

1      **Using Machine Learning Techniques for Data Quality**  
2      **Monitoring at CMS Experiment**

3      by

4      Guillermo A. Fidalgo Rodríguez

5      A thesis presented for the degree of

6      BACHELLOR'S OF SCIENCE??

7      in

8      Physics

9      UNIVERSITY OF PUERTO RICO  
10     MAYAGÜEZ CAMPUS

11     2018

12     Approved by:

13     \_\_\_\_\_  
14     Sudhir Malik, Ph.D.

15     President, Graduate Committee  
Date

16     \_\_\_\_\_  
17     Héctor Méndez, Ph.D.

18     Member, Graduate Committee  
Date

19     \_\_\_\_\_  
20     Samuel Santana Colón, Ph.D.  
21     Member, Graduate Committee

Date

22     \_\_\_\_\_  
23     Rafael A. Ramos, Ph.D.  
24     Chairperson of the Department

Date

# <sup>25</sup> Abstract

<sup>26</sup> The Data Quality Monitoring (DQM) of CMS is a key asset to deliver high-quality  
<sup>27</sup> data for physics analysis and it is used both in the online and offline environment. The cur-  
<sup>28</sup> rent paradigm of the quality assessment is labor intensive and it is based on the scrutiny of  
<sup>29</sup> a large number of histograms by detector experts comparing them with a reference. This  
<sup>30</sup> project aims at applying recent progress in Machine Learning techniques to the automa-  
<sup>31</sup> tion of the DQM scrutiny. In particular the use of convolutional neural networks to spot  
<sup>32</sup> problems in the acquired data is presented with particular attention to semi-supervised  
<sup>33</sup> models (e.g. autoencoders) to define a classification strategy that doesn't assume previ-  
<sup>34</sup>ous knowledge of failure modes. Real data from the hadron calorimeter of CMS are used  
<sup>35</sup> to demonstrate the effectiveness of the proposed approach.

<sup>36</sup> **Acknowledgments**

<sup>37</sup> I wish to thank United States State Department and University of Michigan for pro-  
<sup>38</sup> viding the opportunity to work abroad at CERN during the 2018 Winter Semester. I also  
<sup>39</sup> wish to thank CERN staff, CMS Experiment , Texas Tech University, and the University  
<sup>40</sup> of Puerto Rico at Mayagüez, with special thanks to Dr. Federico de Guio for his local  
<sup>41</sup> mentorship and Dr. Nural Akchurin, and Dr. Sudhir Malik for their guidance. A very  
<sup>42</sup> special thanks to Dr. Jean Krisch for accepting me for this great research opportunity and  
<sup>43</sup> Dr. Steven Goldfarb for being a wonderful host and overall local guidance at CERN.

# <sup>44</sup> List of Figures

|    |  |    |
|----|--|----|
| 45 | Figure 2.1: Schematic representation of the Standard Model particles with<br>46 the added SUSY particles. . . . .  | 4  |
| 47 | Figure 3.1: The CERN Accelerator Complex [1]. . . . .  | 9  |
| 48 | Figure 3.2: The CMS Detector Layout [2]. . . . .   | 11 |
| 49 | Figure 3.3: Overview of the CMS Tracker Layout [3]. . . . .  | 13 |
| 50 | Figure 3.4: Geometrical layout of the CMS Electromagnetic Calorimeter [4]. .   | 14 |
| 51 | Figure 3.5: Layout of the CMS ECAL illustrating its various components [5]. .  | 16 |
| 52 | Figure 3.6: Geometrical layout of the HCAL showing the locations of the<br>53 hadron barrel (HB), endcap (HE), outer (HO) and forward (HF)<br>54 calorimeters [6]. . . . .   | 17 |
| 55 | Figure 3.7: Quarter-view of the CMS muon system (left) and a muon in the<br>56 transverse plane leaves a curved trajectory across the four muon<br>57 detector stations (right). . . . .   | 19 |
| 58 | Figure 3.8: Transverse slice of the CMS detector, showing the individual de-<br>59 tector subsystems and particle signatures in each. The particle<br>60 type can be inferred by combining the detector response in the<br>61 different subdetectors [7]. . . . .  | 22 |
| 62 | Figure 3.9: Phase-2 Pixel Detector Layout[ref]. . . . .  | 24 |
| 63 | Figure 4.1: Sketch of one quarter of the tracker layout in r-z view. In the<br>64 Inner Tracker the green lines correspond to pixel modules made<br>65 of two readout chips and the yellow lines to pixel modules with<br>66 four readout chips. In the Outer Tracker the blue and red lines<br>67 represent the two types of modules $p_T$ modules - (2S and PS). . . | 27 |
| 68 | Figure 4.3: Sketch of four dees (left) forming a double-disc (centre), and<br>69 drawing of a part of a TEDD double-disc (right), illustrating the<br>70 overlap of modules in $\phi$ and $z$ . The upper two dee support struc-<br>71 tures are removed in order to show all layers of modules. The<br>72 modules ob the disc are arranged in 15 or 12 rings. . . . . | 28 |
| 73 | Figure 4.2: Sketch of one quarter of the Outer Tracker in r-z view. Blue (red)<br>74 lines represent PS (2S) modules. The three sub-detectors, named<br>75 TBPS, TB2S, and TEDD, are indicated. All overlapping layers<br>76 are shown separately. . . . .   | 28 |
| 77 | Figure 4.4: Correlation of signals in closely-spaced sensors enables rejection<br>78 of low- $p_T$ particles; the channels shown in green represent the<br>79 selection window to define an accepted stub. . . . .   | 29 |
| 80 | Figure 4.5: Number of stubs per event for standard pixel geometry ( $8s4l$ ). . .  | 32 |

|     |              |   |    |
|-----|--------------|---|----|
| 81  | Figure 4.6:  | Comparison of total number of stubs per event between the standard layout and the other three considered. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right). . . . .  | 32 |
| 84  | Figure 4.7:  | Comparison for number of stubs per event between pixel geometry $8s4l$ with a dead disc (blue) and both pixel geometry $8s4l$ (right plot) and $7s4l$ (left plot) (blue). Both the standard geometry and $7s4l$ are shown in blue in their respective plots. . . . .  | 33 |
| 88  | Figure 4.8:  | Comparison of total number of stubs per $\eta$ value between the standard layout and the other three layouts considered. All four different layouts exhibit the same distribution. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right). . . . .   | 34 |
| 92  | Figure 4.9:  | Comparison of total number of stubs per each of the 6 barrel layers. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right). Tracker Barrel layers shown below. . . . .  | 34 |
| 95  | Figure 4.10: | Comparison of total number of stubs per each of the 5 Endcap Double Discs. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right). Tracker Endcap Discs shown below. . . . .   | 35 |
| 98  | Figure 4.11: | Comparison of total number of stubs for all rings across the five endcap double-discs. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right). The double discs are shown below. . . . .   | 35 |
| 101 | Figure 5.1:  | Simplified model diagrams show of the various Supersymmetry signals considered in this analysis: the T2tt model (top left), the T1tttt model (top right), the T1ttbb model (middle left), the T5tttt (middle right), and the T5ttcc model (bottom). . . . .   | 39 |
| 105 | Figure 5.2:  | Efficiency of the top quark tagger as a function of generator-level top quark $p_T$ for the monojet (red boxes), dijet (magenta triangles), and trijet (green upside-down triangles) categories and for their combination (blue circles), as determined using T2tt signal events with a top squark mass of 850 GeV and an LSP mass of 100 GeV. The vertical bars indicate the statistical uncertainties. . . . .            | 41 |
| 111 | Figure 5.3:  | Search region definitions in the kinematic variables. . . . .   | 44 |
| 112 | Figure 5.4:  | Comparison between the total SM background from simulation and CMS data for $N_t$ (top left), $N_b$ (top right), $m_{T2}$ (bottom left) and $p_T^{miss}$ (bottom right). Total SM backgrounds and signals are scaled to same data yield for a shape comparison. . . . .   | 45 |
| 116 | Figure 5.5:  | Distribution of $p_T^{miss}$ for both the single-electron (left) and single-muon (right) SB data samples compared to predictions for SM processes. The hatched bands indicate the statistical uncertainties in the total SM prediction. Note that the data and the predictions for all backgrounds except that for $t\bar{t}$ , single top quark, and W+jets events are identical between the left and right plots. . . . . | 46 |

|     |             |   |    |
|-----|-------------|---|----|
| 122 | Figure 5.6: | The $p_T^{miss}$ (left) and $N_b$ (right) distributions of data and simulation in the loose dimuon CS after applying both the normalization and shape correction factors. The lower panels show the ratio between data and simulation. Error bars in the plot account only for statistical uncertainties. The values in parentheses indicate the integrated yields for each component. . . . .  | 48 |
| 123 |             |   |    |
| 124 |             |   |    |
| 125 |             |   |    |
| 126 |             |   |    |
| 127 |             |   |    |
| 128 | Figure 5.7: | Observed events compared to the corrected SM background predictions for all 84 search regions in the full data of $35.9 \text{ fb}^{-1}$ collected in 2016. The lower panel shows the ratio between the data and the total SM simulation. The gray bands show the total uncertainty related to the background prediction. . . . .   | 50 |
| 129 |             |   |    |
| 130 |             |   |    |
| 131 |             |   |    |
| 132 |             |   |    |
| 133 | Figure 5.8: | The 95% CL upper limit on the production cross section of the T2tt simplified model as a function of the top squark and LSP masses. The solid black curves represent the observed exclusion contour with respect to NLO+NLL [8] signal cross sections and the change in this contour due to variation of these cross sections within their theoretical uncertainties. The dashed red curves indicate the mean expected exclusion contour and the region containing 68% of the distribution of expected exclusion limits under the background-only hypothesis. . . . . | 51 |
| 134 |             |   |    |
| 135 |             |   |    |
| 136 |             |   |    |
| 137 |             |   |    |
| 138 |             |   |    |
| 139 |             |   |    |
| 140 |             |   |    |
| 141 |             |   |    |
| 142 | Figure 5.9: | The 95% CL upper limit on the production cross section of the T1tttt (upper left), T1ttbb (upper right), T5tttt (bottom left), and T5ttcc (bottom right) simplified models as a function of the gluino and LSP masses. . . . .  | 52 |
| 143 |             |   |    |
| 144 |             |   |    |
| 145 |             |   |    |
| 146 | Figure 6.1: | Leading-order Feynman diagrams for Z+jets and $\gamma+$ jets processes.   | 54 |
| 147 | Figure 6.2: | Shown are both the $p_T^\gamma$ (left) and $p_T^{\gamma(miss)}$ (right) distributions before applying any corrections. $p_T^{\gamma(miss)}$ is obtained by adding the $p_T^\gamma$ to the total $p_T^{miss}$ in every event. . . . .  | 57 |
| 148 |             |   |    |
| 149 |             |   |    |
| 150 | Figure 6.3: | Plots for Fake Rate (left) and Purity (right) as a function of the photon $p_T$ are shown. The events are selected are required to have a $p_T > 200 \text{ GeV}$ , be within the ECAL acceptance range, and pass the loose ID selection cuts. This selection was produced in order to verify the values given by the E/ $\gamma$ POG. As can be seen, the efficiency (purity) is seen to agree with the values of the loose photon ID/isolation selection. . . . .   | 59 |
| 151 |             |   |    |
| 152 |             |   |    |
| 153 |             |   |    |
| 154 |             |   |    |
| 155 |             |   |    |
| 156 |             |   |    |
| 157 | Figure 6.4: | Plots for Fake Rate (left) and Purity (right) as a function of the photon $p_T$ are shown. These plots include photons with the full control region selection. Aside from exhibiting lower statistics, the plots seem to agree with the fake rate and purity before all the control region cuts are applied. . . . .  | 59 |
| 158 |             |   |    |
| 159 |             |   |    |
| 160 |             |   |    |
| 161 |             |   |    |
| 162 | Figure 6.5: | Results of study of the Z+jets to $\gamma+$ jets cross-section ratio for both data and MadGraph simulation. . . . .   | 64 |
| 163 |             |   |    |
| 164 | Figure 6.6: | Jet multiplicity and the associated $S_\gamma$ scale factor in the loose photon control region before any corrections are applied. . . . .  | 65 |
| 165 |             |   |    |
| 166 | Figure 6.7: | $N_{jet}$ distribution in the tight $\mu\mu$ control region after $S_\gamma$ corrections. . . . .   | 65 |

---

|     |  |    |
|-----|--|----|
| 167 | Figure 6.8: $p_T^\gamma$ (left) and $p_T^{\gamma(miss)}$ (right) distributions after applying the $S_\gamma(N_j)$ scale factor. Comparing to Figure 6.2, an improvement in the agreement between data/MC can be observed. . . . .                  | 66 |
| 168 |  |    |
| 169 |  |    |
| 170 | Figure 6.9: $N_b$ distribution before (left) and after (right) applying the $S_\gamma(N_j)$ scale factor. . . . .  | 66 |
| 171 |  |    |
| 172 | Figure 6.10: $H_T$ and $m_{T2}$ distributions applying the $S_\gamma(N_j)$ scale factor. . . . .   | 67 |
| 173 | Figure 6.11: Shown are data/MC comparisons for the $p_T^{miss}$ (left) and $H_T$ (right) distributions after applying both the $N_j$ -dependent shape corrections ( $S_\gamma$ ) and the global normalization scale factor ( $R_{norm}$ ). . . . . | 68 |
| 174 |  |    |
| 175 |  |    |
| 176 | Figure 6.12: Shown are data/MC comparisons for the $N_j$ (left) and $N_b$ (right) distributions after applying both the $N_j$ -dependent shape corrections ( $S_\gamma$ ) and the global normalization scale factor ( $R_{norm}$ ). . . . .        | 68 |
| 177 |  |    |
| 178 |  |    |
| 179 | Figure 6.13: Systematic uncertainty in the final prediction, as a function of the search bin, associated to the MC statistics. . . . .   | 70 |
| 180 |  |    |
| 181 | Figure 6.14: $Z \rightarrow \nu\bar{\nu}$ background prediction for all search bins, including the breakdown of the various uncertainties. . . . .   | 71 |
| 182 |  |    |

# 183 Contents

|     |   |     |
|-----|---|-----|
| 184 | <b>Abstract</b>   | i   |
| 185 | <b>Acknowledgments</b>  | ii  |
| 186 | <b>List of Figures</b>  | iii |
| 187 | <b>1 Introduction</b>   | 1   |
| 188 | <b>2 The CMS Experiment</b>   | 3   |
| 189 | 2.1 Supersymmetric extension of the Standard Model . . . . .                                | 4   |
| 190 | 2.1.1 Simplified Models Approach to Supersymmetry . . . . .                                 | 6   |
| 191 | <b>3 Theoretical Motivation</b>   | 8   |
| 192 | 3.1 The LHC Machine . . . . .   | 8   |
| 193 | 3.2 The CMS Detector . . . . .  | 10  |
| 194 | 3.2.1 Silicon Tracking System . . . . .   | 12  |
| 195 | 3.2.2 Electromagnetic Calorimeter . . . . .   | 14  |
| 196 | 3.2.3 Hadron Calorimeter . . . . .  | 16  |
| 197 | 3.2.4 Magnet . . . . .  | 18  |
| 198 | 3.2.5 Muon Detector . . . . .   | 18  |
| 199 | 3.2.6 Trigger and Data Acquisition . . . . .  | 20  |
| 200 | 3.2.7 Event Reconstruction . . . . .  | 21  |
| 201 | 3.2.8 Future Upgrade of Pixel Detector . . . . .  | 22  |
| 202 | <b>4 Optimizing the Phase-2 Pixel Geometry for the Level-1 Trigger in the Outer Tracker</b> | 26  |
| 204 | 4.1 Upgraded Phase-2 Tracker . . . . .  | 26  |
| 205 | 4.1.1 Phase-2 Tracker Layout . . . . .  | 27  |
| 206 | 4.1.2 Tracker Input to the L1-Trigger . . . . .   | 28  |
| 207 | 4.2 Tracker Stub Simulation . . . . .   | 30  |
| 208 | 4.2.1 Results of Effect of the Phase 2 Pixel Detector on the L1 Stub Rate                   | 32  |
| 209 | 4.3 Results . . . . .   | 36  |
| 210 | <b>5 Search for SUSY in the 0-Lepton Final State</b>  | 37  |
| 211 | 5.1 Introduction . . . . .  | 37  |
| 212 | 5.2 Analysis Description . . . . .  | 37  |
| 213 | 5.2.1 Signal Models . . . . .   | 38  |
| 214 | 5.2.2 Top Tagger . . . . .  | 40  |

---

|     |          |  |           |
|-----|----------|--|-----------|
| 215 | 5.2.3    | Lepton and Track Veto  | 42        |
| 216 | 5.2.4    | Baseline Event Selection   | 43        |
| 217 | 5.2.5    | Search Bin Definition  | 43        |
| 218 | 5.3      | Background Estimation  | 44        |
| 219 | 5.3.1    | Background from $t\bar{t}$ , Single Top Quark and W+jets Events                              | 45        |
| 220 | 5.3.2    | The $Z \rightarrow \nu\bar{\nu}$ Background  | 47        |
| 221 | 5.3.3    | The QCD Multijet Background  | 49        |
| 222 | 5.3.4    | Background From Rare and Other Processes   | 49        |
| 223 | 5.4      | Results and Interpretation   | 50        |
| 224 | <b>6</b> | <b>Estimation of the <math>Z \rightarrow \nu\bar{\nu}</math> Background</b>                  | <b>53</b> |
| 225 | 6.1      | Introduction   | 53        |
| 226 | 6.2      | The Irreducible $Z \rightarrow \nu\bar{\nu}$ Background                                      | 53        |
| 227 | 6.3      | The Loose $\gamma + \text{jets}$ Control Region  | 54        |
| 228 | 6.3.1    | Photon ID and Isolation  | 55        |
| 229 | 6.3.2    | Photon Selection   | 57        |
| 230 | 6.3.3    | Photon Purity and Fake Rate  | 58        |
| 231 | 6.4      | The $Z \rightarrow \mu^+ \mu^-$ Control Region   | 61        |
| 232 | 6.4.1    | Muon ID and Isolation  | 61        |
| 233 | 6.4.2    | Muon Selection in the Tight Control Region   | 62        |
| 234 | 6.5      | Analysis   | 63        |
| 235 | 6.5.1    | Shape Correction Using the $\gamma + \text{jets}$ Control Sample                             | 63        |
| 236 | 6.5.2    | Normalization Correction Using the tight $Z \rightarrow \mu^+ \mu^-$ Control Sample          | 66        |
| 237 | 6.6      | Results  | 69        |
| 238 | 6.6.1    | Systematics  | 69        |
| 239 | 6.6.2    | $Z \rightarrow \nu\bar{\nu}$ Estimation for the Search Bins                                  | 70        |
| 240 | <b>7</b> | <b>Conclusion</b>  | <b>72</b> |
| 241 | 7.1      | L1 Stub Counting for the HL-LHC CMS Tracker Upgrade  | 72        |
| 242 | 7.2      | Search for SUSY in the All-Hadronic Channel  | 73        |
| 243 | 7.3      | Estimation of the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ Background with a Hybrid Method | 74        |
| 244 | <b>A</b> | <b>References</b>  | <b>76</b> |

<sup>245</sup> **Chapter 1**

<sup>246</sup> **Introduction**

<sup>247</sup> The work for this thesis was performed at CERN on CMS Experiment. CERN stands  
<sup>248</sup> for European Organization for Nuclear Research. It was founded in 1954 and is located  
<sup>249</sup> at the Franco-Swiss border near Geneva. At CERN, physicists and engineers are probing  
<sup>250</sup> the fundamental structure of the universe. They use the world's largest and most complex  
<sup>251</sup> scientific instruments to study the basic constituents of matter – the fundamental parti-  
<sup>252</sup> cles. The instruments used at CERN are purpose-built particle accelerators and detectors.  
<sup>253</sup> Accelerators boost beams of particles to high energies before the beams are made to col-  
<sup>254</sup> lide with each other or with stationary targets. Detectors observe and record the results  
<sup>255</sup> of these collisions. The accelerator at CERN is called the Large Hadron Collider (LHC),  
<sup>256</sup> the largest machine ever built by humans and it collides particles (protons) at close to the  
<sup>257</sup> speed of light. The process gives the physicists clues about how the particles interact, and  
<sup>258</sup> provides insights into the fundamental laws of nature. Seven experiments at the LHC use  
<sup>259</sup> detectors to analyze particles produced by proton-proton collisions. The biggest of these  
<sup>260</sup> experiments, ATLAS and CMS, use general-purpose detectors designed to study the fun-  
<sup>261</sup> damental nature of matter and fundamental forces and to look for new physics or evidence  
<sup>262</sup> of particles that are beyond the Standard Model. Having two independently designed de-  
<sup>263</sup> tectors is vital for cross-confirmation of any new discoveries made. The other two major  
<sup>264</sup> detectors ALICE and LHCb, respectively, study a state of matter that was present just  
<sup>265</sup> moments after the Big Bang and preponderance of matter than antimatter. Each experi-

266 ment does important research that is key to understanding the universe that surrounds and  
267 makes us.

268

269     [Chapter 2](#) presents a basic description of the Large Hadron Collider and CMS Detector

270

271     [Chapter 3](#) gives a brief motivation

272

273     [Chapter 4](#) is dedicated to a study optimizing

274

275     [Chapter 5](#) ptimated.

276

277     [Chapter 6](#) details an improvarger production cross-section than Z+jets process used  
278 before.

279

280     The conclusions and results of each chapter are presented in the corresponding chap-  
281 ter.

282

283     This thesis work has been presented at several internal meetings of the CMS Experi-  
284 ment and at the following international meetings and conferences:

285     1. **Andrés Abreu** gave a talk “*Estimation of the Z Invisible Background for Searches*  
286       *for Supersymmetry in the All-Hadronic Channel*” at “APS April 2018: American  
287       Physical Society April Meeting 2018, 14-17 Apr 2018”, Columbus, OH

288     2. **Andrés Abreu** gave a talk “*Phase-2 Pixel upgrade simulations*” at the “USLUA  
289       Annual meeting: 2017 US LHC Users Association Meeting, 1-3 Nov 2017”, Fer-  
290       milab, Batavia, IL

291 **Chapter 2**

292 **The CMS Experiment**

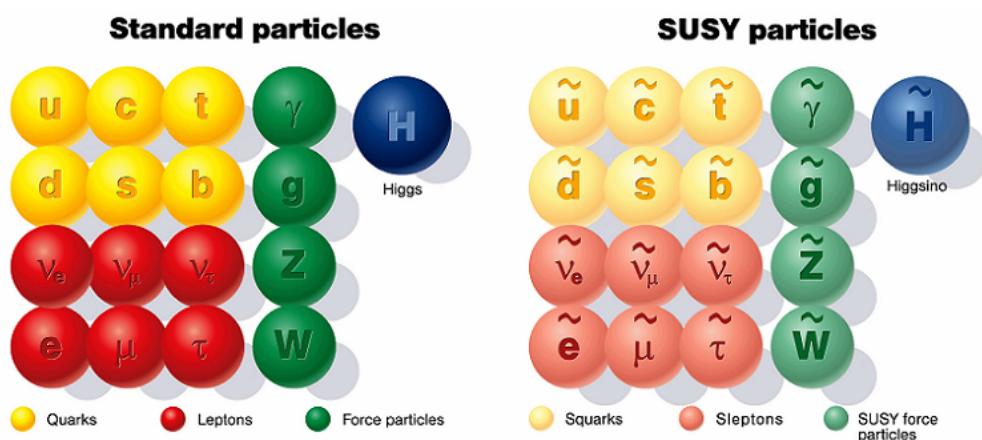
293 The Compact Muon Solenoid (CMS) detector is a general purpose particle detector  
294 designed to investigate various physical phenomena concerning the SM and beyond it,  
295 such as Supersymmetry, Extra Dimensions and Dark Matter. As its name implies, the  
296 detector is a solenoid which is constructed around a superconducting magnet capable of  
297 producing a magnetic field of 3.8 T. The magnetic coil is 13m long with an inner diameter  
298 of 6m, making it the largest superconducting magnet ever constructed. The CMS detector  
299 itself is 21m long with a diameter of 15m and it has a weight of approximately 14,000  
300 tons. The CMS experiment is one of the largest scientific collaborations in the history  
301 of mankind with over 4,000 participants from 42 countries and 182 institutions. CMS is  
302 located at one of these points and it essentially acts as a giant super highspeed camera  
303 that makes 3D images of the collisions that are produced at a rate of 40 MHz (40 million  
304 times per second). The detector has an onion-like structure to capture all the particles that  
305 are produced in these high energy collisions most of them being unstable and decaying  
306 further to stable particles that are detected. CMS detector was designed with the following  
307 features (as shown in [2.1](#) Figure 1) :

- 308 1. A **magnet** with large bending power and high performance muon detector for good  
309 muon identification and momentum resolution over a wide range of momenta and  
310 angles.

- 311    2. An **inner tracking system** capable of high reconstruction efficiency and momen-  
 312    tum resolution requiring **pixel detectors** close to the interaction region.
- 313    3. An **electromagnetic calorimeter** able to provide good electromagnetic energy res-  
 314    olution and a high isolation efficiency for photons and leptons.
- 315    4. A **hadron calorimeter** capable of providing precise missing-transverse-energy and  
 316    dijet-mass resolution.

## 317    2.1 Supersymmetric extension of the Standard Model

318    The supersymmetry theories are based on a symmetry between fermions and bosons.  
 319    It is similar to solving the ‘electron mass hierarchy problem’ in quantum mechanics,  
 320    where the number of particles is doubled: in addition to the electron, there is also a  
 321    positron. The virtual electron–positron contributions solved the problem of electrons  
 322    having a small mass by smearing out the electric charge. Supersymmetry is an analo-  
 323    gous theory where once again the set of particles is doubled, and in doing so the loop  
 324    contributions of one particle to the Higgs are cancelled by the loop contributions of its  
 325    super-partner. It extends space-time symmetry since it relates matter particles to force  
 326    particles. It relates particles with different spins but the same gauge charges.



**Figure 2.1:** Schematic representation of the Standard Model particles with the added SUSY particles.

327    The gluon (QCD) has a fermionic partner, the gluino. Similarly, the spin 1 gauge

328 bosons  $W^+$ ,  $W^0$ ,  $W^-$  and  $B^0$  have their counterparts, called winos and binos, that, af-  
329 ter electroweak symmetry breaking, mix to give the mass eigenstate  $Z^0$  and  $\gamma$ , and the  
330 zino and photino respectively. For every lepton/quark there is a bosonic partner (i.e. a  
331 scalar slepton/squark). There are three families for each of the quark and lepton super-  
332 multiplets. The left-handed and right-handed pieces of the squarks and sleptons are two  
333 separate components, as in the corresponding SM sector. The Higgs doublet is extended  
334 by another doublet leading to other Higgs boson particles (Higgsinos).

335

336 Even though supersymmetry solves many problems in particle physics, it also poses  
337 new problems. If these new particles had the same mass as their counterparts in the stan-  
338 dard model, we would have already observed them. Since none of these particles have  
339 been found yet, SUSY must be a broken symmetry. What makes superpartners heavier  
340 than ordinary particles? Why are superpartners so well hidden in rare phenomena? This  
341 arbitrary mass spectrum for the superpartners would have effects that are far too large in  
342 rare processes that change the flavor of particles. There must be some special reason why  
343 such effects are well hidden. How do we extract information on the mechanism of super-  
344 symmetry breaking? How does supersymmetry impact cosmology? Is the lightest super-  
345 symmetric partner what really composes Dark Matter? Different mechanisms have been  
346 suggested on how this symmetry is broken and have led to different phenomenological  
347 scenarios like the minimal supergravity model (mSUGRA) [9, 10], the gauge-mediated  
348 supersymmetry breaking (GMSB) [11–13] model or R-parity [14, 15]. A discussion of  
349 mSUGRA and GMSB is beyond the scope of this dissertation. R-parity is a new sym-  
350 metry that has been added to the minimal supersymmetry scenario which prevents the  
351 violation of the leptonic and baryonic numbers. All SM fields have even R-parity, while  
352 SUSY particles have odd R-parity. The most obvious experimental constraint comes from  
353 the non-observation of proton decay [16], which would violate both baryonic and leptonic  
354 number conservation. Due to R-parity every interaction vertex in the theory involves an  
355 even number of sparticles, meaning that they must be pair-produced. The decay chains  
356 of the produced particles are characterized by the presence of one stable particle, generi-

357 cally called the Lightest Supersymmetric Particle (LSP), that may account for all the dark  
358 matter in the Universe. However, in other models, like the R-parity violating SUSY this  
359 may not apply. In other scenarios, those LSP can be regarded as vanishing mass and may  
360 not be enough to satisfy the dark matter in the Universe.

361 **2.1.1 Simplified Models Approach to Supersymmetry**

362 Simplified models [17] are a new approach for characterizing LHC supersymmetry  
363 results. In a simplified model only a few new particles and a single decay topology are  
364 introduced, and hence these models are easy to constrain. Traditionally, testing a particle  
365 physics theory against experimental results requires calculating what the theory predicts  
366 in an experimental situation using a particle accelerator detector. This translates into  
367 computing: the masses of new particles, their decay widths, their branching ratios and  
368 production cross-sections. This information is subsequently used in Monte Carlo simula-  
369 tions of passage of proton-proton collisions products and their decays through the CMS  
370 detector. This is done under the hypothesis that the theory is true. The simulated data  
371 is analyzed just like the actual experimental data resulting from proton-proton collisions.  
372 The theoretical and the experimental results can then be compared. Though conceptually  
373 uncomplicated it requires access to the experimental data, and the step of simulating the  
374 detector response requires intimate knowledge of the CMS detector by theorist commu-  
375 nity. It is also restrictive being model/theory dependent. The CMS collaboration, like  
376 ATLAS, does not share experimental details but only the final results via publications.  
377 The Simplified Model approach circumvents this issue by presenting results on very sim-  
378 ple phenomenological models instead of a specific model, such as the constrained MSSM  
379 (cMSSM) [18]. Each simplified models encompasses a specific phenomenological fea-  
380 ture that is common to many different theories. A limited set of hypothetical particles  
381 and decay chains are introduced to produce a given topological signature. The amplitudes  
382 describing the production and decays of these particles are parametrized in terms of the  
383 particle masses and their branching ratios to daughter particles. This makes the analysis

of simplified models less model dependent. The Simplified Models assume that a particular decay signature can be realized without specifying the exact mechanism, offering the possibility to overcome small branching ratios. Some Simplified Models with their corresponding production and decay modes are shown in the [Table 2.1](#). We have used T2tt, T1tttt, T1ttbb, T5tttt and T5ttcc models in analysis in this thesis. Some of these topologies are further discussed in more detail in [chapter 5](#), since they correspond to some of the signals for SUSY in the treated analysis.

| Model name | Production mode        | Decay   | Visibility   |
|------------|------------------------|---|--|
| T1         | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{\text{LSP}}$   | All-Hadronic   |
| T2         | $\tilde{q}\tilde{q}^*$ | $\tilde{q} \rightarrow q\tilde{\chi}_{\text{LSP}}$  | All-Hadronic   |
| T5zz       | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_2^0, \tilde{\chi}_2^0 \rightarrow Z\tilde{\chi}_{\text{LSP}}$   | All-Hadronic<br>Opposite-Sign Dileptons<br>Multileptons                                  |
| T3w        | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{\text{LSP}}$<br>$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \rightarrow W^\pm\tilde{\chi}_{\text{LSP}}$    | Single Lepton + Jets   |
| T5lnu      | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \rightarrow \ell\nu\tilde{\chi}_{\text{LSP}}$   | Same-Sign Dileptons  |
| T3lh       | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{\text{LSP}}$<br>$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_2^0, \tilde{\chi}_2^0 \rightarrow \ell^+\ell^-\tilde{\chi}_{\text{LSP}}$ | Opposite-Sign Dileptons  |
| T1bbbb     | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow b\bar{b}\tilde{\chi}_{\text{LSP}}$   | All-Hadronic (b)   |
| T1tttt     | $\tilde{g}\tilde{g}$   | $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_{\text{LSP}}$   | All-Hadronic (b)<br>Single Lepton + Jets (b)<br>Same-Sign Dileptons (b)<br>Inclusive (b) |
| T2bb       | $\tilde{b}\tilde{b}^*$ | $\tilde{b} \rightarrow b\tilde{\chi}_{\text{LSP}}$  | All-Hadronic (b)   |
| T6ttww     | $\tilde{b}\tilde{b}^*$ | $\tilde{b} \rightarrow t\tilde{\chi}^-, \tilde{\chi}^- \rightarrow W^-\tilde{\chi}_{\text{LSP}}$  | Same-Sign Dileptons (b)  |
| T2tt       | $\tilde{t}\tilde{t}^*$ | $\tilde{t} \rightarrow t\tilde{\chi}_{\text{LSP}}$  | All-Hadronic (b)   |

**Table 2.1:** Some Simplified Models with their corresponding production and decay modes.

391 **Chapter 3**

392 **Theoretical Motivation**

393 **3.1 The LHC Machine**

394 The Large Hadron Collider (LHC) [19], located at the European Organization for  
395 Nuclear Research (CERN) complex in Geneva Switzerland, is the world's largest and  
396 most powerful particle collider as well as the most complex experimental device ever  
397 assembled. The main motivation for its construction was uncovering the nature behind  
398 electroweak symmetry breaking due to the Higgs mechanism as well as to search for new  
399 physical phenomena Beyond the Standard Model (BSM) [20]. Since its completion in  
400 2008, the LHC has helped researchers obtain a plethora of significant results in the field  
401 of particle physics and it successfully achieved its purpose with the discovery of the Higgs  
402 Boson in 2012.

403

404 The LHC consists of several superconducting magnets and accelerating structures ar-  
405 ranged along a 27-kilometer circumference ring which serve to boost and bend the proton  
406 (or other heavy ion) beams along its path. In order for the particle beams to maintain the  
407 orbit while they are accelerated, a magnetic field of 8.3 T is produced by the supercon-  
408 ducting magnets along the circular tunnel, 100 meters below the ground. The magnets are  
409 maintained at a temperature of 1.9 K by using liquid helium. The two beams are made  
410 to travel in opposite directions and in separate beam pipes which are kept at ultrahigh

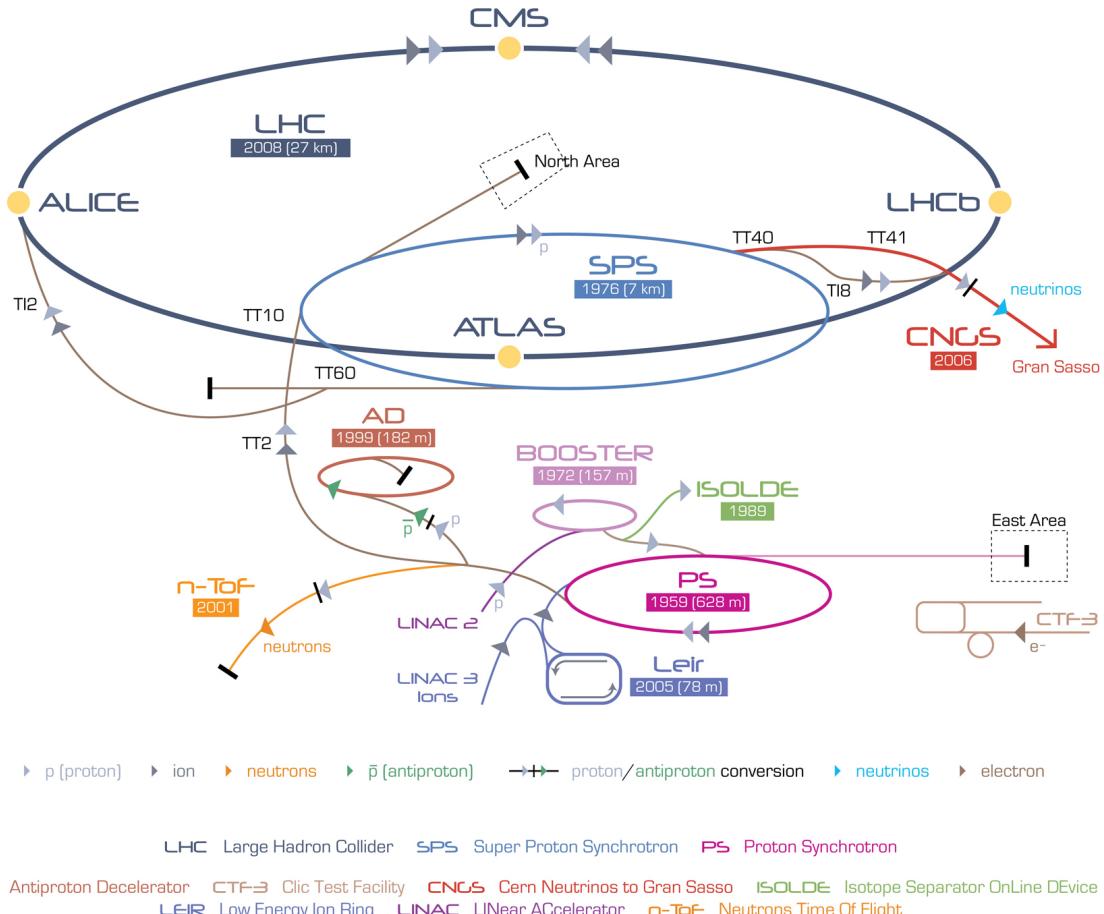


Figure 3.1: The CERN Accelerator Complex [1].

vacuum and are accelerated to near the speed of light. They are then collided head-on at any of the four interaction points located around the LHC ring. Approximately  $1.5 \times 10^{11}$  protons are generated every 25 ns with an energy of 26 GeV at the Proton Synchrotron (PS). The proton beam is then accelerated to an energy of 450 GeV at the Super Proton Synchrotron (SPS) [21] before being delivered to the LHC. At design level, each bunch is accelerated to an energy of 7 TeV and are made to collide at a centre-of-mass energy of up to 14 TeV with a frequency of 40 MHz. The current proton beams achieve an instantaneous luminosity on the order of  $10^{32} \text{ cm}^{-2} \cdot \text{s}^{-1}$  and a total integrated luminosity of  $37 \text{ fb}^{-1}$  was reached during 2016 [4].

420

421 The four main detectors comprising the LHC machine are CMS, ATLAS [22], LHCb  
422 [23] and ALICE [24]. Both CMS and ATLAS are general purpose detectors whose initial

423 designs had the detection of the SM Higgs boson, with its wide range of decay modes,  
 424 in mind. Both detectors managed to accomplish this goal when a 126 GeV scalar boson  
 425 consistent with the SM Higgs was independently verified by both experiments in July of  
 426 2012. Furthermore, the designs for CMS and ATLAS allow for the search of many other  
 427 additional phenomena in BSM physics such as Supersymmetry, Dark Matter [25], Dark  
 428 Sector [26], etc. On the other hand, the LHCb and ALICE detectors focus on more partic-  
 429 ular kinds of searches. The main motivation for the LHCb experiment, where the b stands  
 430 for beauty, concerns itself with the measurement of CP violation parameters in b-hadron  
 431 interactions and studies cover a wide range of aspects of Heavy Flavor Electroweak and  
 432 QCD physics. Meanwhile, the ALICE experiment focuses on the study of heavy ion (Pb-  
 433 Pb) nuclei collisions at a centre-of-mass energy of 2.76 TeV in order to better understand  
 434 the physics behind strongly interacting matter at extreme energy densities.

## 435 3.2 The CMS Detector

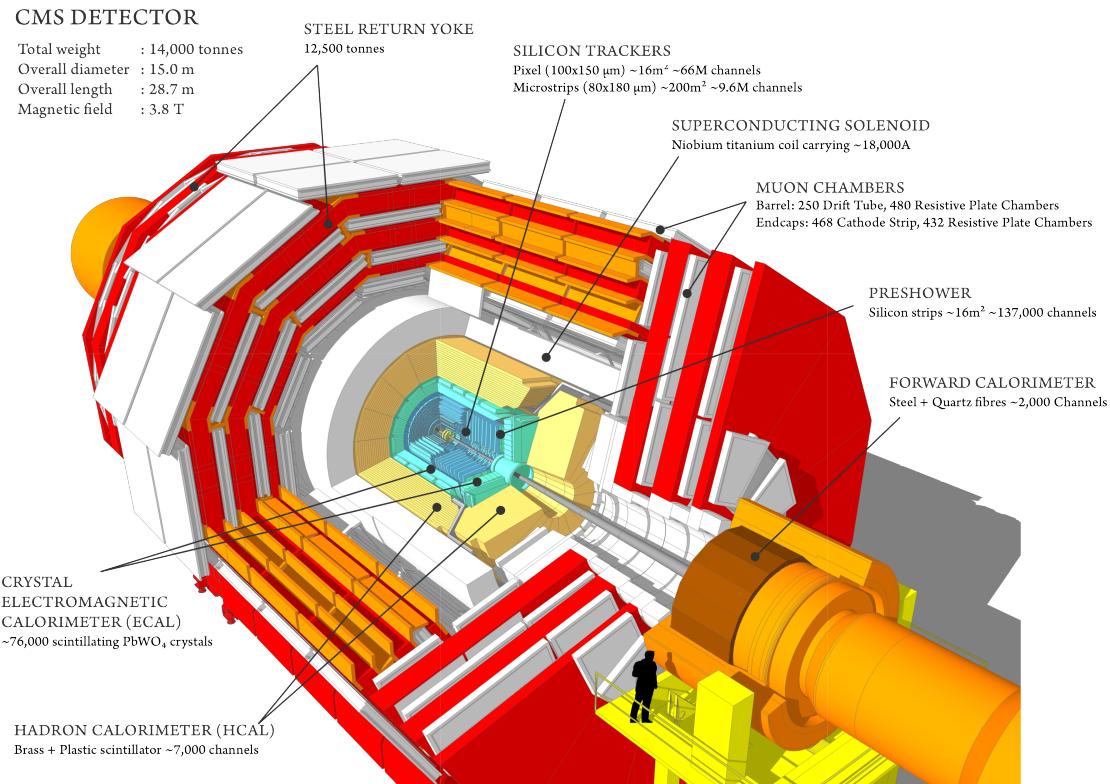
436 The Compact Muon Solenoid (CMS) detector is a general purpose particle detector  
 437 designed to investigate various physical phenomena concerning the SM and beyond it,  
 438 such as Supersymmetry, Extra Dimensions and Dark Matter. As its name implies, the  
 439 detector is a solenoid which is constructed around a superconducting magnet capable of  
 440 producing a magnetic field of 3.8 T. The magnetic coil is 13m long with an inner diameter  
 441 of 6m, making it the largest superconducting magnet ever constructed. The CMS detector  
 442 itself is 21m long with a diameter of 15m and it has a weight of approximately 14,000  
 443 tonnes. The CMS experiment is one of the largest scientific collaborations in the history  
 444 of mankind with over 4,000 participants from 42 countries and 182 institutions.

445

446 In order to meet the many needs of the SM and BSM searches, and the goals of the  
 447 LHC physics program, the CMS detector was designed with the following features:

448

- 449 • A magnet with large bending power and high performance muon detector for good



**Figure 3.2:** The CMS Detector Layout [2].

muon identification and momentum resolution over a wide range of momenta and angles.

- An inner tracking system capable of high reconstruction efficiency and momentum resolution requiring pixel detectors close to the interaction region.
- An electromagnetic calorimeter able to provide good electromagnetic energy resolution and a high isolation efficiency for photons and leptons.
- A hadron calorimeter capable of providing precise missing-transverse-energy ( $p_T^{miss}$ ) and dijet-mass resolution.

A general layout of the CMS detector and all its constituent sub-detectors can be seen in Figure 3.2. The configuration of the CMS sub-detectors follow a cylindrical layer pattern that is symmetrical about the interaction region and consists of a central barrel with endcaps on both ends.

The coordinate system for the CMS detector design uses a right-hand rule convention centered around the ideal interaction point to describe the positions of objects in the experiment. The z-axis is defined along the direction of the LHC beam, with the x-axis pointing towards the center of the LHC ring. In terms of polar coordinates then, r is the radial distance from the center of the pipe, the polar angle  $\theta$  is measured against the z-axis and the azimuthal angle  $\phi$  is measured with respect to the x-axis. However, the pseudorapidity  $\eta$  is generally preferred over the polar angle  $\theta$ . The pseudorapidity is defined as:

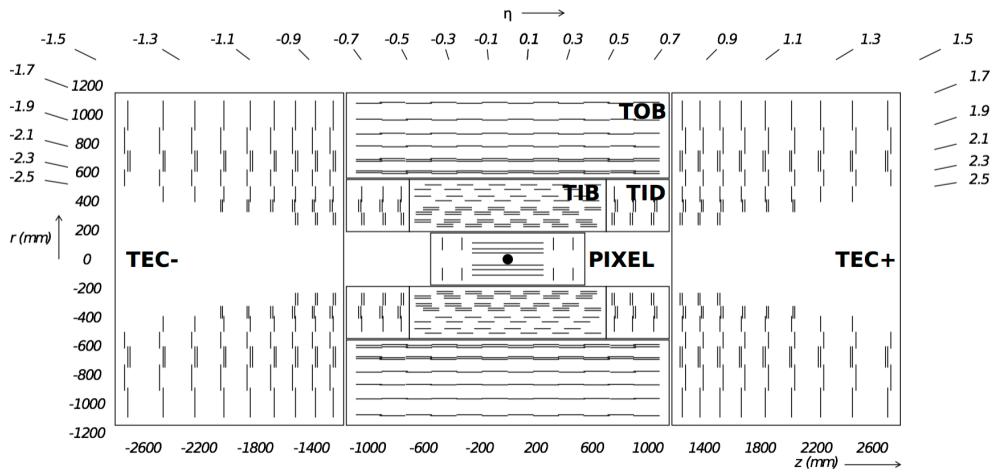
$$\eta = -\ln \tan \frac{\theta}{2}.$$

### 3.2.1 Silicon Tracking System

The CMS tracking system was designed with the goal of obtaining precise and efficient measurements for the trajectories of charged particles resulting from proton-proton collisions at the LHC. In addition, it allows for the precise measurement of secondary vertices and impact parameters needed to efficiently identify the heavy flavours produced in many interesting physics channels. Due to the LHC's design Luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , the currently installed CMS phase-1 tracker is expected to handle an average of 1000 particles from over 20 overlapping proton-proton interactions per bunch crossing, every 25 ns. This required a detector technology capable of achieving a high granularity and fast response as well as being tolerant to the radiation produced from the intense particle flux. These considerations lead to a tracker design composed entirely of silicon detector technology which features an active silicon area of about  $200 \text{ m}^2$ , making it the largest silicon tracker ever built [5].

488

The CMS tracker is built in a cylindrical manner around the interaction point and has a diameter of 2.5 m and a length of 5.8 m. It is comprised of a pixel detector with three barrel layers, positioned at a distance between 4.4 cm and 10.2 cm from the interaction region, and a silicon strip tracker with 10 barrel detection layers extending to a radius of



**Figure 3.3:** Overview of the CMS Tracker Layout [3].

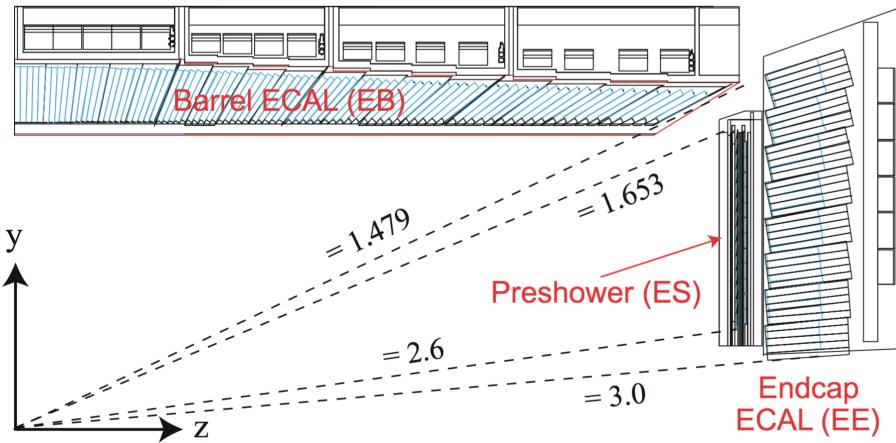
493    1.1 m. Each system is made complete by endcaps at opposite sides of the barrel, consisting  
 494    of 2 discs for the pixel detector and 3 plus 9 discs for the strip tracker, extending the  
 495    acceptance region of the tracker up to a pseudorapidity  $|\eta| < 2.5$ .

496

497    The pixel detector is the part of the CMS tracking system closest to the interaction  
 498    region and consists of 3 barrel layers (BPix) and 2 endcap discs (FPix). It is responsible  
 499    for providing precise tracking points in  $r$ - $\phi$  and  $z$ , a feature that is required for the small  
 500    impact parameter resolution, needed for good secondary vertex reconstruction. The de-  
 501    tector contains 1440 modules covering an area of approximately  $1 \text{ m}^2$  making up a total  
 502    of 66 million pixels. The sensors were designed using an n-on-n concept from  $320 \mu\text{m}$   
 503    thick silicon wafers and are fabricated with read-out chips (ROCs) that are bump-bonded  
 504    to the sensor in standard  $0.25 \mu\text{m}$  CMOS technology. Each of the pixels has a pitch size  
 505    of  $100 \times 150 \mu\text{m}^2$ , which corresponds to an occupancy of about  $10^{-4}$  per bunch crossing.

506

507    The silicon strip tracker is built surrounding the pixel detector and consists of three  
 508    large subsystems. The Tracker Inner Barrel and Disks (TIB/TID) extend to a radius of  
 509    55 cm and are composed of four barrel layers, completed by three disks at each end. The  
 510    TIB/TID employs the use of  $320 \mu\text{m}$  thick silicon micro-strip sensors in order to deliver  
 511    up to 4  $r$ - $\phi$  measurements on a trajectory. Surrounding the TIB/TID is the Tracker Outer



**Figure 3.4:** Geometrical layout of the CMS Electromagnetic Calorimeter [4].

512 Barrel (TOB), which consists of 6 barrel layers and has an outer radius of 116 cm. The  
 513 TOB extends symmetrically in  $z$  between  $\pm 118$  cm and provides an additional  $6 r\phi$   
 514 measurements for a trajectory. Beyond the TOB's  $z$  range lie the Tracker Endcaps (TEC+  
 515 and TEC-, where the sign indicates their position respect to  $z$ ). Each TEC consists of 9  
 516 discs with up to 7 rings of silicon micro-strip detectors, providing up to 9 additional  $\phi$   
 517 measurements per trajectory. The CMS silicon strip tracker has a total active silicon area  
 518 of  $198 \text{ m}^2$  and is composed of 15,148 sensor modules.

### 519 3.2.2 Electromagnetic Calorimeter

520 The CMS Electromagnetic Calorimeter (ECAL) is a hermetic homogeneous calorime-  
 521 ter whose function is to measure the energy of particles that interact via the electromag-  
 522 netic force. With the use of 75,848 scintillator crystals, it is capable of providing good  
 523 energy resolution within the requirements of the ambitious LHC program. In particular,  
 524 the ECAL's design was optimized to search for diphoton events resulting from Higgs bo-  
 525 son decays ( $H \rightarrow \gamma\gamma$ ).

526

527 The ECAL is composed of two main sub-systems – the barrel calorimeter (EB) and the  
 528 endcap calorimeter (EE) – and is completed by a preshower calorimeter (ES), as shown  
 529 in [Figure 3.5](#). It covers a solid angle up to a pseudorapidity of  $|\eta| < 3$ , with the EB ex-

530 tending in the range of  $|\eta| < 1.479$  and the EE covering a range of  $1.479 < |\eta| < 3$ . Both  
531 subsystems are composed of lead tungstate ( $\text{PbWO}_4$ ) scintillator crystals which provide  
532 a fast response time and high radiation tolerance, both crucial requirements for optimal  
533 performance at LHC operating conditions. In addition, the properties of the  $\text{PbWO}_4$  crys-  
534 tals (high density, short radiation length and small Molière radius) led to the design of a compact  
535 calorimeter with fine granularity.

536

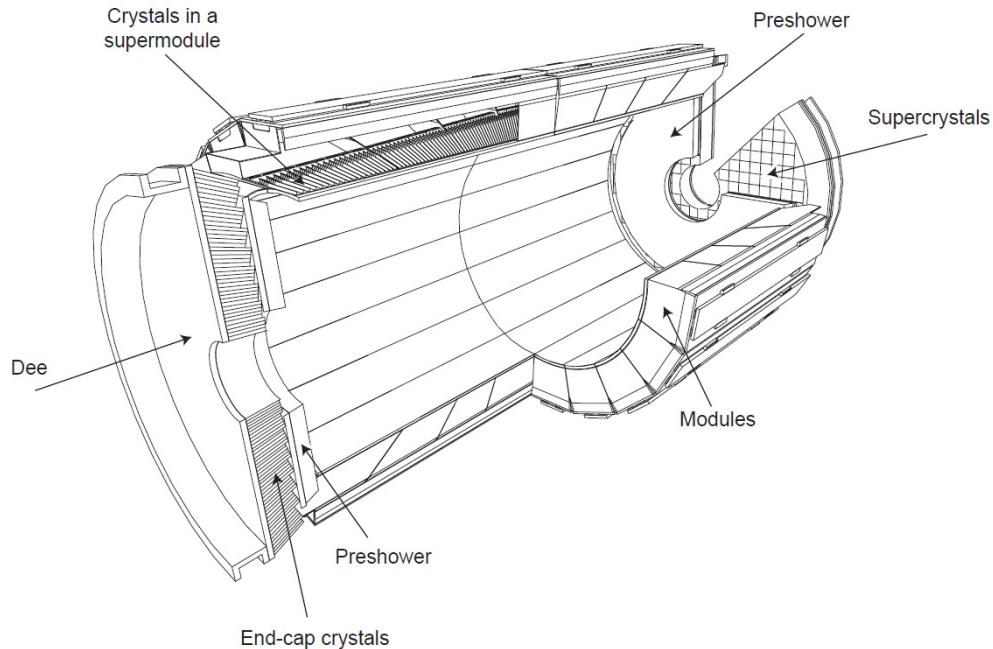
537 The barrel part of the ECAL (EB) is composed of specially designed avalanche pho-  
538 todiodes (APD). It consists of 61,200 crystals, forming a total volume of  $8.14 \text{ m}^3$  and  
539 weighing about 67.4 t. The crystals that form the EB are 230 mm long with a cross-  
540 section of  $22 \times 22 \text{ mm}^2$  at the front face and  $26 \times 26 \text{ mm}^2$  at the back. They are organized  
541 in pairs within thin-walled alveolar structures called submodules. The submodules are  
542 assembled into different types of modules that differ by their location with respect to  $\eta$   
543 and contain 400 or 500 crystals each. These modules are then arranged in sets of four  
544 modules, called supermodules, and contain 1700 crystals each. Thus, the EB is composed  
545 of two half-barrels, each consisting of 18 supermodules.

546

547 In contrast, the photodetectors used in the endcap section of the ECAL are vacuum  
548 phototriodes (VPT) [27]. Each of the endcaps contain 7,324 crystals which, in total, oc-  
549 cupy a volume of  $2.90 \text{ m}^3$  and weigh about 24.0 t. The crystals in the EE are 220 mm in  
550 length with a cross-section of  $28.62 \times 28.62 \text{ mm}^2$  for the front face and  $30 \times 30 \text{ mm}^2$  in the  
551 rear. They are all identical in shape and are arranged in mechanical units of  $5 \times 5$  crystals,  
552 called supercrystals (SC), which consist of carbon-fibre alveola structures. Each of the  
553 endcaps are divided into two semi-circular structures, called *Dees*, which hold a total of  
554 3,662 crystals.

555

556 The ES preshower is located before the EE detector and spans a pseudorapidity range  
557 of  $1.653 < |\eta| < 2.6$ . Its main purpose is to identify neutral pions in the endcaps as  
558 well as to improve the determination of electrons and photons with high granularity. The

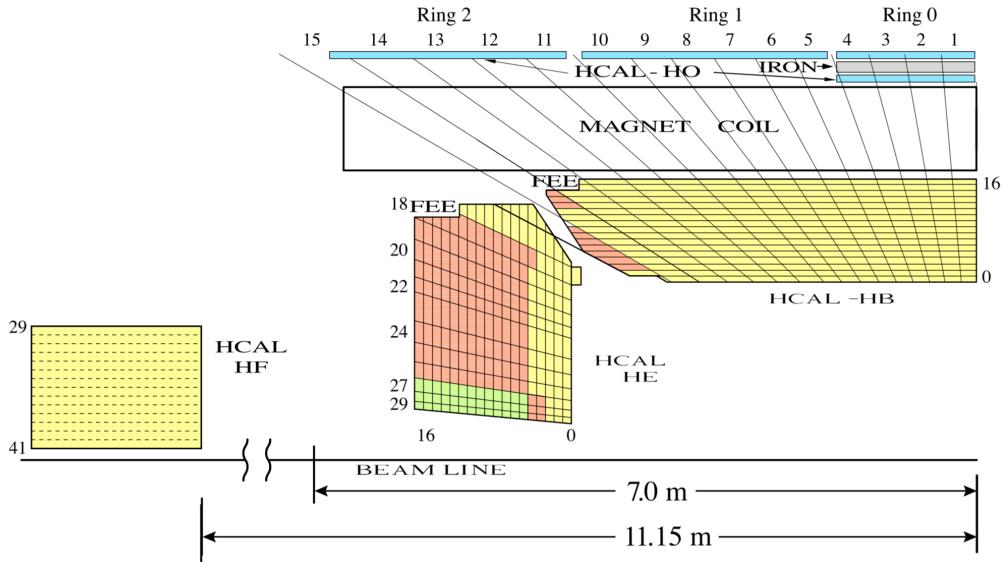


**Figure 3.5:** Layout of the CMS ECAL illustrating its various components [5].

559 preshower consists of two layers and has a total thickness of 20 cm. The first layer is  
 560 conformed by lead radiators which initiate electromagnetic showers from incoming pho-  
 561 tons and electrons. Meanwhile, the second layer is composed of silicon strips which are  
 562 capable of measuring the deposited energy and the transverse shower profiles.

### 563 3.2.3 Hadron Calorimeter

564 The CMS Hadron calorimeter (HCAL) conforms the next layer of the CMS detector.  
 565 It is a sampling calorimeter that consists of alternating layers of massive absorbing brass  
 566 plates and plastic scintillator tiles and is of particular importance for the measurement of  
 567 hadron jet energy and  $p_T^{miss}$ . The HCAL detector is located in between the outer extent of  
 568 the ECAL ( $R = 1.77$  m) and the inner extent of the magnet coil ( $R = 2.95$  m). Similar to  
 569 the other CMS subsystems, it's composed of a barrel part (HB) and an endcap part (HE).  
 570 In addition, it features a tail-catching outer calorimeter (HO), located outside the magnet,  
 571 and a forward calorimeter (HF) in the very forward region near the beam line. A layout  
 572 of the HCAL system can be seen in [Figure 3.6](#).



**Figure 3.6:** Geometrical layout of the HCAL showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [6].

574 The barrel component of the HCAL is a sampling calorimeter which covers the pseu-  
 575 dorapidity range  $|\eta| < 1.3$ . It consists of both the HB and the HO detectors. The reason  
 576 behind separating the barrel detector into the HB and HO is due to the limited amount  
 577 of space available for the barrel detector. The HB is located within the superconducting  
 578 magnet coil and is supplemented by the HO in between the outer solenoid coil and the  
 579 muon chambers. Therefore, the HO acts as a tail-catcher in order to improve the jet en-  
 580 ergy measurements and  $p_T^{miss}$ , with the solenoid in between acting as absorber material.  
 581 The HB consists of two half-barrel sections, identified as HB+ and HB- due to their ge-  
 582 ometrical location, which are composed of 36 identical azimuthal wedges. The wedges,  
 583 which are constructed out of flat brass absorber plates, are aligned parallel to the beam  
 584 axis and are segmented into four azimuthal angle ( $\phi$ ) sections.

585

586 The HE covers a significant amount of the pseudorapidity in the range of  $1.3 < |\eta| <$   
 587 3, a region containing about 34% of the particles produced in the final state. Due to  
 588 the high luminosity of the LHC, the HE is required to have a high radiation tolerance at  
 589  $|\eta| \simeq 3$ , as well as being capable of handling high counting rates. Similar to the HB, the  
 590 HE is also composed of brass absorber plates and scintillator plates which are read out

591 by wavelength shifting fibers. The light captured by the scintillators merges within the  
 592 wavelength shifting fibers and then it's read out by hybrid photo-diodes. The scintillators  
 593 are partitioned in towers with an area of  $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ .

594

### 595 3.2.4 Magnet

596 The CMS superconducting solenoid is one of the driving features of the detector de-  
 597 sign. It is capable of providing a magnetic field with a 3.8T magnitude, which allows  
 598 for the large bending power needed for precise particle transverse momentum ( $p_T$ ) mea-  
 599 surements. The magnet is made of four layers of stabilized reinforced Niobium-Titanium  
 600 (NbTi) and has a cold mass of 220 t. The solenoid consists of a 13 m long coil with an  
 601 internal diameter of 6 m, which houses both the tracking and calorimetric system. This  
 602 design allows for particles to be measured prior to crossing the magnetic coil which sig-  
 603 nificantly improves the energy resolution.

### 604 3.2.5 Muon Detector

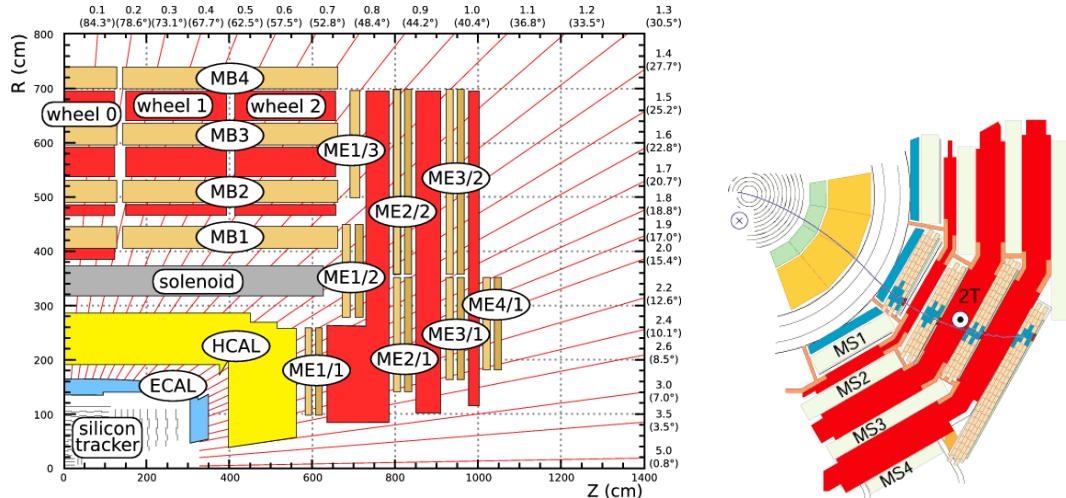
605 As implied by the detector's name, precise and robust muon measurements have been  
 606 a central theme of the CMS experiment since the early stages of its design. The detector  
 607 design takes into account that muons behave as minimum ionizing particles (MIPs) [28]  
 608 and can therefore manage to traverse the tracker and calorimeters with minimal energy  
 609 loss. Furthermore, due to their relatively long lifetime they can be efficiently identified  
 610 by a dedicated system at the outer region of the detector. Consequently, the CMS muon  
 611 systems comprise the outermost layer of the detector, which are integrated into the mag-  
 612 net return yoke that surrounds the solenoid.

613

614 The muon system is capable of three main functions: muon identification, muon  $p_T$   
 615 measurement and triggering. It is composed of three different types of detectors, all of  
 616 which make use of gaseous chamber technology. This choice of detector provides a cost

efficient way of covering most of the full solid angle, featuring a total of 25,000 m<sup>2</sup> of detection plates. As a consequence of the shape of the solenoid magnet, the muon detector was designed to have a cylindrical barrel section as well as two planar endcap regions.

620



**Figure 3.7:** The left diagram shows a quarter-view of CMS with both the muon barrel (MB) and endcap (ME) stations [29]. The right diagram shows a muon in the transverse plane leaves a curved trajectory across the four muon detector stations [30].

For the barrel region of the muon detector, four layers of drift tube (DT) modules are used, covering a pseudorapidity of up to  $|\eta| < 1.2$ . These four layers, called “stations”, are arranged in cylindrical concentric layers around the beam line, where the first three layers have 60 DTs each and the outer cylinder has 70. Each of the DT stations contain 12 individual gas filled tubes, all of which have a 4cm diameter and a center electrode. The use of DTs as tracking detectors for the barrel muon system is possible because of the low expected muon rate and the relatively low strength of the local magnetic field.

628

The endcap regions of the muon system are subject to a higher muon rate, and cover the range  $0.9 < |\eta| < 2.4$  where the magnetic field is stronger and less homogeneous. Considering these conditions, cathode strip chambers (CSCs) are employed, which feature high granularity, fast response time and adequate radiation hardness. CSCs consist of arrays of positively-charged “anode” wires crossed with negatively-charged copper “cathode” strips within a gas volume. They are trapezoidal in shape and can cover either 10° or 20° in  $\phi$ . Furthermore, CSCs have the advantage of featuring both precision muon

636 measurement and muon trigger in a single device.

637

638 The third type of detector used in the CMS muon system are called resistive plate  
 639 chambers (RPCs) and can be found in both the barrel and endcap regions. They consist  
 640 of gaseous parallel-plate detectors capable of providing precise timing information and  
 641 adequate spatial resolution. Because of their excellent time resolution, RPCs provide the  
 642 capability of tagging the time of an ionizing event between 2 consecutive LHC bunch  
 643 crossings (BX) in a much shorter time ( $\sim 1$  ns) than the interval between the BXs (25  
 644 ns). For this reason, an RPC-based dedicated muon trigger device can be implemented to  
 645 unambiguously identify the relevant BX to which a muon track is associated with, despite  
 646 the high rate of events and background expected at the LHC.

### 647 3.2.6 Trigger and Data Acquisition

648 Due to the vast volume of data originating from the proton-proton collisions (deliv-  
 649 ered by the LHC at a rate of 40 MHz), a method of eliminating the majority of the un-  
 650 interesting/unwanted events was a requisite for the CMS detector design. This event rate  
 651 reduction is achieved by the implementation of the so-called trigger system, which man-  
 652 ages to select the potentially interesting interactions and reduce the rate from a staggering  
 653  $40 \text{ TBs}^{-1}$  to a manageable value of just a few hundred Hz.

654

655 The CMS trigger system is implemented using a two-stage rate reduction, which com-  
 656 bines both a hardware and software phase. The combination of both of these triggers is  
 657 designed to reduce the rate by a factor of  $\sim 10^6$ . The first stage used in the rate reduction  
 658 is purely hardware based and it is called the Level 1 (L1) Trigger [31], which consists of  
 659 both Field Programmable Gate Array (FPGA) and Application Specific Integrated Circuit  
 660 (ASIC) technology. During this initial stage, the rate is reduced to about 100 kHz with  
 661 a latency of  $3.2 \mu\text{s}$ . This time interval constrains the trigger decision, allowing only for  
 662 data from the calorimeters and muon system to be processed. Trigger primitives (TP)

663 from these subsystems are processed through a series of steps before the combined event  
664 information is evaluated by the global trigger (GT) where the final decision, whether or  
665 not to accept the event, is made.

666

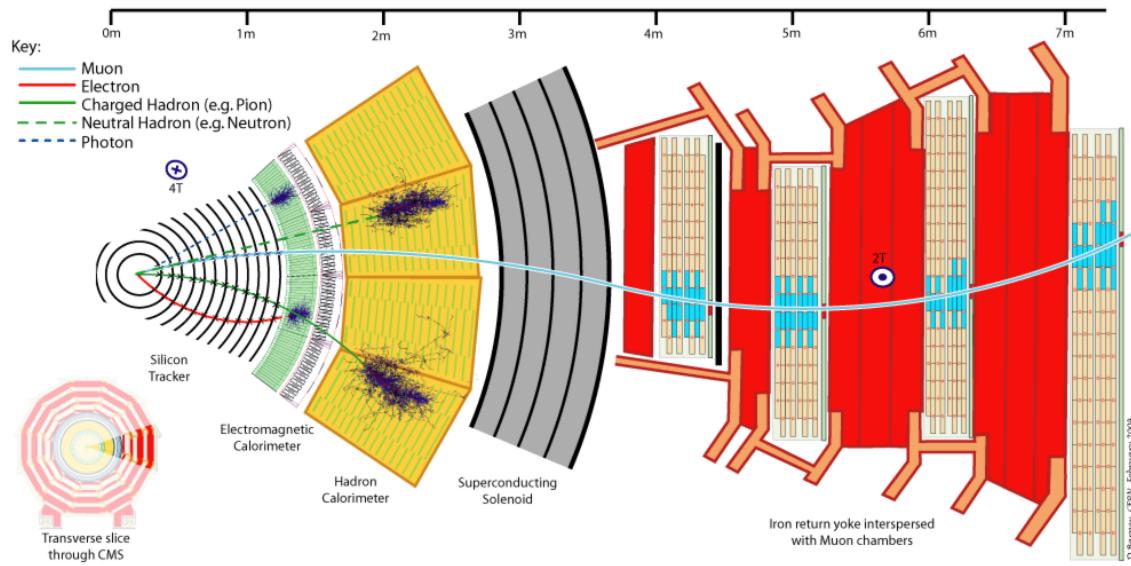
667 The second stage, which implements offline-quality reconstruction algorithms in its  
668 event selection, is referred to as the High-Level Trigger (HLT) [32]. The event selection  
669 process for the HLT requires that physics objects for each event, such as electrons, muons  
670 and jets, are reconstructed and undergo predetermined identification criteria.

### 671 **3.2.7 Event Reconstruction**

672 In order to reconstruct the events the particle flow (PF) algorithm [33] is used. This  
673 algorithm gathers information from all the CMS sub-detectors to reconstruct charged and  
674 neutral hadrons, photons, muons, and electrons. It relies on an efficient and pure track  
675 reconstruction, a clustering algorithm able to distinguish overlapping tracks originated  
676 from different vertices, and an efficient link procedure to associate each particle deposit  
677 in the sub-detectors. Once all the deposits of a particle are associated, it can be correctly  
678 identified and its four-momentum optimally determined from the combined information  
679 of the sub-detectors. The resulting list of particles are then used to reconstruct higher level  
680 objects such as jets, taus, missing transverse energy, and to compute charged lepton and  
681 photon isolation, etc [34]. The CMS experiment is provided with millions of collisions  
682 per second, which need to be triggered, detected, stored and analyzed in a collaboration of  
683 several thousand physicists. This huge amount of data and the complexity of the detector  
684 require a flexible data model that serves all the needs of the collaboration. The data for-  
685 mat is optimized for performance and flexibility of the reconstruction for the end user's  
686 analysis.

687

688 Event information from each step in the simulation and reconstruction chain is logi-  
689 cally grouped into what is called a data tier [35]. From the physicist's point of view the



**Figure 3.8:** Transverse slice of the CMS detector, showing the individual detector subsystems and particle signatures in each. The particle type can be inferred by combining the detector response in the different subdetectors [7].

most important data tiers are RECO, which contains all reconstructed objects and hits, and AOD (a subset of RECO). The AOD will contain a copy of all the high-level physics objects (such as muons, electrons, taus, etc.) and enough information about the event to support all the typical usage patterns of a physics analysis. It also contains a summary of the RECO information sufficient to support typical analysis actions such as track refitting with improved alignment or kinematic constraints, re-evaluation of energy and/or position of ECAL clusters based on analysis-specific corrections. The format of each data tier is ROOT [36]. ROOT is a framework for data processing developed at CERN with the sole purpose of aiding high energy physics research. The various AOD datasets are stored worldwide at various data tier centers. From the AOD's the analysis groups create data structures called NTuples containing only the high-level physics objects needed for their particular analysis.

### 3.2.8 Future Upgrade of Pixel Detector

After the first LHC shutdown called LS1 (2013-2014), and the installation of the Phase-1 Pixel Detector [37] in early 2017, among other things, the LHC is planning another series of upgrades during two major shutdowns, called LS2 and LS3, currently

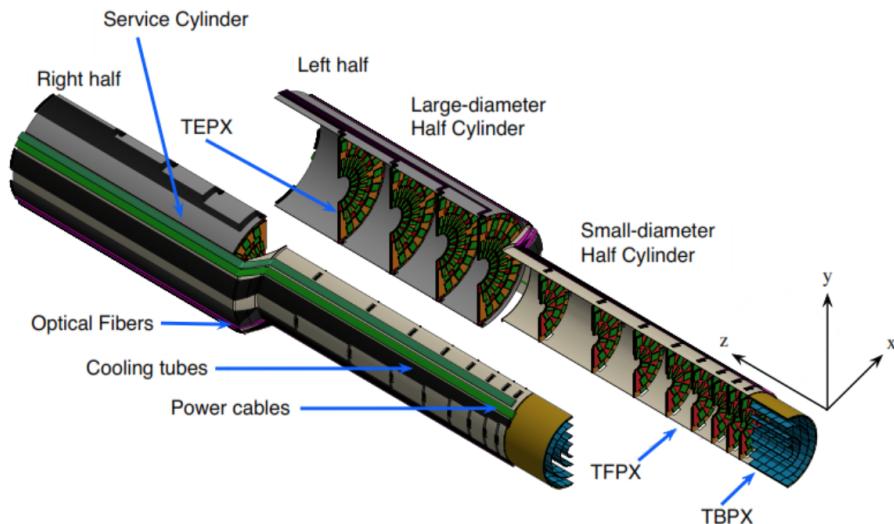
planned for 2019-2020 and  $\sim 2024$ , respectively. LS2 would result in a further increase of the luminosity beyond the original design value, to over  $2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . With the LS3 upgrade of the LHC (called HL-LHC [38]) the luminosity is expected to reach up to  $7.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . Correspondingly, the CMS collaboration has planned a series of further upgrades [39, 40] that will ensure the capabilities of the CMS detector to match to the HL-LHC running conditions, while taking the opportunity to improve the performance and repair any problems uncovered during the data-taking periods. The UPRM group will continue its involvement in the Phase-1 Pixel Detector operations and Phase-2 Pixel Upgrade design.

715

The HL-LHC conditions of instantaneous peak luminosities of up to  $7.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and an integrated luminosity of the order of  $300 \text{ fb}^{-1}$  would result in 1 MeV neutron equivalent fluence of  $2.3 \cdot 10^{16} \text{ neq/cm}^2$  and a total ionizing dose (TID) of 12MGy (1.2 Grad) at the center of CMS, where its innermost component, the Phase-2 Pixel Detector will be installed. The detector should be able to withstand the above radiation dose, handle projected hit rates of  $3\text{GHz}/\text{cm}^2$  at the lowest radius, be able to separate and identify particles in extremely dense collision debris, deal with a pileup of 140-200 collisions per bunch crossing and have a high impact parameter resolution. This translates into requiring a detector design that is highly granular, has thinner sensors and smaller pixels, and faster, radiation hard electronics compared to its Phase-1 counterpart. The selection of interesting physics events at the Level-1 (L1) trigger and inefficiency of selection algorithms in high pileup conditions further require the Tracker to be included in this trigger stage, helping reduce the event rate from 40 MHz rate to 7.5 kHz. The physics goals also require an increase in Pixel Detector coverage to  $|\eta| = 4.0$  which improves the  $p_T^{miss}$  resolution and particle-flow event reconstruction by providing  $p_T$  measurements and trajectories for charged particles entering the calorimeters. The  $p_T^{miss}$  resolution is an essential performance parameter for many BSM physics searches including SUSY and extra dimension models where particles escape undetected from the detector space. The smaller pixel size will further improve b-tagging as well as hadronic reconstruction and track reconstruction

efficiencies within boosted jets, which can be produced from new heavy objects decaying to Higgs, Z bosons, or top quarks – all heavy probes that can be exploited for new physics searches. Improving the b-tagging capabilities directly affects our analysis due to the importance of properly identifying top quark decays.

739



**Figure 3.9:** Phase-2 Pixel Detector Layout[ref].

The Phase-2 Pixel Detector [41, 42] baseline design comprises a barrel part with 4 layers of Tracker Barrel Pixel Detector (TBPX), 8 small double-discs per side of Tracker Forward Pixel Detector (TFPX) and 4 large double-discs per side of Tracker Endcap Pixel Detector (TEPX). This forward and end part is referred to as 8l4s (8 TFPX and 4 TEPX). In the TBPX the pixel modules are arranged in “ladders”. In each layer, neighboring ladders are mounted staggered in radius, so that  $r\text{-}\phi$  overlap between the ladders is achieved. The modules on a ladder do not overlap in  $z$ . A projective gap at  $\eta = 0$  is avoided by mounting an odd number of modules along  $z$ , and by splitting the barrel mechanics in  $z$  into slightly asymmetric halves. In TFPX and TEPX the modules are arranged in concentric rings. Each double-disc is physically made of two discs, which facilitates to mount modules onto four planes, with overlaps in  $r$  as well as  $r\text{-}\phi$ . Each disc is split into two halves, and these D-shaped structures are referred to as “dees”. The TEPX will provide the required luminosity measurement capability, by an appropriate implementation of the

753 readout architecture. In total, the pixel detector will have an active surface of approxi-  
754 mately  $4.9 \text{ m}^2$ . [Figure 3.9](#) shows the layout of the Phase-2 Pixel Detector.

755 **Chapter 4**

756 **Optimizing the Phase-2 Pixel Geometry  
757 for the Level-1 Trigger in the Outer  
758 Tracker**

759 **4.1 Upgraded Phase-2 Tracker**

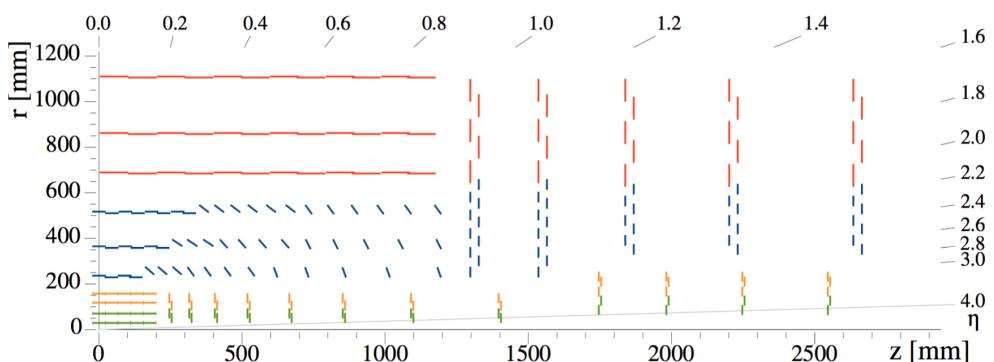
760 This chapter is dedicated to a study that seeks to optimize the geometry of the Phase-2  
761 Pixel Detector (for the HL-LHC) in terms of how the upgraded Inner Tracker (IT) af-  
762 fects the overall stub production (the concept of a stub is defined in [subsection 4.1.2](#)).  
763 The goal of this study was to determine how the removal of discs from the Phase-2 For-  
764 ward Pixel detector would affect the production of L1 stubs in the upgraded CMS Outer  
765 Tracker (OT). More information of the Phase-2 Tracker can be found at [\[43\]](#). This study  
766 has a direct impact on the future of SUSY analyses due to the importance of the tracker  
767 in searches for BSM physics.

768  
769 In the HL-LHC era (2023-2035) an integrated luminosity of  $3,000 \text{ fb}^{-1}$  is expected to  
770 be reached at a center-of-mass energy of 14 TeV. This data will enable us to make precise  
771 measurements of several SM parameters (particularly the scalar sector). This will also  
772 lay the groundwork for new physics searches proposed by many of the BSM models. In

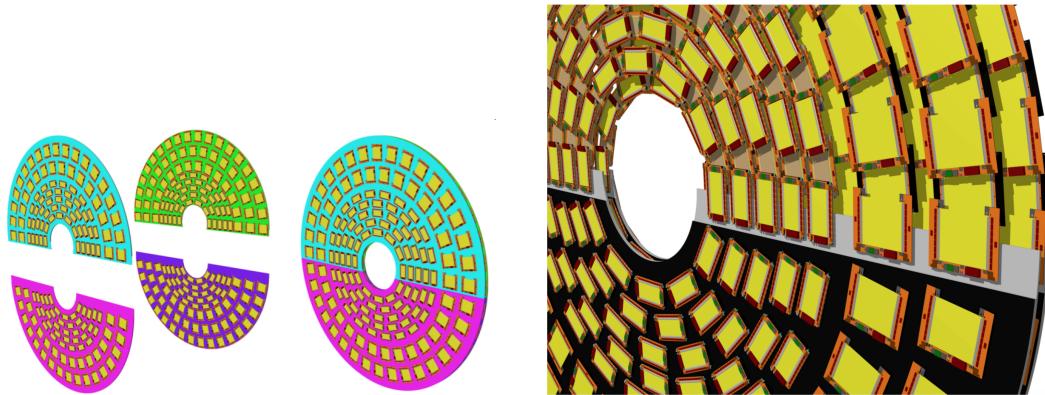
addition, the HL-LHC project is essential to identify any hints of new physics, including SUSY, that could possibly be observed during the LHC Run 2 or 3, or beyond. The bigger, and highly granular pixel and silicon tracker, and its extended  $\eta$  coverage, will improve jet sub-structure, boosted object identification and heavy object tagging to hunt for BSM physics, specifically Supersymmetry. A discovery could happen with various few sigma excesses and then interpretation of the new signal will require a large amount of data. At the HL-LHC, much gain is expected for weakly produced SUSY particles because they are characterized by small production cross-sections. Considering the case that gluinos and squarks are heavy, search for electroweak production channels will offer the best option to discover SUSY. At the HL-LHC, charginos/neutralinos are expected to be discovered (significance of  $5\sigma$ ) up to masses of about 950 GeV. In the case of strong production, the mass reach for gluinos goes up to 2.2 TeV. This study is a contribution in the potential future discovery of SUSY and other BSM physics.

#### 4.1.1 Phase-2 Tracker Layout

The Phase-2 tracker design features an Inner Tracker based on silicon pixel modules and an Outer Tracker made from silicon modules with strip and macro-pixel sensors. The IT is composed of a Tracker Barrel Pixel Detector (TBPX) with four barrel layers, and 8 small Tracker Endcap Pixel Detector (TEPX) discs plus 4 Tracker Forward Pixel Detector (TFPX) large discs in the forward direction.

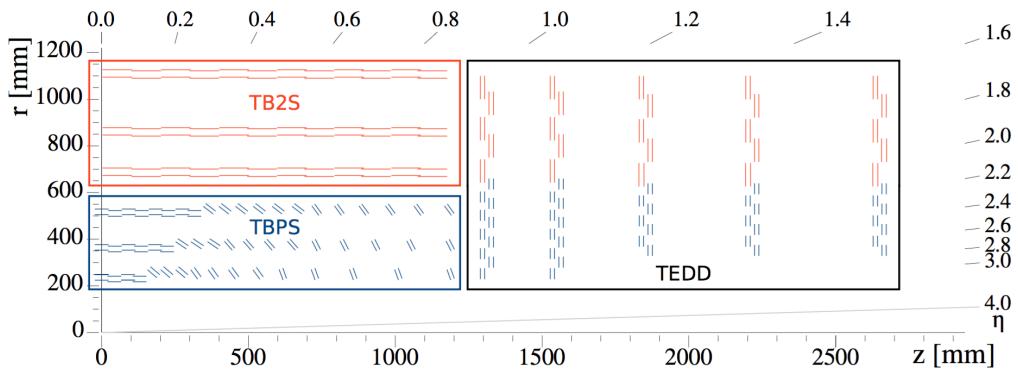


**Figure 4.1:** Sketch of one quarter of the tracker layout in r-z view. In the Inner Tracker the green lines correspond to pixel modules made of two readout chips and the yellow lines to pixel modules with four readout chips. In the Outer Tracker the blue and red lines represent the two types of modules  $p_T$  modules - (2S and PS).



**Figure 4.3:** Sketch of four dees (left) forming a double-disc (centre), and drawing of a part of a TEDD double-disc (right), illustrating the overlap of modules in  $\phi$  and  $z$ . The upper two dee support structures are removed in order to show all layers of modules. The modules ob the disc are arranged in 15 or 12 rings.

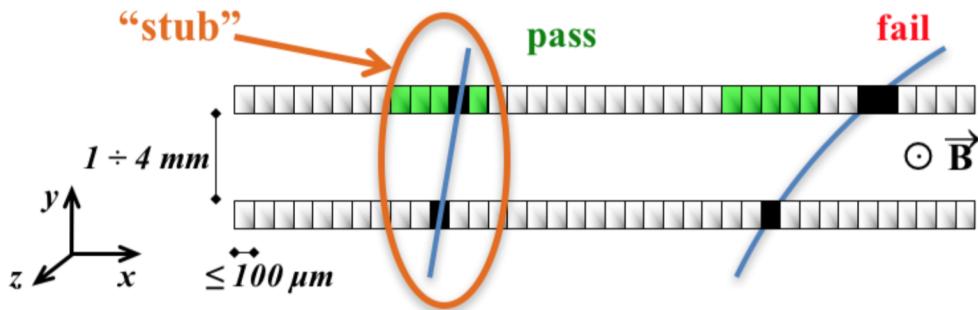
792     The OT will feature so-called “ $p_T$  modules”, which are capable of providing tracking  
 793 information for the L1 trigger. The OT consists of six barrel layers (TBPS and TB2S) and  
 794 five endcap double-discs (TEDD) which are composed of two types of  $p_T$  modules, called  
 795 2S and PS modules. The first two endcap double-discs contain a total of 15 rings and the  
 796 remaining three feature 12 rings each. This is shown in [Figure 4.2](#).



**Figure 4.2:** Sketch of one quarter of the Outer Tracker in  $r$ - $z$  view. Blue (red) lines represent PS (2S) modules. The three sub-detectors, named TBPS, TB2S, and TEDD, are indicated. All overlapping layers are shown separately.

### 797     4.1.2 Tracker Input to the L1-Trigger

798     The enhancement of the trigger performance involves both a higher output rate of in-  
 799 teresting events and an improved discriminating power of the event selection, which is  
 800 more challenging in a high-pileup environment. Improved discriminating power will be  
 801 achieved by using more information in the trigger decision, with a longer latency available



**Figure 4.4:** Correlation of signals in closely-spaced sensors enables rejection of low- $p_T$  particles; the channels shown in green represent the selection window to define an accepted stub.

802 for its processing. The use of tracking information in the L1 trigger will improve the  $p_T$   
 803 resolution of various objects at L1 (e.g. jets), it will allow the exploitation of information  
 804 on track isolation, and will contribute to the mitigation of pileup of 140-200 collisions per  
 805 bunch crossing. CMS plans to enhance the first level trigger rate from presently 100 kHz  
 806 to 750 kHz and to increase the latency from the present value of  $3.2 \mu s$  to  $12.5 \mu s$ . The  
 807 front-end electronics and the L1 trigger track reconstruction need to comply with these  
 808 new requirements. The necessity of providing tracking information to the L1 trigger is a  
 809 main driver for the design of the OT, including its module concept. The use of tracking  
 810 information in the L1 trigger implies that the tracker has to send out self-selected infor-  
 811 mation at every bunch crossing. Such functionality relies upon local data reduction in the  
 812 front-end electronics, in order to limit the volume of data that has to be sent out at 40 MHz.  
 813 This is achieved with modules that are capable of rejecting signals from particles below  
 814 a certain  $p_T$  threshold, referred to as “ $p_T$  modules” [43]. Tracks from charged particles  
 815 are bent in the transverse plane by the 3.8 T field of the CMS magnet, with the bending  
 816 angle depending on the  $p_T$  of the particle. The modules are composed of two single-sided  
 817 closely-spaced sensors read out by a common set of front-end ASICs that correlate the  
 818 signals in the two sensors and select the hit pairs (referred to as “stubs”) compatible with  
 819 particles above the chosen  $p_T$  threshold (as shown in [Figure 4.4](#)). Stubs are defined as  
 820 matching pairs of hits on the adjacent silicon layers of the  $p_T$  modules. Whether or not a  
 821 stub is accepted depends on the  $p_T$  of the particle for a certain  $p_T$  threshold value.

## 4.2 Tracker Stub Simulation

The baseline detector design comprises a barrel part with 4 layers of TBPX, 8 small double-discs per side of TFPX and 4 large double-discs per side of TEPX. This forward and end part is referred to as  $8l4s$  (8 TFPX and 4 TEPX). Any changes made to the detector design or its geometry, due to any circumstance, directly impacts the physics variables, such as tracking efficiency, fake rate,  $p_T$  resolution, b-tagging and primary track reconstruction and needs to be simulated to understand its impact. The removal of discs from the Phase-2 Forward Pixel detector (TFPX/TEPX) would affect the production of L1 stubs in the upgraded CMS OT. A decrease in the total number of stubs is expected when less material is present due to secondary interactions.

832

To perform this study we considered the following four different geometries of the  
834 Pixel Detector:

- 835 • 8 small discs and 4 large discs ( $8s4l$ ) – designated as the default geometry
- 836 • 8 small discs and 3 large discs ( $8s3l$ ) - one large disc removed.
- 837 • 7 small discs and 4 large discs ( $7s4l$ ) - one small disc removed.
- 838 • 6 small discs and 3 large discs ( $6s3l$ ) - one large disc and two small discs removed.
- 839 • A version of the default pixel geometry with a dead disc (used to verify our find-  
840 ings).

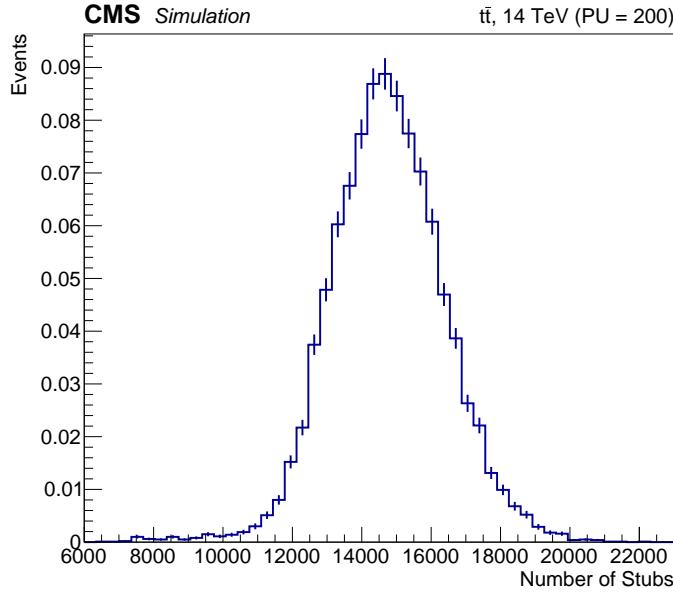
841 The response of the detector due to design modifications is simulated using the offi-  
842 cial CMS software package CMSSW. The detector geometry is generated with a design  
843 tool for innovative silicon tracking detectors called tkLayout [44] and is exported into  
844 the standard CMSSW format along with the materials like sensing silicon elements, elec-  
845 tronics and inactive material. The exported tracker geometry serves as input for the full  
846 simulation and reconstruction performance studies with CMSSW. The design of the CMS

| <b>TEDD1</b> | <b>TEDD2</b> | <b>TEDD3</b> | <b>TEDD4</b> | <b>TEDD5</b> |            |
|--------------|--------------|--------------|--------------|--------------|------------|
| 5.5          | 5.5          | 6.0          | 6.5          | 6.5          | <b>R1</b>  |
| 4.5          | 5.0          | 5.5          | 6.0          | 6.0          | <b>R2</b>  |
| 3.5          | 4.0          | 5.0          | 6.0          | 6.0          | <b>R3</b>  |
| 3.0          | 3.0          | 2.5          | 3.0          | 3.0          | <b>R4</b>  |
| 2.5          | 2.0          | 3.5          | 3.0          | 3.0          | <b>R5</b>  |
| 4.0          | 3.0          | 2.5          | 2.0          | 2.0          | <b>R6</b>  |
| 3.5          | 3.0          | 2.5          | 2.0          | 2.0          | <b>R7</b>  |
| 3.0          | 3.0          | 2.5          | 2.0          | 2.0          | <b>R8</b>  |
| 3.0          | 2.5          | 2.0          | 2.0          | 1.5          | <b>R9</b>  |
| 2.5          | 2.0          | 2.0          | 1.5          | 1.5          | <b>R10</b> |
| 2.0          | 2.0          | 1.5          | 1.5          | 1.5          | <b>R11</b> |
| 2.0          | 2.0          | 1.5          | 1.0          | 1.0          | <b>R12</b> |
| 1.5          | 1.5          |              |              |              | <b>R13</b> |
| 1.5          | 1.5          |              |              |              | <b>R14</b> |
| 1.0          | 1.0          |              |              |              | <b>R15</b> |

|              |           |           |           |           |           |           |           |           |           |            |            |            |
|--------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|
| <b>TB2S3</b> | 6.5       |           |           |           |           |           |           |           |           |            |            |            |
| <b>TB2S2</b> | 5.5       |           |           |           |           |           |           |           |           |            |            |            |
| <b>TB2S1</b> | 4.5       |           |           |           |           |           |           |           |           |            |            |            |
| <b>TBPS3</b> | 3.5       | 4.5       | 4.5       | 4.0       | 4.0       | 4.0       | 4.0       | 3.5       | 3.5       | 3.5        | 3.0        | 3.0        |
| <b>TBPS2</b> | 2.0       | 3.0       | 3.0       | 3.0       | 3.0       | 2.5       | 2.5       | 2.5       | 2.5       | 2.5        | 2.5        | 2.0        |
| <b>TBPS1</b> | 2.0       | 2.5       | 2.5       | 2.5       | 2.5       | 2.5       | 2.5       | 2.0       | 2.0       | 1.5        | 1.5        | 1.0        |
| <b>FLAT</b>  | <b>R1</b> | <b>R2</b> | <b>R3</b> | <b>R4</b> | <b>R5</b> | <b>R6</b> | <b>R7</b> | <b>R8</b> | <b>R9</b> | <b>R10</b> | <b>R11</b> | <b>R12</b> |

**Table 4.1:** Stub window tunings for CMSSW\_9\_3\_2. Shown are the SW values for TEDD (top table) and for the OT barrel (below). Obtained from [this github repository](#).

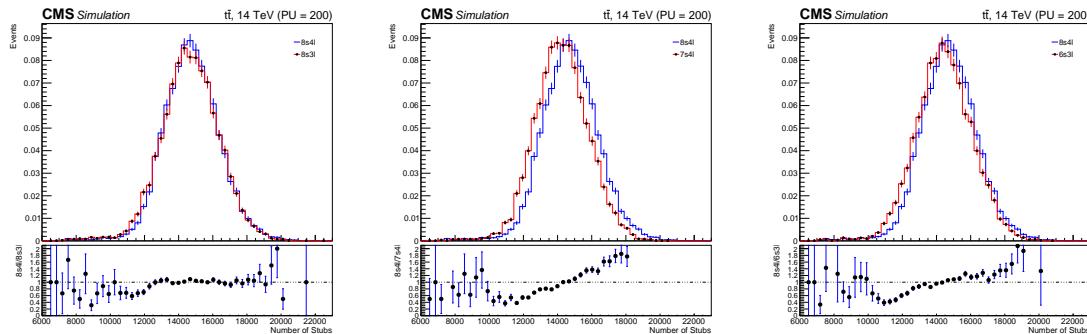
Phase-2 Tracker includes Tracker information for L1 triggering. The study was made from  $t\bar{t}$  samples with a center-of-mass energy of 14 TeV and 200 pileup. The samples were created with the CMS software version referred to as CMSSW\_9\_3\_2. The stub information is used as input for the track fitting algorithm and allows tracks used for L1 trigger to be reconstructed within approximately  $5 \mu s$  (this includes a  $1 \mu s$  estimate for the time required to transmit the stub data from the detector to the counting room) to remain within the overall L1 trigger latency budget of  $12.5 \mu s$ . Stubs are formed from two closely spaced sensors in Tracker measurement layers and provide linked position measurements for each particle passing through them. Table 4.1 shows the stub window  $p_T$  threshold values used for the study.



**Figure 4.5:** Number of stubs per event for standard pixel geometry (8s4l).

### 857 4.2.1 Results of Effect of the Phase 2 Pixel Detector on the L1 Stub 858 Rate

859 (a) **Number of Stubs:** With the current (above table) stub window tuning, the total  
860 number of stubs are on average  $\sim 15,000$  stubs as shown in [Figure 4.5](#).

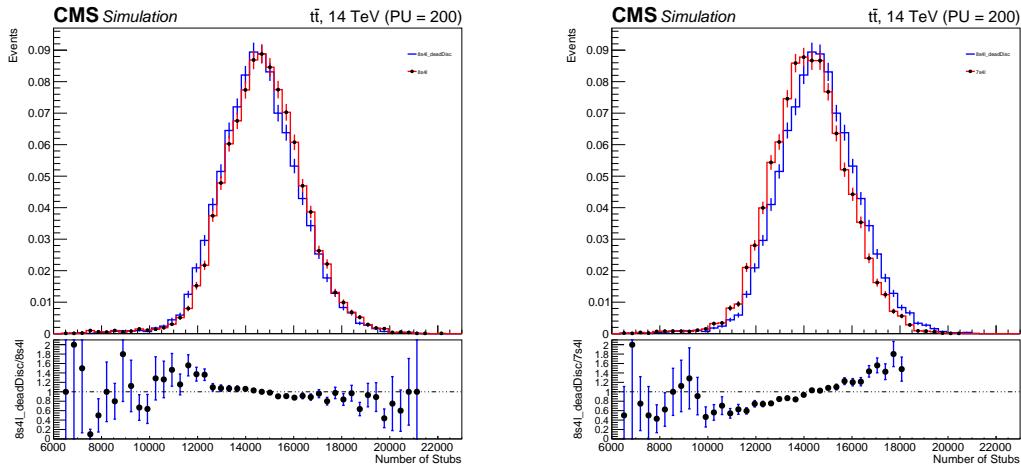


**Figure 4.6:** Comparison of total number of stubs per event between the standard layout and the other three considered. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right).

861 (b) **Comparison of Stubs between default Pixel Geometry (8s4l) to disc removed:**  
862 A comparison of the total number of stubs per event between the standard geometry to  
863 the others that have a disc or more removed (8s3l, 7s4l, 6s3l) was performed ([Figure 4.6](#)).

Out of the three comparisons shown, the largest shift in the stub distribution can be seen from the  $7s4l$  geometry. It was found that the removal of a smaller pixel disc has a more pronounced effect on stub production in the OT than a larger pixel disc.

867



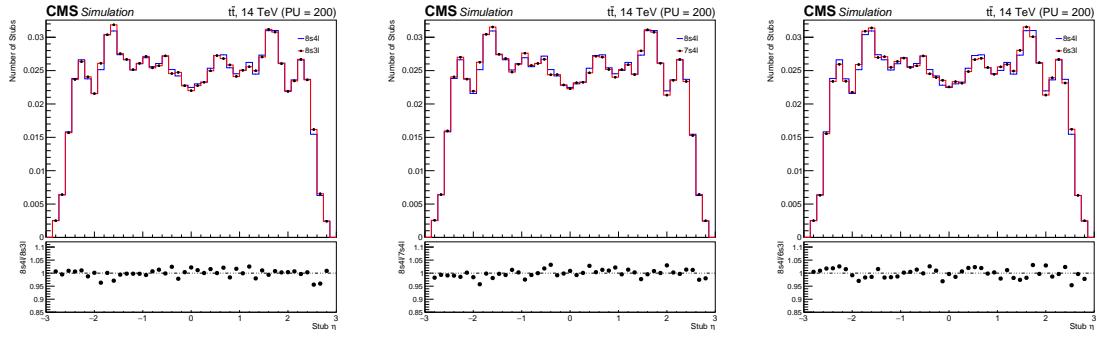
**Figure 4.7:** Comparison for number of stubs per event between pixel geometry  $8s4l$  with a dead disc (blue) and both pixel geometry  $8s4l$  (right plot) and  $7s4l$  (left plot) (blue). Both the standard geometry and  $7s4l$  are shown in blue in their respective plots.

868 (c) **Comparison with a disc dead in default geometry:**

An additional study, using pixel geometry  $8s4l$  with a dead disc, was made to verify the effect observed on the L1 stub counting. Figure 4.7 (right) shows a comparison between 2 geometries that are (in principle) identical but one detects no hits on the second disc in the positive side. This Figure shows a slight deviation from the original  $8s4l$  geometry.

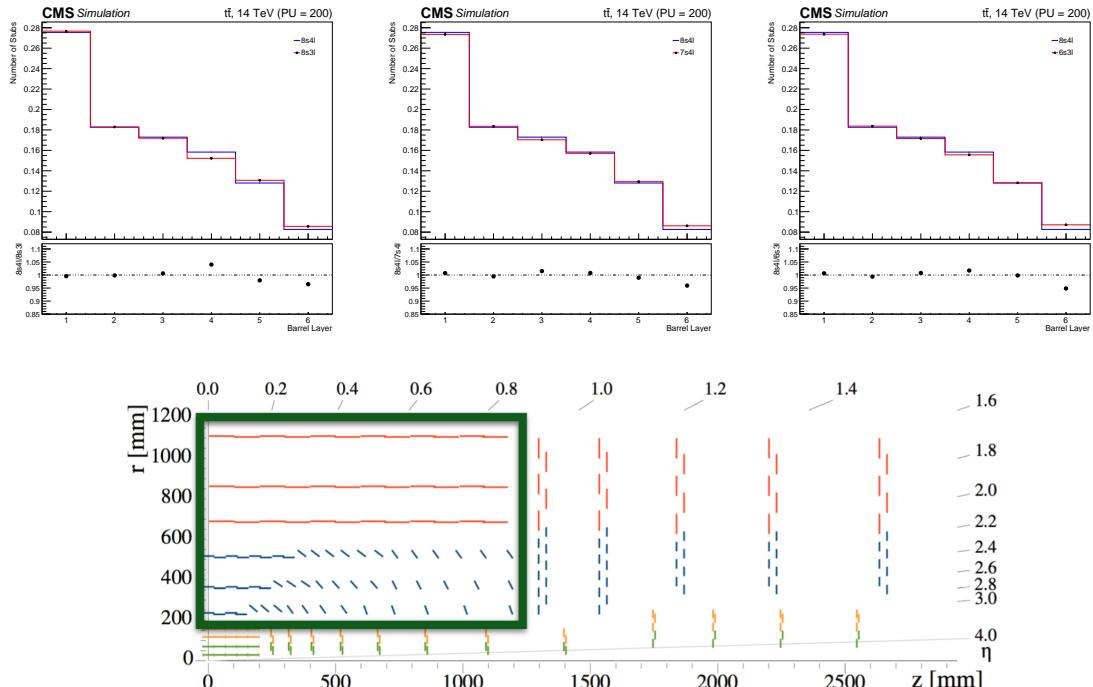
874

Comparing the number of stubs for the geometry with the dead disc (blue) to the one with the removed small disc (red), the same behavior (Figure 4.7, left) is observed as before. As was expected, the stub distribution for pixel geometry  $8s4l$  has a higher mean value than  $7s4l$ , despite the dead disc. The removal of discs in the pixel detector causes the total number of stubs to decrease due to secondary interactions with the material. We also studied the effect of the pixel geometries on the stubs with respect to the pseudorapidity ( $\eta$ ) as shown in Figure 4.8.

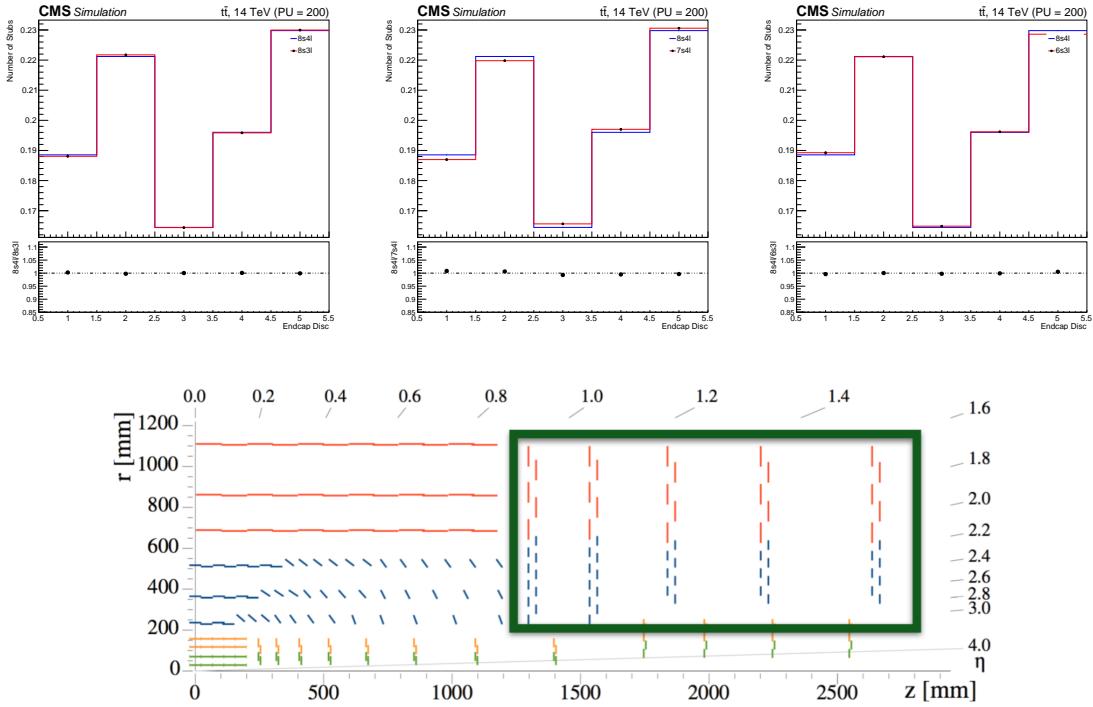


**Figure 4.8:** Comparison of total number of stubs per  $\eta$  value between the standard layout and the other three layouts considered. All four different layouts exhibit the same distribution. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right).

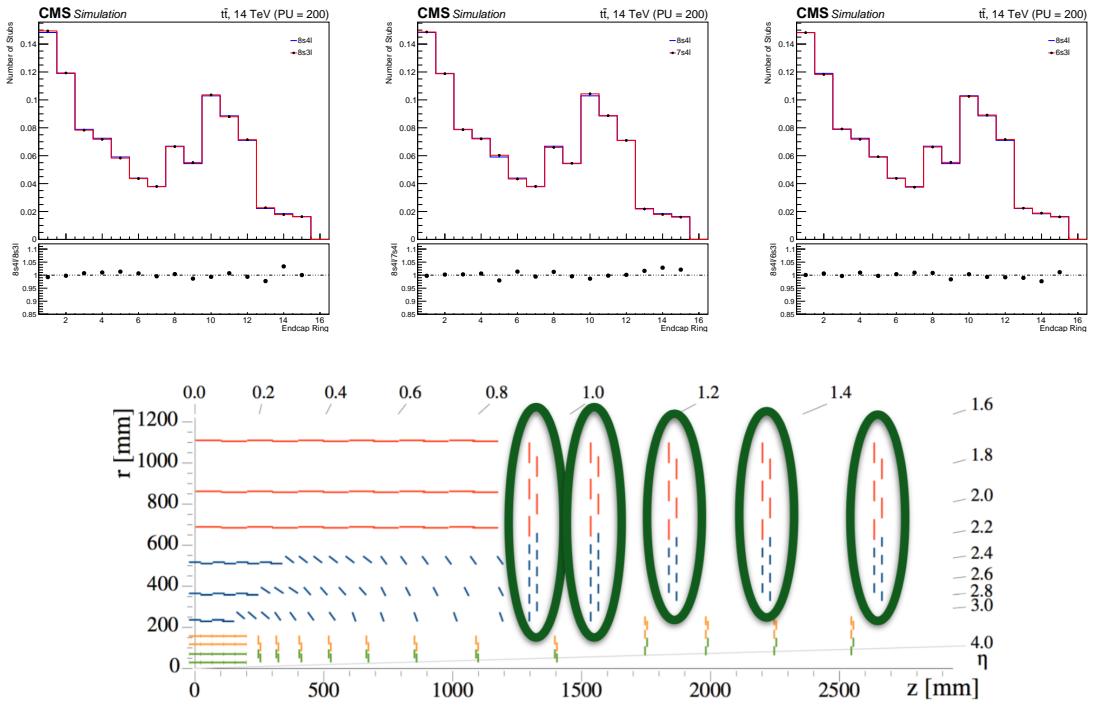
882 We further studied the effect of the pixel geometries on the stubs per barrel (Figure 4.9), end discs (Figure 4.10) of the OT and per ring of the endcap discs (Figure 4.11).  
 883



**Figure 4.9:** Comparison of total number of stubs per each of the 6 barrel layers. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right). Tracker Barrel layers shown below.



**Figure 4.10:** Comparison of total number of stubs per each of the 5 Endcap Double Discs. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right). Tracker Endcap Discs shown below.



**Figure 4.11:** Comparison of total number of stubs for all rings across the five endcap double-discs. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right). The double discs are shown below.

**4.3 Results**

884        The above study shows that the removal of discs in the pixel detector affects the num-  
885        ber of stubs formed in the OT. A reduction in the amount of material (i.e. removal of  
886        discs) translates into a reduced amount of secondary particle interactions and hence, a de-  
887        crease in the net amount of stubs is observed. In other words more material (more discs)  
888        in the pixel layer leads to more secondaries, increasing the total amount of stubs in a given  
889        sample. However, the effect is small and a detailed study with respect to various  $\eta$  regions  
890        is planned in the future, as stub formation is also sensitive to  $\eta$  windows. Knowing the  
891        origin of particles causing stubs is also an important area to look into.

893 **Chapter 5**

894 **Search for SUSY in the 0-Lepton Final**

895 **State**

896 **5.1 Introduction**

897 An overview of a search for Supersymmetry in the all-hadronic channel with missing  
898 transverse momentum and a customized top tagger is presented. The data used for the  
899 analysis was collected in 2016 by the CMS experiment at the LHC, and corresponds to  
900 an integrated luminosity of  $35.9 \text{ fb}^{-1}$  from proton-proton collisions at a center-of-mass  
901 energy ( $\sqrt{s}$ ) of 13 TeV. The results and procedures presented are the product of an analysis  
902 conducted in collaboration with the hadronic SUSY analysis team based at Fermilab in  
903 Batavia, IL, USA. The published results can be obtained from [45].

904 **5.2 Analysis Description**

905 The study consists of an inclusive search for events with final states that contain miss-  
906 ing transverse momentum ( $p_T^{miss}$ ) and reconstructed top quarks. The signal models used  
907 in this study include the production of three different types of SUSY particles. Two of  
908 which are the top squark ( $\tilde{t}$ ) and the gluino ( $\tilde{g}$ ), the supersymmetric partners of the SM top  
909 quark (t) and gluon (g). The third one is the neutralino ( $\tilde{\chi}_1^0$ ), considered to be the lightest

910 SUSY particle (LSP) under the Minimal Supersymmetric Standard Model (MSSM), and  
 911 a possible candidate for Dark Matter.

912

913 The signal data is selected by requiring events to have a minimum amount of jets ( $N_j$ )  
 914 and b-jets ( $N_b$ ) as well as a large  $p_T^{miss}$ . Other factors, such as the number of top jets  
 915 ( $N_t$ ), the scalar sum of the jet transverse momentum ( $H_T$ ), and the so-called stransverse  
 916 mass ( $m_{T2}$ )<sup>1</sup> are also considered in order to select events of interest. The search region is  
 917 then defined in exclusive bins of  $N_t$ ,  $N_b$ ,  $H_T$ ,  $p_T^{miss}$  and  $m_{T2}$ . Additional discriminatory  
 918 variables used in the analysis include the azimuthal angle between  $p_T^{miss}$  and the leading  
 919 jets of an event ( $\Delta\phi$ ), the pseudorapidity ( $\eta$ ) and the cone radius size (defined as  $\Delta R =$   
 920  $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ ) used in the isolation of various physics objects.

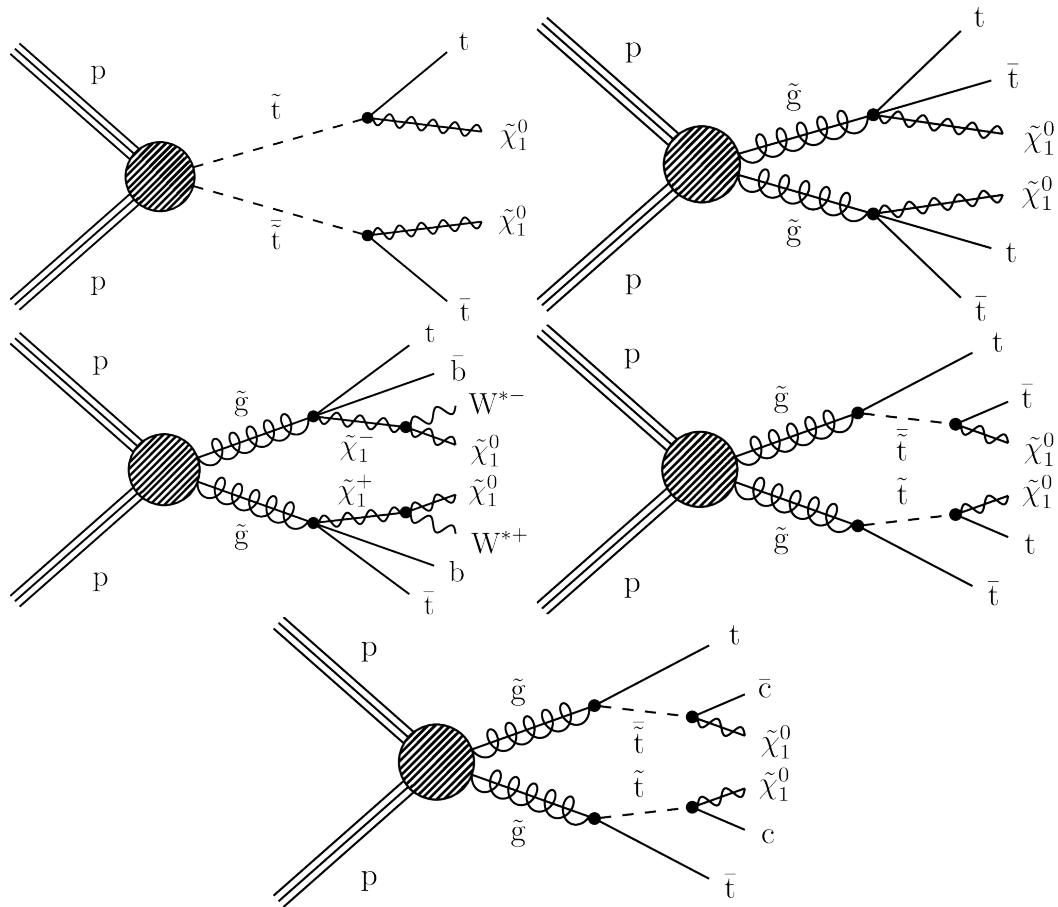
### 921 5.2.1 Signal Models

922 The signal models used in this analysis are based on two different scenarios: direct  
 923 top squark production and gluino mediated top squark production. In the former sce-  
 924 nario we consider one decay within the Simplified Model Spectra (SMS) [46] framework  
 925 called ‘T2tt’ (Figure 5.1, top left), where a top squark pair is produced directly from the  
 926 proton-proton collision and then decays into a pair of top quarks and a pair of neutralinos  
 927 ( $\tilde{t} \rightarrow t\tilde{\chi}_1^0$ ).

928

---

<sup>1</sup> $m_{T2}$  is a minimization of two transverse masses ( $m_T$ ) with a constraint that the sum of the transverse momenta of both  $\tilde{\chi}_1^0$ 's is equal to the transverse momentum of the event, i.e.,  $\vec{p}_T = \vec{q}_T^{(1)} + \vec{q}_T^{(2)}$ . The mathematical definition is given by  $m_{T2} \equiv \min_{\vec{q}_T^{(1)} + \vec{q}_T^{(2)} = \vec{p}_T} \left[ \max \left\{ m_T^2(\vec{p}_T^{(1)}, \vec{q}_T^{(1)}; m_{\tilde{\chi}_1^0}), m_T^2(\vec{p}_T^{(2)}, \vec{q}_T^{(2)}; m_{\tilde{\chi}_1^0}) \right\} \right]$ .



**Figure 5.1:** Simplified model diagrams show of the various Supersymmetry signals considered in this analysis: the T2tt model (top left), the T1ttt model (top right), the T1ttbb model (middle left), the T5tttt (middle right), and the T5ttcc model (bottom).

929 In the case of gluino mediated top squark production there are several different de-  
 930 cay modes considered. The T1ttt model (Figure 5.1, top right) consists of two gluinos  
 931 decaying into a pair of top quarks and a neutralino ( $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ ), accounting for situations  
 932 in which the top squark is too heavy to be produced directly but the gluino is not. In the  
 933 T1ttbb model (Figure 5.1, middle left), a pair of gluinos to either a top or bottom squark  
 934 as  $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$  (25%),  $\tilde{g} \rightarrow b\bar{b}\tilde{\chi}_1^0$  (25%), or  $\tilde{g} \rightarrow \bar{t}b\tilde{\chi}_1^\pm$  (50%), where  $\tilde{\chi}_1^\pm$  denotes the lightest  
 935 positive ( $\tilde{\chi}_1^+$ ) or negative ( $\tilde{\chi}_1^-$ ) chargino. Due to the small difference in mass between the  
 936  $\tilde{\chi}_1^+$  (or its conjugate) and the LSP ( $\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0) = 5$  GeV), the two particles are taken  
 937 to be nearly mass degenerate. The T1ttbb model provides sensitivity to situations where  
 938 there are mixed states of top and bottom squarks.

939

940 The T5ttt model consists of a pair of gluinos, where each decays into a top quark and

941 an on-shell top squark. The top squark then decays into a top quark and a neutralino. For  
 942 this model, a mass difference of  $\Delta m(\tilde{t}, \tilde{\chi}_1^0) = 175$  GeV between the top squark and the  
 943 LSP is expected. This model provides sensitivity to a region of T2tt where the  $t\bar{t}$  back-  
 944 ground is very similar to the signal and  $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$  is very close to the top squark mass.  
 945 In the T5ttcc model ([Figure 5.1](#), bottom) each gluino decays into a final state consisting  
 946 of a top quark, a charm quark and a neutralino ( $\tilde{g} \rightarrow \bar{t}c\tilde{\chi}_1^0$ ). This model is very similar  
 947 to the T5tttt model with the difference that  $\Delta m(\tilde{t}, \tilde{\chi}_1^0) = 20$  GeV and the on-shell top  
 948 squark decays into a charm quark and the LSP. The T5ttcc model provides sensitivity to  
 949 situations where the on-shell top squark is kinematically forbidden from decaying into a  
 950 top quark.

951

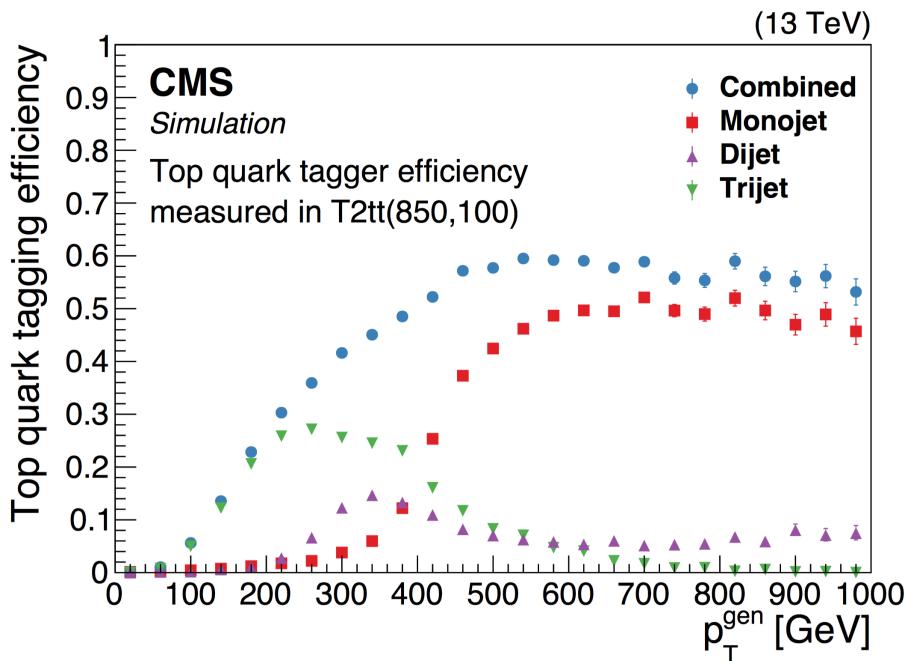
952 All five signal models described share similar final states containing two neutralinos  
 953 and up to four top quarks. Since the neutralino is considered stable and only interacts  
 954 weakly with matter, it cannot be picked up by the detector. Therefore, missing transverse  
 955 momentum ( $p_T^{miss}$ ) is one of the most important variables used to compare event yields  
 956 between the signal and the SM background.

### 957 5.2.2 Top Tagger

958 The distinguishing feature of this analysis is its powerful top-tagging algorithm. It is  
 959 designed to provide high reconstruction efficiency over the full range of the top quark  $p_T$   
 960 for the considered SUSY signal models. This top-tagger combines the use of several dif-  
 961 ferent jet clustering algorithms in order to identify three different categories of top quark  
 962 jets, designated as “monojet”, “dijet” and “trijet”. The tagger makes use of identified  
 963 jets that were reconstructed with the anti- $k_T$  [47] clustering algorithms combined with  
 964 additional corrections and selection criteria provided by the Cambridge–Aachen [48] and  
 965 soft-drop de-clustering [49] algorithms. In addition, multivariate analysis (MVA) tech-  
 966 niques, such as the random forest decision tree algorithm [50], were applied in order to  
 967 further decrease the amount of reconstructed fake tops.

968

969 There are three top-quark jet categories that take into account that top quark decay  
 970 products get closer together as the top quark  $p_T$  becomes higher. For specific  $p_T$  values,  
 971 the decay products of a hadronic process will be reconstructed either as one (“monojet”)  
 972 or two (“dijet”) jets rather than three (“trijet”). The  $p_T$  values for which this is observed  
 973 depends on the size of the jets that are being used. In the case of a highly boosted jet,  
 974 the anti- $k_T$  algorithm is used within a cone size  $\Delta R \sim 0.8$  (called AK8), which captures  
 975 all decay products of the top quark and is reconstructed as a single jet. This technique  
 976 requires that the top quark starts with a  $p_T$  of at least 400 GeV to have the decay products  
 977 fully captured in the 0.8 jet cone. For a  $p_T < 400$  GeV, resolved top-tagging techniques  
 978 are more efficient, which require three jets independently clustered within a cone size of  
 979  $\Delta R \sim 0.4$  (AK4). Both types of algorithms are used in order to obtain a high efficiency  
 980 over a wide range of top quark  $p_T$ .



**Figure 5.2:** Efficiency of the top quark tagger as a function of generator-level top quark  $p_T$  for the monojet (red boxes), dijet (magenta triangles), and trijet (green upside-down triangles) categories and for their combination (blue circles), as determined using T2tt signal events with a top squark mass of 850 GeV and an LSP mass of 100 GeV. The vertical bars indicate the statistical uncertainties.

981 In the case of the dijet top quark decay category, an AK8 jet with a  $p_T > 200$  GeV is  
 982 combined with a loose AK4 jet to form a top quark candidate. Both jets are required to

be within a cone of radius  $\Delta R = 1$ . For the trijet category, three AK4 jets are combined, with all three jets required to be within a radius of  $\Delta R = 1.5$ . Figure 5.2 shows the efficiency of the algorithm for each of the three categories of top jets. The efficiency was determined for the T2tt signal model considering a top squark mass of 850 GeV and an LSP mass of 100 GeV, and is calculated as:

$$\text{Efficiency} = \frac{\text{Number of generator-level top quarks matched to a tagged top}}{\text{Number of generator-level top quarks in an event}}.$$

The top quark candidates are required to be within  $|\eta| < 2.0$ . Furthermore, a cone size of  $\Delta R < 0.4$  is used to match the reconstructed top quark to the generated-level top quark. The average misidentification rate as a function of the  $p_T^{miss}$  is found to be around 20% for simulated  $Z(\nu\bar{\nu})+jets$  events with a selection criteria similar to the one applied to data:  $N_j \geq 4$ ,  $N_b \geq 1$ ,  $p_T^{miss} > 250$  GeV and no isolated electron or muon with  $p_T > 10$  GeV.

### 5.2.3 Lepton and Track Veto

Since the search only involves models with purely hadronic final states, events which contain electrons or muons are vetoed. This requires the isolation of the muon/electron candidates to a cone size of  $\Delta R < 0.2$  for  $p_T \leq 50$  GeV and 0.05 for  $\geq 200$  GeV. This  $\Delta R$  requirement decreases inversely to the lepton  $p_T$  in the range  $50 < p_T < 200$  GeV in order to account for the collimation of a heavy object's decay products as its Lorentz boost increases. The electron and muon objects are considered to be isolated when their relative isolation<sup>2</sup> is less than 0.1 and 0.2, respectively. Due to contributions from pileup, the isolation sum is corrected using an estimate of the amount of pileup energy in the cone.

1004

The events that manage to pass the lepton veto are subjected to undergo an isolated charged-particle track veto. The purpose of this veto is to suppress events which contain  $\tau$  leptons that decay hadronically or misidentified electrons and muons. The track requirements for this veto are  $p_T > 5$  GeV,  $|\eta| < 2.5$  and a relative track isolation less than 0.2.

---

<sup>2</sup>The ratio between the isolation sum and the candidate  $p_T$ .

1009 In addition, the isolated track veto is only applied when the transverse mass of an isolated  
 1010 track- $\vec{p}_T^{miss}$  is  $m_T < 100$  GeV, consistent with W boson decays. This veto successfully  
 1011 reduces the background of leptonically decaying W bosons by about 40%.

### 1012 5.2.4 Baseline Event Selection

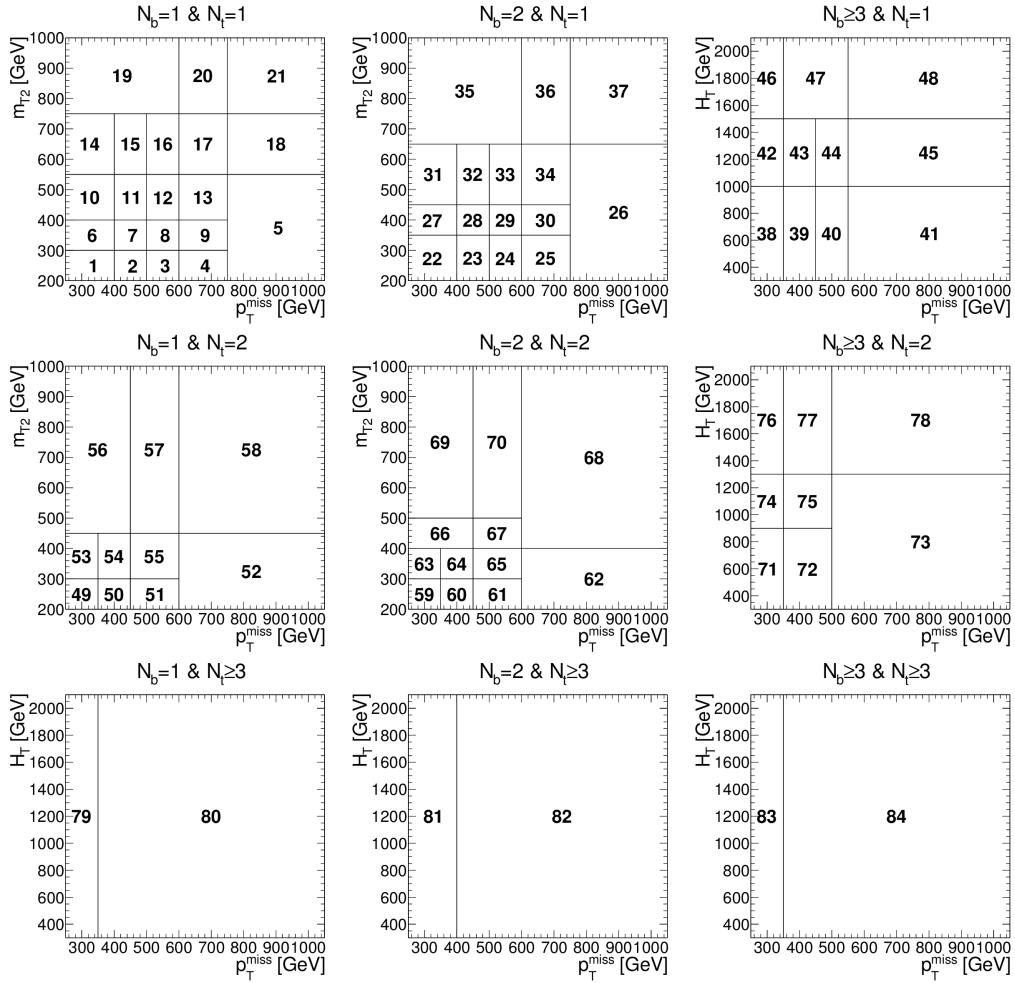
1013 The data selection process begins with the triggers and follows with a pre-selection  
 1014 and the definition of the search bins. The selection criteria applied is optimized for high  
 1015 trigger efficiency as well as being sensitive to a variety of new-physics scenarios. The  
 1016 events selected from data meet the following baseline conditions:

1017

- 1018 • Satisfy the filters designed to remove detector and beam-related noise.
- 1019 • Undergo the lepton, isolated-track, and charged-hadron vetoes.
- 1020 • Have a final state with  $N_j \geq 4$ ,  $N_b \geq 1$ ,  $N_t \geq 1$ ,  $p_T^{miss} > 250$  GeV and  $H_T > 300$   
 1021 GeV.
- 1022 •  $m_{T2} > 200$  GeV, in order to reduce background contributions from  $t\bar{t}$  events.
- 1023 • To reduce the background arising from the QCD multijet background the azimuthal  
 1024 angle between the  $\vec{p}_T^{miss}$  and the three leading jets of an event is required to be  
 1025  $\Delta\phi(p_T^{miss}, j_{1,2,3}) > 0.5, 0.5, 0.3$ , where  $j_1$ ,  $j_2$ , and  $j_3$  indicate the three leading jets  
 1026 in order of decreasing  $p_T$ .

### 1027 5.2.5 Search Bin Definition

1028 The search described in this chapter was performed by defining 84 non-overlapping  
 1029 search regions ([Figure 5.3](#)). Regions that contain events with  $N_b \leq 2$  and  $N_t \leq 2$  use  
 1030  $N_b$ ,  $N_t$ ,  $p_T^{miss}$  and  $m_{T2}$ . In contrast, if the events contain  $N_b \geq 3$  and  $N_t \geq 3$ , then  $N_b$ ,  
 1031  $N_t$ ,  $p_T^{miss}$  and  $H_T$  are used for that region. The reason  $H_T$  is used for these regions instead  
 1032 of  $m_{T2}$  has to do with the fact that for events that contain many jets, the jets arising from  
 1033 the decay of specially heavy objects are not always correctly associated with said object,

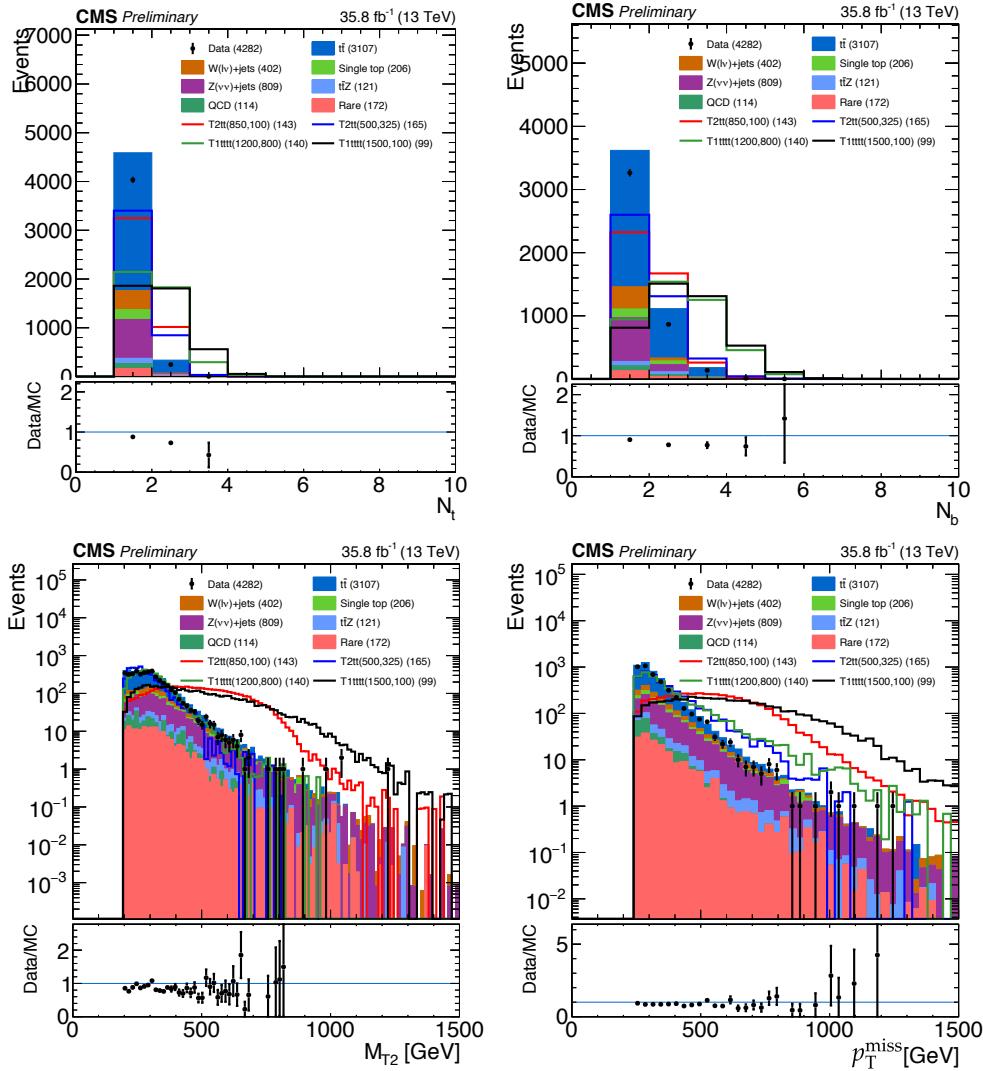


**Figure 5.3:** Search region definitions in the kinematic variables.

giving the  $m_{T2}$  variable a relatively broad and flat distribution. Therefore, in regions with  $N_b \geq 3$  and  $N_t \geq 3$ ,  $H_T$  is found to provide better discrimination between signal and background.

## 5.3 Background Estimation

Since there are many SM processes which exhibit the same characteristics as those of the signal models (e.g. multiple jets, missing energy, etc.) it is of the utmost importance to be able to properly account for their contributions. These interactions, which closely resemble the signal that's being searched for, are called the SM background. However, they cannot be determined solely from the detector's reconstructed data itself, which is why MC simulation is used to recreate them using theoretical predictions from the SM. These



**Figure 5.4:** Comparison between the total SM background from simulation and CMS data for  $N_t$  (top left),  $N_b$  (top right),  $m_{T2}$  (bottom left) and  $p_T^{\text{miss}}$  (bottom right). Total SM backgrounds and signals are scaled to same data yield for a shape comparison.

1044 simulations give us a good idea of what the signal region should look like if there were no  
 1045 signal events, and therefore any significant excess when comparing the MC background to  
 1046 data could signify a potential discovery. The relevant backgrounds for this particular anal-  
 1047 ysis are: the lost-lepton background, the hadronic  $\tau$  background, the  $Z \rightarrow \nu\bar{\nu}$  background  
 1048 and the QCD multijet background.

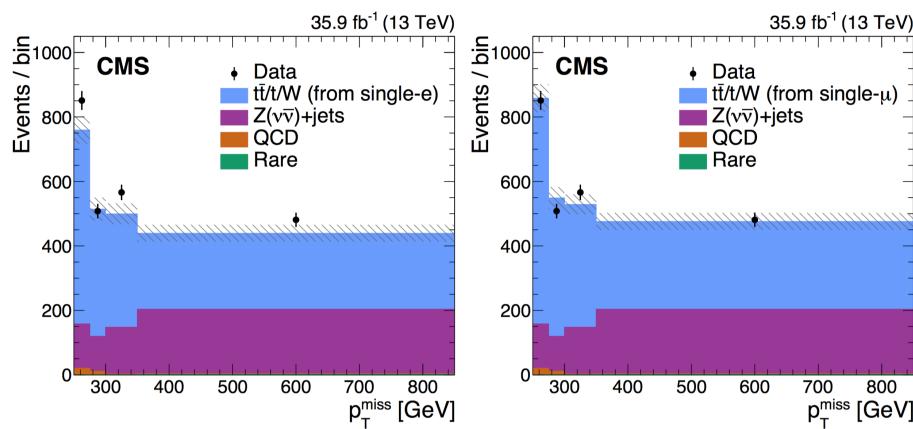
### 1049 5.3.1 Background from $t\bar{t}$ , Single Top Quark and W+jets Events

1050 The vast majority of the expected SM background (around 70%) is due to  $t\bar{t}$  and  
 1051 W+jets events where the W decays into a lepton and a neutrino. Since the expected signal

is purely hadronic, a portion of the leptonic and semi-leptonic decays are vetoed and do not form part of the total SM background. However, there are two different scenarios in which the lepton vetoes can be satisfied. For instance, when the W decays into a  $\tau$  which decays hadronically ( $\tau_h$ ), the  $\tau$  gets reconstructed as a jet and passes the veto. If, on the other hand, the W decays into an electron or a muon, the veto is satisfied when the corresponding lepton is said to be “lost”. In other words, the lepton veto fails to reject events when light leptons (electrons and muons) are not isolated, not identified/reconstructed, or are out of the acceptance region. Both scenarios are evaluated together with a single-lepton data control sample (CS), subjected to the same trigger used for signal events.

1061

The total predicted amount of lost-lepton and  $\tau_h$  events in any given search region is defined by the net sum of single-electron and single-muon events in the respective CS, corrected by a translation factor that is determined from simulation. Due to differences in how they are detected, the single-electron and single-muon samples are determined separately. The translation factor used to correct the search bins is defined as the ratio of simulated  $\tau_h$  and lost-lepton events in the search region over the number of simulated single-electron or single-muon events in the corresponding CS region.



**Figure 5.5:** Distribution of  $p_T^{\text{miss}}$  for both the single-electron (left) and single-muon (right) SB data samples compared to predictions for SM processes. The hatched bands indicate the statistical uncertainties in the total SM prediction. Note that the data and the predictions for all backgrounds except that for  $t\bar{t}$ , single top quark, and W+jets events are identical between the left and right plots.

In order to test this method, a sideband (SB) region is selected which is orthogonal to the search region. This region is defined by the same selection criteria applied to data

1071 except for  $N_t = 0$ ,  $N_b \geq 2$ , and  $\Delta\phi(p_T^{miss}, j_{1,2,3,4}) > 0.5$ , where the last two require-  
 1072 ments are applied in order to suppress contributions from  $Z(\nu\bar{\nu})$ +jets and QCD multijet  
 1073 processes. The SB is divided into four intervals of  $p_T^{miss}$ , where the contribution of lost-  
 1074 lepton and  $\tau_h$  to the corresponding intervals are determined by multiplying the appropriate  
 1075 single -electron and single-muon CS (Figure 5.5) by a translation factor from simulation.  
 1076 The contribution to the SB from other backgrounds, such as  $Z$ +jets, QCD multijet and  
 1077 rare events, is estimated directly from simulation. The total SM background prediction is  
 1078 found to agree with the data within uncertainty, confirming the validity of the translation  
 1079 factor procedure.

1080

1081 Systematic uncertainties stemming from the prediction of  $t\bar{t}$ , single top quark and  
 1082  $W$ +jets background events are determined from various sources: The statistical uncer-  
 1083 tainty associated to the translation factors determined for each search region (1–40%),  
 1084 the lepton reconstruction and isolation efficiency (7–43%), the jet and  $p_T^{miss}$  energy scale  
 1085 and resolution (maximum of 64%), the ISR modeling (maximum of 13%), the PDF's  
 1086 (maximum of 32%) and the b-jet tagging efficiency (1%).

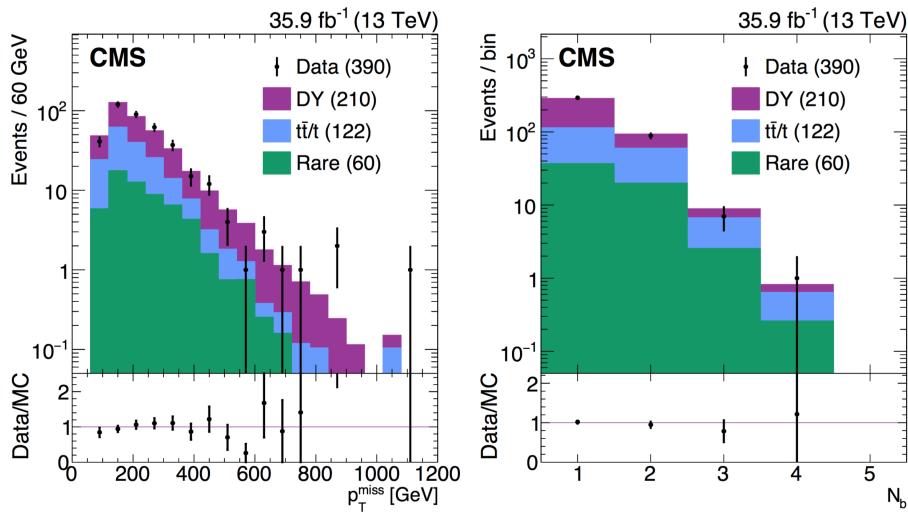
### 1087 5.3.2 The $Z \rightarrow \nu\bar{\nu}$ Background

1088 The  $Z \rightarrow \nu\bar{\nu}$  background is derived using simulated  $Z \rightarrow \nu\bar{\nu}$  events that have been  
 1089 corrected for observed differences between data and simulation. In order to correctly esti-  
 1090 mate this background two scale factors are used to weigh the simulated events:  $R_{norm}$  and  
 1091  $S_{DY}(N_j)$  which correct the normalization of the simulation and the shape of the simulated  
 1092  $N_j$  distribution, respectively. These scale factors are calculated from the dimuon control  
 1093 region which includes events with two muons and no muon or isolated track veto. In the  
 1094 region where  $81 < m < 101$  GeV, the two muons are treated as if they were neutrinos.

1095

1096 The first scale factor,  $R_{norm}$ , is computed by comparing the expected event yield in  
 1097 the dimuon control region from the Drell-Yan (DY) simulation with the observed event

yield in data after subtraction of the other SM processes. The second scale factor,  $S_{DY}$ , depends on the number of jets ( $N_j$ ) in the event and is also computed from the dimuon control region in which the signal region requirements on  $p_T^{miss}$ ,  $N_t$  and  $m_{T2}$  are removed, and the  $H_T$  requirement is relaxed to  $H_T > 200$  GeV. For each  $N_j$  bin this scale factor is calculated from the ratio between data and the DY simulation. Its value ranges between 0.6 and 1.1, depending on the  $N_j$  bin.



**Figure 5.6:** The  $p_T^{miss}$  (left) and  $N_b$  (right) distributions of data and simulation in the loose dimuon CS after applying both the normalization and shape correction factors. The lower panels show the ratio between data and simulation. Error bars in the plot account only for statistical uncertainties. The values in parentheses indicate the integrated yields for each component.

The systematic uncertainties for the  $Z(\nu\bar{\nu})+jets$  background are obtained from the shape differences between data and simulation in the loose dimuon CS in terms of the  $N_b$ ,  $N_t$ ,  $p_T^{miss}$ ,  $m_{T2}$  and  $H_T$  after the normalization factor ( $R_{norm}$ ) has been applied. These include the statistical uncertainty in the  $N_j$  shape correction (1–46%) and in the overall normalization correction (7.6%). Additional systematic uncertainties account for the jet and  $p_T^{miss}$  energy scales (1–71%), the b tagging efficiency (1–23%), the PDFs and the renormalization and factorization scales (1–48%), the statistical uncertainty in the simulation (1–81%, with some search regions as high as 100%), and the trigger (up to 14%). An additional uncertainty is defined from the shift in the central value between the data and simulation in the distributions (14–44% depending on the search region).

### 1114 5.3.3 The QCD Multijet Background

1115 To estimate the QCD Multijet background a signal-depleted, QCD multijet-rich data  
1116 control sample is used. This control sample is defined by inverting the preselection re-  
1117 quirements on  $\Delta\phi(p_T^{miss}, j_{1,2,3})$  and subtracting contributions of other SM backgrounds,  
1118 such as  $t\bar{t}$ , W+jets, and Z+jets. For  $t\bar{t}$  and W+Jets the same methods (lost-lepton and  
1119 hadronic  $\tau$ ) are used to estimate the contributions for this QCD multijet-enriched control  
1120 region. In the case of the  $Z \rightarrow \nu\bar{\nu}$  background, simulation is used for its estimation due to  
1121 its small number. Following that, a translation factor, partly determined by data and partly  
1122 by simulation, is used to convert the number of QCD multijet events measured in the data  
1123 control region into a QCD multijet prediction for each search region bin. This translation  
1124 factor, called  $T_{QCD}$ , is computed as the simulated ratio between the signal region and the  
1125 inverted- $\Delta\phi$  control region, in bins of  $p_T^{miss}$  and  $m_{T2}$  where the bins are bounded in the  
1126 same way as the signal bins.

1127

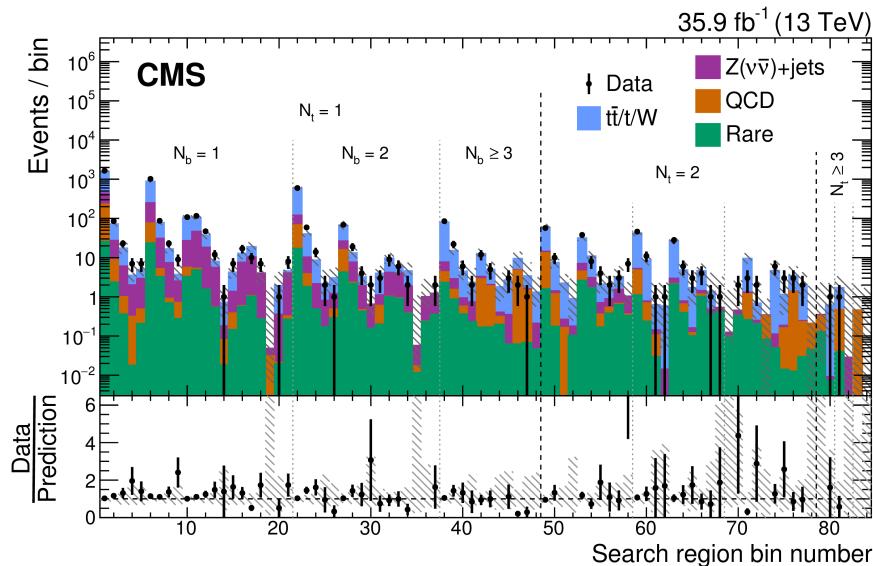
1128 The systematic uncertainty associated with the QCD multijet background for each  
1129 search region is obtained as the difference between the event yield from simulation of  
1130 QCD multijet processes and the prediction obtained by applying the background predic-  
1131 tion procedure to simulated QCD multijet samples (30–500%). Other sources include  
1132 the statistical uncertainty in the translation factors (30–300%) and the subtraction of the  
1133 non-QCD-multijet SM contributions to the QCD control sample (2–50%).

### 1134 5.3.4 Background From Rare and Other Processes

1135 The background contribution from rare processes forms only a tiny fraction of the  
1136 overall SM background and has therefore a small effect on the final result. The estimation  
1137 for these processes are determined directly from simulation, with the biggest contribution  
1138 coming from  $t\bar{t}Z$ . With the exception of the  $t\bar{t}Z$  process, all other remaining backgrounds  
1139 are combined. The comparison between simulation and the obtained data were found to be  
1140 in agreement within a statistical uncertainty of 30%, which is attributed to the systematic

1141 uncertainty in the estimation of the  $t\bar{t}Z$  background.

## 1142 5.4 Results and Interpretation



**Figure 5.7:** Observed events compared to the corrected SM background predictions for all 84 search regions in the full data of  $35.9 \text{ fb}^{-1}$  collected in 2016. The lower panel shows the ratio between the data and the total SM simulation. The gray bands show the total uncertainty related to the background prediction.

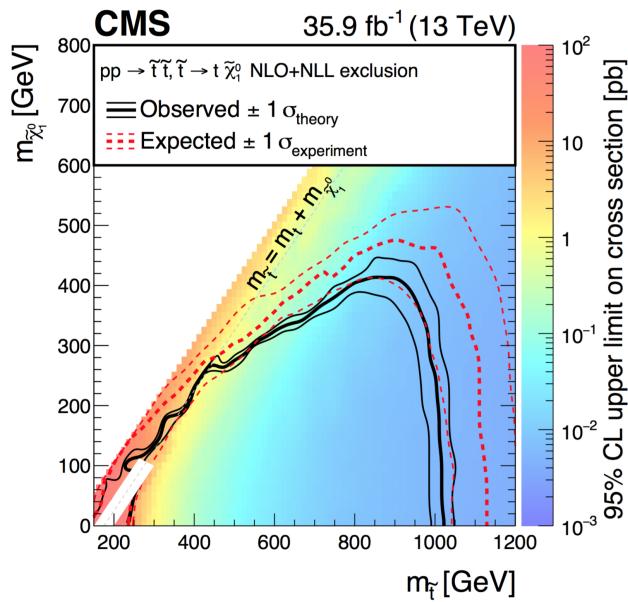
1143 **Figure 5.7** shows the summary of all observed events compared to the corrected SM  
1144 background simulation for all of the 84 search bins. The data obtained shows no statisti-  
1145 cally significant deviation from the predicted SM background. The biggest contribution  
1146 to the background is attributed to the  $t\bar{t}$  and  $W+jets$  processes, followed by  $Z(\nu\bar{\nu})+jets$ ,  
1147 which could be dominant in regions that have a high  $p_T$  threshold. The contributions arising  
1148 from the QCD multijet and rare backgrounds are found to be nearly negligible in all  
1149 of the search regions.

1150

1151 Exclusion limits are calculated for each of the signal models discussed in this chapter  
1152 by applying a binned likelihood fit on the data. The likelihood function is obtained for  
1153 each of the 84 search regions as well as for each of the background data control samples  
1154 (single-electron, single-muon, and QCD) from the product of the Poisson probability den-  
1155 sity function. Exclusion limits were placed on the top squark, gluino and LSP production

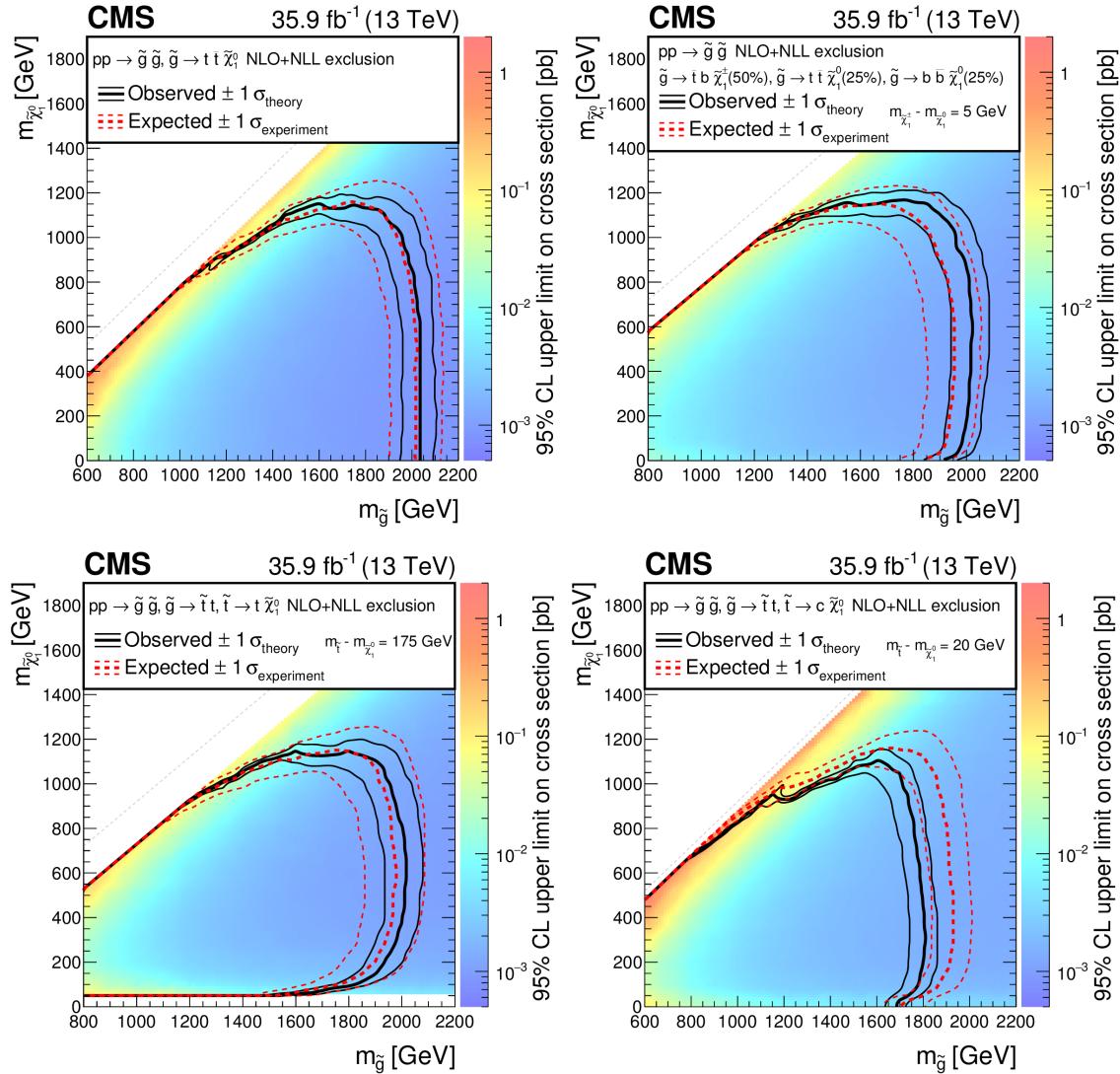
<sup>1156</sup> cross-sections with a 95% confidence level (CL), calculated using a modified frequentist  
<sup>1157</sup> approach with the  $\text{CL}_s$  criterion [51, 52] and asymptotic results for the test statistic [53].

<sup>1158</sup>



**Figure 5.8:** The 95% CL upper limit on the production cross section of the T2tt simplified model as a function of the top squark and LSP masses. The solid black curves represent the observed exclusion contour with respect to NLO+NLL [8] signal cross sections and the change in this contour due to variation of these cross sections within their theoretical uncertainties. The dashed red curves indicate the mean expected exclusion contour and the region containing 68% of the distribution of expected exclusion limits under the background-only hypothesis.

<sup>1159</sup> The uncertainties from the search region modeling are taken into account for each of  
<sup>1160</sup> the 84 bins and arise from several different sources: the statistical uncertainty in the sim-  
<sup>1161</sup> ulated event samples, the integrated luminosity (2.5% [54]), the lepton and isolated-track  
<sup>1162</sup> veto efficiencies (up to 6.8%), the b tagging efficiency (up to 21%), the trigger efficiency  
<sup>1163</sup> (up to 2.6%), the renormalization and factorization scales (up to 3.5%), the ISR modeling  
<sup>1164</sup> (up to 46%), the jet energy scale corrections (up to 34%), the top quark reconstruction  
<sup>1165</sup> efficiency (up to 14%), and the modeling of the fast simulation compared with the full  
<sup>1166</sup> simulation for top quark reconstruction and mis-tagging (up to 24%). All of the uncer-  
<sup>1167</sup> tainties, with the exception of the statistical precision of the simulation, are taken to be  
<sup>1168</sup> fully correlated between the search regions. Contributions from signal contamination in  
<sup>1169</sup> the signal modeling are found to be only significant from the single-lepton control sam-  
<sup>1170</sup> ples and negligible for the rest.



**Figure 5.9:** The 95% CL upper limit on the production cross section of the T1tttt (upper left), T1ttbb (upper right), T5tttt (bottom left), and T5ttcc (bottom right) simplified models as a function of the gluino and LSP masses.

1171

1172     The 95% CL exclusion limits obtained for the T2tt model (Figure 5.8), which consists  
 1173     of direct top squark production, excludes top squark masses up to 1020 GeV and LSP  
 1174     masses up to 430 GeV. Meanwhile, Figure 5.9 shows the results for the gluino pair pro-  
 1175     duction models: T1tttt, T1ttbb, T5tttt, and T5ttcc. For the T1tttt model, gluino masses of  
 1176     up to 2040 GeV and LSP masses up to 1150 GeV are excluded, with corresponding limits  
 1177     of 2020 and 1150 GeV for the T1ttbb model, 2020 and 1150 GeV for the T5tttt model,  
 1178     and 1810 and 1100 GeV for the T5ttcc model.

1179 **Chapter 6**

1180 **Estimation of the  $Z \rightarrow \nu\bar{\nu}$  Background**

1181 **6.1 Introduction**

1182 A detailed explanation for the estimation of the  $Z \rightarrow \nu\bar{\nu}$ +jets background is presented  
1183 in this chapter. The following estimation procedure builds upon the 2016 SUSY analysis  
1184 summarized in [chapter 5](#) and aims to refine the overall background calculation as well  
1185 as reduce the uncertainties associated with the previous method. To accomplish this, an  
1186 additional  $\gamma$ +jets CS is used in conjunction with the tight  $Z \rightarrow \mu^+\mu^-$  control region used in  
1187 the 2016 analysis estimation. The new  $\gamma$ +jets CS provides a more data-driven estimation  
1188 procedure with the added benefit of a substantially larger production cross-section than  
1189  $Z$ +jets processes [[55](#)].

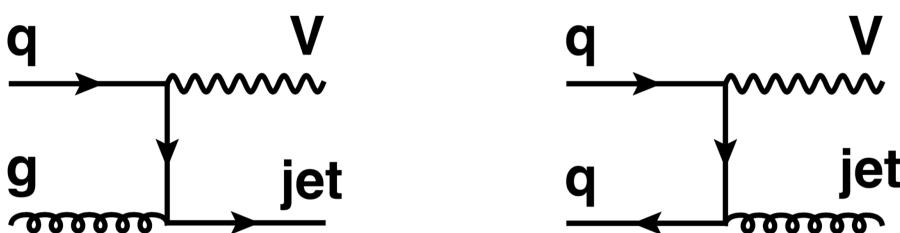
1190 **6.2 The Irreducible  $Z \rightarrow \nu\bar{\nu}$  Background**

1191 An important source of background in searches for SUSY in the 0-lepton final state  
1192 comes from events in which a  $Z$  boson, accompanied by jets, decays into a pair of neu-  
1193 trinos ( $Z \rightarrow \nu\bar{\nu}$ +jets). The resulting final state is comprised of a large  $p_T^{miss}$  (from the  
1194 neutrino pair) and multiple hadron jets, closely mimicking the SUSY signal. For similar  
1195 searches, the  $Z \rightarrow \nu\bar{\nu}$  contribution can make up a large portion of the background in many  
1196 of the search bin regions (higher than 50% in some regions) [[56,57](#)], and can compose up

to about a third of the total SM background. For this particular analysis, the  $Z \rightarrow \nu\bar{\nu}$  accounts for about 17% of the total SM background and owes its low value to the dedicated top-tagging algorithm outlined in subsection 5.2.2.

1200

There are several different methods that have been developed to estimate the  $Z \rightarrow \nu\bar{\nu} +$  jets background [58]. Two of the commonly used methods involve the use of a control region dominated by  $Z \rightarrow ll +$  jets, where the  $l$  stands for lepton (either a muon or an electron, in this case) or  $\gamma +$  jets events. The  $Z \rightarrow ll$  channel has the advantage of having very similar kinematics to the  $Z \rightarrow \nu\bar{\nu}$  region but suffers from low statistics (due in part to its small branching ratio), specially in the tight search regions used in typical SUSY searches. On the other hand, the  $\gamma +$  jets region has a much higher production cross-section but involves a completely different process. The hybrid method described in this chapter makes use of both control regions in order to estimate the  $Z \rightarrow \nu\bar{\nu}$  background corrections, and aims to improve on the results of the 2016 method described in subsection 5.3.2.



**Figure 6.1:** Leading-order Feynman diagrams for  $Z +$  jets and  $\gamma +$  jets processes. The ‘V’ in the figure can represent either  $Z$  or  $\gamma$ .

## 6.3 The Loose $\gamma+$ jets Control Region

The loose  $\gamma +$  jets control region, as well as the photon ID/isolation selection, is described in this section. The  $\gamma +$  jets control region is chosen with the purpose of substituting the loose muon control region used in the 2016 version of this analysis for the calculation of the shape correction scale factors applied to the final estimation of the  $Z \rightarrow \nu\bar{\nu}$  background. The  $\gamma +$  jets control region is presumed to be better suited for this estimation due to having a much higher cross-section and therefore, an expected reduction in the statistical and systematic uncertainties associated to the shape correction factors are

1219 expected.

1220 **6.3.1 Photon ID and Isolation**

1221 Three different working points are provided by the CMS EGM physics object group  
1222 (POG) for simple cut-based photon identification [59]. The three working points, called  
1223 loose, medium and tight, are chosen according to the requirements of the particular analy-  
1224 sis and differ on the amount of background rejection they offer, as well as on their average  
1225 photon selection efficiency. The higher the efficiency of a given ID, the lower the amount  
1226 of background that is rejected, and vice-versa. Table 6.1 shows the cut values that are  
1227 applied to photons that are found within both the ECAL barrel and endcap range. The  
1228 associated values to the efficiency and the background rejection rate are shown for each  
1229 of the three different photon ID selections.

1230

1231 In order to obtain the high efficiency and background rejection rates shown, a robust  
1232 set of identification and isolation criteria are selected. A total of five parameters are  
1233 used for this simple cut based method. For photon identification, the H/E and the  $\sigma I\eta I\eta$   
1234 variables are found to provide the best results. The H/E parameter is defined as the ratio  
1235 of the HCAL tower energy over the ECAL seed cluster energy. A threshold value is  
1236 selected on H/E to remove background from electrons that are detected in both the ECAL  
1237 and HCAL but have no reconstructed track [60]. The  $\sigma I\eta I\eta$  variable is known as the  
1238 photon shower shape variable and is defined in the ECAL as the energy weighted standard  
1239 deviation of a single crystal within the  $5 \times 5$  crystal  $\eta$  range, centered around the crystal  
1240 with maximum energy [61]. This variable is a key component in the identification of both  
1241 electrons and photons since it provides a measure of the shower width where most of the  
1242 energy has been deposited in a given ECAL crystal.

**Table 6.1:** Identification and isolation cut values for photons provided by the CMS EGM POG. Values are provided for the three working points described (loose, medium and tight) for both the ECAL barrel and endcaps. The photon selection efficiency of the three working points, as well as their associated background rejection rate, are provided.

| Barrel  | Loose (90.06%)  | Medium (80.19%)   | Tight (70.01%)  |
|---|---|---|---|
| <b>Background Rejection</b>                   | <b>Loose (85.73%)</b>   | <b>Medium (88.87%)</b>  | <b>Tight (90.66%)</b>   |
| H/E   | 0.105   | 0.035   | 0.020   |
| $\sigma I\eta I\eta$                          | 0.0103  | 0.0103  | 0.0103  |
| $\rho$ -corrected PF charged hadron isolation | 2.839   | 1.416   | 1.158   |
| $\rho$ -corrected PF neutral hadron isolation | $9.188 + 0.0126 \cdot p_T^\gamma + 0.000026 \cdot p_T^{\gamma^2}$ | $2.491 + 0.0126 \cdot p_T^\gamma + 0.000026 \cdot p_T^{\gamma^2}$ | $1.267 + 0.0126 \cdot p_T^\gamma + 0.000026 \cdot p_T^{\gamma^2}$ |
| $\rho$ -corrected PF photon isolation         | $2.956 + 0.0035 \cdot p_T^\gamma$                                 | $2.952 + 0.0040 \cdot p_T^\gamma$                                 | $2.065 + 0.0035 \cdot p_T^\gamma$                                 |

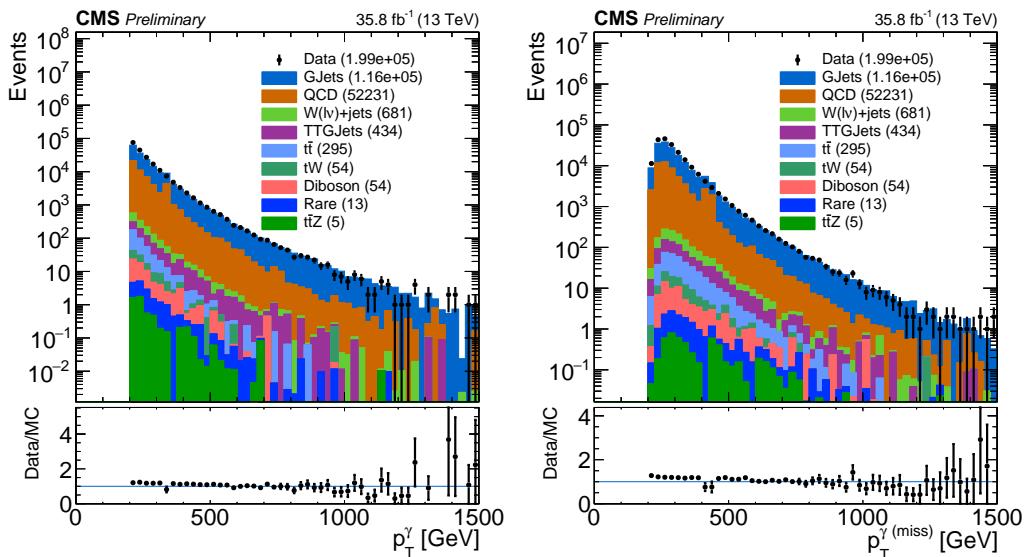
| End Cap                                       | Loose (90.81%)   | Medium (80.06%)   | Tight (70.11%)  |
|---|--|---|---|
| <b>Background Rejection</b>                   | <b>Loose (76.90%)</b>  | <b>Medium (81.50%)</b>  | <b>Tight (84.34%)</b>   |
| H/E   | 0.029  | 0.027   | 0.025   |
| $\sigma I\eta I\eta$                          | 0.0276   | 0.0271  | 0.0271  |
| $\rho$ -corrected PF charged hadron isolation | 2.150  | 1.012   | 0.575   |
| $\rho$ -corrected PF neutral hadron isolation | $10.471 + 0.0119 \cdot p_T^\gamma + 0.000025 \cdot p_T^{\gamma^2}$ | $9.131 + 0.0119 \cdot p_T^\gamma + 0.000025 \cdot p_T^{\gamma^2}$ | $8.916 + 0.0119 \cdot p_T^\gamma + 0.000025 \cdot p_T^{\gamma^2}$ |
| $\rho$ -corrected PF photon isolation         | $4.895 + 0.0040 \cdot p_T^\gamma$                                  | $4.095 + 0.0040 \cdot p_T^\gamma$                                 | $3.272 + 0.0040 \cdot p_T^\gamma$                                 |

1243 The other three parameters considered, comprise the isolation portion of the photon  
 1244 selection cuts. These are the  $\rho$ -corrected particle flow (PF) charged hadron, neutral hadron  
 1245 and photon isolation parameters. As can be seen from Table 6.1, two of these parameters  
 1246 (the neutral hadron and photon isolation) have a dependence on  $p_T^\gamma$ . These cuts are used  
 1247 to ensure that the identified photon is well isolated within its own cone and by rejecting  
 1248 photons that are identified within close proximity to either a charged or a neutral hadron  
 1249 [62]. The value  $\rho$  included in the name of each of these parameters refers to the total

pileup density [63]. Therefore, the term “ $\rho$ -corrected” implies that these values, which are sensitive to the residual contamination that arises from pile-up, have been corrected to include these contributions.

### 6.3.2 Photon Selection

The event selection process for the  $\gamma$ +jets control region starts with photon candidates that have a  $p_T > 200$  GeV and are within the acceptance range of the CMS ECAL (given by  $|\eta| < 1.4442$  for the barrel and  $1.566 < |\eta| < 2.5$  for the endcaps). The photons are subjected to pass the loose ID/isolation cuts described in subsection 6.3.1 in order to remove  $\sim 85\%$  of the background processes and obtain a prompt photon sample that is  $\sim 90\%$  pure, on average. Additional restrictions include some of the same requirements imposed on the signal baseline selection discussed in subsection 5.2.4. These include  $N_j \geq 4$ , an  $H_T > 300$  GeV, the  $\Delta\phi$  requirements for leading jets and the lepton vetoes described in subsection 5.2.3. The lepton veto, in particular, greatly improves the prompt photon selection by removing many of the events in the simulated samples where a lepton gets misidentified as a photon.



**Figure 6.2:** Shown are both the  $p_T^\gamma$  (left) and  $p_T^{\gamma(\text{miss})}$  (right) distributions before applying any corrections.  $p_T^{\gamma(\text{miss})}$  is obtained by adding the  $p_T^\gamma$  to the total  $p_T^{\text{miss}}$  in every event.

To further emulate the  $Z \rightarrow \nu\bar{\nu}$ +jets background, a variable in which the photons are

treated as  $p_T^{miss}$  is defined. We call this variable  $p_T^{\gamma(miss)}$  and we obtain it by adding the  $p_T^\gamma$  for every event to the total  $p_T^{miss}$  in the event. Both the  $p_T^\gamma$  and the resulting  $p_T^{\gamma(miss)}$  distributions are shown in Figure 6.2 as data/MC comparison plots, where the simulated backgrounds are stacked in order of ascending contribution.

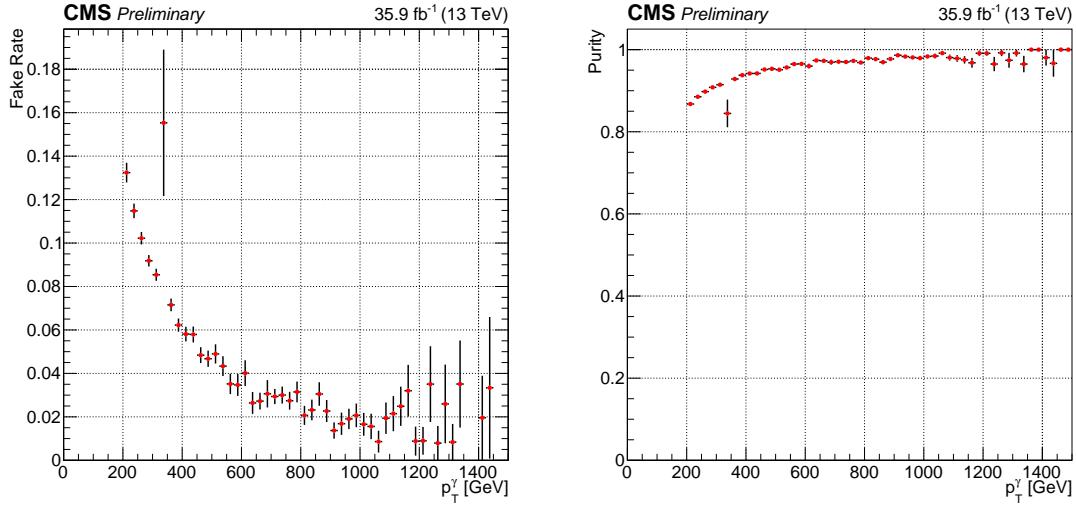
1270

The main contributions from simulation arise from the  $\gamma$ +jets, QCD and to a lesser extent,  $t\bar{t}\gamma$ . Other non-dominant backgrounds in the control region include contributions from  $W(l\nu)+\text{jets}$ ,  $t\bar{t}$ , Diboson,  $tW$ ,  $t\bar{t}Z$  and rare processes. Most of these lesser backgrounds are nearly negligible (several orders of magnitude lower than the dominant backgrounds) and are considered to be mostly composed of fake photons. In addition to the cuts described, all of the simulation samples are subjected to weights that apply corrections to pileup as well as the b-tagging efficiency. Data, on the other hand, is obtained from a sample that contains events with at least one identified photon. Photons in this sample are also subjected to the high-level trigger HLT\_Photon175, which restricts the selection to photons that have a  $p_T > 175$  GeV. Both simulation and data are subjected to the same selection criteria established in this section.

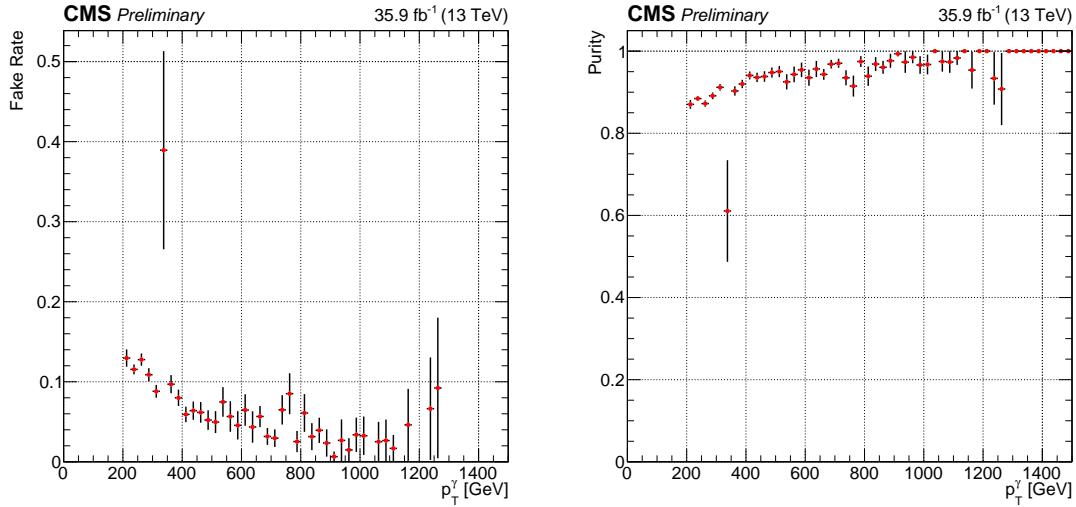
### 6.3.3 Photon Purity and Fake Rate

Three different types of photons make up the  $\gamma$ +jets CS: prompt photons, produced either directly or through fragmentation, and fake photons. Prompt photons are defined as photons which are formed shortly after the proton-proton collision (i.e. before the produced quarks and gluons have had enough time to form hadrons). Two types of photons fit in this category. The first type, which we designate as direct photons, are photons that are produced directly from the proton-proton interaction [64]. A secondary type of prompt photon, that is virtually indistinguishable from the direct photons at the detector level, originates from the decay of  $\pi^0$  mesons and are called fragmentation photons. The final type of photon found in the CS corresponds to fake (or non-prompt) photons. The fake photon contribution typically arises from leptons (mostly electrons) whose tracks are not

<sup>1293</sup> properly reconstructed, yet leave energy measurements in the ECAL.



**Figure 6.3:** Plots for Fake Rate (left) and Purity (right) as a function of the photon  $p_T$  are shown. The events are selected are required to have a  $p_T > 200$  GeV, be within the ECAL acceptance range, and pass the loose ID selection cuts. This selection was produced in order to verify the values given by the E/ $\gamma$  POG. As can be seen, the efficiency (purity) is seen to agree with the values of the loose photon ID/isolation selection.



**Figure 6.4:** Plots for Fake Rate (left) and Purity (right) as a function of the photon  $p_T$  are shown. These plots include photons with the full control region selection. Aside from exhibiting lower statistics, the plots seem to agree with the fake rate and purity before all the control region cuts are applied.

<sup>1294</sup> In order to identify prompt photons, reconstructed photons from the  $\gamma + \text{jets}$  and QCD  
<sup>1295</sup> samples are matched to generator-level photons in space and momentum by requiring  
<sup>1296</sup>  $\Delta R(\gamma_{\text{gen}}, \gamma_{\text{reco}}) < 0.4$  and  $0.5 < p_T^{\text{gen}} / p_T^{\text{reco}} < 2.0$ , respectively. Any reconstructed photon  
<sup>1297</sup> which fails to get matched to a generator level photon is labeled as a fake/non-prompt  
<sup>1298</sup> photon. Direct photons are identified by further requiring that the reconstructed photons

1299 be matched to a parton (a gluon or quark) in space as  $\Delta R(\gamma, \text{parton}) > 0.4$ . This require-  
 1300 ment is intended to distinguish the reconstructed photons from highly boosted  $\pi^0$ 's, which  
 1301 compose a large portion of the experimentally indistinguishable fragmentation photons.  
 1302 Finally, fragmentation photons are obtained exclusively from QCD simulation and are re-  
 1303 quired to have  $\Delta R(\gamma, \text{parton}) < 0.4$  in order to avoid double counting photons from the  
 1304  $\gamma$ +jets sample.

1305

1306 With all three types of photons defined, a study can be carried out from simulation  
 1307 to estimate their respective contributions to the defined control region. The study takes  
 1308 into account that any reconstructed photon in the  $\gamma$ +jets or QCD samples can only be  
 1309 categorized as prompt (through direct production or fragmentation) or non-prompt (fake).  
 1310 The purity and fake rate can then be defined in terms of the relative proportions of prompt  
 1311 or non-prompt photons with respect to the sum of the contributions from all three types of  
 1312 photons. Identified direct photons are taken from the  $\gamma$ +jets sample exclusively. Mean-  
 1313 while, the fragmentation and fake photon contributions are taken from the QCD sample.  
 1314 The three quantities are then added together and their respective contributions are deter-  
 1315 mined in terms of the photon  $p_T$  ([Figure 6.3](#)).

1316

1317 The photon purity ([Figure 6.3](#) and [Figure 6.4](#), right) is defined in terms of the prompt  
 1318 and non-prompt photons as:

1319

$$1320 p_\gamma = \frac{\text{prompt}}{\text{prompt} + \text{fake}} ,$$

1321 where the prompt photon portion comes from the sum of the direct photons (extracted  
 1322 from the  $\gamma$ + jets sample) and the fragmentation photons (extracted from the QCD sample).  
 1323 The remaining non-prompt (or fake) photons all come from photons in the QCD sample  
 1324 that were not matched to truth-level photons in space and momentum with the specified  
 1325 required conditions. Meanwhile, the photon fake rate ([Figure 6.3](#) and [Figure 6.4](#), left) is

1326 defined from this same combination of samples as:

1327

$$\text{f} = \frac{\text{fake}}{\text{prompt+fake}},$$

1329 [Figure 6.3](#) shows the purity and fakerate for photons that pass the loose ID/selection,  
1330 have a  $p_T > 200$  GeV and are within the ECAL acceptance range. A sample is obtained  
1331 in which 77% of the photons are direct, 12% are fragmentation and 11% are fakes. This  
1332 implies an average purity of  $\sim 89\%$  for this sample, well within the value that is expected.  
1333 [Figure 6.4](#) shows the same ratios for the loose  $\gamma +$  jets control region described in [subsection 6.3.2](#). Although the amount of statistics has decreased due to the additional cuts, a  
1334 similar trend can be observed.

## 1336 6.4 The $Z \rightarrow \mu^+\mu^-$ Control Region

1337 The  $Z \rightarrow \mu^+\mu^-$  control region defined in this section is in every respect identical to  
1338 the one applied in the 2016 analysis ([subsection 5.3.2](#)). The only difference between the  
1339 2016 method and the one discussed in this chapter is that the Drell-Yan (DY) sample is  
1340 only used for the normalization correction of the  $Z \rightarrow \nu\bar{\nu}$  background. Therefore, the  
1341 loose  $\mu\mu$  control region is not used or applied in the calculation of the scale factors. In  
1342 the following subsections only the tight  $\mu\mu$  control region, and its usage to obtain the  
1343 normalization scale factor  $R_{norm}$ , is discussed.

### 1344 6.4.1 Muon ID and Isolation

1345 The muons are selected using the “medium muon” selection [65], per the recommen-  
1346 dation of the muon POG. The muon candidates in this selection satisfy  $p_T > 10$  GeV and  
1347  $|\eta| < 2.4$ . Other additional cuts are applied to aid in the muon candidate selection, such as  
1348 an impact parameter cut. Muons are also subjected to a PF relative-isolation (also referred  
1349 to as mini-isolation) in which the cone size is inversely proportional to the muon  $p_T$ . This

requirement enforces the  $p_T$  within the isolation cone to be at most 20% of the muon  $p_T$  in order to eliminate events with an isolated muon. Details of the medium photon selection are included in [Table 6.2](#) and [Table 6.3](#), while details of the impact parameter cut are summarized in [Table 6.4](#).

| Muon Medium ID                                    |              |
|---|--------------|
| Loose muon ID                                     | Yes          |
| Fraction of valid tracker hits >                  | 0.80         |
| Good Global muon OR Tight segment compatibility > | Yes OR 0.451 |

**Table 6.2:** Muon Medium ID 2016 HIP Safe

| Good Global muon                    |       |
|-------------------------------------|-------|
| Global muon                         | Yes   |
| Normalized global-track $\chi^2 <$  | 3     |
| Tracker-Standalone position match < | 12    |
| Kick finder <                       | 20    |
| Segment compatibility >             | 0.303 |

**Table 6.3:** Muon Medium ID HIP Safe Good Global Muon

| Muon Impact Parameter |     |
|-----------------------|-----|
| d0 <                  | 0.2 |
| dz <                  | 0.5 |

**Table 6.4:** Additional Impact Parameter cut on Muons

#### 6.4.2 Muon Selection in the Tight Control Region

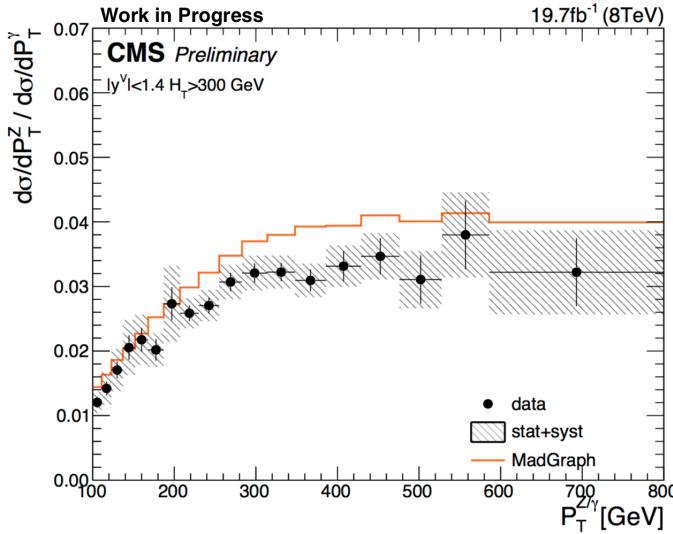
Events are selected from data samples that contain exactly two oppositely charged muons ( $\mu^+ \mu^-$ ), which fall within the invariant mass  $81 < m_{ll} < 101$  GeV window for the  $Z$  boson. Additional cuts for the tight muon selection include baseline requirements such as an  $H_T > 300$  GeV,  $N_j \geq 4$ , the  $\Delta\phi$  baseline cut on leading jets, a  $p_T^{miss} > 250$  GeV, an  $m_{T2} > 200$  GeV and at least 1 top-tagged jet  $N_t \geq 1$ . In addition, the  $p_T$  of the two muons are required to be  $p_T > 50$  GeV for the leading muon and  $p_T > 20$  GeV for the sub-leading one. The only difference, when compared to the signal region is the missing lepton veto, in addition to the dimuon events being treated as  $p_T^{miss}$ . This makes for a region that exhibits very similar kinematics to the  $Z \rightarrow \nu\bar{\nu}$  signal region, yet suffers from a lack of statistics.

1365 **6.5 Analysis**

1366 In this section a detailed explanation of the calculation of the scale factors for both  
 1367 shape and normalization corrections is provided. The following methods make use of  
 1368 the loose  $\gamma$ +jets and the tight  $\mu\mu$  control regions defined in the previous sections. The  
 1369 procedure involves extracting the shape corrections  $S_\gamma$  from the  $\gamma$ +jets control region and  
 1370 afterwards obtain a single normalization correction factor  $R_{norm}$  from the tight  $\mu\mu$  control  
 1371 region. Both factors will then be applied to the final prediction of the  $Z \rightarrow \nu\bar{\nu}$  background  
 1372 in each of the required search bins.

1373 **6.5.1 Shape Correction Using the  $\gamma$  + jets Control Sample**

1374 In this section the validation of the  $\gamma$ +jets simulation is discussed in terms of the  
 1375 shape of the loose photon control region. As it was shown in subsection 6.3.3, this con-  
 1376 trol region has high purity for  $\gamma$ +jets events, particularly in regions of high  $p_T$  ( $\gtrsim 300$ ).  
 1377 In order to apply this correction factor it is assumed that the shape differences between  
 1378 data and simulation are similar between  $Z \rightarrow \nu\bar{\nu}$  and  $\gamma$ +jets events. This assumption is  
 1379 validated in studies which compare the cross-section ratio of Z+jets to  $\gamma$ +jets events [66].  
 1380 Figure 6.5 shows the results of this study, conducted in 2014, for both data and Mad-  
 1381 Graph simulation with an integrated luminosity of  $19.7\text{fb}^{-1}$  and a center-of-mass energy  
 1382 of 8 TeV. It can be seen that for values of  $p_T^{Z/\gamma} \gtrsim 300$  GeV, the ratio of the cross-section of  
 1383 both processes becomes nearly constant. It is then a matter of applying a factor to account  
 1384 for the difference in the amount of events between the Z+jets and  $\gamma$ +jets events in order  
 1385 to obtain the total amount of  $Z \rightarrow \nu\bar{\nu}$ +jets events. This factor is obtained from the tight  
 1386  $\mu\mu$  control region, as shown in subsection 6.5.2.

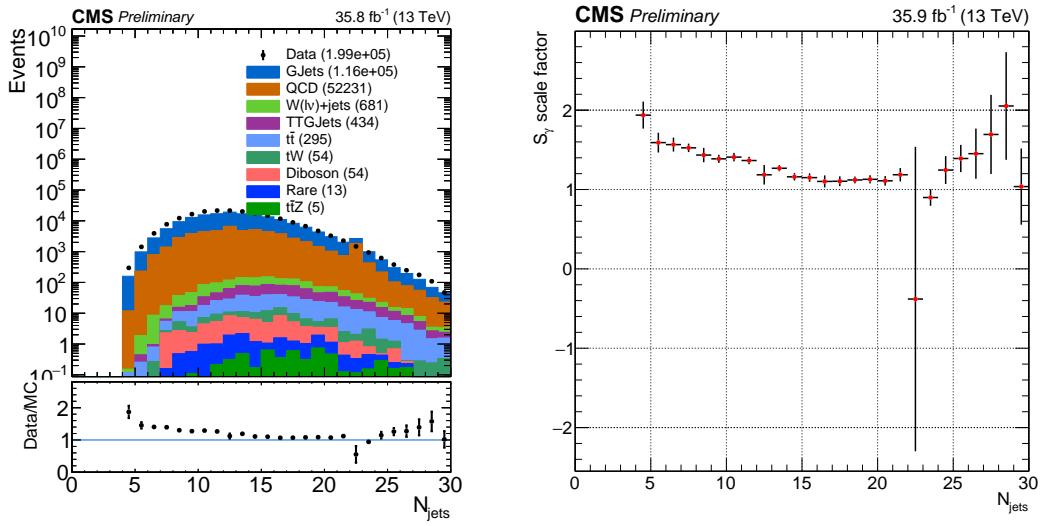


**Figure 6.5:** Results of study of the  $Z$ +jets to  $\gamma$ +jets cross-section ratio for both data and MadGraph simulation. For high values of the vector boson transverse momentum, the ratio between these processes is observed to be nearly constant.

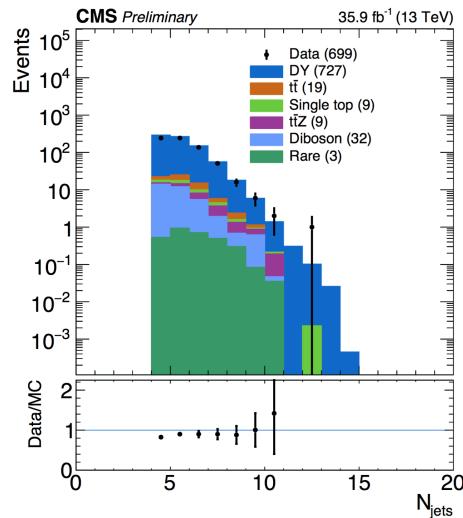
1387 In order to obtain the shape corrections, the ratio between data and simulation of the  
 1388 jet multiplicity distribution is used (Figure 6.6). This is due to it exhibiting the highest  
 1389 difference between the observed data and MC. The re-weight for the  $\gamma$ +jets simulation  
 1390 sample is then accomplished by applying the  $N_{jet}$  dependent factor  $S_\gamma(N_j)$ . This scale  
 1391 factor is determined by taking the ratio of the data and simulation, after subtracting all  
 1392 other MC samples from data events:

$$1393 \quad S_\gamma^i = \frac{\text{Data}^i - \text{MC}_{\text{other}}^i}{\text{MC}_{\gamma+\text{jets}}^i},$$

1394 where  $i$  denotes any given bin in the  $N_j$  distribution. The shape correction factors  $S_\gamma^i$   
 1395 are displayed graphically in Figure 6.6 (right) for each  $N_j$  bin. These factors correct for  
 1396 differences in the jet multiplicity shape, while the overall normalization is estimated from  
 1397 the tight  $\mu\mu$  control region. Figure 6.7 shows the  $N_j$  distribution in the tight  $\mu\mu$  control  
 1398 region after the calculated scale factors have been applied. The  $S_\gamma$  correction will be  
 1399 applied to the  $Z \rightarrow \nu\bar{\nu}$  simulation final prediction for each of the analysis search bins. The  
 1400 uncertainty associated with the scale factor is estimated from the event yields in the loose  
 1401 photon control region. This uncertainty will form part of the total systematic uncertainty  
 1402 in the final prediction.

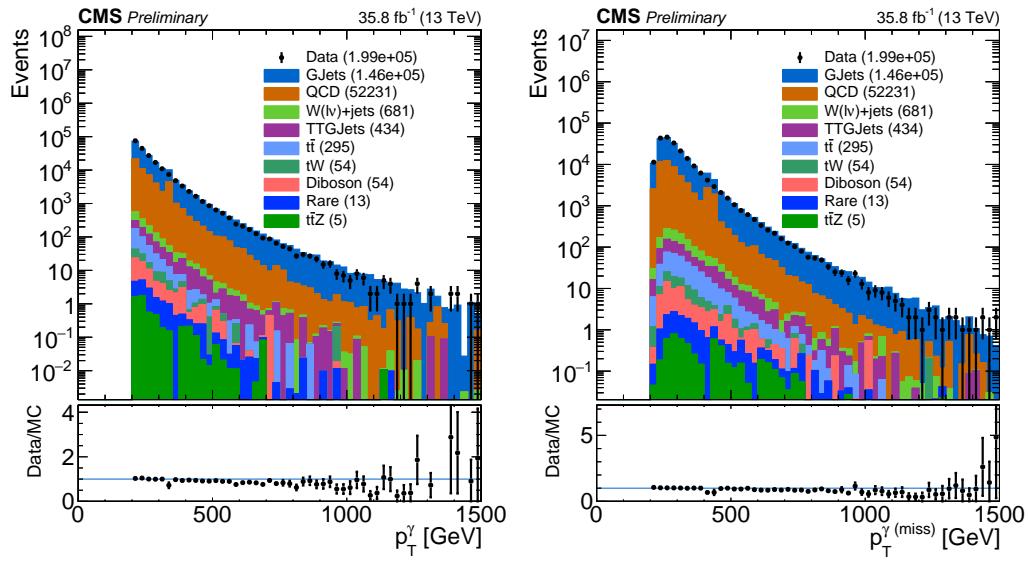


**Figure 6.6:** Jet multiplicity and the associated  $S_\gamma$  scale factor in the loose photon control region before any corrections are applied.

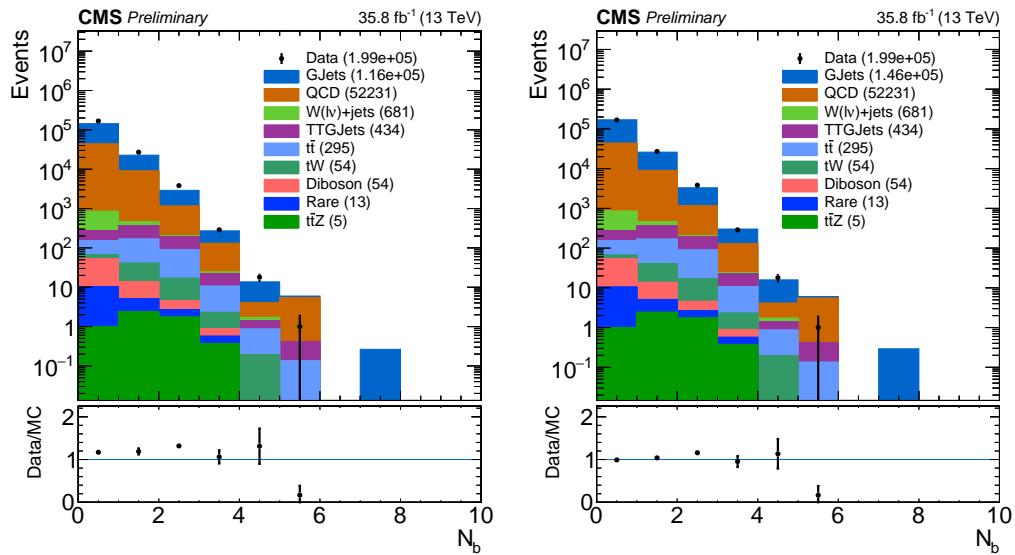


**Figure 6.7:**  $N_{jet}$  distribution in the tight  $\mu\mu$  control region after  $S_\gamma$  corrections.

1403     The effect of the  $S_\gamma(N_j)$  scale factor is shown for various distributions. These results  
 1404     show that the overall agreement between data and simulation improves after applying the  
 1405     corresponding shape corrections.



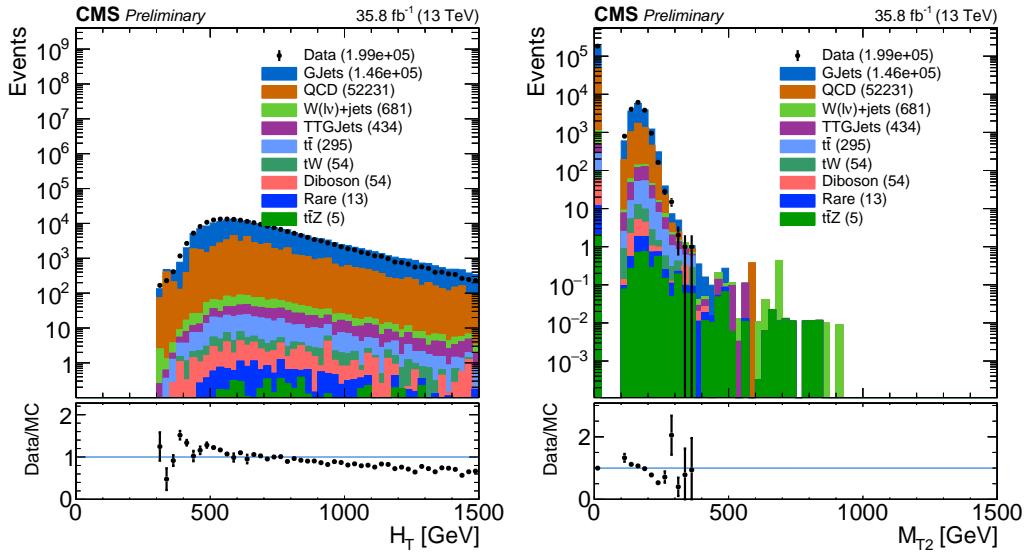
**Figure 6.8:**  $p_T^\gamma$  (left) and  $p_T^{\gamma(\text{miss})}$  (right) distributions after applying the  $S_\gamma(N_j)$  scale factor. Comparing to Figure 6.2, an improvement in the agreement between data/MC can be observed.



**Figure 6.9:**  $N_b$  distribution before (left) and after (right) applying the  $S_\gamma(N_j)$  scale factor.

## 6.5.2 Normalization Correction Using the tight $Z \rightarrow \mu^+\mu^-$ Control Sample

In order to constrain the normalization of the  $Z \rightarrow \nu\bar{\nu}$  simulation sample, a normalization correction factor  $R_{norm}$  is calculated from the tight  $\mu\mu$  control region defined in subsection 6.5.2. Two categories are considered: the zero b-tagged jet category ( $N_b = 0$ ), and the  $\geq 1$  b-tagged jet category ( $N_b \geq 1$ ). Both of these categories are statistically

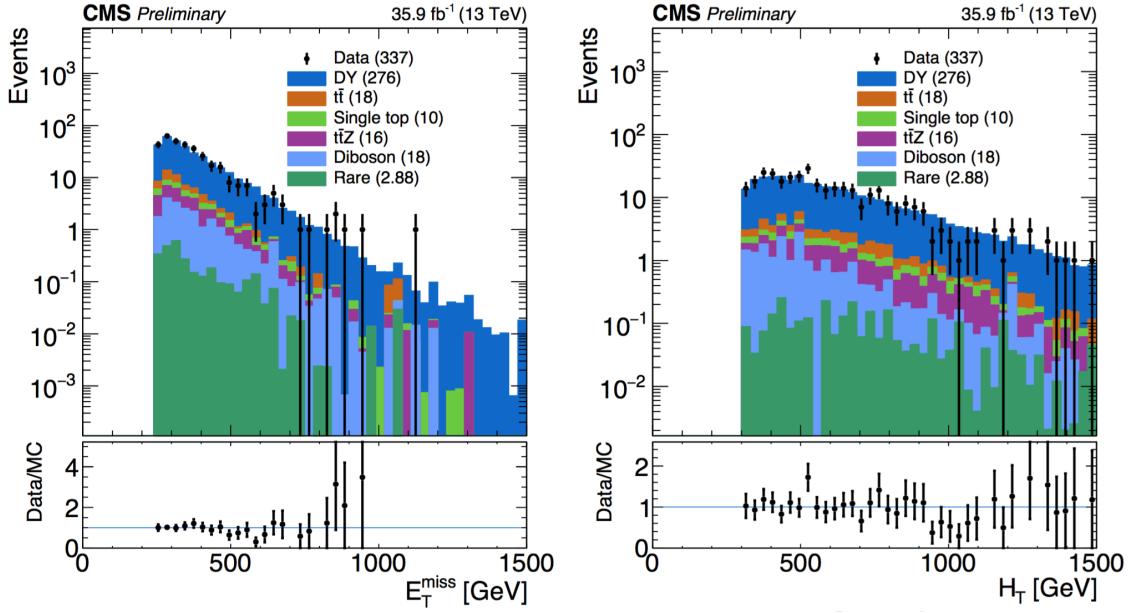


**Figure 6.10:**  $H_T$  and  $m_{T2}$  distributions applying the  $S_\gamma(N_j)$  scale factor.

consistent with each other but the inclusive region ( $N_b \geq 0$ ) has a lower overall uncertainty. The method used to calculate the normalization scale factor requires that the  $N_j$ -dependent shape correction factors already be applied. Then, the  $R_{norm}$  factor can be extracted from the ratio of the total event yield in data to that in the simulation. This factor is found to be:

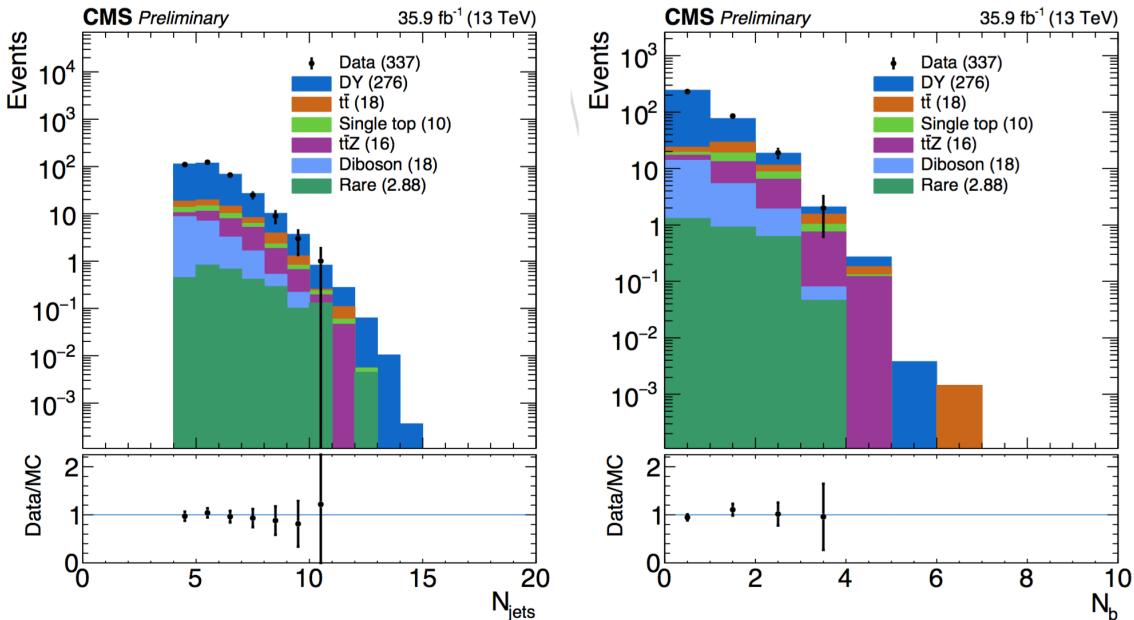
$$R_{norm} = 1.070 \pm 0.085,$$

where the uncertainty includes only the associated statistical uncertainties on data and simulation. This uncertainty is found to be propagated to the final background prediction, see subsection 6.6.1.



**Figure 6.11:** Shown are data/MC comparisons for the  $p_T^{\text{miss}}$  (left) and  $H_T$  (right) distributions after applying both the  $N_j$ -dependent shape corrections ( $S_\gamma$ ) and the global normalization scale factor ( $R_{\text{norm}}$ ).

1422 Data/MC comparisons are shown in [Figure 6.11](#) and [Figure 6.12](#) after applying  $R_{\text{norm}}$   
1423 for several distributions in the study. With this final global scale factor all the required  
1424 ingredients for the central value of the  $Z \rightarrow \nu\bar{\nu}$  background prediction are obtained.



**Figure 6.12:** Shown are data/MC comparisons for the  $N_j$  (left) and  $N_b$  (right) distributions after applying both the  $N_j$ -dependent shape corrections ( $S_\gamma$ ) and the global normalization scale factor ( $R_{\text{norm}}$ ).

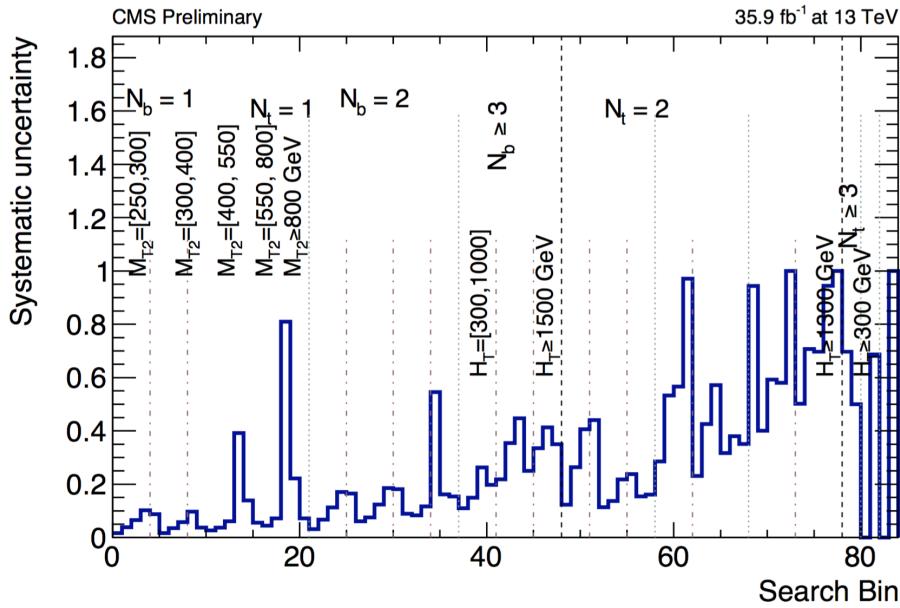
1425 **6.6 Results**

1426 In this section the results for the final estimation of the  $Z \rightarrow \nu\bar{\nu}$  are presented.

1427 The current study includes preliminary results using only data obtained at the CMS de-  
1428 tector during 2016. The results for this study are intended to confirm the assumption that  
1429 the additional  $\gamma + \text{jets}$  control region introduced in this analysis reduce the overall uncer-  
1430 tainties obtained in the 2016 analyses (described in [Chapter 5](#)). Furthermore, this study is  
1431 intended as a benchmark for future analyses of the SUSY stop group based in Fermilab  
1432 and will be the method used for the 2017 CMS data.

1433 **6.6.1 Systematics**

1434 Two categories of uncertainties for the  $Z \rightarrow \nu\bar{\nu}$  prediction are considered: uncertain-  
1435 ties that are associated to the use of MC simulation and the uncertainties specifically  
1436 associated to the background prediction method. Several sources are acknowledged in the  
1437 first category mentioned such as PDF and renormalization/factorization scale choices, jet  
1438 and  $p_T^{\text{miss}}$  energy scale uncertainties b-tag scale factor uncertainties, and trigger efficiency  
1439 uncertainties. Given that the simulation sample is normalized to data in the tight control  
1440 region, uncertainties associated with the luminosity and cross-section are excluded. In  
1441 addition, the overall  $Z \rightarrow \nu\bar{\nu}$  statistical uncertainty from MC simulation is also taken into  
1442 account.



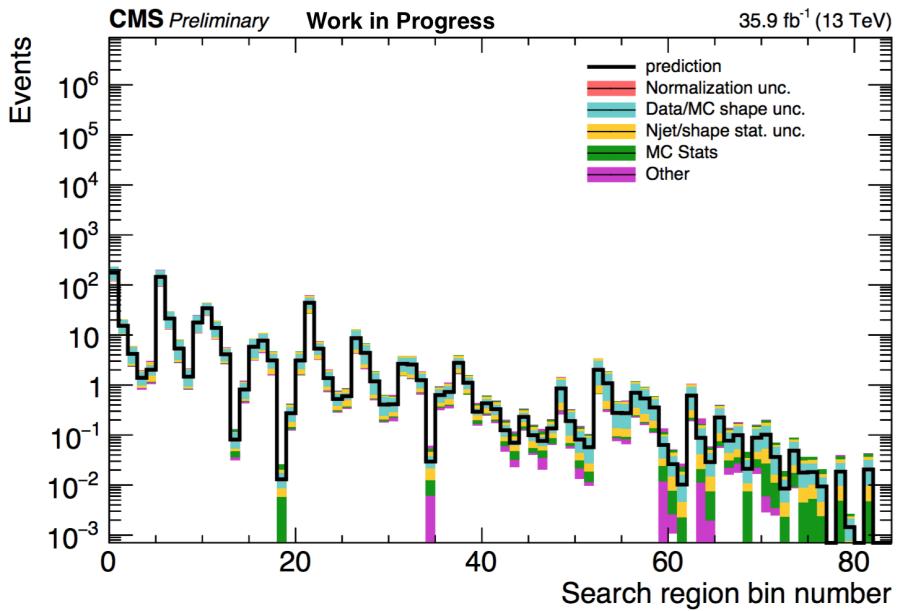
**Figure 6.13:** Systematic uncertainty in the final prediction, as a function of the search bin, associated to the MC statistics.

The statistical uncertainty associated with each bin in the MC is propagated as a systematic uncertainty. The relative uncertainty per bin can be see in Figure 6.13. It shows that the uncertainties for the MC vary from as low as 1% up to 81% and even 100% in some regions. Since the final estimation is scaled using the global normalization factor from the tight  $\mu\mu$  control region ( $R_{norm}$ ), the total uncertainty, due to limited amounts of events in data, is propagated in the final prediction. This is also true for the  $S_\gamma(N_j)$  scale factor, in which the residual differences in search variables other than  $N_j$  are evaluated in the loose photon control region. Both the uncertainty arising from the  $N_j$  re-weighting as well as the residual differences are evaluated together. The uncertainty from  $R_{norm}$  is propagated as a flat value of 7.9% uncertainty per each search bin.

## 6.6.2 $Z \rightarrow \nu\bar{\nu}$ Estimation for the Search Bins

The final estimation for the  $Z \rightarrow \nu\bar{\nu}$  background calculated for all 84 search bins is shown in Figure 6.14. The statistical uncertainty in bins that have zero events is treated as the average weight (the sum of the weights squared over the weight) times the poisson error on 0 which is 1.8. This average weight is calculated on the basis of a relaxed cut in which  $N_b \geq 2$  is required. For comparison, a cut in which  $N_t > 2$  where two tops are

<sub>1460</sub> fake for the  $Z \rightarrow \nu\bar{\nu}$  is used.



**Figure 6.14:**  $Z \rightarrow \nu\bar{\nu}$  background prediction for all search bins, including the breakdown of the various uncertainties.

1461 **Chapter 7**

1462 **Conclusion**

1463 There are three major topics of research that were discussed in this dissertation: The  
1464 simulation studies involving the counting of L1-stubs for the HL-LHC CMS Inner Tracker  
1465 upgrade ([chapter 4](#)), the overall 2016 search for SUSY in the all-hadronic channel us-  
1466 ing a customized top-tagger ([chapter 5](#)) and the improvements made for the estimation  
1467 of the  $Z \rightarrow \nu\bar{\nu} +$  jets background using an additional control region from  $\gamma +$  jets events  
1468 ([chapter 6](#)). These studies were explained in detail in their respective chapters and their  
1469 individual results are provided. A summary of the most important results from each study  
1470 is provided in this chapter.

1471 **7.1 L1 Stub Counting for the HL-LHC CMS Tracker Up-**  
1472 **grade**

1473 Results from this study (detailed in [chapter 4](#)) reflect the overall effects that were ex-  
1474 pected beforehand. The removal of discs from the standard pixel geometry (consisting  
1475 of 8 small and 4 large discs) results in a noticeable reduction of stub hits in the upgraded  
1476 CMS Outer Tracker. This effect is specially apparent if the disc that is removed is closer to  
1477 the interaction point, due to the much larger volume of particles that are present in this re-  
1478 gion. Therefore, the reduction in stubs is more pronounced when a small disc is removed  
1479 (as in the case of the  $7s4l$  geometry) than if a large disc is removed (as in the  $8s3l$  pixel

1480 geometry). The reason for this effect stems from the fact that as particles travel through  
1481 the various layers of the Inner Tracker material, some of them are bound to interact with  
1482 it, producing particles that did not originate from the initial proton-proton collision. The  
1483 stubs produced via such processes are considered to be “fake” stubs. To confirm these  
1484 findings, an additional study was conducted using a sample that was virtually indistin-  
1485 guishable from the standard pixel geometry, but with the second disc on the positive side  
1486 “turned off” or “dead”. The results from this study confirm the initial findings and shows  
1487 that there is indeed a correlation between the average number of stubs detected in the  
1488 Outer Tracker and the total amount of material present in the upgraded Inner Tracker. An  
1489 important factor that needs to be taken into account when interpreting these results is the  
1490 re-optimization of the disc positions after removing a disc in the different pixel geome-  
1491 tries considered. This feature could provide a possible explanation as to why the 6s3l  
1492 geometry, which has two less small discs than the standard geometry (and one less large  
1493 one), was found to have less of an effect on the average number of stubs than the 7s4l  
1494 geometry.

## 1495 7.2 Search for SUSY in the All-Hadronic Channel

1496 The analysis presented in [chapter 5](#) shows the results of a search for SUSY in the  
1497 0-lepton final state using a customized top-tagger. The data was obtained from proton-  
1498 proton collisions at the CMS detector during 2016 with a total integrated luminosity of  
1499  $35.9 \text{ fb}^{-1}$  at a center-of-mass energy of 13 TeV. The search was conducted by speci-  
1500 fying 84 non-overlapping regions of phase space with varying requirements on the  $N_b$ ,  
1501  $N_t$ ,  $p_T^{\text{miss}}$ ,  $H_T$  and  $m_{\text{T2}}$  variables ([subsection 5.2.5](#)). Several dominant and non-dominant  
1502 backgrounds were identified and estimated to account for all the majority of the processes  
1503 that were seen in the collected data. The estimation procedures and their respective sys-  
1504 tematic and statistical uncertainties are discussed in [section 5.3](#). The total background  
1505 prediction vs. data for all 84 search bins ([Figure 5.7](#)) shows no statistically significant de-  
1506 viation from the predicted SM background. The biggest sources background were shown

1507 to be the  $t\bar{t}$  and  $W+\text{jets}$  processes, followed by  $Z(\nu\bar{\nu})+\text{jets}$ , which were seen to be dom-  
 1508 inant in regions with a high  $p_T$  threshold. Meanwhile, the contributions from the QCD  
 1509 multijet and rare backgrounds are found to be nearly negligible in all of the 84 search bins.  
 1510 Exclusion limits were calculated from these results for each of the signal models used, by  
 1511 applying a binned likelihood fit on the data. The likelihood function was obtained for  
 1512 each of the 84 search regions as well as for each of the background data control sam-  
 1513 ples from the product of the Poisson probability density function. Exclusion limits were  
 1514 placed on the top squark, gluino and LSP production cross-sections with a 95% confi-  
 1515 dence level (CL), calculated using a modified frequentist approach with the  $\text{CL}_s$  criterion  
 1516 and asymptotic results for the test statistic. The 95% CL exclusion limits obtained for the  
 1517  $T2tt$  model, which consists of direct top squark production, excludes top squark masses  
 1518 up to 1020 GeV and LSP masses up to 430 GeV. For the  $T1tttt$  model, gluino masses of  
 1519 up to 2040 GeV and LSP masses up to 1150 GeV are excluded, with corresponding limits  
 1520 of 2020 and 1150 GeV for the  $T1ttbb$  model, 2020 and 1150 GeV for the  $T5tttt$  model,  
 1521 and 1810 and 1100 GeV for the  $T5ttcc$  model.

## 1522 7.3 Estimation of the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ Background with a 1523 Hybrid Method

1524 Chapter 6 presents a different method (than the one briefly described in chapter 5)  
 1525 for estimating the total amount of  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  events per search bin. The new method  
 1526 makes use of an additional “loose  $\gamma + \text{jets}$ ” control sample for the estimation of the  $S_\gamma(N_j)$   
 1527 shape correction factor, in addition to the “tight  $Z \rightarrow \mu\mu + \text{jets}$ ” sample (used in the 2016  
 1528 analysis) to calculate the  $R_{norm}$  correction factor (subsection 5.3.2). The  $\gamma + \text{jets}$  sample  
 1529 was chosen to substitute the previously used “loose  $Z \rightarrow \mu\mu + \text{jets}$ ” control region, due to  
 1530 its much higher cross-section and kinematic similarity to the  $Z \rightarrow \nu\bar{\nu} + \text{jets}$  process at high  
 1531  $p_T$ . The implementation of this method sought to refine the results obtained in 2016 by  
 1532 reducing the statistical uncertainties that stem from the low branching fraction of  $Z \rightarrow \mu\mu +$

1533 jets events. The  $\gamma+$  jets control region was found to have a high purity and a low fake-rate  
1534 in the regime of high  $p_T$  that was being studied. The  $N_j$ -dependent  $S_\gamma^i$  was obtained for  
1535 each bin i in  $N_j$  from a comparison between data and the  $\gamma+$  jets MC, after subtracting  
1536 the other backgrounds from data and normalizing both samples to 1. This resulted in the  
1537  $S_\gamma$  scale factor plot depicted in [Figure 6.6](#), on the right. As can be seen from this plot, the  
1538 statistical uncertainty of the various correction factors is small due to the large number  
1539 of events available. Using these values, the  $R_{norm}$  normalization correction factor was  
1540 obtained from the tight  $\mu\mu$  control sample as  $R_{norm} = 1.070 \pm 0.085$ . Both of these  
1541 scale factors were then applied to the final estimation of the  $Z \rightarrow \nu\bar{\nu} +$  jets background  
1542 ([Figure 6.14](#)).

1543     **Appendix A**

1544     **References**

- 1545     [1] C. Lefevre, “The CERN accelerator complex. Complexe des accelerateurs du  
1546       CERN.” Available on the [CERN document server](#), Dec 2008.
- 1547     [2] T. Sakuma and T. McCauley, “Detector and Event Visualization with SketchUp  
1548       at the CMS Experiment,” *J. Phys. Conf. Ser.*, vol. 513, p. 022032, 2014.  
1549       [doi:10.1088/1742-6596/513/2/022032](https://doi.org/10.1088/1742-6596/513/2/022032).
- 1550     [3] S. Spannagel, *CMS Pixel Detector Upgrade and Top Quark Pole Mass Determina-*  
1551       *tion*, vol. 27. Springer, 2017. [doi:10.1007/978-3-319-58880-3](https://doi.org/10.1007/978-3-319-58880-3).
- 1552     [4] CMS Collaboration, “CMS Physics Technical Design Report Volume 1: Detec-  
1553       tor Performance and Software,” Technical Design Report CERN-LHCC-2006-001,  
1554       CERN (2006).
- 1555     [5] CMS Collaboration, “The CMS Experiment at the CERN LHC,” *JINST*, vol. 3,  
1556       p. S08004, 2008. [doi:10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- 1557     [6] CMS Collaboration, “Performance of the CMS Hadron Calorimeter with Cos-  
1558       mic Ray Muons and LHC Beam Data,” *JINST*, vol. 5, p. T03012, 2010.  
1559       [doi:10.1088/1748-0221/5/03/T03012](https://doi.org/10.1088/1748-0221/5/03/T03012).
- 1560     [7] J. Nielsen, “Fundamentals of LHC Experiments,” in *Proceedings, Theoretical Ad-*  
1561       *vanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory*

- 1562       *and Its Applications: From meV to the Planck Scale: Boulder, Colorado, USA, June*  
1563       *1-25, 2010*, pp. 127–152, 2011. [arXiv:1106.2516](https://arxiv.org/abs/1106.2516).
- 1564 [8] C. Borschensky, M. Kramer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and  
1565       X. Portell, “Squark and gluino production cross sections in pp collisions at  $\sqrt{s}$   
1566       = 13, 14, 33 and 100 TeV,” *Eur. Phys. J.*, vol. C74, no. 12, p. 3174, 2014.  
1567       doi:[10.1140/epjc/s10052-014-3174-y](https://doi.org/10.1140/epjc/s10052-014-3174-y).
- 1568 [9] S. Ferrara, R. Kallosh, A. Linde, and M. Poratti, “Minimal Supergrav-  
1569       ity Models of Inflation,” *Phys. Rev.*, vol. D88, no. 8, p. 085038, 2013.  
1570       doi:[10.1103/PhysRevD.88.085038](https://doi.org/10.1103/PhysRevD.88.085038).
- 1571 [10] A. B. Lahanas and D. V. Nanopoulos, “The Road to No Scale Supergravity,” *Phys.*  
1572       *Rept.*, vol. 145, p. 1, 1987. doi:[10.1016/0370-1573\(87\)90034-2](https://doi.org/10.1016/0370-1573(87)90034-2).
- 1573 [11] G. F. Giudice and R. Rattazzi, “Theories with gauge mediated supersymmetry break-  
1574       ing,” *Phys. Rept.*, vol. 322, pp. 419–499, 1999. doi:[10.1016/S0370-1573\(99\)00042-3](https://doi.org/10.1016/S0370-1573(99)00042-3).
- 1576 [12] S. P. Martin and J. D. Wells, “Implications of gauge-mediated supersymmetry break-  
1577       ing with vector-like quarks and a 125 GeV Higgs boson,” *Phys. Rev.*, vol. D86,  
1578       p. 035017, 2012. doi:[10.1103/PhysRevD.86.035017](https://doi.org/10.1103/PhysRevD.86.035017).
- 1579 [13] R. Barbieri and G. F. Giudice, “Upper Bounds on Supersymmetric Particle Masses,”  
1580       *Nucl. Phys.*, vol. B306, pp. 63–76, 1988. doi:[10.1016/0550-3213\(88\)90171-X](https://doi.org/10.1016/0550-3213(88)90171-X).
- 1581 [14] R. Barbieri *et al.*, “R-parity violating supersymmetry,” *Phys. Rept.*, vol. 420, pp. 1–  
1582       202, 2005. doi:[10.1016/j.physrep.2005.08.006](https://doi.org/10.1016/j.physrep.2005.08.006).
- 1583 [15] D. Dercks, H. Dreiner, M. E. Krauss, T. Opferkuch, and A. Reinert, “R-  
1584       Parity Violation at the LHC,” *Eur. Phys. J.*, vol. C77, no. 12, p. 856, 2017.  
1585       doi:[10.1140/epjc/s10052-017-5414-4](https://doi.org/10.1140/epjc/s10052-017-5414-4).
- 1586 [16] P. Langacker, “Grand Unified Theories and Proton Decay,” *Phys. Rept.*, vol. 72,  
1587       p. 185, 1981. doi:[10.1016/0370-1573\(81\)90059-4](https://doi.org/10.1016/0370-1573(81)90059-4).

- 1588 [17] CMS Collaboration, “Interpretation of Searches for Supersymmetry with  
1589 simplified Models,” *Phys. Rev.*, vol. D88, no. 5, p. 052017, 2013.  
1590 [doi:10.1103/PhysRevD.88.052017](https://doi.org/10.1103/PhysRevD.88.052017).
- 1591 [18] P. Bechtle *et al.*, “Killing the cMSSM softly,” *PoS*, vol. EPS-HEP2015, p. 139, 2015.  
1592 [doi:10.22323/1.234.0139](https://doi.org/10.22323/1.234.0139).
- 1593 [19] L. Evans and P. Bryant, “LHC Machine,” *JINST*, vol. 3, p. S08001, 2008.  
1594 [doi:10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001).
- 1595 [20] R. Rosenfeld, “Physics Beyond the Standard Model,” in *Proceedings, 8th CERN–*  
1596 *Latin-American School of High-Energy Physics (CLASHEP2015): Ibarra, Ecuador,*  
1597 *March 05–17, 2015*, pp. 159–164, 2016. [doi:10.5170/CERN-2016-005.159](https://doi.org/10.5170/CERN-2016-005.159).
- 1598 [21] N. Doble, L. Gatignon, K. Hübner, and E. Wilson, *The Super Proton Synchrotron*  
1599 (*SPS*): *A Tale of Two Lives*, pp. 135–177. World Scientific Publishing Co, 2017.
- 1600 [22] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Col-  
1601 linder,” *JINST*, vol. 3, p. S08003, 2008. [doi:10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- 1602 [23] LHCb Collaboration, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005,  
1603 2008. [doi:10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005).
- 1604 [24] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3,  
1605 p. S08002, 2008. [doi:10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- 1606 [25] K. Garrett and G. Duda, “Dark Matter: A Primer,” *Adv. Astron.*, vol. 2011,  
1607 p. 968283, 2011. [doi:10.1155/2011/968283](https://doi.org/10.1155/2011/968283).
- 1608 [26] R. C. G. Landim, *Dark Sector Cosmology*. PhD thesis, Sao Paulo U., 2017.  
1609 [arXiv:1702.06024](https://arxiv.org/abs/1702.06024).
- 1610 [27] K. W. Bell *et al.*, “Vacuum phototriodes for the CMS electromagnetic  
1611 calorimeter endcap,” *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 2284–2287, 2004.  
1612 [doi:10.1109/TNS.2004.836053](https://doi.org/10.1109/TNS.2004.836053).

- 1613 [28] H. Bichsel, D. E. Groom, and S. R. Klein, “Passage of particles through matter,”  
1614 2004.
- 1615 [29] CMS Collaboration, “Alignment of the CMS Muon System with Cosmic-Ray  
1616 and Beam-Halo Muons,” *JINST*, vol. 5, p. T03020, 2010. [doi:10.1088/1748-0221/5/03/T03020](https://doi.org/10.1088/1748-0221/5/03/T03020).
- 1618 [30] “Muon Detectors.” <http://cms.web.cern.ch/news/muon-detectors>.  
1619 Accessed: 2018-04-08.
- 1620 [31] A. Tapper and D. Acosta, “CMS Technical Design Report for the Level-1 Trigger  
1621 Upgrade,” 2013.
- 1622 [32] W. Adam *et al.*, “The CMS high level trigger,” *Eur. Phys. J.*, vol. C46, pp. 605–667,  
1623 2006. [doi:10.1140/epjc/s2006-02495-8](https://doi.org/10.1140/epjc/s2006-02495-8).
- 1624 [33] F. Beaudette, *The CMS Particle Flow Algorithm*, 2013.
- 1625 [34] A. K. Nayak, “Reconstruction of physics objects in the CMS detector,” *PoS*,  
1626 vol. CHARGED2012, p. 010, 2012.
- 1627 [35] CMS Collaboration, “CMS: The computing project. Technical design report,” Tech-  
1628 nical Design Report CERN-LHCC-2005-023, CERN (2005).
- 1629 [36] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis frame-  
1630 work,” *Nucl. Instrum. Meth.*, vol. A389, pp. 81–86, 1997. [doi:10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).
- 1632 [37] M. Lipinski, “The Phase-1 upgrade of the CMS pixel detector,” *JINST*, vol. 12,  
1633 no. 07, p. C07009, 2017. [doi:10.1088/1748-0221/12/07/C07009](https://doi.org/10.1088/1748-0221/12/07/C07009).
- 1634 [38] G. Apollinari, I. Bejar Alonso, O. Bruning, M. Lamont, and L. Rossi, “High-  
1635 Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report,” 2015.  
1636 [doi:10.5170/CERN-2015-005](https://doi.org/10.5170/CERN-2015-005).

- 1637 [39] D. Contardo, M. Klute, J. Mans, L. Silvestris, and J. Butler, “Technical Proposal for  
1638 the Phase-II Upgrade of the CMS Detector,” 2015. [CMS-TDR-15-02](#).
- 1639 [40] J. Pilot, “Future of the CMS Muon System: Upgrades and Aging,” *PoS*,  
1640 vol. ICHEP2016, p. 281, 2016. [doi:10.22323/1.282.0281](#).
- 1641 [41] K. Klein, “The Phase-2 Upgrade of the CMS Tracker,” 2017. [CMS-TDR-014](#).
- 1642 [42] S. Mersi, “Phase-2 Upgrade of the CMS Tracker,” *Nucl. Part. Phys. Proc.*, vol. 273-  
1643 275, pp. 1034–1041, 2016. [doi:10.1016/j.nuclphysbps.2015.09.162](#).
- 1644 [43] C. Foudas, A. Rose, J. Jones, and G. Hall, “A Study for a tracking trigger at first  
1645 level for CMS at SLHC,” in *Proceedings, eleventh Workshop on Electronics for*  
1646 *LHC and Future Experiments, Heidelberg, Germany, 12-16 September 2005*, p. 90,  
1647 2005. [arXiv:physics/0510227](#).
- 1648 [44] N. De Maio, G. Bianchi, S. Martina, Z. Drasal, G. Hugo, and S. Mersi.  
1649 <http://tklayout.web.cern.ch>.
- 1650 [45] CMS Collaboration, “Search for supersymmetry in proton-proton collisions at 13  
1651 tev using identified top quarks,” *Phys. Rev. D*, vol. 97, p. 012007, Jan 2018.  
1652 [doi:10.1103/PhysRevD.97.012007](#).
- 1653 [46] H. E. Haber and L. Stephenson Haskins, “Supersymmetric Theory and Models,” in  
1654 *Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating*  
1655 *the Next Discoveries in Particle Physics (TASI 2016) Boulder, CO, USA, June 6-July*  
1656 *1, 2016*, 2017. [arXiv:1712.05926](#).
- 1657 [47] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti- $k(t)$  jet clustering algorithm,”  
1658 *JHEP*, vol. 04, p. 063, 2008. [doi:10.1088/1126-6708/2008/04/063](#).
- 1659 [48] CMS Collaboration, “A Cambridge-Aachen (C-A) based Jet Algorithm for boosted  
1660 top-jet tagging,” 2009.

- 1661 [49] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop,” *JHEP*, vol. 05,  
1662 p. 146, 2014. [doi:10.1007/JHEP05\(2014\)146](https://doi.org/10.1007/JHEP05(2014)146).
- 1663 [50] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.  
1664 [doi:10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).
- 1665 [51] T. Junk, “Confidence level computation for combining searches with small statis-  
1666 tics,” *Nucl. Instrum. Meth.*, vol. A434, pp. 435–443, 1999. [doi:10.1016/S0168-9002\(99\)00498-2](https://doi.org/10.1016/S0168-9002(99)00498-2).
- 1668 [52] A. L. Read, “Presentation of search results: The CL(s) technique,” *J. Phys.*, vol. G28,  
1669 pp. 2693–2704, 2002. [,11(2002)], [doi:10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- 1670 [53] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for  
1671 likelihood-based tests of new physics,” *Eur. Phys. J.*, vol. C71, p. 1554, 2011. [Er-  
1672 ratum: *Eur. Phys. J.*C73,2501(2013)], [doi:10.1140/epjc/s10052-013-2501-z](https://doi.org/10.1140/epjc/s10052-013-2501-z).
- 1673 [54] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking  
1674 Period,” Tech. Rep. CMS-PAS-LUM-17-001, CERN, Geneva, 2017. [CMS-PAS-  
1675 LUM-17-001](#).
- 1676 [55] CMS Collaboration, “Comparison of the  $Z/\gamma^*$  + jets to  $\gamma$  + jets cross  
1677 sections in  $pp$  collisions at  $\sqrt{s} = 8$  TeV,” *JHEP*, vol. 10, p. 128,  
1678 2015. [Erratum: *JHEP*04,010(2016)], [doi:10.1007/JHEP04\(2016\)010](https://doi.org/10.1007/JHEP04(2016)010),  
1679 [doi:10.1007/JHEP10\(2015\)128](https://doi.org/10.1007/JHEP10(2015)128).
- 1680 [56] CMS Collaboration, “Search for supersymmetry in the multijet and missing trans-  
1681 verse momentum final state in  $pp$  collisions at 13 TeV,” *Phys. Lett.*, vol. B758,  
1682 pp. 152–180, 2016. [doi:10.1016/j.physletb.2016.05.002](https://doi.org/10.1016/j.physletb.2016.05.002).
- 1683 [57] CMS Collaboration, “Search for supersymmetry in multijet events with missing  
1684 transverse momentum in proton-proton collisions at 13 TeV,” *Phys. Rev.*, vol. D96,  
1685 no. 3, p. 032003, 2017. [doi:10.1103/PhysRevD.96.032003](https://doi.org/10.1103/PhysRevD.96.032003).

- 1686 [58] “Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets  
1687 Search,” 2009. [CMS-PAS-SUS-08-002](#).
- 1688 [59] F. Micheli, “Cut Based Photon ID for Run 2.” <https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedPhotonIdentificationRun2>.  
1689 Accessed: 2018-01-27.
- 1690 [60] J. Berryhill, “Electron Detection at CMS.” The Egamma Physics Object Group,  
1691 2009. [ppt](#).
- 1692 [61] CMS Collaboration, “Performance of Photon Reconstruction and Identification with  
1693 the CMS Detector in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV,” *JINST*, vol. 10,  
1694 no. 08, p. P08010, 2015. [doi:10.1088/1748-0221/10/08/P08010](https://doi.org/10.1088/1748-0221/10/08/P08010).
- 1695 [62] A. M. Sirunyan *et al.*, “Particle-flow reconstruction and global event description  
1696 with the CMS detector,” *JINST*, vol. 12, no. 10, p. P10003, 2017. [doi:10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003).
- 1697 [63] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens, and  
1698 M. Selvaggi, “DELPHES 3, A modular framework for fast simulation of a generic  
1699 collider experiment,” *JHEP*, vol. 02, p. 057, 2014. [doi:10.1007/JHEP02\(2014\)057](https://doi.org/10.1007/JHEP02(2014)057).
- 1700 [64] M. Klasen, C. Klein-Bosing, and H. Poppenborg, “Prompt photon produc-  
1701 tion and photon-jet correlations at the LHC,” *JHEP*, vol. 03, p. 081, 2018.  
1702 [doi:10.1007/JHEP03\(2018\)081](https://doi.org/10.1007/JHEP03(2018)081).
- 1703 [65] S. Folgueras, “Baseline muon selections for Run-II.” <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>. Accessed: 2018-05-13.
- 1704 [66] CMS Collaboration, “Measurement of the Z/gamma\*+jets/photon+jets cross section  
1705 ratio in pp collisions at  $\sqrt{s}=8$  TeV,” 2014. [CMS-PAS-SMP-14-005](#).