudorapidity for electrons from $Z \rightarrow e^+e^-$ decays. The energy resolution estimated on Monte Carlo events is shown in red, while in blue and gray the resolution for, respectively, promptly reconstructed data and a later reconstruction of the same sample upon the usage of the best calibrations available are shown.	28
4.6	A schematic diagram of the CMS HCAL showing its individual components. . .	30
4.7	CMS muon system and position of muon subdetectors. The image is taken from Ref. [45].	33
6.1	Event diagrams for gluino pair production with four top final state (left), four b-quark final state (middle) and four light quark final state (right)	43
6.2	Signal vs. stacked backgrounds in H_T (HT, left) and H_T^{miss} (MHT, right). . .	50
6.3	Signal vs. stacked backgrounds in N_{jet} (NJets, left) and $N_{\text{b-jet}}$ (BJet, right). .	50
6.4	Search binning in the 2D phase-space of H_T and H_T^{miss}	51
6.5	Numbers of observed events in the photon control samples in barrel (left) and endcap (right) compared to simulation.	56
6.6	Data-MC agreement in terms search variables H_T (left) and H_T^{miss} (right) for the $\gamma+\text{jets}$ control sample.	57
6.7	Data-MC agreement in terms search variables N_{jet} (left) and $N_{\text{b-jet}}$ (right) for $\gamma+\text{jets}$ control region.	58
6.8	Distribution of σ_{inj} in bins of H_T (left), N_{jet} and H_T^{miss} for barrel photons. .	59

List of Figures

6.9 Fit results for purity determination in EB (top) and EE (bottom) for events with H_T^{miss} in the range 200-300, 300-500, and for $H_T^{\text{miss}} > 500$, respectively. β_γ is the fitted purity within the corresponding region. The red vertical line shows the σ_{inj} requirement for the loose ID working point. The non-prompt PDF comes from the QCD simulation.	60
6.10 $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ for the 18 analysis bins used in the analysis after baseline selection. The units of H_T and H_T^{miss} are in GeV.	61
6.11 The observed data vs the SM backgrounds in the 72 exclusive intervals of the signal region.	65
6.12 The 95% CL upper limits on the production cross sections for four top quark (left), four light quark (middle) and four b-quark (right) in the final state.	65
6.13 Search binning in the 2D plane of H_T and H_T^{miss} used in the 2016 analysis.	67
6.14 Numbers of observed events in the photon control samples in EB (left) and EE (right) compared to simulation. The units of H_T and H_T^{miss} are in GeV.	68
6.15 The observed data vs the SM backgrounds in the 160 exclusive intervals of the signal region.	70
6.16 The 95% CL upper limits on the production cross sections for four top quark (left), four light quark (middle) and four b-quark (right) in the final state.	70
7.1 Data vs. background plots for M_{ee} (left) and S_T (right) after Preselection	75
7.2 Data vs. MC for the LQ mass point 650 GeV in M_{ej}^{min} (left) and S_T (right).	78
7.3 The 95% CL upper limits on the production cross sections as a function of the leptoquark mass.	79

List of Figures

8.1	Pictorial representation of the pruning algorithm. The image is taken from Ref. [64].	83
8.2	Jet mass distributions for PF QCD jets for different grooming parameters.	85
8.3	Jet mass distributions for PF+CHS QCD jets for different grooming parameters.	85
8.4	Pileup dependence of the average jet mass for PF jets (left) and PF+CHS jets (right) for various grooming algorithms and parameters.	86
8.5	Comparison of jet mass resolution for PF and PF+CHS with different grooming algorithms and parameters.	87
9.1	The 95% CL upper limits on the gluino pair production cross sections for four top quark final state. The left plot is taken from Ref. [68]. The right plot is produced by us with 2.3 fb^{-1} of data.	89
9.2	The 95% CL upper limits on the leptoquark production cross section times branching fraction. The left plot is taken from Ref. [56] while the right plot is our result based on 2.6 fb^{-1} of data.	89

1 Introduction

Since we began doing modern science centuries ago, our efforts have been to understand the world around us in a logical and organized manner. Atoms, once thought to be the smallest, indivisible constituents of matter, was proven to be wrong in 1897 by J.J. Thomson who observed and identified the charged electrons in cathode rays. Since then one after the other remarkable experimental discoveries and the formulation of theoretical ideas to understand the physical phenomena have gone hand in hand culminating into a landmark framework namely the “Standard Model of Particle Physics”. Figure. 1.1 shows the sketch of the temporal evolution of fundamental constituents of matter. Today, the ‘Standard Model’ (SM) represents our best understanding about the interactions among these fundamental constituents. Even though the SM has been phenomenally successful in explaining a wide range of physical phenomena, it has several drawbacks. We have a number of laboratory and astronomical findings that are inconsistent with its predictions. These include the observation of nonzero neutrino masses, the presence of dark matter in the universe. Moreover, the SM suffers from the so-called hierarchy problem. More technically, it does not answer why the Higgs boson is so much lighter than the Planck mass. One would expect that the large quantum corrections to the square of the Higgs boson mass making it huge, comparable to the scale at which

1 Introduction

new physics appears, unless there is an incredible fine-tuning cancellation between the quadratic radiative corrections and the bare Higgs mass.

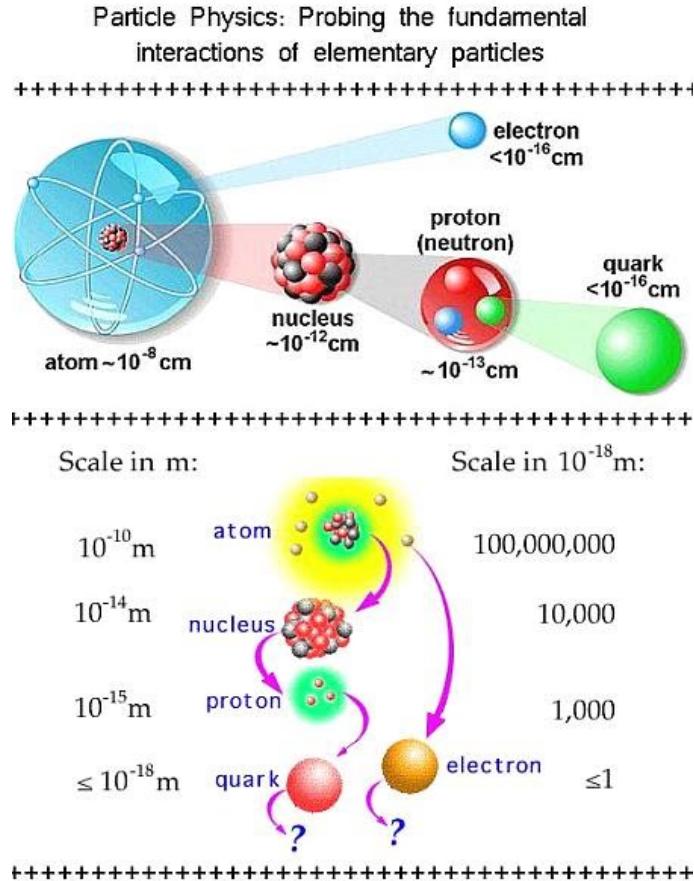


Figure 1.1: Looking deeper and deeper into matter, from atom to quark.

Many proposals going beyond the SM try to address these problems in different ways. Supersymmetry [1] is generally regarded as one of the likely extensions to the SM. The model is based on a unique way to extend the space-time symmetry group underpinning the SM, introducing a relationship between fermions and bosons.

The structure of the SM suggests a possible fundamental relationship between quarks and leptons. In some proposals beyond the SM, such as SU(5) grand-unification [2], Pati-Salam SU(4) [3], composite models [4], technicolor [5; 6; 7], and superstring-inspired

1 Introduction

models [8], the existence of a new symmetry relates the quarks and leptons in a fundamental way. These models predict the existence of a new class of bosons, called leptoquarks (LQs). They are colored particles, have fractional electric charge, can be either scalar or vector, and couple to leptons and quarks.

In the thesis, we present results of our searches for supersymmetry and leptoquarks using pp collision data recorded with the compact muon solenoid (CMS) detector at LHC, CERN. In these searches, jets play a very important role as they constitute a major part of the signal events. In the high luminosity environment i.e. during Run-2, the measurement of the properties of jets become extremely challenging due to large pileup. We study some of the advanced pileup mitigation techniques as part of the thesis.

We start with a brief account of theoretical background behind our study in Chapter 2. We discuss beyond-the-SM theories namely SUSY and LQ in Chapter 3. Relevant details of the CMS detector at LHC are discussed in Chapter 4.

To draw inferences from the observed data, the understanding of correct statistical procedures is crucial. We discuss briefly the LHC recommended statistical procedures in Chapter 5.

We devote Chapter 6 for an extensive report on the search for supersymmetry in all hadronic final state with 13 TeV data. Chapter 7 gives a comparatively brief description of the search analysis we carried out for first-generation scalar leptoquarks. We present results on the advanced pileup mitigation techniques in Chapter 8. We conclude with Chapter 9 by giving summary of the work and its significance in a broader context.

2 The Standard Model of Particle Physics

2.1 Introduction

The standard model (SM) of particle physics [9; 10] is a theory that describes the elementary particle constituents of matter and their interactions. These particles can be classified into two groups, *bosons* and *fermions* with integer and a half-integer spin, respectively. The bosons are responsible for various interactions while fermions mostly make up all of the visible matter in the universe.

The fermions come in two types: the *leptons*, which do not interact via the strong nuclear interaction, and the *quarks* that do interact via strong interactions. For each fermion there exists a corresponding antifermion which carries the same mass but opposite charge(including electric charge). Currently, there are six known leptons, and they are of two varieties. The electron, muon, and tau particles are electrically charged having a charge of $-1e$, while the corresponding three neutrinos namely electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ) are electrically neutral. Except for the difference in mass, the three charged leptons are identical with respect to how they

2 The Standard Model of Particle Physics

Table 2.1: Fermions in the SM with different generation.

Fermions	First Generation	Second Generation	Third Generation	Charge $Q/ e $
Quarks	up (u) down (d)	charm (c) strange (s)	top (t) bottom (b)	+2/3 -1/3
Leptons	electron (e) electron neutrino (ν_e)	muon (μ) muon neutrino (ν_μ)	tau (τ) tau neutrino (ν_τ)	-1 0

interact under the different fundamental forces. According to the way they interact via the weak interaction, the leptons are classified into three generations. Similar to leptons at present there are six known quarks, and they are also of two types. The up (u), charm (c), and top (t) quarks carry an electric charge of $+2/3e$, while the down (d), strange (s), and bottom (b) quarks have an electric charge of $-1/3e$. They are analogously classified into three generations. The fermions in the SM and their corresponding generations are listed in Table 2.1.

There are four fundamental forces in nature namely strong, weak, electromagnetic, and gravitation [11]. They govern interactions among the elementary particles. In our everyday life, we experience only the electromagnetic and gravitational force while the other two occur only at the subatomic or nuclear scale. The electromagnetic interaction is mediated by a massless photon (γ), while three massive gauge bosons (W^+ , W^- and Z^0) are the messengers of the weak interaction that is responsible for the nuclear beta decay. The strong interaction is mediated by eight massless gluons that bind quarks to form nucleons (protons and neutrons). Table 2.2 lists different forces and their mediators. No successful description of gravitation in terms of a quantum field theory exists to date, therefore it is not implemented in the SM. As it also has a minuscule impact at sub-atomic scales, it can be neglected in high energy physics experiments.

Among the fundamental bosons observed so far, the photon (γ) is massless, electri-

2 The Standard Model of Particle Physics

cally neutral, and acts as the only carrier of the electromagnetic force. The W bosons are massive, have $+1e$ or $-1e$ electric charge, and interact via the electromagnetic and weak force, but not via strong force. The Z boson is massive, electrically neutral, and interacts via the weak force, but does not interact via the electromagnetic or strong force. The gluon is massless, electrically neutral and does not interact via the electromagnetic or weak force, but does interact via the strong force. More precisely, there are eight different types of gluons carrying the eight different types of color charges (charges of the strong interaction). Finally, there is a massive ‘boson’ discovered by the CMS [12] and ATLAS [13] experiments at LHC having mass ~ 125 GeV and seems to be consistent with the SM Higgs boson. Figure 2.1 shows the interactions among all SM particles as well as with the Higgs boson (H) that gives mass to the weak-interaction mediators and other elementary particles. Fig. 2.2 shows masses and other properties of all SM particles.

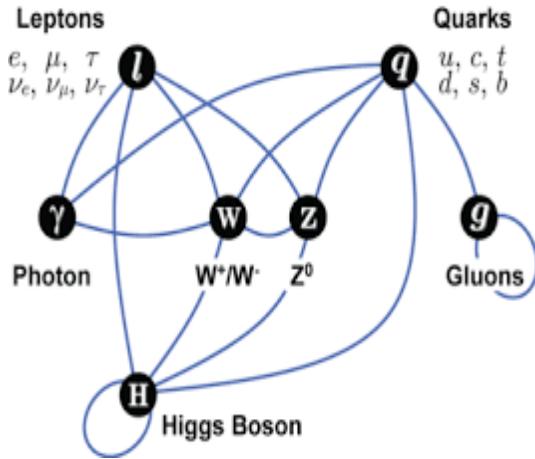


Figure 2.1: Summary of the interactions among the SM particles. The plot is taken from Ref. [14].

2 The Standard Model of Particle Physics

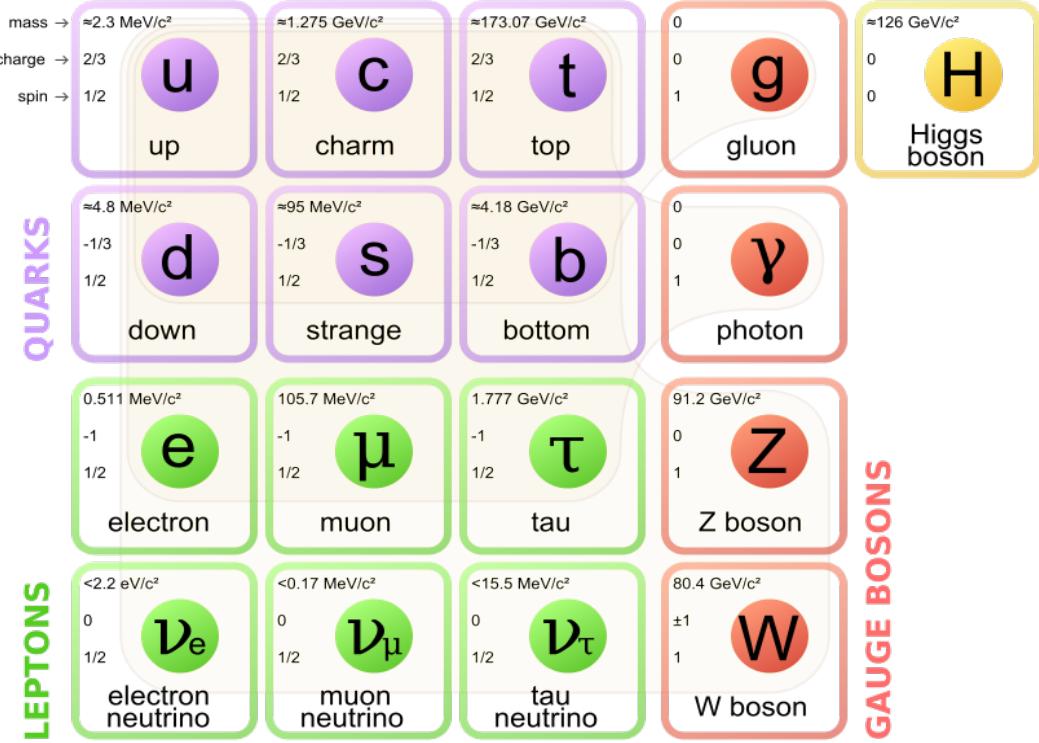


Figure 2.2: Masses and other properties of SM particles. The plot is taken from Ref. [14]

2.2 Gauge Theory Formulation

The SM is a gauge theory, based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which describes strong, weak and electromagnetic interactions, via the exchange of quanta of the corresponding spin-1 gauge fields. As described earlier, it describes the dynamics of six leptons, six quarks, eight carriers of strong force (gluons), weak vector bosons which mediate the weak interaction, and photon which mediates the electromagnetic

Table 2.2: Different interactions and their mediators in nature.

Interactions	Mediators
Strong	gluon (g)
Electromagnetic	photon (γ)
Weak	W^+ , W^- , and Z^0
Gravitational	graviton (not verified by experiment)

2 The Standard Model of Particle Physics

interaction. The three families of quarks and leptons are,

$$\begin{pmatrix} u & \nu_e \\ d & e \end{pmatrix}, \begin{pmatrix} c & \nu_\mu \\ s & \mu \end{pmatrix}, \begin{pmatrix} t & \nu_\tau \\ b & \tau \end{pmatrix}$$

where each quark appears in three different colors and each particle has its antiparticle. The part of the lagrangian that describes the dynamics of strong force is described by the symmetry group $SU(3)_C$ (the subscript C stands for color). We call this part as the quantum chromodynamics(QCD) sector. The $SU(3)_C$ invariant QCD lagrangian is given by

$$L_{\text{QCD}} = -\frac{1}{4}G_a^{\mu\nu}G_a^{\mu\nu} + \sum_f \bar{q}_f(i\gamma^\mu D_\mu - m_f)q_f, \quad (2.1)$$

where f runs over the quark flavors, a denotes the summation over eight gluons and q_f represents the quark field. $G_a^{\mu\nu}$ is defined as $G_a^{\mu\nu}(x) = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f^{abc} G_b^\mu G_c^\nu$ with G_a^μ being the gauge fields. $D^\mu = \partial^\mu + ig_s \frac{\lambda^a}{2} G_a^\mu(x)$ and is the covariant derivative term, where λ^a are the generators of the $SU(3)_C$ algebra. All interactions are expressed in terms of a single universal coupling g_s which we call the strong coupling constant. The types of interactions between quarks and gluons are followed from the above lagrangian. A large body of experimental evidences for QCD has been gathered over the years. The first evidence for quarks as real constituent elements of hadrons was obtained in deep inelastic scattering experiments at SLAC [15; 16]. The first evidence for gluons came in three jet events at PETRA [17].

Two important properties of strong force to be noted are :

Color confinement [18], which means that the force between (colored)quarks does

2 The Standard Model of Particle Physics

not diminish as they are separated. Because of this, when one tries to separate one quark from the other, the energy in the gluon field becomes strong enough to create another quark pair; they are thus forever bound into colorless hadrons such as protons and neutrons. Although analytically unproven, confinement is widely believed to be true as it explains the consistent failure of free quark searches.

Asymptotic freedom [18], which means that in very high-energy reactions(few GeV), quarks and gluons interact very weakly creating a quark-gluon plasma. This prediction of QCD was first proposed in the early 1970s by David Politzer, Frank Wilczek and David Gross [19; 20].

2.3 Electroweak Theory and Higgs Mechanism

Electroweak theory is a unified theory of electromagnetic and weak interactions. It is described by the symmetry group $SU(2)_L \otimes U(1)_Y$. The formulation of this theory takes into account a plethora of experimental observations that have been gathered over years. Parity symmetry is known to be maximally violated in weak interactions. This essentially means that the vector bosons do not interact with right-handed fermions, necessitating different representations within the gauge group for different chiral components. To successfully describe weak interactions, with several fermionic flavors and different properties for left- and right-handed fields, and with the requirement that the left-handed fermions should appear in doublets and we need massive gauge bosons and massless photon. The simplest group with doublet representations is $SU(2)$. To that we need to include also the electromagnetic interactions necessitating the introduction of an additional group. The symmetry group to consider is then $G = SU(2)_L \otimes U(1)_Y$, where L refers to left-handed fields and Y denotes hypercharge. We shall come back this later in the chapter. For time being let us introduce the left- and right-handed quark

2 The Standard Model of Particle Physics

fields as

$$\psi_1(x) = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \psi_2(x) = u_R, \quad \psi_3(x) = d_R. \quad (2.2)$$

We will treat the leptons in an identical fashion. The free lagrangian in terms of these fields would be

$$L_0 = \sum_{i=1}^3 i\bar{\psi}_i(x)\gamma^\mu\partial_\mu\psi_i(x). \quad (2.3)$$

L_0 is invariant under global G transformations in flavor space:

$$\psi_1(x) \rightarrow \psi'_1(x) = e^{iy_1\beta} U_L \psi_1(x)$$

$$\psi_2(x) \rightarrow \psi'_2(x) = e^{iy_2\beta} \psi_2(x)$$

$$\psi_3(x) \rightarrow \psi'_3(x) = e^{iy_3\beta} \psi_3(x)$$

where the $SU(2)_L$ transformation $U_L = e^{i\frac{\sigma^i}{2}\alpha^i}$ ($i = 1, 2, 3$) and y_i are hypercharges.

If we require the Lagrangian to be invariant under local gauge transformations we will need to introduce covariant derivatives in the following way

$$D_\mu\psi_1(x) = [\partial_\mu + ig\bar{W}_\mu(x) + ig'y_1B_\mu(x)]\psi_1(x)$$

$$D_\mu\psi_2(x) = [\partial_\mu + ig'y_2B_\mu(x)]\psi_2(x)$$

$$D_\mu\psi_3(x) = [\partial_\mu + ig'y_2B_\mu(x)]\psi_3(x)$$

where $\bar{W}_\mu = \frac{\sigma^i}{2}W_\mu^i(x)$. There are four different gauge fields, three $W_\mu^i(x)$ s and one $B_\mu(x)$ aimed to describe for four gauge bosons. Now the local gauge invariant Lagrangian would be

$$L = \sum_{i=1}^3 i\bar{\psi}_i(x)\gamma^\mu D_\mu\psi_i(x).$$

We note here that the above Lagrangian is without a mass term as the mass term

2 The Standard Model of Particle Physics

will break the symmetry as left and right components have different transformation properties. It also does not have the kinetic term for the gauge fields. In order to write the gauge-invariant kinetic term for the gauge fields, we introduce the gauge field tensors as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \text{ and}$$

$$\bar{W}_{\mu\nu} = \partial_\mu \bar{W}_\nu - \partial_\nu \bar{W}_\mu + ig[W_\mu, W_\nu] \text{ leading to}$$

$$L_{\text{kin}} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_i^iW_i^{\mu\nu}$$

So the total electroweak Lagrangian is given by $L_{\text{tot}} = L + L_{\text{kin}}$. Still the Lagrangian does not have mass terms for the bosons and fermions. In order to accomplish that we will have to break the symmetry of the lagrangian, so that the gauge bosons (W^+, W^-, Z) become massive and the photon remains massless.

In the quantum field theory, a symmetry is ‘spontaneously’ broken when the Lagrangian itself remains invariant with the Hamiltonian of the theory attaining its minimum. In the context of the electroweak theory, spontaneous symmetry breaking is achieved through the introduction of a complex scalar field which attains a nonzero vacuum expectation value (VEV) [21].

Let us first consider a complex scalar field $\phi(x)$ and the lagrangian

$$L = (\partial_\mu \phi^\dagger)(\partial^\mu \phi) - V(\phi) \quad (2.4)$$

$$V(\phi) = \mu^2 \phi^\dagger \phi + h(\phi^\dagger \phi)^2 \quad (2.5)$$

L is invariant under the global phase transformation of the scalar field. The type of minima of the lagrangian will be different depending on the value of μ . We can have two cases here

- $\mu^2 > 0$, for this the potential has trivial minimum and it describes a massive scalar

2 The Standard Model of Particle Physics

particle with mass μ and quartic coupling h

- $\mu^2 < 0$, in this scenario the minimum of the field is given by

$|\phi_0| = \sqrt{\frac{-\mu^2}{2h}} \equiv \frac{v}{\sqrt{2}} > 0$ or $V(\phi_0) = -\frac{h}{4}v^4$. Now there are infinite states of degenerate states of minimum energy with solutions given by $\phi_0 = \frac{v}{\sqrt{2}}e^{i\theta}$. If we choose $\theta = 0$, the symmetry gets spontaneously broken. Now parameterizing the excitations over the ground state as

$\phi(x) \equiv \frac{1}{\sqrt{2}}[v + \varphi_1(x) + \varphi_2(x)]$, where φ_1 and φ_2 are real fields, the potential takes the form

$$V(\phi) = V(\phi_0) - \mu^2 \varphi_1^2 + h v \varphi_1 (\varphi_1^2 + \varphi_2^2) + \frac{h}{4} (\varphi_1^2 + \varphi_2^2)^2 \quad (2.6)$$

Here φ_1 describes a massive particle with mass $-2\mu^2$ and φ_2 is massless. These bosons are called goldstone bosons. In fact, this is quite general result and follows from the Goldstone theorem [23]. At first look, the Goldstone theorem does little to our mass problem. But the problem could be solved through somewhat related mechanism called the Higgs mechanism [21; 22].

2.3.1 Higgs Mechanism

To discuss the Higgs mechanism, let us consider an $SU(2)_L$ doublet of complex scalar fields

$$\phi(x) = \begin{pmatrix} \phi^{(+)}(x) \\ \phi^{(0)}(x) \end{pmatrix}_L \quad (2.7)$$

Now the gauge invariant lagrangian becomes

2 The Standard Model of Particle Physics

$$L_S = (D_\mu \phi)^\dagger D_\mu \phi - \mu^2 \phi^\dagger \phi - h(\phi^\dagger \phi)^2 \quad (2.8)$$

is invariant under local gauge transformations. Here

$$D^\mu \phi = [\partial^\mu + ig \bar{W}^\mu + ig' y_\phi B^\mu] \phi \quad (2.9)$$

with the vacuum expectation value of ϕ^0 is given by

$$\text{VEV}(\phi^0) = \sqrt{\frac{-\mu^2}{2h}} \equiv \frac{v}{\sqrt{2}} \quad (2.10)$$

Once we choose a particular ground state the symmetry gets spontaneously broken. According to the Goldstone theorem, the mass of three massless states will appear. Now parameterizing the scalar doublet in the general form

$$\phi(x) = e^{i\frac{\sigma_i}{2}\theta^i(x)} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}_L \quad (2.11)$$

where $\theta^i(x)$ and $H(x)$ are four real fields. Now taking a physical gauge $\theta^i = 0$ the kinetic part of the Lagrangian takes the form

$$(D_\mu \phi)^\dagger D_\mu \phi \rightarrow \frac{1}{2} (\partial_\mu H)(\partial^\mu H) + (v + H)^2 \left(\frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2(\theta_W)} Z^\mu Z_\mu \right) \quad (2.12)$$

We could see from the above term that the gauge bosons have acquired mass. The relationship of their masses is given by

$$M_Z \cos(\theta_W) = M_W = \frac{1}{2} v g \quad (2.13)$$

2 The Standard Model of Particle Physics

This is how the Higgs mechanism is responsible for giving masses to vector bosons. Fermions also get masses by Higgs through Yukawa interactions.

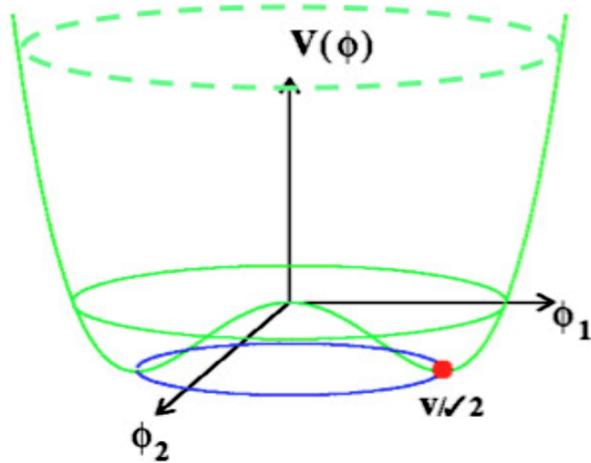


Figure 2.3: The potential $V(\phi)$ for a complex scalar field $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ where $\mu^2 < 0$ and $\lambda > 0$ with a minimum at $v^2 = -\mu^2/\lambda$.

2.4 Experimental Verification of SM

The SM has been phenomenally successful in predicting existence of several particles. All of them have been observed experimentally. The Higgs mechanism has a precise prediction for the W^\pm and Z masses, relating them to the vacuum expectation value of the scalar field through Eq(2.13). It tells M_Z should be bigger than M_W which is experimentally verified [24; 25].

$$M_Z = 91.1875 \pm 0.0021 \text{GeV}, M_W = 80.398 \pm 0.025 \text{GeV} \quad (2.14)$$

The top quark, being the heaviest one among the known elementary particles, constitutes an important ingredient for precision electroweak tests as well as for an indirect

2 The Standard Model of Particle Physics

determination of the Higgs boson mass. It was observed for the first time at Tevatron [26]. The recent world-average top mass, driven LHC measurements is $173.34 \pm 0.27(\text{stat.}) \pm 0.71(\text{syst.})\text{GeV}$ [27].

Another important cornerstone of the SM is the spontaneous electroweak symmetry breaking mechanism proposed almost 50 years ago by Higgs, Brout, Englert, Guralnik, Hagen and Kibble to generate mass for the fermions and gauge bosons [21]. This last missing piece to the jigsaw, the Higgs boson, was discovered by the CMS [12] and ATLAS [13] experiments at LHC in 2012. The first observation of the particle was reported in its high resolution decay channels to $H \rightarrow \gamma\gamma$ and 4-lepton(via ZZ^*) with a local significance of 5σ .

2.5 Drawbacks of the Standard Model

Even though the SM has been very successful in describing many aspects, it cannot be the complete theory due to its inability to explain several experimental observations and mathematical flaws. Some of them can be summarized as below

- The SM does not explain gravitation. It is unable to successfully marry theory of gravitation, general relativity with quantum field theory. One of the implications is that plausible quantum field theories of gravity break down before reaching the Planck scale. As a result, we do not have a theory of early universe.
- The SM has 19 independent parameters : 9 Yukawa couplings, 3 CKM angles and 1 CP violating phase, 3 coupling constants, the μ and λ parameters of the Higgs potential, and QCD vacuum angle θ_{QCD} linked to the strong CP problem. It is natural to consider the SM as a low-energy effective theory, far from a fundamental theory.

2 The Standard Model of Particle Physics

- Experiments show neutrinos have mass while the SM does not allow it. In fact, the framework of the SM is based on mass less neutrinos.
- Dark energy is an unknown form of energy that is hypothesized to permeate all of space, tending to accelerate the expansion of the universe. Several astrophysical observations suggest that universe contains dark matter [28]. The SM does not explain the existence of either dark energy or dark matter.
- The SM also fails to explain preponderance of matter over antimatter. The CP violation content of the SM falls short by several orders of magnitude in explaining the observed matter-antimatter asymmetry.
- The Higgs mechanism in the SM gives rise to hierarchy problem, which is, the large discrepancy between the scale of weak interaction and gravity. It is not understood why the former is 10^{24} times stronger than gravity. More technically, why the Higgs boson is so much lighter than the Planck mass. In the context of the SM, the self-energy correction to the Higgs boson mass should diverge.

3 Beyond the Standard Model

Physics beyond the standard model (BSM) [29] refers to the theoretical proposals put forward to explain deficiencies of the SM. BSM models include various supersymmetric extensions [30], such as the Minimal Supersymmetric Standard Model (MSSM) [31] and Next-to-Minimal Supersymmetric Standard Model (NMSSM) [32], Leptoquark models or entirely novel explanations deriving from string theory such as extra dimensions [33]. In the following sections we will discuss about supersymmetry (SUSY) and Leptoquark models.

3.1 Supersymmetry

Supersymmetry (SUSY) is one of the proposed extensions to the SM that introduces a new kind of space-time symmetry and relates two classes of elementary particles: bosons and fermions. In this model, each SM particle has its so called superpartner differing by spin $\frac{1}{2}$. If supersymmetry were unbroken, all the SUSY partners would have the same mass as the corresponding SM particles. The fact that no SUSY particles have been found, tells us that it must be a broken symmetry. SUSY solves many outstanding problems in particle physics including the hierarchy problem. The simplest realization

3 Beyond the Standard Model

of the broken SUSY is MSSM. As described earlier, a supersymmetric transformation converts a bosonic state to a fermionic state and vice versa. If \mathbf{Q} denotes such an operator then

$$\mathbf{Q}|\text{Boson}\rangle = |\text{Fermion}\rangle, \mathbf{Q}|\text{Fermion}\rangle = |\text{Boson}\rangle$$

with \mathbf{Q} being an anti-commuting spinor. The supersymmetric partner of fermion is called a sfermion. For instance, quarks have squarks and leptons have sleptons. The left- and right-handed components of the quarks and leptons are distinct two-component Weyl fermions with different gauge transformation properties in the SM, so each must have its own complex scalar partner. The symbols for squarks and sleptons are the same as for the corresponding fermion, but with a tilde \sim used to denote the superpartner of a SM particle. As an example, the superpartners of the left- and right-handed components of the electron Dirac field are called left- and right-handed selectrons, and are denoted \tilde{e}_L and \tilde{e}_R respectively. Here, the handedness does not refer to the helicity of selectrons. As the neutrinos are always left-handed they are denoted just by $\tilde{\nu}_e, \tilde{\nu}_\tau$ and $\tilde{\nu}_\mu$. Similarly, for the quarks they are given by \tilde{q}_L and \tilde{q}_R , where q stands for six types of quarks. Theories show that the Higgs cannot reside in a single supermultiplet. So there are two Higgs supermultiplets with $Y=\pm\frac{1}{2}$. We will call the $SU(2)_L$ -doublet complex scalar fields with $Y=+\frac{1}{2}$ and $Y=-\frac{1}{2}$ by the names H_u and H_d , respectively. The superpartners of Higgs are called higgsinos. The weak isospin components of H_u with $T_3 = (\frac{1}{2}, -\frac{1}{2})$ have electric charges 1 and 0, respectively, and are denoted by (H_u^+, H_u^0) . The neutral scalar that corresponds to the physical SM Higgs boson is in a linear combination of H_u^0 and H_d^0 . Higgsinos are denoted as \tilde{H}_u , \tilde{H}_d for the $SU(2)_L$ doublet left-handed Weyl spinor fields, with weak isospin components $(\tilde{H}_u^+ \text{ and } \tilde{H}_u^0)$ and $(\tilde{H}_d^0 \text{ and } \tilde{H}_d^-)$. Similarly the superpartner of gluon(g) is gluino (\tilde{g}). The electroweak gauge symmetry $SU(2)_L \otimes U(1)_Y$ is associated with spin-1 gauge bosons W^\pm , W^0 and B^0 with spin $\frac{1}{2}$ superpartners \tilde{W}^\pm , \tilde{W}^0 and \tilde{B}^0 , called winos and bino. After electroweak symmetry breaking the W^0 and

3 Beyond the Standard Model

B^0 eigenstates mix to give mass eigen states Z^0 and γ . The corresponding gaugino mixtures of \tilde{W}^0 and \tilde{B}^0 are called zino (\tilde{Z}^0) and photino ($\tilde{\gamma}$). A list of SUSY and SM particles in the context of MSSM is given in Table. 3.1. They are also pictorially shown in Fig. 3.1.

Table 3.1: Field content of MSSM in terms of superfields(1st column), sparticles and SM particles(2nd and 3rd column) while in the 4th column the quantum numbers for various fields are shown.

Supermultiplet	S=0 or 1	S=1/2	$SU(3), SU(2)_L, U(1)_Y$
	squarks	quarks	
Q	$(\tilde{u}_L, \tilde{d}_L)$	(u_L, d_L)	$(3, 2, 1/6)$
\bar{U}	\tilde{u}_R^*	u_R^\dagger	$(\bar{3}, 1, -2/3)$
\bar{D}	\tilde{d}_R^*	d_R^\dagger	$(\bar{3}, 1, 1/3)$
	sleptons	leptons	
L	$(\tilde{\nu}_L, \tilde{e}_L)$	(ν_L, e_L)	$(1, 2, -1/2)$
\bar{E}	\tilde{e}_R^*	e_R^\dagger	$(1, 1, 1)$
	Higgs	Higgsino	
H_u	(H_u^+, H_u^0)	$(\tilde{H}_u^+, \tilde{H}_u^0)$	$(1, 2, 1/2)$
H_d	(H_d^0, H_d^-)	$(\tilde{H}_d^0, \tilde{H}_d^-)$	$(1, 2, -1/2)$
	gauge bosons	gaugino	
$V^{(y)}$	B_μ	\tilde{B}_μ	$(1, 1, -)$
$V^{a(2)}$	W_μ^a	\tilde{W}_μ^a	$(3, 1, -)$
$V^{a(3)}$	g_μ^a	\tilde{g}_μ^a	$(8, 1, -)$

In the framework of MSSM with R-parity [34] conservation, SUSY particles are produced in pairs. At LHC, the most abundantly produced SUSY particles are expected to be strongly interacting partners of quarks and gluons, i.e. the squarks and gluinos. The dominant gluino and squark production processes are

$$pp \rightarrow \tilde{g}\tilde{g} + X, \tilde{q}\tilde{q}, \tilde{q}\tilde{q}^*, \tilde{q}\tilde{q}^*$$

In our SUSY analysis, we mainly focus on gluino pair production because it has largest cross section for squark decoupling simplified scenarios [35]. Simplified SUSY models, also called “simplified model space” (SMS), try to reduce the complex parameter

SUPERSYMMETRY

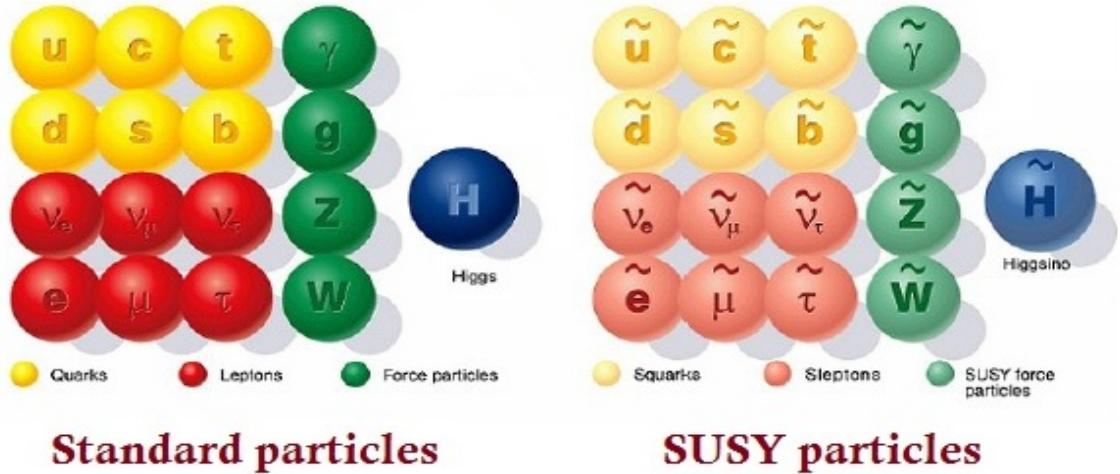


Figure 3.1: SM and their SUSY partners.

space of SUSY to a minimal number so that it would be easier for the experimentalists to interpret data and derive limits on the cross sections. The SMS models used in our analysis have two free parameters, namely the gluino and lightest SUSY particle (LSP) mass. Other particles are considered to be too heavy to participate in the process. [36; 35; 37]

3.2 Leptoquark models

Leptoquarks are scalar or vector particles that carry information back and forth between quarks and leptons of a given generation allowing them to interact. They are color-triplet bosons and carry both lepton and baryon numbers. They appear in various extensions of the SM, such as technicolor [6; 7; 5] or grand unified theories based on Pati-Salam model [3], SU(5) [70] or E6 [8] or composite models [4]. Leptoquarks explain why there are three generations of matter. Therefore, they provide very useful information to either falsify or exclude the parameter space of BSM theories. In the search we will be describing later, we look for first generation scalar leptoquarks. We list different scalar leptoquark models in Table 3.2. We follow the notation from Ref. [38].

Table 3.2: Quantum numbers of scalar leptoquarks and possible renormalizable interactions containing them. We take the table from Ref. [39].

Leptoquark	Renormalizable couplings	$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$
S_0	$S_0 Q_L^\dagger, S_0 u_R^\dagger e_R^\dagger, S_0 Q_L Q_L$	$(\mathbf{3}, \mathbf{1}, -1/3)$
\tilde{S}_0	$\tilde{S}_0 d_R^\dagger e_R^\dagger, \tilde{S}_0 u_R u_R$	$(\mathbf{3}, \mathbf{1}, -4/3)$
S_1	$S_1 Q_L^\dagger, L_L^\dagger, S_1 Q_L Q_L$	$(\mathbf{3}, \mathbf{3}, -1/3)$
$S_{1/2}^\dagger$	$S_{1/2}^\dagger Q_L^\dagger e_R, S_{1/2}^\dagger u_R^\dagger L_L$	$(\mathbf{3}, \mathbf{2}, +7/6)$
$\tilde{S}_{1/2}^\dagger$	$\tilde{S}_{1/2}^\dagger d_R^\dagger L_L$	$(\mathbf{3}, \mathbf{2}, +1/6)$

4 The Compact Muon Solenoid

4.1 CMS detector

The Compact Muon Solenoid (CMS) [40] is one of the general purpose detectors at the Large Hadron Collider (LHC) [41], CERN near Geneva, Switzerland. The LHC is considered to be the most efficient and controlled instrument with which one can probe fundamental physics of high energy collisions. Its main job is to accelerate two energetic beams of charged particles (protons or lead ions) and to let them collide at few well defined points. The CMS detector sits at one point and records all the possible interesting events that results from these collisions. It is 21.6 m long, 15 m in diameter, and weighs about 14 kilotons. Approximately 3,800 people, representing 199 scientific institutes from 43 countries, form the CMS collaboration who built and now operate the gigantic detector. It comprises of subsystems, well designed to measure the energy and momentum of photons, electrons, muons, and other products of the collisions. It is located in an underground cavern at Cessy in France, just across the border from Geneva. Studies presented in this thesis are based on proton-proton collision data collected by the CMS experiment at \sqrt{s} of 13 TeV. Two pictures of the detector are shown in Figs. 4.1 and 4.2. More details about the CMS detector can be found from the Ref. [42; 43].

4 The Compact Muon Solenoid

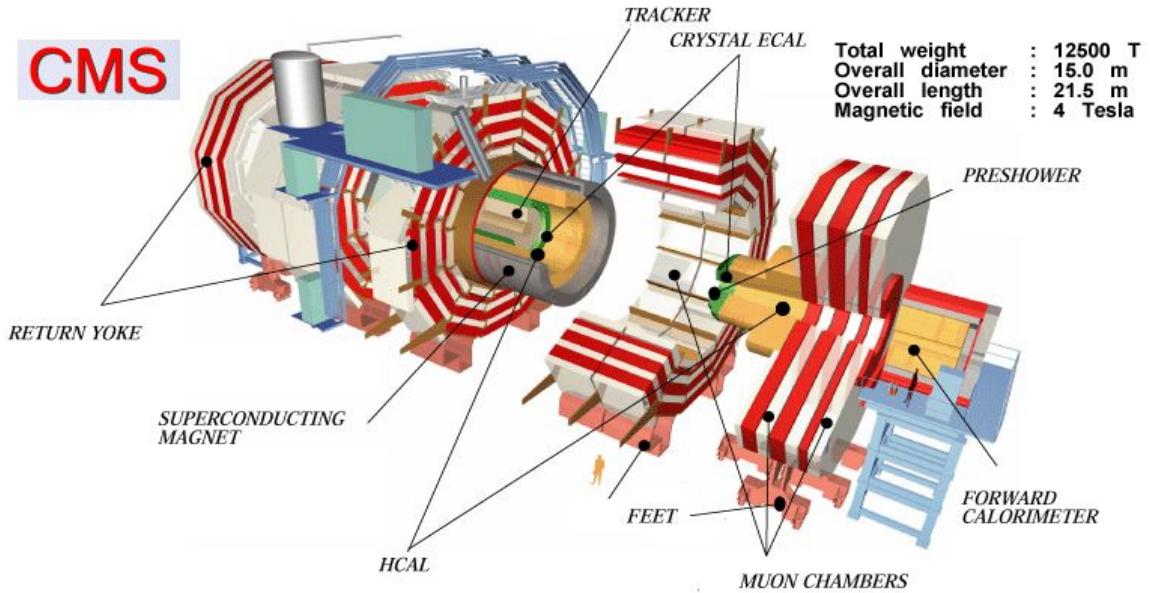


Figure 4.1: An inside view of the CMS detector, showing various subdetectors that are placed around the beam pipe and form a series of cylindrical layers of the experiment. This image is taken from Ref. [44].

CMS has four major subcomponents, namely

- Tracker,
- Electromagnetic calorimeter
- Hadron calorimeter, and
- Muon system.

We shall now explain the individual components of the detector one by one.

4.1.1 Tracker

The innermost part of the CMS detector is a silicon tracking system places at an approximate distance of 10 cm from the IP. It infers the momentum of charged particles by measuring their bending in presence of the magnetic field. It also plays an important

4 The Compact Muon Solenoid

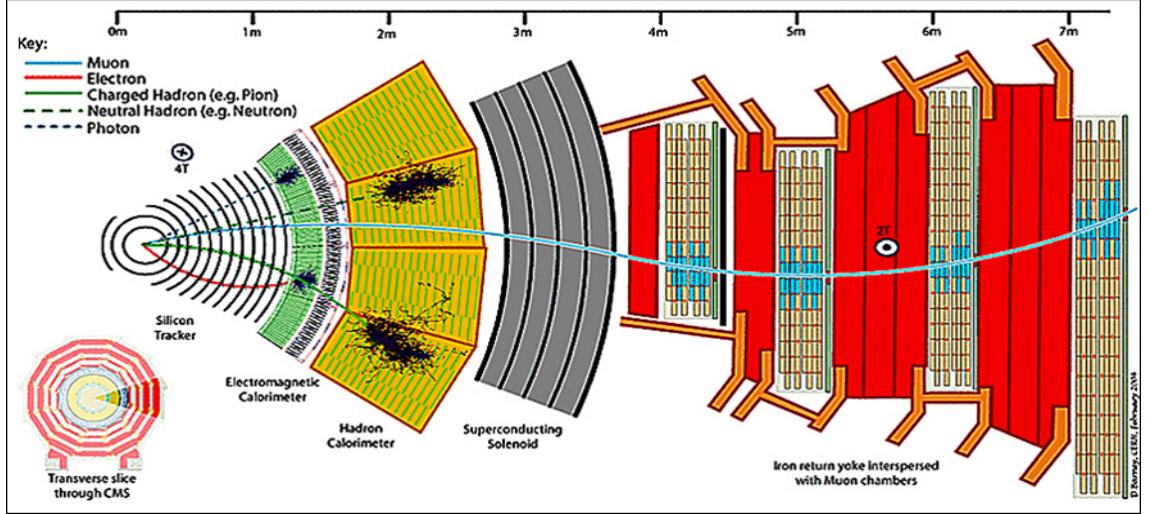


Figure 4.2: Various detector components that contribute to event reconstruction are (from left to right) tracker, electromagnetic calorimeter, hadron calorimeter, superconducting solenoid and muon system. The paths of different particles such as photons, muons, electrons, neutral and charged hadrons passing through the detector indicated by different solid or dashed color lines.

role in identifying the number of primary vertices present in each event and helps reduce the pileup contamination in the event. The Figure. 4.1 shows the position of tracker in the CMS detector. It is 5.8 m long and 2.5 m in diameter, centered around the interaction point (IP). It is designed to provide a robust, efficient and precise measurement of the trajectories of charged particles. Tracker consists of two main subcomponents namely pixel and microstrip detector. The tracker covers up to $|\eta| < 2.5$. The pixel detector helps to detect the secondary vertices and hence to identify jets initiated by a b-quark (b-jets). On the other hand, the microstrip detector is a key to the momentum measurement.

The central part of the tracking detector is called barrel which has a cylindrical shape with the axis of symmetry coinciding with the LHC beamline and consists of three layers. In the forward and backward regions there are two endcaps each comprising two layers of pixels. The pixel detector consists of in total 66 million pixels. It is capable of detecting particles within the pseudorapidity range $|\eta| \leq 2.4$ with average spatial resolution of 10

4 The Compact Muon Solenoid

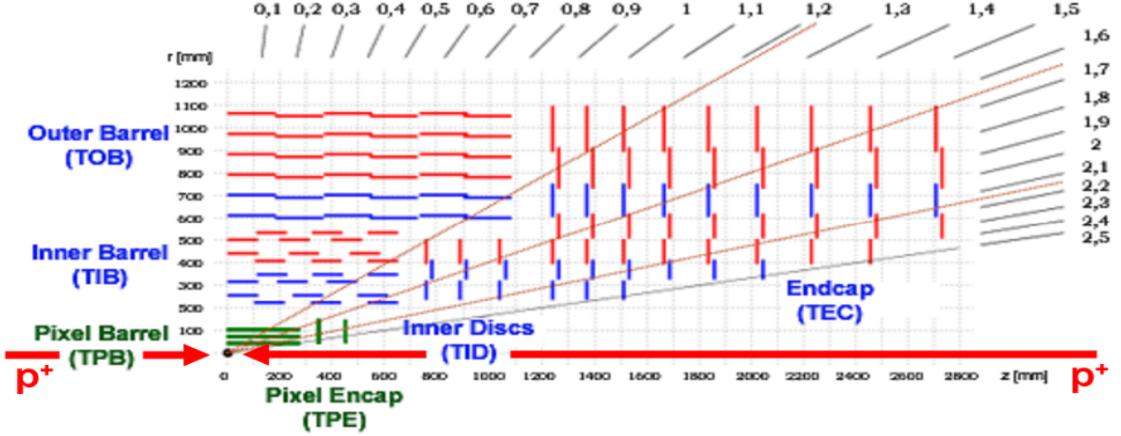


Figure 4.3: A schematic diagram of the CMS Tracker. The plot shows one quadrant of the tracking detector along the $r\phi$ plane.

μm along the transverse $r\phi$ plane and $20 \mu m$ along the z direction.

As the distance from the IP increases the particle flux decreases and tracking job is done by a silicon microstrip detector that surrounds the pixel detector. It consists of 10 barrel layers (4 TIB-Tracker Inner Barrel and 6 TOB (Tracker Outer Barrel), 3 TID (Tracker Inner Disks) and 9 TEC (Tracker Endcap) disks. A schematic diagram of the cross sectional slice of one quarter is shown in Fig. 4.3. More details about tracker can be found in Ref. [42].

4.1.2 Electromagnetic Calorimeter

Calorimeters (energy measuring devices) can be divided into two categories: homogeneous if the whole detector volume is active , and sampling, if the detector mostly comprises a passive absorber with only a fraction consisting of active volume. The CMS electromagnetic calorimeter (ECAL) is a homogeneous type that provides excellent energy and position measurements for photons and electrons. The active materials of the ECAL are lead tungstate ($PbWO_4$) crystals providing around 25 radiation lengths(X_0). A particle entering the ECAL results in an electromagnetic shower caused by the suc-

4 The Compact Muon Solenoid

cessive processes of bremsstrahlung and pair production. At the end, the energy of the particle is deposited in the calorimeter material via ionization and photoelectric effect. An electromagnetic shower is the process through which an energetic electromagnetic particle interacts with matter generating a cascade process composed of large number of secondary particles (photons, electrons and positrons). This occurs because electrons and positrons having energy greater than 1 GeV lose energy mostly via bremsstrahlung while photons via pair production. The secondary particles produced by the primary ones interact through the same process leading to the development of a particle shower inside the calorimeter.

The CMS ECAL consists of 75848 $PbWO_4$ crystals. The radiation length of this crystal is 0.89 cm. The lateral spread of the shower depends on the Moliere radius ($R_M = 2.2\text{cm}$) of the material. The light emission is fast enough to work with the highest LHC bunch crossing rate of 25 ns.

The ECAL has two parts : barrel (EB) extending up to $|\eta| < 1.479$ and ECAL endcap (EE) covering $|\eta|$ from 1.479 to 3.0. For a better discrimination of photons against neutral pions, a preshower device(ES) is mounted in front of EE. In ES silicon millistrip sensors with a resolution of 2mm, placed behind two planes of lead, are able to distinguish single photons from π^0 s decaying into photon pairs.

4 The Compact Muon Solenoid

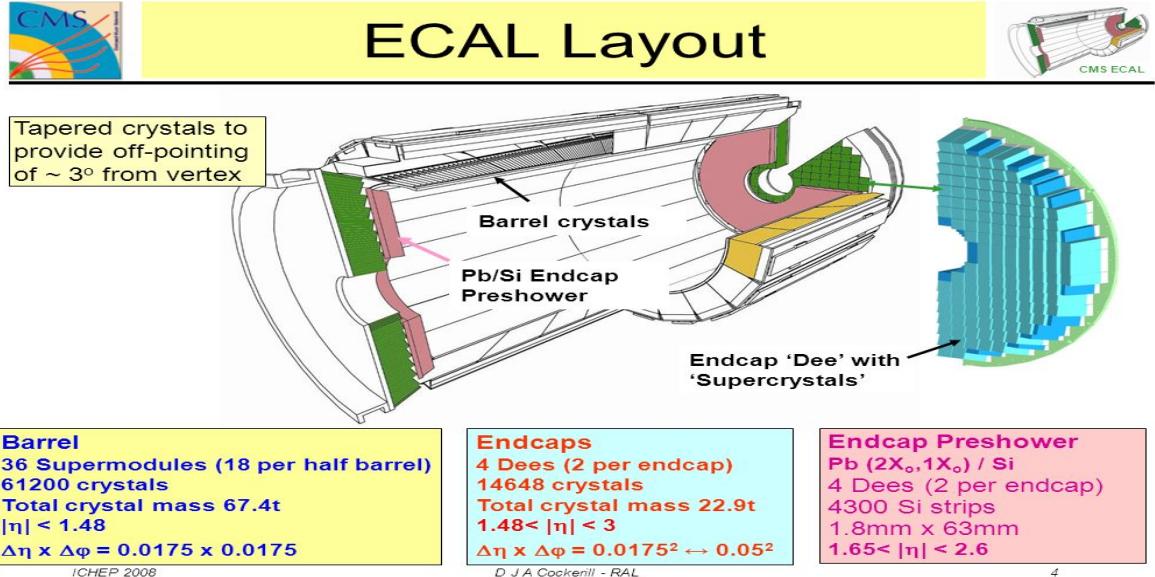


Figure 4.4: A schematic diagram of the CMS ECAL showing its different components.

ECAL Performance

The energy resolution of the CMS ECAL is given by the following expression

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \quad (4.1)$$

where E is the energy in GeV, $S=2.8\%$ is the stochastic term, $N= 124$ MeV is the noise term, and $C = 0.3\%$ is the constant term. Various sources of contributions to these three terms as follows.

Stochastic term

- event-to-event fluctuations in the lateral shower containment
- a photo statistics contribution of 2.1%
- fluctuations in the energy deposited in the ES absorber with respect to what is measured in the preshower silicon detector.

Constant term

- non-uniformity of longitudinal light collection
- intercalibration errors

4 The Compact Muon Solenoid

- leakage of the energy from the back of the crystal

Noise term

- electronics noise
- digitization noise
- pileup noise

The ECAL material budget and energy resolution vs. pseudorapidity is shown in Fig. 4.5.

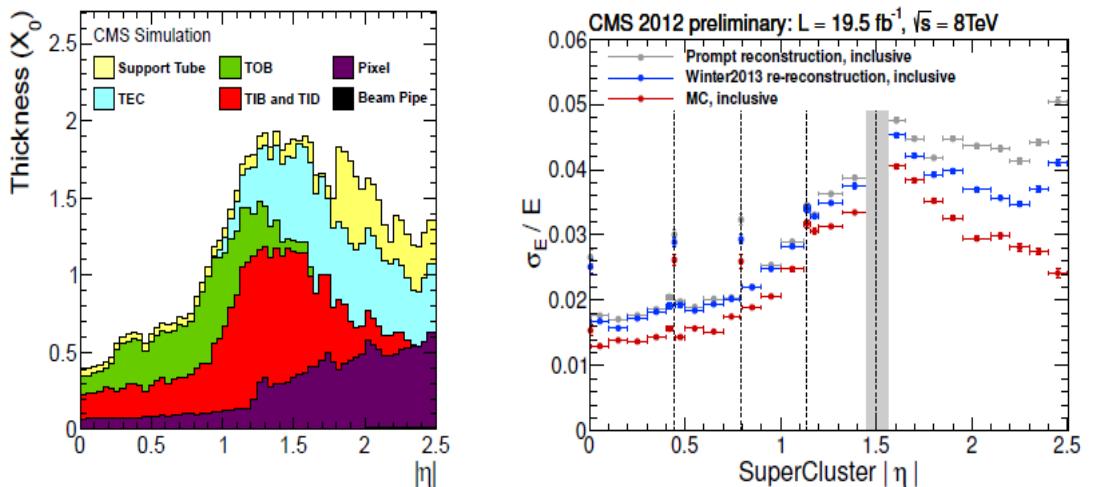


Figure 4.5: Amount of material (in units of radiation lengths) upstream of the ECAL (left). ECAL energy resolution (right) as a function of pseudorapidity for electrons from $Z \rightarrow e^+e^-$ decays. The energy resolution estimated on Monte Carlo events is shown in red, while in blue and gray the resolution for, respectively, promptly reconstructed data and a later reconstruction of the same sample upon the usage of the best calibrations available are shown.

4.1.3 Hadron Calorimeter

The Hadron Calorimeter (HCAL) measures the energy of hadrons, particles made of quarks and gluons, for example protons, neutrons, pions and kaons. Additionally, it

4 The Compact Muon Solenoid

provides an indirect measurement for non-interacting, uncharged particles such as neutrinos. The HCAL consists of layers of dense material (brass or steel) interleaved with tiles of plastic scintillators, which are read out via wavelength-shifting fibres by hybrid photodiodes. This combination was decided to allow the maximum amount of absorbing material inside of the magnet coil. The high pseudorapidity region ($3.0 < |\eta| < 5.0$) is instrumented by the Hadronic Forward (HF) detector. It uses a slightly different technology of steel absorbers and quartz fibres for readout, designed to allow a better separation of particles in the congested forward region. The HF is also used to measure online, the relative luminosity system in CMS.

The HCAL is not only used to measure single hadrons but also jets with good precision. Jets are complex objects made of a mixture of particles such as hadrons, electrons, photons and muons, originating from the interaction of partons during hadron-hadron collisions. Achieving good precision for their energy scale and resolution is of crucial importance for many physics analyses and poses a big challenge for the HCAL design. The CMS HCAL is a sampling calorimeter surrounding the Traker and ECAL. It reconstructs both charged and neutral hadrons produced during pp collisions. Brass has been chosen as the absorber material for the HCAL as it is non-magnetic and has a relatively short interaction length of $\lambda_I = 16$ cm. The large fraction of passive material is cost efficient. The HCAL has four main subcomponents as shown in Fig. 4.6

The subcomponents are Hadron Barrel (HB), Hadron Outer (HO), Hadron Endcap (HE) and Hadron Forward (HF). HB is the inner part of the HCAL barrel with $1.305 < |\eta| < 1.392$. HO helps to measure the energy of hadron showers penetrating the magnetic coil. It extends up to $|\eta|=1.26$. HE has 14 additional calorimeter towers covering the pseudorapidity region $1.3 < |\eta| < 3.0$. To cover higher η range $2.8 < |\eta| < 5.2$, HF is located at $z = \pm 11.2$ m from the IP close to the beam pipe. Jets with very high $|\eta|$ values as well as the hadronization products of beam remnants are detected with this

4 The Compact Muon Solenoid

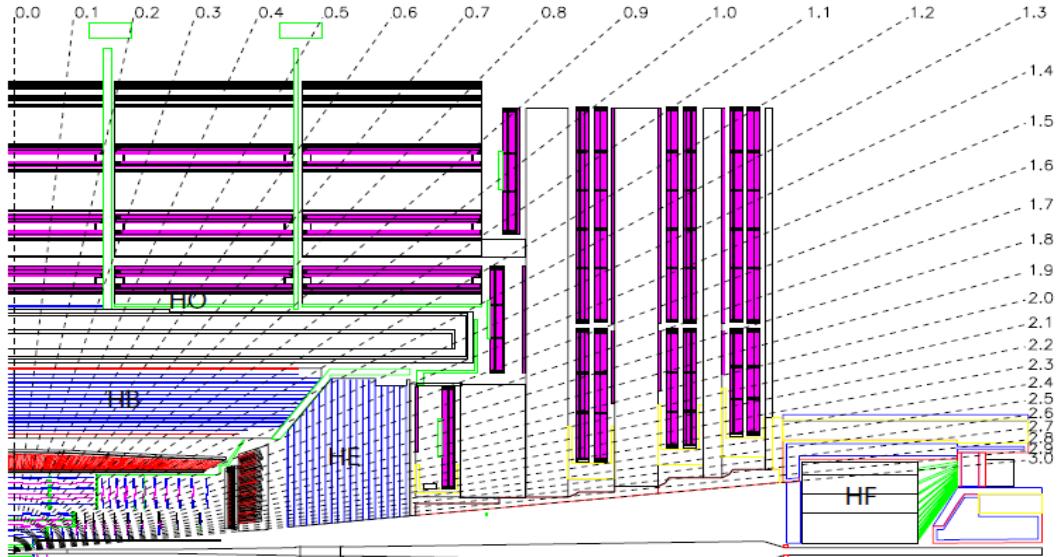


Figure 4.6: A schematic diagram of the CMS HCAL showing its individual components.
system.

HCAL Performance

The hadron energy resolution for the barrel HCAL and ECAL combination can be given as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \quad (4.2)$$

where a is the stochastic term and b is the constant term, measured in the test beam as $0.847 \pm 0.016 \sqrt{\text{GeV}}$ and 0.074 ± 0.008 , respectively.

4.1.4 Muon System

As the name ‘Compact Muon Solenoid’ suggests, detecting muons is one of the most important tasks for CMS. Muons are minimum interacting particles that can pass through several metres of iron without much energy loss and are not stopped by two calorimeters of CMS. Therefore, gas-based chambers to detect muons are placed at the very edge of

4 The Compact Muon Solenoid

the experiment where they are the only particles likely to register a signal.

To identify muons and measure their momenta, CMS uses three types of detector: drift tubes (DTs), cathode strip chambers (CSCs) and resistive plate chambers (RPCs). The DTs are used for precise trajectory measurements in the barrel region, while the CSCs are employed in the endcaps. The RPCs provide a fast (trigger) signal when a muon passes through the muon detector, and are installed in both the barrel and the endcaps.

Drift Tubes

The muon barrel (MB) consists of four layers of DT chambers. Each chamber is 4 cm wide tube that contains a stretched wire within a gas volume. When a muon or any other charged particle passes through the volume, it knocks out electrons off the gas atoms. These electrons follow the electric field ending up at the positively-charged wire. By keeping track of the wire electrons hit as well as by calculating the muons' original distance away from the wire, the DTs give two coordinates for the muons' position.

Cathode Strip Chambers

The muon endcap consists of 468 CSCs in two endcaps. The CSCs are used in the endcap disks where the magnetic field is irregular and the particle rates are very high. They consist of arrays of positively-charged anode wires crossed with negatively-charged copper cathode strips within a gas volume. When muons pass through, they knock electrons off the gas atoms, which flock to the anode wires creating an avalanche of electrons. Positive ions move away from the anode wire and towards the copper cathode, inducing a charge pulse in the strips, at right angles to the wire direction. Because the strips and the wires are perpendicular, we get two position coordinates for each passing particle.

In addition to providing precise space and time information, the closely spaced wires

4 The Compact Muon Solenoid

make the CSCs fast detectors suitable for triggering. Each CSC module contains six layers enabling it to accurately identify muons and match their tracks to those in the tracker.

Resistive Plate Chambers

RPCs are fast gaseous detectors that provide a muon trigger system parallel with those of the DTs and CSCs. They consist of two parallel plates, a positively-charged anode and a negatively-charged cathode, both made of a very high resistivity plastic material and separated by a gas volume.

When a muon passes through the RPC, electrons are knocked out of gas atoms. These electrons in turn hit other atoms causing an avalanche of electrons. The electrodes are transparent to the signal (due to electrons), which are instead picked up by external metallic strips after a small but precise time delay. The pattern of hit strips gives a quick measure of the muon momentum, which is then used by the trigger to make immediate decisions about whether the data are worth retaining. RPCs combine a good spatial resolution with a time resolution of just one nanosecond. Fig. 4.7 shows the schematic diagram of CMS muon system with position of CSCs, DTs and RPCs.

4 The Compact Muon Solenoid

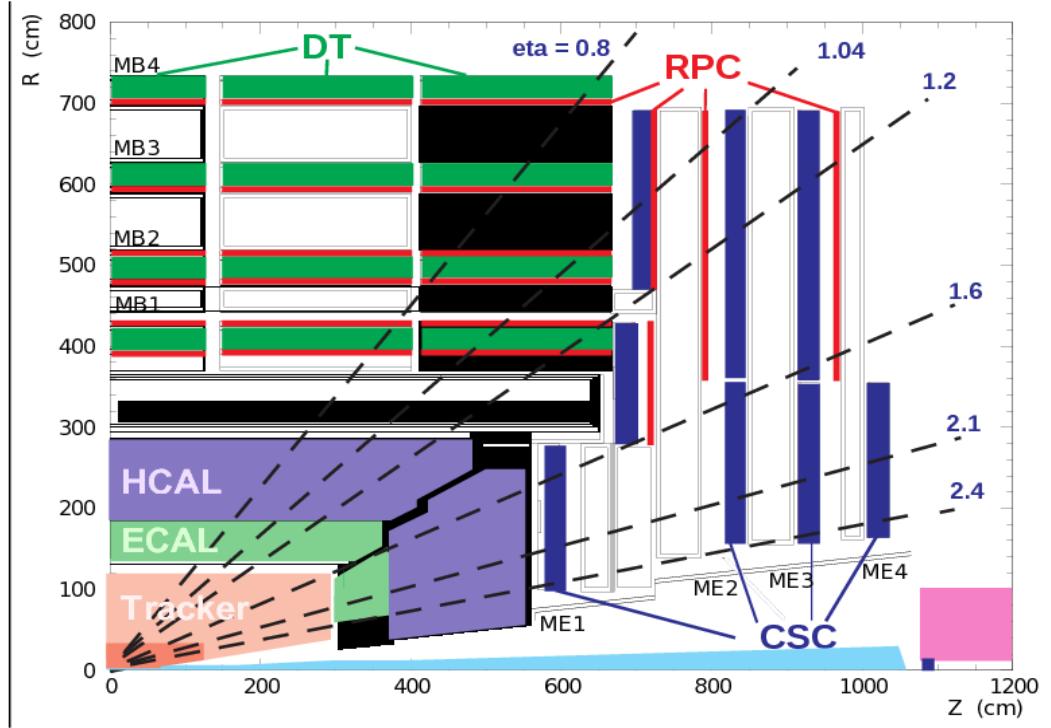


Figure 4.7: CMS muon system and position of muon subdetectors. The image is taken from Ref. [45].

Muon System Performance

The muon momentum is measured in the inner silicon tracker as well as in the muon system. When only used the muon system, a momentum resolution of 10% (20%) is achieved for 40 GeV muons in the barrel (endcaps). However, the resolution of a global muon that combines information from the tracker with the muon system, is about 1% (2%) for muons with $p_T < 100$ GeV in the barrel (endcaps). For high p_T muons (above 100 GeV), the resolution degrades to about 8% for the barrel and 10 % for the endcaps.

4.2 Object Reconstruction

In this section, we discuss generic methods that are used to reconstruct electron, muon, jets and MET in CMS. These methods remain the same for all objects used throughout the analysis works presented here. Basically these objects are reconstructed by studying the distributions of several variables that provide some information about the former. The exact criteria might change from analysis to analysis but the reconstruction methods remain unchanged.

4.2.1 Primary Vertex Reconstruction

The goal of primary vertex reconstruction is to determine the hardest vertex i.e. the precise position of pp interaction point. This task is accomplished with the so called Deterministic Annealing [46] clustering of tracks. From a set of reconstructed vertices, the one with maximal $\sum p_T^2$ is chosen as the primary vertex or the collision point.

4.2.2 Electron Reconstruction

Electrons are identified by combining tracks in the silicon tracker with energy deposits in the ECAL. Track trajectories are reconstructed using a dedicated modeling of the energy-loss due to bremsstrahlung radiation within the tracker material, and are fitted with a gaussian sum filter (GSF) [47]. The electron track reconstruction is performed by collecting the hits and accessing all the track parameters of the tracks with a large spectrum of emitted bremsstrahlung photons.

Electrons are identified using the following track and cluster shape variables.

- The difference between the cluster position and the track extrapolation from the innermost measurement in the η direction ($\Delta\eta_{in}$).

4 The Compact Muon Solenoid

- The difference between the cluster position and the track extrapolation from the innermost measurement in the ϕ direction ($\Delta\phi_{in}$).
- The cluster η width ($\sigma_{i\eta i\eta}$), which is defined from the covariance matrix with logarithmic weights
- The ratio of energy leaked into HCAL over the energy in the ECAL (H/E).
- The impact parameter in the transverse ($r\phi$) plane with respect to the primary vertex
- The impact parameter along the z -axis with respect to the primary vertex.

4.2.3 Muon Reconstruction

Muon reconstruction is performed using the all-silicon inner tracker at the centre of the detector, and with up to four stations of gas-ionization muon detectors installed outside the solenoid and sandwiched between the layers of the steel return yoke [48]. The muon system covers the pseudorapidity region $\eta < 2.4$ and performs three main tasks: triggering on muons, identifying them, and improving their momentum measurement and charge determination.

4.2.4 Jet Reconstruction

All three analyses presented in this thesis use jets. In fact, jets constitute an important part of object reconstruction in CMS. First particles (electrons, muons, photons, charged and neutral hadrons) are reconstructed using the Particle Flow (PF) method. [49]. These PF candidates are then fed to the anti- k_t [50] jet reclustering algorithm to form jets. The cone radius for jet reconstruction in the SUSY and leptoquark analysis is taken to be 0.4. In the pileup mitigation study, however, the same is changed to 0.8. To further improve their quality, a following set of criteria are imposed on our studies.

- Number of constituent particles in the jet should be greater than 1,

4 The Compact Muon Solenoid

- Charged electromagnetic energy fraction in the jet should be less than 0.99,
- Neutral hadronic energy fraction in the jet should be less than 0.99,
- Neutral electromagnetic energy fraction in the jet should be less than 0.99,
- Charged energy fraction in the jet should be greater than 0.0 within the $|\eta| < 2.4$,
and
- Charged multiplicity in the jet should be non-zero within $|\eta| < 2.4$.

4.2.5 Missing Transverse Energy

If a neutrino is present in the final state, it goes undetected in the detector. As a result, when we add up the momentum of all visible particles in the transverse plane, there is an overall imbalance. The amount needed to balance the transverse momentum to zero decides the MET (E_T) of the event. So E_T is computed as the negative vector sum of the transverse momenta of all PF candidates. The raw E_T is corrected for to remove the bias due to non-linearity in the calorimetric response for neutral and charged hadrons, caused by pileup, large bending of low p_T tracks due to the strong magnetic fields in the CMS.

5 Statistical Procedures for Search

In this chapter we discuss the general statistical procedures for interpreting data toward calculations of upper limits and significances. The content of this chapter is a brief collection from Ref. [51]. We will also describe how nuisance parameters are treated while extracting the upper limits. The concepts we will be discussing here will be used both in the SUSY and LQ analysis.

Both Bayesian and frequentist methods are used to calculate upper limits in CMS. We shall focus on the frequentist approach as we have used it in our analyses. This method allow one to quantify the level of incompatibility of the data with the signal hypothesis, which is expressed in terms of confidence level (CL). It is common to require 95% CL for excluding a signal, however this is a convention.

We denote s as the signal yield for a given analysis and b is the background yield. We also name μ to be the signal strength, defined as $\mu = s_0/s$ where s_0 is the signal yield after scaled by μ . Both predicted signal and background yields, prior to the scrutiny of the observed data entering the statistical analysis, are subject to multiple uncertainties that are handled by introducing nuisance parameters θ . As a result, the signal and background expectations become functions of these parameters: $s(\theta)$ and $b(\theta)$.

All sources of uncertainties are taken to be either 100%-correlated (positively or

negatively) or uncorrelated (independent). Partially correlated errors are either broken down to sub-components that can be said to be either 100% correlated or uncorrelated, or declared to be 100% or 0% correlated, whichever is believed to be appropriate or more conservative. This allows us to include all constraints in the likelihood functions in a clean factorized form.

5.1 Observed Limits

To calculate observed limits we use a test statistic that is widely known as the LHC test statistic. In the following we will construct the test statistic by defining few things . Let's first construct a likelihood function

$$L(\text{data}|\mu, \theta) = \text{Poisson}[\text{data}|\mu \cdot s(\theta) + b(\theta)].p(\tilde{\theta}|\theta) \quad (5.1)$$

where data represents either the actual experimental observation or pseudodata used to construct sampling distributions. $\text{Poisson}[\text{data}|\mu \cdot s + b]$ stands for a product of poisson probabilities to observe n_i events in bin i :

$$\prod_i \frac{(\mu \cdot s_i + b_i)^{n_i} \cdot e^{-\mu \cdot s_i - b_i}}{n_i!} \quad (5.2)$$

To compare the compatibility of the data with the background-only and signal-plus-background hypothesis, where signal is allowed to be scaled by some factor μ , we construct the test statistic:

5 Statistical Procedures for Search

$$\tilde{q}_\mu = -2 \ln \frac{L(\text{data}|\mu, \hat{\theta}_\mu)}{L(\text{data}|\hat{\mu}, \hat{\theta})} \quad (5.3)$$

with a constraint $0 \leq \hat{\mu} \leq \mu$. Here, $\hat{\theta}_\mu$ refers to the conditional maximum likelihood estimators of θ , given the signal strength parameter μ and data that, as before, may refer to the actual experimental observation or pseudodata. The pair of parameter estimators $\hat{\mu}$ and $\hat{\theta}$ corresponds to the global maximum of the likelihood. The lower constraint is guided by physics (signal rate should be positive), while the upper constraint is imposed by hand to guarantee a one-sided confidence interval. Physics wise, this means upward fluctuations of the data such that $\hat{\mu} > \mu$ are not considered as evidence against the signal hypothesis.

- Once we define the test statistic , then the observed value of the test statistic $\tilde{q}_\mu^{\text{obs}}$ is found.
- Next step is to find the values of the nuisance parameters best describing the observed data (i.e. maximizing the likelihood as given before), for the background-only and signal-plus-background hypothesis.
- Then the probability distributions of test statistic are generated using toy data for both signal-plus-background and background hypothesis. Lets denote those as $f(\tilde{q}_\mu|\mu, \theta_\mu^{\text{obs}})$ and $f(\tilde{q}_\mu|0, \theta_\mu^{\text{obs}})$ respectively.
- After getting the test statistic distributions we obtain two p-values associated with the two hypotheses as

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \text{signal + background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu^{\text{obs}}) d\tilde{q}_\mu \quad (5.4)$$

and

$$1 - p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \text{background}) = \int_{\tilde{q}_0^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | 0, \hat{\theta}_0^{\text{obs}}) d\tilde{q}_\mu \quad (5.5)$$

- Then we calculate CL_s as the ratio of these two probabilities:

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b} \quad (5.6)$$

- To quote the 95% CL upper limit on μ , denoted as $\mu^{95\%CL}$, we adjust μ until we reach $CL_s = 0.05$.

5.2 Expected Limits

A most straightforward way for defining the expected median upper-limit and $\pm 1\sigma$ as well as $\pm 2\sigma$ bands for the background-only hypothesis is to generate a large set of background only pseudodata and calculate CL_s and $\mu^{95\%CL}$ for all of them as if they were real data. One can then build a cumulative probability distribution of results by starting integration from the side that corresponds to low event yield. The point at which the cumulative probability distribution crosses the quantile of 50% is the median expected value. The $\pm 1\sigma$ (68%) band is defined by the crossings of the 16% and the

84% quantiles. Crossings at 2.5% and 97.5% define the $\pm 2\sigma$ (95%) band.

5.3 Significance

The presence of a signal is quantified by the background-only p-value, i.e. the probability for background to fluctuate and give an excess of events as large or larger than the observed one. The test statistic used to get the p-value for this purpose, is slightly different one and is defined as

$$\tilde{q}_0 = -2 \ln \frac{L(\text{data}|0, \hat{\theta}_0)}{L(\text{data}|\hat{\mu}, \hat{\theta})} \quad (5.7)$$

with a constraint $\hat{\mu} \geq 0$. The p-value is obtained by

$$p_0 = P(q_0 \geq q_0^{\text{obs}} | \text{background}) = \int_{q_0^{\text{obs}}}^{\infty} f(q_0|0, \hat{\theta}_0^{\text{obs}}) d\tilde{q}_0 \quad (5.8)$$

To convert the p-value to significance Z, we use a normal Gaussian function as:

$$p = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (5.9)$$

We have here discussed the basic principles. Details of how systematics are treated will be discussed in the respective analysis section.

6 Search for SUSY with 13 TeV pp

Collision Data

As discussed earlier, SUSY is one BSM theory that provides solutions to many problems of the SM. It provides a way to take care the Higgs mass divergence problem as well as to unify strong, weak and electromagnetic interactions, and has a candidate for dark matter. Natural SUSY models [69] minimize the fine tuning associated with the value of the Higgs boson mass and its radiative corrections. In these models, the top squark, left-handed bottom squark, gluino, and higgsino are required to be light, i.e., to have masses near the electroweak energy scale, making them potentially accessible at the LHC. Among many SUSY processes, the gluino pair production has largest cross-section [36] which increases by a factor of 30 going from 8 to 13 TeV. This fact makes it pertinent to search for SUSY even with a couple of fb^{-1} of data. Since baryon number (B) and lepton number (L) conservation have been tested very precisely, the B-L violating couplings need to be very small in order not to be in conflict with experimental data. R-parity is a symmetry acting on the MSSM fields that forbids these couplings. In R-parity conserving SUSY models, the lightest supersymmetric particle (LSP) is stable and a possible dark-matter candidate. The results of our search are interpreted with

6 Search for SUSY with 13 TeV pp Collision Data

R-parity conserving simplified SUSY scenarios [37].

In this chapter, we first describe the study with 2.3 fb^{-1} of the 2015 data and later with 12.9 fb^{-1} of the 2016 data. We search in all hadronic final states i.e., in the final states of jets and missing transverse momentum. All hadronic SUSY searches are especially interesting, because they often make up a large portion of branching fractions for typical signals and tend to be easily hidden in extensions to SUSY models.

We present a general search for gluino pair production leading to final states with large missing transverse energy (p_T^{miss}), large scalar sum of transverse energy (H_T), large number of jets (N_{jet}) and large number of b-tagged jets ($N_{\text{b-jet}}$). The event diagrams of gluino pair production that we considered in our analysis are shown below. We shall focus more on the $Z \rightarrow \nu\bar{\nu}$ invisible background estimation as this is the part of the analysis where we have contributed to the most.

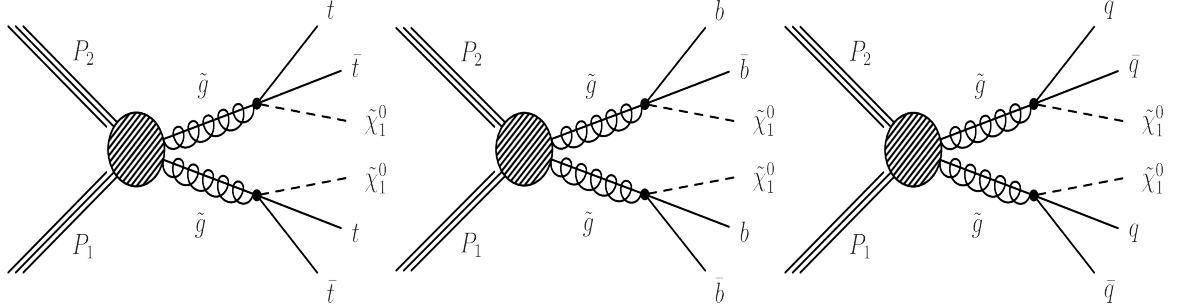


Figure 6.1: Event diagrams for gluino pair production with four top final state (left), four b-quark final state (middle) and four light quark final state (right)

The major backgrounds for this final state come from $t\bar{t}$, $W+\text{jets}$, $Z(\rightarrow \nu\bar{\nu})+\text{jets}$, QCD processes. All these backgrounds are estimated using data driven techniques, keeping usage of MC samples to a minimum.

6.1 Study with 2.3 fb^{-1} of the 2015 Data

In this section, we describe the details of the SUSY search done with 2.3 fb^{-1} of the 2015 data.

6.1.1 Data-MC Samples and Trigger

First we give various MC samples used in the analysis along with their cross sections. The SM samples are listed in Tables 6.1-6.8. The cross sections listed correspond to next-to-leading-order (NLO) calculations unless otherwise noted. These samples correspond to a pileup distribution with an average of 20 interactions per bunch crossing and a 25 ns interval between bunches.

Table 6.1: MC samples for signal SMS (simplified SUSY) model points.

Dataset (masses in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
SMS-T1tttt ($m_{\text{gluino}} = 1500, m_{\text{LSP}} = 100$)	madgraph-pythia8	0.014	7268
SMS-T1tttt ($m_{\text{gluino}} = 1200, m_{\text{LSP}} = 800$)	madgraph-pythia8	0.086	1719
SMS-T1bbbb ($m_{\text{gluino}} = 1500, m_{\text{LSP}} = 100$)	madgraph-pythia8	0.014	3708
SMS-T1bbbb ($m_{\text{gluino}} = 1000, m_{\text{LSP}} = 900$)	madgraph-pythia8	0.325	438.5
SMS-T1qqqq ($m_{\text{gluino}} = 1400, m_{\text{LSP}} = 100$)	madgraph-pythia8	0.025	1958
SMS-T1qqqq ($m_{\text{gluino}} = 1400, m_{\text{LSP}} = 800$)	madgraph-pythia8	0.325	293.0

Table 6.2: SM $t\bar{t}$ MC samples used in the analysis. The cross sections are calculated to NNLO (next-to-next-to leading order).

Dataset(decay mode/ H_T range in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
$t\bar{t} + \text{jets}(\text{inclusive})$	madgraph,pythia8	816.0	13.90
$t\bar{t} + \text{jets} (\text{with leptonic } t \text{ decay})$	madgraph-pythia8	179.3	324.6
$t\bar{t} + \text{jets} (\text{with leptonic } \bar{t} \text{ decay})$	madgraph-pythia8	179.3	335.7
$t\bar{t} + \text{jets}(\text{both } ts \text{ decays leptonically})$	madgraph-pythia8	86.66	351.2
$t\bar{t} + \text{jets}(H_T\text{-}600\text{to}800)$	madgraph-pythia8	2.615	1898
$t\bar{t} + \text{jets}(H_T\text{-}800\text{to}1200)$	madgraph-pythia8	1.077	3198
$t\bar{t} + \text{jets}(H_T\text{-}1200\text{to}2500)$	madgraph-pythia8	0.195	5063
$t\bar{t} + \text{jets}(H_T > 2500)$	madgraph-pythia8	0.002	218575

6 Search for SUSY with 13 TeV pp Collision Data

Table 6.3: SM QCD MC samples used in the analysis. All cross sections are calculated to LO.

Dataset (H_T range in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
QCD (200-300)	madgraph-pythia8	1735000	0.01
QCD (300-500)	madgraph-pythia8	366800	0.05
QCD (500-700)	madgraph-pythia8	29370	0.67
QCD (700-1000)	madgraph-pythia8	6524	2.30
QCD (1000-1500)	madgraph-pythia8	1064	4.67
QCD (1500-2000)	madgraph-pythia8	121.5	31.67
QCD (> 2500)	madgraph-pythia8	25.42	77.17

Table 6.4: SM $Z \rightarrow \nu\nu + \text{jets}$ MC samples used in the analysis. The cross sections are calculated to NNLO.

Dataset(H_T range in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
$Z(\nu\bar{\nu}) + \text{jets}$ (100-200)	madgraph	345.0	14.92
$Z(\nu\bar{\nu}) + \text{jets}$ (200-400)	madgraph	96.38	52.22
$Z(\nu\bar{\nu}) + \text{jets}$ (400-600)	madgraph	13.46	75.34
$Z(\nu\bar{\nu}) + \text{jets}$ (>600)	madgraph	5.170	196.5

Table 6.5: SM $W \rightarrow l\nu + \text{jets}$ MC samples used in the analysis. The cross sections are calculated to NNLO.

Dataset(H_T range in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
$W(\ell\nu) + \text{jets}$ (100-200)	madgraph-pythia8	1635	6.20
$W(\ell\nu) + \text{jets}$ (200-400)	madgraph-pythia8	437.0	11.97
$W(\ell\nu) + \text{jets}$ (400-600)	madgraph-pythia8	59.50	31.96
$W(\ell\nu) + \text{jets}$ (> 600)	madgraph-pythia8	22.80	45.44
$W(\ell\nu) + \text{jets}$ (600-800)	madgraph-pythia8	15.50	257.1
$W(\ell\nu) + \text{jets}$ (800-1200)	madgraph-pythia8	6.366	247.4
$W(\ell\nu) + \text{jets}$ (1200-2500)	madgraph-pythia8	1.614	158.4
$W(\ell\nu) + \text{jets}$ (>-2500)	madgraph-pythia8	0.037	6770

6 Search for SUSY with 13 TeV pp Collision Data

Table 6.6: SM diboson and other rare process MC samples used in the analysis. The cross sections are calculated to NNLO.

Dataset	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
t̄H + jets(H → bb)	amcatnloFXFX-madspin-pythia8	0.293	18269
t̄Z (with 2 ℓ and 2 ν in the final state)	amcatnlo-pythia8	0.228	811.4
t̄Z (hadronic top decays)	amcatnlo-pythia8	0.530	663.4
t̄W + jets(W → $\ell\nu$)	amcatnloFXFX-madspin-pythia8	0.204	635.6
t̄W + jets(W → q̄q̄)	amcatnloFXFX-madspin-pythia8	0.423	1018
ZH(H → bb, Z → $\nu\bar{\nu}$)	amcatnloFXFX madspin pythia8	0.100	12116
WH(H → b̄b, W → $\ell\nu$)	amcatnloFXFX madspin pythia8	0.260	4782
WW(W → qq̄, W → $\ell\nu$)	amcatnloFXFX madspin pythia8	50.00	64.26
WW(W → $\ell\nu$, W → $\ell\nu$)	powheg	12.18	158.5
WZ(W → $\ell\nu$, Z → q̄q̄)	amcatnloFXFX madspin pythia8	10.71	1339
WW(W → $\ell\nu$, Z → $\nu\bar{\nu}$)	amcatnloFXFX madspin pythia8	3.058	305.7
ZZ(Z → q̄q̄, Z → $\nu\bar{\nu}$)	amcatnloFXFX madspin pythia8	4.040	5556
ZZ(Z → $\ell^+\ell^-$, Z → q̄q̄)	amcatnloFXFX madspin pythia8	3.220	3706
WWZ	amcatnlo-pythia8	0.165	1341
WZZ	amcatnlo-pythia8	0.056	3938
ZZZ	amcatnlo-pythia8	0.014	15297

Table 6.7: SM DY+jets MC samples used in the analysis. The cross sections are calculated to NNLO.

Dataset (H_T range in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
$Z(\ell^+\ell^-) + \text{jets}(\text{inclusive})$	madgraph-pythia8	6025	1.50
$Z(\ell^+\ell^-) + \text{jets}, (100-200)$	madgraph-pythia8	171.5	15.31
$Z(\ell^+\ell^-) + \text{jets}, (200-400)$	madgraph-pythia8	52.58	18.18
$Z(\ell^+\ell^-) + \text{jets}, (400-600)$	madgraph-pythia8	6.761	155.0
$Z(\ell^+\ell^-) + \text{jets}, (>600)$	madgraph-pythia8	2.718	363.5

Table 6.8: SM γ +jets MC samples used in the analysis. The cross sections are calculated to LO.

Dataset(H_T range in GeV)	Generator	σ (pb)	$\mathcal{L}_{\text{int}}(\text{fb}^{-1})$
$\gamma + \text{jets} (100-200)$	madgraph-pythia8	22010	0.23
$\gamma + \text{jets} (200-400)$	madgraph-pythia8	9110	1.13
$\gamma + \text{jets} (400-600)$	madgraph-pythia8	273	9.07
$\gamma + \text{jets} (>600)$	madgraph-pythia8	94.5	26.99

6.1.1.1 Trigger

The trigger information about the signal region and Z to invisible background control region are given here. To know about the other background region triggers refer to Ref. [52] .

Signal Region Trigger: The data collected with the CMS experiment at $\sqrt{s}=13$ TeV, corresponding to an integrated luminosity of 2.3 fb^{-1} , are used in the analysis. For the signal region the events are selected using a trigger with online $H_T > 250 \text{ GeV}$ and online MET $> 100 \text{ GeV}$. The technical name of the path is HLT-PFHT350-PFMET100-NoiseCleaned, where, HLT stands for High Level Trigger. This trigger is found to be almost fully efficient for offline $H_T > 500 \text{ GeV}$ and $H_T^{\text{miss}} > 200 \text{ GeV}$.

Control Region Trigger: We estimate the $Z(\rightarrow \nu\bar{\nu})+\text{jets}$ background using $\gamma+\text{jets}$ events in data. The single photon events are selected using a trigger with online $\gamma p_T > 90 \text{ GeV}$ and $H_T > 500 \text{ GeV}$. The trigger is found to be almost 99% efficient in regions of search.

6.1.2 Event Selection Criteria and Exclusive Search Intervals

The following criteria define our baseline selection.

- $N_{\text{jet}} > 4$;

Since each pair of gluinos decays to four quarks, all events are required to contain at least four good jets, satisfying

- $p_T > 30 \text{ GeV}$,
- $|\eta| < 2.4$,

6 Search for SUSY with 13 TeV pp Collision Data

- loose jet ID criteria for PF jets defined by:

- * neutral hadron fraction < 0.99 ,
- * neutral EM fraction < 0.99 ,
- * number of constituents > 1 ,
- * charged hadron fraction > 0 ,
- * charged multiplicity > 0 ,
- * charged EM fraction < 0.99

- $H_T > 500$ GeV, where $H_T = \sum_{\text{jets}} p_T$. The jets must meet the criteria listed above.
- $H_T^{\text{miss}} > 200$ GeV, where $H_T^{\text{miss}} = \left| \sum_{\text{jets}} \vec{p}_T \right|$. All jets included in this sum must satisfy $p_T > 30$ GeV, $|\eta| < 5$. The jets within tracker acceptance ($|\eta| < 2.4$) must also satisfy the “loose” jet ID listed above. The jets outside of tracker acceptance need only to satisfy
 - For $3 < |\eta| < 5$:
 - * neutral EM fraction < 0.90 ,
 - * number of neutral particles > 10 .
 - For $2.4 < |\eta| < 3$:
 - * neutral EM fraction < 0.99 ,
 - * neutral hadron fraction < 0.99 ,
 - * number of constituents > 1 .
- Muon veto:

Muon candidates are selected using the muon POG (Physics Object Group) [53] recommended muon working point for Run2 data.

6 Search for SUSY with 13 TeV pp Collision Data

- Electron veto:

Electron candidates are selected using the Egamma POG recommended Cut Based VETO selection [54].

- Angular cut:

The majority of QCD multijet events in our high- H_T^{miss} search region have jets with under measured momenta leading to a spurious momentum imbalance. A signature of such an event is a jet closely aligned in direction with the H_T^{miss} vector. To suppress this background, we reject all events in which the two highest- p_T jets lie within 0.5 rad of the H_T^{miss} vector in the azimuthal coordinate:

$$\Delta\phi(j_1, H_T^{\text{miss}}) > 0.5 \quad (6.1)$$

$$\Delta\phi(j_2, H_T^{\text{miss}}) > 0.5 \quad (6.2)$$

This requirement is relaxed for the third- and fourth-highest- p_T jets:

$$\Delta\phi(j_3, H_T^{\text{miss}}) > 0.3 \quad (6.3)$$

$$\Delta\phi(j_4, H_T^{\text{miss}}) > 0.3 \quad (6.4)$$

No such requirement is placed on other jets.

Figures . 6.2 and 6.3 show the distributions of search variables for various MC simulated events after the baseline selections are applied.

Search Binning

On top of the baseline we split the available phase-space into 72 exclusive intervals to enhance the sensitivity. The search variables and corresponding binning are as follows:

6 Search for SUSY with 13 TeV pp Collision Data

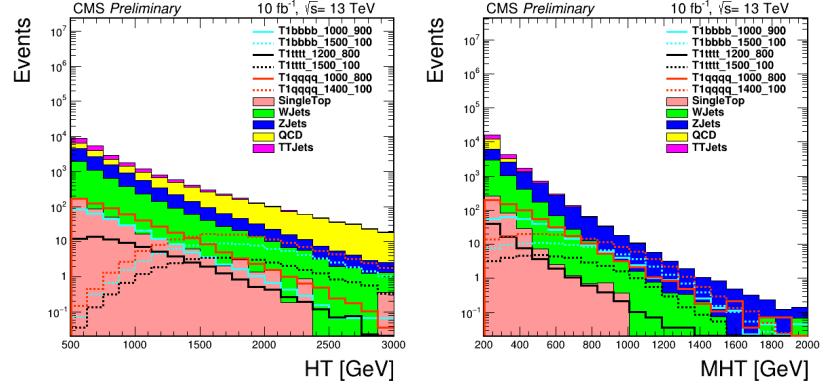


Figure 6.2: Signal vs. stacked backgrounds in H_T (HT, left) and H_T^{miss} (MHT, right).

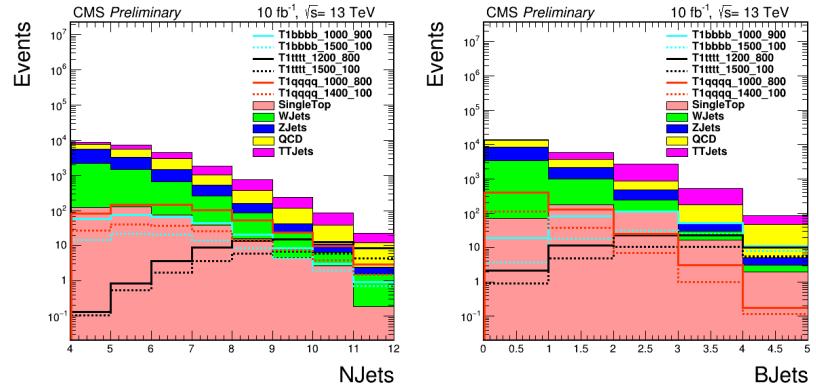


Figure 6.3: Signal vs. stacked backgrounds in N_{jet} (NJets, left) and $N_{b\text{-jet}}$ (BJet, right).

- N_{jet} : 4–6, 7–8, ≥ 9 ;
- $N_{b\text{-jet}}$: 0, 1, 2, ≥ 3 ;
- H_T : 500–800, 800–1200, ≥ 1200 GeV;
- H_T^{miss} : 200–500, 500–750, ≥ 750 GeV.

The analysis is restricted to $N_{\text{jet}} \geq 4$ because of our focus on gluino pair production. The $H_T > 500$ GeV and $H_T^{\text{miss}} > 200$ GeV requirements are dictated by the trigger conditions. The binning is rectangular as above, except for the 12 bins with $H_T < 800$ GeV and $H_T^{\text{miss}} > 750$ GeV are removed since H_T^{miss} cannot exceed (or be on the order of) H_T .

6 Search for SUSY with 13 TeV pp Collision Data

in a physical event. Additionally, for $500 < H_T^{\text{miss}} < 750$ GeV, a single range with $500 < H_T < 1200$ GeV is used, and for $H_T^{\text{miss}} > 750$ GeV, a single range with $H_T > 800$ GeV is used, because of the lower number of events expected at large H_T^{miss} . The merging of these bins further reduces the bin count by 24, to a total of 72 bins. The six bins in the H_T and H_T^{miss} plane are shown visually in Fig. 6.4.

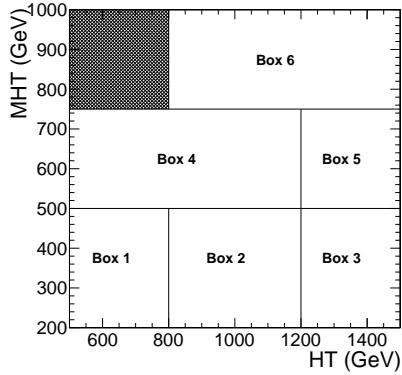


Figure 6.4: Search binning in the 2D phase-space of H_T and H_T^{miss} .

6.1.3 Backgrounds

All the SM backgrounds are categorized into four different groups. These are a)Z to invisible, b)Lostlepton, c)Hadronic Tau, d)QCD. Z($\rightarrow \nu\bar{\nu}$)+jets is an irreducible background having exactly the same final state as the signal. The lost lepton background arises when a top or W decays leptonically and the lepton goes unidentified, or non-isolated or out of acceptance. In case the W decays to taus with taus subsequently decaying to hadrons, it becomes hadronic tau background. QCD processes have fake MET due to jet energy mis measurement, so they enter the signal region. All the backgrounds mentioned here are determined using data driven techniques because the MC samples usually do not have accurate simulations for the higher order processes. Below we will discuss extensively the estimation of Z to invisible background in which we have

6 Search for SUSY with 13 TeV pp Collision Data

taken a key role.

6.1.3.1 $Z(\rightarrow \nu\bar{\nu})$ +jets background estimation

This is the most important backgrounds being an irreducible one. We are talking about events with a Z boson produced in association with jets when the former decays to two neutrinos. The most straightforward way to measure this background is to exploit the decays $Z(\ell^+\ell^-)$ +jets in which the Z boson can be reconstructed from an observed pair of muons or electrons. The efficiency-corrected yields from these decays can be translated directly into the $Z(\nu\bar{\nu})$ +jets background yield by the known branching ratios. The limitation of this approach stems from the rather small branching ratio between the charged and neutral leptons, so that the transfer factor from the control sample measurement to the predicted background is larger than one ;ignoring efficiencies, the branching ratio itself is approximately 3 when both muon and electron pairs are used.

An alternative approach is to exploit the similarity to Z boson radiation of the more copious radiation of photons. Here the challenge is to obtain validation in data of the MC predictions connecting the two processes.

Our baseline strategy is to use the γ +jets sample to determine the yields in the 18 bins corresponding to $N_{b\text{-jet}}=0$. These are then compared with the $Z(\ell^+\ell^-)$ +jets yields in the low- N_{jet} bins to establish the systematic uncertainty of the physics modeling of γ +jets, and the normalization corrected if necessary. The extrapolation to bins with $N_{b\text{-jet}} > 0$ is performed to the extent possible with the $Z(\ell^+\ell^-)$ +jets data sample, supplemented with MC information where necessary. We use simulation to correct for the resulting distribution to the higher N_{jet} bins. The master formula that gives the prediction in the 0 b-tagged bins is

$$N_{Z(\nu\nu)+\text{jets}}^{\text{prediction}} = \frac{R_{Z \rightarrow \ell^+\ell^-/\gamma}^{\text{obs}}}{R_{Z \rightarrow \ell^+\ell^-/\gamma}^{\text{MC}}} \cdot \frac{N_{Z \rightarrow \nu\bar{\nu}}^{\text{MC}}}{N_{\gamma}^{\text{MC}}} \cdot (\beta_{\text{purity}}^{\text{EB}} \cdot N_{\gamma}^{\text{data,EB}} + \beta_{\text{purity}}^{\text{EE}} \cdot N_{\gamma}^{\text{data,EE}}) \cdot \frac{1}{C_{\text{data/MC}}}$$

6 Search for SUSY with 13 TeV pp Collision Data

where

- $N_{Z(\nu\nu)+\text{jets}}^{\text{prediction}}$ is the number of predicted $Z(\nu\nu)$ +jets events in the search bins.
- $N_\gamma^{\text{data,EB}}$ and $N_\gamma^{\text{data,EE}}$ are the number of single photon events observed in data in barrel and endcap, respectively.
- $\beta_{\text{purity}}^{\text{EB}}$ and $\beta_{\text{purity}}^{\text{EE}}$ are the photon purity in the barrel and endcap, respectively. Single photon events get contaminated from $\pi^0 \rightarrow \gamma\gamma$ processes. We measure the purity also using data driven techniques. This will be discussed in the following sections.
- $R_{Z \rightarrow \ell^+\ell^-/\gamma}^{\text{MC}}$ is the ratio of production rates at the reconstruction level. This is calculated using MC events. As we do not trust simulations fully, we derive a correction to this ratio by calculating a double ratio, which is defined as the ratio of $Z \rightarrow \ell^+\ell^-$ events to $\gamma + \text{jets}$ events calculated in data and MC respectively.

Photon Reconstruction

For the photon+jets control sample, we require at least one well identified and isolated photon candidate having a minimum p_T of 100 GeV and $|\eta| < 2.5$, excluding the barrel-endcap transition, $1.4442 < |\eta| < 1.566$. Identification and isolation requirements, intended to reject electrons and pions misreconstructed as photons, are adopted from the Egamma POG's recommendations for 8 TeV and 13 TeV analyses. The identification criteria, which include requirements for low hadronic activity (H/E), a shower shape ($\sigma_{inj\eta}$) consistent with a photon, and an associated pixel seed veto, are listed in Table 6.9 for events with a photon in the barrel or endcap.

The isolation requirements restrict the energy sum from PF candidates within a cone of $\Delta R < 0.3$ around the momentum vector of the photon candidate. Specifically, we require that the energy from charged hadrons ($\text{Iso}_{\text{pfCh.}}$), neutral hadrons ($\text{Iso}_{\text{pfNu.}}$), and

6 Search for SUSY with 13 TeV pp Collision Data

electromagnetic particles $\text{Iso}_{\text{pfGa.}}$ not to exceed the p_{T} -dependent thresholds listed in Table 6.9. The isolation energy from each of the three particle species is corrected for pileup with the per-event average pileup energy density (ρ) and per-photon η -dependent effective area given in Table 6.10.

Table 6.9: Identification and isolation requirements for photons in the barrel and endcap.
EA stands for effective area.

Variable	Barrel	Endcap
Pixel Seed Veto	Yes	Yes
H/E	<0.05	<0.05
$\sigma_{inj\eta}$	<0.011	<0.031
$\max(\text{Iso}_{\text{pfCh.}} - \text{EA} \cdot \rho, 0)$	<0.7	<0.5
$\max(\text{Iso}_{\text{pfNu.}} - \text{EA} \cdot \rho, 0)$	$<0.4 + 0.04 \times p_{\text{T}}$	$<1.5 + 0.04 \times p_{\text{T}}$
$\max(\text{Iso}_{\text{pfGa.}} - \text{EA} \cdot \rho, 0)$	$<0.5 + 0.005 \times p_{\text{T}}$	$<1.0 + 0.005 \times p_{\text{T}}$

Table 6.10: Effective areas used in pileup correction as a function of photon pseudorapidity. Three columns shows different values used for calculating $\text{Iso}_{\text{pfCh.}}$, $\text{Iso}_{\text{pfNu.}}$ and $\text{Iso}_{\text{pfGa.}}$, respectively, in various η bins.

$ \eta $	EA(pfCh.)	EA(pfNu.)	EA(pfGa.)
0.0 - 1.0	0.012	0.030	0.148
1.0 - 1.479	0.010	0.057	0.130
1.479 - 2.0	0.014	0.039	0.112
2.0 - 2.2	0.012	0.015	0.216
2.2 - 2.3	0.016	0.024	0.262
2.3 - 2.4	0.020	0.039	0.260
> 2.4	0.012	0.072	0.266

N_{γ}^{obs} and photon categories:

The photons in the control sample come from three sources: direct prompt photons, fragmentation prompt photons, and non-prompt photons. Prompt photons are radiated from a quark while non-prompt photons come from the decay of a meson. Direct (fragmentation) prompt photons are radiated with relatively (small) large transverse momentum relative to the quark;

6 Search for SUSY with 13 TeV pp Collision Data

Direct prompt photons, which make up about 85% of the control sample, are the most useful in predicting the $Z \rightarrow \nu\nu$ background because their production processes most neatly map onto the set of Z boson production processes. As processes leading to non-prompt photons have no corresponding process for those leading to Z bosons, non-prompt photons do not contribute to the $Z \rightarrow \nu\nu$ background estimation . At the most basic level, the correspondence (or lack thereof) between fragmentation photon and Z processes is less straightforward, and the special treatment of this photon category is instead motivated by practical considerations.

We categorize photons in the simulation as

- prompt photons: stable final state photons with its parent either a photon, quark or proton.
 - direct prompt photons: prompt photons with $\Delta R > 0.4$ with respect to any quarks or gluons,
 - fragmentation prompt photons: all prompt photons that are not direct, and
- non-prompt photons: all photons that are not prompt.

Any photon in the $\gamma+\text{jets}$ or QCD samples is either non-prompt, direct prompt, or fragmentation prompt.

The numbers of observed photons in ECAL barrel (EB) and endcap (EE) are shown in Fig. 6.5. There is no systematic uncertainty on N_γ^{obs} because it is a simple observation. The statistical uncertainty related to this observation is included in the analysis when a Poisson PDF is assigned to this variable. The plots in Figs. 6.6, 6.7 shows the data-MC agreement in terms of search variables for the $\gamma+\text{jets}$ control sample.

6 Search for SUSY with 13 TeV pp Collision Data

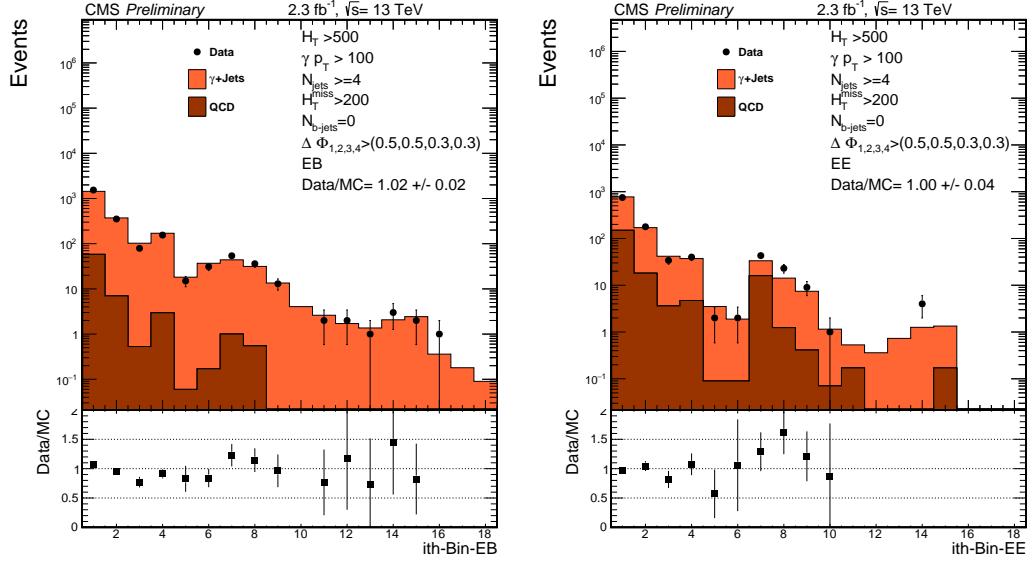


Figure 6.5: Numbers of observed events in the photon control samples in barrel (left) and endcap (right) compared to simulation.

Purity:

The purity β is defined as the fraction of all photons (prompt+non-prompt) that are prompt: $\beta = N_{\text{prompt}} / (N_{\text{prompt}} + N_{\text{non-prompt}})$. In practice, we allow the purity to depend on photon p_T , which is essentially equivalent to event H_T^{miss} . We also explore differences in purity for ECAL barrel and endcap separately.

Prompt photons can be distinguished from non-prompt ones by the differences in shape of their showers in the ECAL, as described by the well known quantity $\sigma_{i\eta i\eta}$. The purity is determined with a two-component fit to the $\sigma_{i\eta i\eta}$ distribution in the photon control sample. We have done an extensive study of $\sigma_{i\eta i\eta}$ distribution to validate the model to determine the purity of photons from data. The shape of prompt PDF is taken from $\gamma+$ jets MC sample. The shape of the prompt PDF changes in the bins of H_T^{miss} but not in H_T and N_{jet} . So we parameterize the purity as a function of H_T^{miss} inclusive of all H_T and N_{jet} bins. The shapes of the prompt PDF in H_T and N_{jet} bins for barrel is shown in Fig. 6.8

6 Search for SUSY with 13 TeV pp Collision Data

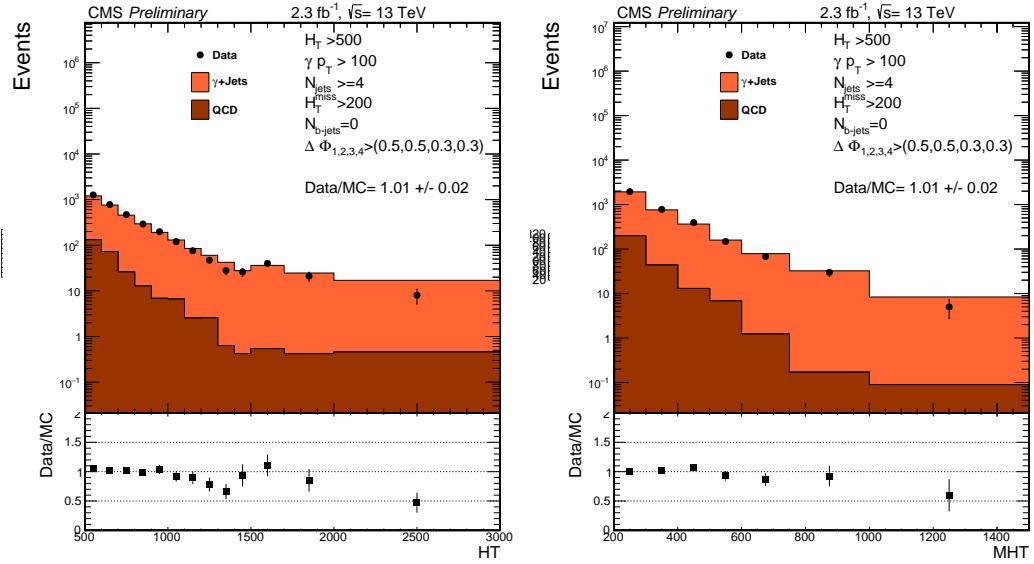


Figure 6.6: Data-MC agreement in terms search variables H_T (left) and H_T^{miss} (right) for the $\gamma + \text{jets}$ control sample.

The PDF for the prompt component is fit directly in data using a Gaussian function, which was motivated by the $\gamma + \text{jets}$ MADGRAPH simulation, and the PDF for the non-prompt component comes from the QCD MADGRAPH simulation. The fit result is $(96.3 \pm 0.6)\%$ for H_T^{miss} in the range 200-500 GeV and $(90 \pm 3)\%$ for $H_T^{\text{miss}} > 500$ GeV. Results of the EB and EE purity fits are shown in Fig. 6.9.

To obtain the systematic uncertainty on purity we perform the purity fits additional times with alternative prompt and non-prompt PDFs, and we take the difference in resulting purity as the systematic uncertainty. The photon control sample is defined with a requirement on the charge isolation of photon candidates. For the alternative non-prompt PDF, we obtain a sample of mostly non-prompt photons by inverting the charge isolation requirement, and we use the σ_{inj} shape from these photons as the PDF. We also considered using a template derived from $\gamma + \text{jets}$ simulation as an alternative prompt PDF. The variation in purity result for the alternative prompt PDF is negligible. The variation in purity result for the alternative non-prompt PDF is about 4%. We

6 Search for SUSY with 13 TeV pp Collision Data

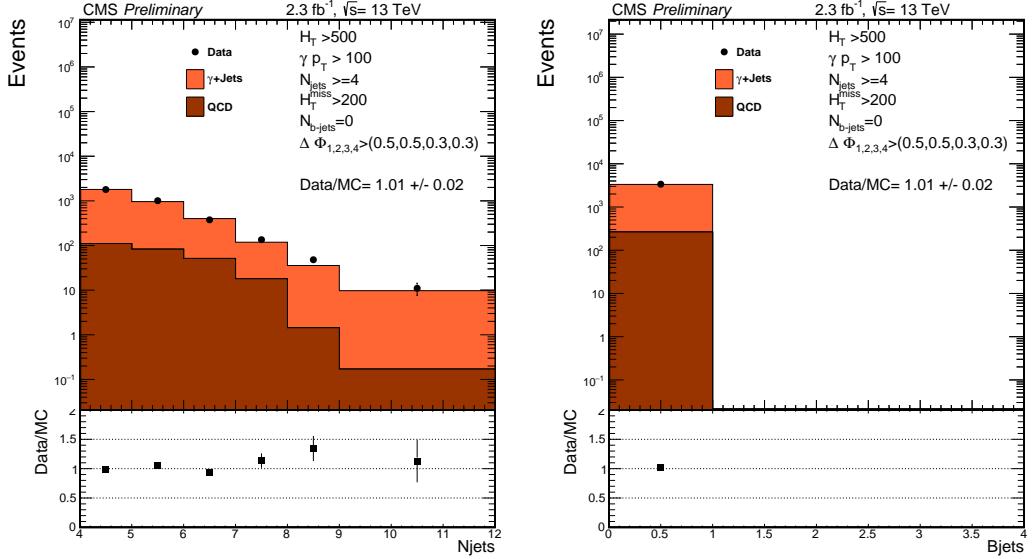


Figure 6.7: Data-MC agreement in terms search variables N_{jet} (left) and $N_{\text{b-jet}}$ (right) for $\gamma+\text{jets}$ control region.

take the quadrature sum of the difference in fit and the statistical uncertainty on the nominal fit as the systematic uncertainty on the purity; the systematic uncertainty is uncorrelated across different H_T^{miss} bins.

$$\mathcal{R}_{Z(\nu\bar{\nu})/\gamma} \text{ and } \mathcal{F}_{\text{dir}}$$

To account for the difference in cross sections and small differences in kinematic distributions of Z and γ events, we apply a scale factor $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ which is the ratio of the number of $Z \rightarrow \nu\nu + \text{jets}$ and $\gamma + \text{jets}$ events computed with LO generator-level simulation (MADGRAPH+PYTHIA) in each analysis bin. The numerator and denominator are both determined after full selection at the reconstruction, instead of generation, physics objects. Fig. 6.10 shows $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ in the 18 0-btagged bins.

As $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ is determined at reconstruction level, we correct $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ for potential data-MC scale factors (SFs) for efficiency related to photon reconstruction, identification, isolation, and trigger. These SFs have been computed for our selection as functions of photon E_T , photon η , N_{jet} , and $\Delta R(\gamma, \text{nearest jet})$. Results show that the SFs are

6 Search for SUSY with 13 TeV pp Collision Data

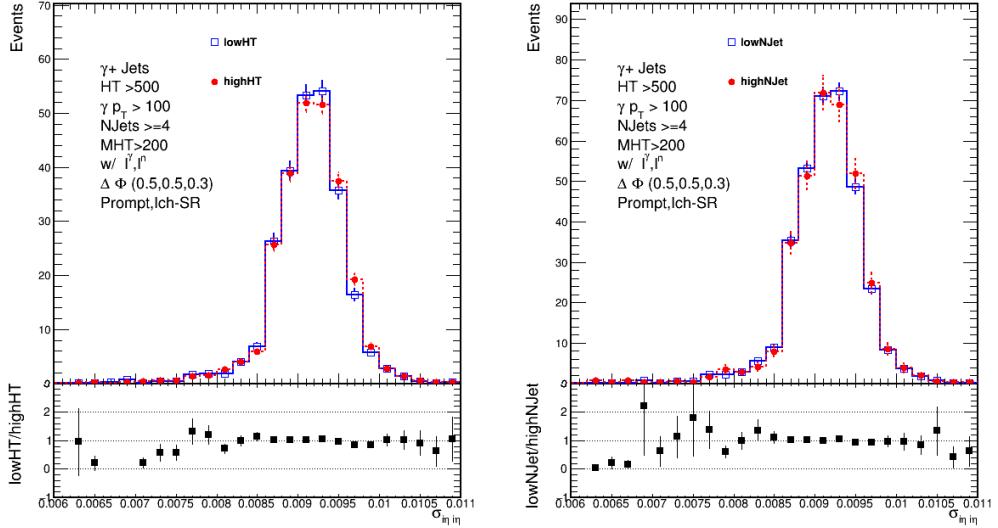


Figure 6.8: Distribution of $\sigma_{i\eta i\eta}$ in bins of H_T (left), N_{jet} and H_T^{miss} for barrel photons.

independent of the analysis variables, so we correct for the efficiency by the single value of 0.98 ± 0.05 when computing $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$, and vary the SF by a 5% to determine the systematic uncertainty.

Two different kinds of uncertainties are shown in Fig. 6.10. The statistical ones due to finite MC statistics and the systematic ones are obtained by propagating the ID/Iso SF to $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$.

As introduced above, fragmentation photons should be treated carefully when calculating $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ primarily because the MADGRAPH $\gamma + \text{jets}$ events that make up the $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ denominator are generated with the requirement that $\Delta R(\gamma, \text{quarks, gluons}) > 0.4$ and $\Delta R(\gamma, \text{gluons}) > 0.4$. As these fragmentation prompt photons are included in the photon control sample and are experimentally indistinguishable from direct prompt photons (unlike non-prompt ones), we must correct for this missing component of the simulation. We use prompt photons included in the QCD simulation sample to determine the fraction with $\Delta R > 0.4$. The value of fragmentation fraction is found to be

6 Search for SUSY with 13 TeV pp Collision Data

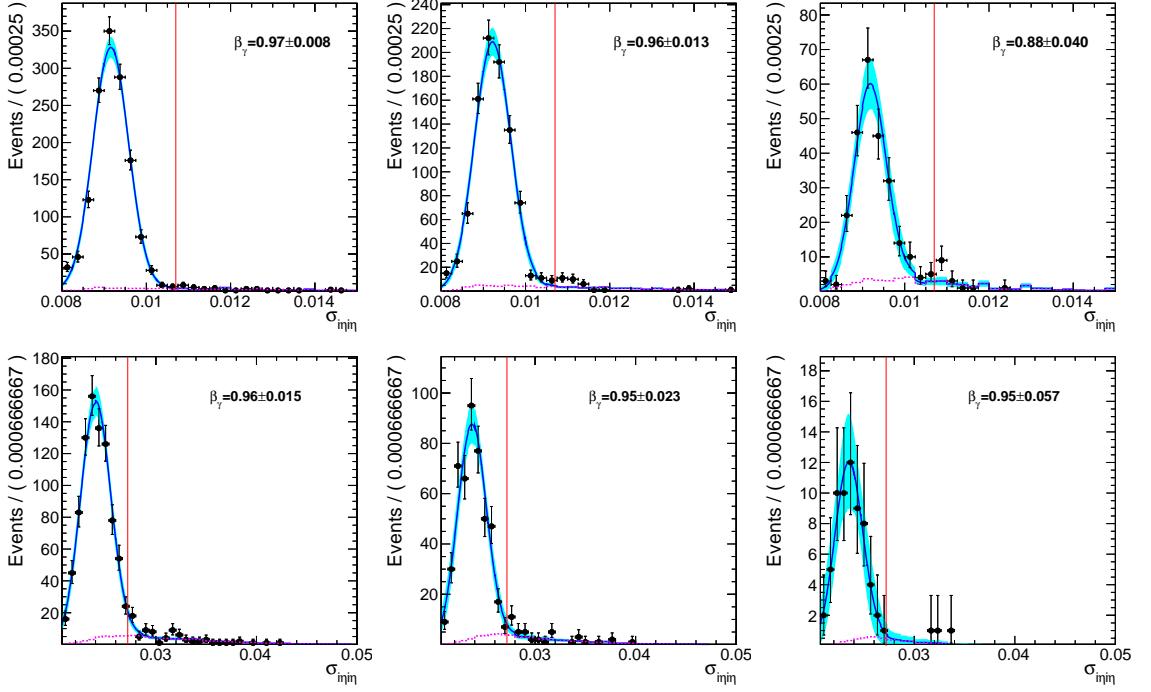


Figure 6.9: Fit results for purity determination in EB (top) and EE (bottom) for events with H_T^{miss} in the range 200-300, 300-500, and for $H_T^{\text{miss}} > 500$, respectively. β_γ is the fitted purity within the corresponding region. The red vertical line shows the $\sigma_{\text{inj}\eta\eta}$ requirement for the loose ID working point. The non-prompt PDF comes from the QCD simulation.

0.92 ± 0.08 .

Lost lepton: The SM events (mostly $t\bar{t}$ and W+jets) with muons or electrons can satisfy event selection and enter the signal sample as the lost-lepton background if the requirements for any of the following analysis steps are not satisfied:

- Kinematic acceptance,
- Reconstruction, or
- Isolation.

6 Search for SUSY with 13 TeV pp Collision Data

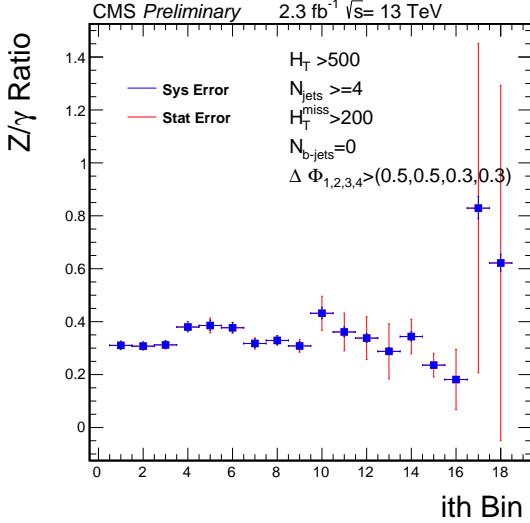


Figure 6.10: $\mathcal{R}_{Z(\nu\bar{\nu})/\gamma}$ for the 18 analysis bins used in the analysis after baseline selection. The units of H_T and H_T^{miss} are in GeV.

The basic idea behind our data-driven method for calculating the lost-lepton background is to select single-lepton control sample (CS) in the data, through inversion of the lepton vetoes, and to weight each CS event by a factor that represents the probability for a lost-lepton event to appear with the corresponding search-variables: H_T , H_T^{miss} , N_{jet} , and $N_{\text{b-jet}}$. The weights are determined through evaluation of the efficiencies for each analysis step. The weighted distributions of the search variables, summed over all CS events, define the predicted lost-lepton background in the respective search regions.

QCD: The H_T^{miss} in QCD multijet events is almost always due to a mismeasured jet in the event, thus the H_T^{miss} direction is usually close to the jet. The $\Delta\phi$ variable is the minimum azimuthal (ϕ) difference between H_T^{miss} and one of the four highest p_T jets.

The low $\Delta\phi$ region is significantly enriched in QCD events. The sample of events with the $\Delta\phi$ requirement inverted (i.e., $\Delta\phi_1 < 0.5$ or $\Delta\phi_2 < 0.5$ or $\Delta\phi_3 < 0.3$ or $\Delta\phi_4 < 0.3$) serves as the QCD control sample. The background at high $\Delta\phi$, is estimated from the QCD yield at low $\Delta\phi$ and a high-to-low ratio R^{QCD} for the QCD component. The

6 Search for SUSY with 13 TeV pp Collision Data

$\Delta\phi$ distribution shows that the R^{QCD} has some dependence on the search variables H_T , H_T^{miss} , and N_{jet} . We model this dependence by assuming it factorizes. That is, we assume the H_T dependence does not depend on H_T^{miss} or N_{jet} and vice versa.

Hadronic tau: To evaluate the $t\bar{t}$, single-top and $W+jets$ backgrounds that arises when a W boson decays to a hadronically decaying τ lepton (τ^h) and a neutrino, we employ a tau-template method. In this approach, the τ^h background is estimated from a control sample (CS) of $\mu+jets$ events, which we select by requiring exactly one muon with $p_T > 20$ GeV and $|\eta| < 2.1$. This single-muon CS is mainly composed of $t\bar{t}(\rightarrow \mu\nu)$ and $W(\rightarrow \mu\nu)+jets$ events. Since both $\mu+jets$ and τ^h+jets events arise from the same underlying process, the hadronic components of the two event classes are expected to be the same, aside from the response of the detector to a muon or a τ^h jet. The latter effect is incorporated by smearing the muon p_T in the CS events, with MC-derived response functions (the “templates”), in order to emulate the τ^h jet response. Global hadronic variables such as N_{jet} , H_T , and H_T^{miss} are then recomputed, with the full analysis procedure being applied.

6.1.3.2 Uncertainties

Various kinds of systematic and statistical uncertainties that are considered in the analysis. The uncertainties that result from the background estimations including data control region statistics, purity of control region , CR trigger efficiency etc. are discussed in details in Ref. [52]. Below we briefly describe various systematic uncertainties that affect the expected signal yield.

- **Luminosity:** A flat 4.6 % uncertainty on luminosity is propagated to the signal yield.

6 Search for SUSY with 13 TeV pp Collision Data

- **b-tag efficiency:** The b-tagging and mistagging scale factors are functions of the jet p_T and η . The scale factors are varied by their uncertainties and these variations are propagated as migrations between the different signal bins.
- **MC statistics:** The MC statistical uncertainties are propagated to the signal yield.
- **Trigger efficiency:** The trigger efficiencies are measured in the data. The effect of the statistical and systematic uncertainties is at most 1.1 % at low H_T^{miss} .
- **Pileup reweighting:** The uncertainties in the pileup reweighting correction are derived from the uncertainties in the minimum bias cross section and the difference between the expected and observed number of interactions in the data. The minimum bias cross section in the 13 TeV is estimated to be 69 mb with an uncertainty of $\pm 5\%$. The correction is varied according to these uncertainties, with a maximum effect of 0.5 %.
- **Scale:** The uncertainty is calculated from using the envelope of weights from varying the renormalization and factorization scales. The effect on the yield of non-compressed samples is less than 0.1% and on compressed samples ranges from 1% to 3%.
- **ISR:** The effect on the yield of non-compressed samples is less than 0.1 % and on compressed samples ranges from 3% to 11%.
- **Jet Energy Corrections:** The jet energy corrections (JECs) are varied using the p_T and η dependent jet energy scale uncertainties from the official database, with a separate set of corrections for the fast simulation samples. The overall effect ranges from 0.5% to 4%.

- **PDF:** The LHC4PDF prescription for the uncertainty on the total cross section is included as $\pm 1 \sigma$ bands in the results plots.

6.1.4 Results and Statistical Interpretation

Here we shall describe the interpretation of the data observed in the signal region. Fig. 6.11 shows the data vs. SM backgrounds in the 72 exclusive intervals as defined before. The ratio of likelihood functions, \mathcal{L} is profiled to compute the expected signal significance in units of standard deviations. The likelihood is constructed as the product of Poisson PDFs of observing N events, given a mean n, which depends on the floating parameters in the likelihood. The test statistic used in the profile likelihood calculations is given by

$$q_\mu = -2 \ln (\mathcal{L}_\mu / \mathcal{L}_{max})$$

where the \mathcal{L}_μ is the maximum likelihood at a given signal strength μ and \mathcal{L}_{max} is the maximum floating all parameters. This test statistic is the same as described in Chapter 5. The profile likelihood signal significance is given by $\sqrt{q_0}$ (described earlier in Sec 5.3) in terms of standard deviations.

For setting limits, we use the LHC-style CL_s [51] approach. The CL_s ratio is the ratio of confidence intervals:

$$CL_s = \frac{CL_{s+b}}{CL_b}$$

where CL_b is the confidence interval for the background only region and CL_{s+b} is the same in the background-plus-signal. The 95% CL is given by the probability that a test statistic Q is less than the observed value in data: $P_{95\%}(Q < Q_{obs})$, so it gives a probability of the data being discrepant with the background-only hypothesis. The test statistic Q comes from a modified Profile Likelihood with constraints on the upper bound

6 Search for SUSY with 13 TeV pp Collision Data

of the signal strength, so that the method gives an upper limit instead of a two-sided limit. The sensitivity of the analysis will scale with the exclusion power, so the 95% CL upper exclusion limits will decrease with increasing sensitivity to signal.

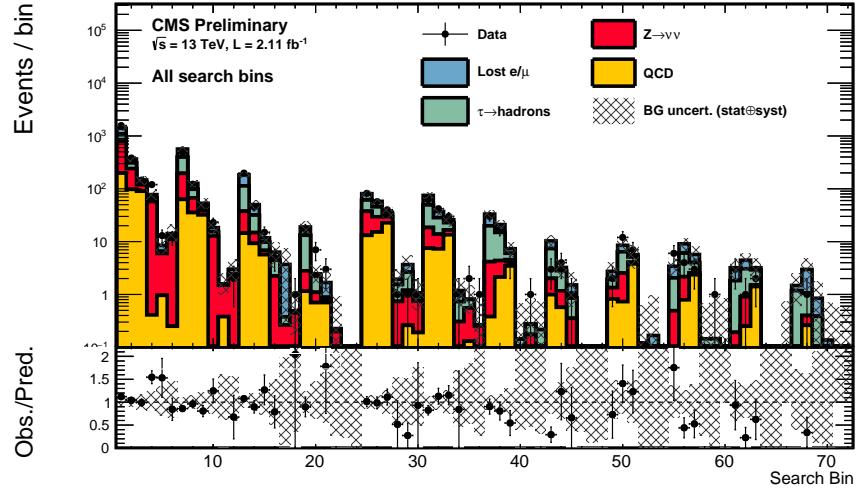


Figure 6.11: The observed data vs the SM backgrounds in the 72 exclusive intervals of the signal region.

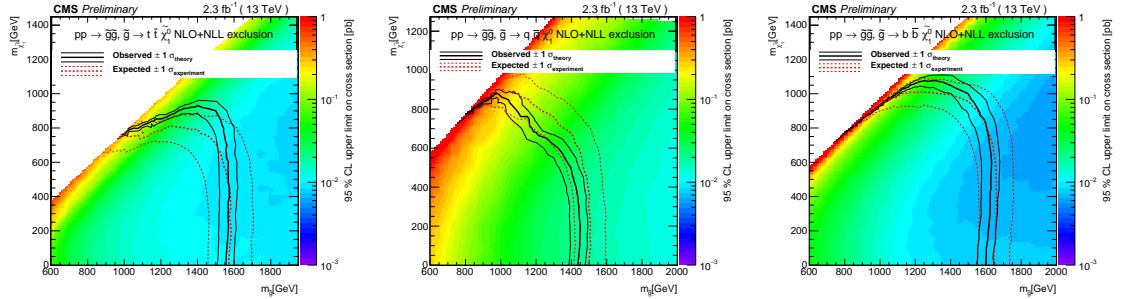


Figure 6.12: The 95% CL upper limits on the production cross sections for four top quark (left), four light quark (middle) and four b-quark (right) in the final state.

6.2 Study with 12.9 fb^{-1} of 2016 data

With the collection of more data in 2016 we update our analysis strategy to enhance the sensitivity. We perform the search in 2016 data using a strategy similar to that used in the 2015 analysis. Therefore, we do not discuss everything in detail, but rather focus on the major changes before moving over to results. A detailed exposition could be found in Ref. [55]. The three most significant changes to the analysis strategy are:

- Expanding the search region to include events with 3 or more jets. The 2015 analysis considered events with 4 or more jets. A more inclusive selection allows us to target models of direct squark production as well as other low-multiplicity signatures.
- Adopting triggers that make online requirements only on H_T^{miss} and MET. The trigger has online PF based MET $> 100 \text{ GeV}$ and $H_T^{\text{miss}} > 100 \text{ GeV}$. These triggers allow us to lower the offline H_T threshold from 500 to 300 GeV, giving us increased sensitivity to SUSY models with low jet multiplicity or compressed spectra.
- Using a finer binning in H_T , H_T^{miss} , and N_{jet} to take advantage of a much larger dataset in 2016.

Apart from these, there are a number of less important changes in the background estimation methods. Below we discuss in detail about the changes that occurred to the $Z(\rightarrow \nu\bar{\nu}) + \text{jets}$ background.

6.2.1 Search Binning

As with more data finer bins are possible , we further split the available phase-space into 160 bins in the following way. A schematic diagram in the 2D plane of H_T and H_T^{miss} is

6 Search for SUSY with 13 TeV pp Collision Data

shown in Fig. 6.2.1.

- **4** N_{jet} bins: 3-4, 5-6, 7-8, ≥ 9 ;
- **4** $N_{\text{b-jet}}$ bins: 0, 1, 2, ≥ 3 ; and
- **10** bins in H_T and H_T^{miss} as defined below:

Bin	H_T range [GeV]	H_T^{miss} range [GeV]
1:	300-500	300-350
2:	500-1000	300-350
3:	>1000	300-350
4:	350-500	350-500
5:	500-1000	350-500
6:	>1000	350-500
7:	500-1000	500-750
8:	>1000	500-750
9:	750-1500	>750
10:	>1500	>750

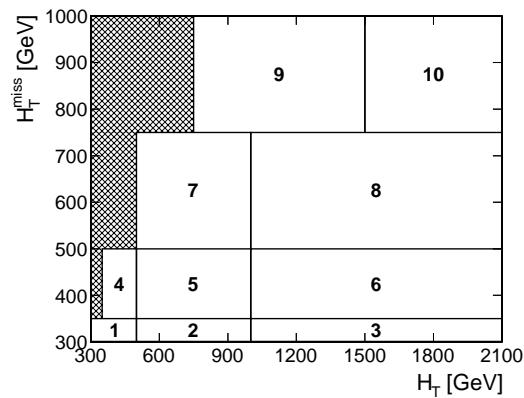


Figure 6.13: Search binning in the 2D plane of H_T and H_T^{miss} used in the 2016 analysis.

6.2.2 $Z(\rightarrow \nu\bar{\nu})$ +jets Background Estimation

Our baseline strategy is the same as that used in the 2015 search discussed earlier. We use the γ +jets sample to determine the yields in the 40 bins corresponding to $N_{b\text{-jet}} = 0$. These are then compared with the $Z(\rightarrow \ell^+\ell^-)$ yields in the low- N_{jet} bin to establish the systematic uncertainty of the physics modeling of γ +jets, and the normalization corrected if necessary. The extrapolation to bins with $N_{b\text{-jet}} > 0$ is performed to the extent possible with the $Z(\rightarrow \ell^+\ell^-)$ data sample, supplemented with MC information where necessary. The photon (dilepton) samples require a photon (dilepton) candidate with $p_T > 200\text{GeV}$. The increase in the p_T cut is due to trigger reasons. We show the data vs. MC expectation in the 40 0-btagged bins in the Fig. 6.14. In Table 6.11 we present the estimated Z background in those 40 bins.

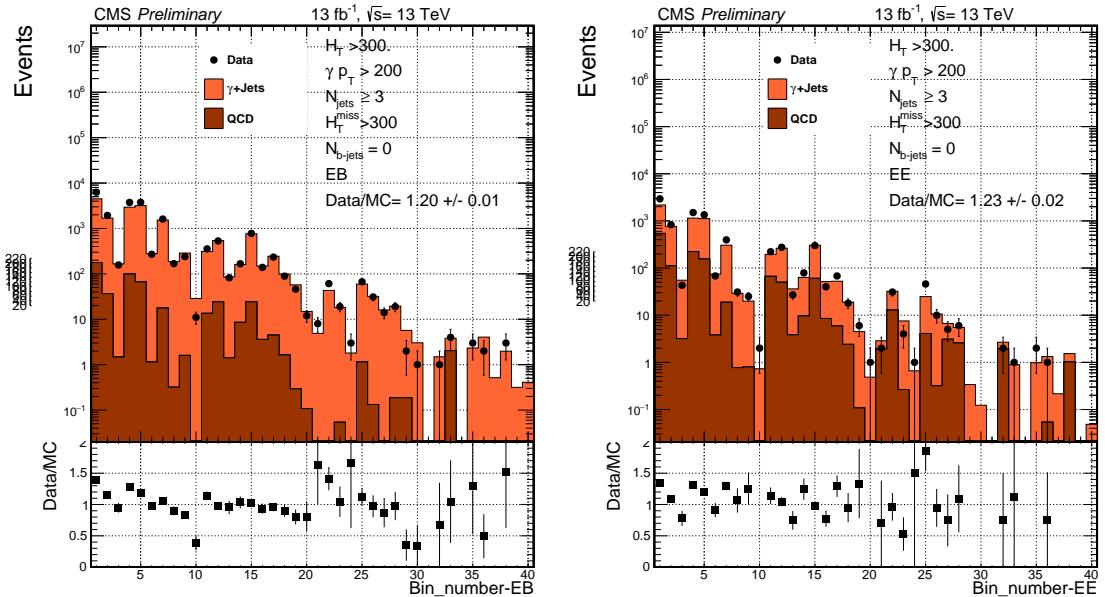


Figure 6.14: Numbers of observed events in the photon control samples in EB (left) and EE (right) compared to simulation. The units of H_T and H_T^{miss} are in GeV.

The background estimated from γ +jets are used to extrapolate to higher b-tagged

6 Search for SUSY with 13 TeV pp Collision Data

bins using a method described in Ref. [55]. Without discussing further on other background methods, we will go to result directly.

6.2.3 Results and Statistical Interpretation

This section contains the results of the search. The observations in the signal regions are in general good agreement with the predicted backgrounds. For the 160 search bins, the observed data and the pre-fit predictions for each background component are shown in Fig. 6.15. The statistical procedures used to calculate limits are the same as described in Chapter 5. The upper limits on the signal strengths are shown in Fig. 6.16. As could be seen from Fig. 6.15, no significant excess is found beyond the SM expectations. For the simplified models considered here, we exclude gluino masses up to the range 1600-1700 GeV as could be seen in Fig. 6.16. In general, the observed exclusion is weaker than expected because of the small excess (not significant) of events observed in several bins.

6 Search for SUSY with 13 TeV pp Collision Data

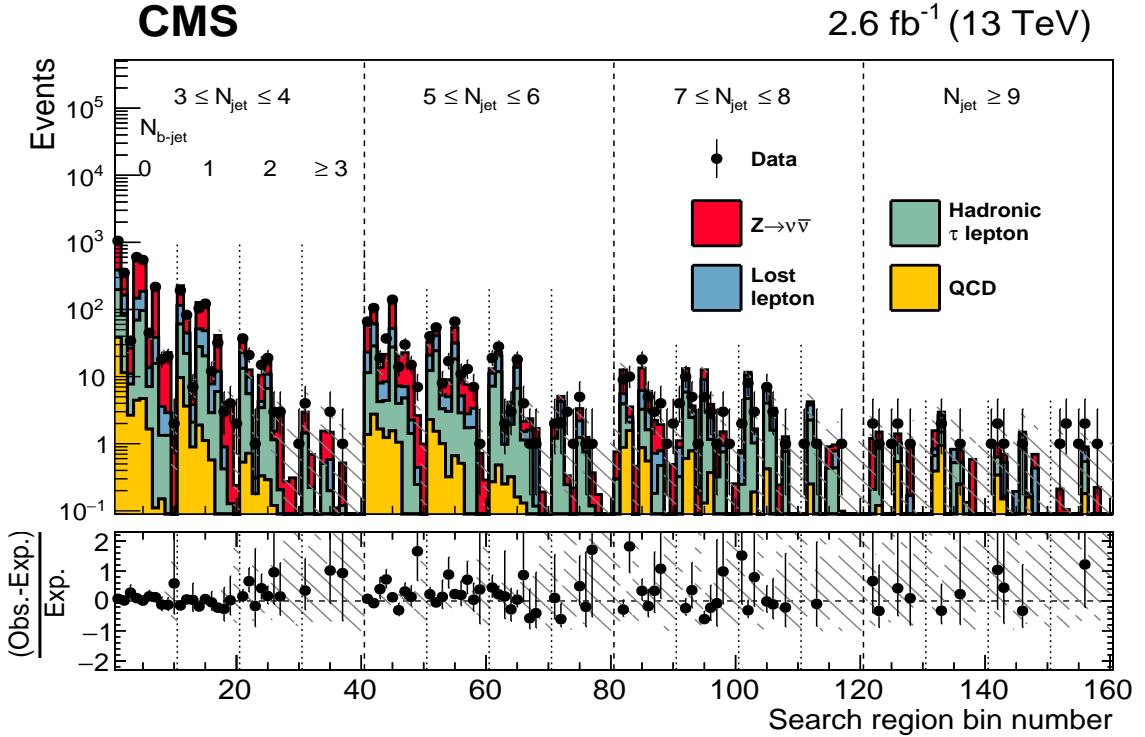


Figure 6.15: The observed data vs the SM backgrounds in the 160 exclusive intervals of the signal region.

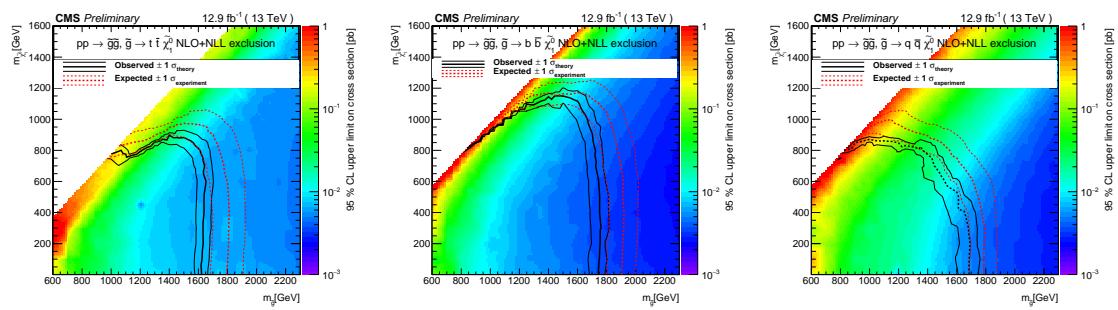


Figure 6.16: The 95% CL upper limits on the production cross sections for four top quark (left), four light quark (middle) and four b-quark (right) in the final state.

6 Search for SUSY with 13 TeV pp Collision Data

Table 6.11: $Z \rightarrow \nu\nu + \text{jets}$ prediction ($N_{Z \rightarrow \nu\nu}^{\text{pred}}$) in the 40 0-btag analysis bins as determined with the $\gamma + \text{jets}$ method. We show the inputs to the determination of $N_{Z \rightarrow \nu\nu}^{\text{pred}}$ as that vary for the 40 bins, including N_{γ}^{obs} in EB and EE, $R_{Z/\gamma}$, and double ratio. The first uncertainty on $R_{Z/\gamma}$ comes from the statistical uncertainty on the simulation and the second from variation of the data-MC scale factors for photon reconstruction, ID, and isolation.

Bin	$N_{\gamma, \text{EB}}^{\text{obs}}$	$N_{\gamma, \text{EE}}^{\text{obs}}$	$R_{Z/\gamma}$	Double ratio	$N_{Z \rightarrow \nu\nu}^{\text{pred}}$
1 (N_{jet} 3-4, MHT_HT_1)	6287	2935	$0.413 \pm 0.004^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.024}_{-0.000}$	$3146.2 \pm 32.8^{+216.6}_{-189.5}$
2 (N_{jet} 3-4, MHT_HT_2)	1936	830	$0.374 \pm 0.004^{+0.020}_{-0.018}$	$0.966 \pm 0.032^{+0.022}_{-0.000}$	$856.5 \pm 16.3^{+58.6}_{-51.7}$
3 (N_{jet} 3-4, MHT_HT_3)	157	43	$0.378 \pm 0.007^{+0.020}_{-0.018}$	$0.966 \pm 0.032^{+0.122}_{-0.089}$	$63.0 \pm 4.5^{+9.0}_{-7.0}$
4 (N_{jet} 3-4, MHT_HT_4)	3747	1502	$0.430 \pm 0.005^{+0.023}_{-0.021}$	$0.966 \pm 0.032^{+0.021}_{-0.000}$	$1910.2 \pm 26.4^{+122.5}_{-106.8}$
5 (N_{jet} 3-4, MHT_HT_5)	3749	1347	$0.407 \pm 0.004^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.018}_{-0.000}$	$1757.4 \pm 24.6^{+110.2}_{-97.2}$
6 (N_{jet} 3-4, MHT_HT_6)	270	68	$0.403 \pm 0.006^{+0.021}_{-0.019}$	$0.966 \pm 0.032^{+0.122}_{-0.089}$	$115.9 \pm 6.3^{+16.3}_{-12.5}$
7 (N_{jet} 3-4, MHT_HT_7)	1612	396	$0.429 \pm 0.004^{+0.023}_{-0.021}$	$0.966 \pm 0.032^{+0.062}_{-0.033}$	$749.9 \pm 16.7^{+63.8}_{-46.3}$
8 (N_{jet} 3-4, MHT_HT_8)	168	31	$0.451 \pm 0.008^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.122}_{-0.089}$	$78.3 \pm 5.6^{+10.9}_{-8.3}$
9 (N_{jet} 3-4, MHT_HT_9)	240	25	$0.443 \pm 0.006^{+0.023}_{-0.021}$	$0.966 \pm 0.032^{+0.160}_{-0.102}$	$102.6 \pm 6.3^{+17.9}_{-12.1}$
10 (N_{jet} 3-4, MHT_HT_10)	11	2	$0.515 \pm 0.025^{+0.027}_{-0.025}$	$0.966 \pm 0.032^{+0.209}_{-0.156}$	$5.8 \pm 1.6^{+1.3}_{-1.0}$
11 (N_{jet} 5-6, MHT_HT_1)	355	222	$0.393 \pm 0.015^{+0.021}_{-0.019}$	$0.966 \pm 0.032^{+0.059}_{-0.033}$	$186.4 \pm 7.8^{+17.8}_{-14.5}$
12 (N_{jet} 5-6, MHT_HT_2)	530	276	$0.382 \pm 0.008^{+0.020}_{-0.018}$	$0.966 \pm 0.032^{+0.059}_{-0.033}$	$254.0 \pm 8.9^{+23.0}_{-18.2}$
13 (N_{jet} 5-6, MHT_HT_3)	82	27	$0.375 \pm 0.009^{+0.020}_{-0.018}$	$0.966 \pm 0.032^{+0.122}_{-0.089}$	$34.0 \pm 3.3^{+4.9}_{-3.8}$
14 (N_{jet} 5-6, MHT_HT_4)	167	79	$0.357 \pm 0.020^{+0.019}_{-0.017}$	$0.966 \pm 0.032^{+0.059}_{-0.033}$	$74.1 \pm 4.7^{+7.5}_{-6.3}$
15 (N_{jet} 5-6, MHT_HT_5)	780	300	$0.410 \pm 0.007^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.059}_{-0.033}$	$374.8 \pm 11.4^{+32.5}_{-25.0}$
16 (N_{jet} 5-6, MHT_HT_6)	139	40	$0.415 \pm 0.008^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.122}_{-0.089}$	$63.1 \pm 4.7^{+8.9}_{-6.9}$
17 (N_{jet} 5-6, MHT_HT_7)	234	68	$0.422 \pm 0.009^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.062}_{-0.033}$	$110.8 \pm 6.4^{+9.7}_{-7.2}$
18 (N_{jet} 5-6, MHT_HT_8)	89	18	$0.449 \pm 0.011^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.122}_{-0.089}$	$41.9 \pm 4.0^{+5.9}_{-4.5}$
19 (N_{jet} 5-6, MHT_HT_9)	46	6	$0.442 \pm 0.015^{+0.023}_{-0.021}$	$0.966 \pm 0.032^{+0.160}_{-0.102}$	$20.1 \pm 2.8^{+3.6}_{-2.4}$
20 (N_{jet} 5-6, MHT_HT_10)	12	1	$0.462 \pm 0.031^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.209}_{-0.156}$	$5.3 \pm 1.5^{+1.2}_{-1.0}$
21 (N_{jet} 7-8, MHT_HT_1)	8	2	$0.454 \pm 0.126^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.127}_{-0.085}$	$3.8 \pm 1.2^{+1.2}_{-1.1}$
22 (N_{jet} 7-8, MHT_HT_2)	61	31	$0.375 \pm 0.025^{+0.020}_{-0.018}$	$0.966 \pm 0.032^{+0.127}_{-0.085}$	$28.5 \pm 3.0^{+4.6}_{-3.6}$
23 (N_{jet} 7-8, MHT_HT_3)	19	4	$0.410 \pm 0.021^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.127}_{-0.089}$	$7.9 \pm 1.6^{+1.2}_{-1.0}$
24 (N_{jet} 7-8, MHT_HT_4)	3	1	$0.688 \pm 0.319^{+0.036}_{-0.033}$	$0.966 \pm 0.032^{+0.127}_{-0.085}$	$2.3 \pm 1.2^{+1.1}_{-1.1}$
25 (N_{jet} 7-8, MHT_HT_5)	67	46	$0.407 \pm 0.022^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.127}_{-0.085}$	$38.6 \pm 3.6^{+6.0}_{-4.5}$
26 (N_{jet} 7-8, MHT_HT_6)	31	10	$0.421 \pm 0.017^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.127}_{-0.089}$	$14.6 \pm 2.3^{+2.2}_{-1.7}$
27 (N_{jet} 7-8, MHT_HT_7)	14	5	$0.411 \pm 0.032^{+0.022}_{-0.020}$	$0.966 \pm 0.032^{+0.127}_{-0.085}$	$6.8 \pm 1.6^{+1.1}_{-0.9}$
28 (N_{jet} 7-8, MHT_HT_8)	19	6	$0.474 \pm 0.026^{+0.025}_{-0.023}$	$0.966 \pm 0.032^{+0.127}_{-0.089}$	$10.3 \pm 2.1^{+1.6}_{-1.2}$
29 (N_{jet} 7-8, MHT_HT_9)	2	0	$0.502 \pm 0.053^{+0.027}_{-0.024}$	$0.966 \pm 0.032^{+0.160}_{-0.102}$	$0.9 \pm 0.6^{+0.2}_{-0.1}$
30 (N_{jet} 7-8, MHT_HT_10)	1	0	$0.458 \pm 0.066^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.209}_{-0.156}$	$0.4 \pm 0.4^{+0.1}_{-0.1}$
31 (N_{jet} 9+, MHT_HT_1)	0	0	$0.455 \pm 0.202^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$0.4 \pm 0.4^{+0.2}_{-0.2}$
32 (N_{jet} 9+, MHT_HT_2)	1	2	$0.268 \pm 0.054^{+0.014}_{-0.013}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$0.6 \pm 0.4^{+0.2}_{-0.2}$
33 (N_{jet} 9+, MHT_HT_3)	4	1	$0.565 \pm 0.086^{+0.030}_{-0.027}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$2.4 \pm 1.1^{+0.6}_{-0.5}$
34 (N_{jet} 9+, MHT_HT_4)	0	0	$0.455 \pm 0.202^{+0.024}_{-0.022}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$0.4 \pm 0.4^{+0.2}_{-0.2}$
35 (N_{jet} 9+, MHT_HT_5)	3	2	$0.364 \pm 0.064^{+0.019}_{-0.017}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$1.5 \pm 0.7^{+0.4}_{-0.4}$
36 (N_{jet} 9+, MHT_HT_6)	2	1	$0.430 \pm 0.049^{+0.023}_{-0.021}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$1.1 \pm 0.6^{+0.3}_{-0.2}$
37 (N_{jet} 9+, MHT_HT_7)	0	0	$0.258 \pm 0.082^{+0.014}_{-0.012}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$0.2 \pm 0.2^{+0.1}_{-0.1}$
38 (N_{jet} 9+, MHT_HT_8)	3	0	$0.484 \pm 0.79^{+0.026}_{-0.023}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$1.3 \pm 0.7^{+0.3}_{-0.3}$
39 (N_{jet} 9+, MHT_HT_9)	0	0	$0.265 \pm 0.109^{+0.014}_{-0.013}$	$0.966 \pm 0.032^{+0.193}_{-0.133}$	$0.2 \pm 0.2^{+0.1}_{-0.1}$
40 (N_{jet} 9+, MHT_HT_10)	0	0	$0.350 \pm 0.134^{+0.019}_{-0.017}$	$0.966 \pm 0.032^{+0.209}_{-0.156}$	$0.3 \pm 0.3^{+0.1}_{-0.1}$

7 Search for Pair Production of First Generation Leptoquarks

Leptoquark pair-production at the LHC proceeds primarily through gluon-gluon fusion, with a smaller contribution from quark-antiquark annihilation. Here, the leptoquark mass (M_{LQ}) and the Yukawa coupling λ to leptons and quarks are relevant production and decay parameters. The production cross section is nearly independent of λ .

A leptoquark will decay into a lepton and a quark, giving rise to a final state containing high-momentum leptons and jets. The decay into a charged lepton and a quark has branching fraction β , while the decay into a neutrino and a quark has branching fraction $1 - \beta$ (β is a free parameter). Previous searches for rare processes mediated by lepton number violation, and flavor-changing neutral currents prefer leptoquarks at LHC-accessible mass ranges that couple to leptons and quarks of the same generation without intergenerational mixing. Therefore, we assume first-generation leptoquarks decay to electrons and electron neutrinos, leading to following three possible final states:

- Two electrons and two jets: each leptoquark decays into an electron and a quark;
- One electron, MET, and two jets: one leptoquark decays into an electron and a quark, while the other decays into a an electron neutrino and a quark; and

7 Search for Pair Production of First Generation Leptoquarks

- No electrons, MET, and two jets: each leptoquark decays into a neutrino and a quark.

The final state branching fractions are β^2 , $2 \times \beta(1 - \beta)$, and $(1 - \beta)^2$, respectively. We consider two values of β : 1, where leptoquarks always decaying to the first final state; and 0.5, where 50% of leptoquarks decay to the second final state, with 25% decaying to the first and 25% decaying to the third final states. The first and second final states are denoted as $eejj$ and $e\nu jj$. In this analysis, we search for leptoquark decays in the first final state only, $eejj$.

Previous searches have not detected any leptoquarks. Recently, CMS published a search [56] using about 20 fb^{-1} of 2012 LHC data recorded at $\sqrt{s} = 8 \text{ TeV}$, which gave first generation leptoquark mass limits of less than 1010 (850) GeV for $\beta = 1$ (0.5). Similarly, ATLAS used the 2012 LHC data to set first generation leptoquark mass limits of less than 1050 (900) GeV for $\beta = 1$ (0.5), and also looked at the possibility of $\beta = 0.2$, which yielded a mass limit of 900 GeV [57]. All of the above limits are set at the 95% CL.

7.1 Event Selection

The $eejj$ final state is the result of two pair-produced leptoquarks each decaying to an electron and a jet, where the branching fraction β as defined previously is equal to 1. Events containing at least two electrons and two jets are selected, where the two leading p_T electrons and jets are used in the analysis.

7 Search for Pair Production of First Generation Leptoquarks

7.1.0.1 Preselection

The preselection is aimed at selecting a background-dominated event sample in order to compare MC simulated background predictions with the data. The corresponding selection requirements that the candidate event should

- Pass the signal trigger satisfying the following criteria
 - Single electron with online p_T greater than 27 GeV, and
 - it has to satisfy loose identification criteria (loose means, loose compared to the offline HEEP criteria which will be defined later),
- Contain exactly two electrons passing HEEP (High Energy Electrons) ID, additionally having ECAL cluster $E_T > 50$ GeV and ECAL pseudorapidity $|\eta| < 2.5$ (HEEP selection criteria as listed in Table 7.1 gives high efficiency for high energy electrons),
- Contains no muons, where muons are selected with $p_T > 45$ GeV and $|\eta| < 2.1$,
- Contains at least two jets with $p_T > 50$ GeV and $|\eta| < 2.4$,
- Have $m_{ee} > 50$ GeV, where m_{ee} is the dielectron invariant mass, and
- Have $S_T > 300$ GeV, where S_T is the scalar sum of p_T of the two electrons and the two leading jets.

After applying these criteria, the overall MC background agreement with the data is examined in a series of plots, we find a quite good agreement between the two. The data-MC agreement in terms of the last two variables above could be seen in Fig. 7.1. The ratio plots have been provided for a better view of the data-MC agreement.

7 Search for Pair Production of First Generation Leptoquarks

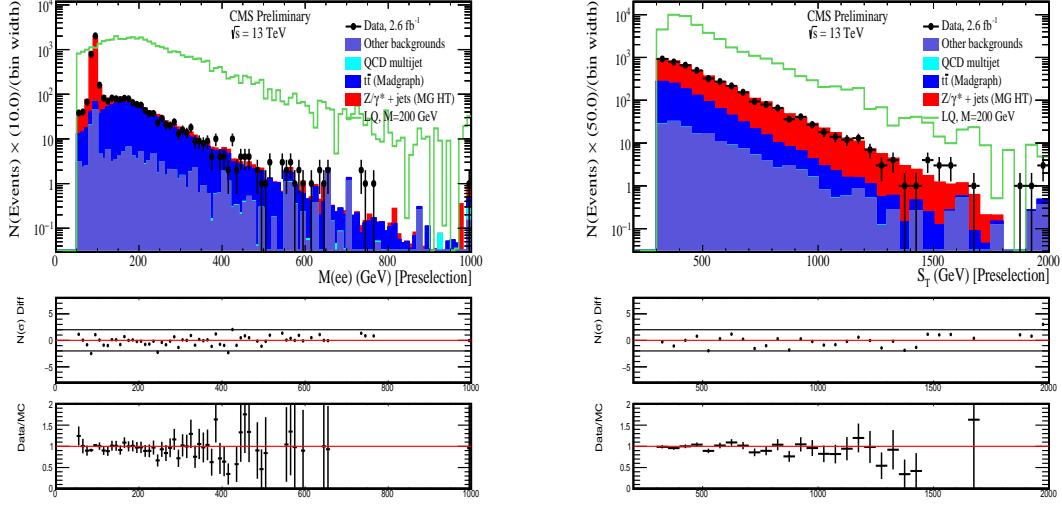


Figure 7.1: Data vs. background plots for M_{ee} (left) and S_T (right) after Preselection

7.1.0.2 Final Selection

In order to maximize the signal sensitivity, we have optimized a set of cuts on three variables. They are

- S_T ,
- m_{ee} and
- m_{ej}^{min} : As each leptoquark decays to an electron and a quark, when we choose the correct e and j pair, mass difference from the two reconstructed LQs should be very small. Out of two electrons and two jets we make four combinations to see which combinations gives minimum difference between the $m_{e,j}$ mass, the invariant mass of that particular combination is m_{ej}^{min} .

The variables S_T , m_{ee} , and m_{ej} are varied independently over a broad range of values, and the thresholds which maximize the optimization metric are obtained for each leptoquark mass point. The considered metric is called ‘Punzi’s significance [58] as given in Eq. 7.1, where t represents an attempted set of cuts, $\epsilon(t)$ is the efficiency times

7 Search for Pair Production of First Generation Leptoquarks

Table 7.1: HEEP electron identification criteria used for the analysis.

Variable	Barrel Threshold	Endcap Threshold
η	$ \eta_{SC} < 1.4442$	$1.566 < \eta_{SC} < 2.5$
isEcalDriven	yes	yes
Inner layer lost hits	≤ 1	≤ 1
$ d_{xy} $ [cm]	< 0.02	< 0.05
$ \Delta\eta_{in}^{seed} $	< 0.004	< 0.006
$ \Delta\phi_{in} $	< 0.06	< 0.06
H/E	$< 1/E + 0.05$	$< 5/E + 0.05$
$\sigma_{in\eta}^{full5\times 5}$	n/a	< 0.03
Shower shape	$E^{2\times 5}/E^{5\times 5} > 0.94$ OR $E^{1\times 5}/E^{5\times 5} > 0.83$	–
Calo iso. [GeV]	$< 2 + 0.03 \times E_T + 0.28 \times \rho$	$\begin{cases} < 2.5 + 0.28 \times \rho & \text{if } E_T < 50. \\ < 2.5 + 0.28 \times \rho + 0.03 \times (E_T - 50.) & \text{if } E_T \geq 50. \end{cases}$
Tracker iso. [GeV]	< 5	< 5

acceptance of the signal, a is the one-sided significance required in standard deviations, and $B(t)$ is the expected number of background events.

$$\frac{\epsilon(t)}{a/2 + \sqrt{B(t)}} \quad (7.1)$$

We choose $a = 5$ to require 5σ discovery significance. A unique property of the Punzi significance is that it does not depend on the signal cross section, but only on its acceptance times efficiency. The optimized set of cuts are shown in the Table 7.2.

7.2 Backgrounds

The major backgrounds from SM processes are $Z+jets$ and $t\bar{t}$, whereas single top, $W+jets$, diboson, and $\gamma+jets$ contribute at a lower level. There is also an instrumental background from QCD multijet events with jets faking as electrons and passing the HEEP ID. Below we describe briefly about the estimation of QCD background as this

7 Search for Pair Production of First Generation Leptoquarks

is the part where we have maximally contributed. The QCD contribution to the signal region is small compared to $Z + \text{jets}$ and $t\bar{t}$, however as QCD is one the less understood process, so a careful attention is needed to estimate background. We employ a fully data driven QCD estimation.

7.2.1 QCD multijet:

QCD events enters our signal region because of misidentification of jets are electrons. In other word, our HEEP criteria have a finite fake rate (FR). We define FR as the probability of a jet to be misidentified as an electron. Mathematically,

$$FR = \frac{N_{\text{pass,HEEP,jets}}}{N_{\text{total,jets}}} \quad (7.2)$$

As our signal region requires two electrons, we take a control sample of two electrons with very loose criteria imposed on them. Basically, the control sample becomes a jet dominated region. Then we estimate the number of QCD events that contribute to the signal region

$$N_{\text{QCD}}^{eejj} = \sum_{\text{events}} FR_1 \otimes FR_2 \quad (7.3)$$

where FR_1 and FR_2 are the fake rates of the two electrons. More details on the fake rate calculation can be found in Ref. [74].

7.3 Results and Statistical Interpretation

Similar to the two SUSY analyses [52; 55], we observe no significant excess of events as compared to the SM backgrounds. A broad excess of the events that was seen in the 8

7 Search for Pair Production of First Generation Leptoquarks

TeV analysis [56] has disappeared with the 2015 data. The data-MC agreement after the final selection for the LQ mass point of 650 could be seen from Fig. 7.2. We set the upper limits on the cross section times branching fraction using the LHC style approach. Fig. 7.3 shows the observed and expected limits for different LQ masses. We exclude LQ masses up to 1130 GeV from this study.

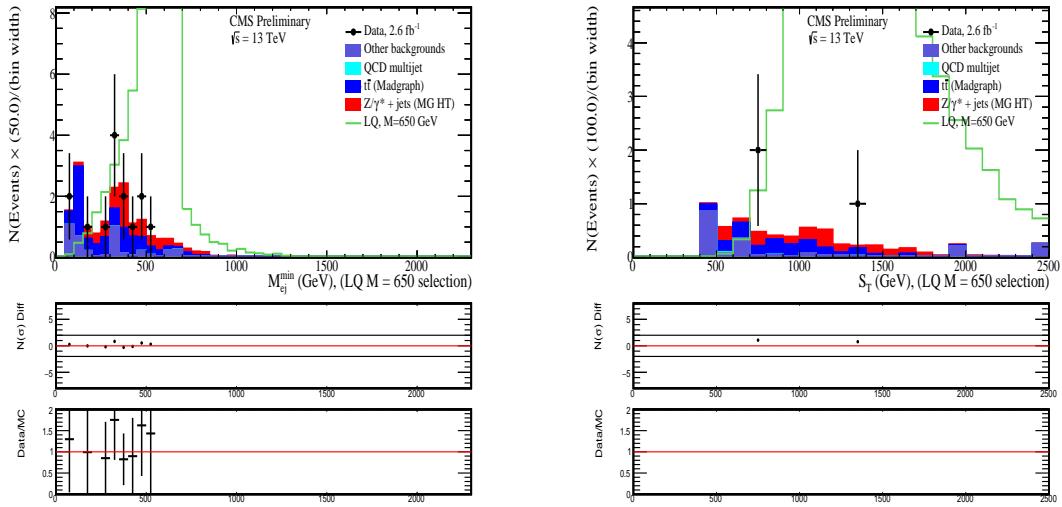


Figure 7.2: Data vs. MC for the LQ mass point 650 GeV in M_{ej}^{min} (left) and S_T (right).

7 Search for Pair Production of First Generation Leptoquarks

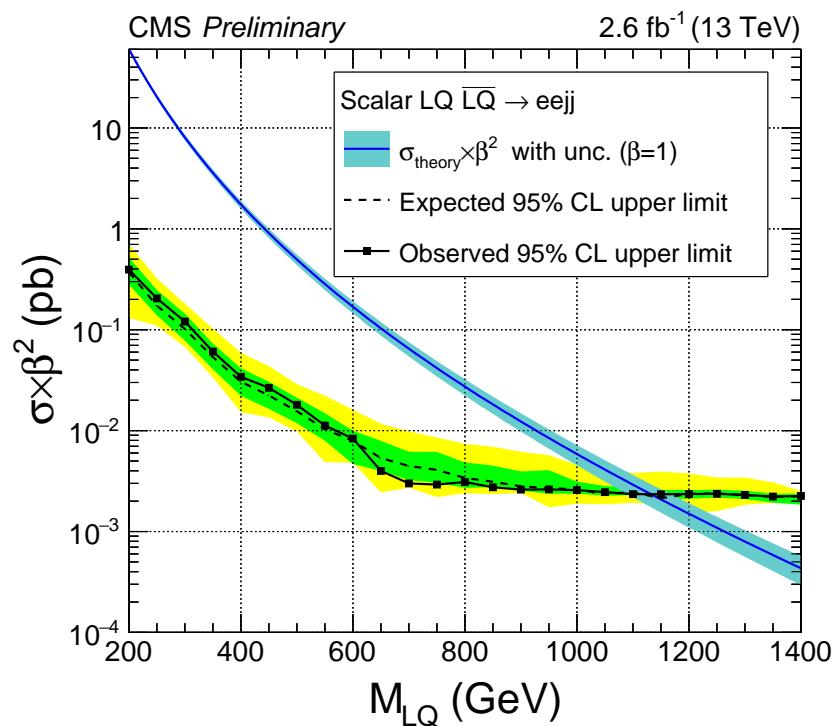


Figure 7.3: The 95% CL upper limits on the production cross sections as a function of the leptoquark mass.

7 Search for Pair Production of First Generation Leptoquarks

Table 7.2: Optimized final selection thresholds for the $eejj$ analysis.

M_{LQ} [GeV]	S_T [GeV] >	m_{ee} [GeV] >	m_{ej}^{min} [GeV] >
200	340	130	160
250	405	140	205
300	470	155	245
350	535	165	285
400	595	175	325
450	660	185	360
500	720	195	400
550	780	205	435
600	840	210	470
650	900	220	500
700	960	230	535
750	1015	235	565
800	1075	245	595
850	1130	250	625
900	1190	255	650
950	1245	265	675
1000	1300	270	705
1050	1355	275	725
1100	1410	280	750
1150	1460	285	775
1200	1515	285	795
1250	1565	290	815
1300	1615	295	830
1350	1670	300	850
1400	1720	300	865
1450	1770	300	880
≥ 1500	1815	305	895

8 Advanced Pileup Mitigation

Techniques

The performance of jets is extremely important to the success of most physics analyses. One of the outstanding challenges of the LHC run is the increase of instantaneous luminosity, which results in a large number of additional proton-proton collisions in each event(pileup). In such high pileup environment, the accurate reconstruction of jet properties and shapes becomes more and more demanding. The most common observables for jets are their p_T , η , and ϕ . Pileup changes the structure of jets by adding a lot of unwanted soft and wide-angle particles to them. In those cases the internal substructure of the jet contains information about the pileup contamination. In this part of the study, we have exploited the internal structure of the jet to get rid of the unwanted contaminations coming from pileup. The commonly used term ‘jet grooming’ has become increasingly popular in recent times. Grooming has been introduced for the first time in Ref. [75] and is intended to remove soft and wide-angle radiations from the jet. It is typically used to reduce the overall jet mass of quark- and gluon-initiated jets while retaining the larger jet mass for jets originating from decays of heavy particles such as the top quark and W/Z/H bosons. In presence of pileup, grooming can be used to

8 Advanced Pileup Mitigation Techniques

reduce the dependence of jet mass on pileup. It should be noted that jet grooming, in general, alters the soft structure of the jet while other jet structure observables may rely on this soft structure.

In this section, we explore the effect of grooming as a pileup mitigation method on large-radius jets ($R=0.8, 1.2$) summarizing our work from Ref. [59]. We consider the following grooming methods: pruning [60], trimming [61], modified mass drop tagger [62], and soft drop [63].

The pruning algorithm reclusters the constituents of the jet through the Cambridge-Aachen (CA) algorithm [76], using the same distance parameter as of the jet. At each step in the clustering algorithm, the softer of the two particles i and j to be merged is removed when the following conditions are met:

$$z_{ij} = \frac{\min(p_T_i, p_T_j)}{p_T_i + p_T_j} < z_{\text{cut}} \quad (8.1)$$

$$\Delta R_{ij} = \frac{2 \times r_{\text{cut}} \times m_J}{p_T} > D_{\text{cut}} \quad (8.2)$$

where m_J and p_T are the mass and transverse momentum of the originally-clustered jet, and z_{cut} and r_{cut} are parameters of the algorithm. A pictorial of pruning algorithm is shown in Fig. 8.1. A point to be noted is that in our case the reclustering is done using CA, even though in the picture two options of reclustering are mentioned.

Trimming ignores particles within a jet that fall below a dynamic threshold in p_T . It reclusters the constituents of the jet using the k_t algorithm [65] with a radius R_{sub} , accepting only the subjets that have $p_{T_{\text{sub}}} > f_{\text{cut}} \lambda_{\text{hard}}$, where f_{cut} is a dimensionless cutoff

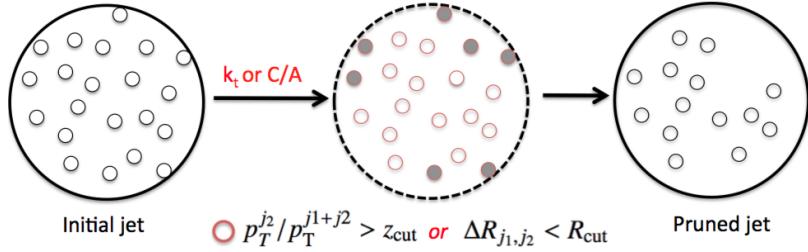


Figure 8.1: Pictorial representation of the pruning algorithm. The image is taken from Ref. [64].

parameter, and λ_{hard} is some hard QCD scale chosen to equal the p_T of the original jet.

Soft-drop and modified mass drop tagger decluster the jet recursively removing soft and wide-angle radiation from the jet. The jet is reclustered using the CA algorithm. Then the jet is declustered and at each step, subjets j_1 and j_2 are defined to test the following condition:

$$\frac{\min(p_{T,j_1}, p_{T,j_2})}{p_{T,j_1} + p_{T,j_2}} > z_{\text{cut}} \times \left(\frac{\Delta R_{12}}{R_0} \right)^\beta$$

where the algorithm parameters are z_{cut} and β . If the condition is met, the declustering continues, otherwise only the leading p_T subjet is kept. In the case when $\beta = 0$, soft drop can be considered a generalization of the modified mass drop tagger.

The (groomed) masses are corrected for pileup using a four-vector safe subtraction [66]. In the cases of soft drop, modified mass drop tagger, and trimming, the four-vector subtraction corrects the jet p_T and mass at each step in the algorithm. For pruning, the correction is applied to the final product using the pruned jet area. The parameters for the grooming algorithms explored in this study are listed in Table 8.1.

8.0.1 Simulation Results at High Pileup

The performance of various grooming algorithms is obtained at higher pileup scenarios expected in LHC Run II. For a given p_T bin, we obtain the jet mass resolution for

8 Advanced Pileup Mitigation Techniques

Table 8.1: Summary of grooming parameters.

Grooming algorithm	Parameters
Pruning	$z_{\text{cut}} = 0.1, r_{\text{cut}} = 0.5$ $z_{\text{cut}} = 0.05, r_{\text{cut}} = 0.5$ $z_{\text{cut}} = 0.1, r_{\text{cut}} = 0.75$ $z_{\text{cut}} = 0.05, r_{\text{cut}} = 0.75$
Trimming	$r_{\text{sub}} = 0.2, \text{pT}_{\text{frac}} = 0.05$ $r_{\text{sub}} = 0.2, \text{pT}_{\text{frac}} = 0.03$ $r_{\text{sub}} = 0.1, \text{pT}_{\text{frac}} = 0.03$ $r_{\text{sub}} = 0.3, \text{pT}_{\text{frac}} = 0.03$
Soft drop/MMDT	$z_{\text{cut}} = 0.1, \beta = -1$ $z_{\text{cut}} = 0.1, \beta = 0$ $z_{\text{cut}} = 0.1, \beta = 1$ $z_{\text{cut}} = 0.1, \beta = 2$

both background QCD and signal W jets. In the signal case, the jet mass resolution is calculated after matching the jet to the W boson direction at the particle level. In the background case, a ΔR matching is done between the particle-level jet and the jet after full detector reconstruction where $\Delta R < 0.3$. In each case, the *groomed* particle-level jet is compared against the corresponding matched detector-level reconstructed jet.

We first evaluate the performance of groomers based on PF input collection with and without **C**harged **H**adron **S**ubtraction (CHS) to understand the effect. The CHS is a standard method in CMS to remove the charged hadrons coming from pileup vertices. We show the average jet mass distributions for PF and PF+CHS background QCD jets in Fig. 8.4. The left column shows the average jet mass as a function of pileup for jets using PF inputs. The right column shows the same for jets using PF inputs including CHS. We see in general that the CHS jet masses to be more stable against pileup. The trimming algorithm is generally stable against pileup regardless of inputs while pruning shows the most pileup dependence, particularly when using PF inputs. The soft drop has a mild pileup dependence.

8 Advanced Pileup Mitigation Techniques

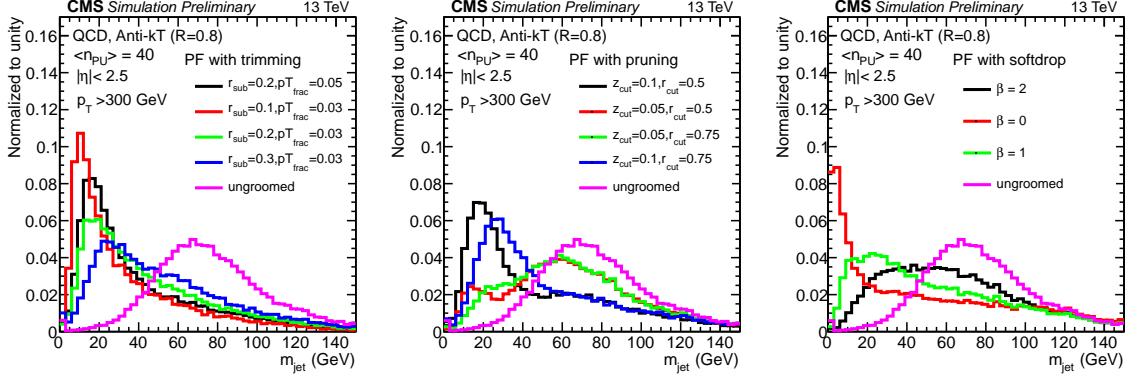


Figure 8.2: Jet mass distributions for PF QCD jets for different grooming parameters.

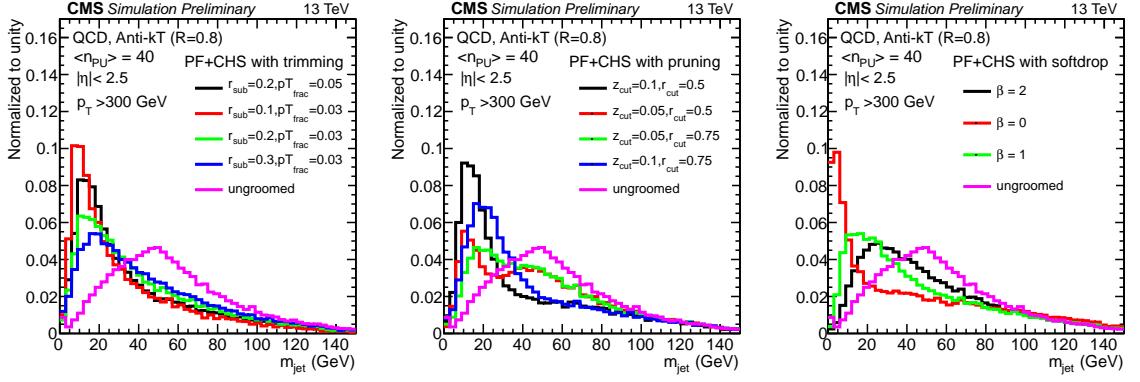


Figure 8.3: Jet mass distributions for PF+CHS QCD jets for different grooming parameters.

We also compare the resolution of jets after the application of various grooming algorithms by plotting the RMS values from the jet response templates. Fig. 8.5 shows that trimming has better mass resolution followed by soft-drop and pruning.

8 Advanced Pileup Mitigation Techniques

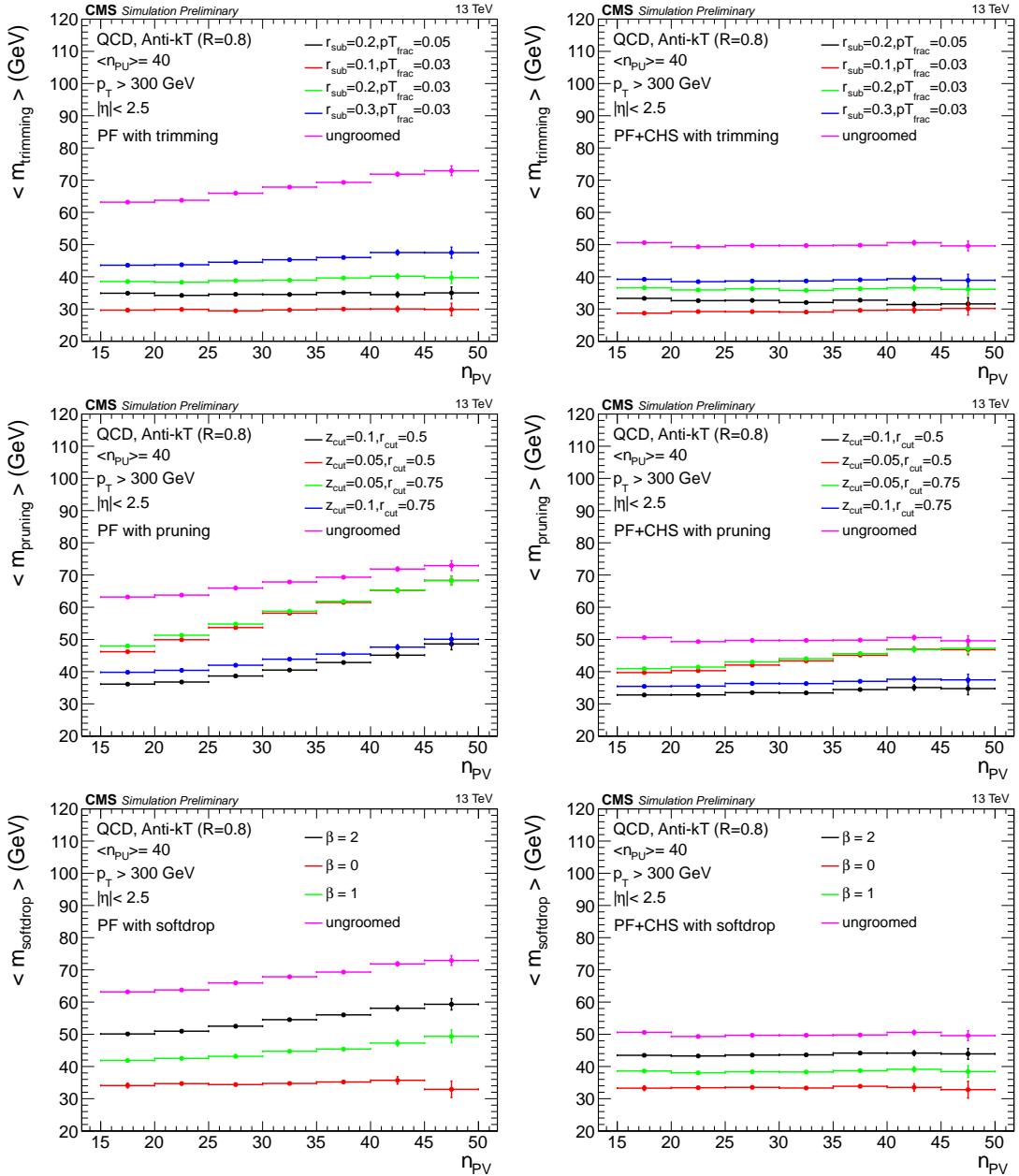


Figure 8.4: Pileup dependence of the average jet mass for PF jets (left) and PF+CHS jets (right) for various grooming algorithms and parameters.

8 Advanced Pileup Mitigation Techniques

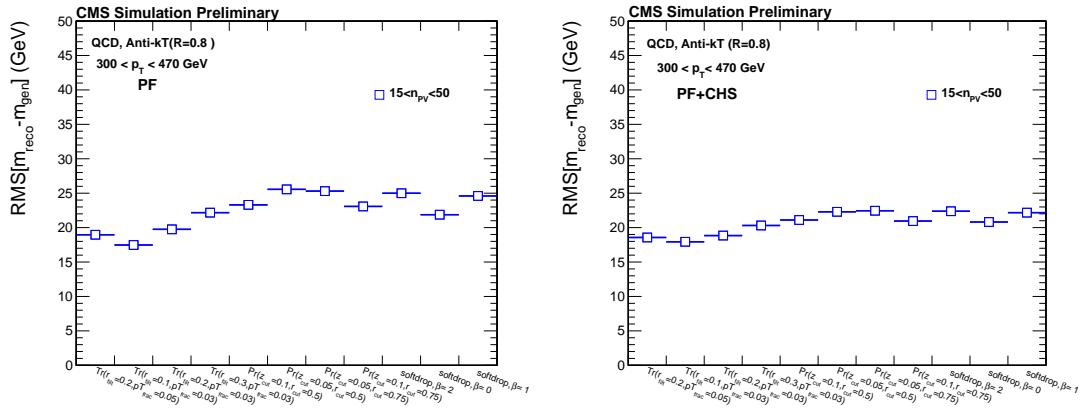


Figure 8.5: Comparison of jet mass resolution for PF and PF+CHS with different grooming algorithms and parameters.

9 Summary and Conclusions

In this thesis we have presented our studies on three different topics. The first two are on the search for physics beyond the standard model, especially for supersymmetry and leptoquarks. We did not find any evidence of new physics in these studies, however the limits we obtained are significantly more restrictive compared to the previous 8-TeV results. To drive the point home, we compare in Fig. 9.1 the upper limits on the cross section just from one channel that has four top quarks in the final state shows the limits). We see that the upper limit for a zero LSP mass hypothesis has been extended from around 1.2 to 1.6 TeV with a small amount of 13 TeV data. Although we have not shown comparison plots from four light-quark or four b-quark final states (all these results for 8 and 13 TeV could be found at the CMS public page [67]) the message is similar i.e., our results provide more stringent limit on the new physics parameter space. Moreover in the 2016 analysis with 12.9 fb^{-1} of data, we extended the 2.3 fb^{-1} limits.

Similarly, in the leptoquark analysis we have extended the previous limits. However, these are not improved as in SUSY because the increase in SUSY production cross section is much higher going from 8 to 13 TeV compared to the LQ models. In Fig. 9.2, we present a comparison of the 8 TeV limit vis-a-vis the 13 TeV one. Further, a small excess seen for an LQ mass of 650 GeV in the 8 TeV analysis [56] was not confirmed in

9 Summary and Conclusions

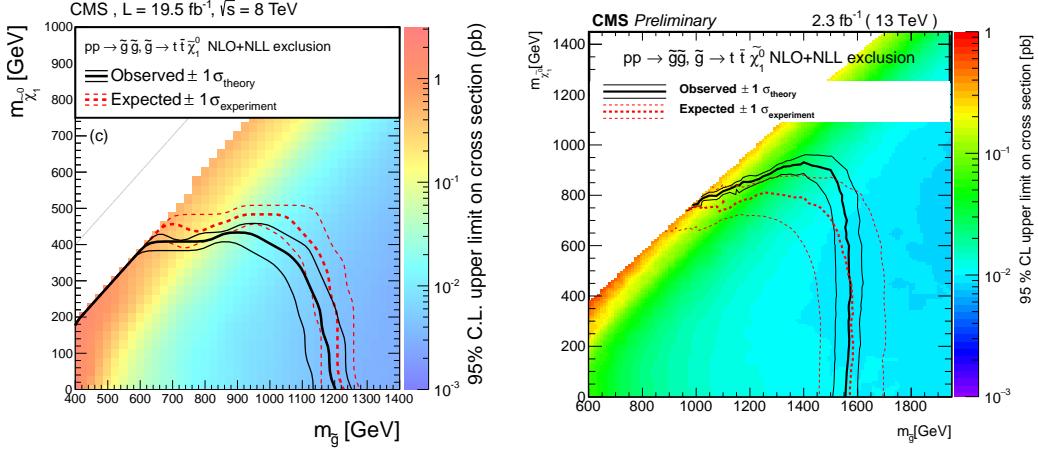


Figure 9.1: The 95% CL upper limits on the gluino pair production cross sections for four top quark final state. The left plot is taken from Ref. [68]. The right plot is produced by us with 2.3 fb^{-1} of data.

our analysis indicating the prior signal could be a statistical fluctuation.

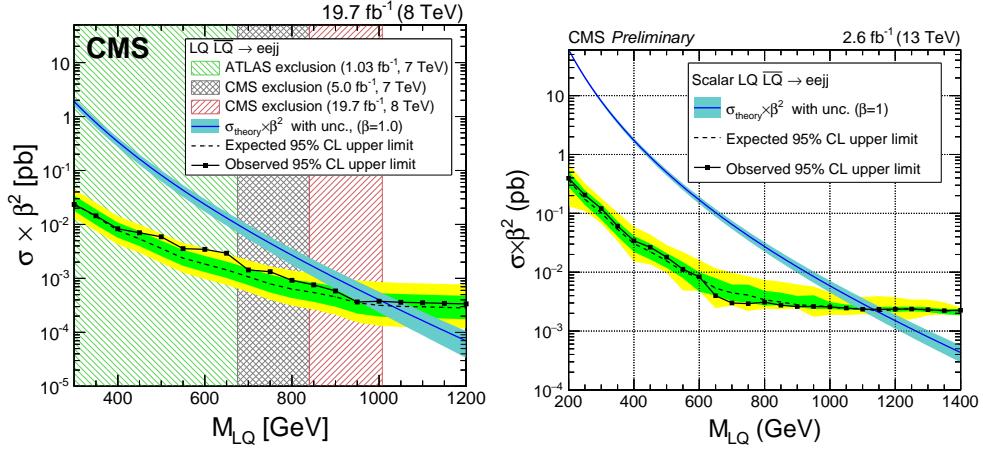


Figure 9.2: The 95% CL upper limits on the leptoquark production cross section times branching fraction. The left plot is taken from Ref. [56] while the right plot is our result based on 2.6 fb^{-1} of data.

At the end, we have presented a study on the various advanced pileup mitigation techniques. It turns out that grooming is very useful when jets are contaminated with pileup. Motivated by our study, many of the current CMS analyses are using grooming techniques for pileup mitigation.

9 Summary and Conclusions

Though our searches yielded null results for new physics, the results were useful in excluding the parameter space for SUSY and leptoquarks. As we understand the SM is incomplete, there must be new physics at a scale currently not known to us. We believe, with collection of more data with higher energies in the coming years could potentially reveal new physics.

On this positive note, we conclude the thesis.

9 Summary and Conclusions

References

- [1] H. Nilles, “Supersymmetry, supergravity and particle physics,” *Phys. Rep.* **110** (1984) 1.
- [2] H. Georgi and S. Glashow, “Unity of all elementary-particle forces,” *Phys. Rev. Lett.* **32** (1974) 438–441.
- [3] J. C. Pati and A. Salam, “Lepton Number as the Fourth Color,” *Phys. Rev. D* **10** (1974) 275.
- [4] B. Schrempp and F. Schremp, “Light Leptoquark,” *Phys. Lett. B* **153** (1985) 101.
- [5] S. Dimopoulos and L. Susskind, “Mass Without Scalar,” *Nucl. Phys. B* **155** (1979) 237.
- [6] S. Dimopoulos, “Technicolored signatures,” *Nucl. Phys. B* **168**, (1980) 69 .
- [7] E. Eichten and K. Lane.“Dynamical breaking of the weak interaction symmetries,” *Nucl. Phys. B* **90** (1980) 125 .
- [8] J. L. Hewett and T. G. Rizzo.“Low-energy phenomenology of superstring-inspired E₆ models,” *Phys. Rep.* **183** (1989) 193 .
- [9] M. J. Herrero, “The Standard Model”, <http://arxiv.org/abs/hep-ph/9812242>.
- [10] A. Pich, “The Standard model of electroweak interactions,” *Proceedings of the lectures given at the 2006 European School of High Energy Physics, Aronsborg, Sweden, June 18 - July 1, 2006* **1**, (2007) 1-49.

References

- [11] F. Halzen and A. D. Martin, “Quarks and Leptons: An Introductory Course in Modern Particle Physics” *1st edition*.
- [12] 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.
- [13] 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) 129.
- [14] Wikipedia “Standard Model” https://en.wikipedia.org/wiki/Standard_Model.
- [15] E. D. Bloom *et al.* “High-energy inelastic $e-p$ scattering at 6^0 and 10^0 ”, *Phys. Rev. Lett.* **23** (1969) 930–934.
- [16] M. Breidenbach *et al.* “Observed behavior of highly inelastic electron-proton scattering”, *Phys. Rev. Lett.* **23** (1969) 935–939.
- [17] W. Bartel *et al.* “Observation of planar three-jet events in e+e annihilation and evidence for gluon bremsstrahlung”, *Phys. Lett. B* **91** (1980) 1:142-147
- [18] A. Pich, Quantum chromodynamics. (hep-ph/9505231. FTUV-95-19. IFIC-95-19):51 p, May 1995.
- [19] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Non-Abelian Gauge Theories”, *Phys. Rev. Lett.* **30** (1973) 1343
- [20] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions?”, *Phys. Rev. Lett.* **30** (1973) 1346.
- [21] P. W. Higgs, “Broken symmetries and the mass of gauge vector mesons”, *Phys. Rev. Lett.* **13** (1964) 508.
- [22] F. Englert and R. Brout, “Broken symmetries and the mass of gauge vector mesons”, *Phys. Rev. Lett.* **13** (1964) 321.
- [23] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, *Phys. Rev.* **127** (1962) 965.

References

- [24] “The LEP Collaborations ALEPH, DELPHI, L3 and OPAL and the LEP Electroweak Working Group”, <http://arxiv.org/abs/hep-ex/0612034>.
- [25] “The ALEPH, DELPHI, L3, OPAL and SLD Collaborations, The LEP Electroweak Working Group and the SLD Electroweak and Heavy Flavour Group”, *Phys. Rept.* **427** (2006) 257.
- [26] F. Fiedler, “Top quark production and properties at the Tevatron”, http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-05-226-E.
- [27] (The ATLAS, CDF, CMS, D0 Collaborations), “First combination of Tevatron and LHC measurements of the top-quark mass”, <https://arxiv.org/abs/1403.4427>.
- [28] Planck Collaboration, “Planck 2013 results. XVI Cosmological parameters”, <https://arxiv.org/abs/1303.5076>.
- [29] A. Pomarol, “Beyond the Standard Model”, CERN Yellow Report CERN-2012-001, pp. 115-151, <http://arxiv.org/abs/1202.1391>.
- [30] S. P. Martin, “A Supersymmetry Primer”, *Adv. Ser. Direct. High Energy Physics* **18** (1998) 1
- [31] C. Csaki, “The Minimal supersymmetric standard model (MSSM)”, *Mod. Phys. Lett.* **A11** (1996) 599
- [32] P. Fayet, “Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and its Neutrino”, *Nucl. Phys.* **B90** (1975) 104–124.
- [33] T. Appelquist, H. C. Cheng and B. A. Dobrescu, “Bounds on Universal Extra Dimensions”, *Phys. Rev.* **D64** (2001) 035002
- [34] G. Farrar and P. Fayet, “Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry,” *Phys. Lett. B* **76** (1978) 575.
- [35] C. Borschensky *et al.*, “Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV”, *Eur. Phys. J.* **C74(12)** (2014) 3174
- [36] LHC SUSY Cross Section Working Group , “SUSY Cross Sections”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections>.

References

- [37] P. S. Alwall and N. Toro, “Simplified models for a first characterization of new physics at the LHC,” *Phys. Rev. D* **79** (2009) 075020.
- [38] K. Olive *et al.*, “Review of Particle Physics”, *Chin. Phys. C* **38** (2014) 090001.
- [39] J. L. Evans and N. Nagata, “Signatures of Leptoquarks at the LHC and Right-handed Neutrinos”, <https://arxiv.org/pdf/1505.00513.pdf>.
- [40] R. Adolphi *et al.*, CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **0803:S08004** (2008).
- [41] The Large Hadron Collider, “The Large Hadron Collider”, <http://home.cern/topics/large-hadron-collider>.
- [42] CMS collaboration, “Physics Technical Design Report”, Volume I : Detector Performance and software, CMS/LHCC 2006-001, CMS TDR 8.1.<https://cdsweb.cern.ch/record/922757>
- [43] CMS collaboration, “Physics Technical Design Report”, Volume II : Physics performance, CMS/LHCC 2006-021, CMS TDR 8.2.<https://cdsweb.cern.ch/record/922757>
- [44] CMS Collaboration, “Recent Results of the CMS Experiment”, <https://inspirehep.net/record/1193345/files/cms.png>.
- [45] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, <https://arxiv.org/abs/1206.4071>.
- [46] E. Chabanat and N. Estre , “Deterministic annealing for vertex finding at CMS”, <http://dx.doi.org/10.5170/CERN-2005-002.287>.
- [47] CMS Collaboration ,“Electromagnetic Physics Objects Commisioning with first LHC data”, <http://dx.doi.org/10.5170/CERN-2005-002.287>.
- [48] CMS Collaboration, S.Chatrchyan *et al* “Performance of CMS Muon reconstruction in pp collision events at $\sqrt{s}= 7$ TeV ”, *JINST* **7** (2012) P10002, <http://arxiv.org/abs/1206.4071>.
- [49] CMS Collaboration, “Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV”, CMS-PAS-PFT-2010-002 (2010), <https://cds.cern.ch/record/127934>.

References

- [50] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, JHEP 04 (2008) 063, <http://arxiv.org/abs/0802.1189>.
- [51] The ATLAS and CMS Collaborations and The LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011,” Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011, <https://cds.cern.ch/record/1379837>.
- [52] CMS Collaboration, “Search for supersymmetry in the multijet and missing transverse momentum channel in pp collisions at 13 TeV,” Tech. Rep. CMS-PAS-SUS-15-002, CERN, Geneva, 2015. <https://cds.cern.ch/record/2114817>.
- [53] CMS Collaboration, “Muon Physics Object Group,” <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonPOG>
- [54] CMS Collaboration, “E-gamma Physics Object Group,” <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>
- [55] CMS Collaboration, “Search for supersymmetry in events with jets and missing transverse momentum in proton-proton collisions at 13 TeV,” Tech. Rep. CMS-PAS-SUS-16-014, CERN, Geneva, 2016. <https://cds.cern.ch/record/2205158>.
- [56] CMS Collaboration, “Search for Pair-production of First Generation Scalar Leptoquarks in pp Collisions at $\sqrt{s} = 8$ TeV,” Tech. Rep. CMS-PAS-EXO-12-041, CERN, Geneva, 2014. <https://cds.cern.ch/record/1742179>.
- [57] G. Aad *et al.*, “Searches for scalar leptoquarks in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”, *Eur. Phys. J.*, **C76(1):5** (2016).
- [58] G. Punzi, “Sensitivity of searches for new signals and its optimization”, <https://arxiv.org/abs/physics/0308063>
- [59] CMS Collaboration, “Study of Pileup Removal Algorithms for Jets ,” Tech. Rep. CMS-PAS-JME-14-001, CERN, Geneva, 2016. <https://cds.cern.ch/record/1751454>.
- [60] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches,” *Phys. Rev.* **D81** (2010) 094023.

References

- [61] D. Krohn, J. Thaler, and L.-T. Wang, “Jet Trimming,” *JHEP* **02** (2010) 084.
- [62] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, “Towards an understanding of jet substructure,” *JHEP* **09** (2013) 029.
- [63] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop,” *JHEP* **05** (2014) 146.
- [64] M. Peruzzi, “Jet and photon physics”, *Proceedings, 2nd Conference on Large Hadron Collider Physics Conference (LHCP 2014): New York, USA, June 2-7, 2014* <https://arxiv.org/abs/1409.8539>.
- [65] S. Cataniand, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, “Longitudinally invariant k_t clustering algorithms for hadron hadron collisions,” *Nucl. Phys.* **B 406:187** (1993).
- [66] M. Cacciari, G. P. Salam, and G. Soyez, “On the use of charged-track information to subtract neutral pileup”, <https://arxiv.org/abs/1404.7353>
- [67] CMS Collaboration , “CMS SUSY public results”, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSYSMSSummaryPlots8TeV>.
- [68] CMS Collaboration, Chatrchyan *et. al*, “Search for new physics in the multijet and missing transverse momentum final state in proton-proton collisions at $\sqrt{s} = 8$ TeV”, *J. High Energy Phys.*, **06** (2014) 055
- [69] G. G. S. Dimopoulos, “Naturalness constraints in supersymmetric theories with nonuniversal soft terms,” *Phys. Lett. B* **357** (1995) 573.
- [70] W. Buchmuller and D. Wyler, “Constraints on SU(5) Type Leptoquarks,” *Phys. Lett. B* **177** (1986) 377–382.
- [71] CMS Collaboration, S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,” *JINST* **3** (2008) S08004.
- [72] CMS Collaboration, F. Beaudette, “The CMS Particle Flow Algorithm,” in *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pp. 295–304. 2013. arXiv:1401.8155 [hep-ex]. <https://inspirehep.net/record/1279774/files/arXiv:1401.8155.pdf>.

References

- [73] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm,” *JHEP* **04** (2008) 063, [arXiv:0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189).
- [74] CMS Collaboration, “Search for pair-production of first generation scalar lepto-quarks in pp collisions at $\sqrt{s} = 13$ TeV with 2.6 fb^{-1} ,” Tech. Rep. CMS-PAS-EXO-16-043, CERN, Geneva, 2016. <https://cds.cern.ch/record/2205285>.
- [75] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, “Jet substructure as a new Higgs search channel at the LHC,” *Phys. Rev. Lett.* **100** (2008) 242001.
- [76] Y. L. Dokshitzer, G.D. Leder, S. Moretti, and B.R. Webber, “Better jet clustering algorithms”, *JHEP*, **08** (1997) 001.

References