

1 **Measurement of the Z Invisible background for Stop**
2 **Quark Searches in all Hadronic Channel at $\sqrt{s} = 13$ TeV**

3 by

4 Andrés J. Abreu Nazario

5 A thesis presented for the degree of

6 MASTER 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 _____
16 Date

17 President, Graduate Committee

18 _____
19 Héctor Méndez, Ph.D.

20 _____
21 Date

22 Member, Graduate Committee

23 _____
24 Samuel Santana Colón, Ph.D.

25 _____
26 Date

27 Member, Graduate Committee

28 _____
29 Rafael A. Ramos, Ph.D.

30 _____
31 Date

32 Chairperson of the Department

²⁵ Abstract

²⁶ The Large Hadron Collider (LHC) is the world's largest and most powerful particle ac-
²⁷celerator designed to discover physics Beyond the Standard Model (BSM). The Standard
²⁸Model (SM) of fundamental particles and their interactions has been extremely successful
²⁹in describing phenomena in the atomic and subatomic realms in the last 100 years. After
³⁰the discovery of a boson with properties consistent with the SM Higgs boson at the LHC
³¹(by CMS and ATLAS Experiments in 2012), attention has shifted to searches for new
³²physics that can explain the fine-tuned cancellation of large quantum corrections required
³³to stabilize the Higgs boson mass at 125 GeV. Among BSM scenarios, Supersymmetry
³⁴(SUSY) is a new symmetry between fermions and bosons that provides an elegant mech-
³⁵anism to mitigate this hierarchy problem and overcome shortcomings of the Standard
³⁶Model. However, the data from the LHC has so far shown no signs of BSM, however,
³⁷stringent limits have been placed on its possible signatures. The future High Luminosity
³⁸run of the LHC (HL-LHC) from 2023-2035 will enable us to make precise measurements
³⁹of several of the SM parameters, in particular, the scalar sector and open the search for
⁴⁰new physics proposed by many BSM models and identify any new physics hint, including
⁴¹SUSY, that could also possibly be observed during the current run. This requires the CMS
⁴²collaboration to perform a series of upgrades that will ensure the capabilities of the CMS
⁴³detector to match the HL-LHC running conditions. One of these upgrades is to build an
⁴⁴entire new silicon Tracker that includes Pixel Detector and Outer Tracker. The work pre-
⁴⁵sented in this thesis is aligned with evolution of the CMS experiment. A simulation study
⁴⁶is performed to determine the effect of HL-LHC Pixel Detector on stubs (matching hit
⁴⁷pairs from the same particle track on two adjacent silicon layers of the modules in Outer
⁴⁸Tracker) which is key to future possible BSM discoveries and SUSY in particular. A
⁴⁹new method to measure the Z-invisible ($Z \rightarrow \nu\bar{\nu}$) background for the stop quark searches
⁵⁰in all-hadronic channel is presented. The method based on γ +jets provides an improve-
⁵¹ment over the method used for Z-invisible background in analysis presented here that
⁵²searches for SUSY in the all-hadronic channel with missing transverse momentum and a
⁵³customized top tagger. The data was collected by the CMS Detector from proton-proton
⁵⁴collisions at the LHC at a center-of-mass energy of 13 TeV with an integrated luminosity
⁵⁵of 35 fb^{-1} collected in 2016.

56 **Resumen**

57 El Large Hadron Collider (LHC) constituye el acelerador de partículas más poderoso
58 y más grande del mundo, diseñado con el propósito de descubrir evidencia de fenómenos
59 físicos que trascienden el modelo estándar. El modelo estándar de partículas elemen-
60 tales, y las interacciones que describe, han sido extremadamente exitosos en explicar
61 fenómenos de la física al nivel atómico y sub-atómico durante el pasado siglo. Luego
62 del descubrimiento de un bosón con propiedades consistentes con el bosón de Higgs
63 (como se ha descrito por el modelo estándar), el enfoque se ha tornado hacia búsquedas de
64 fenómenos físicos que sean capaces de explicar la cancelación de correcciones cuánticas
65 grandes, necesarias para obtener un bosón de Higgs con una masa de 125 GeV. Dentro de
66 los modelos propuestos, la Supersimetría (SUSY) consiste de un tipo nuevo de simetría
67 entre fermiones y bosones que provee una solución elegante para este problema de jer-
68 arquía, entre otras deficiencias. Sin embargo, hasta el momento no se ha logrado observar
69 ninguna evidencia (de data obtenida del LHC) de fenómenos que van más allá del mod-
70 elo estándar, aunque si se han impuesto límites sobre las posibles señales de búsqueda.
71 La implementación del LHC de alta luminosidad (HL-LHC, por sus siglas en inglés) en-
72 tre 2023-2035, va a proveer la posibilidad de realizar medidas precisas de varios de los
73 parámetros del modelo estándar facilitando la búsqueda de varios modelos propuestos, in-
74 cluyendo SUSY. Esto está sujeto a que la colaboración de CMS actualice el detector para
75 que pueda funcionar bajo las nuevas condiciones de funcionamiento del LHC. Una de es-
76 tas actualizaciones incluye la construcción total de un nuevo sistema de rastreo de silicio,
77 el cual incluye un detector de pixeles y un rastreador externo. El trabajo presentado en
78 esta tesis está alineado con la evolución del detector CMS. Se realizó un estudio basado
79 en simulación con el propósito de determinar el efecto que va tener el nuevo detector de
80 pixeles sobre la formación de “stubs” (definido como la coincidencia de un par de señales
81 en capas de silicio adyacentes en los componentes del rastreador externo), los cuales for-
82 man la base para el descubrimiento de física nueva, en particular SUSY. Se presenta un
83 método nuevo para estimar las interacciones de fondo relacionadas el decaimiento de Z
84 en un par de neutrinos ($Z \rightarrow \nu\bar{\nu}$) para búsquedas del top squark en estados finales con cero
85 leptones. El método está basado en la realización que una muestra de eventos de $\gamma+jets$
86 debe exhibir mejores resultados que el método utilizado previamente para búsquedas de
87 SUSY en estados finales de cero leptones, con momentum transverso desaparecido y un
88 identificador de top-quarks. La data utilizada para el estudio fue obtenida de colisiones
89 de protones generadas por el LHC en el detector CMS con un centro de masa de 13 TeV
90 y una luminosidad integrada de 35 fb^{-1} durante el 2016.

⁹¹ Acknowledgements

⁹² In the first place I would like to extend my thanks to the UPRM CMS group. In partic-
⁹³ ular, I want to thank my advisor Prof. Sudhir Malik for all the help, assistance and support
⁹⁴ he provided throughout these past two years, without which none of this would have been
⁹⁵ possible. I want to thank our UPRM postdoc, Scarlet Norberg, who helped me build up
⁹⁶ my high-energy physics and computational tools knowledge from scratch throughout my
⁹⁷ stay at Fermilab and who made it possible to finish this work to completion. I would also
⁹⁸ like to thank Prof. Juan Eduardo Ramirez and the other UPRM physics professors and
⁹⁹ faculty members. Thanks to them, this research dissertation has been made a reality.

¹⁰⁰
¹⁰¹ I would also like to extend my thanks to all the postdocs and students in the SUSY stop
¹⁰² search group in Fermilab for all their help and understanding: Prof. Kenichi Hatakeyama,
¹⁰³ Joe Pastika, Zhenbin Wu, V. Daniel Elvira, Ken Call, Hui Wang, Koushik Mandal.

¹⁰⁴
¹⁰⁵ Many thanks also go to the members of the LPCFPix group who helped me with my
¹⁰⁶ simulation work: Prof. Corrinne Mills, Juliette Alimena, Susan Dittmer. Their knowledge
¹⁰⁷ and assistance in the subject of the Phase-2 simulations was very helpful and it would not
¹⁰⁸ have been possible to complete my studies without them.

¹⁰⁹
¹¹⁰ I would also like to thank the National Science Foundation (NSF Award ID 1506168)
¹¹¹ for funding various of the physics conferences I attended as well as several of my trips to
¹¹² Fermilab, allowing me to make this thesis a reality.

¹¹³
¹¹⁴ Finally, I want to thank the LHC Physics Center (LPC) at Fermilab for providing the
¹¹⁵ necessary funding for my research during a full year at Fermilab (May 2017 - May 2018).
¹¹⁶ This opportunity was probably one of the most influential in the completion of this final
¹¹⁷ work, and it would have definitely been impossible to achieve without it. Special thanks
¹¹⁸ go to Cecilia Gerber and Boaz Klima for this opportunity, and to all of the other people at
¹¹⁹ the LPC whom I had the pleasure to work with and learn from.

¹²⁰ List of Figures

121	Figure 2.1:	The CERN Accelerator Complex [?].	4
122	Figure 2.2:	The CMS Detector Layout [?].	5
123	Figure 2.3:	Overview of the CMS Tracker Layout [?].	6
124	Figure 2.4:	Geometrical layout of the CMS Electromagnetic Calorimeter [?].	8
125	Figure 2.5:	Layout of the CMS ECAL illustrating its various components [?].	9
126	Figure 2.6:	Geometrical layout of the HCAL showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [?].	10
127	Figure 2.7:	Quarter-view of the CMS muon system (left) and a muon in the transverse plane leaves a curved trajectory across the four muon detector stations (right).	11
128	Figure 2.8:	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 [?].	13
129	Figure 2.9:	Phase-2 Pixel Detector Layout[ref].	14
130	Figure 3.1:	Schematic representation of the Standard Model particles.	16
131	Figure 3.2:	Schematic representation of the Standard Model particles with the added SUSY particles.	18
132	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).	22
133	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.	22
134	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 ϕ 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.	23
135	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.	24

157	Figure 4.5:	Number of stubs per event for standard pixel geometry ($8s4l$) . . .	25
158	Figure 4.6:	Comparison of total number of stubs per event between the stan-	
159		dard layout and the other three considered. $8s4l$ vs $8s3l$ (left),	
160		$8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right).	26
161	Figure 4.7:	Comparison for number of stubs per event between pixel geom-	
162		etry $8s4l$ with a dead disc (blue) and both pixel geometry $8s4l$	
163		(right plot) and $7s4l$ (left plot) (blue). Both the standard geom-	
164		etry and $7s4l$ are shown in blue in their respective plots.	26
165	Figure 4.8:	Comparison of total number of stubs per η value between the	
166		standard layout and the other three layouts considered. All four	
167		different layouts exhibit the same distribution. $8s4l$ vs $8s3l$ (left),	
168		$8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right).	27
169	Figure 4.9:	Comparison of total number of stubs per each of the 6 barrel	
170		layers. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$	
171		vs $6s3l$ (right). Tracker Barrel layers shown below.	27
172	Figure 4.10:	Comparison of total number of stubs per each of the 5 Endcap	
173		Double Discs. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$	
174		vs $6s3l$ (right). Tracker Endcap Discs shown below.	28
175	Figure 4.11:	Comparison of total number of stubs for all rings across the five	
176		endcap double-discs. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle)	
177		and $8s4l$ vs $6s3l$ (right). The double discs are shown below.	28
178	Figure 5.1:	Simplified model diagrams show of the various Supersymmetry	
179		signals considered in this analysis: the T2tt model (top left), the	
180		T1tttt model (top right), the T1ttbb model (middle left), the T5tttt	
181		(middle right), and the T5ttcc model (bottom).	31
182	Figure 5.2:	Efficiency of the top quark tagger as a function of generator-level	
183		top quark p_T for the monojet (red boxes), dijet (magenta trian-	
184		gles), and trijet (green upside-down triangles) categories and for	
185		their combination (blue circles), as determined using T2tt signal	
186		events with a top squark mass of 850 GeV and an LSP mass of	
187		100 GeV. The vertical bars indicate the statistical uncertainties. .	33
188	Figure 5.3:	Search region definitions in the kinematic variables.	35
189	Figure 5.4:	Comparison between the total SM background from simulation	
190		and CMS data for N_t (top left), N_b (top right), m_{T2} (bottom left)	
191		and p_T^{miss} (bottom right). Total SM backgrounds and signals are	
192		scaled to same data yield for a shape comparison.	36
193	Figure 5.5:	Distribution of p_T^{miss} for both the single-electron (left) and single-	
194		muon (right) SB data samples compared to predictions for SM	
195		processes. The hatched bands indicate the statistical uncertain-	
196		ties in the total SM prediction. Note that the data and the predic-	
197		tions for all backgrounds except that for $t\bar{t}$, single top quark, and	
198		W+jets events are identical between the left and right plots.	37

199	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.	38
200			
201			
202			
203			
204			
205	Figure 5.7:	Observed events compared to the corrected SM background predictions for all 84 search regions in the full data of 35.9 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.	39
206			
207			
208			
209			
210	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 [?] 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.	40
211			
212			
213			
214			
215			
216			
217			
218			
219	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.	41
220			
221			
222			
223	Figure 6.1:	Leading-order Feynman diagrams for Z+jets and γ +jets processes.	43
224	Figure 6.2:	Shown are both the p_T^γ (left) and $p_T^{\gamma(miss)}$ (right) distributions before applying any corrections. $p_T^{\gamma(miss)}$ is obtained by adding the p_T^γ to the total p_T^{miss} in every event.	45
225			
226			
227	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/ γ POG. As can be seen, the efficiency (purity) is seen to agree with the values of the loose photon ID/isolation selection.	46
228			
229			
230			
231			
232			
233			
234	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.	46
235			
236			
237			
238			
239	Figure 6.5:	Results of study of the Z+jets to γ +jets cross-section ratio for both data and MadGraph simulation.	49
240			
241	Figure 6.6:	Jet multiplicity and the associated S_γ scale factor in the loose photon control region before any corrections are applied.	50
242			
243	Figure 6.7:	N_{jet} distribution in the tight $\mu\mu$ control region after S_γ corrections.	50

244	Figure 6.8: p_T^γ (left) and $p_T^{\gamma(miss)}$ (right) distributions after applying the $S_\gamma(N_j)$	
245	scale factor. Comparing to Figure 6.2, an improvement in the	
246	agreement between data/MC can be observed.	51
247	Figure 6.9: N_b distribution before (left) and after (right) applying the $S_\gamma(N_j)$	
248	scale factor.	51
249	Figure 6.10: H_T and m_{T2} distributions applying the $S_\gamma(N_j)$ scale factor.	52
250	Figure 6.11: Shown are data/MC comparisons for the p_T^{miss} (left) and H_T (right)	
251	distributions after applying both the N_j -dependent shape correc-	
252	tions (S_γ) and the global normalization scale factor (R_{norm}).	52
253	Figure 6.12: Shown are data/MC comparisons for the N_j (left) and N_b (right)	
254	distributions after applying both the N_j -dependent shape correc-	
255	tions (S_γ) and the global normalization scale factor (R_{norm}).	53
256	Figure 6.13: Systematic uncertainty in the final prediction, as a function of the	
257	search bin, associated to the MC statistics.	54
258	Figure 6.14: $Z \rightarrow \nu\bar{\nu}$ background prediction for all search bins, including the	
259	breakdown of the various uncertainties.	55

260 **Contents**

261	List of Figures	iv
262	1 Introduction	1
263	2 Large Hadron Collider and the Compact Muon Solenoid	3
264	2.1 The LHC Machine	3
265	2.2 The CMS Detector	4
266	2.2.1 Silicon Tracking System	6
267	2.2.2 Electromagnetic Calorimeter	7
268	2.2.3 Hadron Calorimeter	8
269	2.2.4 Magnet	10
270	2.2.5 Muon Detector	10
271	2.2.6 Trigger and Data Acquisition	11
272	2.2.7 Event Reconstruction	12
273	2.2.8 Future Upgrade of Pixel Detector	13
274	3 Theoretical Motivation	16
275	3.1 The Standard Model	16
276	3.2 Indications of Physics Beyond the Standard Model	17
277	3.3 Supersymmetric extension of the Standard Model	18
278	3.3.1 Simplified Models Approach to Supersymmetry	19
279	4 Optimizing the Phase-2 Pixel Geometry for the Level-1 Trigger in the Outer Tracker	21
280	4.1 Upgraded Phase-2 Tracker	21
281	4.1.1 Phase-2 Tracker Layout	22
282	4.1.2 Tracker Input to the L1-Trigger	22
283	4.2 Tracker Stub Simulation	23
284	4.2.1 Results of Effect of the Phase 2 Pixel Detector on the L1 Stub Rate	25
285	4.3 Results	29
286	5 Search for SUSY in the 0-Lepton Final State	30
287	5.1 Introduction	30
288	5.2 Analysis Description	30
289	5.2.1 Signal Models	31
290	5.2.2 Top Tagger	32
291	5.2.3 Lepton and Track Veto	33
292	5.2.4 Baseline Event Selection	34

294	5.2.5	Search Bin Definition	34
295	5.3	Background Estimation	34
296	5.3.1	Background from $t\bar{t}$, Single Top Quark and W+jets Events	35
297	5.3.2	The $Z \rightarrow \nu\bar{\nu}$ Background	37
298	5.3.3	The QCD Multijet Background	38
299	5.3.4	Background From Rare and Other Processes	39
300	5.4	Results and Interpretation	39
301	6	Estimation of the $Z \rightarrow \nu\bar{\nu}$ Background	42
302	6.1	Introduction	42
303	6.2	The Irreducible $Z \rightarrow \nu\bar{\nu}$ Background	42
304	6.3	The Loose γ +jets Control Region	43
305	6.3.1	Photon ID and Isolation	43
306	6.3.2	Photon Selection	44
307	6.3.3	Photon Purity and Fake Rate	45
308	6.4	The $Z \rightarrow \mu^+\mu^-$ Control Region	47
309	6.4.1	Muon ID and Isolation	48
310	6.4.2	Muon Selection in the Tight Control Region	48
311	6.5	Analysis	49
312	6.5.1	Shape Correction Using the γ + jets Control Sample	49
313	6.5.2	Normalization Correction Using the tight $Z \rightarrow \mu^+\mu^-$ Control Sample	51
314	6.6	Results	53
315	6.6.1	Systematics	53
316	6.6.2	$Z \rightarrow \nu\bar{\nu}$ Estimation for the Search Bins	54
317	7	Conclusion	56
318	7.1	L1 Stub Counting for the HL-LHC CMS Tracker Upgrade	56
319	7.2	Search for SUSY in the All-Hadronic Channel	57
320	7.3	Estimation of the $Z \rightarrow \nu\bar{\nu}$ + jets Background with a Hybrid Method	57

³²¹ Chapter 1

³²² Introduction

³²³ The recently discovered Higgs boson at the LHC could be the final particle required
³²⁴ in the highly successful theory of the Standard Model. However, the SM is not with-
³²⁵ out its shortcomings, for instance one requires fine-tuned cancellations of large quantum
³²⁶ corrections for the Higgs boson to have a mass at the electroweak symmetry breaking
³²⁷ scale known as the hierarchy problem. The magnitude of this fine-tuning makes one sus-
³²⁸ pect that there is some dynamical mechanism which makes this fine-tuning “natural”.
³²⁹ SUSY, a popular extension of the SM offers a well-motivated explanation: the dominant
³³⁰ role in cancelling the quantum effects would come from scalar partners to the bottom/top
³³¹ quarks (sbottoms/stops), due to the large 3rd generation Yukawa couplings in the SM,
³³² and fermionic partners to the Higgs boson (higgsinos). The lightest SUSY particle, the
³³³ neutralino, could also be a candidate for Dark Matter. Alternatively, the naturalness of
³³⁴ electroweak symmetry breaking could be effectively explained by non-SUSY BSM mod-
³³⁵ els involving strong dynamics. Searching for sparticles in the all-hadronic mode is very
³³⁶ attractive because it makes up a large portion of branching fractions for typical signals.
³³⁷ The final state involves events with high jet multiplicity and missing transverse energy
³³⁸ and this search method could eventually decipher the mysteries surrounding the Higgs
³³⁹ boson.

³⁴⁰

³⁴¹ The work presented in this thesis is about discovering physics beyond the Standard
³⁴² Model and is three fold: (1) A simulation study is performed to determine the effect of
³⁴³ CMS Phase-2 Pixel Detector on stubs (matching hit pairs from the same particle track on
³⁴⁴ two adjacent layers of silicon modules) (2) An analysis involving a search for Supersym-
³⁴⁵ metry in the all-hadronic channel with missing transverse momentum and a customized
³⁴⁶ top tagger (3) An improved new method based on $\gamma + \text{jets}$ to determine $Z \rightarrow \nu\bar{\nu}$ background
³⁴⁷ in search for the SUSY signature.

³⁴⁸

³⁴⁹ This thesis work was carried out at the LPC (LHC Physics Center) at Fermilab. The
³⁵⁰ funding for this work was provided by the **NSF (National Science Foundation - NSF**
³⁵¹ **Award ID 1506168) and LPC (Guest and Visitor) Fellowship.**

³⁵²

³⁵³ [Chapter 2](#) presents a basic description of the Large Hadron Collider and CMS Detec-
³⁵⁴ tor (including trigger, data acquisition and Phase-2 upgrade plans), which are the basic
³⁵⁵ tools used to collect data in order to successfully complete the thesis work.

³⁵⁶

357 Chapter 3 gives a brief motivation to look for extensions for physics Beyond the Stan-
358 dard Model for which Supersymmetry (SUSY) is the most popular choice. It describes
359 SUSY briefly and in relation to Simplified Models few of which are used in the analysis
360 described in the thesis.

361
362 Chapter 4 is dedicated to a study optimizing the geometry of Phase-2 Pixel Detector
363 for the High Luminosity-LHC. A simulation is performed using the $t\bar{t}$ sample to estimate
364 the impacting of pixel detector on the formation of stubs used in the outer tracker to de-
365 termine particle tracks that are used in the Level-1 (L1) trigger. This work was done as
366 part of service work for the CMS experiment.

367
368 Chapter 5 presents an analysis on the search for SUSY in the 0-Lepton final state
369 where an inclusive search for events with final states that contain missing transverse mo-
370 mentum and reconstructed top quarks is performed. The signal models used in this study
371 include the production of three different types of SUSY particles. Two of which are the
372 top squark and the gluino, the supersymmetric partners of the SM top and gluon, respec-
373 tively. The third one is the neutralino that is considered to be the lightest SUSY particle
374 (LSP) under the Minimal Supersymmetric Standard Model (MSSM) and a possible can-
375 didate for Dark Matter. Background from $t\bar{t}$, single top quark, W+jets and $Z \rightarrow \nu\bar{\nu}$ events
376 was estimated.

377
378 Chapter 6 details an improved new method based on γ +jets to determine $Z \rightarrow \nu\bar{\nu}$
379 background in search for the SUSY signature. This estimation procedure builds upon the
380 $Z \rightarrow \nu\bar{\nu}$ method used in SUSY analysis described in Chapter 5 (also referred to as 2016
381 analysis) and aims to refine the overall background calculation as well as reduce the un-
382 certainties associated with the previous method. To accomplish this, an additional γ +jets
383 control sample (CS) is used in conjunction with the tight $Z \rightarrow \mu^+\mu^-$ control region used in
384 the 2016 analysis estimation. The new γ +jets CS provides a more data-driven estimation
385 procedure with the added benefit of a substantially larger production cross-section than
386 Z +jets process used before.

387
388 The conclusions and results of each chapter are presented in the corresponding chap-
389 ter.

390
391 This thesis work has been presented at several internal meetings of the CMS Experi-
392 ment and at the following international meetings and conferences:

- 393 1. **Andrés Abreu** gave a talk “*Estimation of the Z Invisible Background for Searches*
394 *for Supersymmetry in the All-Hadronic Channel*” at “APS April 2018: American
395 Physical Society April Meeting 2018, 14-17 Apr 2018”, Columbus, OH
- 396 2. **Andrés Abreu** gave a talk “*Phase-2 Pixel upgrade simulations*” at the “USLUA
397 Annual meeting: 2017 US LHC Users Association Meeting, 1-3 Nov 2017”, Fer-
398 milab, Batavia, IL
- 399 3. **Andrés Abreu** gave a talk, “*Direct production of top squark pairs in all-hadronic*
400 *channel*” presentation at “FNAL50: Fermilab 50th Anniversary Symposium and
401 Users Meeting and New Perspectives Workshop, Fermilab, Batavia, June 5-8, 2017”,
402 Batavia, IL

403 **Chapter 2**

404 **Large Hadron Collider and the
405 Compact Muon Solenoid**

406 **2.1 The LHC Machine**

407 The Large Hadron Collider (LHC) [?], located at the European Organization for Nu-
408 clear Research (CERN) complex in Geneva Switzerland, is the world's largest and most
409 powerful particle collider as well as the most complex experimental device ever assem-
410 bled. The main motivation for its construction was uncovering the nature behind elec-
411 troweak symmetry breaking due to the Higgs mechanism as well as to search for new
412 physical phenomena Beyond the Standard Model (BSM) [?]. Since its completion in
413 2008, the LHC has helped researchers obtain a plethora of significant results in the field
414 of particle physics and it successfully achieved its purpose with the discovery of the Higgs
415 Boson in 2012.

416

417 The LHC consists of several superconducting magnets and accelerating structures ar-
418 ranged along a 27-kilometer circumference ring which serve to boost and bend the proton
419 (or other heavy ion) beams along its path. In order for the particle beams to maintain the
420 orbit while they are accelerated, a magnetic field of 8.3 T is produced by the supercon-
421 ducting magnets along the circular tunnel, 100 meters below the ground. The magnets are
422 maintained at a temperature of 1.9 K by using liquid helium. The two beams are made
423 to travel in opposite directions and in separate beam pipes which are kept at ultrahigh
424 vacuum and are accelerated to near the speed of light. They are then collided head-on at
425 any of the four interaction points located around the LHC ring. Approximately 1.5×10^{11}
426 protons are generated every 25 ns with an energy of 26 GeV at the Proton Synchrotron
427 (PS). The proton beam is then accelerated to an energy of 450 GeV at the Super Proton
428 Synchrotron (SPS) [?] before being delivered to the LHC. At design level, each bunch is
429 accelerated to an energy of 7 TeV and are made to collide at a centre-of-mass energy of
430 up to 14 TeV with a frequency of 40 MHz. The current proton beams achieve an instan-
431 taneous luminosity on the order of $10^{32} \text{ cm}^{-2} \cdot \text{s}^{-1}$ and a total integrated luminosity of 37
432 fb^{-1} was reached during 2016 [?].

433

434 The four main detectors comprising the LHC machine are CMS, ATLAS [?], LHCb
435 [?] and ALICE [?]. Both CMS and ATLAS are general purpose detectors whose initial
436 designs had the detection of the SM Higgs boson, with its wide range of decay modes,

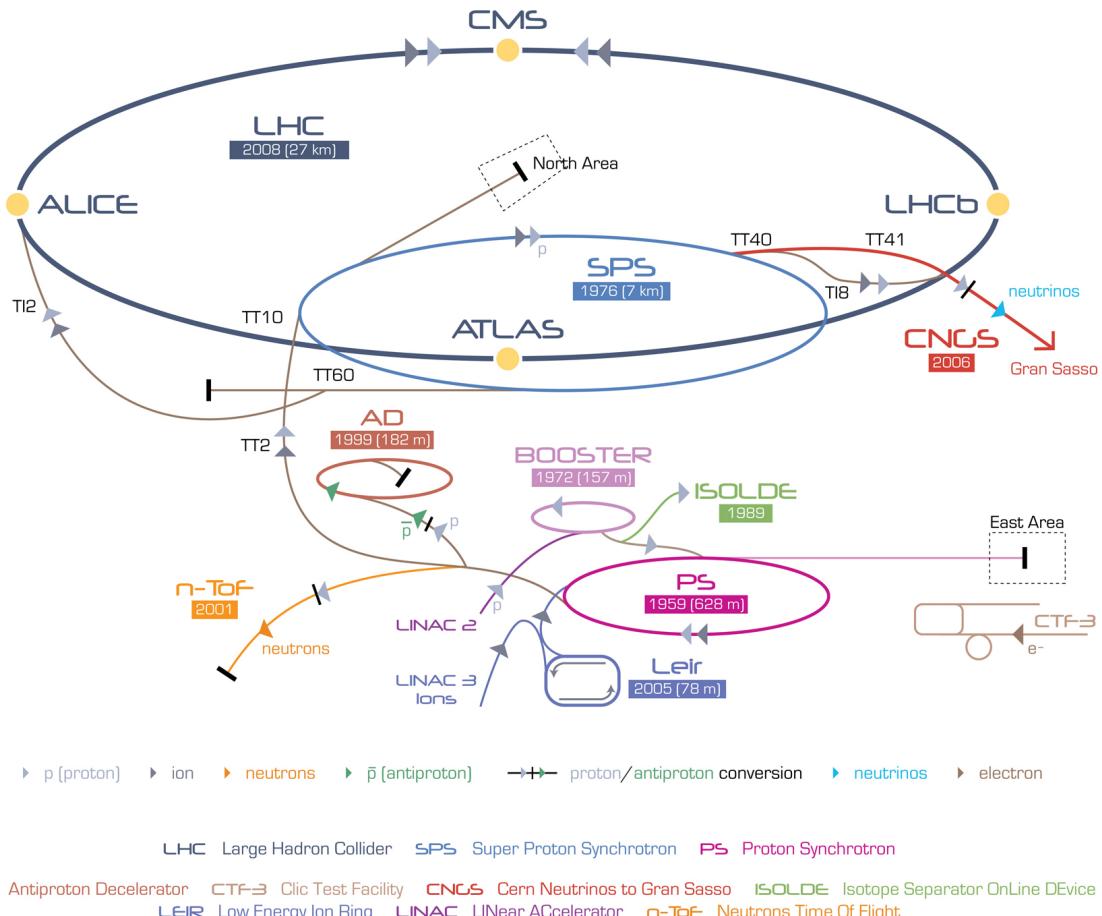


Figure 2.1: The CERN Accelerator Complex [?].

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

2.2 The CMS Detector

The Compact Muon Solenoid (CMS) detector is a general purpose particle detector designed to investigate various physical phenomena concerning the SM and beyond it, such as Supersymmetry, Extra Dimensions and Dark Matter. As its name implies, the detector is a solenoid which is constructed around a superconducting magnet capable of producing a magnetic field of 3.8 T. The magnetic coil is 13m long with an inner diameter

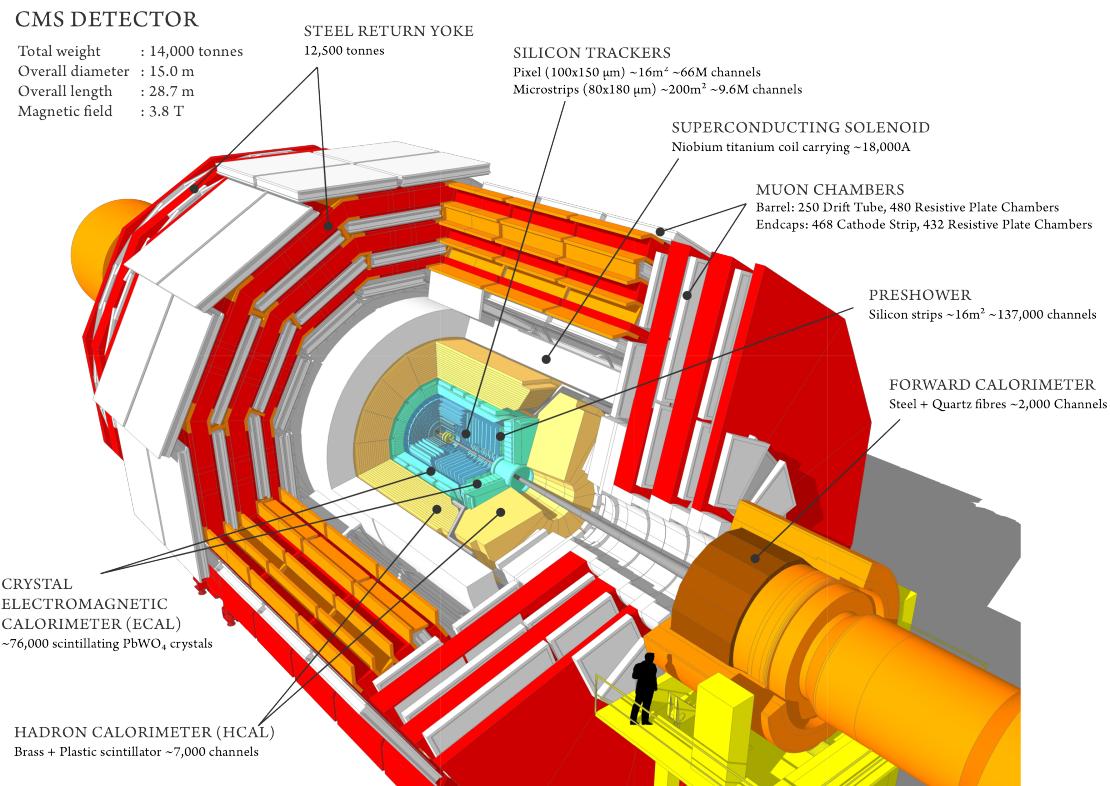


Figure 2.2: The CMS Detector Layout [?].

of 6m, making it the largest superconducting magnet ever constructed. The CMS detector itself is 21m long with a diameter of 15m and it has a weight of approximately 14,000 tonnes. The CMS experiment is one of the largest scientific collaborations in the history of mankind with over 4,000 participants from 42 countries and 182 institutions.

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

- A magnet with large bending power and high performance muon detector for good 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 2.2. 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.

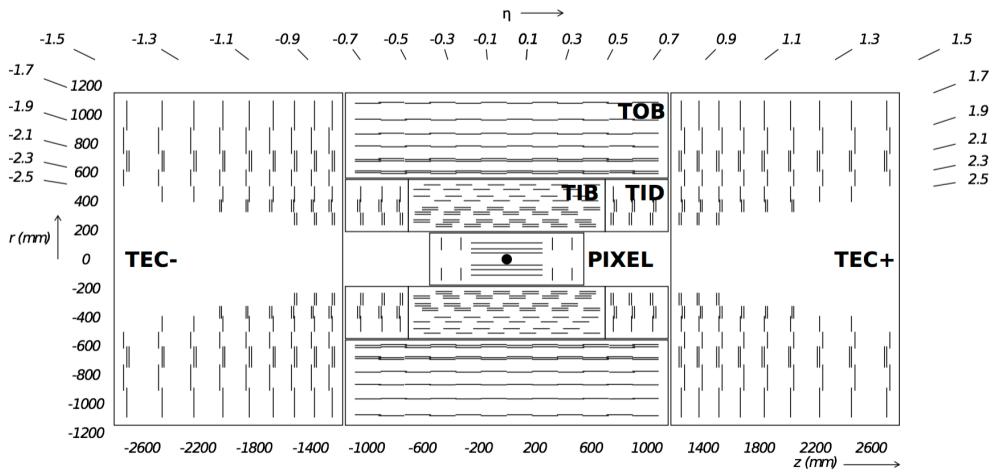


Figure 2.3: Overview of the CMS Tracker Layout [?].

479 The coordinate system for the CMS detector design uses a right-hand rule convention
 480 centered around the ideal interaction point to describe the positions of objects in the experiment.
 481 The z-axis is defined along the direction of the LHC beam, with the x-axis
 482 pointing towards the center of the LHC ring. In terms of polar coordinates then, r is the
 483 radial distance from the center of the pipe, the polar angle θ is measured against the z-axis
 484 and the azimuthal angle ϕ is measured with respect to the x-axis. However, the pseudo-
 485 rapidity η is generally preferred over the polar angle θ . The pseudorapidity is defined
 486 as:

$$487 \quad \eta = -\ln \tan \frac{\theta}{2}.$$

488 2.2.1 Silicon Tracking System

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

501 The CMS tracker is built in a cylindrical manner around the interaction point and has
 502 a diameter of 2.5 m and a length of 5.8 m. It is comprised of a pixel detector with three
 503 barrel layers, positioned at a distance between 4.4 cm and 10.2 cm from the interaction
 504 region, and a silicon strip tracker with 10 barrel detection layers extending to a radius of
 505 1.1 m. Each system is made complete by endcaps at opposite sides of the barrel, consist-
 506 ing of 2 discs for the pixel detector and 3 plus 9 discs for the strip tracker, extending the

508 acceptance region of the tracker up to a pseudorapidity $|\eta| < 2.5$.

509

510 The pixel detector is the part of the CMS tracking system closest to the interaction
511 region and consists of 3 barrel layers (BPix) and 2 endcap discs (FPix). It is responsible
512 for providing precise tracking points in $r\phi$ and z , a feature that is required for the small
513 impact parameter resolution, needed for good secondary vertex reconstruction. The de-
514 tector contains 1440 modules covering an area of approximately 1 m^2 making up a total of
515 66 million pixels. The sensors were designed using an n-on-n concept from $320 \mu\text{m}$
516 thick silicon wafers and are fabricated with read-out chips (ROCs) that are bump-bonded
517 to the sensor in standard $0.25 \mu\text{m}$ CMOS technology. Each of the pixels has a pitch size
518 of $100 \times 150 \mu\text{m}^2$, which corresponds to an occupancy of about 10^{-4} per bunch crossing.

519

520 The silicon strip tracker is built surrounding the pixel detector and consists of three
521 large subsystems. The Tracker Inner Barrel and Disks (TIB/TID) extend to a radius of
522 55 cm and are composed of four barrel layers, completed by three disks at each end. The
523 TIB/TID employs the use of $320 \mu\text{m}$ thick silicon micro-strip sensors in order to deliver
524 up to 4 $r\phi$ measurements on a trajectory. Surrounding the TIB/TID is the Tracker Outer
525 Barrel (TOB), which consists of 6 barrel layers and has an outer radius of 116 cm. The
526 TOB extends symmetrically in z between $\pm 118 \text{ cm}$ and provides an additional 6 $r\phi$
527 measurements for a trajectory. Beyond the TOB's z range lie the Tracker Endcaps (TEC+
528 and TEC-, where the sign indicates their position respect to z). Each TEC consists of 9
529 discs with up to 7 rings of silicon micro-strip detectors, providing up to 9 additional ϕ
530 measurements per trajectory. The CMS silicon strip tracker has a total active silicon area
531 of 198 m^2 and is composed of 15,148 sensor modules.

532 2.2.2 Electromagnetic Calorimeter

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

539

540 The ECAL is composed of two main sub-systems – the barrel calorimeter (EB) and the
541 endcap calorimeter (EE) – and is completed by a preshower calorimeter (ES), as shown
542 in [Figure 2.5](#). It covers a solid angle up to a pseudorapidity of $|\eta| < 3$, with the EB ex-
543 tending in the range of $|\eta| < 1.479$ and the EE covering a range of $1.479 < |\eta| < 3$. Both
544 subsystems are composed of lead tungstate (PbWO_4) scintillator crystals which provide
545 a fast response time and high radiation tolerance, both crucial requirements for optimal
546 performance at LHC operating conditions. In addition, the properties of the PbWO_4 crys-
547 tals (high density, short radiation length and small Moli  re radius) led to the design of a compact
548 calorimeter with fine granularity.

549

550 The barrel part of the ECAL (EB) is composed of specially designed avalanche pho-
551 todiodes (APD). It consists of 61,200 crystals, forming a total volume of 8.14 m^3 and
552 weighing about 67.4 t. The crystals that form the EB are 230 mm long with a cross-

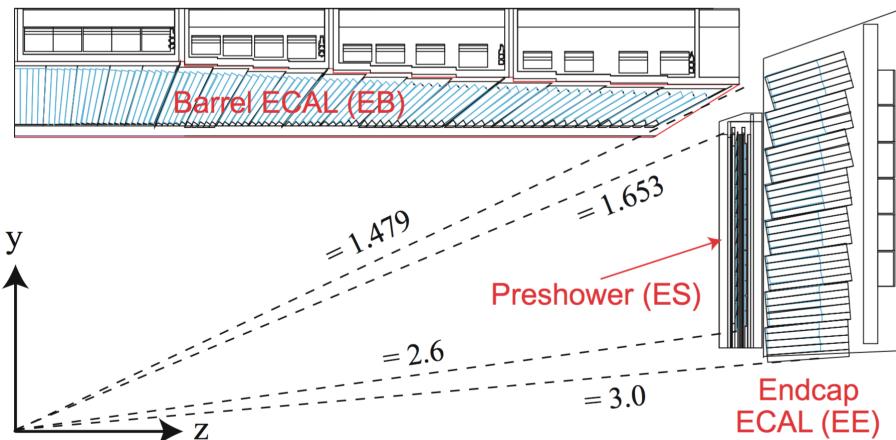


Figure 2.4: Geometrical layout of the CMS Electromagnetic Calorimeter [?].

553 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
 554 in pairs within thin-walled alveolar structures called submodules. The submodules are
 555 assembled into different types of modules that differ by their location with respect to η
 556 and contain 400 or 500 crystals each. These modules are then arranged in sets of four
 557 modules, called supermodules, and contain 1700 crystals each. Thus, the EB is composed
 558 of two half-barrels, each consisting of 18 supermodules.

559

560 In contrast, the photodetectors used in the endcap section of the ECAL are vacuum
 561 phototriodes (VPT) [?]. Each of the endcaps contain 7,324 crystals which, in total, oc-
 562 cupy a volume of 2.90 m^3 and weigh about 24.0 t. The crystals in the EE are 220 mm in
 563 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
 564 rear. They are all identical in shape and are arranged in mechanical units of 5×5 crystals,
 565 called supercrystals (SC), which consist of carbon-fibre alveola structures. Each of the
 566 endcaps are divided into two semi-circular structures, called *Dees*, which hold a total of
 567 3,662 crystals.

568

569 The ES preshower is located before the EE detector and spans a pseudorapidity range
 570 of $1.653 < |\eta| < 2.6$. Its main purpose is to identify neutral pions in the endcaps as
 571 well as to improve the determination of electrons and photons with high granularity. The
 572 preshower consists of two layers and has a total thickness of 20 cm. The first layer is
 573 conformed by lead radiators which initiate electromagnetic showers from incoming pho-
 574 tons and electrons. Meanwhile, the second layer is composed of silicon strips which are
 575 capable of measuring the deposited energy and the transverse shower profiles.

576 2.2.3 Hadron Calorimeter

577 The CMS Hadron calorimeter (HCAL) conforms the next layer of the CMS detector.
 578 It is a sampling calorimeter that consists of alternating layers of massive absorbing brass
 579 plates and plastic scintillator tiles and is of particular importance for the measurement of
 580 hadron jet energy and p_T^{miss} . The HCAL detector is located in between the outer extent of
 581 the ECAL ($R = 1.77 \text{ m}$) and the inner extent of the magnet coil ($R = 2.95 \text{ m}$). Similar to
 582 the other CMS subsystems, it's composed of a barrel part (HB) and an endcap part (HE).

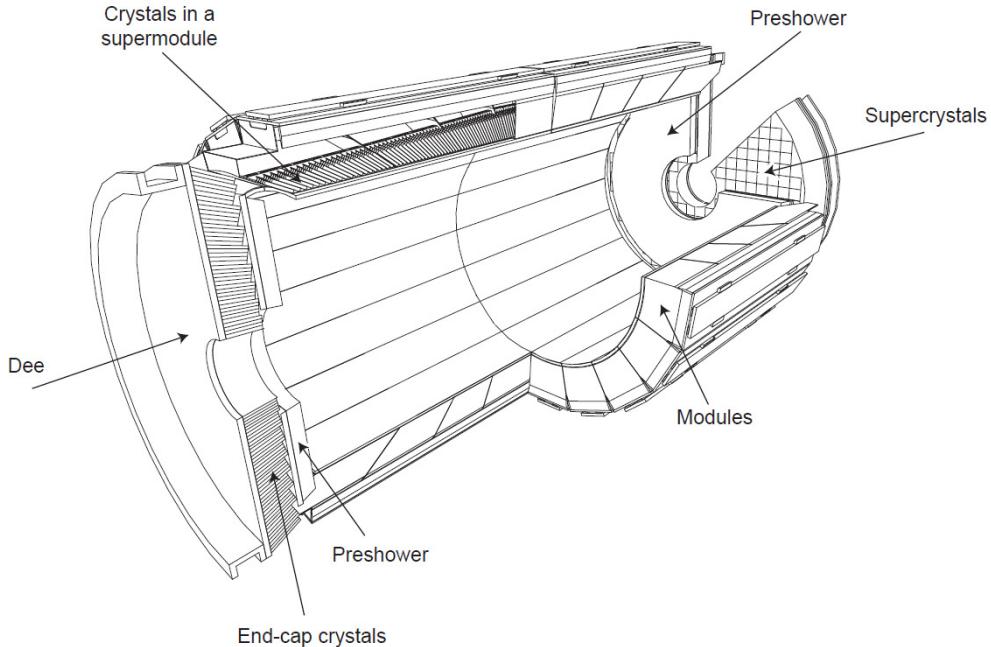


Figure 2.5: Layout of the CMS ECAL illustrating its various components [?].

583 In addition, it features a tail-catching outer calorimeter (HO), located outside the magnet,
 584 and a forward calorimeter (HF) in the very forward region near the beam line. A layout
 585 of the HCAL system can be seen in [Figure 2.6](#).

586

587 The barrel component of the HCAL is a sampling calorimeter which covers the pseudo-
 588 rapidity range $|\eta| < 1.3$. It consists of both the HB and the HO detectors. The reason
 589 behind separating the barrel detector into the HB and HO is due to the limited amount
 590 of space available for the barrel detector. The HB is located within the superconducting
 591 magnet coil and is supplemented by the HO in between the outer solenoid coil and the
 592 muon chambers. Therefore, the HO acts as a tail-catcher in order to improve the jet en-
 593 ergy measurements and p_T^{miss} , with the solenoid in between acting as absorber material.
 594 The HB consists of two half-barrel sections, identified as HB+ and HB- due to their ge-
 595ometrical location, which are composed of 36 identical azimuthal wedges. The wedges,
 596 which are constructed out of flat brass absorber plates, are aligned parallel to the beam
 597 axis and are segmented into four azimuthal angle (ϕ) sections.

598

599 The HE covers a significant amount of the pseudorapidity in the range of $1.3 < |\eta| <$
 600 3, a region containing about 34% of the particles produced in the final state. Due to
 601 the high luminosity of the LHC, the HE is required to have a high radiation tolerance at
 602 $|\eta| \simeq 3$, as well as being capable of handling high counting rates. Similar to the HB, the
 603 HE is also composed of brass absorber plates and scintillator plates which are read out
 604 by wavelength shifting fibers. The light captured by the scintillators merges within the
 605 wavelength shifting fibers and then it's read out by hybrid photo-diodes. The scintillators
 606 are partitioned in towers with an area of $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$.

607

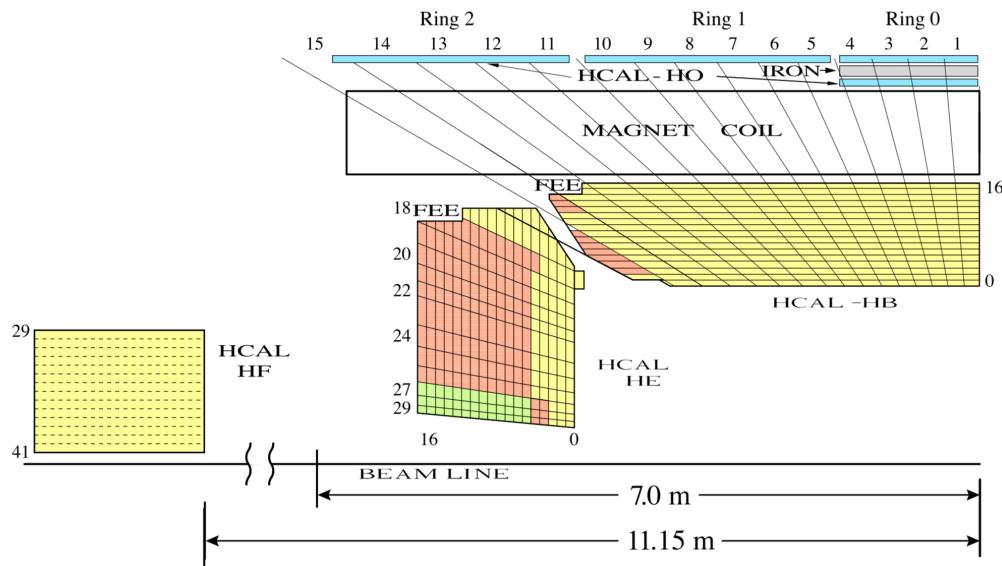


Figure 2.6: Geometrical layout of the HCAL showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [?].

2.2.4 Magnet

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

2.2.5 Muon Detector

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

The muon system is capable of three main functions: muon identification, muon p_T measurement and triggering. It is composed of three different types of detectors, all of 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 m² 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.

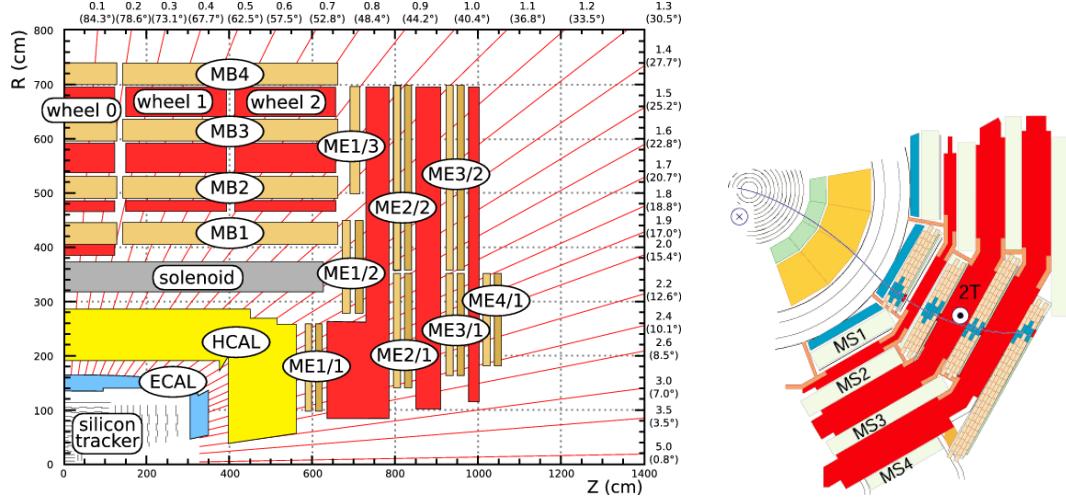


Figure 2.7: The left diagram shows a quarter-view of CMS with both the muon barrel (MB) and endcap (ME) stations [?]. The right diagram shows a muon in the transverse plane leaves a curved trajectory across the four muon detector stations [?].

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.

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 ϕ . Furthermore, CSCs have the advantage of featuring both precision muon measurement and muon trigger in a single device.

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

2.2.6 Trigger and Data Acquisition

Due to the vast volume of data originating from the proton-proton collisions (delivered by the LHC at a rate of 40 MHz), a method of eliminating the majority of the un-

interesting/unwanted events was a requisite for the CMS detector design. This event rate reduction is achieved by the implementation of the so-called trigger system, which manages to select the potentially interesting interactions and reduce the rate from a staggering 40 TBs^{-1} to a manageable value of just a few hundred Hz.

667

The CMS trigger system is implemented using a two-stage rate reduction, which combines both a hardware and software phase. The combination of both of these triggers is designed to reduce the rate by a factor of $\sim 10^6$. The first stage used in the rate reduction is purely hardware based and it is called the Level 1 (L1) Trigger [?], which consists of both Field Programmable Gate Array (FPGA) and Application Specific Integrated Circuit (ASIC) technology. During this initial stage, the rate is reduced to about 100 kHz with a latency of $3.2 \mu\text{s}$. This time interval constrains the trigger decision, allowing only for data from the calorimeters and muon system to be processed. Trigger primitives (TP) from these subsystems are processed through a series of steps before the combined event information is evaluated by the global trigger (GT) where the final decision, whether or not to accept the event, is made.

679

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

684

2.2.7 Event Reconstruction

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

700

Event information from each step in the simulation and reconstruction chain is logically grouped into what is called a data tier [?]. From the physicist's point of view the 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 re-

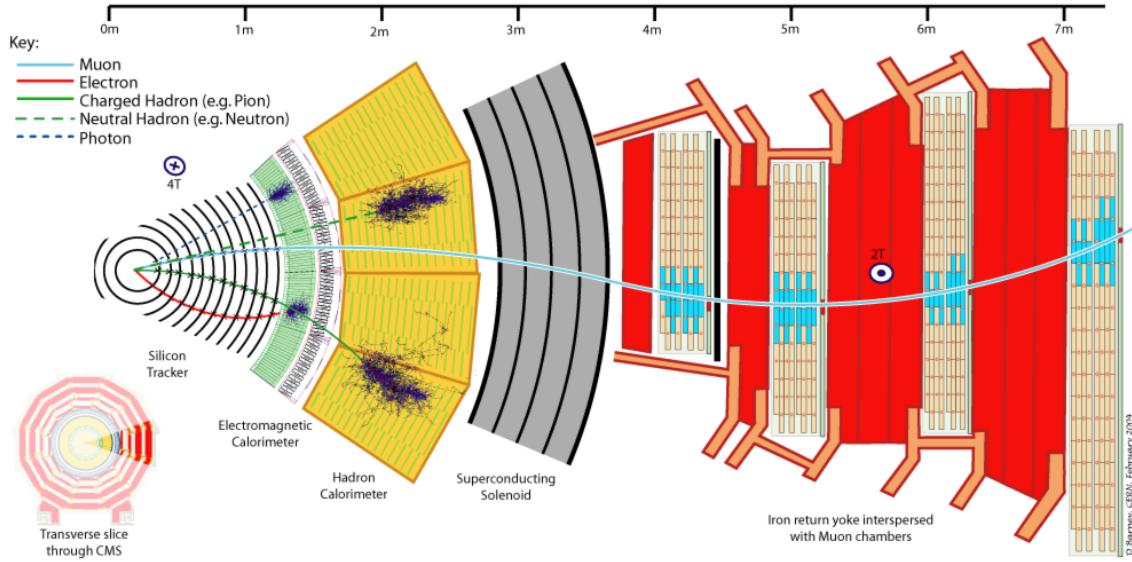


Figure 2.8: Transverse slice of the CMS detector, showing the individual detector subsystems and particle signatures in each. The particle type can be inferred by combining the detector response in the different subdetectors [?].

708 fitting with improved alignment or kinematic constraints, re-evaluation of energy and/or
 709 position of ECAL clusters based on analysis-specific corrections. The format of each data
 710 tier is ROOT [?]. ROOT is a framework for data processing developed at CERN with the
 711 sole purpose of aiding high energy physics research. The various AOD datasets are stored
 712 worldwide at various data tier centers. From the AOD's the analysis groups create data
 713 structures called NTuples containing only the high-level physics objects needed for their
 714 particular analysis.

715 2.2.8 Future Upgrade of Pixel Detector

716 After the first LHC shutdown called LS1 (2013-2014), and the installation of the
 717 Phase-1 Pixel Detector [?] in early 2017, among other things, the LHC is planning an-
 718 other series of upgrades during two major shutdowns, called LS2 and LS3, currently
 719 planned for 2019-2020 and \sim 2024, respectively. LS2 would result in a further increase
 720 of the luminosity beyond the original design value, to over $2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. With the
 721 LS3 upgrade of the LHC (called HL-LHC [?]) the luminosity is expected to reach up to
 722 $7.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Correspondingly, the CMS collaboration has planned a series of fur-
 723 ther upgrades [?, ?] that will ensure the capabilities of the CMS detector to match to the
 724 HL-LHC running conditions, while taking the opportunity to improve the performance
 725 and repair any problems uncovered during the data-taking periods. The UPRM group
 726 will continue its involvement in the Phase-1 Pixel Detector operations and Phase-2 Pixel
 727 Upgrade design.

728
 729 The HL-LHC conditions of instantaneous peak luminosities of up to $7.5 \cdot 10^{34} \text{ cm}^{-2}$
 730 s^{-1} and an integrated luminosity of the order of 300 fb^{-1} would result in 1 MeV neutron
 731 equivalent fluence of $2.3 \cdot 10^{16} \text{ neq/cm}^2$ and a total ionizing dose (TID) of 12MGy (1.2
 732 Grad) at the center of CMS, where its innermost component, the Phase-2 Pixel Detector
 733 will be installed. The detector should be able to withstand the above radiation dose, han-

734 dle projected hit rates of $3\text{GHz}/\text{cm}^2$ at the lowest radius, be able to separate and identify
 735 particles in extremely dense collision debris, deal with a pileup of 140-200 collisions per
 736 bunch crossing and have a high impact parameter resolution. This translates into requiring
 737 a detector design that is highly granular, has thinner sensors and smaller pixels, and faster,
 738 radiation hard electronics compared to its Phase-1 counterpart. The selection of interesting
 739 physics events at the Level-1 (L1) trigger and inefficiency of selection algorithms in
 740 high pileup conditions further require the Tracker to be included in this trigger stage, helping
 741 reduce the event rate from 40 MHz rate to 7.5 kHz. The physics goals also require
 742 an increase in Pixel Detector coverage to $|\eta| = 4.0$ which improves the p_T^{miss} resolution
 743 and particle-flow event reconstruction by providing p_T measurements and trajectories for
 744 charged particles entering the calorimeters. The p_T^{miss} resolution is an essential perfor-
 745 mance parameter for many BSM physics searches including SUSY and extra dimension
 746 models where particles escape undetected from the detector space. The smaller pixel size
 747 will further improve b-tagging as well as hadronic reconstruction and track reconstruction
 748 efficiencies within boosted jets, which can be produced from new heavy objects decaying
 749 to Higgs, Z bosons, or top quarks – all heavy probes that can be exploited for new physics
 750 searches. Improving the b-tagging capabilities directly affects our analysis due to the im-
 751 portance of properly identifying top quark decays.

752

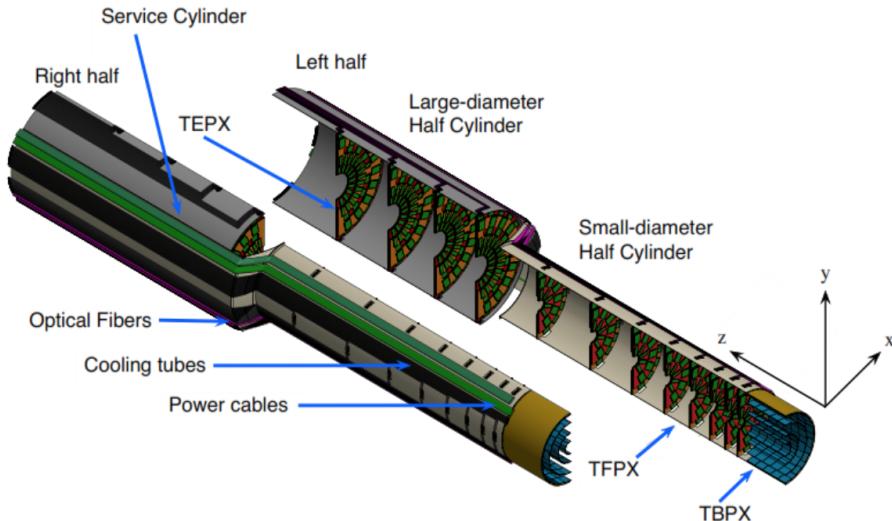


Figure 2.9: Phase-2 Pixel Detector Layout[ref].

753 The Phase-2 Pixel Detector [?,?] baseline design comprises a barrel part with 4 layers
 754 of Tracker Barrel Pixel Detector (TBPX), 8 small double-discs per side of Tracker For-
 755 ward Pixel Detector (TFPX) and 4 large double-discs per side of Tracker Endcap Pixel
 756 Detector (TEPX). This forward and end part is referred to as 8l4s (8 TFPX and 4 TEPX).
 757 In the TBPX the pixel modules are arranged in “ladders”. In each layer, neighboring ladders
 758 are mounted staggered in radius, so that $r\text{-}\phi$ overlap between the ladders is achieved.
 759 The modules on a ladder do not overlap in z . A projective gap at $\eta = 0$ is avoided by
 760 mounting an odd number of modules along z , and by splitting the barrel mechanics in z
 761 into slightly asymmetric halves. In TFPX and TEPX the modules are arranged in concen-
 762 tric rings. Each double-disc is physically made of two discs, which facilitates to mount
 763 modules onto four planes, with overlaps in r as well as $r\text{-}\phi$. Each disc is split into two

764 halves, and these D-shaped structures are referred to as “dees”. The TEPX will provide
765 the required luminosity measurement capability, by an appropriate implementation of the
766 readout architecture. In total, the pixel detector will have an active surface of approxi-
767 mately 4.9 m^2 . [Figure 2.9](#) shows the layout of the Phase-2 Pixel Detector.

768 Chapter 3

769 Theoretical Motivation

770 3.1 The Standard Model

771 Quantum Mechanics and the Theory of Special Relativity, along with the plethora of
772 new particles discovered over the past century or so, has led to the formulation of the
773 Standard Model of Particle Physics (SM) [?]. It is a theory concerning the electromagnetic,
774 weak [?], and strong nuclear interactions [?], as well as classifying all the subatomic
775 particles known. It provides our best current understanding of the universe, what matter
776 is made of and how it is held together. It rests on two basic ideas: all matter is made of
777 particles, and these particles interact with each other by exchanging other particles associated
778 with the fundamental forces. The basic components of matter are fermions and the
779 force carriers are bosons and fermions have half-integer values of spin whereas bosons
780 have integer values.

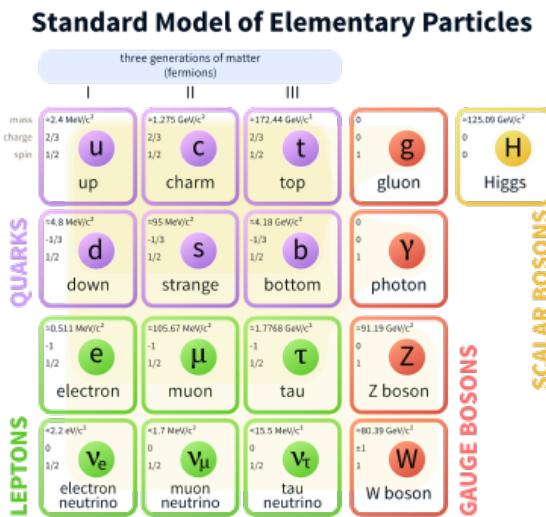


Figure 3.1: Schematic representation of the Standard Model particles. Shown are the three generations of matter (formed by fermions), the gauge bosons and the Higgs boson.

781 Fermions come in two families as shown in Figure 3.1 - the leptons and quarks. The
782 lepton family has six members, meanwhile the quark family contains six quarks. The
783 up and down quarks are found inside protons and neutrons. The twelve fermions are the
784 building blocks of matter and each one has a spin value of $1/2$. These particles interact

785 with each other through fundamental forces. Each force comes with one or more force
786 carriers. The nuclear force comes with the gluon and binds the quarks within the proton
787 and neutrons. The photon is associated with the electromagnetic force and is responsible
788 for mediating many of the forces humans encounter on a daily basis. The weak interaction
789 is responsible for radioactivity and is mediated by the Z and W bosons. The gluons,
790 photons and the Z and W bosons all have a spin of 1. The Standard Model is both remark-
791 ably simple and very powerful with nearly every measured quantity in particle physics
792 laboratories over the past five decades falling right on the predicted value (within exper-
793 imental error margins) [?, ?]. Until recently, the only missing piece of the SM was the
794 Higgs boson [?]. The Higgs boson is a particle corresponding to the Higgs field which,
795 according to the SM, gives mass to all the fundamental particles. It was finally discovered
796 at the LHC in 2012 [?, ?].

797 3.2 Indications of Physics Beyond the Standard Model

798 The Standard Model gives an incomplete picture of the makeup of the Universe as
799 listed below:

- 800 1. The Standard Model does not include gravity and does not explain why gravity is
801 so much weaker than the electromagnetic or nuclear forces [?]. The weak force
802 is 10^{24} times as strong as gravity. It is also incompatible with the most successful
803 theory of gravity to date, general relativity [?, ?].
- 804 2. There is a wide spectrum of masses among the building blocks of matter [?]. Why?
805 The electron is about 200 times lighter than the muon and 3,500 times lighter than
806 the τ -lepton, meanwhile the top quark is 75,000 times heavier than the up quark.
807 Why is there such a wide spectrum of masses among the building blocks of matter?
- 808 3. The SM introduces particle masses through a process known as spontaneous sym-
809 metry breaking [?, ?, ?] caused by the Higgs field. Within the standard model, the
810 mass of the Higgs gets some very large quantum corrections due to the presence of
811 virtual particles (mostly virtual top quarks). These corrections are much larger than
812 the actual mass of the Higgs. This means that the bare mass parameter of the Higgs
813 in the standard model must be fine-tuned in such a way that almost completely
814 cancels the quantum corrections. This level of fine-tuning is deemed unnatural.
- 815 4. It explains only about 5% of matter present in the universe. However, 26% exists
816 as dark matter [?] that behaves just like other matter (in terms of gravity) but which
817 only interacts weakly (if at all) with the Standard Model fields. Yet, the Standard
818 Model does not supply any fundamental particles that are good dark matter candi-
819 dates. The rest (69%) is dark energy [?], a constant energy density for the vacuum.
820 Attempts to explain dark energy in terms of the standard model have failed.
- 821 5. Neutrinos are predicted to be massless particles by the Standard Model [?]. How-
822 ever, neutrino oscillation experiments have shown that neutrinos do have mass [?, ?].
823 Mass terms for the neutrinos can be added to the standard model by hand, but these
824 lead to new theoretical problems as these mass terms need to be extraordinarily
825 small and it is doubtful if the neutrino masses would arise in the same way than the
826 masses of other fundamental particles do in the Standard Model.

- 827 6. The universe is made out of mostly matter but the standard model predicts that matter
 828 and antimatter should have been created in almost equal amounts if the initial
 829 conditions of the universe did not involve disproportionate matter relative to anti-
 830 matter [?]. But no mechanism, sufficient to explain matter–antimatter asymmetry,
 831 exists in the Standard Model.

832 Thus, as long as we observe various phenomena at low energy the Standard Model
 833 behaves properly, however, it lacks robustness at higher energy. To overcome its short-
 834 comings several theories exists beyond the Standard Model that include various exten-
 835 sions of the standard model through supersymmetry, such as the Minimal Supersym-
 836 metric Standard Model (MSSM) [?, ?] and Next-to-Minimal Supersymmetric Standard
 837 Model (NMSSM) [?], or entirely new ones like such as string theory [?] and extra di-
 838 mensions [?]. As these theories tend to reproduce the entirety of current phenomena, the
 839 question of which theory is the right one can only be determined by experiments. With the
 840 advent of even more powerful accelerators like the LHC, we are within reach of energy
 841 levels which existed only shortly after the Big Bang where the Standard Model has issues.

842 3.3 Supersymmetric extension of the Standard Model

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

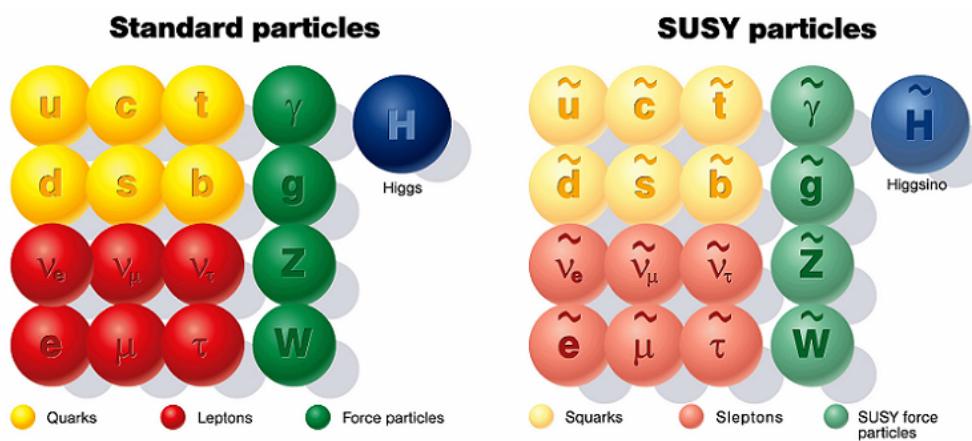


Figure 3.2: Schematic representation of the Standard Model particles with the added SUSY particles.

852 The gluon (QCD) has a fermionic partner, the gluino. Similarly, the spin 1 gauge
 853 bosons W^+ , W^0 , W^- and B^0 have their counterparts, called winos and binos, that, af-
 854 ter electroweak symmetry breaking, mix to give the mass eigenstate Z^0 and γ , and the

855 zino and photino respectively. For every lepton/quark there is a bosonic partner (i.e. a
856 scalar slepton/squark). There are three families for each of the quark and lepton super-
857 multiplets. The left-handed and right-handed pieces of the squarks and sleptons are two
858 separate components, as in the corresponding SM sector. The Higgs doublet is extended
859 by another doublet leading to other Higgs boson particles (Higgsinos).

860

861 Even though supersymmetry solves many problems in particle physics, it also poses
862 new problems. If these new particles had the same mass as their counterparts in the stan-
863 dard model, we would have already observed them. Since none of these particles have
864 been found yet, SUSY must be a broken symmetry. What makes superpartners heavier
865 than ordinary particles? Why are superpartners so well hidden in rare phenomena? This
866 arbitrary mass spectrum for the superpartners would have effects that are far too large
867 in rare processes that change the flavor of particles. There must be some special reason
868 why such effects are well hidden. How do we extract information on the mechanism of
869 supersymmetry breaking? How does supersymmetry impact cosmology? Is the lightest
870 supersymmetric partner what really composes Dark Matter? Different mechanisms have
871 been suggested on how this symmetry is broken and have led to different phenomenologi-
872 cal scenarios like the minimal supergravity model (mSUGRA) [?, ?], the gauge-mediated
873 supersymmetry breaking (GMSB) [?, ?, ?] model or R-parity [?, ?]. A discussion of
874 mSUGRA and GMSB is beyond the scope of this dissertation. R-parity is a new sym-
875 metry that has been added to the minimal supersymmetry scenario which prevents the
876 violation of the leptonic and baryonic numbers. All SM fields have even R-parity, while
877 SUSY particles have odd R-parity. The most obvious experimental constraint comes from
878 the non-observation of proton decay [?], which would violate both baryonic and leptonic
879 number conservation. Due to R-parity every interaction vertex in the theory involves an
880 even number of sparticles, meaning that they must be pair-produced. The decay chains
881 of the produced particles are characterized by the presence of one stable particle, generi-
882 cally called the Lightest Supersymmetric Particle (LSP), that may account for all the dark
883 matter in the Universe. However, in other models, like the R-parity violating SUSY this
884 may not apply. In other scenarios, those LSP can be regarded as vanishing mass and may
885 not be enough to satisfy the dark matter in the Universe.

886 3.3.1 Simplified Models Approach to Supersymmetry

887 Simplified models [?] are a new approach for characterizing LHC supersymmetry re-
888 sults. In a simplified model only a few new particles and a single decay topology are
889 introduced, and hence these models are easy to constrain. Traditionally, testing a particle
890 physics theory against experimental results requires calculating what the theory predicts
891 in an experimental situation using a particle accelerator detector. This translates into
892 computing: the masses of new particles, their decay widths, their branching ratios and
893 production cross-sections. This information is subsequently used in Monte Carlo simula-
894 tions of passage of proton-proton collisions products and their decays through the CMS
895 detector. This is done under the hypothesis that the theory is true. The simulated data
896 is analyzed just like the actual experimental data resulting from proton-proton collisions.
897 The theoretical and the experimental results can then be compared. Though conceptually
898 uncomplicated it requires access to the experimental data, and the step of simulating the
899 detector response requires intimate knowledge of the CMS detector by theorist commu-

nity. It is also restrictive being model/theory dependent. The CMS collaboration, like ATLAS, does not share experimental details but only the final results via publications. The Simplified Model approach circumvents this issue by presenting results on very simple phenomenological models instead of a specific model, such as the constrained MSSM (cMSSM) [?]. Each simplified models encompasses a specific phenomenological feature that is common to many different theories. A limited set of hypothetical particles and decay chains are introduced to produce a given topological signature. The amplitudes describing the production and decays of these particles are parametrized in terms of the 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 3.1](#). We have used T2tt, T1ttt, T1ttbb, T5ttt 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
			Opposite-Sign Dileptons Multileptons
T3w	$\tilde{g}\tilde{g}$	$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{\text{LSP}}$ $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \rightarrow W^\pm\tilde{\chi}_{\text{LSP}}$	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	$\tilde{g}\tilde{g}$	$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_{\text{LSP}}$ $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_2^0, \tilde{\chi}_2^0 \rightarrow \ell^+\ell^-\tilde{\chi}_{\text{LSP}}$	Opposite-Sign Dileptons
T1bbbb	$\tilde{g}\tilde{g}$	$\tilde{g} \rightarrow b\bar{b}\tilde{\chi}_{\text{LSP}}$	All-Hadronic (b)
T1ttt	$\tilde{g}\tilde{g}$	$\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_{\text{LSP}}$	All-Hadronic (b) Single Lepton + Jets (b) Same-Sign Dileptons (b) 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 3.1: Some Simplified Models with their corresponding production and decay modes.

916 **Chapter 4**

917 **Optimizing the Phase-2 Pixel Geometry
918 for the Level-1 Trigger in the Outer
919 Tracker**

920 **4.1 Upgraded Phase-2 Tracker**

921 This chapter is dedicated to a study that seeks to optimize the geometry of the Phase-2
922 Pixel Detector (for the HL-LHC) in terms of how the upgraded Inner Tracker (IT) affects
923 the overall stub production (the concept of a stub is defined in [subsection 4.1.2](#)). The
924 goal of this study was to determine how the removal of discs from the Phase-2 Forward
925 Pixel detector would affect the production of L1 stubs in the upgraded CMS Outer Tracker
926 (OT). More information of the Phase-2 Tracker can be found at [\[?\]](#). This study has a direct
927 impact on the future of SUSY analyses due to the importance of the tracker in searches
928 for BSM physics.

929

930 In the HL-LHC era (2023-2035) an integrated luminosity of $3,000 \text{ fb}^{-1}$ is expected to
931 be reached at a center-of-mass energy of 14 TeV. This data will enable us to make precise
932 measurements of several SM parameters (particularly the scalar sector). This will also
933 lay the groundwork for new physics searches proposed by many of the BSM models. In
934 addition, the HL-LHC project is essential to identify any hints of new physics, including
935 SUSY, that could possibly be observed during the LHC Run 2 or 3, or beyond. The
936 bigger, and highly granular pixel and silicon tracker, and its extended η coverage, will
937 improve jet sub-structure, boosted object identification and heavy object tagging to hunt
938 for BSM physics, specifically Supersymmetry. A discovery could happen with various
939 few sigma excesses and then interpretation of the new signal will require a large amount
940 of data. At the HL-LHC, much gain is expected for weakly produced SUSY particles
941 because they are characterized by small production cross-sections. Considering the case
942 that gluinos and squarks are heavy, search for electroweak production channels will offer
943 the best option to discover SUSY. At the HL-LHC, charginos/neutralinos are expected to
944 be discovered (significance of 5σ) up to masses of about 950 GeV. In the case of strong
945 production, the mass reach for gluinos goes up to 2.2 TeV. This study is a contribution in
946 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.

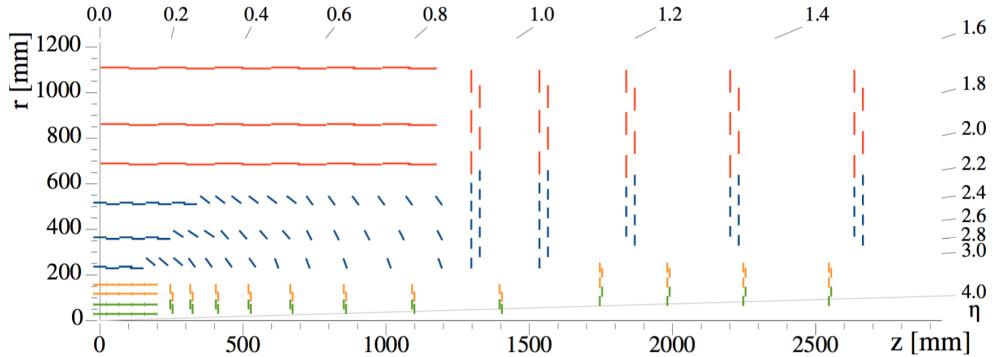


Figure 4.1: Sketch of one quarter of the tracker layout in r-z view. In the Inner Tracker the green lines correspond to pixel modules made of two readout chips and the yellow lines to pixel modules with four readout chips. In the Outer Tracker the blue and red lines represent the two types of modules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

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

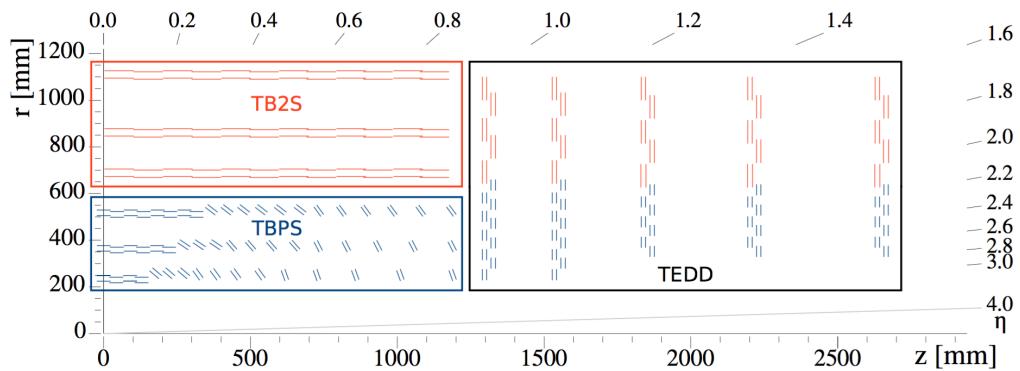


Figure 4.2: Sketch of one quarter of the Outer Tracker in r-z view. Blue (red) lines represent PS (2S) modules. The three sub-detectors, named TBPS, TB2S, and TEDD, are indicated. All overlapping layers are shown separately.

4.1.2 Tracker Input to the L1-Trigger

The enhancement of the trigger performance involves both a higher output rate of interesting events and an improved discriminating power of the event selection, which is

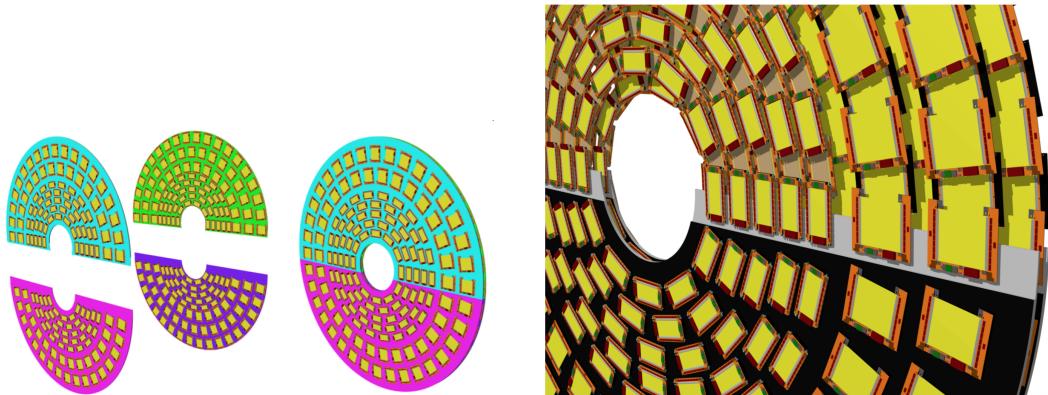


Figure 4.3: Sketch of four dees (left) forming a double-disc (centre), and drawing of a part of a TEDD double-disc (right), illustrating the overlap of modules in ϕ and z . The upper two dee support structures are removed in order to show all layers of modules. The modules ob the disc are arranged in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

more challenging in a high-pileup environment. Improved discriminating power will be achieved by using more information in the trigger decision, with a longer latency available for its processing. The use of tracking information in the L1 trigger will improve the p_T resolution of various objects at L1 (e.g. jets), it will allow the exploitation of information on track isolation, and will contribute to the mitigation of pileup of 140-200 collisions per bunch crossing. CMS plans to enhance the first level trigger rate from presently 100 kHz to 750 kHz and to increase the latency from the present value of $3.2 \mu s$ to $12.5 \mu s$. The front-end electronics and the L1 trigger track reconstruction need to comply with these new requirements. The necessity of providing tracking information to the L1 trigger is a main driver for the design of the OT, including its module concept. The use of tracking information in the L1 trigger implies that the tracker has to send out self-selected information at every bunch crossing. Such functionality relies upon local data reduction in the front-end electronics, in order to limit the volume of data that has to be sent out at 40 MHz. This is achieved with modules that are capable of rejecting signals from particles below a certain p_T threshold, referred to as “ p_T modules” [?]. Tracks from charged particles are bent in the transverse plane by the 3.8 T field of the CMS magnet, with the bending angle depending on the p_T of the particle. The modules are composed of two single-sided closely-spaced sensors read out by a common set of front-end ASICs that correlate the signals in the two sensors and select the hit pairs (referred to as “stubs”) compatible with particles above the chosen p_T threshold (as shown in Figure 4.4). Stubs are defined as matching pairs of hits on the adjacent silicon layers of the p_T modules. Whether or not a 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 TEPEX. This forward and end part is referred to as 8l4s (8 TFPX and 4 TEPEX). 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/TEPEX) would affect the production of L1 stubs

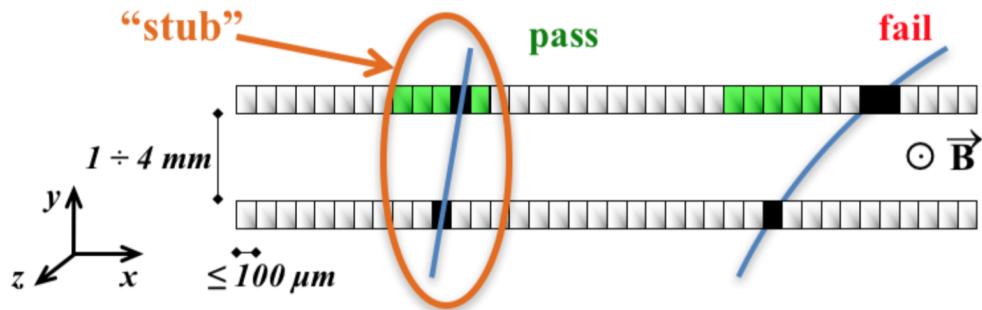


Figure 4.4: Correlation of signals in closely-spaced sensors enables rejection of low- p_T particles; the channels shown in green represent the selection window to define an accepted stub.

in the upgraded CMS OT. A decrease in the total number of stubs is expected when less material is present due to secondary interactions.

993

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

- 996 • 8 small discs and 4 large discs ($8s4l$) – designated as the default geometry
- 997 • 8 small discs and 3 large discs ($8s3l$) - one large disc removed.
- 998 • 7 small discs and 4 large discs ($7s4l$) - one small disc removed.
- 999 • 6 small discs and 3 large discs ($6s3l$) - one large disc and two small discs removed.
- 1000 • A version of the default pixel geometry with a dead disc (used to verify our find-
1001 ings).

1002 The response of the detector due to design modifications is simulated using the offi-
1003 cial CMS software package CMSSW. The detector geometry is generated with a design
1004 tool for innovative silicon tracking detectors called tkLayout [?] and is exported into the
1005 standard CMSSW format along with the materials like sensing silicon elements, elec-
1006 tronics and inactive material. The exported tracker geometry serves as input for the full
1007 simulation and reconstruction performance studies with CMSSW. The design of the CMS
1008 Phase-2 Tracker includes Tracker information for L1 triggering. The study was made
1009 from $t\bar{t}$ samples with a center-of-mass energy of 14 TeV and 200 pileup. The samples
1010 were created with the CMS software version referred to as CMSSW_9_3_2. The stub in-
1011 formation is used as input for the track fitting algorithm and allows tracks used for L1
1012 trigger to be reconstructed within approximately $5 \mu\text{s}$ (this includes a $1 \mu\text{s}$ estimate for
1013 the time required to transmit the stub data from the detector to the counting room) to
1014 remain within the overall L1 trigger latency budget of $12.5 \mu\text{s}$. Stubs are formed from
1015 two closely spaced sensors in Tracker measurement layers and provide linked position
1016 measurements for each particle passing through them. [Table 4.1](#) shows the stub window
1017 p_T threshold values used for the study.

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

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

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](#).

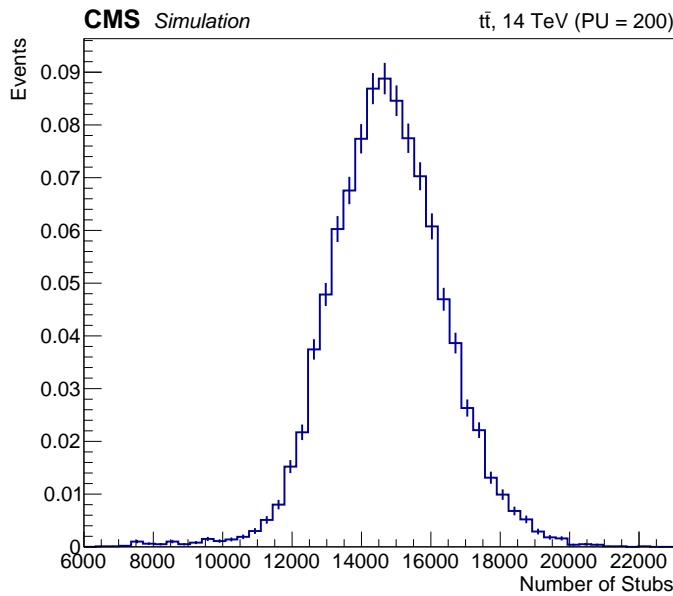


Figure 4.5: Number of stubs per event for standard pixel geometry (8s4l).

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

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

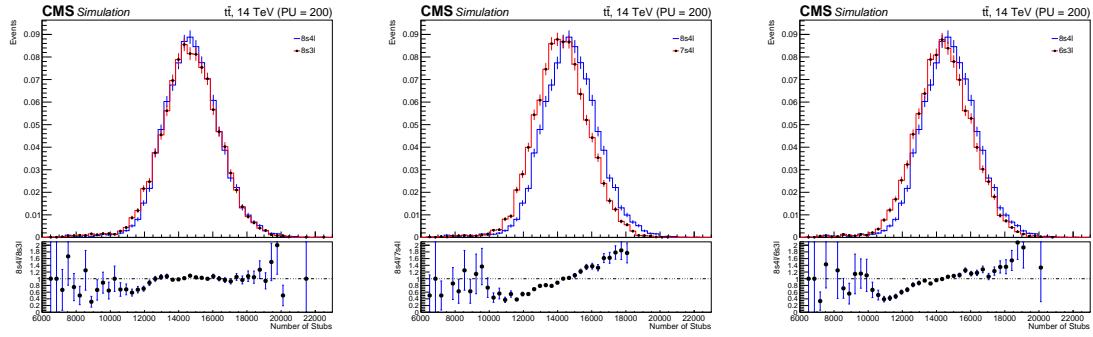


Figure 4.6: Comparison of total number of stubs per event between the standard layout and the other three considered. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right).

1022 (b) **Comparison of Stubs between default Pixel Geometry (8s4l) to disc removed:**

1023 A comparison of the total number of stubs per event between the standard geometry to
 1024 the others that have a disc or more removed (8s3l, 7s4l, 6s3l) was performed ([Figure 4.6](#)).
 1025 Out of the three comparisons shown, the largest shift in the stub distribution can be seen
 1026 from the 7s4l geometry. It was found that the removal of a smaller pixel disc has a more
 1027 pronounced effect on stub production in the OT than a larger pixel disc.
 1028

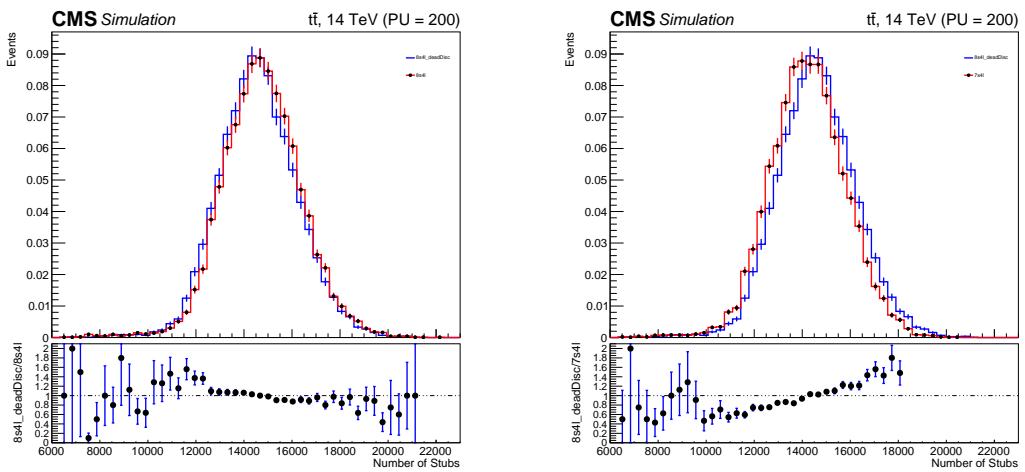


Figure 4.7: Comparison for number of stubs per event between pixel geometry 8s4l with a dead disc (blue) and both pixel geometry 8s4l (right plot) and 7s4l (left plot) (blue). Both the standard geometry and 7s4l are shown in blue in their respective plots.

1029 (c) **Comparison with a disc dead in default geometry:**

1030 An additional study, using pixel geometry 8s4l with a dead disc, was made to verify
 1031 the effect observed on the L1 stub counting. [Figure 4.7](#) (right) shows a comparison be-
 1032 tween 2 geometries that are (in principle) identical but one detects no hits on the second
 1033 disc in the positive side. This Figure shows a slight deviation from the original 8s4l ge-
 1034 ometry.
 1035

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 (η) as shown in Figure 4.8.

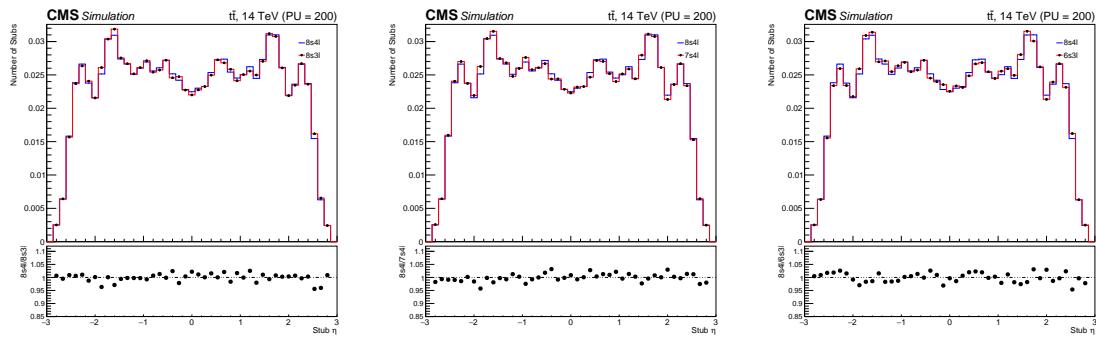


Figure 4.8: Comparison of total number of stubs per η value between the standard layout and the other three layouts considered. All four different layouts exhibit the same distribution. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right).

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).

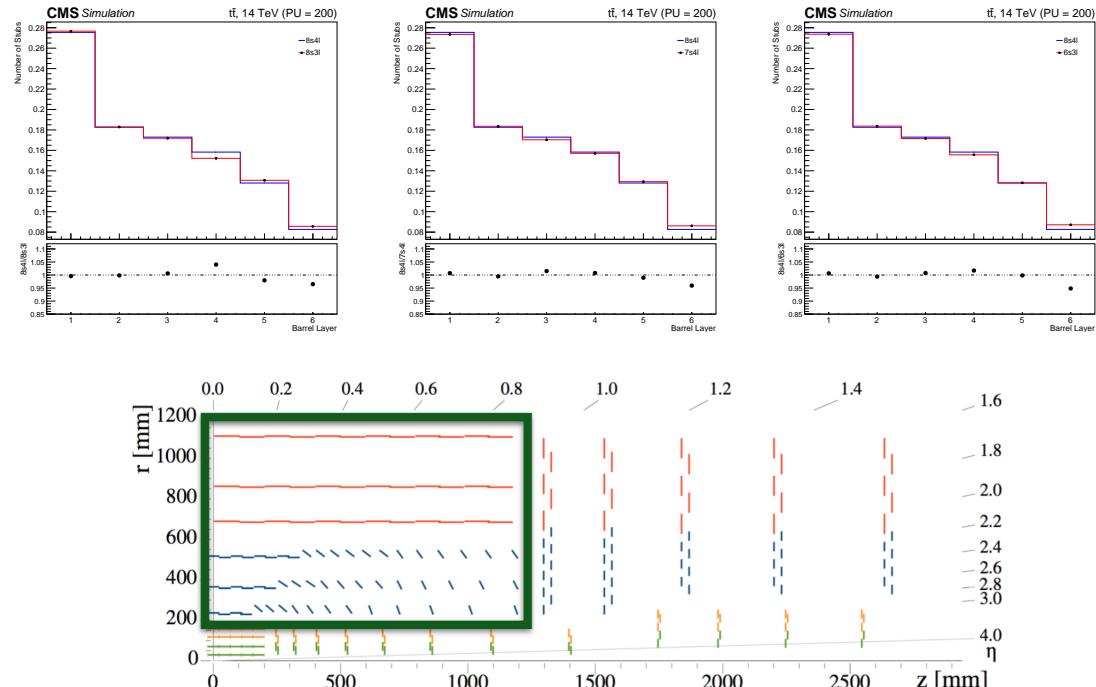


Figure 4.9: Comparison of total number of stubs per each of the 6 barrel layers. $8s4l$ vs $8s3l$ (left), $8s4l$ vs $7s4l$ (middle) and $8s4l$ vs $6s3l$ (right). Tracker Barrel layers shown below.

4.2. Tracker Stub Simulation

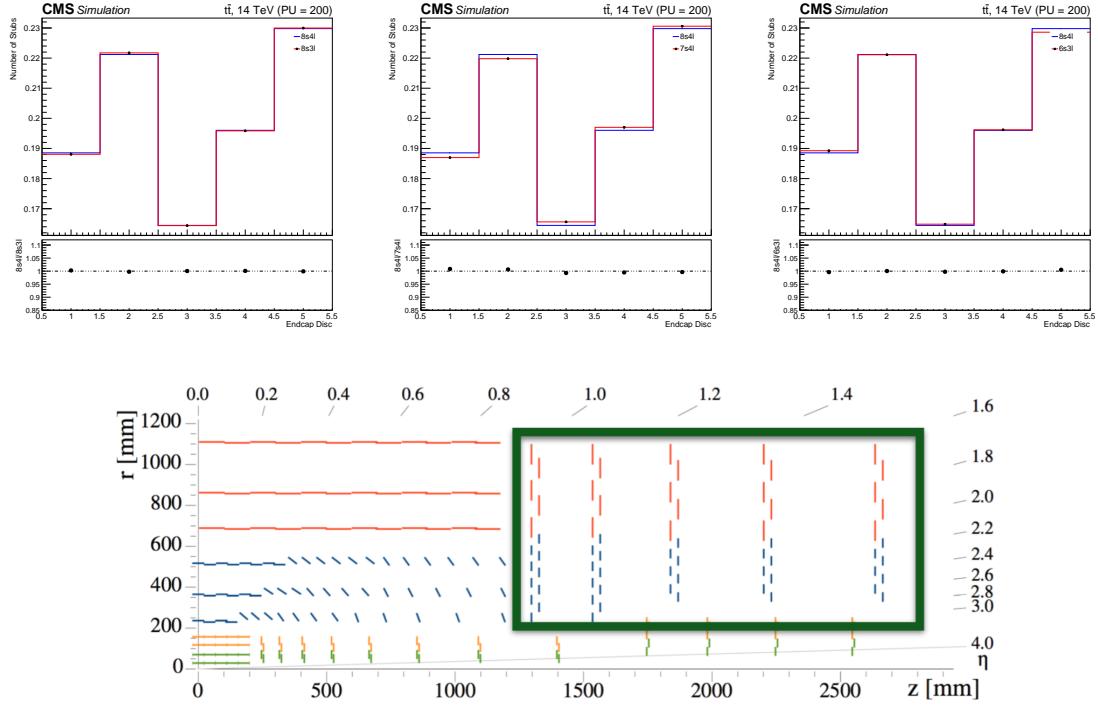


Figure 4.10: Comparison of total number of stubs per each of the 5 Endcap Double Discs. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right). Tracker Endcap Discs shown below.

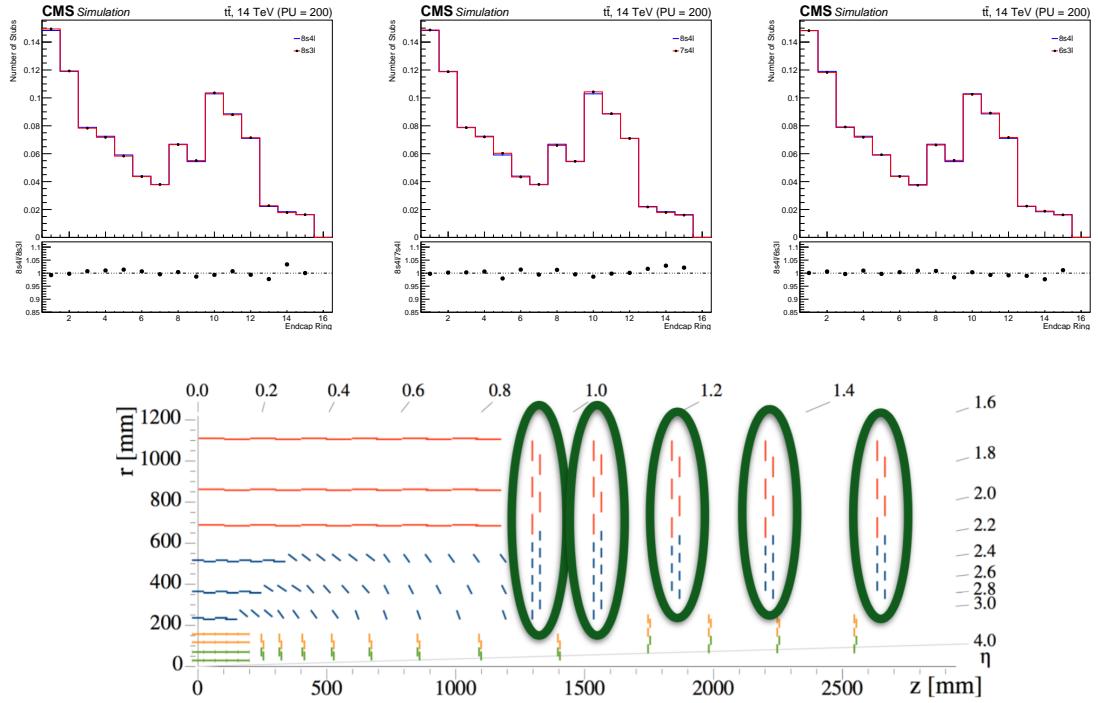


Figure 4.11: Comparison of total number of stubs for all rings across the five endcap double-discs. 8s4l vs 8s3l (left), 8s4l vs 7s4l (middle) and 8s4l vs 6s3l (right). The double discs are shown below.

1045 **4.3 Results**

1046 The above study shows that the removal of discs in the pixel detector affects the num-
1047 ber of stubs formed in the OT. A reduction in the amount of material (i.e. removal of
1048 discs) translates into a reduced amount of secondary particle interactions and hence, a de-
1049 crease in the net amount of stubs is observed. In other words more material (more discs)
1050 in the pixel layer leads to more secondaries, increasing the total amount of stubs in a given
1051 sample. However, the effect is small and a detailed study with respect to various η regions
1052 is planned in the future, as stub formation is also sensitive to η windows. Knowing the
1053 origin of particles causing stubs is also an important area to look into.

1054 **Chapter 5**

1055 **Search for SUSY in the 0-Lepton Final
1056 State**

1057 **5.1 Introduction**

1058 An overview of a search for Supersymmetry in the all-hadronic channel with missing
1059 transverse momentum and a customized top tagger is presented. The data used for the
1060 analysis was collected in 2016 by the CMS experiment at the LHC, and corresponds to
1061 an integrated luminosity of 35.9 fb^{-1} from proton-proton collisions at a center-of-mass
1062 energy (\sqrt{s}) of 13 TeV. The results and procedures presented are the product of an analysis
1063 conducted in collaboration with the hadronic SUSY analysis team based at Fermilab in
1064 Batavia, IL, USA. The published results can be obtained from [?].

1065 **5.2 Analysis Description**

1066 The study consists of an inclusive search for events with final states that contain miss-
1067 ing transverse momentum (p_T^{miss}) and reconstructed top quarks. The signal models used
1068 in this study include the production of three different types of SUSY particles. Two of
1069 which are the top squark (\tilde{t}) and the gluino (\tilde{g}), the supersymmetric partners of the SM top
1070 quark (t) and gluon (g). The third one is the neutralino ($\tilde{\chi}_1^0$), considered to be the lightest
1071 SUSY particle (LSP) under the Minimal Supersymmetric Standard Model (MSSM), and
1072 a possible candidate for Dark Matter.

1073 The signal data is selected by requiring events to have a minimum amount of jets (N_j)
1074 and b-jets (N_b) as well as a large p_T^{miss} . Other factors, such as the number of top jets
1075 (N_t), the scalar sum of the jet transverse momentum (H_T), and the so-called stransverse
1076 mass (m_{T2})¹ are also considered in order to select events of interest. The search region is
1077 then defined in exclusive bins of N_t , N_b , H_T , p_T^{miss} and m_{T2} . Additional discriminatory
1078 variables used in the analysis include the azimuthal angle between p_T^{miss} and the leading
1079 jets of an event ($\Delta\phi$), the pseudorapidity (η) and the cone radius size (defined as $\Delta R =$

¹ 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 $m_{T2} \equiv \min_{\vec{q}_T^{(1)} + \vec{q}_T^{(2)} = \vec{p}_T} \left[\max \left\{ m_T^2(\vec{p}_T^{(1)}, \vec{q}_T^{(1)}; m_{\tilde{\chi}_1^0}), (\vec{p}_T^{(2)}, \vec{q}_T^{(2)}; m_{\tilde{\chi}_1^0}) \right\} \right]$.

¹⁰⁸¹ $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ used in the isolation of various physics objects.

¹⁰⁸² 5.2.1 Signal Models

¹⁰⁸³ The signal models used in this analysis are based on two different scenarios: direct
¹⁰⁸⁴ top squark production and gluino mediated top squark production. In the former sce-
¹⁰⁸⁵ nario we consider one decay within the Simplified Model Spectra (SMS) [?] framework
¹⁰⁸⁶ called ‘T2tt’ (Figure 5.1, top left), where a top squark pair is produced directly from the
¹⁰⁸⁷ proton-proton collision and then decays into a pair of top quarks and a pair of neutralinos
¹⁰⁸⁸ ($\tilde{t} \rightarrow t\tilde{\chi}_1^0$).
¹⁰⁸⁹

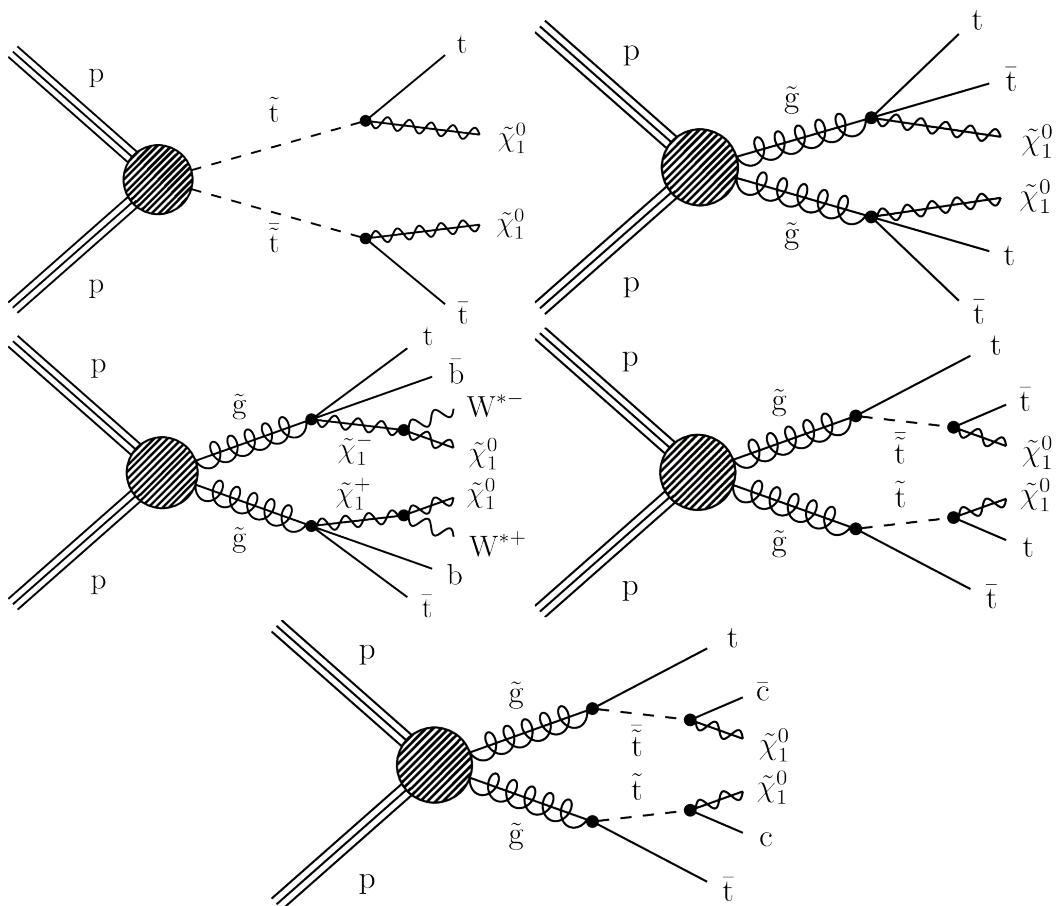


Figure 5.1: Simplified model diagrams show of the various Supersymmetry signals considered in this analysis: the T2tt model (top left), the T1tttt model (top right), the T1ttbb model (middle left), the T5tttt (middle right), and the T5ttcc model (bottom).

¹⁰⁹⁰ In the case of gluino mediated top squark production there are several different de-
¹⁰⁹¹ cay modes considered. The T1tttt model (Figure 5.1, top right) consists of two gluinos
¹⁰⁹² decaying into a pair of top quarks and a neutralino ($\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$), accounting for situations
¹⁰⁹³ in which the top squark is too heavy to be produced directly but the gluino is not. In the
¹⁰⁹⁴ T1ttbb model (Figure 5.1, middle left), a pair of gluinos to either a top or bottom squark
¹⁰⁹⁵ 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
¹⁰⁹⁶ positive ($\tilde{\chi}_1^+$) or negative ($\tilde{\chi}_1^-$) chargino. Due to the small difference in mass between the
¹⁰⁹⁷ $\tilde{\chi}_1^+$ (or its conjugate) and the LSP ($\Delta m(\tilde{\chi}_1^+, \tilde{\chi}_1^0) = 5$ GeV), the two particles are taken

1098 to be nearly mass degenerate. The T1ttbb model provides sensitivity to situations where
 1099 there are mixed states of top and bottom squarks.

1100

1101 The T5tttt model consists of a pair of gluinos, where each decays into a top quark and
 1102 an on-shell top squark. The top squark then decays into a top quark and a neutralino. For
 1103 this model, a mass difference of $\Delta m(\tilde{t}, \tilde{\chi}_1^0) = 175$ GeV between the top squark and the
 1104 LSP is expected. This model provides sensitivity to a region of T2tt where the $t\bar{t}$ back-
 1105 ground is very similar to the signal and $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$ is very close to the top squark mass.
 1106 In the T5ttcc model (Figure 5.1, bottom) each gluino decays into a final state consisting
 1107 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
 1108 to the T5tttt model with the difference that $\Delta m(\tilde{t}, \tilde{\chi}_1^0) = 20$ GeV and the on-shell top
 1109 squark decays into a charm quark and the LSP. The T5ttcc model provides sensitivity to
 1110 situations where the on-shell top squark is kinematically forbidden from decaying into a
 1111 top quark.

1112

1113 All five signal models described share similar final states containing two neutralinos
 1114 and up to four top quarks. Since the neutralino is considered stable and only interacts
 1115 weakly with matter, it cannot be picked up by the detector. Therefore, missing transverse
 1116 momentum (p_T^{miss}) is one of the most important variables used to compare event yields
 1117 between the signal and the SM background.

1118 5.2.2 Top Tagger

1119 The distinguishing feature of this analysis is its powerful top-tagging algorithm. It
 1120 is designed to provide high reconstruction efficiency over the full range of the top quark
 1121 p_T for the considered SUSY signal models. This top-tagger combines the use of sev-
 1122 eral different jet clustering algorithms in order to identify three different categories of
 1123 top quark jets, designated as “monojet”, “dijet” and “trijet”. The tagger makes use of
 1124 identified jets that were reconstructed with the anti- k_T [?] clustering algorithms combined
 1125 with additional corrections and selection criteria provided by the Cambridge–Aachen [?]
 1126 and soft-drop de-clustering [?] algorithms. In addition, multivariate analysis (MVA) tech-
 1127 niques, such as the random forest decision tree algorithm [?], were applied in order to
 1128 further decrease the amount of reconstructed fake tops.

1129

1130 There are three top-quark jet categories that take into account that top quark decay
 1131 products get closer together as the top quark p_T becomes higher. For specific p_T values,
 1132 the decay products of a hadronic process will be reconstructed either as one (“monojet”)
 1133 or two (“dijet”) jets rather than three (“trijet”). The p_T values for which this is observed
 1134 depends on the size of the jets that are being used. In the case of a highly boosted jet,
 1135 the anti- k_T algorithm is used within a cone size $\Delta R \sim 0.8$ (called AK8), which captures
 1136 all decay products of the top quark and is reconstructed as a single jet. This technique
 1137 requires that the top quark starts with a p_T of at least 400 GeV to have the decay products
 1138 fully captured in the 0.8 jet cone. For a $p_T < 400$ GeV, resolved top-tagging techniques
 1139 are more efficient, which require three jets independently clustered within a cone size of
 1140 $\Delta R \sim 0.4$ (AK4). Both types of algorithms are used in order to obtain a high efficiency
 1141 over a wide range of top quark p_T .

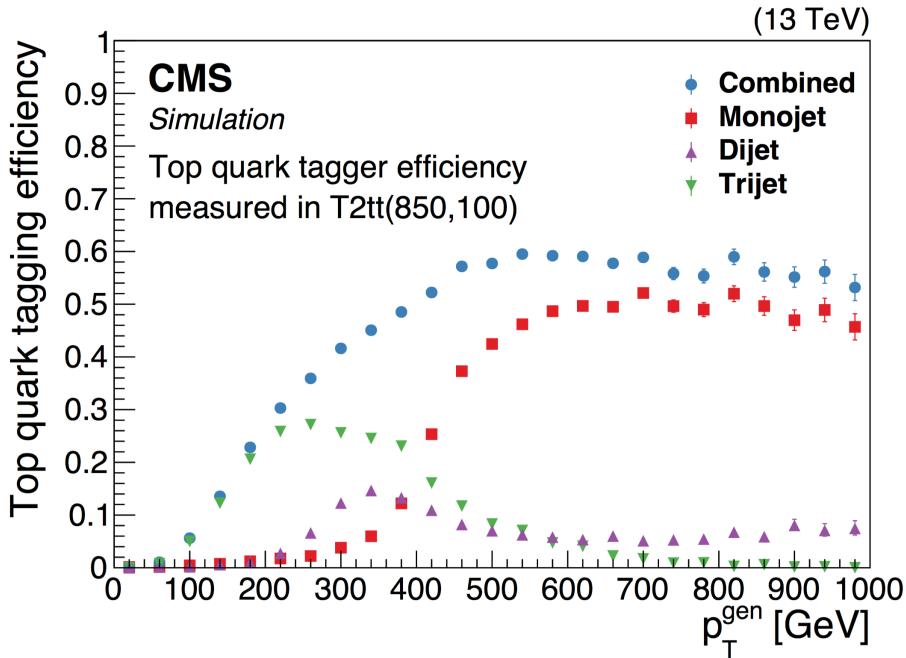


Figure 5.2: Efficiency of the top quark tagger as a function of generator-level top quark p_T for the monojet (red boxes), dijet (magenta triangles), and trijet (green upside-down triangles) categories and for their combination (blue circles), as determined using T2tt signal events with a top squark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

In the case of the dijet top quark decay category, an AK8 jet with a $p_T > 200$ GeV is 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 $Z(\nu\bar{\nu})+\text{jets}$ events with a selection criteria similar to the one applied to data: $N_j \geq 4$, $N_b \geq 1$, $p_T^{\text{miss}} > 250$ 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 ≥ 200 GeV. This ΔR 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 isolation² 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.

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 τ leptons that decay hadronically or misidentified electrons and muons. The track requirements for this veto are $p_T > 5 \text{ GeV}$, $|\eta| < 2.5$ and a relative track isolation less than 0.2. In addition, the isolated track veto is only applied when the transverse mass of an isolated track- \bar{p}_T^{miss} is $m_T < 100 \text{ GeV}$, consistent with W boson decays. This veto successfully reduces the background of leptonically decaying W bosons by about 40%.

5.2.4 Baseline Event Selection

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

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

5.2.5 Search Bin Definition

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

²The ratio between the isolation sum and the candidate p_T .

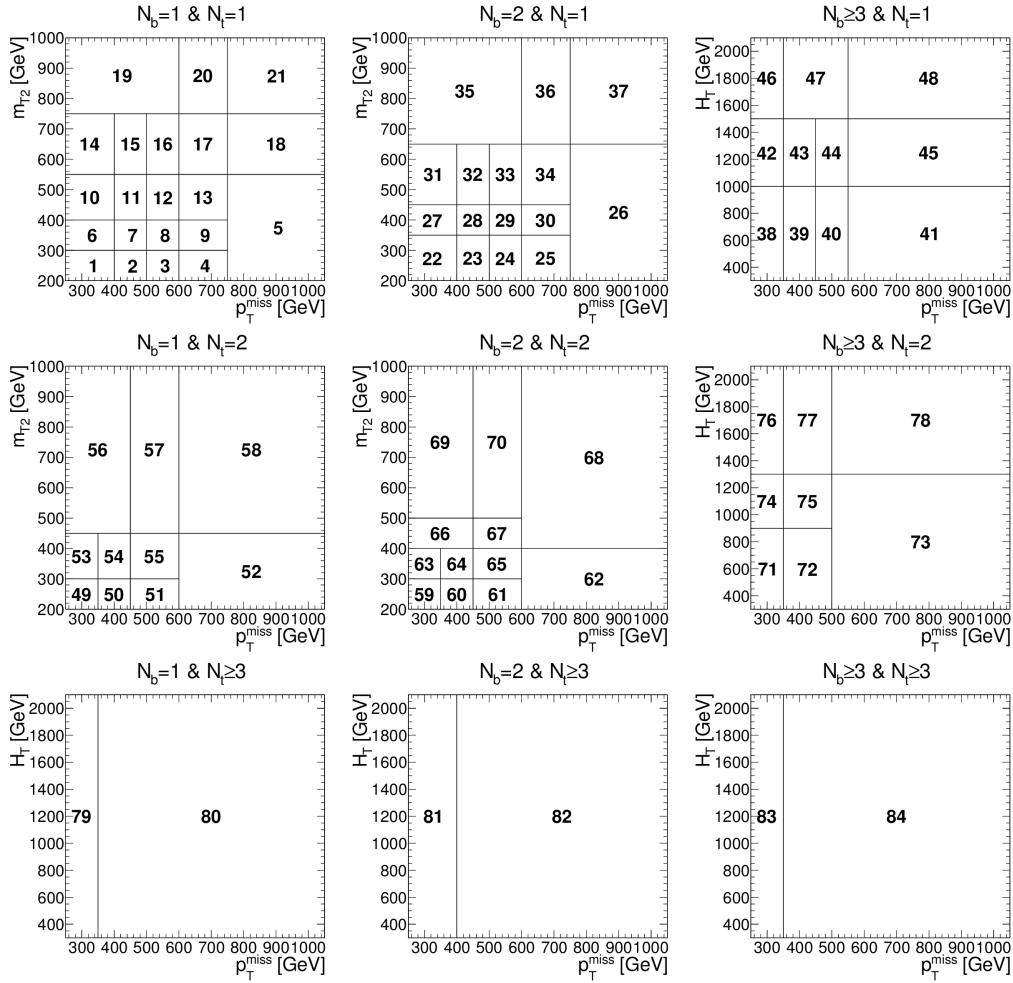


Figure 5.3: Search region definitions in the kinematic variables.

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

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

The vast majority of the expected SM background (around 70%) is due to $t\bar{t}$ and 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

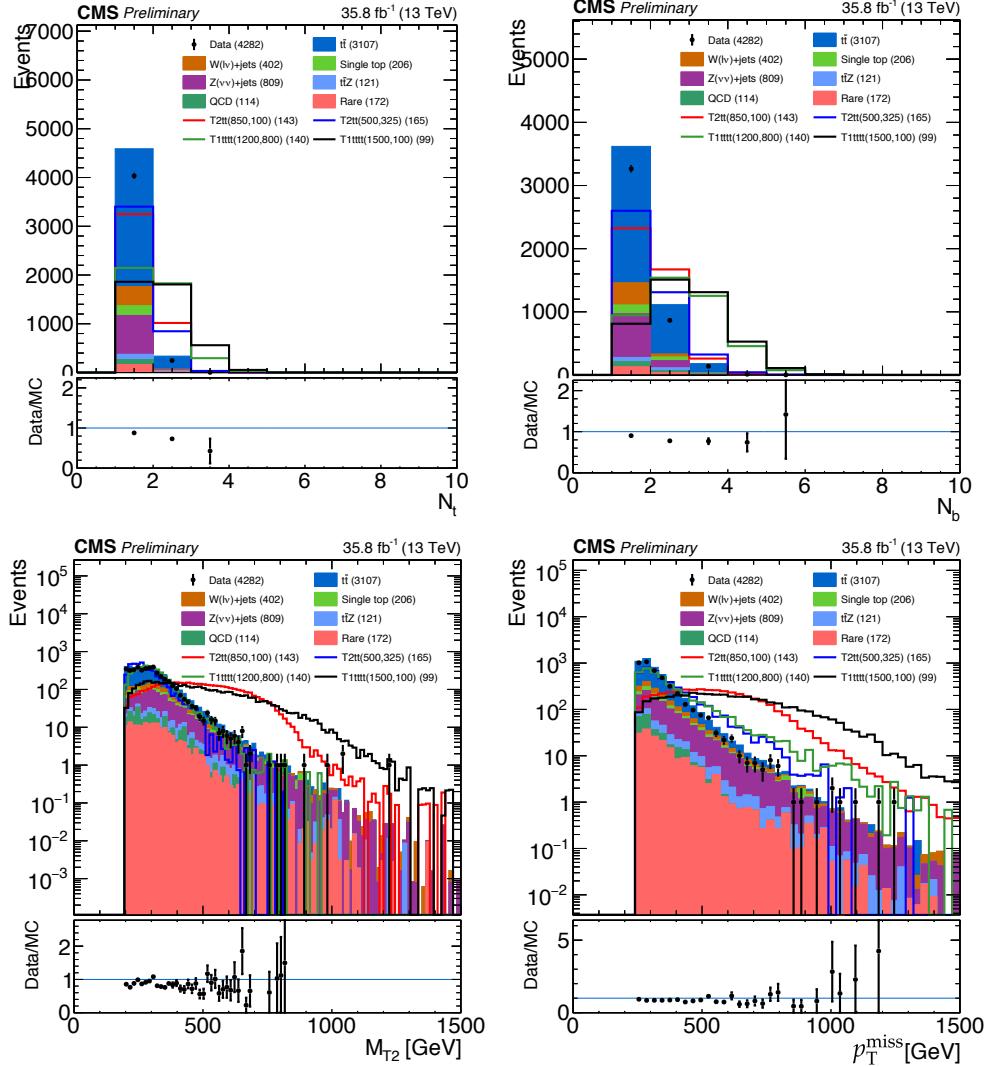


Figure 5.4: Comparison between the total SM background from simulation and CMS data for N_t (top left), N_b (top right), m_{T2} (bottom left) and p_T^{miss} (bottom right). Total SM backgrounds and signals are scaled to same data yield for a shape comparison.

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 τ which decays hadronically (τ_h), the τ 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.

1222

1223 The total predicted amount of lost-lepton and τ_h events in any given search region is
 1224 defined by the net sum of single-electron and single-muon events in the respective CS,
 1225 corrected by a translation factor that is determined from simulation. Due to differences
 1226 in how they are detected, the single-electron and single-muon samples are determined
 1227 separately. The translation factor used to correct the search bins is defined as the ratio
 1228 of simulated τ_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.

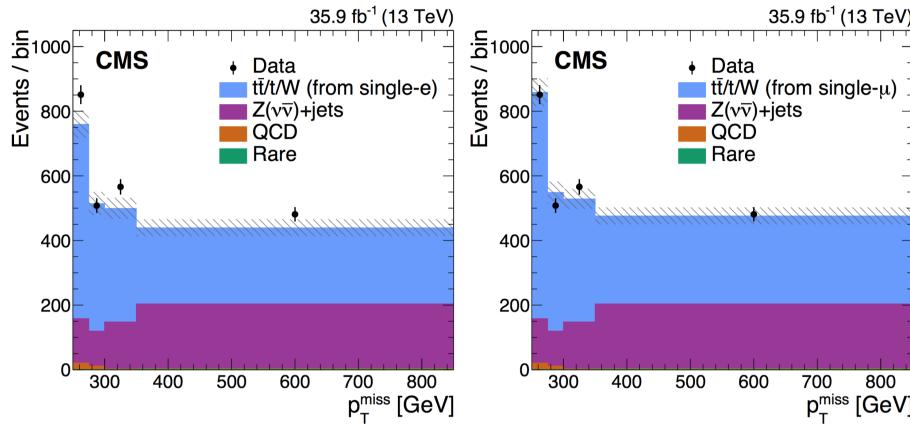


Figure 5.5: Distribution of p_T^{miss} for both the single-electron (left) and single-muon (right) SB data samples compared to predictions for SM processes. The hatched bands indicate the statistical uncertainties in the total SM prediction. Note that the data and the predictions for all backgrounds 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 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 requirements are applied in order to suppress contributions from $Z(\nu\bar{\nu})+jets$ and QCD multijet processes. The SB is divided into four intervals of p_T^{miss} , where the contribution of lost-lepton and τ_h to the corresponding intervals are determined by multiplying the appropriate single-electron and single-muon CS (Figure 5.5) by a translation factor from simulation. The contribution to the SB from other backgrounds, such as $Z+jets$, QCD multijet and rare events, is estimated directly from simulation. The total SM background prediction is found to agree with the data within uncertainty, confirming the validity of the translation factor procedure.

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

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

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

The first scale factor, R_{norm} , is computed by comparing the expected event yield in 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.

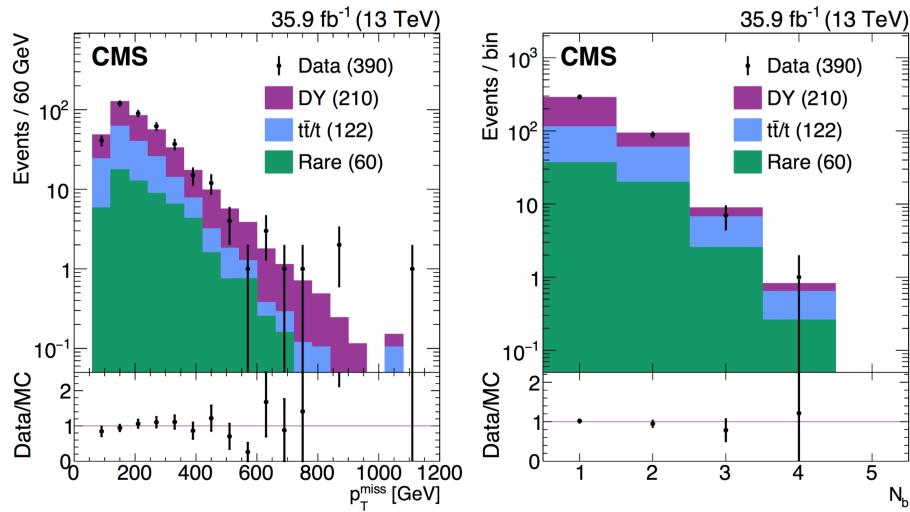


Figure 5.6: The p_T^{miss} (left) and N_b (right) distributions of data and simulation in the loose dimuon CS after applying both the normalization and shape correction factors. The lower panels show the ratio between data and simulation. Error bars in the plot account only for statistical uncertainties.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

The systematic uncertainties for the $Z(\nu\bar{\nu})+j$ ets 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).

5.3.3 The QCD Multijet Background

To estimate the QCD Multijet background a signal-depleted, QCD multijet-rich data control sample is used. This control sample is defined by inverting the preselection requirements on $\Delta\phi(p_T^{miss}, j_{1,2,3})$ and subtracting contributions of other SM backgrounds, such as $t\bar{t}$, $W+j$ ets, and $Z+j$ ets. For $t\bar{t}$ and $W+j$ ets the same methods (lost-lepton and hadronic τ) are used to estimate the contributions for this QCD multijet-enriched control region. In the case of the $Z \rightarrow \nu\bar{\nu}$ background, simulation is used for its estimation due to its small number. Following that, a translation factor, partly determined by data and partly

by simulation, is used to convert the number of QCD multijet events measured in the data control region into a QCD multijet prediction for each search region bin. This translation factor, called T_{QCD} , is computed as the simulated ratio between the signal region and the inverted- $\Delta\phi$ control region, in bins of p_T^{miss} and m_{T2} where the bins are bounded in the same way as the signal bins.

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

5.3.4 Background From Rare and Other Processes

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

5.4 Results and Interpretation

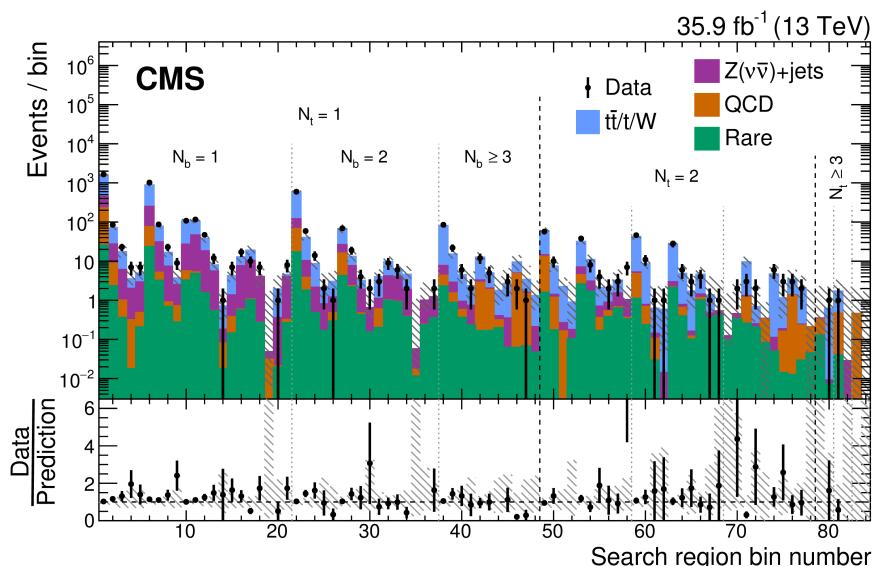


Figure 5.7: Observed events compared to the corrected SM background predictions for all 84 search regions in the full data of 35.9 fb^{-1} collected in 2016. The lower panel shows the ratio between the data and the total SM simulation. The gray bands show the total uncertainty related to the background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

Figure 5.7 shows the summary of all observed events compared to the corrected SM background simulation for all of the 84 search bins. The data obtained shows no statisti-

1306 cally significant deviation from the predicted SM background. The biggest contribution
 1307 to the background is attributed to the $t\bar{t}$ and $W+jets$ processes, followed by $Z(\nu\bar{\nu})+jets$,
 1308 which could be dominant in regions that have a high p_T threshold. The contributions arising
 1309 from the QCD multijet and rare backgrounds are found to be nearly negligible in all
 1310 of the search regions.

1311

1312 Exclusion limits are calculated for each of the signal models discussed in this chapter
 1313 by applying a binned likelihood fit on the data. The likelihood function is obtained for
 1314 each of the 84 search regions as well as for each of the background data control samples
 1315 (single-electron, single-muon, and QCD) from the product of the Poisson probability den-
 1316 sity function. Exclusion limits were placed on the top squark, gluino and LSP production
 1317 cross-sections with a 95% confidence level (CL), calculated using a modified frequentist
 1318 approach with the CL_s criterion [?, ?] and asymptotic results for the test statistic [?].

1319

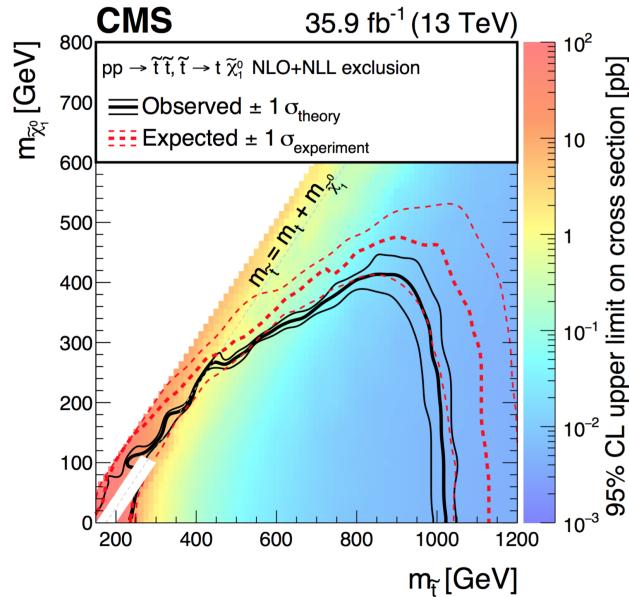


Figure 5.8: The 95% CL upper limit on the production cross section of the T2tt simplified model as a function of the top squark and LSP masses. The solid black curves represent the observed exclusion contour with respect to NLO+NLL [?] signal cross sections and the change in this contour due to variation 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.

1320

1321 The uncertainties from the search region modeling are taken into account for each of
 1322 the 84 bins and arise from several different sources: the statistical uncertainty in the sim-
 1323 ulated event samples, the integrated luminosity (2.5% [?]), the lepton and isolated-track
 1324 veto efficiencies (up to 6.8%), the b tagging efficiency (up to 21%), the trigger efficiency
 1325 (up to 2.6%), the renormalization and factorization scales (up to 3.5%), the ISR modeling
 1326 (up to 46%), the jet energy scale corrections (up to 34%), the top quark reconstruction
 1327 efficiency (up to 14%), and the modeling of the fast simulation compared with the full
 1328 simulation for top quark reconstruction and mis-tagging (up to 24%). All of the uncer-
 1329 tainties, with the exception of the statistical precision of the simulation, are taken to be
 fully correlated between the search regions. Contributions from signal contamination in

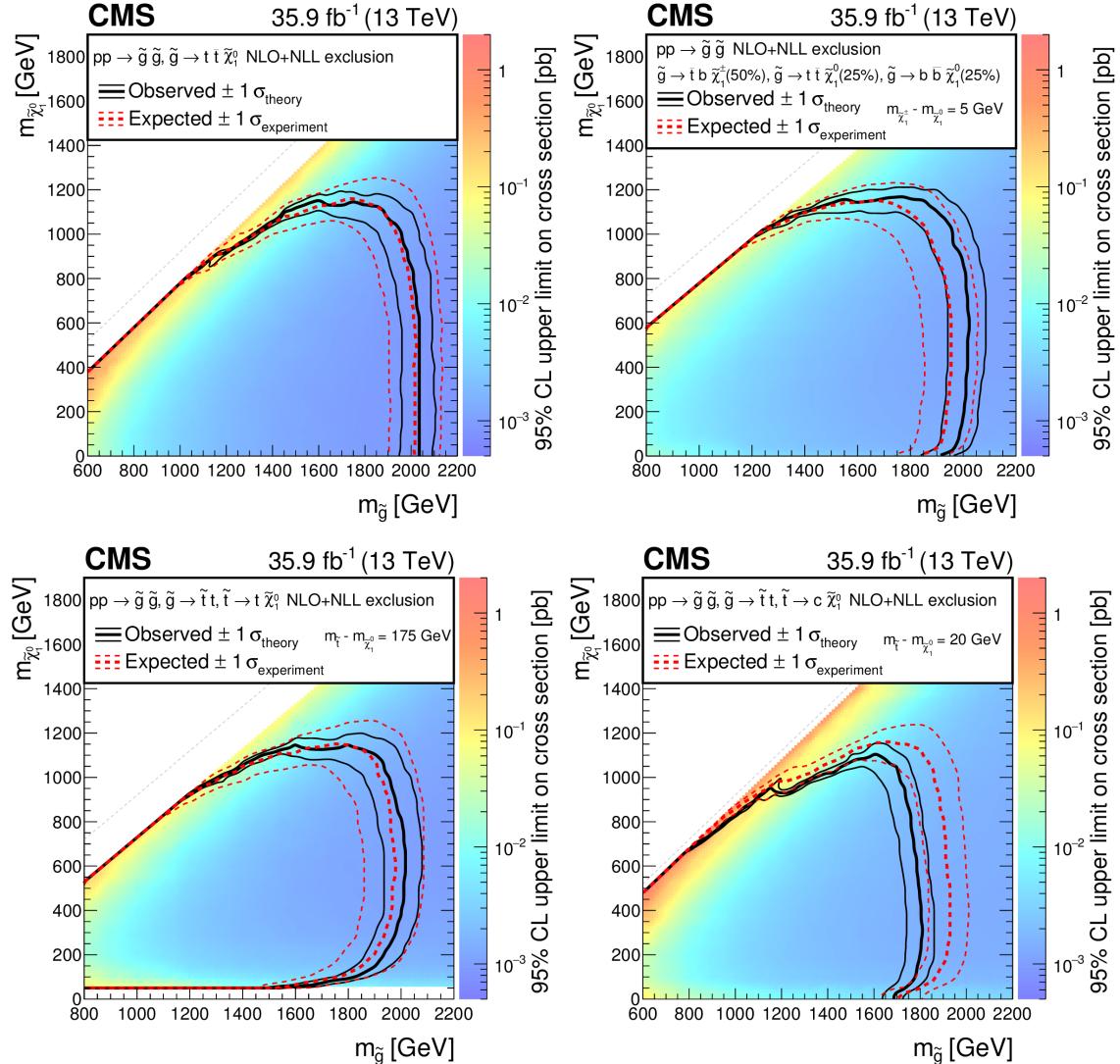


Figure 5.9: The 95% CL upper limit on the production cross section of the T1tttt (upper left), T1ttbb (upper right), T5tttt (bottom left), and T5ttcc (bottom right) simplified models as a function of the gluino and LSP masses.

the signal modeling are found to be only significant from the single-lepton control samples and negligible for the rest.

1332

1333 The 95% CL exclusion limits obtained for the T2tt model (Figure 5.8), which consists
 1334 of direct top squark production, excludes top squark masses up to 1020 GeV and LSP
 1335 masses up to 430 GeV. Meanwhile, Figure 5.9 shows the results for the gluino pair pro-
 1336 duction models: T1tttt, T1ttbb, T5tttt, and T5ttcc. For the T1tttt model, gluino masses of
 1337 up to 2040 GeV and LSP masses up to 1150 GeV are excluded, with corresponding limits
 1338 of 2020 and 1150 GeV for the T1ttbb model, 2020 and 1150 GeV for the T5tttt model,
 1339 and 1810 and 1100 GeV for the T5ttcc model.

1340 **Chapter 6**

1341 **Estimation of the $Z \rightarrow \nu\bar{\nu}$ Background**

1342 **6.1 Introduction**

1343 A detailed explanation for the estimation of the $Z \rightarrow \nu\bar{\nu}$ +jets background is presented
1344 in this chapter. The following estimation procedure builds upon the 2016 SUSY analysis
1345 summarized in [chapter 5](#) and aims to refine the overall background calculation as well
1346 as reduce the uncertainties associated with the previous method. To accomplish this, an
1347 additional γ +jets CS is used in conjunction with the tight $Z \rightarrow \mu^+\mu^-$ control region used in
1348 the 2016 analysis estimation. The new γ +jets CS provides a more data-driven estimation
1349 procedure with the added benefit of a substantially larger production cross-section than
1350 Z +jets processes [?].

1351 **6.2 The Irreducible $Z \rightarrow \nu\bar{\nu}$ Background**

1352 An important source of background in searches for SUSY in the 0-lepton final state
1353 comes from events in which a Z boson, accompanied by jets, decays into a pair of neu-
1354 trinos ($Z \rightarrow \nu\bar{\nu}$ +jets). The resulting final state is comprised of a large p_T^{miss} (from the
1355 neutrino pair) and multiple hadron jets, closely mimicking the SUSY signal. For similar
1356 searches, the $Z \rightarrow \nu\bar{\nu}$ contribution can make up a large portion of the background in many
1357 of the search bin regions (higher than 50% in some regions) [?, ?], and can compose up
1358 to about a third of the total SM background. For this particular analysis, the $Z \rightarrow \nu\bar{\nu}$ ac-
1359 counts for about 17% of the total SM background and owes its low value to the dedicated
1360 top-tagging algorithm outlined in [subsection 5.2.2](#).

1361

1362 There are several different methods that have been developed to estimate the $Z \rightarrow$
1363 $\nu\bar{\nu}$ +jets background [?]. Two of the commonly used methods involve the use of a control
1364 region dominated by $Z \rightarrow ll$ +jets, where the l stands for lepton (either a muon or an
1365 electron, in this case) or γ +jets events. The $Z \rightarrow ll$ channel has the advantage of having
1366 very similar kinematics to the $Z \rightarrow \nu\bar{\nu}$ region but suffers from low statistics (due in part
1367 to its small branching ratio), specially in the tight search regions used in typical SUSY
1368 searches. On the other hand, the γ +jets region has a much higher production cross-section
1369 but involves a completely different process. The hybrid method described in this chapter
1370 makes use of both control regions in order to estimate the $Z \rightarrow \nu\bar{\nu}$ background corrections,
1371 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 $Z + \text{jets}$ and $\gamma + \text{jets}$ processes. The ‘ V ’ in the figure can represent either Z or γ .

1372 6.3 The Loose $\gamma + \text{jets}$ Control Region

1373 The loose $\gamma + \text{jets}$ control region, as well as the photon ID/isolation selection, is de-
 1374 scribed in this section. The $\gamma + \text{jets}$ control region is chosen with the purpose of substitut-
 1375 ing the loose muon control region used in the 2016 version of this analysis for the calcu-
 1376 lation of the shape correction scale factors applied to the final estimation of the $Z \rightarrow \nu\bar{\nu}$
 1377 background. The $\gamma + \text{jets}$ control region is presumed to be better suited for this estima-
 1378 tion due to having a much higher cross-section and therefore, an expected reduction in
 1379 the statistical and systematic uncertainties associated to the shape correction factors are
 1380 expected.

1381 6.3.1 Photon ID and Isolation

1382 Three different working points are provided by the CMS EGM physics object group
 1383 (POG) for simple cut-based photon identification [?]. The three working points, called
 1384 loose, medium and tight, are chosen according to the requirements of the particular analy-
 1385 sis and differ on the amount of background rejection they offer, as well as on their average
 1386 photon selection efficiency. The higher the efficiency of a given ID, the lower the amount
 1387 of background that is rejected, and vice-versa. [Table 6.1](#) shows the cut values that are
 1388 applied to photons that are found within both the ECAL barrel and endcap range. The
 1389 associated values to the efficiency and the background rejection rate are shown for each
 1390 of the three different photon ID selections.

1391
 1392 In order to obtain the high efficiency and background rejection rates shown, a robust
 1393 set of identification and isolation criteria are selected. A total of five parameters are
 1394 used for this simple cut based method. For photon identification, the H/E and the $\sigma I\eta I\eta$
 1395 variables are found to provide the best results. The H/E parameter is defined as the ratio
 1396 of the HCAL tower energy over the ECAL seed cluster energy. A threshold value is
 1397 selected on H/E to remove background from electrons that are detected in both the ECAL
 1398 and HCAL but have no reconstructed track [?]. The $\sigma I\eta I\eta$ variable is known as the
 1399 photon shower shape variable and is defined in the ECAL as the energy weighted standard
 1400 deviation of a single crystal within the 5×5 crystal η range, centered around the crystal
 1401 with maximum energy [?]. This variable is a key component in the identification of both
 1402 electrons and photons since it provides a measure of the shower width where most of the
 1403 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%)
Background Rejection	Loose (85.73%)	Medium (88.87%)	Tight (90.66%)
H/E	0.105	0.035	0.020
$\sigma I\eta I\eta$	0.0103	0.0103	0.0103
ρ -corrected PF charged hadron isolation	2.839	1.416	1.158
ρ -corrected PF neutral hadron isolation	$9.188 + 0.0126 \cdot p_T^\gamma + 0.000026 \cdot p_T^{\gamma^2}$	$2.491 + 0.0126 \cdot p_T^\gamma + 0.000026 \cdot p_T^{\gamma^2}$	$1.267 + 0.0126 \cdot p_T^\gamma + 0.000026 \cdot p_T^{\gamma^2}$
ρ -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%)
Background Rejection	Loose (76.90%)	Medium (81.50%)	Tight (84.34%)
H/E	0.029	0.027	0.025
$\sigma I\eta I\eta$	0.0276	0.0271	0.0271
ρ -corrected PF charged hadron isolation	2.150	1.012	0.575
ρ -corrected PF neutral hadron isolation	$10.471 + 0.0119 \cdot p_T^\gamma + 0.000025 \cdot p_T^{\gamma^2}$	$9.131 + 0.0119 \cdot p_T^\gamma + 0.000025 \cdot p_T^{\gamma^2}$	$8.916 + 0.0119 \cdot p_T^\gamma + 0.000025 \cdot p_T^{\gamma^2}$
ρ -corrected PF photon isolation	$4.895 + 0.0040 \cdot p_T^\gamma$	$4.095 + 0.0040 \cdot p_T^\gamma$	$3.272 + 0.0040 \cdot p_T^\gamma$

1404 The other three parameters considered, comprise the isolation portion of the photon
 1405 selection cuts. These are the ρ -corrected particle flow (PF) charged hadron, neutral hadron
 1406 and photon isolation parameters. As can be seen from Table 6.1, two of these parameters
 1407 (the neutral hadron and photon isolation) have a dependence on p_T^γ . These cuts are used
 1408 to ensure that the identified photon is well isolated within its own cone and by rejecting
 1409 photons that are identified within close proximity to either a charged or a neutral hadron
 1410 [?]. The value ρ included in the name of each of these parameters refers to the total
 1411 pileup density [?]. Therefore, the term “ ρ -corrected” implies that these values, which are
 1412 sensitive to the residual contamination that arises from pile-up, have been corrected to
 1413 include these contributions.

1414 6.3.2 Photon Selection

1415 The event selection process for the γ +jets control region starts with photon candidates
 1416 that have a $p_T > 200$ GeV and are within the acceptance range of the CMS ECAL (given
 1417 by $|\eta| < 1.4442$ for the barrel and $1.566 < |\eta| < 2.5$ for the endcaps). The photons
 1418 are subjected to pass the loose ID/isolation cuts described in subsection 6.3.1 in order to
 1419 remove $\sim 85\%$ of the background processes and obtain a prompt photon sample that is
 1420 $\sim 90\%$ pure, on average. Additional restrictions include some of the same requirements
 1421 imposed on the signal baseline selection discussed in subsection 5.2.4. These include
 1422 $N_j \geq 4$, an $H_T > 300$ GeV, the $\Delta\phi$ requirements for leading jets and the lepton vetoes
 1423 described in subsection 5.2.3. The lepton veto, in particular, greatly improves the prompt
 1424 photon selection by removing many of the events in the simulated samples where a lepton
 1425 gets misidentified as a photon.

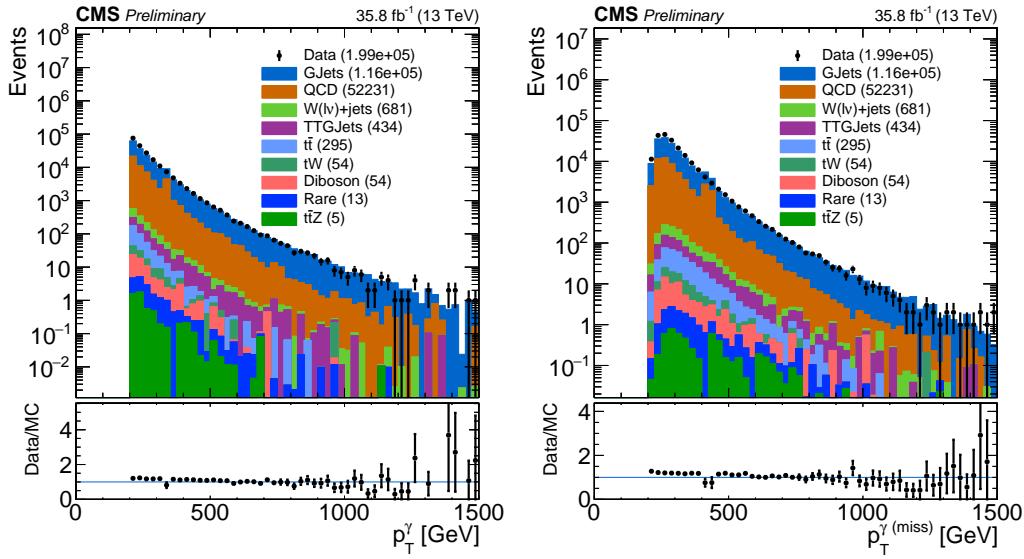


Figure 6.2: Shown are both the p_T^γ (left) and $p_T^{\gamma(\text{miss})}$ (right) distributions before applying any corrections. $p_T^{\gamma(\text{miss})}$ is obtained by adding the p_T^γ to the total p_T^{miss} in every event.

1426 To further emulate the $Z \rightarrow \nu\bar{\nu} + \text{jets}$ background, a variable in which the photons are
 1427 treated as p_T^{miss} is defined. We call this variable $p_T^{\gamma(\text{miss})}$ and we obtain it by adding the
 1428 p_T^γ for every event to the total p_T^{miss} in the event. Both the p_T^γ and the resulting $p_T^{\gamma(\text{miss})}$
 1429 distributions are shown in [Figure 6.2](#) as data/MC comparison plots, where the simulated
 1430 backgrounds are stacked in order of ascending contribution.

1431
 1432 The main contributions from simulation arise from the $\gamma + \text{jets}$, QCD and to a lesser
 1433 extent, $t\bar{t}\gamma$. Other non-dominant backgrounds in the control region include contributions
 1434 from $W(l\nu) + \text{jets}$, $t\bar{t}$, Diboson, tW , $t\bar{t}Z$ and rare processes. Most of these lesser back-
 1435 grounds are nearly negligible (several orders of magnitude lower than the dominant back-
 1436 grounds) and are considered to be mostly composed of fake photons. In addition to the
 1437 cuts described, all of the simulation samples are subjected to weights that apply correc-
 1438 tions to pileup as well as the b-tagging efficiency. Data, on the other hand, is obtained
 1439 from a sample that contains events with at least one identified photon. Photons in this
 1440 sample are also subjected to the high-level trigger HLT_Photon175, which restricts the
 1441 selection to photons that have a $p_T > 175$ GeV. Both simulation and data are subjected to
 1442 the same selection criteria established in this section.

1443 6.3.3 Photon Purity and Fake Rate

1444 Three different types of photons make up the $\gamma + \text{jets}$ CS: prompt photons, produced
 1445 either directly or through fragmentation, and fake photons. Prompt photons are defined
 1446 as photons which are formed shortly after the proton-proton collision (i.e. before the pro-
 1447 duced quarks and gluons have had enough time to form hadrons). Two types of photons
 1448 fit in this category. The first type, which we designate as direct photons, are photons that
 1449 are produced directly from the proton-proton interaction [?]. A secondary type of prompt
 1450 photon, that is virtually indistinguishable from the direct photons at the detector level,
 1451 originates from the decay of π^0 mesons and are called fragmentation photons. The final
 1452 type of photon found in the CS corresponds to fake (or non-prompt) photons. The fake

1453 photon contribution typically arises from leptons (mostly electrons) whose tracks are not
 1454 properly reconstructed, yet leave energy measurements in the ECAL.

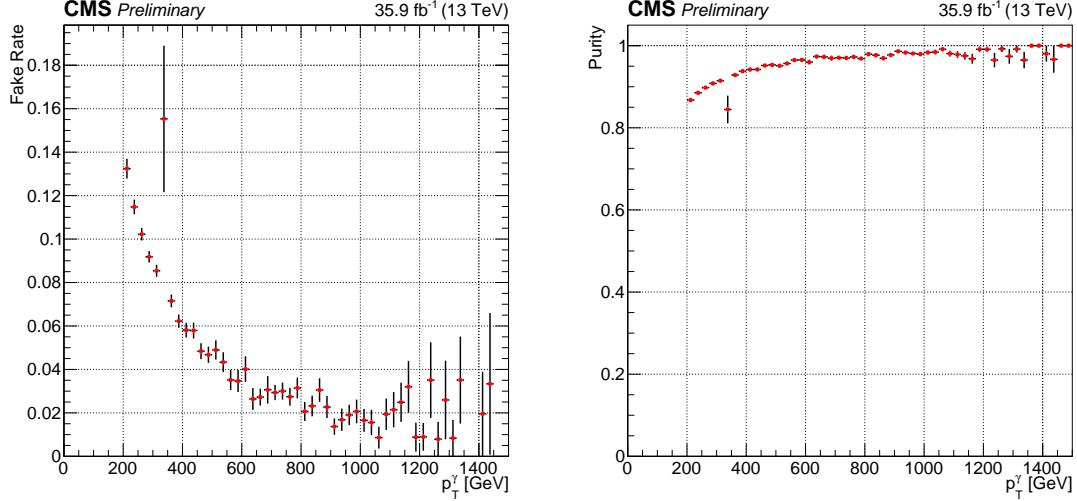


Figure 6.3: Plots for Fake Rate (left) and Purity (right) as a function of the photon p_T are shown. The events selected are required to have a $p_T > 200$ GeV, be within the ECAL acceptance range, and pass the loose ID selection cuts. This selection was produced in order to verify the values given by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

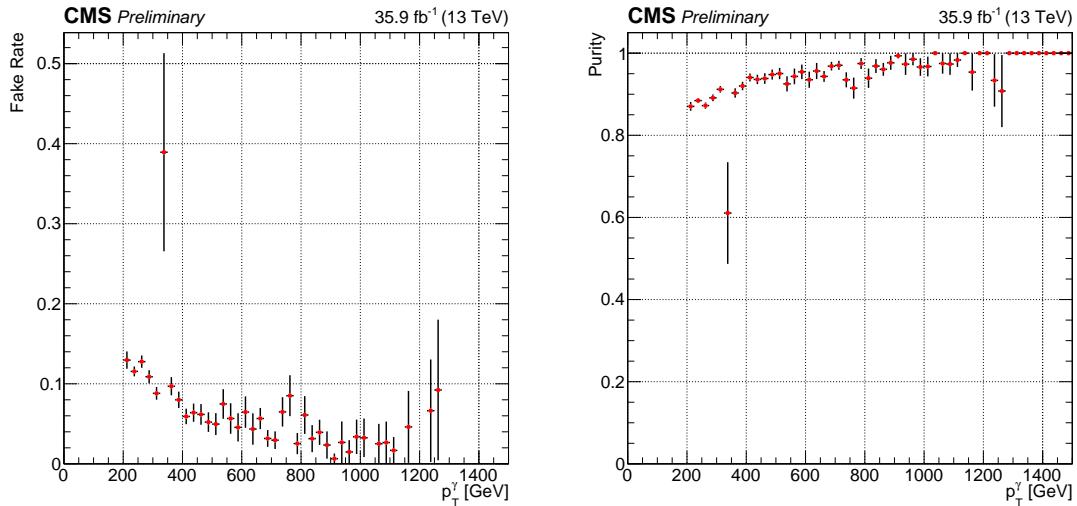


Figure 6.4: Plots for Fake Rate (left) and Purity (right) as a function of the photon p_T are shown. These plots include photons with the full control region selection. Aside from exhibiting lower statistics, the plots seem to agree with the fake rate and purity before all the control region cuts are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

1455 In order to identify prompt photons, reconstructed photons from the $\gamma+jets$ and QCD
 1456 samples are matched to generator-level photons in space and momentum by requiring
 1457 $\Delta R(\gamma_{\text{gen}}, \gamma_{\text{reco}}) < 0.4$ and $0.5 < p_T^{\text{gen}}/p_T^{\text{reco}} < 2.0$, respectively. Any reconstructed photon
 1458 which fails to get matched to a generator level photon is labeled as a fake/non-prompt
 1459 photon. Direct photons are identified by further requiring that the reconstructed photons
 1460 be matched to a parton (a gluon or quark) in space as $\Delta R(\gamma, \text{parton}) > 0.4$. This require-
 1461 ment is intended to distinguish the reconstructed photons from highly boosted π^0 's, which

compose a large portion of the experimentally indistinguishable fragmentation photons. Finally, fragmentation photons are obtained exclusively from QCD simulation and are required to have $\Delta R(\gamma, \text{parton}) < 0.4$ in order to avoid double counting photons from the $\gamma+\text{jets}$ sample.

With all three types of photons defined, a study can be carried out from simulation to estimate their respective contributions to the defined control region. The study takes into account that any reconstructed photon in the $\gamma+\text{jets}$ or QCD samples can only be categorized as prompt (through direct production or fragmentation) or non-prompt (fake). The purity and fake rate can then be defined in terms of the relative proportions of prompt or non-prompt photons with respect to the sum of the contributions from all three types of photons. Identified direct photons are taken from the $\gamma+\text{jets}$ sample exclusively. Meanwhile, the fragmentation and fake photon contributions are taken from the QCD sample. The three quantities are then added together and their respective contributions are determined in terms of the photon p_T (Figure 6.3).

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

1480

$$1481 \quad p_\gamma = \frac{\text{prompt}}{\text{prompt} + \text{fake}},$$

where the prompt photon portion comes from the sum of the direct photons (extracted from the $\gamma+\text{jets}$ sample) and the fragmentation photons (extracted from the QCD sample). The remaining non-prompt (or fake) photons all come from photons in the QCD sample that were not matched to truth-level photons in space and momentum with the specified required conditions. Meanwhile, the photon fake rate (Figure 6.3 and Figure 6.4, left) is defined from this same combination of samples as:

1488

$$1489 \quad f = \frac{\text{fake}}{\text{prompt} + \text{fake}},$$

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

1497 6.4 The $Z \rightarrow \mu^+ \mu^-$ Control Region

1498 The $Z \rightarrow \mu^+ \mu^-$ control region defined in this section is in every respect identical to
 1499 the one applied in the 2016 analysis (subsection 5.3.2). The only difference between the
 1500 2016 method and the one discussed in this chapter is that the Drell-Yan (DY) sample is

1501 only used for the normalization correction of the $Z \rightarrow \nu \bar{\nu}$ background. Therefore, the
 1502 loose $\mu\mu$ control region is not used or applied in the calculation of the scale factors. In
 1503 the following subsections only the tight $\mu\mu$ control region, and its usage to obtain the
 1504 normalization scale factor R_{norm} , is discussed.

1505 6.4.1 Muon ID and Isolation

1506 The muons are selected using the “medium muon” selection [?], per the recommen-
 1507 dation of the muon POG. The muon candidates in this selection satisfy $p_T > 10$ GeV and
 1508 $|\eta| < 2.4$. Other additional cuts are applied to aid in the muon candidate selection, such as
 1509 an impact parameter cut. Muons are also subjected to a PF relative-isolation (also referred
 1510 to as mini-isolation) in which the cone size is inversely proportional to the muon p_T . This
 1511 requirement enforces the p_T within the isolation cone to be at most 20% of the muon p_T
 1512 in order to eliminate events with an isolated muon. Details of the medium photon selec-
 1513 tion are included in [Table 6.2](#) and [Table 6.3](#), while details of the impact parameter cut are
 1514 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

1515 6.4.2 Muon Selection in the Tight Control Region

1516 Events are selected from data samples that contain exactly two oppositely charged
 1517 muons ($\mu^+ \mu^-$), which fall within the invariant mass $81 < m_{ll} < 101$ GeV window for
 1518 the Z boson. Additional cuts for the tight muon selection include baseline requirements
 1519 such as an $H_T > 300$ GeV, $N_j \geq 4$, the $\Delta\phi$ baseline cut on leading jets, a $p_T^{miss} > 250$
 1520 GeV, an $m_{T2} > 200$ GeV and at least 1 top-tagged jet $N_t \geq 1$. In addition, the p_T of
 1521 the two muons are required to be $p_T > 50$ GeV for the leading muon and $p_T > 20$ GeV
 1522 for the sub-leading one. The only difference, when compared to the signal region is the
 1523 missing lepton veto, in addition to the dimuon events being treated as p_T^{miss} . This makes
 1524 for a region that exhibits very similar kinematics to the $Z \rightarrow \nu \bar{\nu}$ signal region, yet suffers
 1525 from a lack of statistics.

1526 6.5 Analysis

1527 In this section a detailed explanation of the calculation of the scale factors for both
 1528 shape and normalization corrections is provided. The following methods make use of
 1529 the loose γ +jets and the tight $\mu\mu$ control regions defined in the previous sections. The
 1530 procedure involves extracting the shape corrections S_γ from the γ +jets control region and
 1531 afterwards obtain a single normalization correction factor R_{norm} from the tight $\mu\mu$ control
 1532 region. Both factors will then be applied to the final prediction of the $Z \rightarrow \nu\bar{\nu}$ background
 1533 in each of the required search bins.

1534 6.5.1 Shape Correction Using the γ + jets Control Sample

1535 In this section the validation of the γ +jets simulation is discussed in terms of the
 1536 shape of the loose photon control region. As it was shown in subsection 6.3.3, this con-
 1537 trol region has high purity for γ +jets events, particularly in regions of high p_T ($\gtrsim 300$).
 1538 In order to apply this correction factor it is assumed that the shape differences between
 1539 data and simulation are similar between $Z \rightarrow \nu\bar{\nu}$ and γ +jets events. This assumption is
 1540 validated in studies which compare the cross-section ratio of Z +jets to γ +jets events [?].
 1541 Figure 6.5 shows the results of this study, conducted in 2014, for both data and Mad-
 1542 Graph simulation with an integrated luminosity of 19.7fb^{-1} and a center-of-mass energy
 1543 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
 1544 both processes becomes nearly constant. It is then a matter of applying a factor to account
 1545 for the difference in the amount of events between the Z +jets and γ +jets events in order
 1546 to obtain the total amount of $Z \rightarrow \nu\bar{\nu}$ +jets events. This factor is obtained from the tight
 1547 $\mu\mu$ control region, as shown in subsection 6.5.2.

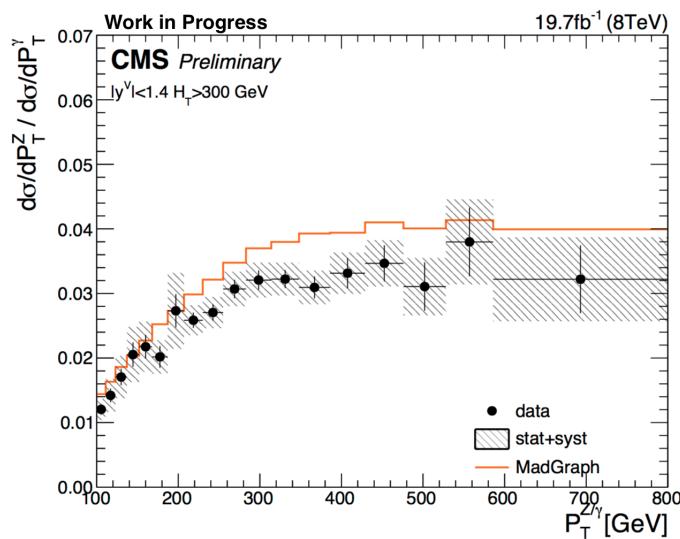


Figure 6.5: Results of study of the Z +jets to γ +jets cross-section ratio for both data and MadGraph sim-
 ulation. For high values of the vector boson transverse momentum, the ratio between these processes is
 observed to be nearly constant.

1548 In order to obtain the shape corrections, the ratio between data and simulation of the
 1549 jet multiplicity distribution is used (Figure 6.6). This is due to it exhibiting the highest
 1550 difference between the observed data and MC. The re-weight for the γ +jets simulation

sample is then accomplished by applying the N_{jet} dependent factor $S_\gamma(N_j)$. This scale factor is determined by taking the ratio of the data and simulation, after subtracting all other MC samples from data events:

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

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

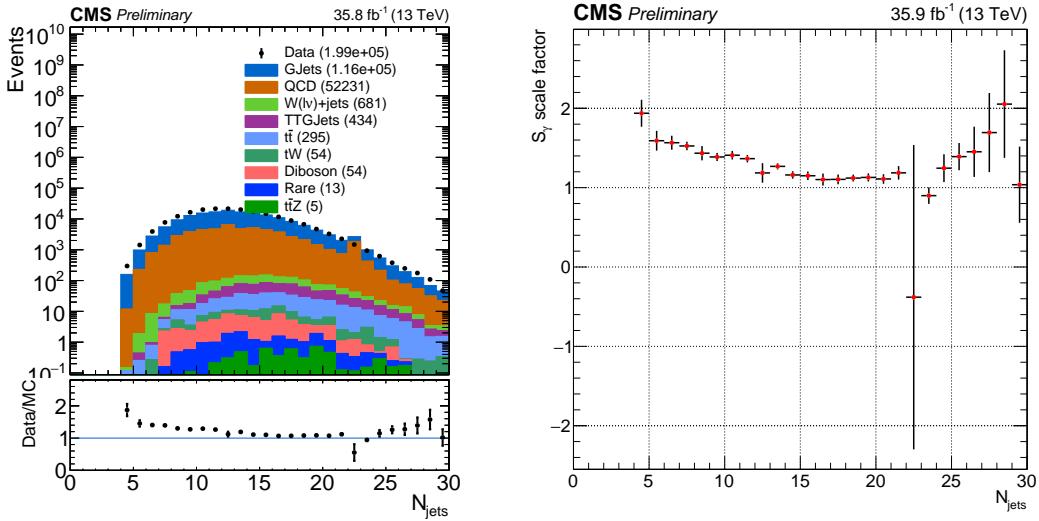


Figure 6.6: Jet multiplicity and the associated S_γ scale factor in the loose photon control region before any corrections are applied.

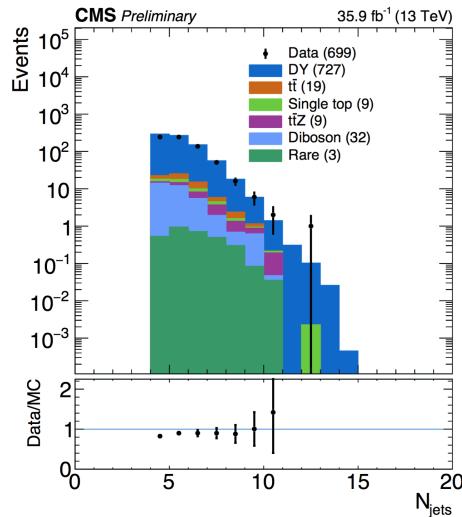


Figure 6.7: N_{jet} distribution in the tight $\mu\mu$ control region after S_γ corrections.

1564 The effect of the $S_\gamma(N_j)$ scale factor is shown for various distributions. These results
 1565 show that the overall agreement between data and simulation improves after applying the
 1566 corresponding shape corrections.

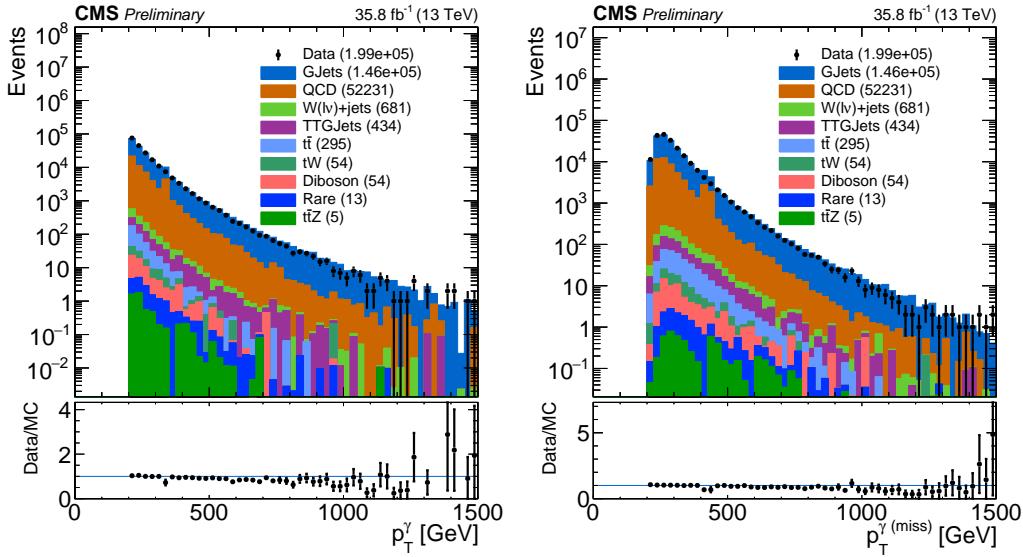


Figure 6.8: p_T^γ (left) and $p_T^{\gamma(\text{miss})}$ (right) distributions after applying the $S_\gamma(N_j)$ scale factor. Comparing to Figure 6.2, an improvement in the agreement between data/MC can be observed.

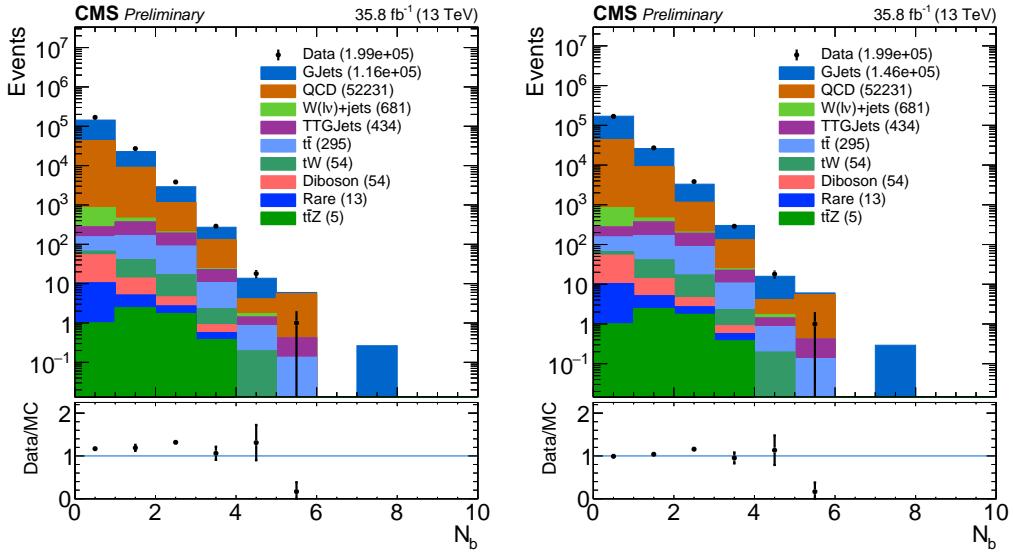


Figure 6.9: N_b distribution before (left) and after (right) applying the $S_\gamma(N_j)$ scale factor.

1567 6.5.2 Normalization Correction Using the tight $Z \rightarrow \mu^+\mu^-$ Control 1568 Sample

1569 In order to constrain the normalization of the $Z \rightarrow \nu\bar{\nu}$ simulation sample, a normaliza-
 1570 tion correction factor R_{norm} is calculated from the tight $\mu\mu$ control region defined in
 1571 subsection 6.5.2. Two categories are considered: the zero b-tagged jet category ($N_b = 0$),
 1572 and the ≥ 1 b-tagged jet category ($N_b \geq 1$). Both of these categories are statistically

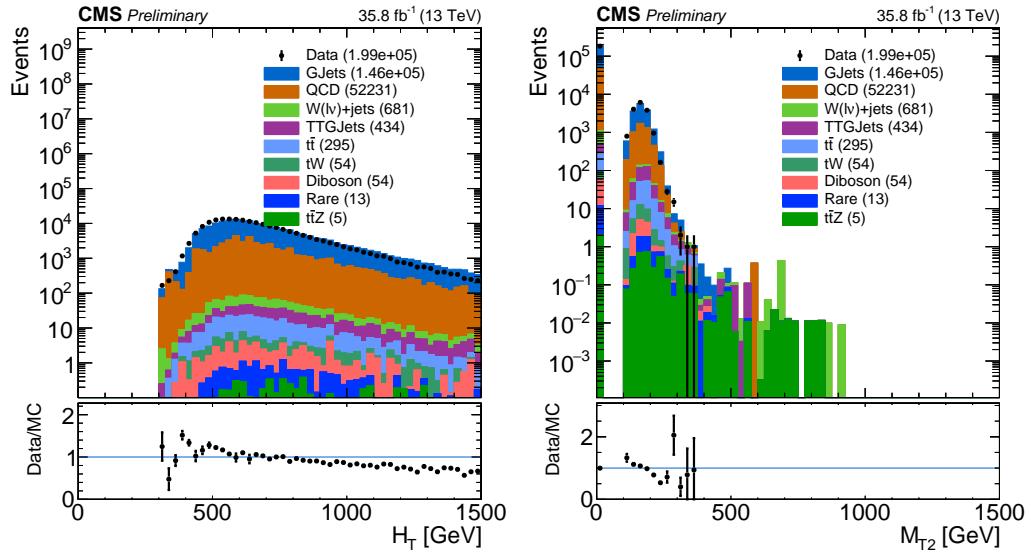


Figure 6.10: H_T and m_{T2} distributions applying the $S_\gamma(N_j)$ scale factor.

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.

1582

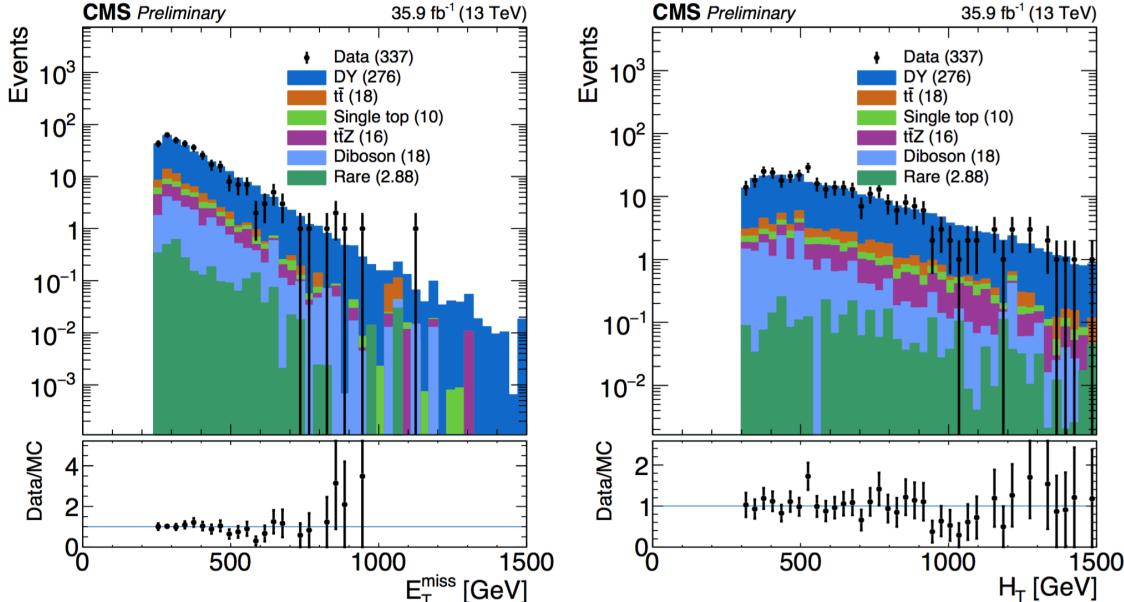


Figure 6.11: Shown are data/MC comparisons for the p_T^{miss} (left) and H_T (right) distributions after applying both the N_j -dependent shape corrections (S_γ) and the global normalization scale factor (R_{norm}).

1583 Data/MC comparisons are shown in [Figure 6.11](#) and [Figure 6.12](#) after applying R_{norm}
 1584 for several distributions in the study. With this final global scale factor all the required
 1585 ingredients for the central value of the $Z \rightarrow \nu\bar{\nu}$ background prediction are obtained.

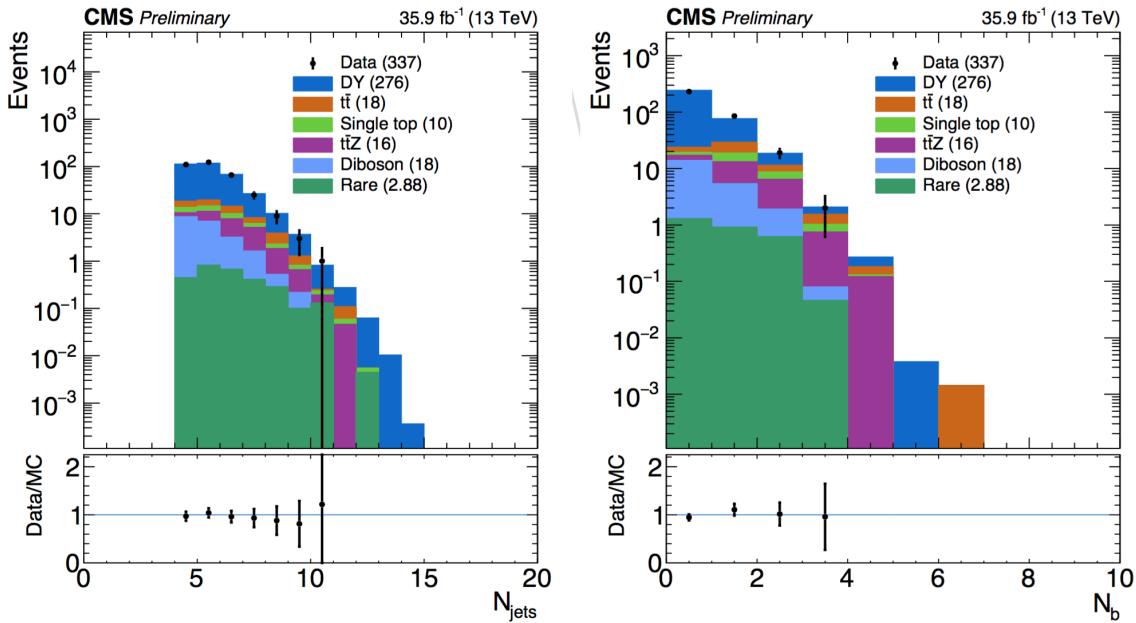


Figure 6.12: Shown are data/MC comparisons for the N_j (left) and N_b (right) distributions after applying both the N_j -dependent shape corrections (S_γ) and the global normalization scale factor (R_{norm}).

1586 6.6 Results

1587 In this section the results for the final estimation of the $Z \rightarrow \nu\bar{\nu}$ are presented.
 1588 The current study includes preliminary results using only data obtained at the CMS de-
 1589 tector during 2016. The results for this study are intended to confirm the assumption that
 1590 the additional $\gamma + \text{jets}$ control region introduced in this analysis reduce the overall uncer-
 1591 tainties obtained in the 2016 analyses (described in [chapter 5](#)). Furthermore, this study is
 1592 intended as a benchmark for future analyses of the SUSY stop group based in Fermilab
 1593 and will be the method used for the 2017 CMS data.

1594 6.6.1 Systematics

1595 Two categories of uncertainties for the $Z \rightarrow \nu\bar{\nu}$ prediction are considered: uncertain-
 1596 ties that are associated to the use of MC simulation and the uncertainties specifically
 1597 associated to the background prediction method. Several sources are acknowledged in the
 1598 first category mentioned such as PDF and renormalization/factorization scale choices, jet
 1599 and p_T^{miss} energy scale uncertainties b-tag scale factor uncertainties, and trigger efficiency
 1600 uncertainties. Given that the simulation sample is normalized to data in the tight control
 1601 region, uncertainties associated with the luminosity and cross-section are excluded. In
 1602 addition, the overall $Z \rightarrow \nu\bar{\nu}$ statistical uncertainty from MC simulation is also taken into
 1603 account.

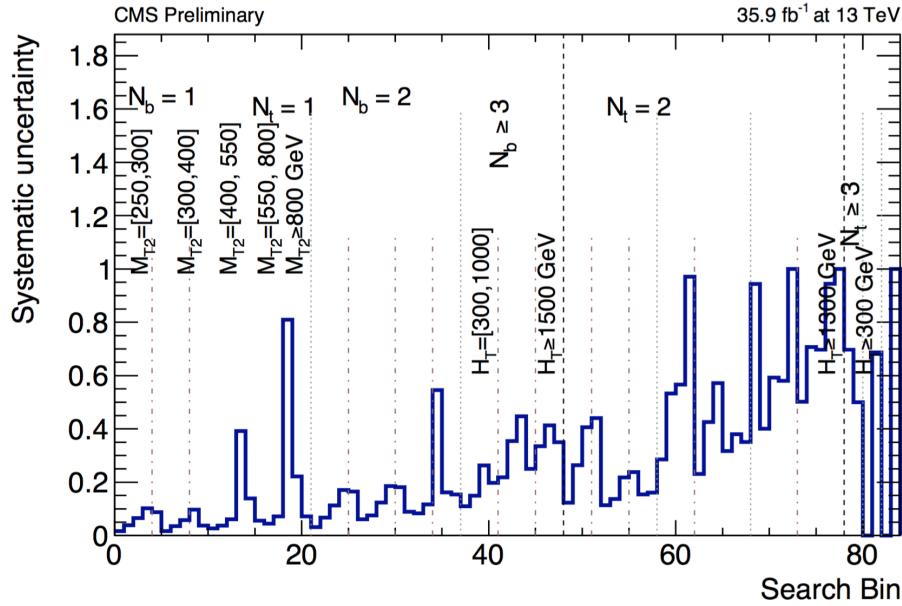


Figure 6.13: Systematic uncertainty in the final prediction, as a function of the search bin, associated to the MC statistics.

1605 The statistical uncertainty associated with each bin in the MC is propagated as a sys-
 1606 tematic uncertainty. The relative uncertainty per bin can be see in [Figure 6.13](#). It shows
 1607 that the uncertainties for the MC vary from as low as 1% up to 81% and even 100% in
 1608 some regions. Since the final estimation is scaled using the global normalization factor
 1609 from the tight $\mu\mu$ control region (R_{norm}), the total uncertainty, due to limited amounts of
 1610 events in data, is propagated in the final prediction. This is also true for the $S_\gamma(N_j)$ scale
 1611 factor, in which the residual differences in search variables other than N_j are evaluated in
 1612 the loose photon control region. Both the uncertainty arising from the N_j re-weighting
 1613 as well as the residual differences are evaluated together. The uncertainty from R_{norm} is
 1614 propagated as a flat value of 7.9% uncertainty per each search bin.

1615 6.6.2 $Z \rightarrow \nu\bar{\nu}$ Estimation for the Search Bins

1616 The final estimation for the $Z \rightarrow \nu\bar{\nu}$ background calculated for all 84 search bins is
 1617 shown in [Figure 6.14](#). The statistical uncertainty in bins that have zero events is treated
 1618 as the average weight (the sum of the weights squared over the weight) times the poisson
 1619 error on 0 which is 1.8. This average weight is calculated on the basis of a relaxed cut in
 1620 which $N_b \geq 2$ is required. For comparison, a cut in which $N_t > 2$ where two tops are
 1621 fake for the $Z \rightarrow \nu\bar{\nu}$ is used.

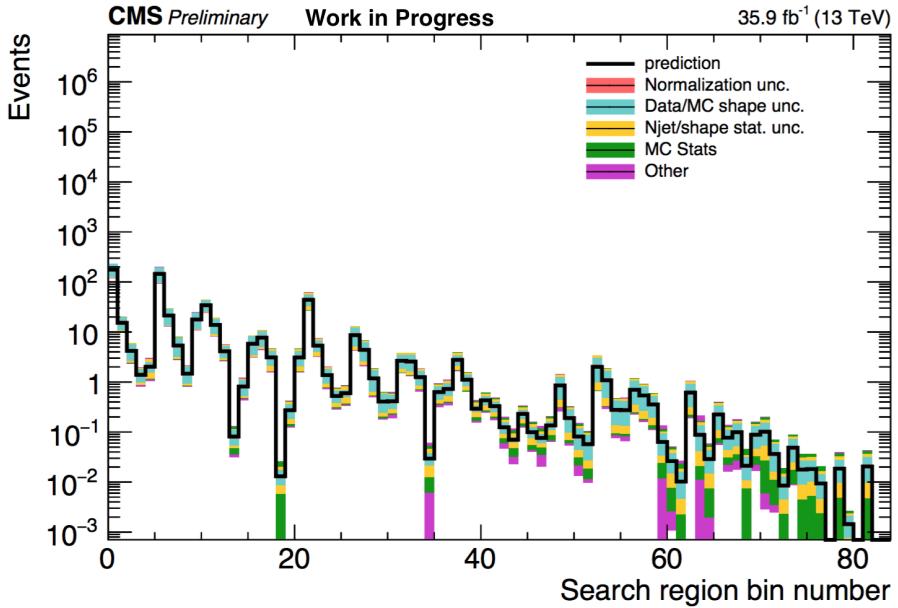


Figure 6.14: $Z \rightarrow \nu\bar{\nu}$ background prediction for all search bins, including the breakdown of the various uncertainties.

1622

Chapter 7

1623

Conclusion

1624 There are three major topics of research that were discussed in this dissertation: The
1625 simulation studies involving the counting of L1-stubs for the HL-LHC CMS Inner Tracker
1626 upgrade ([chapter 4](#)), the overall 2016 search for SUSY in the all-hadronic channel us-
1627 ing a customized top-tagger ([chapter 5](#)) and the improvements made for the estimation
1628 of the $Z \rightarrow \nu\bar{\nu} +$ jets background using an additional control region from $\gamma +$ jets events
1629 ([chapter 6](#)). These studies were explained in detail in their respective chapters and their
1630 individual results are provided. A summary of the most important results from each study
1631 is provided in this chapter.

1632

7.1 L1 Stub Counting for the HL-LHC CMS Tracker Up- 1633 grade

1634 Results from this study (detailed in [chapter 4](#)) reflect the overall effects that were ex-
1635 pected beforehand. The removal of discs from the standard pixel geometry (consisting
1636 of 8 small and 4 large discs) results in a noticeable reduction of stub hits in the upgraded
1637 CMS Outer Tracker. This effect is specially apparent if the disc that is removed is closer to
1638 the interaction point, due to the much larger volume of particles that are present in this re-
1639 gion. Therefore, the reduction in stubs is more pronounced when a small disc is removed
1640 (as in the case of the $7s4l$ geometry) than if a large disc is removed (as in the $8s3l$ pixel
1641 geometry). The reason for this effect stems from the fact that as particles travel through
1642 the various layers of the Inner Tracker material, some of them are bound to interact with
1643 it, producing particles that did not originate from the initial proton-proton collision. The
1644 stubs produced via such processes are considered to be “fake” stubs. To confirm these
1645 findings, an additional study was conducted using a sample that was virtually indistin-
1646 guishable from the standard pixel geometry, but with the second disc on the positive side
1647 “turned off” or “dead”. The results from this study confirm the initial findings and shows
1648 that there is indeed a correlation between the average number of stubs detected in the
1649 Outer Tracker and the total amount of material present in the upgraded Inner Tracker. An
1650 important factor that needs to be taken into account when interpreting these results is the
1651 re-optimization of the disc positions after removing a disc in the different pixel geome-
1652 tries considered. This feature could provide a possible explanation as to why the $6s3l$
1653 geometry, which has two less small discs than the standard geometry (and one less large
1654 one), was found to have less of an effect on the average number of stubs than the $7s4l$

1655 geometry.

1656 7.2 Search for SUSY in the All-Hadronic Channel

1657 The analysis presented in [chapter 5](#) shows the results of a search for SUSY in the
 1658 0-lepton final state using a customized top-tagger. The data was obtained from proton-
 1659 proton collisions at the CMS detector during 2016 with a total integrated luminosity of
 1660 35.9 fb^{-1} at a center-of-mass energy of 13 TeV. The search was conducted by speci-
 1661 fying 84 non-overlapping regions of phase space with varying requirements on the N_b ,
 1662 N_t , p_T^{miss} , H_T and m_{T2} variables ([subsection 5.2.5](#)). Several dominant and non-dominant
 1663 backgrounds were identified and estimated to account for all the majority of the processes
 1664 that were seen in the collected data. The estimation procedures and their respective sys-
 1665 tematic and statistical uncertainties are discussed in [section 5.3](#). The total background
 1666 prediction vs. data for all 84 search bins ([Figure 5.7](#)) shows no statistically significant de-
 1667 viation from the predicted SM background. The biggest sources background were shown
 1668 to be the $t\bar{t}$ and $W+jets$ processes, followed by $Z(\nu\bar{\nu})+jets$, which were seen to be dom-
 1669 inant in regions with a high p_T threshold. Meanwhile, the contributions from the QCD
 1670 multijet and rare backgrounds are found to be nearly negligible in all of the 84 search bins.
 1671 Exclusion limits were calculated from these results for each of the signal models used, by
 1672 applying a binned likelihood fit on the data. The likelihood function was obtained for
 1673 each of the 84 search regions as well as for each of the background data control sam-
 1674 ples from the product of the Poisson probability density function. Exclusion limits were
 1675 placed on the top squark, gluino and LSP production cross-sections with a 95% confi-
 1676 dence level (CL), calculated using a modified frequentist approach with the CL_s criterion
 1677 and asymptotic results for the test statistic. The 95% CL exclusion limits obtained for the
 1678 $T2tt$ model, which consists of direct top squark production, excludes top squark masses
 1679 up to 1020 GeV and LSP masses up to 430 GeV. For the $T1tttt$ model, gluino masses of
 1680 up to 2040 GeV and LSP masses up to 1150 GeV are excluded, with corresponding limits
 1681 of 2020 and 1150 GeV for the $T1ttbb$ model, 2020 and 1150 GeV for the $T5tttt$ model,
 1682 and 1810 and 1100 GeV for the $T5ttcc$ model.

1683 7.3 Estimation of the $Z \rightarrow \nu\bar{\nu} + jets$ Background with a 1684 Hybrid Method

1685 [Chapter 6](#) presents a different method (than the one briefly described in [chapter 5](#))
 1686 for estimating the total amount of $Z \rightarrow \nu\bar{\nu} + jets$ events per search bin. The new method
 1687 makes use of an additional ‘loose $\gamma + jets$ ’ control sample for the estimation of the $S_\gamma(N_j)$
 1688 shape correction factor, in addition to the ‘tight $Z \rightarrow \mu\mu + jets$ ’ sample (used in the 2016
 1689 analysis) to calculate the R_{norm} correction factor ([subsection 5.3.2](#)). The $\gamma + jets$ sample
 1690 was chosen to substitute the previously used ‘loose $Z \rightarrow \mu\mu + jets$ ’ control region, due to
 1691 its much higher cross-section and kinematic similarity to the $Z \rightarrow \nu\bar{\nu} + jets$ process at high
 1692 p_T . The implementation of this method sought to refine the results obtained in 2016 by
 1693 reducing the statistical uncertainties that stem from the low branching fraction of $Z \rightarrow \mu\mu +$
 1694 jets events. The $\gamma + jets$ control region was found to have a high purity and a low fake-rate
 1695 in the regime of high p_T that was being studied. The N_j -dependent S_γ^i was obtained for

1696 each bin i in N_j from a comparison between data and the $\gamma +$ jets MC, after subtracting
1697 the other backgrounds from data and normalizing both samples to 1. This resulted in the
1698 S_γ scale factor plot depicted in [Figure 6.6](#), on the right. As can be seen from this plot, the
1699 statistical uncertainty of the various correction factors is small due to the large number
1700 of events available. Using these values, the R_{norm} normalization correction factor was
1701 obtained from the tight $\mu\mu$ control sample as $R_{norm} = 1.070 \pm 0.085$. Both of these
1702 scale factors were then applied to the final estimation of the $Z \rightarrow \nu\bar{\nu} +$ jets background
1703 ([Figure 6.14](#)).