

SEARCH FOR SUPERSYMMETRY, LEPTOQUARKS AND  
STUDY OF ADVANCED PILEUP MITIGATION  
 TECHNIQUES WITH THE ~~COMPACT MUON SOLENOID~~  
EXPERIMENT

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  
Colaba, ~~MUMBAI~~ - 400005  
~~FEBRUARY~~, 2017

© 2017 by Bibhuprasad Mahakud

All rights reserved.

## DEDICATION

*To all the physics lovers ~~and~~ my parents and wife.*

 **CERTIFICATE**

This is to certify that the thesis entitled **SEARCH FOR SUPERSEMIMETRY, LEPTO-QUARKS AND STUDY OF ADVANCED PILEUP MITIGATION TECHNIQUES WITH COMPACT MUON SOLENOID**, being submitted by **Bibhuprasad Mahakud** for the award of the degree of **Doctor of Philosophy** in Physics of Tata Institute of Fundamental Research, Mumbai is a record of bonafide research work carried out by him. He has carried out this investigation for last five years on the subject matter of the thesis under our supervision and we consider it worthy of consideration for the award of the degree of Doctor of Philosophy of the University. The works embodied in this thesis have not been submitted to any other university or institution for any degree or diploma.

---

**Prof. Gagan Bihari Mohanty**

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

## **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<sup>of</sup> collaborative research and discussions.

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

---

**Mr. Bibhuprasad Mahakud**

Date:

## Acknowledgements

Firstly, I ~~would like to~~ 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, 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 ~~would like to~~ thank the rest of my thesis committee: Prof. Tariq Aziz, Prof. Sudeshna Banerjee, and Prof. Shashikant Dugad, Prof. Indranil Muzumdar and Dr. Rishi Sharma for their insightful comments and encouragement, ~~but also~~ for the hard question which incited me to widen my research from various perspectives.

My sincere thanks also ~~goes~~ to Dr. Andrew Whitbeck, Dr. Jim Hirsch and Dr. Nhan Viet Tran who provided me an opportunity to join their team at Fermilab, and ~~who~~ gave access to the laboratory and research facilities. ~~Without their precious support it would not have been~~ ~~be possible to conduct this research.~~ I am also thankful for useful discussions and guidance that they provided which have been very crucial to ~~me~~. I ~~would like to~~ thank Bibhuti Paria, my senior, for starting ~~new~~ collaboration with Fermilab, USA ~~so that I got~~ an opportunity to work in ~~a~~ Big group.

I ~~would like to~~ thank my friends Jacky, Varghese, Bajarang, Muzamil, Soureek, Nairit, Gouranga, Saurabh, Saranya, Soham, Suman, Pallavi, Arvind, Deepanwita, Sandhya, Adiba, Debashish for useful discussions that ~~has~~ helped my understanding through ~~out~~ my PhD.

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

Last but not the least, I ~~would like to~~ thank my family: my parents, my wife and my brother

for supporting me morally throughout writing this thesis and my life in general.

~~Finally, I am thankful to TIFR for providing me such an excellent lab to learn physics.~~



# Contents

<b>Title page</b>	<b>i</b>
<b>Certificate by the Supervisor</b>	<b>iii</b>
<b>Declaration</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>Contents</b>	<b>vii</b>
<b>Synopsis</b>	<b>1</b>
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 $2.3 \text{ fb}^{-1}$ of 2015 Data . . . . .	12
0.4.2    Study on $12.9 \text{ fb}^{-1}$ of 2016 Data . . . . .	20
0.5    Search for Pair Production of First Generation Leptoquarks . . . . .	23
0.5.1    Background Estimation . . . . .	24

## *Contents*

0.5.2	Results	24
0.6	Advanced Pileup Mitigation Techniques	25
0.6.1	Jet Grooming	25
0.6.2	Results	26
0.7	Summary	28
<b>List of Symbols</b>		<b>29</b>
<b>List of Tables</b>		<b>30</b>
<b>List of Figures</b>		<b>32</b>
<b>Chapter 1 Introduction</b>		<b>1</b>
<b>Chapter 2 The Standard Model of Particle Physics</b>		<b>4</b>
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 Standard Model	14
<b>Chapter 3 Beyond the Standard Model</b>		<b>16</b>
3.1	Supersymmetry	16
3.2	Lepto quark models	19
<b>Chapter 4 The Compact Muon Solenoid</b>		<b>20</b>
4.1	CMS detector	20
4.1.1	Tracker	21

## Contents

4.1.2	Electromagnetic Calorimeter . . . . .	23
4.1.3	Hadron Calorimeter . . . . .	26
4.1.4	Muon System . . . . .	29
<b>Chapter 5</b>	<b>Statistical procedures for Search</b>	<b>32</b>
5.1	Observed limits . . . . .	33
5.2	Expected limits limits . . . . .	35
5.3	Significance . . . . .	36
<b>Chapter 6</b>	<b>Search for Supersymmetry with 13 TeV pp collision data</b>	<b>37</b>
6.0.1	Object reconstruction . . . . .	38
6.1	Study with $2.3 \text{ fb}^{-1}$ of 2015 data . . . . .	38
6.1.1	Data samples and Trigger . . . . .	39
6.1.2	Event selection criterias and exclusive search intervals . . . . .	41
6.1.3	Backgrounds . . . . .	45
6.1.4	Results and Statistical interpretation . . . . .	53
6.2	Study with $12.9 \text{ fb}^{-1}$ of 2016 data . . . . .	55
6.2.1	Search binning . . . . .	56
6.2.2	$Z(\rightarrow \nu\bar{\nu}) + jets$ background estimation . . . . .	56
6.2.3	Results and Statistical Interpretation . . . . .	56
<b>Chapter 7</b>	<b>Search for Pair Production of First Generation Leptoquarks</b>	<b>57</b>
7.1	Event Selection . . . . .	58
7.2	SM Backgrounds . . . . .	60
7.3	Results and statistical interpretation . . . . .	60
<b>Chapter 8</b>	<b>Advanced Pilup Mitigation Techniques</b>	<b>63</b>
8.0.1	Simulation-only results at high pileup . . . . .	65

*Contents*

<b>Chapter 9 Summary and Conclusion</b>	<b>69</b>
<b>References</b>	<b>70</b>

*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  $2.3 \text{ fb}^{-1}$ , 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  $2.6 \text{ fb}^{-1}$  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, [PLB 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  $2.6 \text{ fb}^{-1}$ , [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

- Search for supersymmetry in multijet final states in proton-proton collisions at 13 TeV, [CMS-AN-16-188](#).

*Contents*

- **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.**

## *Contents*

### **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 [2] 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 [3], jets initiated by top and bottom quarks.

In this search, we consider SUSY scenarios in the context of simplified models [4] 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. In the first one, shown in Fig. 1 (left), the gluino pair production is followed by the decay of each gluino

## Contents

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 gluino decay to a light-flavored quark and off-shell squark up, down, strange, and charm with equal probability, for each gluino separately.

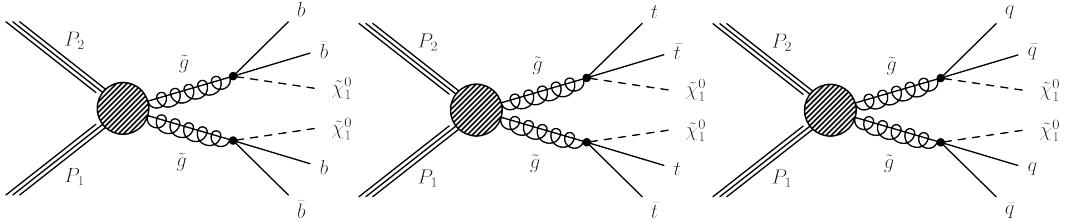


Figure 0.1: Event diagrams for the main new-physics scenarios considered in this study.

The SM hints that its lepton and quark sectors could be related by a fundamental symmetry. Many beyond the SM theories, include such a symmetry [5; 6; 7; 8], which gives rise to new class of bosons called leptoquarks (LQ). Leptoquarks 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 such as lepton number violation and flavor-changing neutral currents suggest that leptoquarks 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-quark-lepton vertex, the scalar LQ pair-production has a negligible dependence on the Yukawa LQ-quark-lepton coupling, usually denoted as  $\lambda$ . LQ searches are essentially independent of  $\lambda$ . In our search we consider pair produced leptoquarks with each of them decaying to an electron and a quark. Therefore we look for signature in two electron

## *Contents*

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 ~~is~~ become more and more demanding. The most common observables for jets at the LHC are primarily the jet  $p_T$ ,  $\eta$ , and  $\phi$ . 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 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. [14]. The

## *Contents*

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 [15], 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<sup>1</sup> 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<sup>1</sup>. 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 [16] 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 individual jets  $p_T$  from pileup interactions are subtracted, and corrections are applied as a function of jet  $p_T$  and  $\eta$  to account for residual effects of nonuniform detector response.

## *Contents*

## 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  $H_T^{\text{miss}}$ , 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  $H_T^{\text{miss}}$  as well as large jet multiplicity. The data are examined in bins of  $N_{\text{jet}}$ ,  $N_{\text{b-jet}}$ ,  $H_T$ , and  $H_T^{\text{miss}}$ , where  $N_{\text{jet}}$  is the number of jets and  $N_{\text{b-jet}}$  the number of tagged bottom quark jets (b jets). ~~The search is performed in exclusive bins of these four observables.~~

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 ~~latter~~ 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  $H_T^{\text{miss}}$  can arise if the W boson decays leptonically, producing a neutrino and an undetected charged lepton, while Z+jets events can exhibit significant  $H_T^{\text{miss}}$  if the Z boson decays to two neutrinos. For QCD multijet events, significant  $H_T^{\text{miss}}$  ~~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 0.3.

## Contents

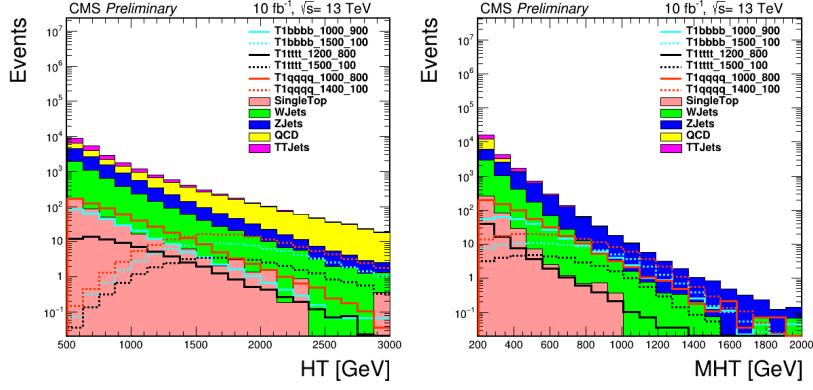


Figure 0.2: Signal vs. stacked backgrounds in  $H_T$ (-HT, left) and  $H_T^{\text{miss}}$ (-MHT, right)

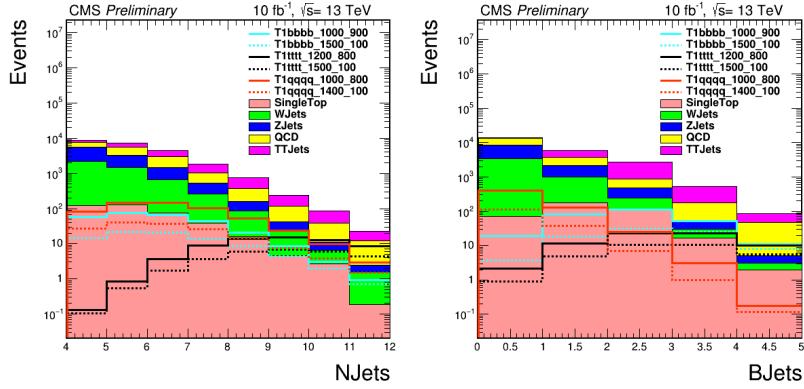


Figure 0.3: Signal vs. stacked backgrounds in  $N_{\text{jets}}$ (NJets, left) and  $N_{b\text{-jet}}$ (MHT, right)

### 0.4.1 Study on 2.3 fb<sup>-1</sup> 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<sup>-1</sup> data we search for gluino pair production [in terms of](#) the above four variables. To get maximum signal and minimum background we select events with [following](#) criteria defined below.

#### 0.4.1.1 Event Selection and Search Region

The following requirements define the signal event candidates:

## Contents

- $N_{\text{jet}} \geq 4$ , 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  $H_T^{\text{miss}}$  is the magnitude of  $\vec{H}_T^{\text{miss}}$ , the negative of the vector  $p_T$  sum of jets with  $|\eta| < 5$ ; the  $\eta$  range is extended in this case so that  $\vec{H}_T^{\text{miss}}$  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  $\vec{p}_T^{\text{miss}}$  and isolated-track  $p_T$  vector, with  $\vec{H}_T^{\text{miss}}$  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  $H_T^{\text{miss}}$  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**  $N_{\text{jet}}$  bins: 4, 6, 7, 8 and  $\geq 9$ ;
- **4**  $N_{\text{b-jet}}$  bins: 0, 1, 2 and  $\geq 3$ ;
- **3**  $H_T$  bins: 500, 800, 1200 and  $> 1200 \text{ GeV}$ ;
- **3**  $H_T^{\text{miss}}$  bins : 200, 500, 750 and  $> 750 \text{ GeV}$ .

## *Contents*

### **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 background 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  $Z(v\bar{v})$ +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  $\gamma$ +jets sample to determine the yields in the 18 bins corresponding to  $N_{b-jet}=0$ . 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  $\gamma$ +jets, and the normalization corrected if necessary. The extrapolation to bins with  $N_{b-jet} > 0$  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.

## Contents

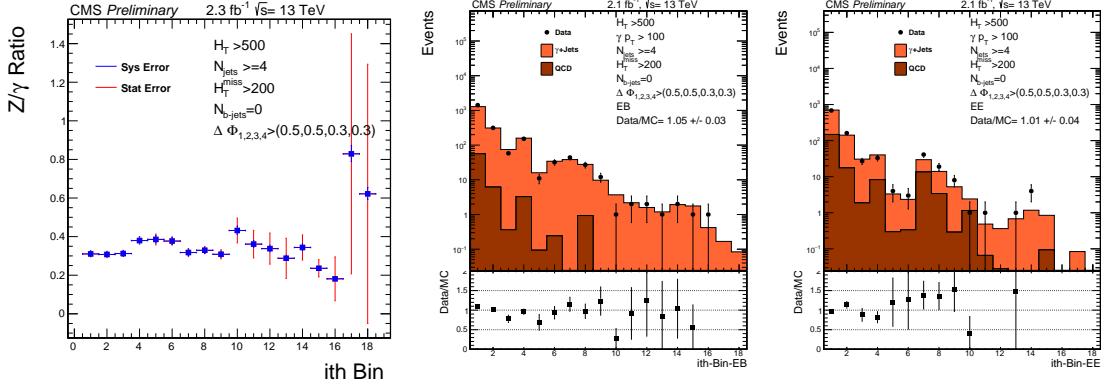


Figure 0.4: Physics ratio  $R_{Z(\nu\bar{\nu})/\gamma}$  (left), numbers of observed events in the photon control samples in barrel(EB, middle) and endcap(EE, right) compared to simulation.

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

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

## *Contents*

**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:  $H_T$ ,  $H_T^{\text{miss}}$ ,  $N_{\text{jet}}$ , 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  $H_T^{\text{miss}}$  in QCD multijet events is almost always due to a mismeasured jet in the event, thus the  $H_T^{\text{miss}}$  direction is usually close to the jet. The  $\Delta\phi$  variable is the minimum  $\phi$  difference between  $H_T^{\text{miss}}$  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/low ratio  $\kappa^{QCD}$  for the QCD component. The  $\Delta\phi$  distribution shows that the high/low ratio has some dependence on the search variables  $H_T$ ,  $H_T^{\text{miss}}$ , and  $N_{\text{jet}}$ . We model this dependence by assuming that it factorizes. That is, we assume the  $H_T$  dependence does not depend on  $H_T^{\text{miss}}$  or  $N_{\text{jet}}$  and similarly for  $H_T^{\text{miss}}$  and  $N_{\text{jet}}$ .

## Contents

**Hadronic tau:** To evaluate the  $t\bar{t}$ , single-top and  $W+jets$  backgrounds that arises when a  $W$  boson decays to a neutrino and a hadronically decaying  $\tau$  lepton ( $\tau^h$ ), we employ a tau-template method. In this approach, the  $\tau^h$  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  $\tau^h+jets$  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  $\tau^h$  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  $\tau^h$  jet response. Global hadronic variables such as  $N_{jet}$ ,  $H_T$ , and  $H_T^{\text{miss}}$  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 results from the background estimations include data control region statistics, Purity of control region , CR trigger efficiency etc. are discussed in the respective sections Ref. [18]. The signal systematics are discussed briefly below.

- **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  $\eta$ . 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  $H_T^{\text{miss}}$ .

## *Contents*

- **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 ~~actual number of interactions~~ and ~~the~~ observed number of interactions in the data. The minimum bias cross section in the 13 TeV is ~~currently~~ 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  $\eta$  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.

### **0.4.1.4 Results**

The data in the signal regions are found to be in ~~generally~~ good agreement with the predicted backgrounds. Therefor 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<sup>CL</sup> 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

## Contents

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  $\mu$  to vary and  $L_\mu$  is the maximum likelihood for a fixed signal strength . The details of the statistical procedure can be found in Ref. [17]. For an explanation of the treatment of uncertainties, we refer to Ref. [18]. 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.

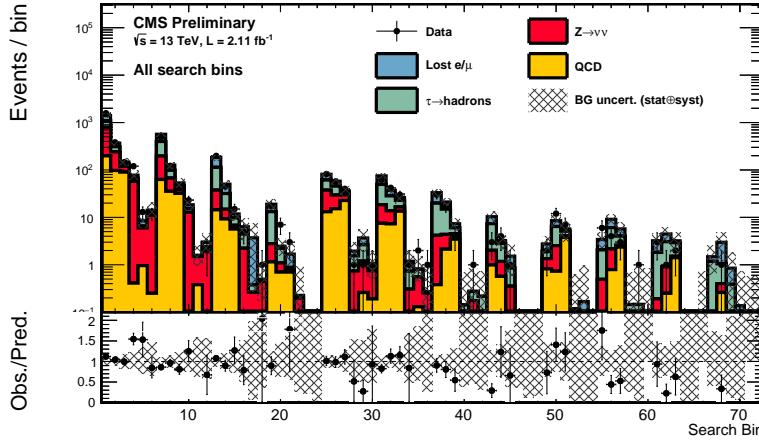


Figure 0.5: Data vs. the SM background before fit

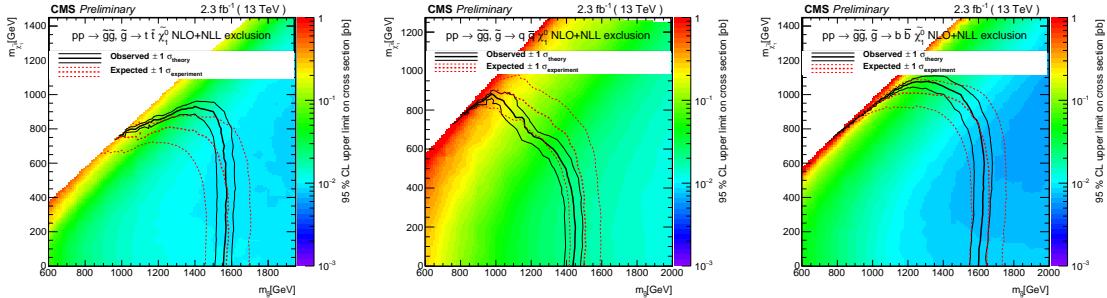


Figure 0.6: The 95% CL upper limits on the production cross sections for four top quark(left),four light quark(middle) and four b-quark(right) in the final state.



## *Contents*

### **0.4.2 Study on $12.9 \text{ fb}^{-1}$ of 2016 Data**

In 2016 we lowered the  $H_T$  and  $N_{\text{jet}}$  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 from the 2015 analysis are given below.

#### **0.4.2.1 Search Region and Event Selection**

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

- $N_{\text{jet}} \geq 3$ , 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**  $N_{\text{jet}}$  bins: 3-4, 5-6, 78,  $\geq 9$ ;
- **4**  $N_{\text{b-jet}}$  bins: 0, 1, 2,  $\geq 3$ ;
- **10** bins in  $H_T$  and  $H_T^{\text{miss}}$ : defined below

## Contents

Bin	$H_T$ range [GeV]	$H_T^{\text{miss}}$ range [GeV]
1:	300-500	300-350
2:	500-1000	300, <del>350</del>
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

### 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. [19]

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

- 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  $\sigma_{inj\eta}$  as we find it to have a better performance.
- Trigger for the photon control region is also changed.

## Contents

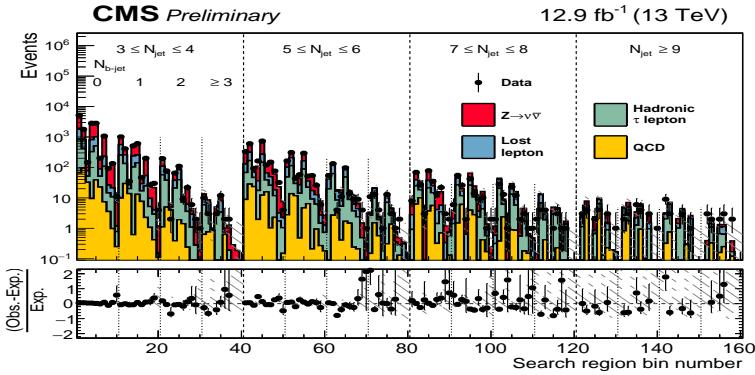


Figure 0.7: Data vs. the SM background before fit

Note that due to the change in  $H_T^{\text{miss}}$  and  $N_{\text{jet}}$  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 when compared with 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 a lot of improvements in terms of expected limits when compared to the previous results with  $2.3 \text{ fb}^{-1}$  of data (Fig. 6.9). But the mismatch between the observed and expected curves owing to the deficit of events in some of the bins.

## Contents

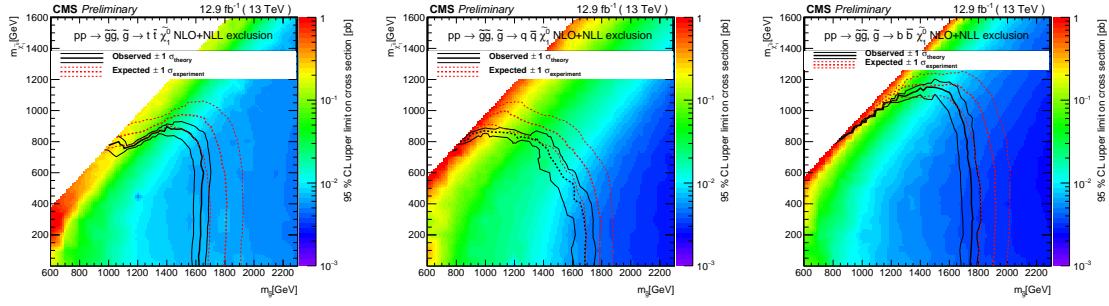


Figure 0.8: The 95% CL upper limits on the production cross sections for four top quark(left),four light quark(middle) and four b-quark(right) in the final state

## 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
- $M_{\ell,j}^{\min}$  is ~~defined as~~ the smaller lepton-jet invariant mass for the assignment of jets and leptons to leptoquarks that minimizes the  $LQ - \overline{LQ}$  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. [20].

The data vs. backgrounds plots for two variables are shown in ~~the following figure~~.

## Contents

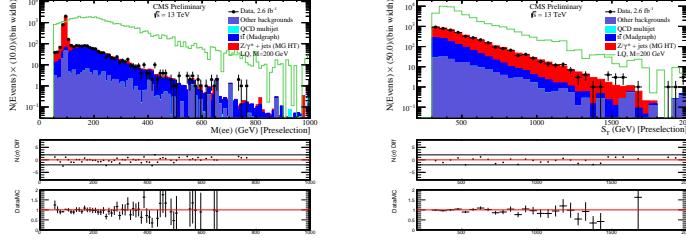


Figure 0.9: Data vs. background plots for  $M_{ee}$  (left) and  $S_T$  (right).



### 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 the backgrounds are determined in this analysis.

- The  $Z+jets$  and  $t\bar{t}$  background shapes are taken from MC simulation, and normalized to data using the eejj preselection. More details are given in Ref. [20].
- Single top,  $W+jets$ , diboson, and  $\gamma+jets$  backgrounds are derived completely from MC. The samples are scaled to the cross sections.
- QCD background is determined using a data-driven fake rate method, as described in Ref. [20].

### 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 [21] 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

## Contents

this study.

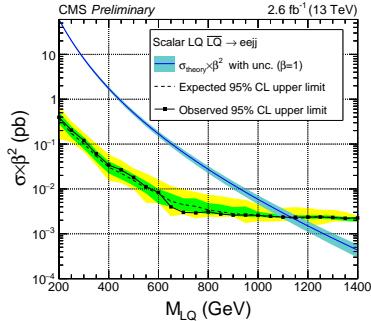


Figure 0.10: The 95% CL upper limits on the production cross sections as a function of lepto-quark mass.

## 0.6 Advanced Pileup Mitigation Techniques

### 0.6.1 Jet Grooming

Grooming, introduced in Ref. [22], 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 ~~jet structure~~ 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 [23] reclusters the constituents of the jet through the Cambridge-Aachen (CA) algorithm [24] 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:

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

## Contents

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

**2.** Trimming [25] 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 [16] with a radius parameter  $r_{\text{sub}}$ , accepting only the subjets that have  $p_{T_{\text{sub}}} > p_{T_{\text{frac}}} \lambda_{\text{hard}}$ , where  $p_{T_{\text{frac}}}$  is a dimensionless cutoff parameter, and  $\lambda_{\text{hard}}$  is some hard QCD scale chosen to be equal the  $p_T$  of the original jet.

**3.** Soft-drop [26] 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:

$$\frac{\min(p_{T_{j1}}, p_{T_{j2}})}{p_{T_{j1}} + p_{T_{j2}}} > z_{\text{cut}} \times \left( \frac{\Delta R_{12}}{R_0} \right)^\beta$$

where the algorithm parameters are  $z_{\text{cut}}$  and  $\beta$ . 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) [24].

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. [27].

### 0.6.2 Results

The groomed mass distributions of charged hadron subtracted PF (PFCHS) [15; 27] jets are shown in Fig. 0.11. One sees that after the application of grooming owing 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 one-dimensional mass distributions do not tell everything but it says because of the application of grooming the mass

## Contents

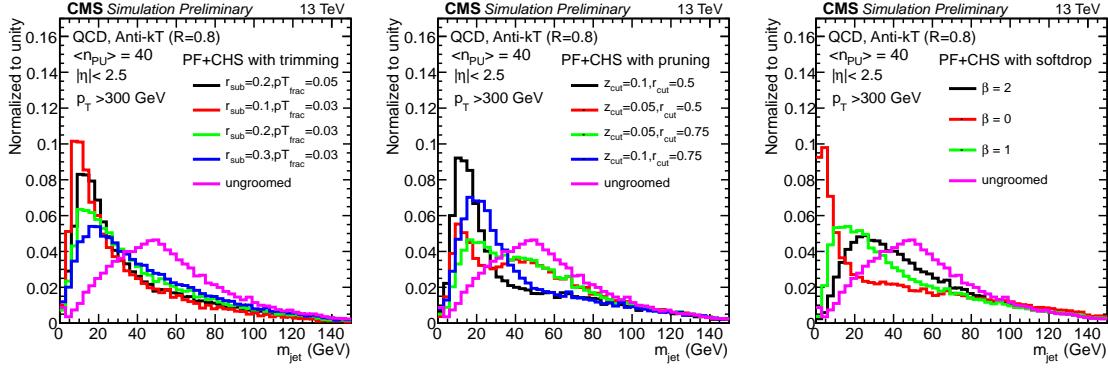


Figure 0.11: The mass distributions of PFCHS jets after application of trimming (left), pruning (pruning) and soft-drop (right) for different set of parameters. The ungroomed mass distribution is shown to compare the aggressiveness of grooming methods.

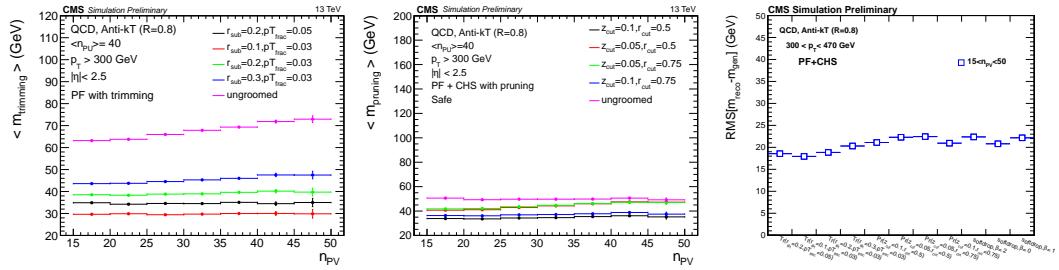


Figure 0.12: The plots shows variation of average jet mass vs. number of primary vertices for trimming (left), pruning (middle) and RMS vs. various grooming techniques(right).

distributions change aggressively.

Average jet mass and jet resolutions are taken as quality parameters for determining 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 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  $2.3 \text{ fb}^{-1}$  of 2015 and  $12.9 \text{ fb}^{-1}$  of 2016 data yield null results of new physics. Similarly no signatures of any signal are found in the first generation scalar lepto-quark search performed with  $2.6 \text{ fb}^{-1}$  of the 2015 data. Upper limits on the signal cross section are obtained in all cases significantly extending the previous 8-TeV limits.

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

## List of Symbols

Symbols	Definition
 $e$	Electron
$\mu$	Muon
$\tau$	Tau
$\nu_e$	Electron neutrino
$\nu_\mu$	Muon neutrino
$\nu_\tau$	Tau neutrino
$\gamma$	Photon
LHC	Large Hadron Collider
CMS	Compact Muon Solenoid
PV	Primary vertex
$\sqrt{s}$	Center-of-mass-energy
MET	Missing transverse energy
$m_H$	Mass of the Higgs
$m_J$	Mass of the jet
$\Gamma$	Decay width
$\text{fb}^{-1}$	Inverse femtobarn (Unit of Luminosity)
GeV	Gigaelectronvolt
TeV	Teraelectronvolt

## *Contents*

---

Symbols	Definition
$p_T$	Momentum in the transverse plane
$\eta$	Pseudorapidity
$\phi$	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/c<sup>2</sup> 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. . . . .	6
3.1	Field Content of MSSM in terms of superfileds(Ist coulmn), sparticles and particles(2nd and 3rd coulmn) and in 4th column the quantum numbers for various fields are shown. . . . .	18
6.1	MC FullSim samples for signal SMS model points. . . . .	39
6.2	SM $t\bar{t}$ MC samples used in the analysis. The cross sections are calculated to NNLO. . . . .	39
6.3	SM QCD MC samples used in the analysis. All cross sections are calculated to LO. . . . .	39
6.4	SM $Z \rightarrow \nu\nu + \text{jets}$ MC samples used in the analysis. The cross sections are calculated to NNLO. . . . .	40
6.5	SM $W \rightarrow l\nu + \text{jets}$ MC samples used in the analysis. The cross sections are calculated to NNLO. . . . .	40
6.6	SM diboson and other rare process MC samples used in the analysis. The cross sections are calculated to NNLO. . . . .	41

*List of Tables*

6.7	SM DY+jets MC samples used in the analysis. The cross sections are calculated to NNLO. . . . .	41
6.8	SM $\gamma$ +jets MC samples used in the analysis. The cross sections are calculated to LO. . . . .	42
6.9	Photon identification and isolation requirements for events with photons in the barrel and endcap. . . . .	48
6.10	Effective areas used in pileup correction as a function of photon pseudorapidity. . . . .	48
7.1	Optimized final selection thresholds for the $eejj$ analysis. . . . .	61
8.1	Summary of grooming parameters . . . . .	65

# 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 $H_T^{\text{miss}}$ ( MHT, right) . . . . .	12
0.3	Signal vs. stacked backgrounds in $N_{\text{jet}}$ (NJets, 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 . . . . .	19
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. . . . .	19
0.7	Data vs. the SM background before fit . . . . .	22
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. . . . .	25

## List of Figures

0.11 The mass distributions of PFCHS jets after application of trimming (left), pruning (pruning) and soft-drop (right) for different set of parameters. The ungroomed mass distribution is shown to compare the aggressiveness of grooming methods.	27
0.12 The plots shows variation of average jet mass vs. number of primary vertices for trimming (left), pruning (middle) and RMS vs. various grooming techniques(right).	27
 1.1 Looking deeper and deeper into matter, from atom to quark.	2
 2.1 Summary of the interactions between the SM constituents. The plot is taken from Ref. [33]	6
2.2 Masses and other useful properties of SM family. The plot is taken from Ref. [33]	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 $v^2 = -\mu^2/\lambda$ .	13
 3.1 SM and their SUSY partners.	19
 4.1 An inside view of CMS detector, showing various subdetectors that are placed around the beam pipe and form a series of cylindrical layers of the experiment.	21
4.2 Various detector components that contribute to the event reconstruction are (from left to right) tracker, electromagnetic calorimeter, superconducting solenoid and muon system. The paths of different particles passing through the detector such as photons, muons, electrons, neutral hadrons and charged hadrons are indicated by different solid or dashed color lines.	22
4.3 A schematic diagram of the CMS Tracker. The plot shows the quadrants of the tracking detector along the rz plane.	22
4.4 A schematic diagram of the CMS ECAL sub-detector showing its different components.	24

## List of Figures

4.5 Amount of material (in units of radiation lengths) upstream of the ECAL (left). Electron energy resolution as a function of pseudorapidity for electrons from $Z \rightarrow e^+e^-$ decays. In red, the energy resolution estimated on Monte Carlo events is shown, while in blue and gray the resolution for respectively promptlyreconstructed data and a later reconstruction of the same data sample upon the usage of the best calibrations available are shown (right). . . . .	26
4.6 A schematic diagram of the CMS HCAL showing its individual components. . . . .	28
4.7 Muon resistive plate chamber . . . . .	31
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) . . . . .	38
6.2 Plot shows signal vs. stacked backgrounds in $H_T$ ( HT, left) and $H_T^{\text{miss}}$ ( MHT, right) . . . . .	44
6.3 Plot signal vs. stacked backgrounds in $N_{\text{jet}}(NJets$ , left) and $N_{\text{b-jet}}(MHT$ , right) . . . . .	44
6.4 Search binning in 2D phase space of $H_T$ and $H_T^{\text{miss}}$ . . . . .	45
6.5 Numbers of observed events in the photon control samples in EB (left) and EE (right) compared to simulation. . . . .	50
6.6 Fit results for purity determination in EB (top) and EE (bottom) for events with $H_T^{\text{miss}}$ in the range 200-300, 300-500, and for $H_T^{\text{miss}} > 500$ , respectively. $\beta_\gamma$ is the fitted purity within the corresponding region. The red vertical line shows the $\sigma_{i\eta i\eta}$ requirement for the loose ID working point. The non-prompt PDF comes from the QCD simulation. . . . .	51
6.7 $\mathcal{R}_{Z(v\bar{v})/\gamma}$ for the 18 analysis bins used in the analysis. . . . .	52
6.8 The observed data vs the standard model background in the 72 exclusive intervals of signal region . . . . .	54

## *List of Figures*

6.9	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. . . . .	55
6.10	Search binning in 2D phase space of $H_T$ and $H_T^{\text{miss}}$ in 2016 analysis. . . . .	56
6.11	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. . . . .	56
7.1	Data vs. background plots for $M_{ee}$ (left) and $S_T$ (right) . . . . .	62
7.2	The 95% CL upper limits on the production cross sections as a function of leptoquark mass. . . . .	62
8.1	Jet mass distribution for PF QCD jets for different grooming parameters. The PF jets are safe subtracted. . . . .	66
8.2	Jet mass distribution for PF+CHS QCD jets for different grooming parameters. The PFCHS jets are safe subtracted. . . . .	66
8.3	Pileup dependence of the average jet mass for PF jets (left) and PFCHS jets (right) for several grooming algorithms and parameters. Both PF and PFCHS jets are safe subtracted. . . . .	67
8.4	Comparison of jet mass resolution for PF, CHS and Puppi with different grooming variables . . . . .	68

# 1 Introduction

Since we began doing science centuries ago, our effort has 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 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 formulation of theoretical frameworks to understand the physical phenomena has gone hand in hand culminating into land mark theories like "Standard Model of Particle Physics"(SM). Figure 1.1 shows the sketch of the temporal evolution of the fundamental constituents of matter. Today, 'Standard Model of Particle Physics' represents our best theoretical understanding about the interactions between the fundamental constituents of matter. Even though SM has been phenomenally successful for explaining a wide range of physical phenomena, it has several draw backs. Inconsistencies arises from a handful of laboratory and astronomical findings that are inconsistent with predictions. These findings include the discovery of neutrino masses, presence of dark matter in the universe. Moreover standard model suffers from 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 contributions to the square of the Higgs boson mass would make the mass huge, comparable to the scale at which new physics appears, unless there is an incredible fine-tuning cancellation between the

## 1 Introduction

quadratic radiative corrections and the bare mass.

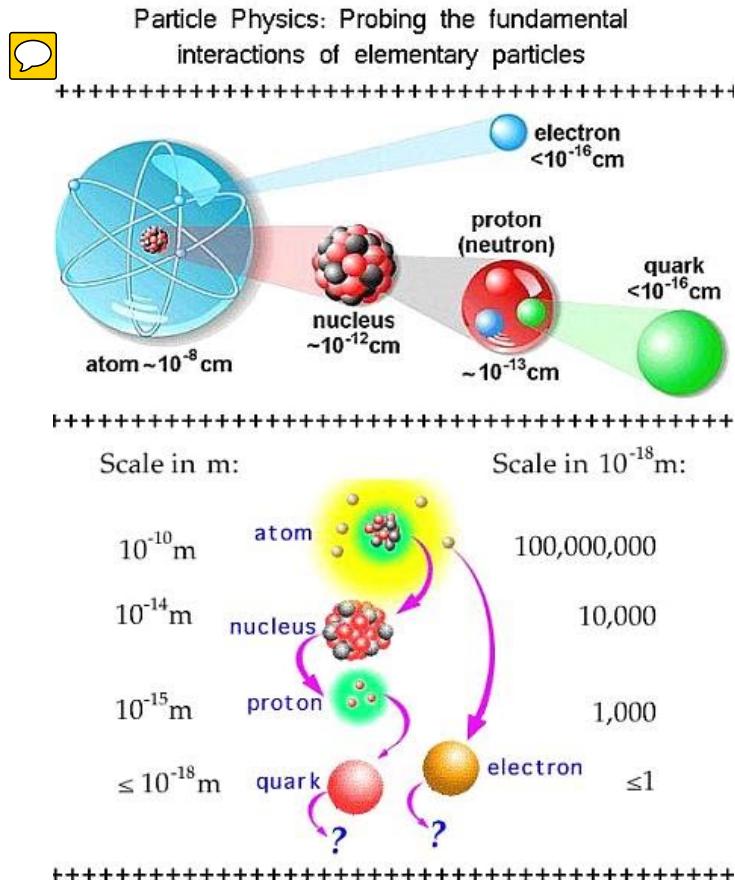


Figure 1.1: Looking deeper and deeper into matter, from atom to quark.

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

Moreover the structure of the standard model (SM) of particle physics suggests a fundamental relationship between quarks and leptons. In some theories beyond the SM, such as SU(5) grand-unification [9], Pati-Salam SU(4) [7], composite models [10], technicolor [5; 6; 11; 12], and superstring-inspired models [13], the existence of a new symmetry relates the

## 1 Introduction

quarks and leptons in a fundamental way. These models predict the existence of new bosons, called leptoquarks. The leptoquark (LQ) is colored, has fractional electric charge, can be either a scalar or vector particle, and couples to a lepton and a quark.

In the thesis, we present results of our searches for supersymmetry and leptoquarks with pp collision data recorded in compact muon solenoid (CMS) detector at LHC, CERN.

 In the searches the role of jets is very important as they constitute a major part of the events in the analysis. In the high luminosity environment i.e. is in runII, the measurement of the properties of jets become extremely challenging because large pileup. We also study some advanced pileup mitigation techniques presented in the last part of the thesis.

We start with a brief account of the SM and the theoretical background behind the Higgs mechanism in Chapter 2. We discuss beyond the standard model theories SUSY and LQ models in chapter3. The details of the CMS detector along with other experiments at LHC are discussed in Chapter 4.

For drawing inferences from the observed data, the understanding of correct statistical procedures become very important. We discuss briefly the LHC recommended statistical procedures in chapter5.

We devote chapter 6 for extensive report of the search for supersymmetry in all hadronic final state with 13 TeV data. Chapter7 gives a comparatively brief description of the search analysis we carried out for first generation scalar leptoquarks. And finally we present our results on the advanced pileup mitigation techniques in chapter8. We conclude with chapter9 by giving summary of the work and its signifying the importance of the work.

# 2 The Standard Model of Particle Physics

## 2.1 Introduction

The SM of particle physics [30; 28] is a theory that describes the elementary particles of the universe and their interactions. These particles can be classified into two groups: *bosons* and *fermions*. The bosons have an integer spin while the fermions have a half-integer spin. The bosons are responsible for various interactions while fermions make up all of the visible matter in the universe.

The fermions exist in two types: the *leptons*, which do not interact via the strong interaction, and the *quarks* that do interact via strong interactions. Also, for each fermion there exists a corresponding antifermion which carries same mass but opposite 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 ( $\nu_e$ ), muon neutrino ( $\nu_\mu$ ) and tau neutrino ( $\nu_\tau$ ) are electrically neutral. Except for the difference in mass, the three charged leptons are identical with respect to how they interact under the relevant fundamental forces. According to the way they interact via

## 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 ( $\nu_e$ )	muon ( $\mu$ ) muon neutrino ( $\nu_\mu$ )	tau ( $\tau$ ) tau neutrino ( $\nu_\tau$ )	-1 0

the weak interaction, the leptons are classified into three generations. Similarly, at present there are six known quarks, and they are also of two types. The up ( $u$ ), charm ( $c$ ), and top ( $t$ ) quarks have 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 shown in Table 2.1.

There are four fundamental interactions in nature namely strong, weak, electromagnetic and gravitation [29]. 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 ( $\gamma$ ), while the weak interaction that is responsible for the nuclear beta decay is mediated by the massive gauge bosons ( $W^+$ ,  $W^-$  and  $Z^0$ ). The strong interaction is mediated by massless gluons that bind quarks to form nucleons (protons and neutrons). Table 2.2 lists different forces and their mediators. There is no successful description of gravitation in terms of a quantum field theory to this date, so it is not implemented in the SM. As gravitation has a minuscule impact at sub-atomic scales, it can be neglected in high energy physics experiments.

There are six fundamental bosons that have been observed so far. The photon ( $\gamma$ ) is massless, electrically neutral, and only acts as the 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 do not interact via the strong force. The  $Z$  boson is massive, electrically neu-

## 2 The Standard Model of Particle Physics

Table 2.2: Different interactions and their mediators in nature.

Interactions	Mediators
Strong	gluon (g)
Electromagnetic	photon ( $\gamma$ )
Weak	$W^+$ , $W^-$ , and $Z^0$
Gravitational	graviton (?)

tral, 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. Finally, there is a massive ‘boson’ recently discovered by the CMS [32] and ATLAS [31] experiments at LHC having mass  $\sim 125$  GeV, which seems to be consistent with the SM Higgs boson. Figure. 2.1 shows the interactions between all SM particles along with the Higgs boson (H) that gives mass to the weak interaction mediators and other elementary particles. Figure [?] shows masses and other properties of all SM particles.

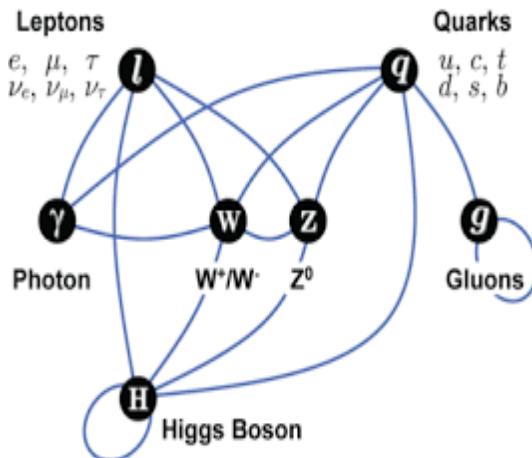


Figure 2.1: Summary of the interactions between the SM constituents. The plot is taken from Ref. [33]

---

In Table 2.2, we put “(?)” after “graviton” because its presence is not experimentally verified.

## 2 The Standard Model of Particle Physics

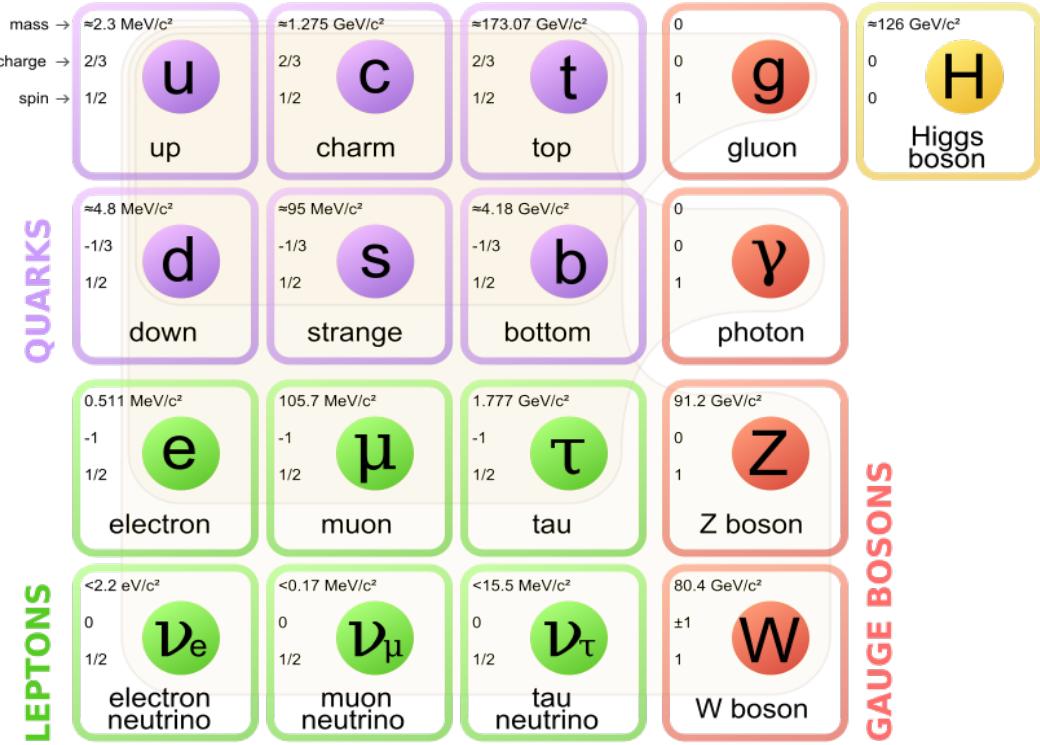


Figure 2.2: Masses and other useful properties of SM family. The plot is taken from Ref. [33]

## 2.2 Gauge Theory Formulation

The Standard Model (SM) of particle physics is a gauge theory, based on the symmetry group  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  [28] which describes strong, weak and electromagnetic interactions, via the exchange 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 interactions and photon which mediate the electromagnetic interactions. The three families of the 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 antiparticles.

## 2 The Standard Model of Particle Physics

The part of the lagrangian that describes the dynamics of strong force is described by the symmetry group  $SU(3)_C$ . Here C stands for color. We call this part as the QCD sector. The  $SU(3)_C$  invariant lagrangian of Quantum Chromodynamics is given by

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

Where  $f$  runs over the quark flavours.  $a$  denotes the summation over eight gluons.  $q_f$  denotes the quark field.  $G_a^{\mu\nu}$  is defined as  $G_a^{\mu\nu}(x) = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f^{abc} G_b^\mu G_c^\nu$  and  $G_a^\mu$  are the gauge fields.  $D^\mu = \partial^\mu + ig_s \frac{\lambda^a}{2} G_a^\mu(x)$  and is the covariant derivative term. Here  $\lambda^a$  are the generators of the  $SU(3)_C$  algebra. All interactions are given in terms of a single universal coupling  $g_s$  which is called strong coupling constant. The types of interactions between quarks and gluons is followed from the above lagrangian. A large body of experimental evidence 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**. The first evidence for gluons came in three jet events at **PETRA**.

Two important properties of strong force to be notes are :

**Confinement** [34], which means that the force between quarks does not diminish as they are separated. Because of this, when you do separate a quark from other quarks, the energy in the gluon field is enough to create another quark pair; they are thus forever bound into hadrons such as the proton and the neutron or the pion and kaon. Although analytically unproven, confinement is widely believed to be true because it explains the consistent failure of free quark searches.

**Asymptotic freedom** [34], which means that in very high-energy reactions, quarks and gluons interact very weakly creating a quarkgluon plasma. This prediction of QCD was first discovered in the early 1970s by David Politzer, Frank Wilczek and David Gross [35; 36].

## 2.3 Electroweak Theory and Higgs Mechanism

Electroweak theory is the unified theory of electromagnetic and weak interactions. The theory is described by the symmetry group  $SU(2)_L \otimes U(1)_Y$  [28]. The formulation of this theory takes into account large number of experimental observations that has been gathered over years. Parity symmetry is known to be maximally violated in weak interactions, we know that from various experiments that the vector bosons do not interact with right handed fermions, so different chiral components must be kept in different representations of the gauge group to have different interactions for them. To describe weak interactions, with several fermionic flavours and different properties for left- and right-handed fields; moreover, the left-handed fermions should appear in doublets, and we would like to have massive gauge bosons and  $Z$  in addition to the photon. The simplest group with doublet representations is  $SU(2)$ . We want to include also the electromagnetic interactions; thus we need 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  $Y$  refers to hypercharge. We will describe this later in the chapter. For the discussion in this section let us denote the following quantities

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

We will treat the leptons in the exactly similar way. The free lagrangian in terms of these fields would be

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

## 2 The Standard Model of Particle Physics

$L_0$  is invariant under global G transformations in flavour 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  $U_L = e^{i\frac{\sigma^i}{2}\alpha^i}$  ( $i = 1, 2, 3$ ). If we require the lagrangian to be invariant under local gauge transformations we will have to introduce covariant derivatives as defined in the following way

$$D_\mu \psi_1 x = [\partial_\mu + ig\bar{W}_\mu(x) + ig'y_1 B_\mu(x)]\psi_1(x)$$

$$D_\mu \psi_2 x = [\partial_\mu + ig'y_2 B_\mu(x)]\psi_2(x)$$

$$D_\mu \psi_3 x = [\partial_\mu + ig'y_3 B_\mu(x)]\psi_3(x)$$

where  $\bar{W}_\mu = \frac{\sigma^i}{2} W_\mu^i(x)$ . There are four different gauge fields( $W_\mu^i(x)$ s and  $B_\mu(x)$ ) aimed to describe for 4 gauge bosons. So now the local gauge invariant lagrangian would be given by

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

We note here that this above lagrangian is without a mass term. Mass term 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

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

$$\bar{W}_{\mu\nu} = \partial_\mu \bar{W}_\nu - \partial_\nu \bar{W}_\mu + ig[W_\mu, W_\nu]. \text{ So now the kinetic term is given by}$$

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

So the total electroweak lagrangian is given by  $L^{tot} = L + L_{kin}$ . As we understand still the lagrangian does not have mass term for the bosons and fermions. In order to have mass term we will have to break the symmetry of the lagrangian somehow, so that the gauge bosons ( $W^+, W^-, Z$ ) becomes massive and the photon field should remain massless.

## 2 The Standard Model of Particle Physics

In the quantum field theory, a symmetry is ‘spontaneously’ broken when the Lagrangian itself remains invariant and the Hamiltonian of the theory attains 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) [38].

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  $\mu$ . We have two cases here

- $\mu^2 > 0$ , For this the potential has trivial minimum and it describes a massive scalar particle with mass  $\mu$  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 parametrizing the excitations over the ground state as

$\phi(x) \equiv \frac{1}{\sqrt{2}}[v + \varphi_1(x) + \varphi_2(x)]$ . Here  $\varphi_1$  and  $\varphi_2$  are real fields. Now the potential takes the form

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

So here  $\varphi_1$  describes a massive particle with mass  $-2\mu^2$  and  $\varphi_2$  is massless. These bosons are called goldstone bosons. Infact this is quite general result and follows from Goldstone

## 2 The Standard Model of Particle Physics

Theorem [37]. At first look, the Goldstone theorem does little to our mass problem. But problem could be solved through another mechanism called Higgs mechanism [38].

### 2.3.1 Higgs Mechanism

To discuss higgs mechanism [38] let us consider a  $SU(2)_L$  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 in SM becomes

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

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  $\phi^0$  is given by

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

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

$$\phi(x) = e^{i\frac{\sigma_2}{2}\theta^i(x)} \frac{I}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}_L \quad (2.11)$$

## 2 The Standard Model of Particle Physics

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

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

So now we could see from the above term that the gauge bosons has aquared masses. The relationship of the masses are given by

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

So this is how the higgs mechanism is responsible for giving masses to vector bosons. Fermions also get masses by higgs through Yukawa term [28].

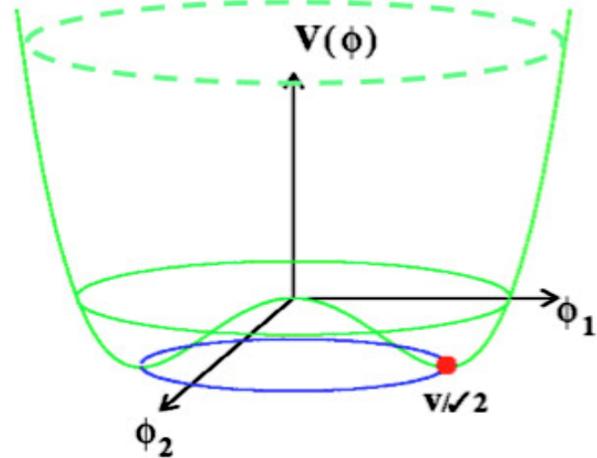


Figure 2.3: The potential  $V(\phi)$  for a complex scalar field  $\phi = (\phi_1 + i\phi_2) / \sqrt{2}$  where  $\mu^2 < 0$  and  $\lambda > 0$  with a minimum at  $v^2 = -\mu^2/\lambda$ .

## 2.4 Experimental verification of SM

SM has been phenomenally successful in predicting existence of several particles. All of them have been observed experimentally. Higgs-Kibble mechanism has produced 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. So it tells  $M_Z$  would be bigger than  $M_W$  and is experimentally verified [39; 40].

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

The top quark, being the heaviest one among the fermions, constitutes an important ingredient for precision electroweak tests and indirect determination of higgs boson mass. It was observed for the first time at Tevatron. The recent world-average top mass, measured at LHC and Tevatron is  $173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})$  GeV [41] .

Another important aspect of SM is the spontaneous electroweak symmetry breaking mechanism which was proposed almost 50 years ago by Higgs, Brout, Guralnik, Hagen and Kibble to generate the SM fermions and gauge bosons [38]. This last missing piece Higgs boson was discovered on July 4th by LHC(both CMS [32] and ATLAS [31]) in 2012. The higgs boson was discovered from its high resolution decay channels to  $\gamma\gamma$  and four lepton (through ZZ) with local significance of  $5\sigma$ .

## 2.5 Drawbacks of Standard Model

Even though SM has been very successful in describing many aspects but it is not the complete theory as evidenced by several experimental observations and mathematical flaws. Some of them can be summarized as

## 2 The Standard Model of Particle Physics

- Standard model does not explain gravitation. It does not explain the theory of gravitation , general relativity in terms of quantum field theory. One of the implication is that quantum field theories of gravity generally break down before reaching the planck scale. As a result we do not have a theory of early universe.
- Further there are 19 independent parameters in the SM: 9 Yukawa couplings, 3 CKM angles and 1 CP violating phase, 3 coupling constants, the  $\mu$  and  $\lambda$  parameter of higgs potential and QCD vacuum angle  $\theta_{QCD}$  linked to the strong CP problem. Given these , it is natural to consider the SM as a low-energy effective theory, rather being a fundamental theory.
- Experiments show that neutrino has mass but the SM does not allow mass in it. In the mathematical framework of standard model neutrinos are taken to be mass less.
- Dark energy is an unknown form of energy which is hypothesized to permeate all of space, tending to accelerate the expansion of the universe. Several astrophysical observations suggest that universe contains dark matter [42]. Dark energy is the most accepted hypothesis to explain the observations since the 1990s indicating that the universe is expanding at an accelerating rate. SM does not explain the existence of dark energy and dark matter.
- SM also fails to explain matter-antimatter asymmetry.
- Higgs mechanism in the standard model gives rise to hierarchy problem. Hierarchy problem is the large discrepancy between the aspects weak force and gravity. It is not understood why weak force is  $10^{24}$  times stronger than gravity. More techically why the higgs boson is so much ligher than the planck mass. In the strict context of standard model the self enegy correction of higgs diverges. One of the proposed solution to this is given by Supersymmetry. We will discuss about this in the following chapter.

# 3 Beyond the Standard Model

Physics beyond the standard model(BSM) [43] refers to the theoretical developments needed to explain the deficiencies of the SM. Theories that lie beyond the Standard Model include various extensions of the standard model through supersymmetry [44], such as the Minimal Supersymmetric Standard Model(MSSM) [45] and Next-to-Minimal Supersymmetric Standard Model(NMSSM) [46], Lepto-quark models or entirely novel explanations, such as string theory, M-theory, and extra dimensions [47]. However we will discuss about supersymmetry(SUSY) and Lepto-quark models in the following sections.

## 3.1 Supersymmetry

Supersymmetry(SUSY) [44] is one of the proposed solution to the standard model that extends SM by introducing a new type of space-time symmetry that relates two basic types of elementary particles: bosons and fermions. As a result in SUSY each SM particle has its so called superpartner differing by spin  $\frac{1}{2}$ . In a theory where supersymmetry is unbroken, all the SUSY partners will have the same mass as their SM particles. But since no supersymmetric partners have been found, tells that SUSY must be a broken symmetry. Broken SUSY could solve many problems in the particle physics including the hierarchy problem. The simplest

### 3 Beyond the Standard Model

realization of this broken SUSY is minimal supersymmetric standard model(MSSM). A supersymmetric transformation converts a bosonic state to a fermionic state and vice versa. If  $Q$  denotes the operator which does that then

$$Q|Boson\rangle = |Fermion\rangle, Q|Fermion\rangle = |Boson\rangle \quad (3.1)$$

with  $Q$  being an anti-commuting spinor. The supersymmetric partner(SP) of fermion is called sfermion. For example quarks have SP called as squarks. Leptons have SP called as sleptons. The left-handed and right-handed pieces of the quarks and leptons are separate two-component Weyl fermions with different gauge transformation properties in the Standard Model, so each must have its own complex scalar partner. The symbols for the squarks and sleptons are the same as for the corresponding fermion, but with a tilde  $\sim$  used to denote the superpartner of a Standard Model particle. For example, the superpartners of the left-handed and right-handed parts of the electron Dirac field are called left- and right-handed selectrons, and are denoted  $\tilde{e}_L$  and  $\tilde{e}_R$ . But the handedness does not refer to the helicity of selectrons. But as the neutrinos are always lefthanded they are denoted just by  $\tilde{\nu}_e, \tilde{\nu}_\tau$  and  $\tilde{\nu}_\mu$ . Similarly for the quarks it is denoted by  $\tilde{q}_L$  and  $\tilde{q}_R$ , where q stands for 6 types of quarks. Theories show that Higgs cannot reside in one 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 supersymmetric partners 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 Standard Model 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  $\tilde{H}_u^+$  and  $\tilde{H}_u^0$  and  $\tilde{H}_d^0$  and  $\tilde{H}_d^-$ . Similarly the supersymmetric partner of gluon( $g$ ) is gluino( $\tilde{g}$ ). The electroweak gauge symmetry  $SU(2)_L \times U(1)_Y$  is associated with spin-1 gauge bosons  $W^\pm, W^0$  and  $B^0$  with spin  $\frac{1}{2}$

### 3 Beyond the Standard Model

superpartners  $\tilde{W}^\pm$ ,  $\tilde{W}^0$  and  $\tilde{B}^0$ , called winos and bino. After electroweak symmetry breaking the  $W^0$  and  $B^0$  eigen states mix to give mass eigen states  $Z^0$  and  $\gamma$ . The corresponding gaugino mixtures of  $\tilde{W}^0$  and  $\tilde{B}^0$  are called zino( $\tilde{Z}^0$ ) and photino ( $\tilde{\gamma}$ ). The list of SUSY and SM particles in terms of superfields in the context of MSSM are given in the table. 3.1. The particles and their SUSY partners are shown in Fig. 3.1.

Super Multiplet	S=0 or 1	S=1/2	$SU(3), SU(2)_L, U(1)_Y$
Q	squarks $(\tilde{u}_L, \tilde{d}_L)$	quarks $(u_L, d_L)$	$(3, 2, 1/6)$
	$\tilde{u}_R^*$	$u_R^\dagger$	$(\bar{3}, 1, -2/3)$
	$\tilde{d}_R^*$	$d_R^\dagger$	$(\bar{3}, 1, 1/3)$
L	sleptons $(\tilde{\nu}_L, \tilde{e}_L)$	leptons $(\nu_L, e_L)$	$(1, 2, -1/2)$
	$\tilde{e}_R^*$	$e_R^\dagger$	$(1, 1, 1)$
	Higgs $(H_u^+, H_u^0)$	Higgsino $(\tilde{H}_u^+, \tilde{H}_u^0)$	$(1, 2, 1/2)$
$H_d$	$(H_d^0, H_d^-)$	$(\tilde{H}_d^0, \tilde{H}_d^-)$	$(1, 2, -1/2)$
gauge bosons		gaugino	
$V^{(y)}$	$B_\mu$	$\tilde{B}_\mu$	$(1, 1, -)$
$V^{a(2)}$	$W_\mu^a$	$\tilde{W}_\mu^a$	$(3, 1, -)$
$V^{a(3)}$	$g_\mu^a$	$\tilde{g}_\mu^a$	$(8, 1, -)$

Table 3.1: Field Content of MSSM in terms of superfields(Ist coulmn), sparticles and particles(2nd and 3rd coulmn) and in 4th column the quantum numbers for various fields are shown.

## SUPERSYMMETRY

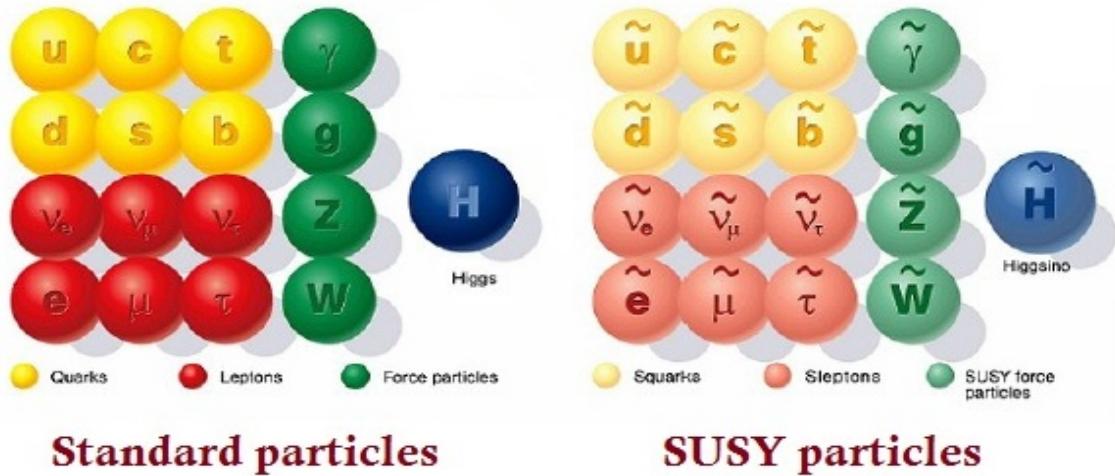


Figure 3.1: SM and their SUSY partners.

### 3.2 Lepto quark models

The leptoquark model

# 4 The Compact Muon Solenoid

## 4.1 CMS detector

The Compact Muon Solenoid (CMS) is one of the general purpose detector present at Large Hadron Collider(LHC), CERN near Geneva Switzerland. LHC is considered as 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 and to let them collide at few well defined points. CMS detector sits at one point and stores all the possible interesting events that results from proton-proton collisions. CMS is 21.6 metres long, 15 m in diameter, and weighs about 14,000 tonnes. Approximately 3,800 people, representing 199 scientific institutes and 43 countries, form the CMS collaboration who built and now operate the detector. It contains subsystems which are 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. All the study presented in this thesis is from the data collected by CMS experiment at a ~~centre of mass~~ energy of 13 TeV. Two diagrams of the detector are shown in figure

CMS has four major subcomponents, these are

- Tracker

## 4 The Compact Muon Solenoid

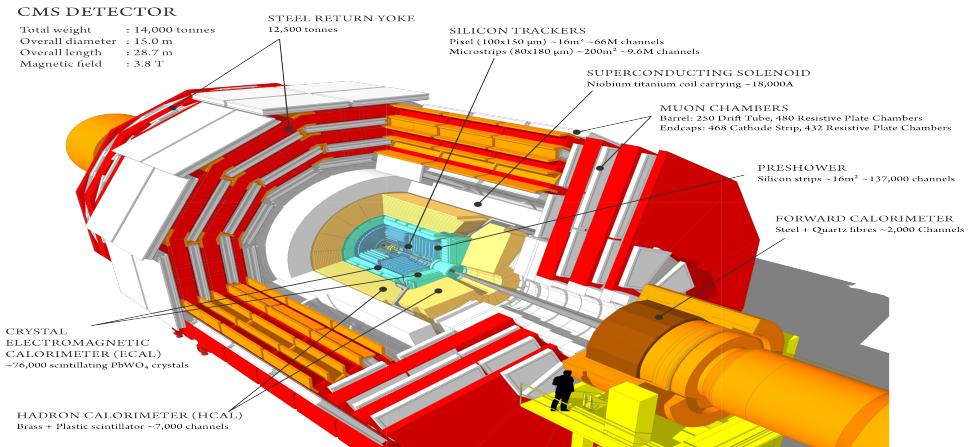


Figure 4.1: An inside view of CMS detector, showing various subdetectors that are placed around the beam pipe and form a series of cylindrical layers of the experiment.

- Electromagnetic Calorimeter
- Hadron Calorimeter
- Muon Chamber

We explain the individual components of the detector one by one

### 4.1.1 Tracker

The innermost part of the CMS detector is tracking system. It measures the momentum of charged particles by measuring the bending of charged particles in the presence of magnetic fields. 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 types of two main subcomponents, these are pixel detector and silicon microstrip detector. Pixel detector helps to detect the secondary vertices and hence in identifying the bjets. The coverage of Tracker if up to  $\eta < 2.5$ .

#### 4 The Compact Muon Solenoid

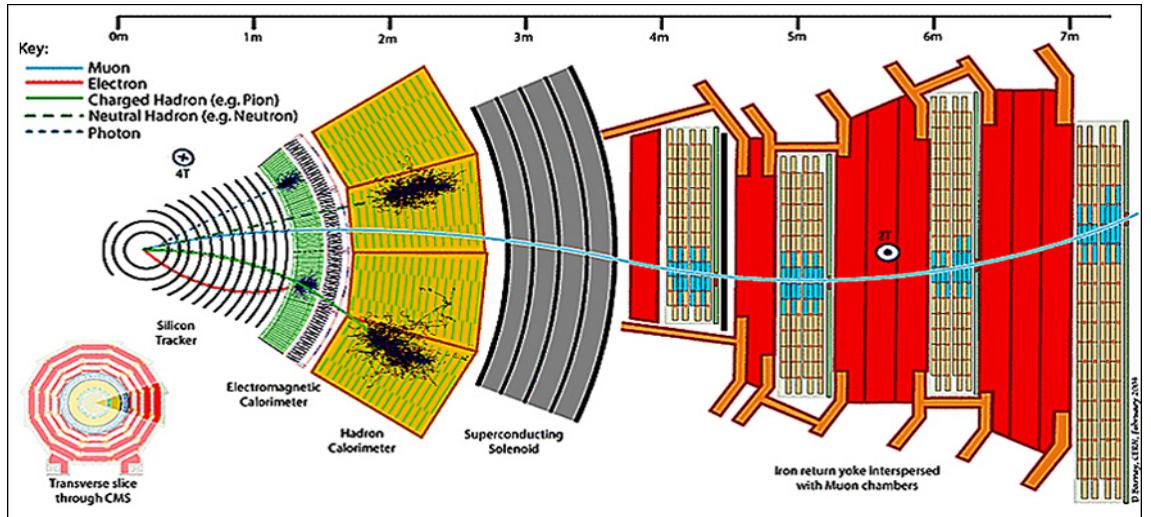


Figure 4.2: Various detector components that contribute to the event reconstruction are (from left to right) tracker, electromagnetic calorimeter, superconducting solenoid and muon system. The paths of different particles passing through the detector such as photons, muons, electrons, neutral hadrons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

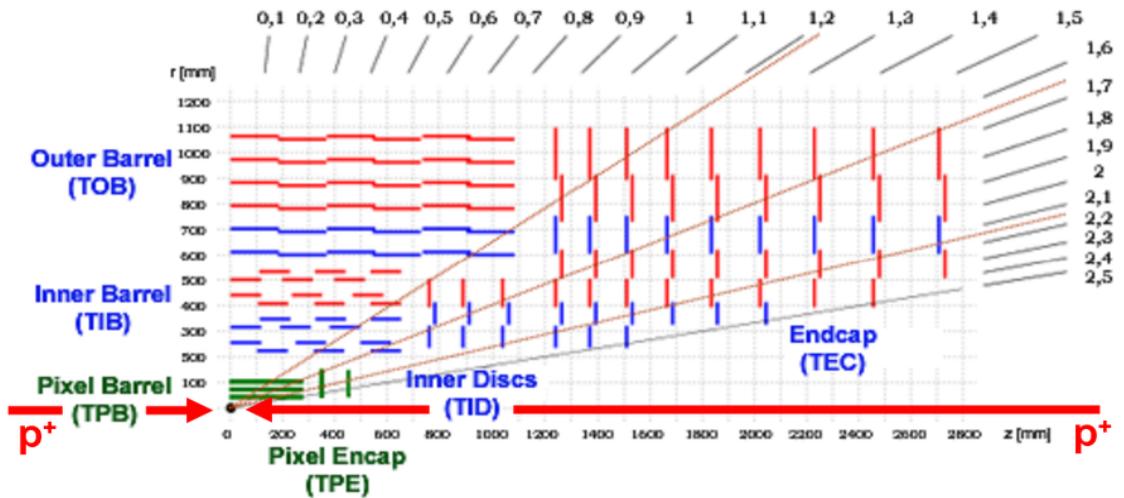


Figure 4.3: A schematic diagram of the CMS Tracker. The plot shows the quadrants of the tracking detector along the rz plane.

Pixel detector covers the innermost region of the system at an approximate distance of 10 cm from the IP. Hence pixel detector plays an important role in identifying number of primary

## 4 The Compact Muon Solenoid

vertices present in each event. This helps to reduce the pileup contamination in the event. The central part of the detector is called barrel, has cylindrical shape with its 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 in total consists of 66 million pixels. It is capable of detecting particles having pseudorapidity  $|\eta| \leq 2.4$  with average spatial resolution of  $10 \mu m$  along the  $r\phi$  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 silicon micro strip detectors that surrounds the pixel detector. It consists of 10 barrel layers(4 TIB-Tracker Inner Barrel and 6 TOBTracker-Outer-Barrel), 3 TIDTrackerInnerDisks and 9 TEC disks. A schematic diagram of cross sectional slice of one quarter is shown in the Figure. 4.3

### 4.1.2 Electromagnetic Calorimeter

Calorimeters 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 of it consisting of active volume. The CMS electromagnetic calorimeter(ECAL) is a homogeneous one that provides excellent energy and position resolution for photons and electrons. The active materials of the ECAL are lead tungstate crystals( $PbWO_4$ ) providing around 25 radiation lengths( $X_0$ ). A particle entering the ECAL results in an electromagnetic shower caused by the successive 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 hav-

## 4 The Compact Muon Solenoid

ing energy higher than 1 GeV lose energy mostly via bremsstrahlung and photons via pair production. The secondary particles produced by the primary one interact through the same process leading to the development of a particle shower inside the calorimeter.

The CMS ECAL consists of 75848 lead tungstate ( $PbWO_4$ ) crystals. The radiation length of this crystal is 0.89 cm. The lateral spread of the shower is 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.

ECAL has two parts : ECAL 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. Silicon strip sensors with a resolution of 2mm, placed behind two planes of lead are able to distinguish single photons from  $\pi^0$ 's decaying into photon pairs that would not be possible with a typical size of 30 mm.

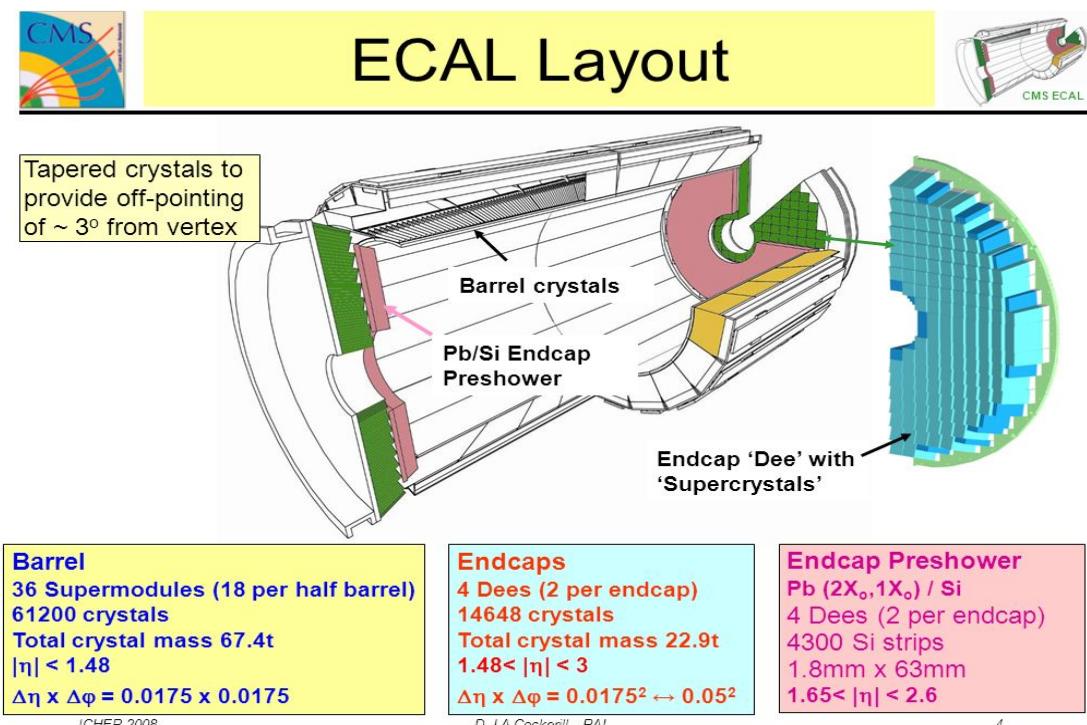


Figure 4.4: A schematic diagram of the CMS ECAL sub-detector showing its different components.

## 4 The Compact Muon Solenoid

### ECAL Performance

The energy resolution of 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 are follows. There are three basic contributions to the stochastic term:

#### 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 preshower absorber with respect to what is measured in the preshower silicon detector.

#### Constant term

- non uniformity of longitudinal light collection
- intercalibration errors
- leakage of the energy from the back of the crystal

#### Noise term

There are three basic sources of noise that contribute to the noise term

- electronics noise
- digitization noise

## 4 The Compact Muon Solenoid

- pileup noise

The energy resolution(vs. pseudorapidity) of ECAL is show in Figure . 4.5

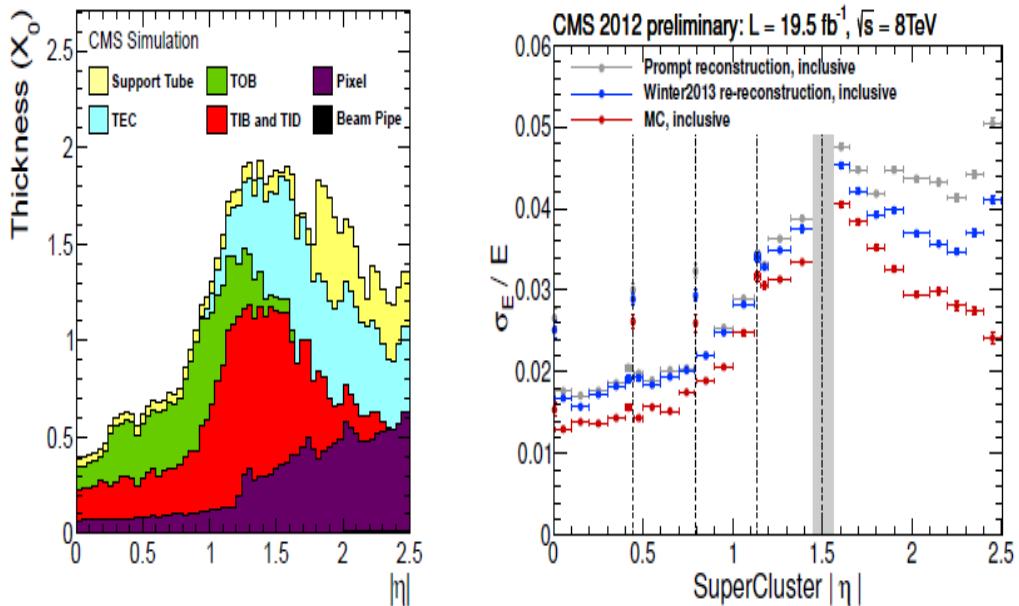


Figure 4.5: Amount of material (in units of radiation lengths) upstream of the ECAL (left). Electron energy resolution as a function of pseudorapidity for electrons from  $Z \rightarrow e^+e^-$  decays. In red, the energy resolution estimated on Monte Carlo events is shown, while in blue and gray the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

### 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 provides indirect measurement of the presence of non-interacting, uncharged particles such as neutrinos. The

#### *4 The Compact Muon Solenoid*

HCAL consists of layers of dense material (brass or steel) interleaved with tiles of plastic scintillators, read out via wavelength-shifting fibres by hybrid photodiodes. This combination was determined 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. Located 11 m either side of the interaction point, this uses a slightly different technology of steel absorbers and quartz fibres for readout, designed to allow better separation of particles in the congested forward region. The HF is also used to measure the relative online luminosity system in CMS.

Hadron calorimeter 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 transverse momentum resolution is of crucial importance for many physics analyses and poses a big challenge for design of hadronic calorimeter. The CMS HCAL is a sampling calorimeter surrounding the aforementioned parts of the detector. 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  $\lambda_l = 16$  cm. The large fraction of passive material is cost efficient. CMS HCAL has four main subcomponents as shown in the Figure. 4.6

#### 4 The Compact Muon Solenoid

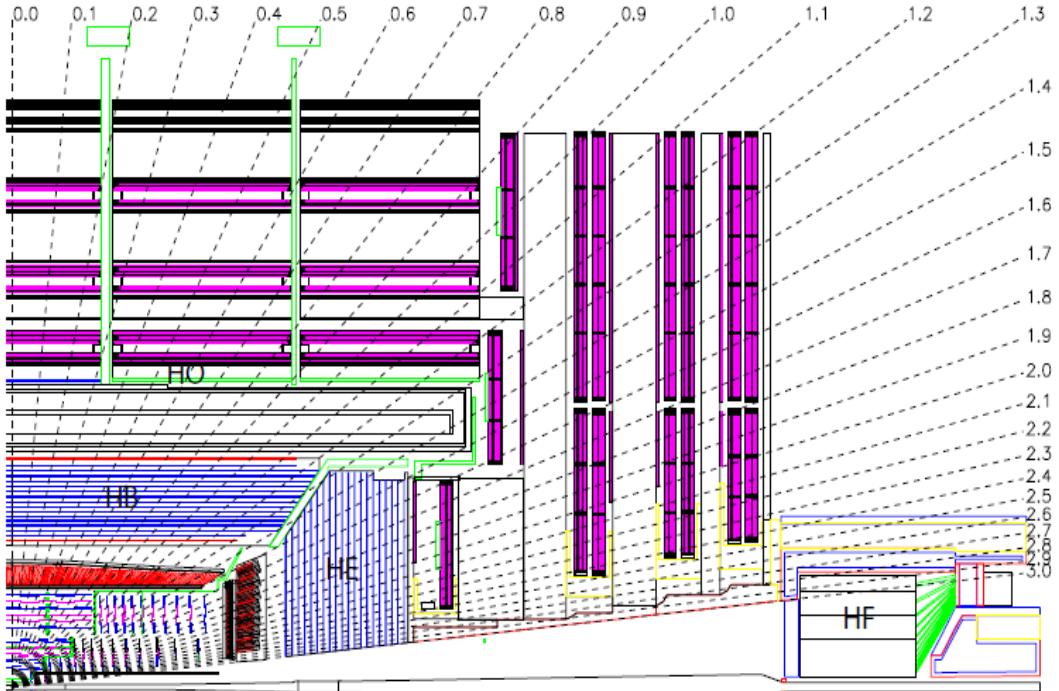


Figure 4.6: A schematic diagram of the CMS HCAL showing its individual components.

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  $\eta$  range  $2.8 < |\eta| < 5.2$ , HF is located at  $z = \pm 11.2m$  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 system.

**HCAL Performance** The hadron energy resolution for the barrel HCAL and ECAL combination can be parametrized as

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

## 4 The Compact Muon Solenoid

where a is the stochastic term and b is the constant term. These two terms are measured in the test beam as  $0.847 \pm 0.016 \sqrt{GeV}$  and  $0.074 \pm 0.008$  respectively.

### 4.1.4 Muon System

As the name suggests, detecting muons is one of CMSs most important tasks. Muons are less interacting particles and pass through several metres of iron without much ionization loss and bremsstrahlung and are not stopped by any of the calorimeters of CMS. Therefore, chambers to detect muons are placed at the very edge of 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 (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The DTs are used for precise trajectory measurements in the central barrel region, while the CSCs are used in the end caps. The RPCs provide a fast signal when a muon passes through the muon detector, and are installed in both the barrel and the end caps.

#### Drift Tubes

The muon barrel (MB) consists of four layers of drift tube chambers (DTs). 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 atoms of the gas. 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's original distance away from the wire, the DTs give two coordinates for the muons's position.

#### Cathode Strip Chambers

The muon endcap consists of 468 cathode strip chambers(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.CSCs consist of arrays of positively-charged anode wires crossed with

## *4 The Compact Muon Solenoid*

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 wire and towards the copper cathode, also 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 make the CSCs fast detectors suitable for triggering. Each CSC module contains six layers making it able to accurately identify muons and match their tracks to those in the tracker.

### **Resistive Plate Chambers**

Resistive plate chambers (RPC) are fast gaseous detectors that provide a muon trigger system parallel with those of the DTs and CSCs.

RPCs 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 chamber, 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 (the 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 keeping. RPCs combine a good spatial resolution with a time resolution of just one nanosecond. Figure. 4.7 shows the schematic diagram of RPC.

#### 4 The Compact Muon Solenoid



Figure 4.7: Muon resistive plate chamber

#### Muon System Performance

The muon momentum is measured in the inner tracker as well as in the return yoke flux. 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. For high  $p_T$  muons (above 100 GeV), the resolution degrades to about 8% for the barrel and 10 % for the endcaps as shown.

## 5 Statistical procedures for Search

In this section we discuss the general statistical procedures for interpreting data by calculating the upper limits and significances. We will also discuss how nusances are treated while calculating the upper limits. The concepts we will be discussing here will be used both in the SUSY and LeptoQuark analysis to calculate the upper limits.

Both Bayesian and frequentist methods of calculating upperlimits are used in CMS. However we will discuss more about the frequentist approach of calculating the upper limits and significances as in this analyses we use a frequentist method. These methods allow one to quantify the level of incampatibility of the data with the signal hypothesis which is expressed as confidence level (CL). It is common to require 95% CL for excluding a signal, however this is a convention.

In this section we denote  $s$  as the signal yield for any analysis ,  $b$  is the background yield. We also name  $\mu$  to be the signal strength that is defined as the  $\mu = s_0/s$  where  $s$  is our nominal signal yield and  $s_0$  is change signal yield after scaled by  $\mu$ . Predictions for both 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  $\theta$ , so that signal and background expectations become functions of the nuisance parameters:  $s(\theta)$  and  $b(\theta)$ .

## 5 Statistical procedures for Search

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% / 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 factorised form.

### 5.1 Observed limits

For calculating observed limits we use a test statistic that is widely known as LHC test statistic. In the following we will construct the test statistic by defining few things . First lets construct a likelihood function

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

where data represents either the actual experimental observation or pseudodata used to construct sampling distributions.  $\mu$  is the signal strength and  $\theta$  represents all the nuisance parameters associated.  $\text{Poisson}[\text{data} | \mu s + b]$  stands for a product of probabilities to observe  $n_i$  events in bin  $i$ :

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

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  $\mu$ , we construct the test statistic  $\tilde{q}_\mu$ :

## 5 Statistical procedures for Search

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

with a constraint  $0 \leq \hat{\mu} \leq \mu$ . Where  $\hat{\theta}_\mu$  refers to the conditional maximum likelihood estimators of  $\theta$ , given signal strength parameter  $\mu$  and data that, as before, may refer to the actual experimental observation or pseudo data. 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 ), and 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 signal hypothesis.

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

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

and

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

- Then we calculate  $CL_s$  as the ratio of these two probabilities i.e.

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

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

## 5.2 Expected limits limits

The most straightforward way for defining the expected median upper-limit and  $\pm 1\sigma$  and  $\pm 2\sigma$  bands for the background-only hypothesis is to generate a large set of background only pseudo-data and calculate  $CL_s$  and  $\mu^{95\%CL}$  for all of them as if they were real data. hen, one can build a cumulative probability distribution of results by starting integration from the side corresponding 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 the signal is quantified by the background-only p-value, i.e. the probability for the background to fluctuate and give an excess of events as large or larger than the observed one. The test statistic that is used to get the p-value for the purpose of significance is slightly different one and is defined as the following

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

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

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

To convert the p-value to significance Z, we use a normal gaussian and calculate according to the following

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

The basic principles of approach are what we discussed here. Details of how systematics are treated will be discussed in the respective analysis result section.

# **6 Search for Supersymmetry with 13 TeV pp collision data**

As discussed earlier SUSY is one the BSM theory that provides solutions to the problems of standard model. It provides a way to cancel the higgs mass divergence problem, provide a candidate for dark matter and unification of strong , weak and electromagnetic forces. Natural SUSY models minimize the fine tuning associated with the value of the Higgs boson mass and its radiative corrections. In natural SUSY, 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 gluino pair production has largest cross-section and the cross section increases in the order of 30 from 8 TeV to 13 TeV. This fact makes it pertinent to search for SUSY even with couple of femtobarn inverse of data. The results are interpreted with R-parity conserving simplified SUSY scenarios.

Here in this chapter we describe the study first with  $2.3 \text{ fb}^{-1}$  of 2015 data and with  $12.9 \text{ fb}^{-1}$  of 2016 data later. 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

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

to be easily hidden in extensions to canonical SUSY models.

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

The event diagrams of gluino pair production that we considered in our analysis are shown below.

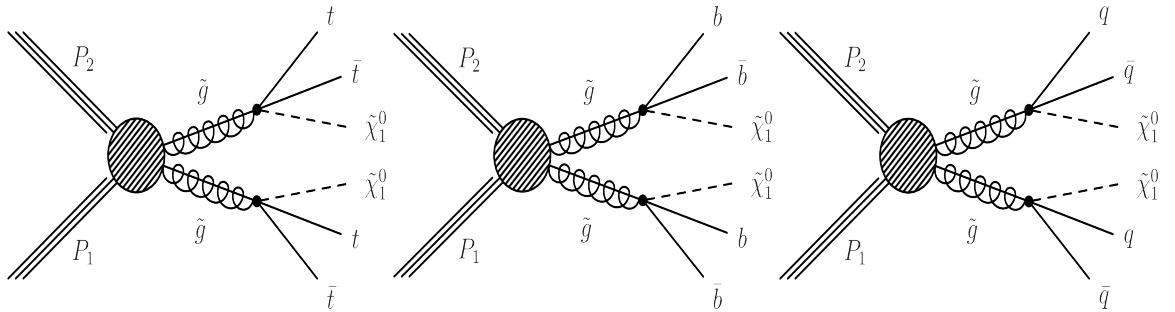


Figure 6.1: Event diagrams for gluino pair production with four top final state(left), four b-quark final state(middle) and four light quark final state(right)

The major backgrounds for this type of final state comes from  $\text{top}/W + \text{jets}$ ,  $Z(\rightarrow \nu\bar{\nu}) + \text{jets}$ , QCD processes. All the backgrounds are estimated using data driven techniques. Use of montecarlo is minimized as possible. The details about the cross sections of the simulated samples are given in the following section.

### 6.0.1 Object reconstruction

## 6.1 Study with $2.3 \text{ fb}^{-1}$ of 2015 data

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

### 6.1.1 Data samples and Trigger

Monte Carlo (MC) samples that are used in the analysis where ever required are given 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. The samples simulates a pileup distribution with an average of 20 interactions per bunch crossing and a 25 ns interval between bunches.

Table 6.1: MC FullSim samples for signal SMS model points.

Dataset	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
SMS-T1tttt ( $m_{\text{gluino}} = 1500, m_{\text{LSP}} = 100$ )	madgraphMLM-pythia8	0.014	7268
SMS-T1tttt ( $m_{\text{gluino}} = 1200, m_{\text{LSP}} = 800$ )	madgraphMLM-pythia8	0.086	1719
SMS-T1bbbb ( $m_{\text{gluino}} = 1500, m_{\text{LSP}} = 100$ )	madgraphMLM-pythia8	0.014	3708
SMS-T1bbbb ( $m_{\text{gluino}} = 1000, m_{\text{LSP}} = 900$ )	madgraphMLM-pythia8	0.325	438.5
SMS-T1qqqq ( $m_{\text{gluino}} = 1400, m_{\text{LSP}} = 100$ )	madgraphMLM-pythia8	0.025	1958
SMS-T1qqqq ( $m_{\text{gluino}} = 1400, m_{\text{LSP}} = 800$ )	madgraphMLM-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.

Dataset	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
$t\bar{t} + \text{jets}$	madgraph, pythia8	816.0	13.90
$t\bar{t} + \text{jets}$ (SingleLeptonFrom t)	madgraph, pythia8	179.3	324.6
$t\bar{t} + \text{jets}$ (SingleLeptonFrom $\bar{t}$ )	madgraph, pythia8	179.3	335.7
$t\bar{t} + \text{jets}$ (DiLepton)	madgraph, pythia8	86.66	351.2
$t\bar{t} + \text{jets}$ ( $H_T$ -600to800)	madgraph, pythia8	2.615	1898
$t\bar{t} + \text{jets}$ ( $H_T$ -800to1200)	madgraph, pythia8	1.077	3198
$t\bar{t} + \text{jets}$ ( $H_T$ -1200to2500)	madgraph, pythia8	0.195	5063
$t\bar{t} + \text{jets}$ ( $H_T$ -2500toInf)	madgraph, pythia8	0.002	218575

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

Dataset	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
QCD ( $H_T$ )-200to300)	madgraph, pythia8	1735000	0.01
QCD ( $H_T$ )-300to500	madgraph, pythia8	366800	0.05
QCD ( $H_T$ )-500to700	madgraph, pythia8	29370	0.67
QCD ( $H_T$ )-700to1000	madgraph, pythia8	6524	2.30
QCD ( $H_T$ )-1000to1500	madgraph, pythia8	1064	4.67
QCD ( $H_T$ )-1500to2000	madgraph, pythia8	121.5	31.67
QCD ( $H_T$ )-2000toInf	madgraph, pythia8	25.42	77.17

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

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

Dataset	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
$Z(\nu\bar{\nu}) + \text{jets}, H_T\text{-}100\text{To}200$	madgraph	345.0	14.92
$Z(\nu\bar{\nu}) + \text{jets}, H_T\text{-}200\text{To}400$	madgraph	96.38	52.22
$Z(\nu\bar{\nu}) + \text{jets}, H_T\text{-}400\text{To}600$	madgraph	13.46	75.34
$Z(\nu\bar{\nu}) + \text{jets}, H_T\text{-}600\text{ToInf}$	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	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
$W(\ell\nu) + \text{jets}_{H_T\text{-}100\text{To}200}$	madgraph, pythia8	1635	6.20
$W(\ell\nu) + \text{jets}_{H_T\text{-}200\text{To}400}$	madgraph, pythia8	437.0	11.97
$W(\ell\nu) + \text{jets}_{H_T\text{-}400\text{To}600}$	madgraph, pythia8	59.50	31.96
$W(\ell\nu) + \text{jets}_{H_T\text{-}600\text{ToInf}}$	madgraph, pythia8	22.80	45.44
$W(\ell\nu) + \text{jets}_{H_T\text{-}600\text{To}800}$	madgraph, pythia8	15.50	257.1
$W(\ell\nu) + \text{jets}_{H_T\text{-}800\text{To}1200}$	madgraph, pythia8	6.366	247.4
$W(\ell\nu) + \text{jets}_{H_T\text{-}1200\text{To}2500}$	madgraph, pythia8	1.614	158.4
$W(\ell\nu) + \text{jets}_{H_T\text{-}2500\text{ToInf}}$	madgraph, pythia8	0.037	6770

### 6.1.1.1 Trigger

The information about the signal region trigger and  $Z$  to invisible background control region trigger are given here. We do not discuss about the other background region triggers here.

**Signal Region Trigger** The data collected with the CMS experiment at the center-of-mass energy ( $\sqrt{s}$ ) of 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 with a trigger with online  $H_T > 250$  GeV and online MET(missing transverse energy)  $> 100$  GeV. The technical name of the path is HLT-PFHT350-PFMET100-NoiseCleaned. Here HLT stands for High Level Trigger. The trigger is found to be almost fully efficient for offline  $H_T > 500$  GeV and  $H_T^{\text{miss}} > 200$  GeV.

### Control Region Trigger

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

## 6 Search for Supersymmetry 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	$\sigma$ (pb)	(fb $^{-1}$ )
ttHJetTobb M125 13TeV amcatnloFXFX madspin pythia8	0.293	18269
TTZToLLNuNu M-10 TuneCUETP8M1 13TeV-amcatnlo-pythia8	0.228	811.4
TTZToQQ TuneCUETP8M1 13TeV-amcatnlo-pythia8	0.530	663.4
TTWJetsToLNu TuneCUETP8M1 13TeV-amcatnloFXFX-madspin-pythia8	0.204	635.6
TTWJetsToQQ TuneCUETP8M1 13TeV-amcatnloFXFX-madspin-pythia8	0.423	1018
ZH HToBB ZToNuNu M125 13TeV amcatnloFXFX madspin pythia8	0.100	12116
WH HToBB WToLNu M125 13TeV amcatnloFXFX madspin pythia8	0.260	4782
WWTo1L1Nu2Q 13TeV amcatnloFXFX madspin pythia8	50.00	64.26
WWTo2L2Nu 13TeV-powheg	12.18	158.5
WZTo1L1Nu2Q 13TeV amcatnloFXFX madspin pythia8	10.71	1339
WZTo1L3Nu 13TeV amcatnloFXFX madspin pythia8	3.058	305.7
ZZTo2Q2Nu 13TeV amcatnloFXFX madspin pythia8	4.040	5556
ZZTo2L2Q 13TeV amcatnloFXFX madspin pythia8	3.220	3706
TTTT TuneCUETP8M1 13TeV-amcatnlo-pythia8	0.009	57031
WWZ TuneCUETP8M1 13TeV-amcatnlo-pythia8	0.165	1341
WZZ TuneCUETP8M1 13TeV-amcatnlo-pythia8	0.056	3938
ZZZ TuneCUETP8M1 13TeV-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	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
$Z(\ell^+\ell^-) + jets$	madgraphMLM-pythia8	6025	1.50
$Z(\ell^+\ell^-) + jets, H_T\text{-}100\text{to}200$	madgraphMLM-pythia8	171.5	15.31
$Z(\ell^+\ell^-) + jets, H_T\text{-}200\text{to}400$	madgraphMLM-pythia8	52.58	18.18
$Z(\ell^+\ell^-) + jets, H_T\text{-}400\text{to}600$	madgraphMLM-pythia8	6.761	155.0
$Z(\ell^+\ell^-) + jets, H_T\text{-}600\text{to}Inf$	madgraphMLM-pythia8	2.718	363.5

### 6.1.2 Event selection criterias and exclusive search intervals

The following criterias define our baseline selection.

- $N_{jet} > 4$ ;

Since every pair of gluinos decays to four quarks, all events are required to contain at least four good jets, defined by

- $p_T > 30$  GeV,
- $|\eta| < 2.4$ ,
- satisfying loose jet ID criteria for PF jets defined by the following

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

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

Dataset	Generator	$\sigma$ (pb)	(fb $^{-1}$ )
$\gamma + jets, H_T\text{-}100\text{To}200$	madgraphMLM-pythia8	22010	0.23
$\gamma + jets, H_T\text{-}200\text{To}400$	madgraphMLM-pythia8	9110	1.13
$\gamma + jets, H_T\text{-}400\text{To}600$	madgraphMLM-pythia8	273	9.07
$\gamma + jets, H_T\text{-}600\text{ToInf}$	madgraphMLM-pythia8	94.5	26.99

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

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

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

- Muon veto:

Muon candidates are selected using the POG-recommended Medium Muon

- Electron veto:

Electron candidates are selected using the POG-recommended Cut Based VETO selection

- Angular cut  $\Delta\phi(j_1, H_T^{\text{miss}}) > 0.5$ ,  $\Delta\phi(j_2, H_T^{\text{miss}}) > 0.5$ ,  $\Delta\phi(j_3, H_T^{\text{miss}}) > 0.3$ ,  $\Delta\phi(j_4, H_T^{\text{miss}}) > 0.3$

The majority of QCD multijet events in our high- $H_T^{\text{miss}}$  search region have jets with undermeasured momenta and thus a spurious momentum imbalance. A signature of such an event is a jet closely aligned in direction with the  $H_T^{\text{miss}}$  vector. To suppress this background, we reject all events in which the two highest-  $p_T$  jets lie within 0.5 radians of the  $H_T^{\text{miss}}$  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.

The plots below shows the distribution of search variables after the base line selection. Simulated MCs are used for these plots.

## 6 Search for Supersymmetry with 13 TeV pp collision data

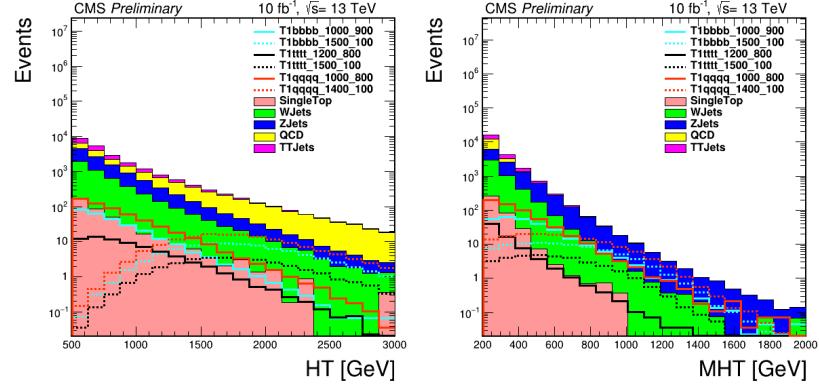


Figure 6.2: Plot shows signal vs. stacked backgrounds in  $H_T$  ( HT, left) and  $H_T^{\text{miss}}$  ( MHT, right)

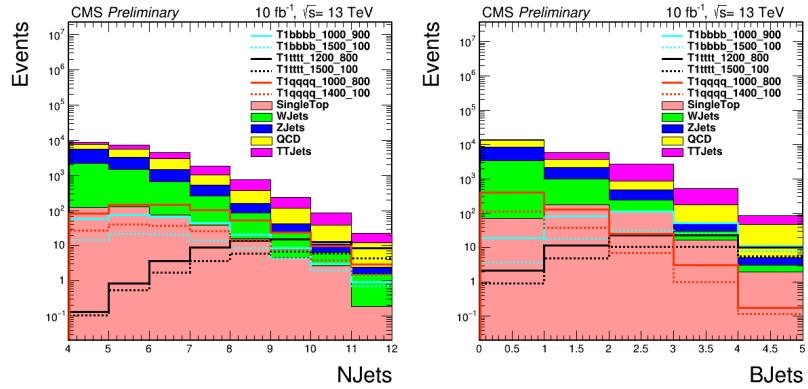


Figure 6.3: Plot signal vs. stacked backgrounds in  $N_{\text{jet}}$ (NJets, left) and  $N_{b\text{-jet}}$ (MHT, right)

### Search Binning

On top of baseline we split the phase space into 72 exclusive intervals to enhance the sensitivity.

The search variables and corresponding binning are as follows:

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

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

The analysis is restricted to  $N_{\text{jet}} \geq 4$  because of the focus on gluino pair production. The  $H_T > 500\text{GeV}$  and  $H_T^{\text{miss}} > 200\text{GeV}$  restrictions are dictated by the trigger conditions.

The binning is rectangular as above, except the 12 bins with  $H_T < 800\text{GeV}$  and  $H_T^{\text{miss}} > 750\text{GeV}$  are removed since  $H_T^{\text{miss}}$  cannot exceed (or be on the order of)  $H_T$  in a physical event. Additionally, for  $500 < H_T^{\text{miss}} < 750\text{GeV}$ , a single range with  $500 < H_T < 1200\text{GeV}$  is used, and for  $H_T^{\text{miss}} > 750\text{GeV}$ , a single range with  $H_T > 800\text{GeV}$  is used, because of the lower expected number of events at large  $H_T^{\text{miss}}$ . The merging of these bins further reduces the bin count by 24, for a total of 72 bins. The six bins in the  $H_T$  and  $H_T^{\text{miss}}$  plane are shown visually in Fig. 6.4.

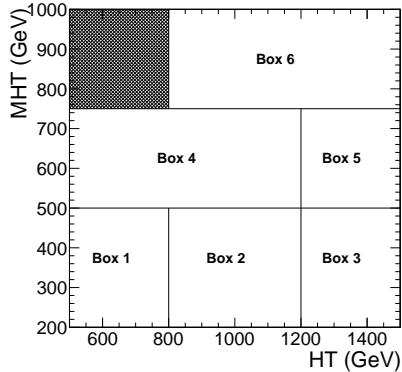


Figure 6.4: Search binning in 2D phase space of  $H_T$  and  $H_T^{\text{miss}}$

### 6.1.3 Backgrounds

All the standard model backgrounds have been categorized into four different groups. These are 1.Z to invisible, 2.Lostlepton, 3.Hadronic Tau, 4.QCD.  $Z(\rightarrow \nu\bar{\nu})+\text{jets}$  is an irreducible background. It has exactly the same final state as the signal. Lost lepton background arises when any top or W decays leptonically and the lepton goes unidentified, unisolated or out of acceptance. In case the W decays to taus and taus decay hadronically, it becomes hadronic tau background. QCD processes have fake MET due to jet energy mismeasurement, so it enters

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

the signal region. All the backgrounds mentioned here are determined using data driven techniques because the montecarlos usually do not simulate the higher order processes correctly. We will discuss the estimation of Z to invisible background extensively.

### 6.1.3.1 $Z(\rightarrow \nu\bar{\nu})$ 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 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  $\gamma$ +jets sample to determine the yields in the 18 bins corresponding to  $N_{b-jet}=0$ . 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  $\gamma$ +jets, and the normalization corrected if necessary. The extrapolation to bins with  $N_{b-jet} > 0$  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 given by the following

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

expression.

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

Where

- $N_{Z(\nu\nu)+\text{jets}}^{\text{prediction}}$  is the number of predicted  $Z(\nu\nu) + \text{jets}$  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 gets contaminated from  $\pi^0 \rightarrow \gamma\gamma$  processes. We measure the purity also using data driven techniques. This will be discussed in later in the following sections.
- $R_{Z \rightarrow \ell^+ \ell^- / \gamma}^{\text{MC}}$  is the ratio of production rates at reconstructed level. This is calculated using montecarlo. But as we do not trust montecarlo fully, we derive a correction to this ratio by calculating the double ratio. Double ratio is defined as the ratio of ratios of  $Z \rightarrow \ell^+ \ell^-$  events to  $\gamma + \text{jets}$  events calculated in data and montecarlo respectively.

### Photon Reconstruction

For the single photon control sample, we require at least one well identified and isolated photon candidate having at least 100 GeV of transverse momentum and pseudorapidity  $|\eta| < 2.5$ , 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 data and 13 TeV data/simulation, where available. The identification criteria, which include requirements for low hadronic activity ( $H/E$ ), a shower shape ( $\sigma_{inj}$ ) consistent with a photon, and an associated pixel seed veto, are shown in Table 6.9 for events with a photon in the barrel or endcap.

The isolation requirements restrict the energy sum from particle flow (PF) candidates within a cone of  $\Delta R < 0.3$  around the momentum vector of the photon candidate. Specif-

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

ically, we require that the energy from charged hadrons ( $\text{Iso}_{\text{pfCh.}}$ ), neutral hadrons ( $\text{Iso}_{\text{pfNu.}}$ ), and electromagnetic particles  $\text{Iso}_{\text{pfGa.}}$  not exceed the  $p_{\text{T}}$ -dependent thresholds shown 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 ( $\rho$ ) and per-photon  $\eta$ -dependent effective areas shown in Table 6.10.

Table 6.9: Photon identification and isolation requirements for events with photons in the barrel and endcap.

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

Table 6.10: Effective areas used in pileup correction as a function of photon pseudorapidity.

$ \eta $	pfCh.	pfNu.	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

$N_{\gamma}^{\text{obs}}$  and photon categories

The observed 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; non-prompt photons come from the decay of a meson. Direct prompt photons are radiated with relatively large transverse momentum relative to the quark; fragmentation prompt photons are radiated with low transverse momentum relative to the quark.

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

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

In practice, we categorize photons in the simulation as

- prompt photons: `status==1` photons with `abs(motherID)<= 22` (either a photon or a quark) or `motherID == 2212` (the incoming proton)
  - direct prompt photons: prompt photons with  $\Delta R > 0.4$  with respect to any status=23 quarks or gluons
  - fragmentation prompt photons: all prompt photons that are not direct
- non-prompt photons: all photons that are not prompt

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

The numbers of observed photons in ECAL barrel (EB) and endcap (EE) are shown in Table 6. There is no systematic uncertainty on  $N_\gamma^{\text{obs}}$  because it is a simple observation. The statistical uncertainty related to this observation is included in the statistical analysis when a Poisson PDF is assigned to this variable.

## Purity

The purity  $\beta$  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  $p_T$ ,

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

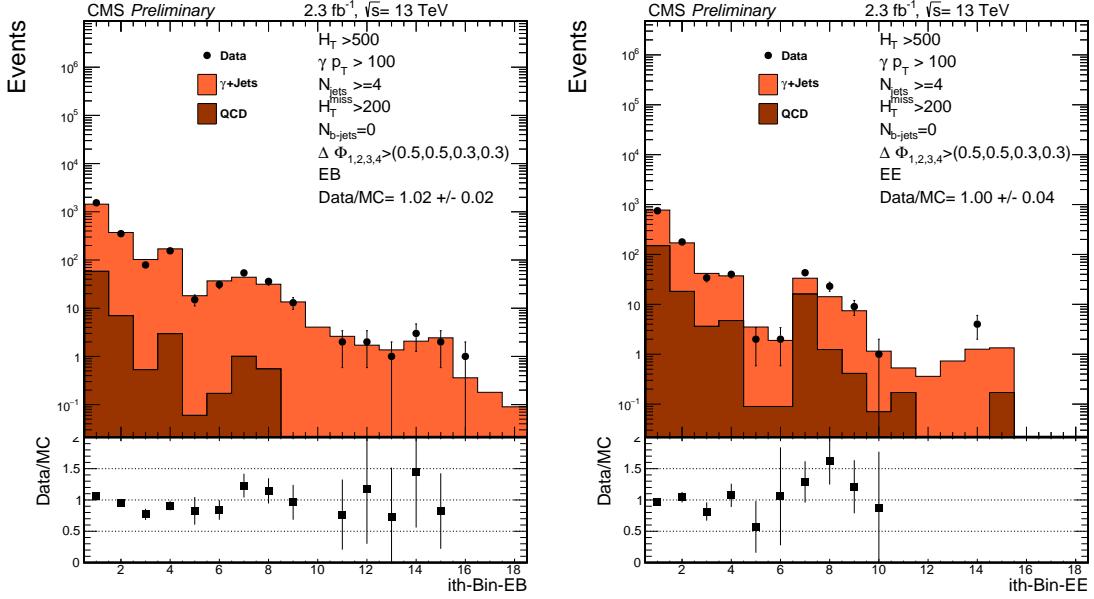


Figure 6.5: Numbers of observed events in the photon control samples in EB (left) and EE (right) compared to simulation.

which is essentially equivalent to event  $H_T^{\text{miss}}$ . We also explore differences in purity for ECAL barrel and endcap, but find that the EB and EE purities agree.

Prompt photons can be distinguished from non-prompt photons by the shapes of their showers in the ECAL, as described by the well known  $\sigma_{i\eta i\eta}$  quantity. The purity is determined with a two-component fit to the  $\sigma_{i\eta i\eta}$  distribution in the photon control sample. The PDF for the prompt component is fit directly in data using a gaussian distribution, which was motivated by  $\gamma$ +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  $H_T^{\text{miss}}$  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. 6.6.

To obtain the systematic uncertainty on the 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. As described in the photon control sample is defined with

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

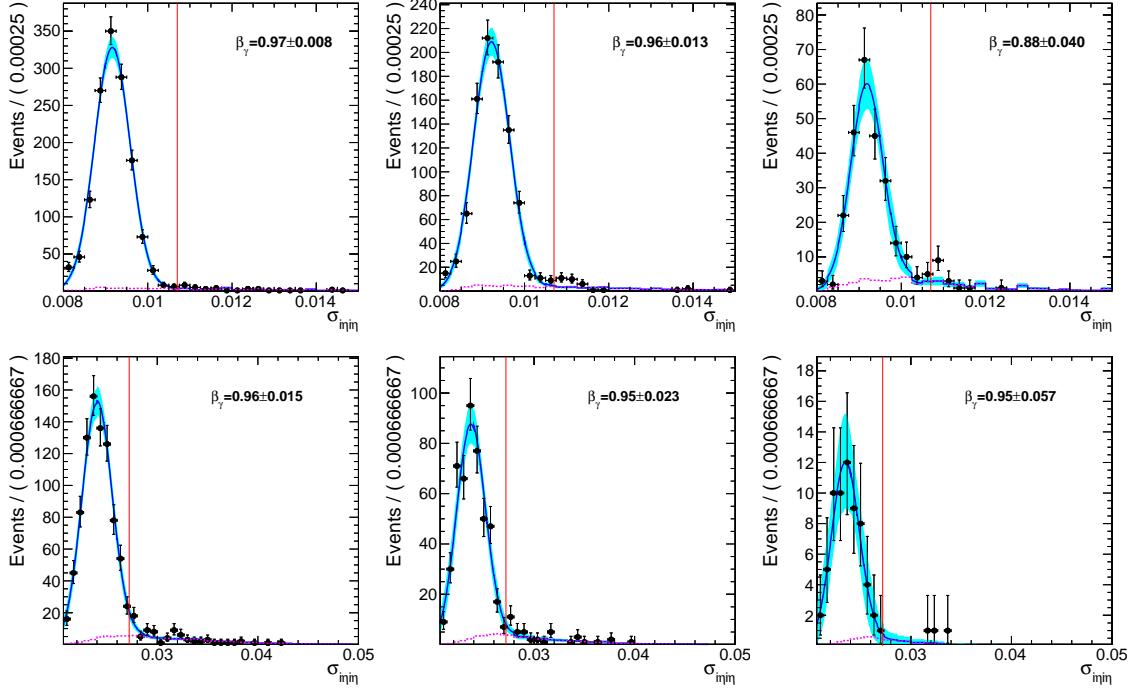


Figure 6.6: Fit results for purity determination in EB (top) and EE (bottom) for events with  $H_T^{\text{miss}}$  in the range 200-300, 300-500, and for  $H_T^{\text{miss}} > 500$ , respectively.  $\beta_\gamma$  is the fitted purity within the corresponding region. The red vertical line shows the  $\sigma_{\text{inj}\eta\eta}$  requirement for the loose ID working point. The non-prompt PDF comes from the QCD simulation.

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  $\sigma_{\text{inj}\eta\eta}$  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 adopt the quadrature sum of the fit result differences and the statistical uncertainty on the nominal fit as the systematic uncertainty on the purity; the systematic uncertainty is uncorrelated across different MHT 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 distribu-

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

tions of  $Z$  and  $\gamma$  events, we apply a scale factor  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  which is the ratio of the number of  $Z \rightarrow \nu\nu + \text{jets}$  and  $\gamma + \text{jets}$  events computed with leading order generator-level simulation (MADGRAPH+PYTHIA) in each analysis bin. The numerator and denominator are both determined after full selection with reco-level (rather than gen-level) physics objects. The Fig. 6.7 shows  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  in the 18 0-btagged bins.

Because  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  is determined at reco-level after full selection, we correct  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  for 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  $\eta$ , , and  $\Delta R(\gamma, \text{nearestjet})$ . Preliminary results show that the SF are independent of the analysis variables, so we correct the efficiency by the single value of  $0.98 \pm 0.05$  when computing  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$ , and we vary the SF by its 5% uncertainty to determine the systematic uncertainty.

Two different kinds of uncertainties are shown in Fig. 6.7. The statistical ones from MonteCarlo statistics and the systematic ones are obtained by propagating the ID/Iso SF to the  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$ .

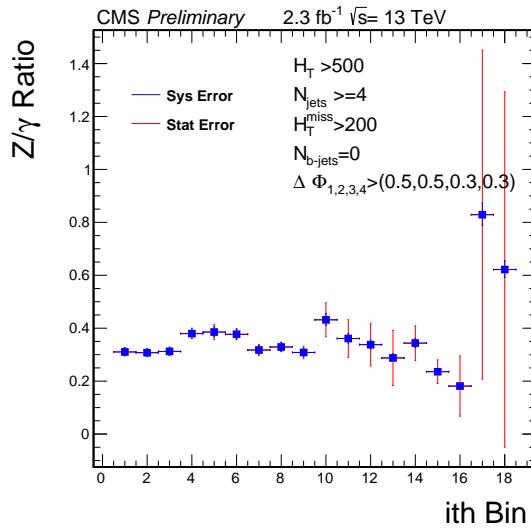


Figure 6.7:  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  for the 18 analysis bins used in the analysis.

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

As introduced above, fragmentation photons must be treated carefully when computing  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  primarily because the **MADGRAPH**  $\gamma+$ jets events that make up the  $\mathcal{R}_{Z(\bar{v}\bar{v})/\gamma}$  denominator are generated with the requirement that  $\Delta R(\gamma, \text{quarks}) > 0.4$  and  $\Delta R(\gamma, \text{gluons}) > 0.4$ . Because these fragmentation prompt photons are included in the photon control sample and are experimentally indistinguishable from direct prompt photons (unlike non-prompt photons), we must correct for this missing component of the simulation. We use the prompt photons included in QCD simulation sample to determine the fraction with  $\Delta R > 0.4$ . The value of fragmentation fraction is found to be  $0.92 \pm 0.08$ .

### 6.1.4 Results and Statistical interpretation

This section describes the interpretation of the observed data in the signal region vs. all the standard model background. The Fig. 6.8 shows the data vs. all the standard model 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  $n$ , where  $n$  depends on the floating parameters in the likelihood. The test statistic used in the profile likelihood calculations is given by

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

where the  $\mathcal{L}_\mu$  is the maximum likelihood at a given signal strength  $\mu$  and  $\mathcal{L}_{max}$  is the maximum floating all parameters. The profile likelihood signal significance is given by  $\sqrt{q_0}$  as a number of standard deviations. The likelihood ratio is profiled using the `combine` tool and is cross-checked with a stand alone signal strength scan in `RooStats`. The likelihood profiling in both algorithms shows good agreement and validates the significance computation done using the `combine` tool.

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

For setting limits, we use the LHC-style  $CL_s$  approach in the Higgs Combine tool. The CLs ratio is the ratio of confidence intervals:

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

where  $CL_b$  is the confidence interval in the background only region and  $CL_{s+b}$  is the confidence interval in the background region with signal. The 95% confidence level is given by the probability that a test statistic  $Q$  is less than the observed value in data:  $P_{95\%}(Q < Q_{obs})$ , so gives a probability of the data being discrepant with the background only hypothesis. The implementation of the fit technique in the Higgs Combination tool is tested against a standalone implementation in RooStats for a single bin, and found to demonstrate good agreement. The test statistic  $Q$  comes from a modified Profile Likelihood with constraints on the upper bound 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% confidence level upper exclusion limits will decrease with increasing sensitivity to signal.

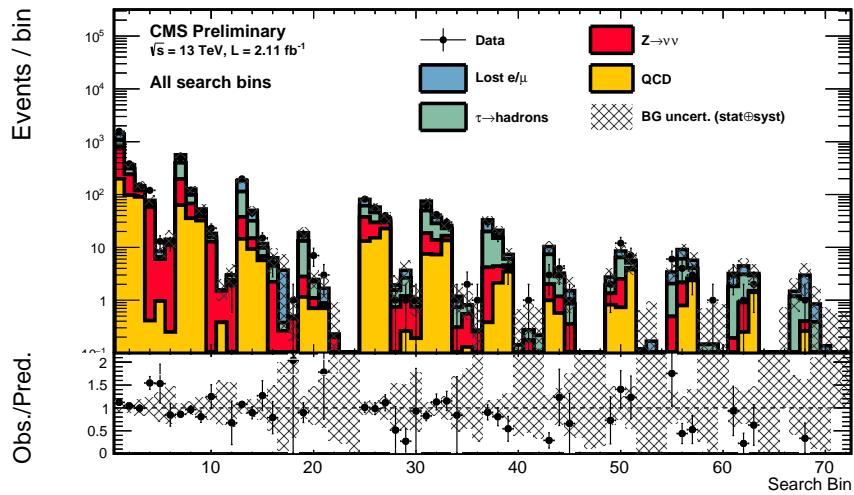


Figure 6.8: The observed data vs the standard model background in the 72 exclusive intervals of signal region

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

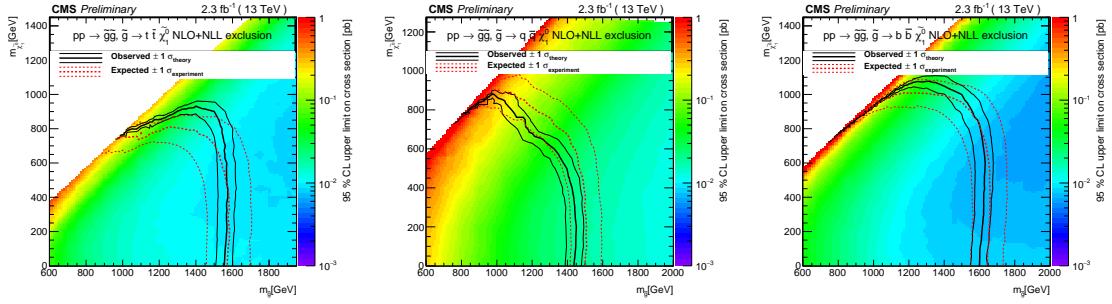


Figure 6.9: The 95% CL upper limits on the production cross sections for four top quark(left),four light quark(middle) and four b-quark(right) in the final state.

## 6.2 Study with $12.9 \text{ fb}^{-1}$ of 2016 data

With collection of more data in 2016 we update our analysis strategy to gain more sensitivity. We approach the search for new physics in 2016 data with a strategy similar to that used in the 2015 analysis. The three most significant changes to the strategy are

- Expanding the search region to include events with 3 or more jets. The 2015 analysis only considered events with 4 or more jets. This more inclusive selection allows us to target more models of direct squark production and other low-multiplicity signatures.
- Adopting triggers that make online requirements only on  $H_T^{\text{miss}}$  and MET, either  $HLT - PFMET(NoMu)100(120) - PFMHT(NoMu)100(120) - IDTight$ . These triggers allow us to lower our minimum offline  $H_T$  threshold from 500 to 300 GeV, which gives us increased sensitivity to SUSY models with low jet multiplicity or compressed spectra.
- Using finer binning in  $H_T$ ,  $H_T^{\text{miss}}$ , and  $N_{\text{jet}}$  to take advantage of a much larger dataset in 2016.

Other than this there are several less important changes in each background estimation methods. We will discuss in detail about the changes that occurred in  $Z(\rightarrow v\bar{v}) + jets$ .

## 6 Search for Supersymmetry with 13 TeV $pp$ collision data

### 6.2.1 Search binning

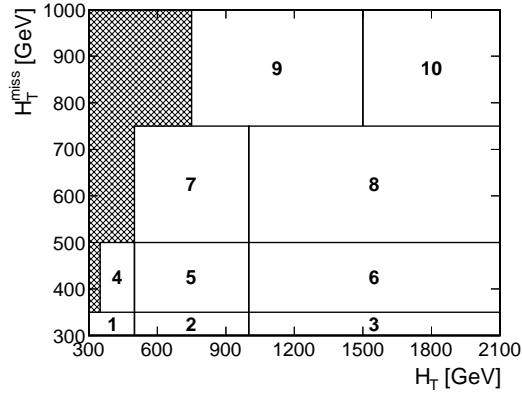


Figure 6.10: Search binning in 2D phase space of  $H_T$  and  $H_T^{\text{miss}}$  in 2016 analysis.

### 6.2.2 $Z(\rightarrow \nu\bar{\nu}) + \text{jets}$ background estimation

### 6.2.3 Results and Statistical Interpretation

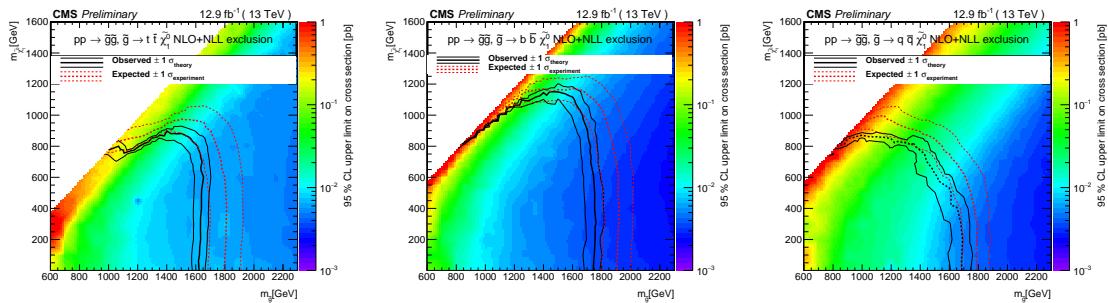


Figure 6.11: The 95% CL upper limits on the production cross sections for four top quark(left),four light quark(middle) and four b-quark(right) in the final state.

# 7 Search for Pair Production of First Generation Leptoquarks

Leptoquark pair-production at the Large Hadron Collider (LHC) at CERN proceeds primarily through gluon-gluon ( $gg$ ) fusion, with a smaller contribution from quark-antiquark ( $q\bar{q}$ ) annihilation. Here, the leptoquark mass ( $M_{LQ}$ ) and the Yukawa coupling  $\lambda$  to leptons and quarks are relevant production and decay parameters. However, the production cross section is nearly independent of  $\lambda$ .

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 quark has branching fraction  $\beta$ , while the decay into a neutrino and a quark has branching fraction  $1 - \beta$ , where  $\beta$  is a free parameter. Previous searches for rare processes such as lepton number violation, proton decay, and flavor-changing neutral currents prefer leptoquarks at LHC-accessible mass ranges which couple to leptons and quarks of same generation. Thus in this search for first generation leptoquarks, we assume leptoquarks decay to electrons and electron neutrinos, and they can do so in three final states:

- Two electrons and two jets: each leptoquark decays into an electron and a quark
- One electron, missing transverse energy, and two jets: one leptoquark decays into an

## 7 Search for Pair Production of First Generation Leptoquarks

electron and a quark, while the other decays into a neutrino and a quark

- No electrons, missing transverse energy, and two jets: each leptoquark decays into a neutrino and a quark

The final state branching fractions are  $\beta^2$ ,  $2 \times \beta(1 - \beta)$ , and  $(1 - \beta)^2$ , respectively. In general, we consider two values of  $\beta$ : 1, which corresponds to 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  $evjj$ . In this analysis, we search for leptoquark decays in the first final state only,  $eejj$ .

Previous searches have not detected any leptoquarks. CMS recently published a search using about  $20 \text{ fb}^{-1}$  of 2012 LHC data 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  $\beta = 0.2$ , which gave a mass limit of 900 GeV. All of the above limits are set at the 95% confidence level.

### 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 *beta* as defined previously is equal to 1. Events containing at least two electrons and two jets are selected, where the two  $p_T$ -leading 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 a background-dominated event sample used to compare Monte Carlo simulated background with the data. The event selection criteria are as follows

- Passes the signal trigger HLT\_Ele27\_WP Loose\_Gsf\_v\*
- Contains exactly two electrons passing HEEP ID, additionally having ECAL supercluster  $E_T > 50$  GeV and ECAL supercluster eta  $|\eta_{SC}| < 2.5$
- 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  $|\eta| < 2.4$
- $m_{ee} > 50$  GeV, where  $m_{ee}$  is the dielectron invariant mass
- $S_T > 300$  GeV, where  $S_T$  is the scalar sum of the  $p_T$  of the two electrons and leading two jets

After applying these criteria, the MC background agreement with the data is examined in a series of plots. Those shows a quite good agreement of the data with the SM backgrounds. Other than this cuts on the above last three variables are optimized for different LQ mass points.

### 7.1.0.2 final selection

Here are the final cuts on  $M_{LQ}$ ,  $m_{ee}$  and  $m_{ej}^{min}$ . The cuts are optimized for different LQ mass points.

The event selections are optimized for each leptoquark mass hypothesis in order to better separate signal and background events. Three variables,  $S_T$ ,  $m_{ee}$ , and  $m_{ej}$  are varied independently over a range of values, and the thresholds which maximize the optimization metric

## 7 Search for Pair Production of First Generation Leptoquarks

are obtained for each leptoquark mass hypothesis. One optimization metric considered is an asymptotic formula, which gives the approximate discovery significance in standard deviations

$$\sqrt{2((s+b)\ln(1+s/b)-s)} \quad (7.1)$$

Here, for a given set of selection thresholds,  $b$  represents the number of predicted background events, and  $s$  the number of predicted signal events. The formula is an improvement with respect to  $s/\sqrt{b}$  or  $s/\sqrt{s+b}$  which tend to overestimate the significance when  $b$  is low. However, it was found that using this method, the  $m_{ej}$  cuts are higher than expected, which leads to a significant efficiency loss. Another optimization criterion was therefore considered, via

$$\frac{\epsilon(t)}{a/2 + \sqrt{B(t)}} \quad (7.2)$$

## 7.2 SM Backgrounds

The major backgrounds from Standard Model processes are  $Z + jets$  and  $t\bar{t}$ , with single top,  $W + jets$ , diboson, and  $\gamma+jets$  contributing at a lower level. There is also an instrumental background from QCD events with jets faking electrons and passing the HEEP ID. Here is how the backgrounds are determined in this analysis:

## 7.3 Results and statistical interpretation

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 [21] has dissapeared with the 2015 data. We set the upper limits on the cross section times branching fraction using the same tool as before. Fig. 7.2 shows both the observed and expected

## 7 Search for Pair Production of First Generation Leptoquarks

Table 7.1: 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
$\geq 1500$	1815	305	895

## 7 Search for Pair Production of First Generation Leptoquarks

limits for different leptoquark masses. We exclude leptoquark masses up to 1130 GeV from this study.

The data vs. backgrounds plots for two variables are shown in the following figure.

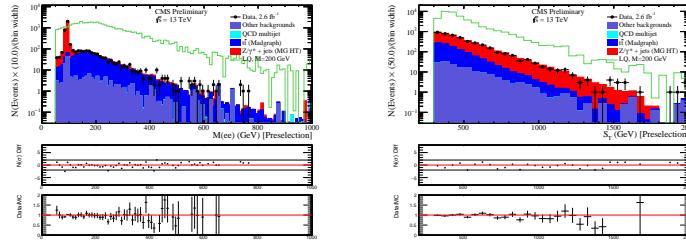


Figure 7.1: Data vs. background plots for  $M_{ee}$ (left) and  $S_T$ (right)

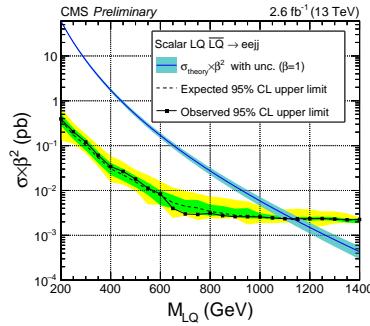


Figure 7.2: The 95% CL upper limits on the production cross sections as a function of leptoquark mass.

# 8 Advanced Pilup Mitigation Techniques

Grooming has been introduced in [?] and is intended to remove soft and wide-angle radiation from the 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 bosons. Additionally, grooming, in the presence of pileup, can be used to reduce the dependence on pileup of jet mass. 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-R jets ( $R=0.8,1,2$ ). We consider the grooming methods: pruning [23], trimming [25], modified mass drop tagger [24] and soft drop [26].

The pruning algorithm reclusters the constituents of the jet through the Cambridge-Aachen (CA) algorithm [?], 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:

## 8 Advanced Pilup Mitigation Techniques

$$z_{ij} = \frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}} < z_{\text{cut}} \quad (8.1)$$

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

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

Trimming ignores particles within a jet that fall below a dynamic threshold in  $p_T$ . It reclusters the constituents of the jet using the kT algorithm with a radius  $R_{\text{sub}}$ , accepting only the subjets that have  $p_{T_{\text{sub}}} > f_{\text{cut}} \lambda_{\text{hard}}$ , where  $f_{\text{cut}}$  is a dimensionless cutoff parameter, and  $\lambda_{\text{hard}}$  is some hard QCD scale chosen to equal the pT of the original jet.

Soft-drop and modified mass drop tagger decluster the jet recursively 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 defined and the following condition is defined:

where the algorithm parameters are  $z_{\text{cut}}$  and  $\beta$ . 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 [? ]. In the cases of soft drop, modified mass drop tagger, and trimming, the four-vector subtraction corrects the jet pT 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 summarized in Table 8.1.

## 8 Advanced Pilup Mitigation Techniques

grooming algorithm	parameter(s)
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	$r_{\text{sub}} = 0.2, pT_{\text{frac}} = 0.05$
	$r_{\text{sub}} = 0.2, pT_{\text{frac}} = 0.03$
	$r_{\text{sub}} = 0.1, pT_{\text{frac}} = 0.03$
	$r_{\text{sub}} = 0.3, pT_{\text{frac}} = 0.03$
Soft drop/MMDT	$z_{\text{cut}} = 0.1, \beta = -1$
	$z_{\text{cut}} = 0.1, \beta = 0$
	$z_{\text{cut}} = 0.1, \beta = 1$
	$z_{\text{cut}} = 0.1, \beta = 2$

Table 8.1: Summary of grooming parameters

### 8.0.1 Simulation-only results at high pileup

The performance of various groomers is evaluated at higher pileup scenarios which expected in LHC Run II. For a given  $p_T$  bin, we evaluate the jet mass resolution for both background QCD jets and signal W jets. In the signal case, the jet mass resolution is evaluated after matching the jet to the W boson direction at the particle level. In the background case, a  $\Delta R$  matching is done between the particle level jet and the jet after full detector reconstruction where  $\Delta R < 0.3$ . In each comparison, the *groomed* particle level jet is compared against its corresponding matched detector-level reconstructed jet.

We first evaluate the performance of groomers based on particle flow input collection with and without CHS to understand the effect. We show the average jet mass for PF and CHS background QCD jets in Fig. 8.3. The left column shows the average jet mass as a function of pileup for jets using particle flow inputs. The right column shows the average jet mass as a function of the pileup for jets using particle flow inputs including CHS. We see generally that the CHS jet masses are more stable against pileup. The trimming algorithm is generally stable against pileup regardless of input while the pruning algorithm shows the most pileup

## 8 Advanced Pilup Mitigation Techniques

dependence, particularly using particle flow inputs. The soft drop algorithm has a mild pileup dependence.

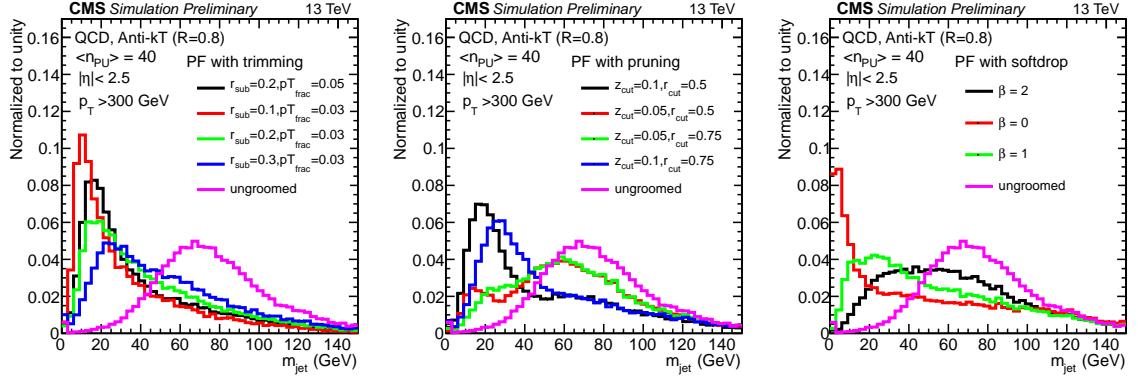


Figure 8.1: Jet mass distribution for PF QCD jets for different groming parameters. The PF jets are safe subtracted.

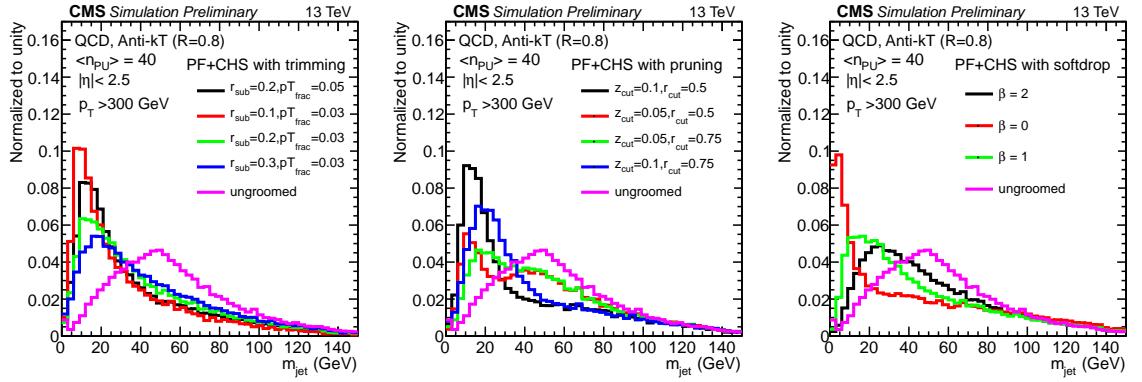


Figure 8.2: Jet mass distribution for PF+CHS QCD jets for different groming parameters. The PFCHS jets are safe subtracted.

## 8 Advanced Pilup Mitigation Techniques

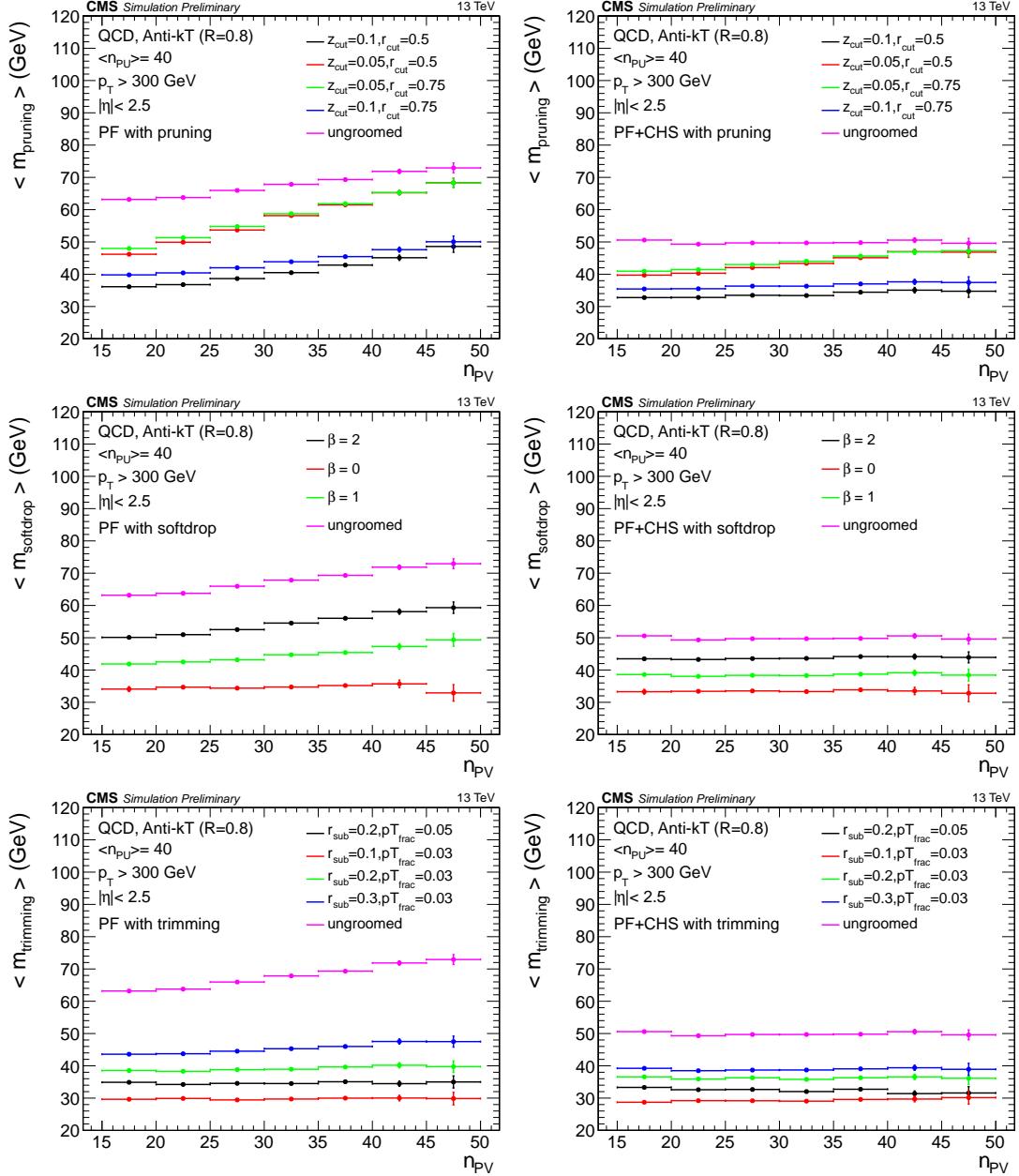


Figure 8.3: Pileup dependence of the average jet mass for PF jets (left) and PFCHS jets (right) for several grooming algorithms and parameters. Both PF and PFCHS jets are safe subtracted.

## 8 Advanced Pilup Mitigation Techniques

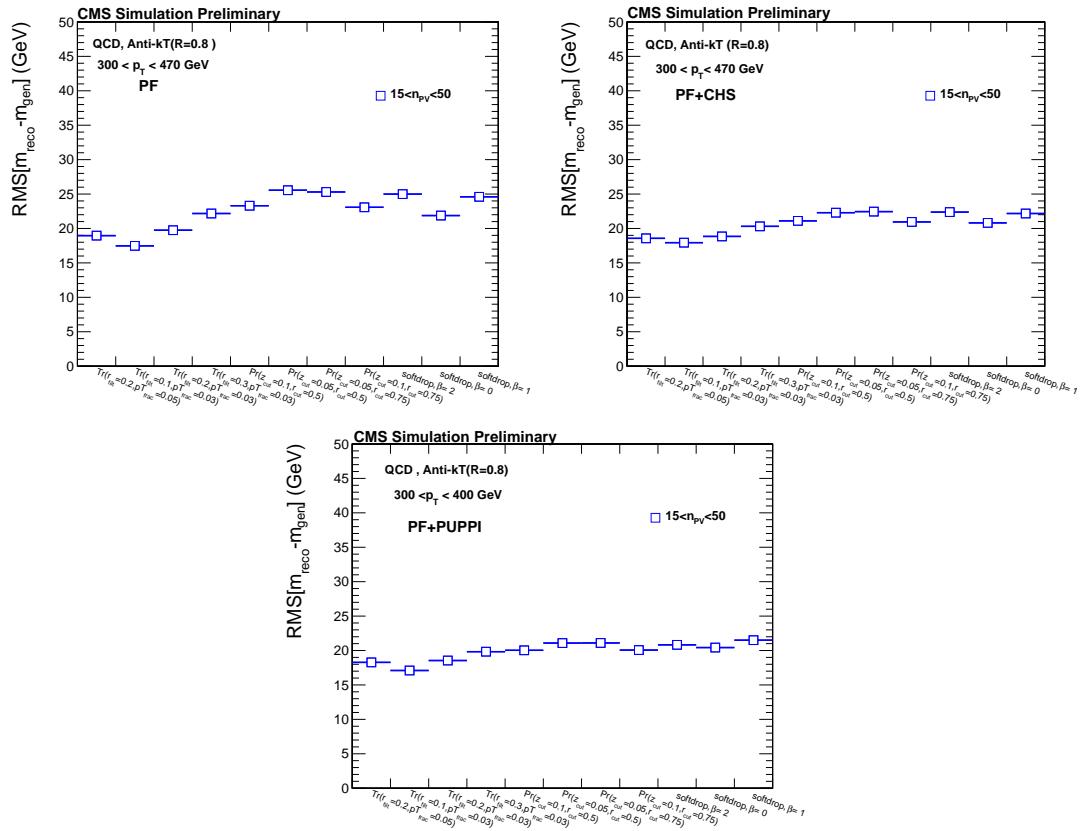


Figure 8.4: Comparison of jet mass resolution for PF, CHS and Puppi with different grooming variables

## **9 Summary and Conclusion**

## *9 Summary and Conclusion*

# References

- [1] H. Nilles, “Supersymmetry, supergravity and particle physics,” *Phys. Rep.* **110** (1984) 1.
- [2] 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.
- [3] G. G. S. Dimopoulos, “Naturalness constraints in supersymmetric theories with nonuniversal soft terms,” *Phys. Lett. B* **357** (1995) 573.
- [4] P. S. Alwall and N. Toro, “Simplified models for a first characterization of new physics at the LHC,” *Phys. Rev. D* **79** (2009) 075020.
- [5] B. Schrempp and F. Schrempp, “Light Leptoquark,” *Phys. Lett. B* **153** (1985) 101.
- [6] S. Dimopoulos and L. Susskind, “Mass Without Scalar,” *Nucl. Phys. B* **155** (1979) 237.
- [7] J. C. Pati and A. Salam, “Lepton Number as the Fourth Color,” *Phys. Rev. D* **10** (1974) 275.
- [8] W. Buchmuller and D. Wyler, “Constraints on SU(5) Type Leptoquarks,” *Phys. Lett. B* **177** (1986) 377–382.
- [9] H. Georgi and S. Glashow. Unity of all elementary-particle forces. *Phys. Rev. Lett.*, 32:438, 1974.
- [10] B. Schrempp and F. Schrempp. Light leptoquarks. *Phys. Lett. B*, 153:101, 1985.
- [11] S. Dimopoulos. Technicolored signatures. *Nucl. Phys. B*, 168:69, 1980.

## References

- [12] E. Eichten and K. Lane. Dynamical breaking of the weak interaction symmetries. *Phys. Lett. B*, 90:85, 1980.
- [13] J. L. Hewett and T. G. Rizzo. Low-energy phenomenology of superstring-inspired  $E_6$  models. *Phys. Lett.*, 183:193, 1989.
- [14] **CMS** Collaboration, S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,” *JINST* **3** (2008) S08004.
- [15] **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://arxiv.org/abs/1401.8155). <https://inspirehep.net/record/1279774/files/arXiv:1401.8155.pdf>.
- [16] 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).
- [17] **The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group** Collaboration, “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>.
- [18] **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>.
- [19] **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>.
- [20] **CMS** Collaboration, “Search for pair-production of first generation scalar leptoquarks in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with  $2.6 \text{ fb}^{-1}$ ,” Tech. Rep. CMS-PAS-EXO-16-043, CERN, Geneva, 2016. <https://cds.cern.ch/record/2205285>.
- [21] **CMS** Collaboration, “Search for Pair-production of First Generation Scalar Leptoquarks in  $pp$  Collisions at  $\text{sqrt } s = 8$  TeV,” Tech. Rep. CMS-PAS-EXO-12-041, CERN, Geneva, 2014. <https://cds.cern.ch/record/1742179>.

## References

- [22] 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.
- [23] 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.
- [24] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, “Towards an understanding of jet substructure,” *JHEP* **09** (2013) 029.
- [25] D. Krohn, J. Thaler, and L.-T. Wang, “Jet Trimming,” *JHEP* **02** (2010) 084.
- [26] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop,” *JHEP* **05** (2014) 146.
- [27] 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>.
- [28] Antonio Pich. The Standard model of electroweak interactions. In *High-energy physics. Proceedings, European School, Aronsborg, Sweden, June 18-July 1, 2006*, pages 1–49, 2007. [,1(2007)].
- [29] F. Halzen and A. D. Martin, “Quarks and Leptons: An Introductory Course in Modern Particle Physics”.
- [30] M. J. Herrero, “The Standard Model”, <http://arxiv.org/abs/hep-ph/9812242>.
- [31] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Physics Letters B* **716** (2012) 129.
- [32] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Physics Letters B* **716** (2012) 3061.
- [33] Wikipedia “Standard Model” [https://en.wikipedia.org/wiki/Standard\\_Model](https://en.wikipedia.org/wiki/Standard_Model).
- [34] Antonio Pich. Quantum chromodynamics. (hep-ph/9505231. FTUV-95-19. IFIC-95-19):51 p, May 1995.
- [35] David J. Gross and Frank Wilczek, “Ultraviolet Behavior of Non-Abelian Gauge Theories”, *Phys. Rev. Lett.* **30** (1973) 1343.

## References

- [36] H. David Politzer, “Reliable Perturbative Results for Strong Interactions?”, Phys. Rev. Lett. 30 (1973) 1346.
- [37] J. Goldstone, A. Salam, and S. Weinberg, Phys. Rev. 127, 965 (1962).
- [38] P. W. Higgs, “Broken symmetries and the masses of gauge bosons”, Phys. Rev. Lett. 13 (1964) 508 ; F. Englert and R. Brout, “Broken symmetries and the mass of gauge vector mesons”, Phys. Rev. Lett. 13 (1964) 321.
- [39] “The LEP Collaborations ALEPH, DELPHI, L3 and OPAL and the LEP Electroweak Working Group”, <http://arxiv.org/abs/hep-ex/0612034>.
- [40] “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.
- [41] (ATLAS, CDF, CMS, D0 Collaborations), P. A. R. Adeet al. ‘First combination of Tevatron and LHC measurements of the top-quark mass’, <https://arxiv.org/abs/1403.4427>.
- [42] (Planck Collaboration) ‘Planck 2013 results. XVI Cosmological parameters’, <https://arxiv.org/abs/1303.5076>.
- [43] A. Pomarol, “Beyond the Standard Model”, CERN Yellow Report CERN-2012-001, pp. 115-151, <http://arxiv.org/abs/1202.1391>.
- [44] Stephen P. Martin. A Supersymmetry primer. 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [45] Csaba Csaki. The Minimal supersymmetric standard model (MSSM). *Mod. Phys. Lett.*, A11:599, 1996.
- [46] Pierre Fayet. Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino. *Nucl. Phys.*, B90:104–124, 1975.
- [47] T. Appelquist, H. C. Cheng, B. A. Dobrescu, “Bounds on Universal Extra Dimensions”, Phys. Rev. D64 035002 (2001).
- [48] Guido Altarelli, “Collider Physics within the Standard Model: a Primer”, <http://arxiv.org/abs/1303.2842>.

## References

- [49] G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B122 (1983) 103.
- [50] G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B126 (1983) 398.
- [51] M. Banner *et al.* (UA2 Collaboration), Phys. Lett. B122 (1983) 476.
- [52] P. Bagnaia *et al.* (UA2 Collaboration), Phys. Lett. B129 (1983) 130.
- [53] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. 74 (1995) 2626.
- [54] S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. 74 (1995) 2632.
- [55] J. Beringer *et al.* (Particle Data Group), “REVIEW OF PARTICLE PHYSICS”, Phys. Rev. D86,010001 (2012).
- [56] G. Abbiendi, *et al.* (the ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration and the OPAL Collaboration, The LEP Working Group for Higgs Boson Searches) ”Search for the Standard Model Higgs Boson at LEP” <http://arxiv.org/abs/hep-ex/0306033v1>.
- [57] “Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to  $10.0 \text{ fb}^{-1}$  of Data”. <http://arxiv.org/abs/1203.3774>.
- [58] “The LEP Electroweak Working Group”, <http://lepewwg.web.cern.ch/LEPEWWG/>.
- [59] S. Dittmaier *et. al.* “Handbook of LHC Higgs Cross Sections: 2. Differential Distributions”, <http://arxiv.org/abs/arXiv:1201.3084>.
- [60] “LHC Higgs Cross Section Working Group (2012-2013)”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>.
- [61] J. Wess and J. Bagger, “Supersymmetry and Supergravity”, Princeton Series in Physics.
- [62] Nima Arkani-Hamed, Savas Dimopoulos, Gia Dvali, “The Hierarchy Problem and New Dimensions at a Millimeter”, Phys. Lett. B 429:263-272, 1998, <http://arxiv.org/abs/hep-ph/9803315>.
- [63] W. de Boer “Grand Unified Theories and Supersymmetry in Particle Physics and Cosmology” <http://arxiv.org/abs/hep-ph/9402266>.
- [64] L. Randall, R. Sundrum, “A Large Mass Hierarchy from a Small Extra Dimension”, Phys. Rev. Lett. 83, 3370, <http://arxiv.org/abs/hep-ph/9905221>.

## References

- [65] “The LHC Design Report v.1”, CERN-2004-003-V-1 <http://cdsweb.cern.ch/record/782076>.
- [66] ALICE Collaboration, “Technical Proposal for a Large Ion Collider Experiment at the CERN LHC”, CERN-LHCC-95-71 (1995).
- [67] ATLAS Collaboration, “Detector and Physics Performance Technical Design Report, vol. 1”, CERN-LHCC-99-14 (1999).
- [68] R. Adolphi *et al.* (CMS Collaboration), “The CMS experiment at the CERN LHC”, JINST 0803:S08004, (2008).
- [69] LHCb Collaboration, “LHCb Technical Proposal”, CERN-LHCC-98-04 (1998).
- [70] CMS collaboration, “Physics Technical Design Report”, Volume I : Detector Performance and software, CMS/LHCC 2006-001, CMS TDR 8.1.
- [71] CMS collaboration, “Physics Technical Design Report”, Volume II : Physics performance, CMS/LHCC 2006-021, CMS TDR 8.2.
- [72] C. Seez, “The CMS trigger system”, Eur Phys J C 34, s01, s151s159 (2004).
- [73] The CMS Offline Workbook. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBook>.
- [74] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting”, Nucl. Instrum. Meth. A262 (1987) 444-450.
- [75] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, CMS-TRK-11-001, <http://arxiv.org/abs/1405.6569>.
- [76] E. Chabanat *et al.* Deterministic Annealing for Vertex Finding at CMS. <https://cds.cern.ch/record/865587/files/p287.pdf> .
- [77] A. Giiammanco, ”Particle Identification with Energy Loss in the CMS Silicon Strip Tracker, CMS AN 2007/008, (2007).
- [78] CMS Collaboration, “Performance of muon identification in pp collisions at  $\sqrt{s} = 7$  TeV”. CMS-PAS-MUO-10-002.

## References

- [79] CMS Physics, Technical Design Report, Detector Performance and Software, Vol-1, CMS/LHCC 2006-001, CMS TDR 8.1.
- [80] G. Abbiendi *et al.*, CMS-AN-2008-097 (2008).
- [81] R. Fruhwirth, “Track fitting with non-Gaussian noise”, Computer Physics Communications 100 (1997) 116.
- [82] W. Adam *et. al.*, “Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc, CMS NOTE-2005/001 (2005).
- [83] “Electron performance with  $19.6 \text{ fb}^{-1}$  of data collected at  $\sqrt{s} = 8 \text{ TeV}$  with the CMS detector”, CMS-DP-2013-003 ; CERN-CMS-DP-2013-003 (2013).
- [84] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, Phys.Lett. B 659 (2008). 119126, <http://arxiv.org/abs/0707.1378>.
- [85] S.D. Ellis *et. al.* “Jets in hadronhadron collisions”, Progress in Particle and Nuclear Physics 60 (2008) 484551.
- [86] CMS Collaboration, “Performance of Jet Algorithms in CMS”, CMS-PAS-JME-07-003.
- [87] V. Chetluru *et. al.* CMS-AN-2009/067 (2009).
- [88] CMS Collaboration, “Performance of Jet Reconstruction with Charged Tracks only”, CMS-PAS-JME-08-001 (2008).
- [89] CMS Collaboration, “Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS”, CMS-PAS-JME-09-002 (2009).
- [90] CMS Collaboration, “ParticleFlow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_T^{\text{miss}}$ ”, CMS-PAS-PFT-09-001 (2009).
- [91] CMS Collaboration, “Determination of the jet energy scale in CMS with pp collisions at  $\sqrt{s} = 7 \text{ TeV}$ ”, CMS-PAS-JME-10-010 (2010).
- [92] B. Parida *et. al.* “Jet energy resolution with  $Z \rightarrow (\mu^+ \mu^-) + \text{jet}$  events at  $\sqrt{s} = 7 \text{ TeV}$ ” CMS-AN-12-366 (2012).
- [93] 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>.

## References

- [94] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, JINST 8 (2013) P04013, <http://arxiv.org/abs/1211.4462>.
- [95] CMS b-tag POG twiki, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>.
- [96] CMS Collaboration, “Missing transverse energy performance of the CMS detector”, JINST 6 (2011) 09001, <http://arxiv.org/abs/1106.5048>.
- [97] S. Cataniand, Y. L. Dokshitzerand, M. H. Seymour, and B. R. Webber, “Longitudinally invariant K-t clustering algorithms for hadron hadron collisions,” Nucl. Phys. B 406:187 (1993).
- [98] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, Phys. Rev. D 48 3160 (1993).
- [99] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, “Better jet clustering algorithms”, JHEP 9708 (1997) 001, <http://arxiv.org/abs/hep-ph/9707323>.
- [100] G. S. Matteo Cacciari and G. Soyez, “FastJet 2.3 User Manual”, Phys. Lett. B 641:57 (2006).
- [101] S.D. Ellis, C.K. Vermilion and J.R. Walsh, “Techniques for improved heavy particle searches with jet substructure”, Phys. Rev. D 80 (2009) 051501, <http://arxiv.org/abs/0903.5081>.
- [102] J. M. Butterworth *et al.*, “Jet substructure as a new Higgs-search channel at the LHC”, Phys. Rev. Lett. 100 (2008) 242001, <http://arxiv.org/abs/0802.2470>.
- [103] D. Krohn, J. Thaler, and L.-T. Wang, “Jet Trimming”, JHEP 02 (2010) 084, <http://arxiv.org/abs/0912.1342>.
- [104] 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. D 81 (2010) 094023, <http://arxiv.org/abs/0912.0033>.
- [105] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop”, JHEP 1405 (2014) 146, <http://arxiv.org/abs/1402.2657>.

## References

- [106] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, “Towards an understanding of jet substructure”, JHEP 1309 (2013) 029, <http://arxiv.org/abs/1307.0007>.
- [107] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, “Jet substructure as a new Higgs search channel at the LHC”, <http://arxiv.org/abs/0810.0409>.
- [108] CMS Collaboration, “Studies of jet mass in dijet and W/Z + jet events”, CMS-PAS-SMP-12-019 (2012).
- [109] Y. Cui, Z. Han, and M. D. Schwartz, “W-jet Tagging: Optimizing the Identification of Boosted Hadronically-Decaying W Bosons”, Phys.Rev. D83 (2011), <http://arxiv.org/abs/1012.2077>.
- [110] J. Thaler and K. Van Tilburg, “Identifying Boosted Objects with N-subjettiness”, JHEP 1103 (2011) 015, <http://arxiv.org/abs/1011.2268>.
- [111] S. D. Ellis *et al.*, “Qjets: A Non-Deterministic Approach to Tree-Based Jet Substructure”, Phys.Rev.Lett. 108 (2012) 182003, <http://arxiv.org/abs/1201.1914>.
- [112] B. Parida *et al.* “Semi-leptonic decay of HWW at high mass in exclusive jet bins”, CMS-AN-13-414 (2014).
- [113] CMS Collaboration, “Search for a Standard Model Higgs-like boson decaying into  $WW \rightarrow \ell\nu q\bar{q}'$  in exclusive jet bins in pp collisions at  $\sqrt{s} = 8$  TeV”, CMS-PAS-HIG-14-008, (2014).
- [114] B. Parida *et. al.* “Searches for new physics in the  $WW \rightarrow \ell\nu j$  final state with merged W bosons”, CMS-AN-12-381, (2013).
- [115] CMS Collaboration, “Search for a Standard Model-like Higgs boson decaying into  $WW \rightarrow \ell\nu q\bar{q}'$  in pp collisions at  $\sqrt{s} = 8$  TeV”, CMS-PAS-HIG-13-008 (2013).
- [116] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, JHEP 11 (2007) 070, <http://arxiv.org/abs/0709.2092>.
- [117] T. Sjostrand, S. Mrenna, and P.Z. Skands, “PYTHIA 6.4 Physics and Manual”, JHEP, 05, 422 (2006).

## References

- [118] G. Corcella *et. al.* “HERWIG 6: an event generator for hadron emission reactions with interfering gluons” JHEP, 01, 010 (2001).
- [119] <https://twiki.cern.ch/twiki/bin/view/CMS/HiggsWWToInnujjCommonPATtuples>.
- [120] CMS Collaboration, “Electron Reconstruction and Selection Studies with First CMS 7 TeV Data”, CMS-PAS-EGM-10-004 (2010).
- [121] CMS Collaboration, “Performance of CMS muon identification in pp collisions at  $\sqrt{s} = 7$  TeV”, CMS-PAS-MUO-2010-002 (2010).
- [122] CMS Collaboration, “CMS MET Performance in Events Containing Electroweak Bosons from pp Collisions at  $\sqrt{s}= 7$  TeV”, CMS-PAS-JME-10-005.
- [123] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, CMS-PAS-PFT-2009-001 (2009).
- [124] 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).
- [125] CMS Collaboration, “Jet Performance in pp Collisions at  $\sqrt{s}= 7$  TeV”, CMS-PAS-JME-10-003, (2010).
- [126] CMS Collaboration, “Jet identification in high pile-up environment”, CMS twiki PileupJetID twiki (2012) <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJetID>.
- [127] A. Benaglia, L. Di Matteo, and P. Govoni, “Search for Higgs Boson Decays into Two W Bosons in the Semi-leptonic Final State with the Full 2011 pp Dataset at  $\sqrt{s} = 7$  TeV”, CMS-AN-12-029 (2012).
- [128] N. Akchurin *et al.*, “Computation of Offline and Trigger Efficiencies for Standard Model Higgs boson search in  $H \rightarrow WW \rightarrow \ell\nu jj$  decay”, CMS-AN-12-021 (2012).
- [129] CMS Collaboration, “Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at  $\sqrt{s}= 8$  TeV”, <http://arxiv.org/abs/1405.3447>.

## References

- [130] CMS Collaboration, “Studies of jet mass in dijet and W/Z + jet events”, JHEP 1305 (2013) 090, <http://arxiv.org/abs/1303.4811>.
- [131] CMS Collaboration, “Identifying Hadronically Decaying Vector Bosons Merged into a Single Jet”, CMS-PAS-JME-13-006 (2013).
- [132] CMS Collaboration, “b-Jet Identification in the CMS Experiment”, CMS-PAS-BTV-11-004, (2012).
- [133] CMS Collaboration, Measurement of  $W^+W^-$  and ZZ production cross sections in pp collisions at  $\sqrt{s} = 8$  TeV, Phys.Lett. B721 (2013) 190211, <http://arxiv.org/abs/1301.4698>.
- [134] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, <http://arxiv.org/abs/1307.1347>.
- [135] CLs “Procedure for the LHC Higgs boson search combination in summer 2011”, ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005 (2011).
- [136] A. L. Read, “Presentation of search results: the CLs technique”, Journal of Physics G: Nuclear and Particle Physics 28 (2002), no. 10, 2693.
- [137] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, CMS-PAS-LUM-13-001 (2013).
- [138] A. Ballestrero, G. Bevilacqua, D. B. Franzosi, and E. Maina, “How well can the LHC distinguish between the SM light Higgs scenario, a composite Higgs and the Higgsless case using VV scattering channels?”, JHEP 11 (2009) 126, <http://arxiv.org/abs/0909.3838>.
- [139] B. Parida *et. al.* “Study of Pileup Removal Algorithms for Jets”, CMS-AN-14-175 (2014).
- [140] CMS Collaboration, “Study of Pileup Removal Algorithms for Jets”, CMS-PAS-JME-14-001 (2014).
- [141] G. Soyez *et al.*, “Pileup subtraction for jet shapes”, Phys. Rev. Lett. 110 (2013), no. 16, 162001, <http://arxiv.org/abs/1211.2811>.

### *References*

- [142] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Techniques for improved heavy particle searches with jet substructure”, Phys. Rev. D 80 (2009) 051501, <http://arxiv.org/abs/0903.5081>.
- [143] M. Cacciari, G. P. Salam, and G. Soyez, “On the use of charged-track information to subtract neutral pileup”, <http://arxiv.org/abs/1404.7353>.
- [144] Peter Berta, Martin Spousta, David W. Miller, Rupert Leitner, “Particle-level pileup subtraction for jets and jet shapes” JHEP 1406 (2014) 092, <http://arxiv.org/abs/1403.3108>.
- [145] D. Krohn, M. Schwartz, M. Low, and L.-T. Wang, “Jet Cleansing: Pileup Removal at High Luminosity”, <http://arxiv.org/abs/1309.4777>.
- [146] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup Per Particle Identification”, <http://arxiv.org/abs/1407.6013>.
- [147] R. Fernow, Introduction to experimental particle physics.
- [148] D. Green, The Physics of Particle Detectors.

*References*