

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Search for Supersymmetry in Proton-Proton Collisions at a Center-of-Mass Energy of 13 TeV with a Customized Top Quark Tagger in the All-Hadronic Final State

Permalink

<https://escholarship.org/uc/item/55n2j7fm>

Author

Wei, Hua

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Search for Supersymmetry in Proton-Proton Collisions at a Center-of-Mass Energy
of 13 TeV with a Customized Top Quark Tagger in the All-Hadronic Final State

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Hua Wei

December 2017

Dissertation Committee:

Prof. John William Gary, Chairperson
Prof. Haibo Yu
Prof. Owen Long

Copyright by
Hua Wei
2017

The Dissertation of Hua Wei is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

Now it is the end of this adventure. It is a just a small one, comparing with the big world. But, I can hardly survive without the supports of my colleagues and friends. I owe a lot of thanks since I left CERN in a hurry.

First, I would like to give a big thank you to my advisor, Bill. I can hardly to have such a good chance to work at CERN, with CMS HCAL guys and on SUSY analysis, without his support. He also put huge efforts to help me to improve the quality of thesis, which I am really appreciated. He helps a lot on both academic work and real life issue. Our UCR sub-group meeting, hosting by Bill and Owen on weekly basis, is a great place to share working progress and ideas. I benefit a lot from the reports and discussions from Florent, Amin and Georgia in the meeting.

Working on CMS detector is not easy. A new member working on the detector is just like a new swimmer struggling in the sea. But, I am lucky enough to have some nice guys supporting me when I am drowning. Jared is the guy who helps me to start the HCAL related work at very beginning. I work under Dick's guidance after Jared left. Dick is a real expert on HCAL. It is indeed a great honor to work with him. I hope I can be a person like him in future: leading the work, helping young people. Josh, I cannot remember how many times we work together to deal with HCAL in the midnight. I bet you will miss me also, especially when you are trying to find someone to practice Chinese in the intervals of operation work. Edmund helps a lot on dealing with the damn HO mapping. I was impressed by his extremely high profile skill in coding. Ted L helps me a lot on the map fiber id checking. His cool head and professional attitude is always valuable. Seth,

John and Martin, I will never forget the days we work together: we are waked up in the mid-night, debugging on the HCAL, and finally fixed the issue.

The analysis work is also challenging, especially to a beginner. However, I am just so lucky to have Florent and Hongxuan as analysis conveners. They answered tons of rudimental questions from me when I was starting the analysis work. They also gave me a lot of good advises when I was getting stuck on work. Kenichi's serious and tireless attitudes on the work are highly respectful in the group. I benefit a lot from his wide knowledge on both physics analysis and detector. It is a great honor to work with all the stop team alpha members: Zhenbin, Joe, Nadja, Scarlet, Koushik, Joshi, Hui, etc. I indeed learned a lot from your folks.

Life in Geneva is wonderful, peaceful and productive. However, it was a bit sad to leave my friend in Riverside and move to Geneva at the beginning. Mingjian and Peng are my primary roommates at CERN. I knew them since I was an undergraduate student. It was a great reunion, until they passed away in a car accident in 2015. There are lots of words I want to but have no chance to tell them.

Fortunately I still have friends to support each other, in both daily life and work. Menglei is such a great friend. It is always pleasant when talking with her, on both daily life and physics topics (ECAL, DQM, SUSY, etc.). Manuel is my last office mate. I was lucky to catch his defense before he move on to his new post-doc job. Huilin is a friend and an "enemy". He works on stop search team A, a great competitor. We are did not exchange much idea on physics until my last year. I benefit a lot from him on the discussion of convolutional neural network. Hualin is a reliable friend. It is always a pleasure to chat

with him on various topics, like CSC, Higgs to ZZ 4 lepton analysis, etc. He also help on to babysitter my car when I left the town.

One of the most interesting things at CERN is that you can always find someone to have physics night after party. I always remember Qi (Atlas), Xiangyang (Atlas), and Jingyu (CMS) had such a night together. We discuss various high-energy physics topics (and drinking) until 4:00 a.m. in the morning. I would also give a special thanks to Jingyu, for hosting me when I visit fermilab, it was indeed a good night for me (only) to watch Cavalry beat Warrior in the buffalo chicken wing, haha.

Yali is my dearest friend. She always covered my back when I was on fire. Her professional knowledge on feminism (although I never take the whole package, haha) and other aspects of expertise in social science give me a brand new view of the world. I never feel tired when talking with her.

Finally, my deepest appreciation belongs to my family, especially to my parents. Sometimes, it is just so supportive to know that there is a place I can always come back.

To my beloved grandpa, who always encourages me to pursue science. But I lost
him in my second year of Ph.D.

ABSTRACT OF THE DISSERTATION

Search for Supersymmetry in Proton-Proton Collisions at a Center-of-Mass Energy of 13 TeV with a Customized Top Quark Tagger in the All-Hadronic Final State

by

Hua Wei

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, December 2017
Prof. John William Gary, Chairperson

This thesis presents a search for direct and gluino-mediated production of supersymmetric scalar top-quark pairs in the all-hadronic final state using top quark tagging. The results of search are based on a sample of proton-proton collision events collected at a center-of-mass energy of 13 TeV with the CMS detector at the CERN Large Hadron Collider, corresponding to an integrated luminosity of 36 fb^{-1} . The results of the search are interpreted in several simplified models for supersymmetric particle production.

Contents

List of Figures	xi
List of Tables	xx
1 Introduction	1
2 The standard model and supersymmetry	3
2.1 The standard model	3
2.1.1 Description	4
2.1.2 Challenges	7
2.2 The supersymmetric extension of the standard model	8
2.2.1 Minimum supersymmetric standard model (MSSM)	8
2.2.2 SUSY naturalness	10
2.2.3 SUSY simplified models	12
3 The LHC and CMS experiment	14
3.1 The LHC	14
3.1.1 LHC: Machine layout and performance	17
3.1.2 LHC: From an operational point of view	20
3.2 The CMS experiment	24
3.2.1 CMS detector system	24
3.2.2 Event reconstruction	41
3.2.3 Event simulation	51
4 SUSY Search with a customized top tagger	56
4.1 Introduction	56
4.2 Customized top tagger	58
4.2.1 Motivation	58
4.2.2 Description of the method	59
4.3 Event selection and search bin design	66
4.3.1 Trigger	69
4.3.2 Baseline selection	77

4.3.3	Search regions	82
4.3.4	MC samples for background and signal Studies	85
4.4	Background estimation	89
4.4.1	Backgrounds from top and W decays	89
4.4.2	Backgrounds from neutrinos in Z decays	120
4.4.3	Backgrounds from QCD multi-jet	122
4.4.4	Backgrounds from TTZ and other standard model rare process	135
4.5	Results	143
5	Conclusions	152
	Bibliography	153
A	Supplementary material for analysis	158
A.1	Quark-gluon discriminator study	158
A.2	Di-Top jets event display	165
A.3	Simplified top tagger, aggregate search bin and results	166
B	Service work	170
B.1	HCAL logical map	170
B.1.1	Remapped phase 1 HB in 2018	172
B.1.2	Phase 2 HB in 2020	186
B.1.3	Phase 1 HE in 2018	199
B.1.4	Phase 1 HF in 2017	212
B.1.5	HO since 2015	225

List of Figures

3.1	The LHC full injection chain.	15
3.2	The PS layout.	16
3.3	Schematic of the Bunch Disposition around an LHC Ring for the 25 ns Filling Scheme.	17
3.4	The LHC layout.	18
3.5	Two dimensional CMS inner tracking system layout, phase 0.	26
3.6	Phase 1 HCAL tower segmentation in the r,z plane.	30
3.7	Backend coordinates in Frontend coordinates, HCAL endcap phase 1 backend plus phase 0 frontend in 2016 operation.	33
3.8	Detector coordinates in Frontend coordinates, HCAL endcap phase 1 backend plus phase 0 frontend in 2016 operation.	34
3.9	Two dimensional CMS inner tracking system layout, phase 0.	36
3.10	Dataflow for the overall phase 1 trigger system.	40
3.11	CMS offline computing structure.	42
3.12	Left: Illustration of the underlying event modeled by PYTHIA, together with hard process. Right: Illustration of the correlations in the azimuthal angle relative to the direction of the leading charged jet.	53
4.1	Signal models of interest in this search: top squark pair production with the top squark decaying into a top quark and neutralino (top), and top squarks from cascade decays of gluinos (middle and bottom). The SUSY simplified model topology shown at the top is referred to as T2tt, the middle left model as T1tttt, middle right model as T1ttbb the bottom left one as T5tttt and the bottom right one as T5ttcc.	57
4.2	Left: $t\bar{t}$ events final states; Right: Hadronically decayed tops.	59
4.3	Top left: Quark Gluon likelihood for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.	64
4.4	Top left: Quark Gluon likelihood for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.	65

4.5	Efficiency of the top quark tagger as a function of generator-level top quark p_T for the monojet (red boxes), dijet (magenta upper-triangles), and trijet (green lower-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.	67
4.6	Top-jet reconstruction mistag rate, measured in Z +jets simulation samples. Left plot is the distribution in p_T^{miss} , Right plot is in N_{jets} . The red curve is the mistag rate for cut-based top tagger, the blue curve is the mistag rate for multivariate top tagger. The green, black and yellow are the mistag rate for tri-jet, di-jet and mono-jet, respectively.	68
4.7	The trigger efficiency, denote by the black point, as a function of the offline p_T^{miss} for (left) $300 < H_T < 1000$ and (right) $H_T > 1000$. The error bar indicates the statistical uncertainty of the trigger efficiency. The dash blue line represents the denominator passing the selection, while the solid blue histogram represents the numerator where the denominator events also trigger the search triggers.	71
4.8	The trigger efficiency, denote by the black point, as a function of the offline p_T^{miss} for (left) $300 < H_T < 1000$ and (right) $H_T > 1000$. The error bar indicates the statistical uncertainty of the trigger efficiency. The blue square represents efficiency measured with single-muon dataset. The red point represents efficiency measured with single-electron dataset while the green triangle represents efficiency measured with HT dataset.	73
4.9	The trigger efficiency, denote by the black point, as a function of the offline (left) number of jets with $p_T > 30$ GeV and (right) number of b -tagged jets with $p_T > 30$ GeV. The error bar indicates the statistical uncertainty of the trigger efficiency. The blue square represents efficiency measured with single-muon dataset. The red point represents efficiency measured with single-electron dataset while the green triangle represents efficiency measured with HT dataset.	74
4.10	The trigger efficiency, denote by the black point, as a function of the offline p_T^{miss} for (left) $300 < H_T < 1000$ and (right) $H_T > 1000$. The error bar indicates the statistical uncertainty of the trigger efficiency. The dash blue line represents the denominator passing the selection, while the solid blue histogram represents the numerator where the denominator events also trigger the search triggers.	75
4.11	The trigger efficiency as a function of the offline leading Muon (left) p_T and (right) η .	77
4.12	The trigger efficiency as a function of the offline (left) number of jets with $p_T > 30$ GeV and (right) number of b -tagged jets with $p_T > 30$ GeV, with leading reconstructed muon p_T above 50GeV from the single-muon trigger. The red point denote the measurement from single-electron dataset. The green triangle represents measurement from HT dataset, while blue square represents measurement from MET dataset.	78

4.13	Comparison of the distributions between total SM backgrounds from simulation and several signal points for N_{tops} at the left and $N_{\text{b-jets}}$ at the right. Total SM backgrounds and signals are scaled to same data yield for a shape comparison. The yields for the Data and the SM backgrounds are in the legend. The scale is included in the legend for the signal points.	84
4.14	Comparison of the distributions between total SM backgrounds from simulation and several signal points for $p_{\text{T}}^{\text{miss}}$ at the top left and $M_{\text{T}2}$ at the top right with H_{T} at the bottom. Total SM backgrounds and signals are scaled to same data yield for a shape comparison. The yields for the Data and the SM backgrounds are in the legend. The scale is included in the legend for the signal points.	91
4.15	Original search bin definitions after baseline cuts (in total 84 search bins).	92
4.16	Background pie chart, from simulation.	93
4.17	Translation factors for the τ_{h} background prediction with their uncertainties from limited MC statistics for both muon and electron CS.	94
4.18	Translation factors for the lost lepton background prediction with their uncertainties from limited MC statistics for both muon and electron CS. . . .	95
4.19	Predicted hadronic tau background yield for a 35.9 fb^{-1} data for all the search regions. Right plot is a zoomed version of left plot. Both statistical and total systematic uncertainties are shown.	95
4.20	Predicted lost lepton background yield for a 35.9 fb^{-1} data for all the search regions. Right plot is a zoomed version of left plot. Both statistical and total systematic uncertainties are shown.	96
4.21	Sketch of the requirements electrons and muons from W decays must meet in order to be rejected by the explicit lepton veto.	98
4.22	Muon isolation efficiencies as a function of the muon p_{T} and the activity around the muon. Only statistical uncertainties are shown.	101
4.23	Electron isolation efficiencies as a function of the electron p_{T} and the activity around the electron. Only statistical uncertainties are shown.	101
4.24	Muon reconstruction efficiencies as a function of the muon p_{T} and Eta . Only statistical uncertainties are shown.	102
4.25	Electron reconstruction efficiencies as a function of the electron p_{T} and Eta . Only statistical uncertainties are shown.	102
4.26	Acceptance efficiencies for muons and electrons in each of the search bins. Only the statistical uncertainties are shown.	103
4.27	Isolated track veto efficiencies for each search bin.	105
4.28	The lost lepton background in all the search regions of the analysis as predicted directly from $t\bar{t}$, single top and W +jets simulation (in red) and as predicted by applying the lost lepton background-determination procedure to simulated muon control sample (in black). The lower panel shows the ratio between the true and predicted yields. The top plot shows the prediction computed from the muon control sample. The bottom plot shows the prediction from the electron control sample.	107

4.29	Lost lepton background predictions on muon control sample, in red. The blue points are the results obtained with the average TF method. The uncertainties include both the statistical and systematic uncertainties. Bottom left plot is a zoom of the top plot and the bottom right plot is log scale. . .	112
4.30	Lost lepton background predictions on electron control sample, in red. The blue points are the results obtained with the average TF method. The uncertainties include both the statistical and systematic uncertainties. Bottom left plot is a zoom of the top plot and the bottom right plot is log scale. . .	113
4.31	Tau jet visible energy fraction templates in tau lepton p_T bins.	120
4.32	$Z \rightarrow \nu\nu$ background prediction for all search bins, including the breakdown of the various uncertainties.	123
4.33	(a) Example of an event passing the $\Delta\phi_{\text{jets}}$ and p_T^{miss} cut. p_T^{miss} is well separated from the leading three jets. (b) Example of an event failing the $\Delta\phi_{\text{jets}}$ and p_T^{miss} cut. p_T^{miss} is well aligned with one of the leading jets and most likely arises from jet mismeasurement.	125
4.34	Basic kinematic distributions in the QCD control sample. The uncertainty relates to the total statistical uncertainty. The variables are p_T^{miss} on top, M_{T2} (left) and H_T (right) in middle and number of b-jets (left) and number of top candidates (right) on bottom.	126
4.35	T_{QCD}^{MC} versus p_T^{miss} and M_{T2} distributions for N_{tops} and $N_{\text{b-jets}} < 3$ requirements. Both the sideband point, $200 \text{ GeV} < p_T^{\text{miss}} < 250 \text{ GeV}$, and the signal region are included.	129
4.36	T_{QCD}^{Scale} versus p_T^{miss} and M_{T2} distributions Both the sideband point, $200 \text{ GeV} < p_T^{\text{miss}} < 250 \text{ GeV}$, and the signal region are included.	130
4.37	Closure test of the method: prediction compared to expectation in H_T binned QCD MC samples. The uncertainty relates to the total statistical uncertainty of the prediction. The shown variables are p_T^{miss} on top, M_{T2} (left) and H_T (right) in middle and number of b jets (left) and number of top jets (right) on bottom.	132
4.38	QCD background Closure tests in all search bins with T_{QCD}^{MC}	133
4.39	QCD background predictions in all search bins.	135
4.40	Dominant leading order Feynman diagrams for $t\bar{t}W^+$ and $t\bar{t}Z$ production at the LHC. The charge conjugate process of $t\bar{t}W^+$ produces $t\bar{t}W^-$	137
4.41	$t\bar{t}Z$ validation in the three-leptons channels from the Single Muon data. The error bar denotes the statistical uncertainty.	142
4.42	Systematic uncertainties from different sources that contribute to the $t\bar{t}Z$ background prediction, normalized to 36 fb^{-1}	142
4.43	Yield of the $t\bar{t}Z$ and rare background prediction normalized to 36 fb^{-1}	143
4.44	Observed event yields in data (black points) and predicted SM background (filled solid area) for the 84 search bins. The lower panel shows the ratio of data over total background prediction in each search bin. For both panels, the error bars show the statistical uncertainty associated with the observed data counts, and the grey (blue) hatched bands indicate the statistical (systematic) uncertainties in the total predicted background.	144

4.45	Exclusion limit at 95% CL for the signal models in this search: top squark pair production with the top squark decaying into a top quark and neutralino (top), and top squarks from cascade decays of gluinos (middle and bottom). The top is T2tt, the middle left model as T1tttt, middle right model as T1ttbb the bottom left one as T5tttt and the bottom right one as T5ttcc.	151
A.1	Top left: Quark Gluon $p_t D$ for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.	159
A.2	Top left: Quark Gluon $p_t D$ for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.	160
A.3	Top left: Quark Gluon multiplicity for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.	161
A.4	Top left: Quark Gluon multiplicity for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.	162
A.5	Top left: Quark Gluon Axis2 for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.	163
A.6	Top left: Quark Gluon Axis2 for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.	164
A.7	Event display for di-topjets event in data. Left plot: Event display for a SUSY candidate event (from Run 274250, Event number 425322875) with two reconstructed top-tagged candidates (one is in di-jet category, the other is in tri-jet category) passed the search region selection, in $r - \phi$ view of the CMS detector. The momenta of the AK4 jets that tagged as b-jet in tri-jet category are marked in green, while other AK4 jets are marked in blue. The momentum of the AK8 jets are marked in yellow and missing transverse momentum is marked in purple. Right plot: Event display for a SUSY candidate event (from Run 274969, Event number 766814305) with two reconstructed top-tagged candidates (one is in mono-jet category, the other is in tri-jet category) passed the search region selection, in $r - \phi$ view of the CMS detector. The momenta of the AK4 jets that tagged as b-jet in tri-jet category are marked in green, while other AK4 jets are marked in blue. The momentum of the AK8 jets are marked in yellow and missing transverse momentum is marked in purple.	165
A.8	Observed event yields (black points) and predicted SM background (filled solid areas) for the 10 aggregate search regions. The lower panel shows the ratio of the data to the total background prediction. The hatched bands correspond to the total uncertainty in the background prediction.	168

B.1	HCAL (phase 1 HB, plus side) detector η distribution in the frontend electronic coordinates.	174
B.2	HCAL (phase 1 HB, plus side) detector ϕ distribution in the frontend electronic coordinates.	175
B.3	HCAL (phase 1 HB, plus side) detector depth distribution in the frontend electronic coordinates.	176
B.4	HCAL (phase 1 HB, minus side) detector η distribution in the frontend electronic coordinates.	177
B.5	HCAL (phase 1 HB, minus side) detector ϕ distribution in the frontend electronic coordinates.	178
B.6	HCAL (phase 1 HB, minus side) detector depth distribution in the frontend electronic coordinates.	179
B.7	HCAL (phase 1 HB, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	180
B.8	HCAL (phase 1 HB, plus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	181
B.9	HCAL (phase 1 HB, plus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	182
B.10	HCAL (phase 1 HB, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	183
B.11	HCAL (phase 1 HB, minus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	184
B.12	HCAL (phase 1 HB, minus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	185
B.13	HCAL (phase 2 HB, plus side) detector η distribution in the frontend electronic coordinates.	187
B.14	HCAL (phase 2 HB, plus side) detector ϕ distribution in the frontend electronic coordinates.	188
B.15	HCAL (phase 2 HB, plus side) detector depth distribution in the frontend electronic coordinates.	189
B.16	HCAL (phase 2 HB, minus side) detector η distribution in the frontend electronic coordinates.	190
B.17	HCAL (phase 2 HB, minus side) detector ϕ distribution in the frontend electronic coordinates.	191
B.18	HCAL (phase 2 HB, minus side) detector depth distribution in the frontend electronic coordinates.	192
B.19	HCAL (phase 2 HB, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	193
B.20	HCAL (phase 2 HB, plus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	194
B.21	HCAL (phase 2 HB, plus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	195
B.22	HCAL (phase 2 HB, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	196

B.23	HCAL (phase 2 HB, minus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	197
B.24	HCAL (phase 2 HB, minus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	198
B.25	HCAL (phase 1 HE, plus side) detector η distribution in the frontend electronic coordinates.	200
B.26	HCAL (phase 1 HE, plus side) detector ϕ distribution in the frontend electronic coordinates.	201
B.27	HCAL (phase 1 HE, plus side) detector depth distribution in the frontend electronic coordinates.	202
B.28	HCAL (phase 1 HE, minus side) detector η distribution in the frontend electronic coordinates.	203
B.29	HCAL (phase 1 HE, minus side) detector ϕ distribution in the frontend electronic coordinates.	204
B.30	HCAL (phase 1 HE, minus side) detector depth distribution in the frontend electronic coordinates.	205
B.31	HCAL (phase 1 HE, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	206
B.32	HCAL (phase 1 HE, plus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	207
B.33	HCAL (phase 1 HE, plus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	208
B.34	HCAL (phase 1 HE, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	209
B.35	HCAL (phase 1 HE, minus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	210
B.36	HCAL (phase 1 HE, minus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	211
B.37	HCAL (phase 1 HF, plus side) detector η distribution in the frontend electronic coordinates.	213
B.38	HCAL (phase 1 HF, plus side) detector ϕ distribution in the frontend electronic coordinates.	214
B.39	HCAL (phase 1 HF, plus side) detector depth distribution in the frontend electronic coordinates.	215
B.40	HCAL (phase 1 HF, minus side) detector η distribution in the frontend electronic coordinates.	216
B.41	HCAL (phase 1 HF, minus side) detector ϕ distribution in the frontend electronic coordinates.	217
B.42	HCAL (phase 1 HF, minus side) detector depth distribution in the frontend electronic coordinates.	218
B.43	HCAL (phase 1 HF, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	219
B.44	HCAL (phase 1 HF, plus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	220

B.45	HCAL (phase 1 HF, plus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	221
B.46	HCAL (phase 1 HF, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	222
B.47	HCAL (phase 1 HF, minus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.	223
B.48	HCAL (phase 1 HF, minus side) backend electronic coordinate uHTR fiber distribution in the frontend electronic coordinates.	224
B.49	HCAL (phase 1 HO, sector 0) detector η distribution in the frontend electronic coordinates.	226
B.50	HCAL (phase 1 HO, sector 0) detector ϕ distribution in the frontend electronic coordinates.	227
B.51	HCAL (phase 1 HO, sector 1, plus side) detector η distribution in the frontend electronic coordinates.	228
B.52	HCAL (phase 1 HO, sector 1, plus side) detector ϕ distribution in the frontend electronic coordinates.	229
B.53	HCAL (phase 1 HO, sector 1, minus side) detector η distribution in the frontend electronic coordinates.	230
B.54	HCAL (phase 1 HO, sector 1, minus side) detector ϕ distribution in the frontend electronic coordinates.	231
B.55	HCAL (phase 1 HO, sector 2, plus side) detector η distribution in the frontend electronic coordinates.	232
B.56	HCAL (phase 1 HO, sector 2, plus side) detector ϕ distribution in the frontend electronic coordinates.	233
B.57	HCAL (phase 1 HO, sector 2, minus side) detector η distribution in the frontend electronic coordinates.	234
B.58	HCAL (phase 1 HO, sector 2, minus side) detector ϕ distribution in the frontend electronic coordinates.	235
B.59	HCAL (phase 1 HO, sector 0) backend electronic coordinate crate distribution in the frontend electronic coordinates.	236
B.60	HCAL (phase 1 HO, sector 0) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.	237
B.61	HCAL (phase 1 HO, sector 0) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.	238
B.62	HCAL (phase 1 HO, sector 0) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.	239
B.63	HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	240
B.64	HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.	241
B.65	HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.	242
B.66	HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.	243

B.67	HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	244
B.68	HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.	245
B.69	HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.	246
B.70	HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.	247
B.71	HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	248
B.72	HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.	249
B.73	HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.	250
B.74	HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.	251
B.75	HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.	252
B.76	HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.	253
B.77	HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.	254
B.78	HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.	255

List of Tables

2.1	The Fundamental Interactions.	4
2.2	List of the fields of the MSSM and their irreducible representations.	9
3.1	Accelerator Mode.	21
3.2	Beam Mode.	23
3.3	Numbers of RPCs for different wheels.	38
3.4	Average Jet components in GEANT4 Simulation.	48
4.1	Jet bin for quark-gluon discriminator study.	63
4.2	Muon Medium ID 2016 HIP Safe.	80
4.3	Muon Medium ID HIP Safe Good Global Muon.	80
4.4	Additional Impact Parameter cut on Muon.	80
4.5	Electron Cut Based Veto 2016 Data in 80X CMSSW offline reconstruction condition.	81
4.6	Standard model Monte Carlo samples used in the analysis.	86
4.7	Standard model Monte Carlo samples used in the analysis.	87
4.8	Cross sections for a couple of mass points for the T2tt and T1tttt Simplified Models. The selected points were generated with full simulation.	88
4.9	Number of expected lost lepton events (not isolated, not identified/reconstructed and out of the acceptance) in $t\bar{t}$, single top and W+jets simulation events, for a luminosity of 35.9 fb^{-1}	97
4.10	M_T^W correction factors obtained from $t\bar{t}$, single top and W+jets simulation, after baseline selection. Only statistical uncertainties are shown.	100
4.11	Source and contribution of the different components of the lost lepton method systematic uncertainties.	110
4.12	Predicted lost lepton background yield from the muon control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.	114
4.13	Predicted lost lepton background yield from the muon control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.	115
4.14	Predicted lost lepton background yield from the muon control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.	116

4.15	Predicted lost lepton background yield from the electron control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.	117
4.16	Predicted lost lepton background yield from the electron control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.	118
4.17	Predicted lost lepton background yield from the electron control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.	119
4.18	T_{QCD}^{MC} versus p_T^{miss} distributions for N_{tops} or $N_{\text{b-jets}} \geq 2$ requirements.	128
4.19	T_{QCD}^{Scale} versus p_T^{miss} distributions for N_{tops} or $N_{\text{b-jets}} \geq 3$ requirements.	130
4.20	Contributions from different sources of systematic uncertainty to the QCD background prediction.	134
4.21	Predicted QCD backgrounds corresponding to the full 35.9 fb^{-1} data sample in first 37 search bins.	136
4.22	Predicted QCD backgrounds corresponding to the full 35.9 fb^{-1} data sample in following 11 search bins.	137
4.23	Predicted QCD backgrounds corresponding to the full 35.9 fb^{-1} data sample in remaining 36 search bins.	138
4.24	The list of MC samples for $t\bar{t}Z$ and other rare backgrounds.	141
4.25	Expected Exclusion summary table.	149
4.26	Observed Exclusion summary table.	150
A.1	Definition of the aggregate search regions.	167
A.2	The observed number of events and the total background prediction for the aggregate search regions. The first uncertainty in the background prediction is statistical and the second is systematic.	169
B.1	HCAL Logical Map status in 2015.	171
B.2	HCAL Logical Map status in 2016.	171
B.3	HCAL Logical Map status in 2017.	172
B.4	HCAL Logical Map status in 2018.	172
B.5	HCAL Logical Map status in 2023.	173

Chapter 1

Introduction

Particle physics is a branch of the physics that studies the nature of the particles that constitute matter and radiation. There has been a glorious history of discoveries in particle physics in the past 50 years. The standard model is a beautiful assemblage of these remarkable breakthroughs. In 2012, July 4th, the discovery of the Higgs boson provided the last piece of the standard model puzzle.

However, the standard model is not perfect. The hierarchy issue arises with the “small” mass of the Higgs boson. The absence of an explanation in the standard model for cold dark matter is also a blemish on the theory.

The supersymmetric extension of the standard model offers a promising solution to these problems. The hierarchy problem in the standard model can be cured with the introduction of light higgsinos, top squarks, bottom squarks or gluinos, which can decay into lightest supersymmetric particle (LSP). The LSP can be neutral, weakly interacting, and

stable under R-parity [38] conservation, providing a candidate for dark matter. Therefore, supersymmetry represents an exciting theory for us to explore.

The CERN Large Hadron Collider (LHC) was designed and built with the goal of discovering the Higgs boson and making new-physics discoveries. The LHC has a design center-of-mass energy of 14 TeV, which is the highest energy that a human designed experiment can reach so far. This high-energy, high-intensity beast provides us the opportunity to reveal the mask of new physics. The Compact Muon Solenoid (CMS) is a general-purpose detector installed at collision point 5 of LHC. The high quality data is collected with high efficiency in the operation. My analysis is based on the LHC-CMS data set collected in 2016, targeting supersymmetry in the all-hadronic final state.

The results of a search based on missing transverse momentum, extended transverse mass, bottom quark jets (b jets) and top quarks in the final state are presented in this thesis. The backgrounds from standard model processes are carefully estimated with robust methods in this all-hadronic final state search. In addition, our analysis group designed a customized top-quark tagging algorithm in order to obtain high efficiency for all relevant values of the top quark transverse momentum.

The organization of this thesis is described as follows. Chapter 2 gives an overview of the theory of the standard model and its supersymmetric extension, with a discussion of natural SUSY and simplified models. The LHC machine and CMS experiment are described in chapter 3. The analysis method and the physics interpretations are presented in chapter 4. Finally, a conclusion is given in Chapter 5.

Chapter 2

The standard model and supersymmetry

2.1 The standard model

The standard model of particle physics is a mathematical model that describes weak, electromagnetic and strong interactions (Table 2.1) between the quarks and leptons. The standard model represents our best understanding of the subatomic world so far. The development of the standard model has been ongoing since the middle of the last century, through the work of many theoretical and experimental physicists. In the first part of this chapter, we will give a brief review of the standard model. In the second part, the challenges of the standard model will be discussed.

Table 2.1: The Fundamental Interactions.

Interaction	Gravity	Weak	Electromagnetic	Strong
Source	Mass-Energy	Weak charge	Electric charge	Color charge
Strength at the scale of quarks	10^{-41}	10^{-4}	1	60
Mediator	Graviton (hypothetical)	W^+, W^- and Z	Photon	Gluons

2.1.1 Description

The standard model states that the material in the universe is made up of elementary fermions interacting through fields. The particles associated with the interaction fields are bosons. The standard model is described in the quantum field theory framework. The Lagrangian formalism is used in quantum field theory to describe the motion of a particle under the influence of field. The standard model Lagrangian is shown in Eq 2.1:

$$\begin{aligned}
L = & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{8}\text{tr}(F_{\mu\nu}F^{\mu\nu}) - \frac{1}{2}\text{tr}(G_{\mu\nu}G^{\mu\nu}), \text{ (Gauge terms)} \\
& + \begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \bar{\sigma}^\mu i D_\mu \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} + \bar{e}_R \sigma^\mu i D_\mu e_R + \bar{\nu}_R \sigma^\mu i D_\mu \nu_R, \text{ (Lepton dynamical terms)} \\
& - \frac{\sqrt{2}}{v} \left[\begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \phi M^e e_R + \bar{e}_R \bar{M}^e \bar{\phi} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \right], \text{ (Electron, muon, Tau mass terms)} \\
& - \frac{\sqrt{2}}{v} \left[\begin{pmatrix} -\bar{e}_L & \bar{\nu}_L \end{pmatrix} \phi^* M^\nu \nu_R + \bar{\nu}_R \bar{M}^\nu \phi^T \begin{pmatrix} -e_L \\ \nu_L \end{pmatrix} \right], \text{ (Neutrino mass terms)} \\
& + \begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \bar{\sigma}^\mu i D_\mu \begin{pmatrix} u_L \\ d_L \end{pmatrix} + \bar{u}_R \sigma^\mu i D_\mu u_R + \bar{d}_R \sigma^\mu i D_\mu d_R, \text{ (quark dynamical terms)} \\
& - \frac{\sqrt{2}}{v} \left[\begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \phi M^d d_R + \bar{d}_R \bar{M}^d \bar{\phi} \begin{pmatrix} u_L \\ d_L \end{pmatrix} \right], \text{ (Down, strange, bottom mass terms)} \\
& - \frac{\sqrt{2}}{v} \left[\begin{pmatrix} -\bar{d}_L & \bar{u}_L \end{pmatrix} \phi^* M^u u_R + \bar{u}_R \bar{M}^u \phi^T \begin{pmatrix} -d_L \\ u_L \end{pmatrix} \right], \text{ (Up, charm, top mass terms)} \\
& + D_\mu \bar{\phi} D^\mu \phi - m_h^2 [\bar{\phi} \phi - v^2/2]^2 / 2v^2, \text{ (Higgs dynamical and mass terms)}
\end{aligned} \tag{2.1}$$

The definition of derivative operators in the Eq 2.1 is:

$$D_\mu \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} = \left[\partial_\mu - \frac{ig_1}{2} B_\mu + \frac{ig_2}{2} W_\mu \right] \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{2.2}$$

$$D_\mu \nu_R = \partial_\mu \nu_R, \quad D_\mu e_R = [\partial_\mu - ig_1 B_\mu] e_R$$

$$D_\mu \begin{pmatrix} u_L \\ d_L \end{pmatrix} = [\partial_\mu + \frac{ig_1}{6}B_\mu + \frac{ig_2}{2}W_\mu + igG_\mu] \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad (2.3)$$

$$D_\mu u_R = [\partial_\mu + \frac{i2g_1}{3}B_\mu + igG_\mu]u_R, \quad D_\mu d_R = [\partial_\mu - \frac{ig_1}{3}B_\mu + igG_\mu]d_R$$

$$D_\mu \phi = [\partial_\mu + \frac{ig_1}{2}B_\mu + \frac{ig_2}{2}W_\mu]\phi \quad (2.4)$$

As mentioned, the standard model describes the strong, weak and electromagnetic interactions. These three interactions can be understood as arising from gauge symmetries [71]. Therefore, the standard model Lagrangian (Eq 2.1) is invariant under gauge transformations. However, these non-Abelian gauge symmetries are conserved only for massless gauge bosons. The gauge bosons can acquire mass through spontaneous symmetry breaking [59, 57, 58]. In order to preserve the gauge symmetry, a scalar boson (Higgs boson) is introduced to give mass to gauge bosons through spontaneous symmetry breaking. This is called the Higgs mechanism [37, 44]. The last piece of the standard model, the Higgs boson, was discovered at CERN in 2012 [7, 28]. This discovery represents a great triumph for the standard model [69, 43, 61].

According to Eq 2.1, if we ignore the neutrino mass term, there are 19 free parameters in the standard model. The 19 free parameters are: three lepton masses, six quark masses, three mixing angles and one CP phase in the CKM matrix [21, 49], three gauge couplings, QCD vacuum angle, Higgs vacuum expectation value and Higgs boson mass. All the parameters have now been measured directly or indirectly by different high-energy experiments. However, this is not the end of the story.

2.1.2 Challenges

The standard model is very successful, but not perfect. There are issues the standard model does not resolve:

- The hierarchy problem: Another important property of the standard model is renormalizability. A renormalizable quantum field theory has only a finite number of divergent terms in the expansion of a perturbative calculation. Those divergences can be absorbed by the free parameters in the Lagrangian. Therefore, the physical variable will be finite in all orders of calculation. In the standard model, the free parameters may or may not vary with the evolution of the renormalization group, depending on the nature of the higher-order corrections. There are three types of particles in the standard model: fermions, vector gauge bosons, and the scalar Higgs boson. The Higgs boson mass term has a quadratic divergence in the renormalization evolution. The Higgs boson mass (about 125 GeV) input (experimental observation) has to be tuned very carefully with respect to the Planck scale cut off (10^{19} GeV) if we want to cancel the Higgs boson mass evolution in renormalization group. This is the infamous hierarchy problem of the standard model. The other fermions and bosons do not have this fine-tuning concern because they are protected by the symmetries. There are several approaches to solve this issue, for example, Supersymmetry (SUSY) or little Higgs models [64].
- The observation of cold dark matter: The existence of cold dark matter is supported by astrophysical observations. However, the neutrino (even with mass) in the standard model can only explain a very small fraction of dark matter. There are several cold

dark matter candidates, such the Axion introduced to explain strong CP violation, the neutrino in SUSY, etc. But none of them have been detected so far.

SUSY is a promising candidate to explain both of these two problems, and will be discussed in the next section.

2.2 The supersymmetric extension of the standard model

A natural idea to solve the fine-tuning problem is to introduce a symmetry to protect the Higgs field. A basic symmetry relating bosons and fermions can be introduced to solve the hierarchy problem. This is so-called supersymmetry. Furthermore, in R-parity [56] conserved SUSY, we can have a stable, neutral, massive lightest SUSY particle, which is an ideal candidate for cold dark matter. We will use the minimal supersymmetric standard model as an example to demonstrate the basics of this theory.

2.2.1 Minimum supersymmetric standard model (MSSM)

The minimal supersymmetric standard model (MSSM) is the simplest possible supersymmetric extension of the standard model. In the MSSM, each standard model particle has a supersymmetric partner with the same quantum numbers except for spin, which differs by $\frac{1}{2}$. The particles in the MSSM are listed in Table 2.2.

There are two important features of the MSSM. The first is the spontaneous breaking of supersymmetry. The particles in the same super-partner pair must have the same mass in unbroken supersymmetry. However, no super partners have been observed so far. Spontaneously broken supersymmetry is introduced to explain this observation. Heavy spar-

Table 2.2: List of the fields of the MSSM and their irreducible representations.

Super-multiplets	Boson field	Fermionic partners	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Gluon/Gluino	g	\tilde{g}	8	1	0
Gauge/Gaugino	W^+, W^-, Z	$\tilde{W}^+, \tilde{W}^-, \tilde{Z}$	1	3	0
	B^0	\tilde{B}^0	1	1	0
Slepton/Lepton	$(\tilde{\nu}_e, \tilde{e})_L$	$(\nu_e, e)_L$	1	2	-1
	\tilde{e}_R	e_R	1	1	-2
Squark/Quark	$(\tilde{u}, \tilde{d})_L$	$(u, d)_L$	3	2	1/3
	\tilde{u}_R	u_R	3	1	4/3
	\tilde{d}_R	d_R	3	1	-2/3
Higgs/Higgsino	(H_u^+, H_u^0)	$(\tilde{H}_u^+, \tilde{H}_u^0)$	1	2	1
	(H_d^0, H_d^-)	$(\tilde{H}_d^0, \tilde{H}_d^-)$	1	2	-1

ticles, the super-partners of the standard model particles, thus emerge from supersymmetry breaking. Therefore, the Higgs boson mass correction can be expressed by Eq 2.5:

$$\Delta m_H^2 = \frac{1}{8\pi}(\lambda_S - |\lambda_f|^2)\Lambda_{UV} + m_s^2\left(\frac{\lambda}{16\pi^2}\ln(\Lambda_{UV}/m_s)\right) + \dots \quad (2.5)$$

The relation $\lambda_S = |\lambda_f|^2$ occurs in unbroken supersymmetry and still holds in soft supersymmetry breaking [56]. Therefore, the quadratic sensitivity of the Higgs boson mass disappears.

The other property is R-parity. The R-parity is defined as $P_R = (-1)^{3(B-L)+3s}$, where B is the baryon number, L is the lepton number and s is the spin. All the standard model particles have R-parity +1, while supersymmetric particles have R-parity -1. The lightest supersymmetry particle is stable if R-parity is conserved. As a result, the gauginos, higgsinos and sneutrinos can be the cold dark matter candidates. This feature provides a strong motivation to search for the gauginos (also called neutralino) in the R-parity conserved SUSY model.

2.2.2 SUSY naturalness

The naturalness of a theory means that the ratios between free parameters or physical constants in this theory should be on the order of 1, so that the parameters are not fine-tuned. In the hierarchy problem, the scalar Higgs boson mass is fine tuned without symmetry protection. We can define a measure of the fine-tuning by Eq 2.6:

$$\Delta = \frac{2\delta m_H^2}{m_h^2} \quad (2.6)$$

The m_h is the physical neutral CP-even Higgs boson mass, and the m_H is a general linear combination of the various masses of the Higgs fields with coefficients that depend on the mixing angle (β in MSSM). This fine-tuning measure can be used for the sparticle mass constraint.

One can take the top squark mass constraint as an example. The Higgs boson potential in SUSY is corrected by both gauge and Yukawa interactions. The major contribution comes from the top squark loop. The Higgs boson mass correction from the top squark loop can be expressed by Eq 2.7:

$$\delta m_{H_u}^2|_{topsquark} = -\frac{3}{8\pi^2} y_t^2 (m_{Q_3}^2 + m_{u_3}^2 + |A_t|^2) \log\left(\frac{\Lambda_{UV}}{TeV}\right) \quad (2.7)$$

We can re-express this result using top squark mass eigenvalues in Eq 2.8:

$$\delta m_{H_u}^2|_{topsquark} \approx -\frac{3}{8\pi^2} y_t^2 (m_{t_1}^2 + m_{t_2}^2 - 2m_t^2 + \frac{m_{t_1}^2 - m_{t_2}^2}{m_t^2} \cos^2\theta_t \sin^2\theta_t) \log\left(\frac{\Lambda_{UV}}{TeV}\right) \quad (2.8)$$

The requirement of a natural Higgs boson potential sets an upper bound on the top squark mass [60], described by Eq 2.9:

$$\sqrt{m_{t_1}^2 + m_{t_2}^2} \leq 600 \text{ GeV} \frac{\sin\beta}{(1+x_t)^{1/2}} \left(\frac{\log(\Lambda/TeV)}{3}\right)^{-1/2} \left(\frac{m_h}{120 \text{ GeV}}\right) \left(\frac{\Delta^{-1}}{20\%}\right)^{-1/2} \quad (2.9)$$

We can set the upper bound for gluino and higgsino mass in similar way [60]. The higgsino and gluino mass upper bounds can be described by Eq 2.9 and Eq 2.11 respectively.

$$\mu \leq 200 \text{ GeV} \left(\frac{m_h}{120 \text{ GeV}} \right) \left(\frac{\Delta^{-1}}{20\%} \right)^{-1/2} \quad (2.10)$$

$$M_3 \leq 900 \text{ GeV} \sin\beta \left(\frac{\log(\Lambda/\text{TeV})}{3} \right)^{-1} \left(\frac{m_h}{120 \text{ GeV}} \right) \left(\frac{\Delta^{-1}}{20\%} \right)^{-1/2} \quad (2.11)$$

To summarize, the requirements for the natural SUSY are:

- Two top squarks and one bottom squark, all below 500-700 GeV. The “left-handed” bottom squark is in the same SU(2) weak multiplet as the “left-handed” top squark and thereby get a mass constraint from naturalness, since the masses of particles in the same multiplet should not be too different. Since the “right-handed” bottom squark is not in an SU(2) weak multiplet with a top squark, there is no naturalness mass constraint on the right-handed bottom squark.
- Two higgsinos, below 200-350 GeV.
- Not too heavy gluinos, below 900-1500 GeV.

Therefore, the top squarks, bottom squark, higgsino and gluino are the interesting search targets in natural SUSY. The top squarks provide the leading contribution to the Higgs boson mass correction. We have a relatively high probability to find top squark and gluino in hadronic channel since the hadronic process cross section is relatively high at the LHC.

Then we need to consider R-parity. R-parity violating (RPV) SUSY models can also be natural SUSY models. However, since the SUSY LSP can decay into standard model particle in RPV models. There is no cold dark matter candidate in RPV SUSY.

Now, we can put all facts together to motivate a search strategy:

- Naturalness suggests: top squarks, bottom squarks, gluinos, and higgsinos may be light enough to be detected at the LHC.
- Cold dark matter candidate: R-parity conservation.
- Experimental sensitivity: Large cross section at the LHC for gluino and top squark production (for a given SUSY mass value) and large branching fraction to all-hadronic final states.

Therefore, the motivation to search for top squarks and gluinos in the all-hadronic final state is clearly established putting all these thoughts together. The search is performed in all-hadronic final states since this final state has the largest expected branching fraction.

2.2.3 SUSY simplified models

As mentioned in the previous section, hadronic decays to top squarks or gluinos are ideal SUSY search channels at the LHC. However, there are too many free parameters in the MSSM. Setting limits on the parameters can be tricky when there is a high degree of freedom in the parameter phase space.

Simplified models [11] provide a solution to this issue. A simplified model represents a specific event topology, with SUSY mass values treated as free parameters. They

usually contain only 2-4 free parameters (the masses of the SUSY particles), making it much simpler to set limits. Simplified models are widely used in SUSY searches at the LHC [30].

In this thesis, the model topologies we are interested in are top squark pair production with decays of each top squark into a top quark and the LSP (T2tt), gluino pair production with three-body decay of the gluinos into a top quark-antiquark pair and the LSP (T1 models), and gluino pair production followed by the decay of the gluinos into a top quark and a top squark, followed by the decay of the top squark to a top quark and the LSP (T5 models). The signal topologies will be discussed again in the chapter 4.

Now, we have a motivated SUSY strategy and an effective interpretation method. Before presenting the analysis, we need to have a detour to discuss the experimental instruments and the basic physics object definitions. This is presented in the next chapter.

Chapter 3

The LHC and CMS experiment

3.1 The LHC

The Large Hadron Collider (LHC) is the largest and most powerful particle accelerator in the world. The LHC was installed in an existing 26.7 km tunnel that was constructed for the LEP machine between 1984 and 1989. The LHC tunnel has 8 straight sections and 8 arcs, and is located between 45m and 170m below the surface. The LHC hosts 4 experiments currently: CMS (Point 5), ATLAS (Point 1), ALICE (Point 2) and LHCb (Point 8). The LHC can accelerate the proton beam to 7 TeV (running at 6.5 TeV during the LHC Run 2, the data used for this thesis).

However, the LHC is not the only CERN accelerator needed to boost the proton to such a high energy. The LHC is the last and biggest accelerator in the chain. A simplified injection chain is shown in Fig 3.1. Hydrogen gas is injected into a metal cylinder, and then placed in a strong electrical field to strip the electron from hydrogen to obtain bare protons. The protons are accelerated to 90 keV with a DC voltage power supply. Then

a Radio Frequency Quadrupole (QRF) focuses and boosts the protons to 750 keV. After that, the proton beam is injected into a linear accelerator (LINAC2) and accelerated to 50 MeV. The LINAC2 will be replaced by the LINAC4 with a negative hydrogen ion source (H^-), and a higher beam intensity and energy (160 MeV) in 2019-2020.

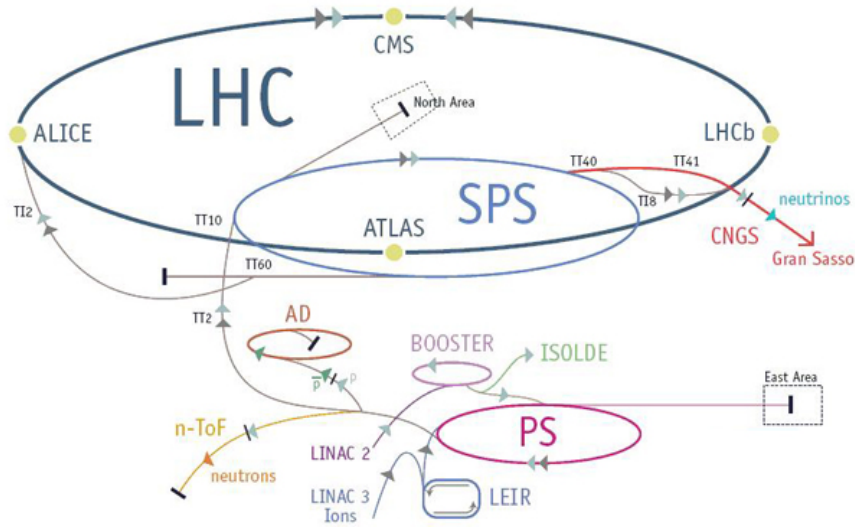


Figure 3.1: The LHC full injection chain.

The proton beam is boosted to 6.5 TeV with four circular accelerators from the linear accelerator. The first one in the chain is the proton synchrotron booster (PSB, Fig 3.2), a four-ring slow-cycling synchrotron. The PSB was inserted in between the LINAC and proton synchrotron in 1972 to increase the beam intensity. The PSB increases the proton energy up to 1.4 GeV in 530 ms and then injects the beam into the proton synchrotron (PS). The protons are accelerated to 25 GeV inside the PS. The super proton synchrotron (SPS) takes the beam from the PS and boosts it to 450 GeV in 4.3 seconds. And finally, the

protons are injected into the LHC and are accelerated up to 6.5 TeV in 25 minutes during so-called beam ramp-up.

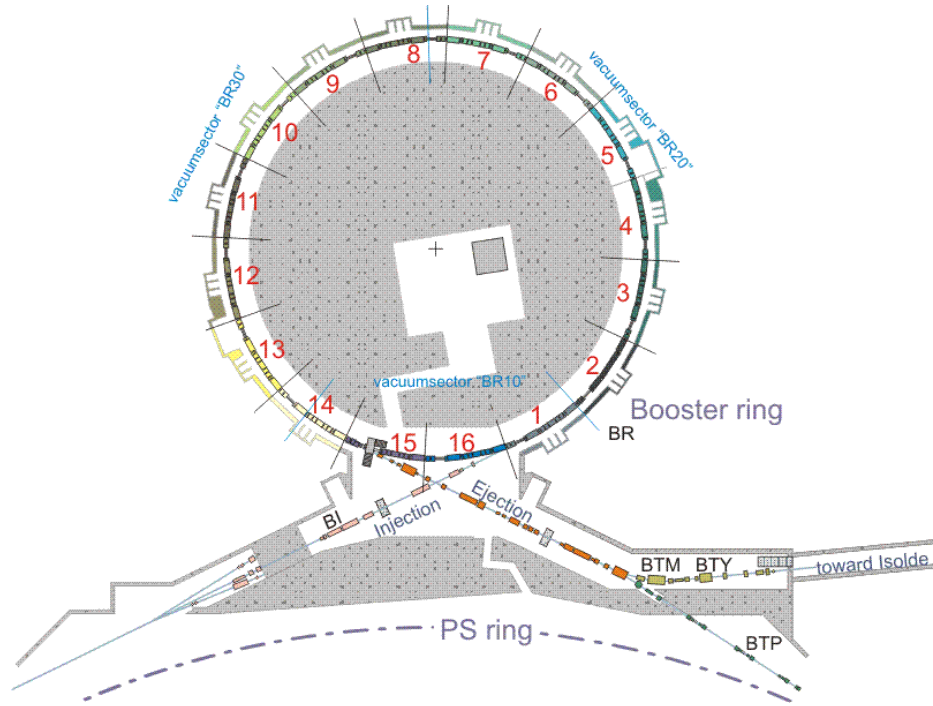


Figure 3.2: The PS layout.

We can perform a simple calculation to estimate how long it takes protons to travel from the LINAC2 until they have 450 GeV of energy in the SPS: $0.53+1.025+4.3=5.86$ seconds. However, we also split one fill of the beam into several bunches and batches with certain time spacing to obtain a desired luminosity. Therefore, we have latency for different batch and the 5.86 seconds is just a lower limit for pre-LHC acceleration time.

The LHC fill scheme can be different for various purposes. A common fill scheme has 2808 bunches in the LHC ring. This fill scheme is illustrated in the Fig 3.3. The SPS will have 12 injections to the LHC. The number of 72-bunch batches are in the pattern of

234 334 334 334. There are different luminosities for different fill schemes. In the detector operation, it is important to know the fill scheme and expected luminosity in advance to in order to design the trigger menu.

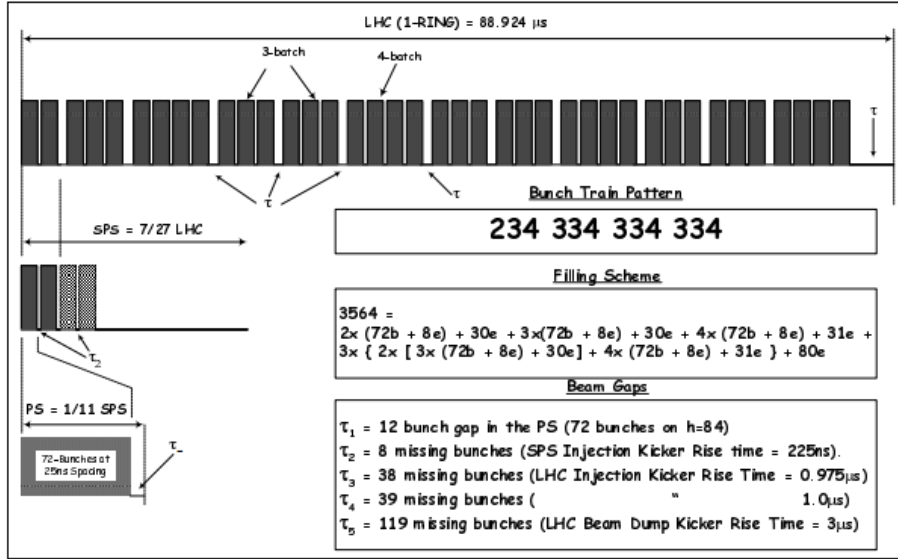


Figure 3.3: Schematic of the Bunch Disposition around an LHC Ring for the 25 ns Filling Scheme.

3.1.1 LHC: Machine layout and performance

The LHC is a nearly circular ring with 8 arcs and 8 straight sections (Fig 3.4). Each straight section is around 528 meters long and can host a detector system or infrastructure related to the accelerator. Two general purpose experiments are located at opposite straight sections: ATLAS experiment at point 1 and CMS experiment at point 5. The ALICE experiment and the clockwise circulating beam (beam1) injection system are located at

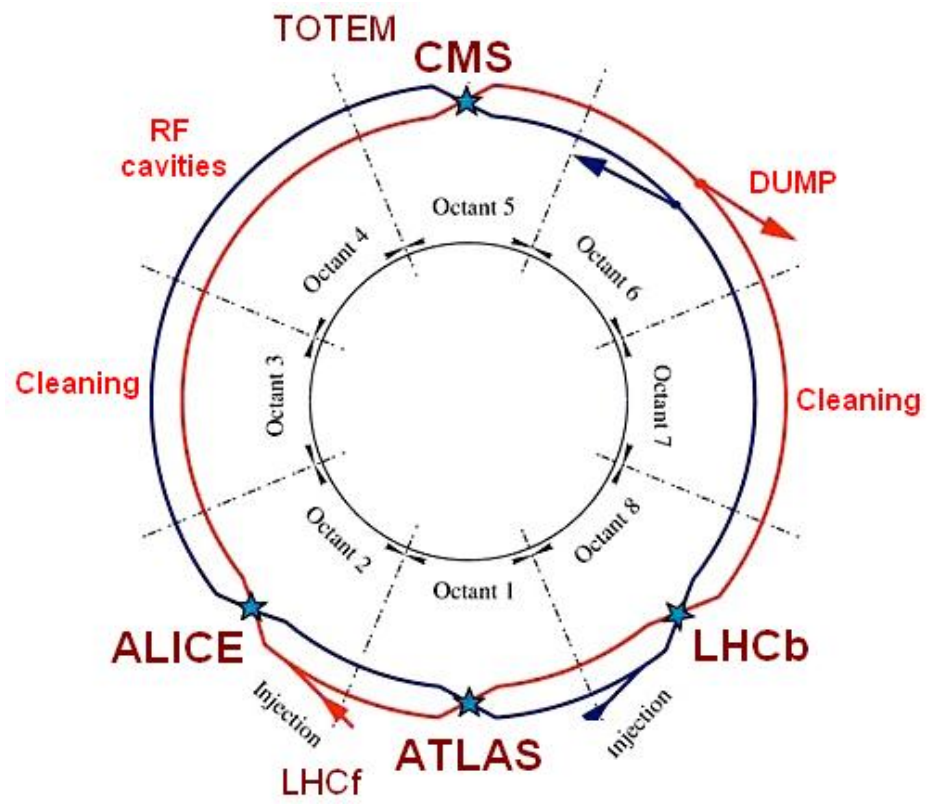


Figure 3.4: The LHC layout.

point 2, while the LHCb experiment and the counter clockwise circulating beam (beam2) injection system are at point 8.

Collimation systems [14] are located at points 3 and 7. The collimation system protects the machine against unavoidable regular or irregular beam loss. The collimator at point 3 is designed to clean beam transverse momentum dispersion and the one at point 7 is for beam betatron emission. The materials of the collimators are required to be extremely radiation hard because they are installed in the most active regions of the LHC.

The beam dump insertion is sits at the point 6. There are three steps to dump the beam: extraction, dilution and absorption. The magnet kicker can be turned on in 3000 ns, which is the time interval required for each beam to orbit around the LHC ring. Then the beam is deflected at an angle of 0.27 mrad. The extracted beam is swept in a quasi-circular orbit by two sets of deflecting dilution kickers. Finally, a 7m long segmented carbon cylinder absorbs the beam.

Point 4 is the only straight section that contains acceleration systems. Two radio frequency systems are installed for both beam1 and beam2.

One of the main aims of the LHC is to detect physics beyond the standard model. The study of the rare events in the LHC requires both high beam energies and high beam intensities. The energy of the beam is limited by the machine geometry. The luminosity of the machine can be described by Eq 3.1 with the Gaussian beam distribution. The N_b term is the number of particles per bunch and n_b is the number of bunches per beam. This indicates that we can increase the luminosity through optimization of the fill scheme. The f_{rev} term is the revolution frequency, γ_r is the relativistic gamma factor, ε_n is the normalized

beam emittance and β^* is the impact parameter at the collision point. The F term is the geometric luminosity reduction factor due to the crossing angle at the interaction point.

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (3.1)$$

The definition of F is given in Eq 3.2. The θ_c term is the full crossing angle at the interaction point, the σ_z is the RMS bunch length and σ^* is the transverse RMS beam size at the interaction point. The LHC experts are trying to reduce the crossing angle from 370 to 280 mrad, which will increase the peak luminosity by 15 %.

$$F = \left(1 + \frac{\theta_c \sigma_z}{2\sigma^*}\right)^{-1/2} \quad (3.2)$$

In the CMS and ATLAS experiments, the design peak luminosity at the collision point is approximately $10^{34} \text{cm}^{-2} \text{s}^{-1}$. The LHC achieved this goal in 2016. Attempts are now being made to increase to increase the luminosity through further optimization.

3.1.2 LHC: From an operational point of view

The LHC is an extremely complex machine and it is almost impossible to grasp every detail. However, the LHC provides a summary of status for its operational activities, which are very useful in for detector operation.

3.1.2.1 Accelerator mode

The accelerator mode indicator provides a summary of the status of the LHC machine. All the accelerator modes are listed in Table 3.1. The accelerator modes can be divided into two groups by the presence or not of the beam. For example, ACCESS mode

means that the LHC group or at least one of the detector system groups needs access to the machine area, with the beams off, to investigate a problem. The beam must be dumped or already be off to allow this access period. Another example is the BEAM SETUP mode, which means the beam is under preparation. The beam modes are described in the following section. The detector systems (e.g., CMS) make daily operational decisions depending on the accelerator modes and beam modes provided by the LHC.

Table 3.1: Accelerator Mode.

Mode Name	Description	Beam exist
SHUTDOWN	Machine not running	NO BEAM
COOLDOWN	Machine comes back from shutdown, cryogenics related activities going on	NO BEAM
MACHINE CHECKOUT	Checking out LHC subsystems	NO BEAM
ACCESS	Access going on	NO BEAM
MACHINE TEST	Operation tests without beam	NO BEAM
CALIBRATION	Power converter calibration	NO BEAM
WARM-UP	Sectors warm up for repair	NO BEAM
RECOVERY	Quench recovery	NO BEAM
SECTOR DEPENDENT	Sector activities going on	NO BEAM
BEAM SETUP	Machine setup with 1 or 2 beams, usually a signal of next physics fill when taking data	BEAM
PROTON PHYSICS	Beam on for proton physics	BEAM
ION PHYSICS	Beam on for ion physics	BEAM
TOTEM PHYSICS	Beam on for TOTEM physics	BEAM
MACHINE DEVELOPMENT	Beam on machine development	BEAM

3.1.2.2 Beam mode

The beam modes are shown in the Table 3.2. The beam modes describe standard injection procedures and accidental abort issues of the two beams in the LHC. A standard

successful injection for physics should have the following beam modes in sequence during the BEAM SETUP accelerator mode:

- BEAM SETUP: The SPS is injecting beam bunches into the transfer line. The beam is not circulating yet in the LHC.
- INJECTION PROBE BEAM: A test beam with relatively small intensity is being injected into the LHC from the transfer line. Although we have a machine protection system to shield the LHC from the beam, we still need to make sure the system is safe for beam circulation.
- INJECTION SETUP BEAM and INJECTION PHYSICS BEAM: The LHC is measuring the beam properties. Then a full intensity beam will be injected for physics data taking.
- PRERAMP and RAMP: The LHC is ramping up the beam energy up with the radio frequency system.
- FLAT TOP and SQUEEZE: The LHC is checking the system. Then the beam will be focused to reduce the impact parameter.
- ADJUST and STABLE BEAM: The LHC is adjusting the beams before collision. Then we can begin to enjoy the data run.

The beam modes are used to estimate the time remaining before the beginning of the subsequent data run. For example, CMS requires all the subsystem to be ready for data taking when the LHC declares injection physics beam. Some of the subsystems need

Table 3.2: Beam Mode.

Mode Name	Description
SETUP	Beam in transfer line, but not in the ring
ABORT	Recovery mode following beam drop
INJECTION PROBE BEAM	Ring is injected with test beam for safe circulating
INJECTION SETUP BEAM	Beam measurement going on after probe beam but before injection physics beam
INJECTION PHYSICS BEAM	Beam for physics is injected in the ring
PRERAMP	Injection done, prepare for ramp
RAMP	Ramp up the beam energy
FLAT TOP	Ramp done, pre-squeeze checks
SQUEEZE	Squeezing the beam size
ADJUST	Preparing for collision or after collision
STABLE BEAMS	Stable collision, detector should taking data
UNSTABLE BEAMS	Unstable beam because of sudden beam degradation
BEAM DUMP WARNING	Beam dump warning in case of emergency beam dump
BEAM DUMP	End of physics collision
RAMP DOWN	Ramp down beam energy after programmed dump
CYCLING	Pre-cycle before injection following access, recovery, etc
NO BEAM	No beams exist

to do the alignment or calibration during this period. The time estimates allowed through use of the LHC beam status is important for smooth data taking.

3.2 The CMS experiment

The CMS Experiment is a particle physics experiment at the LHC. It consists of the CMS detector system and event reconstruction, supported by the detector operation team, the computing/data storage team and , the software and event reconstruction team, the detector performance groups (DPGs), the physics objects groups (POGs), the physics analysis groups (PAGs), the Publications Committee, the Conference Committee, the Authorship Committee, the spokesperson and deputies, and many other people. In total, there are around 2500 individuals listed on a typical physics publication, and around 1000 more who contribute to the operation of the experiment.

3.2.1 CMS detector system

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors at the LHC, along with the ATLAS detector. To fulfill its “general purpose”, CMS is designed as a combination of several subsystems: Silicon Pixels and Strips for tracking information, Electromagnetic and Hadron Calorimeters for “light” particle energy deposition and drift tubes and cathode strip/resistive plate chambers for muon detection. The locations of these subdetectors in the CMS are not random: The silicon tracking subsystems are closest to the collision point. The calorimeters come after the silicon subdetectors, because the calorimeter absorbs all SM particles except for muons and neutrinos. Finally the muon

system is the outermost detector. To obtain high performance, CMS is immersed in a 4-T magnetic field, which is provided by a super-conducting solenoid (“S” in CMS). The magnet contributes a large fraction of the CMS total weight: 12500 tons out of 14000. However, compared with its heavy weight, the size of the CMS is relatively small: only 5000 m³. In comparison, the ATLAS detector weighs 7000 tons and encompasses a volume of 22500 m³. This explains the name “Compact” placed in front of “Muon Solenoid” since its density is nearly to 3000 kg/m³.

3.2.1.1 Inner tracking system

As indicated by its name — inner tracking system, there are two main characteristics of this CMS sub-detector: it is the closest sub-detector to the LHC beam and its purpose is to reconstruct the trajectories of charged particles (“tracks”) that traverse it. The main challenge is a consequence of the first feature: the inner tracking systems have to be radiation hard so that they can survive over years of operation in the intense radiation environment near the beam lines. Moreover, since the LHC has high intensity and a small time interval between bunch crossings (25 nano seconds), we also need the inner tracking systems to have a fine granularity and fast response. The silicon based technology detector is the best option to satisfy these requirements. On the other hand, the second feature actually comes from the physics requirement. We need to reconstruct the tracks of charged particles and secondary vertices in the event from the three dimensional hits. The track information is not only used in the charged particle reconstruction, but also applied in the particle flow algorithm [2] as a basis of all physics object reconstruction in CMS.

The secondary vertices are also important in the heavy flavor object reconstruction, and in new-physics search of long-lived particles.

As a result of budget-performance balance, the CMS inner tracking systems are designed as a combination of two sub-systems: the silicon pixel detector and the silicon strip detector. A schematic drawing of the CMS tracking system is shown in Fig 3.5. More details are given in the following sections.

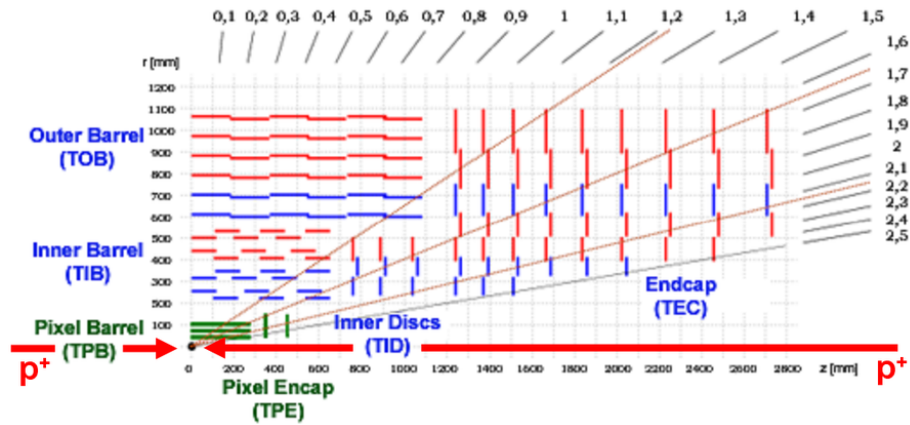


Figure 3.5: Two dimensional CMS inner tracking system layout, phase 0.

3.2.1.1.1 Silicon pixels The pixel system is the part of inner tracking system that is closest to the collision point. It provides a precise measurement of the tracking and therefore makes a major contribution to the secondary vertex reconstruction. The pixel cells are distributed with a size of $100 \mu\text{m} \times 150 \mu\text{m}$ in three-dimensional space, which allows a 3D reconstruction method for the secondary vertex. The signals are read by sensors and packed by readout chip. Then the signal is delivered to the front-end driver through the analog chain once a positive bit is received from the L1 trigger.

The CMS detector system has three phases, corresponding to different stages of the LHC. Phase 0 CMS is the baseline detectors for the LHC with 7 or 8 TeV and integrated luminosity around 25 fb^{-1} . Phase 1 CMS is the first detector upgrades for 13 or 14 TeV and integrated luminosity around 300 fb^{-1} . Phase 2 CMS is the major upgrade for high-lumi LHC (HL-LHC) with 3000 fb^{-1} . The phase 0 pixel detector contains three barrel layers (BPix) and two endcap disks (FPix), which cover a pseudorapidity range $-2.5 < \eta < 2.5$. The pseudorapidity η is a spatial coordinate describing the angle angle of a particle relative to beam axis. It is defined as $\eta = -\ln[\tan(\frac{\theta}{2})]$, where θ is the angle between the particle three-momentum and beam direction. In the phase 1 upgrade [48], because of radiation damage, all layers and disks were replaced with new ones, and an additional layer was added both in the barrel region and in both endcap regions.

3.2.1.1.2 Silicon strips The particles pass through ten layers of silicon strip detectors after the pixels. The silicon strip tracker is composed of 15148 detector modules. Each module is composed of either one $320 \mu\text{m}$ or two $500 \mu\text{m}$ silicon sensors from a total of 24244 sensors. The signal from a sensor is amplified, shaped and stored by a custom integrated circuit. Once a positive L1 trigger decision is received, the analog signal is multiplexed and transmitted via an optical link to the front end driver, then digitalized.

The inner strip detector contains 4 barrel layers (TIB) and 3 endcap disks (TID) on each side. The outer strip detector has 6 barrel layers (TOB, 2 double-sided, 4 single-sided) and the endcap strip has 9 disks (TEC) on each side.

A laser alignment system [65] is used to monitor the stability and the alignment of the strip detector mechanical structure. An infrared laser light is delivered directly to the

sensor on the 434 silicon modules (3%) to trigger a signal pulse. The alignment data can be taken during commissioning, an inter-fill period or in the orbit gap with stable beams. The pixel detector is aligned with strip detector using the reconstructed tracks, and therefore the strip alignment is the only online and direct alignment needed for the inner tracking system.

3.2.1.2 Calorimeters

3.2.1.2.1 Electromagnetic calorimeter The CMS electromagnetic calorimeter (ECAL) is a hermetic homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals. Lead tungstate is an appropriate choice for the ECAL because the crystal is both radiation hard and has a rapid response (same level as the LHC bunch crossing). In total 61200 lead tungstate crystals are installed in the ECAL barrel region while 7324 are installed in each endcap. Lead tungstate emits 80% of the light in 25 ns once an electron (or photon) strikes on the crystal.

A photon detector is needed to collect the light yield from the crystals. As for the crystal, the photon detector needs to be both radiation tolerant and fast. To satisfy these requirements, avalanche photodiodes (APD) are installed in barrel and the vacuum phototriodes in the endcap. All the photon detectors have been tested in a harsh environment (high radiation plus 4T magnetic field) before actually being installed. The light yield from the crystal is weak, so it needs to be amplified with a photomultiplier. Multi-gain pre-amplifiers (MGPA), and an ASIC (Application-specific integrated circuit) specifically designed for ECAL, are installed to reshape and amplify the signal from photon detectors.

The amplified signals are integrated and digitized by on-detector front end electronics and then sent to the central data acquisition system from the off-detector back end electronics. The ECAL electronics, especially for the off-detector electronics, are very similar to the phase 0 HCAL off-detector electronics. The HCAL electronics are discussed in the next section.

A last component of ECAL is the preshower system. The two photons from neutral pion decay are almost collinear when the pion energy is high. The main purpose of the preshower detector is to trigger an electromagnetic shower with high spatial resolution before the ECAL endcap. As a result, the almost collinear photons can be distinguished by the ECAL reconstruction algorithm.

3.2.1.2.2 Hadron calorimeter The CMS HCAL is a set of detectors that measure the energy of hadronic particles, i.e., particles composed of quarks and gluons. It contains four parts: Hadron Barrel (HB), Hadron Endcap (HE), Hadron Outer (HO) and Hadron Forward (HF). The overall geometry scheme is indicated in Fig 3.6. The CMS HCAL is a sampling calorimeter with brass absorber to trigger a shower and scintillator for the energy measurement. We take advantage of the dense absorber to fit the HCAL in the limited space inside the solenoid (Except for HO). However, the HCAL suffers from a relatively large fluctuations in energy deposit, for a given energy of an incident hadron, because of the invisible energy loss in the absorber and uncompensated design of the calorimeter. The following paragraphs provide more description of HB, HE, HO and HF.

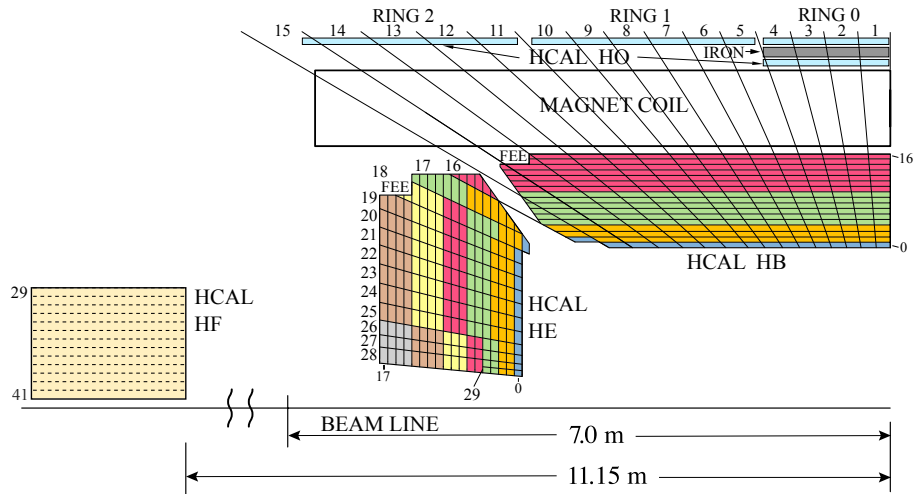


Figure 3.6: Phase 1 HCAL tower segmentation in the r,z plane.

The HB and HE are the major components of the HCAL. They are typical sampling calorimeters with brass from recycled Russian naval shell casings and scintillator for 70000 tiles. Lights from the scintillators is collected and amplified by hybrid photodiodes (HPDs) in the phase 0 HCAL. The HPDs will be replaced with Silicon photomultipliers (SiPMs) for Phase 1 because the SiPMs are less noisy and more stable in heavy radiated area.

After the photon detectors, the analog signal is delivered to the FrontEnd electronics. The charge pulse is integrated and digitized on the charge integrator and encoder (QIE). In the phase 1 upgrade, we will replace the QIE8 with QIE11 for all frontend readout modules in HE. The QIE11 have a much better time resolution, which is on the level of 0.5 ns.

After the FrontEnd, the digitalized signal is delivered from the commission cavern to the service cavern through a long bonded fibers bundle. The backend electronics located in the service cavern receives those signals, generates the trigger primitives and delivers the

signals to the central DAQ link. The HCAL backend electronics were upgraded from VME to uTCA on HB and HE during the 2015-2016 year-end technical stop (YETS) [55].

However, due to the limitations of the “compact” cavern, the EB and HB cannot provide complete containment for the hadronic showers. The HO is inserted just between the solenoid and muon system to ensure adequate sampling depth in the barrel region. The solenoid coil is used as the absorber for HO, which measures the shower deposit after HB.

The HO detector construction is the same as the phase 0 HB and HE, except for the photon detectors. The HO has the SiPMs instead of HPDs since the LHC long shut down 1. The SiPMs have proved to be very reliable in the current operation except for a small leakage current drift monitoring issue at the beginning of LHC Run 2.

The last component of HCAL is the Forward calorimeter (HF). The calorimeter design is severely challenged by the extremely high radiation environment. Therefore, quartz fibers are used as the active medium for HF. When a shower occurs, Cherenkov light is emitted in the quartz fiber and transported to photomultiplier tubes (PMT). The HF also provides input to the CMS luminosity measurement.

In summary, there are four subdetectors in HCAL: HB, HE, HO and HF. There are 9216 readout channels for HB, 6912 for HE, 2376 for HO and 3456 for HF in the HCAL phase 1 scheme. The backend electronics coordinates are compressed into the RAW event data after the event builder in the HLT. On one hand, the offline reconstruction software and online data quality monitor need to know the detector geometry coordinates in order to map the digitalized output into physical space. We need to build a database object to map the backend coordinates and detector coordinates. On the other hand, the backend

electronics are not connected directly to the HCAL tiles: we still have photomultiplier and frontend electronics in between. HCAL has several different upgrade projects for different parts of different sub detectors during the LHC Run 2. Therefore, we need a dynamic map, which reflects the real connections among all HCAL components. This map is called the HCAL logical Map. The logical map starts with frontend coordinates, maps upwards to the photomultipliers, and then detector tiles, and downwards to the backend coordinates and trigger primitive channels. We avoid re-designing everything after the each upgrade with this dynamic structure of the map.

To build a robust HCAL logical map, we need inputs from experts in different fields. For example, we need to follow a special symmetric design in the frontend to backend optical patch. This symmetry is highly dependent on the firmware design. Although nowadays FPGA (Field Programmable Gate Array) can be programmed with less restriction, we still need to keep the algorithm maintainable. We also plot the backend and detector coordinates respectively in frontend coordinates in order to check the symmetry in the mapping (Fig 3.7 and Fig 3.8).

Subsets of the HCAL logical map are critical in the HCAL software. One of the most important applications is the electronics map (EMap). The HCAL electronics map is a subset of the logical map that contains backend coordinates and detector coordinates. As was mentioned, the reconstruction software needs the EMap database to obtain the detector coordinates from the backend coordinates. This is a necessary step in the RAW to DIGI reconstruction and in the DIGI to RAW conversion in MC simulation. Another application is the QIE calibration table. In the DIGI to RECO step, we need to translate from ADC

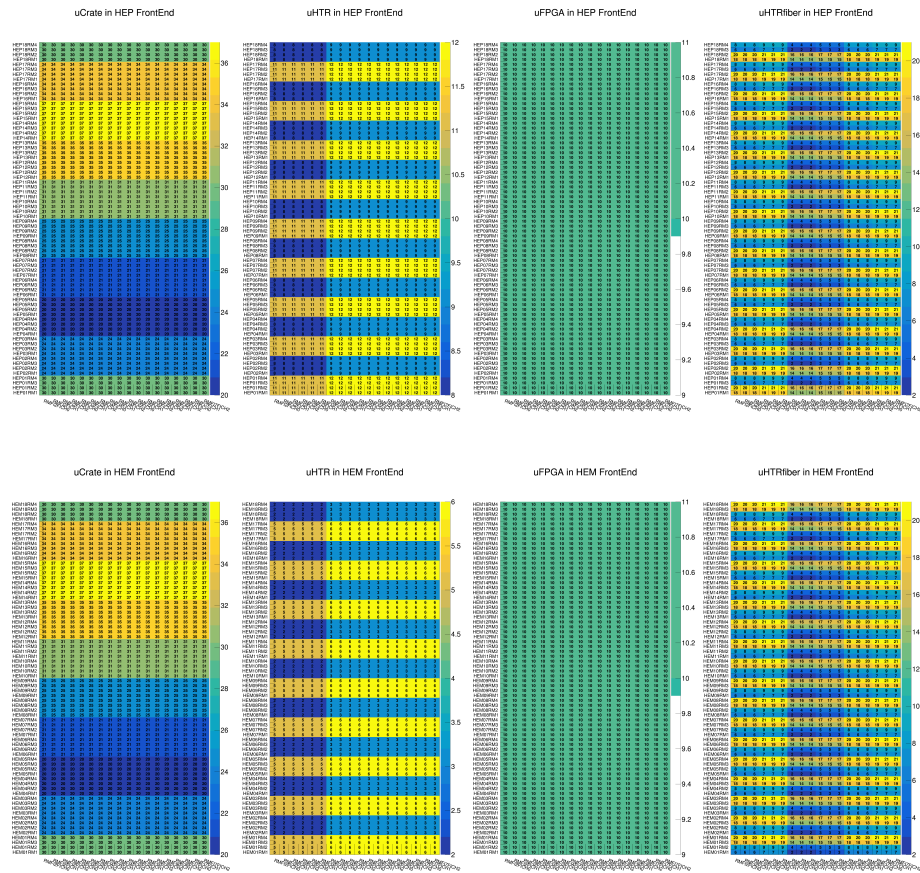


Figure 3.7: Backend coordinates in Frontend coordinates, HCAL endcap phase 1 backend plus phase 0 frontend in 2016 operation.

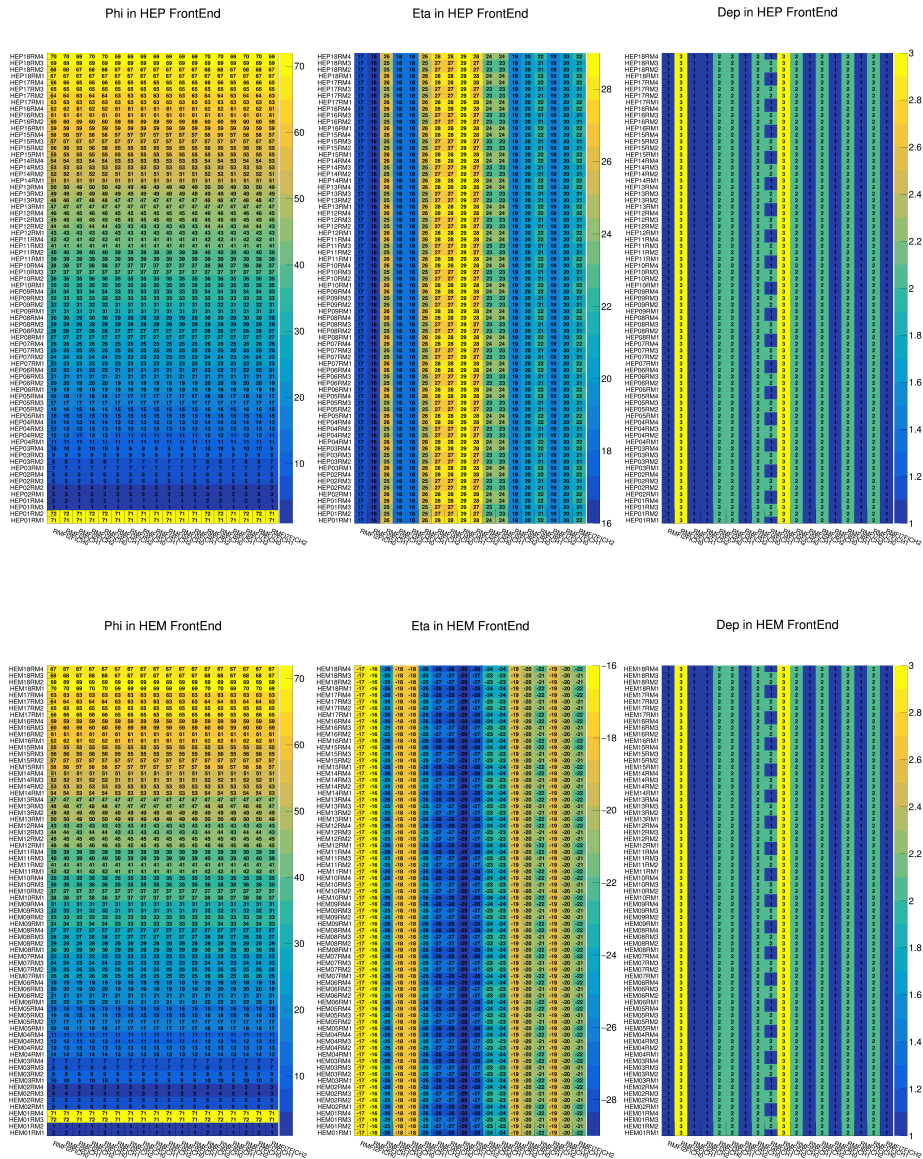


Figure 3.8: Detector coordinates in Frontend coordinates, HCAL endcap phase 1 backend plus phase 0 frontend in 2016 operation.

(Analog Digital count) to the fC with the piecewise linear QIE gain function. This map is produced from the logical map. The robustness of the logical map is critical for the HCAL offline reconstruction.

3.2.1.3 Muon detectors

The importance of the muon detectors is implied by the experiments middle name (“M” in CMS). The CMS muon detectors provide measurements of muon track coordinates. The measurement of muons is important in both in standard model physics (e.g. Higgs to ZZ) and in new physics searches (e.g. searches for supersymmetric particles that can decay to leptons). Moreover, the muon system provides a muon-related level-1 trigger that is used to reduce the data rate.

The CMS muon detectors consist of three sub-detectors: Drift tubes in the barrel region, cathode strip chambers in the endcap and resistive plate chambers in both the barrel and endcap. A layout of CMS muon detectors is shown in Fig 3.9. More details are given in the following sections.

3.2.1.3.1 Drift tubes The drift tubes (DT) are the part of CMS muon detectors that measure the muon tracks in the barrel region. The basic unit of the drift tubes is the drift cell. A drift cell is a 42mm-wide tube containing a thin conductive wire at high positive voltage within a gas volume. The muon knocks electrons off the atom of the gas when it traverses the chamber. The muon tracks can be reconstructed by measuring the drift time for different cells.

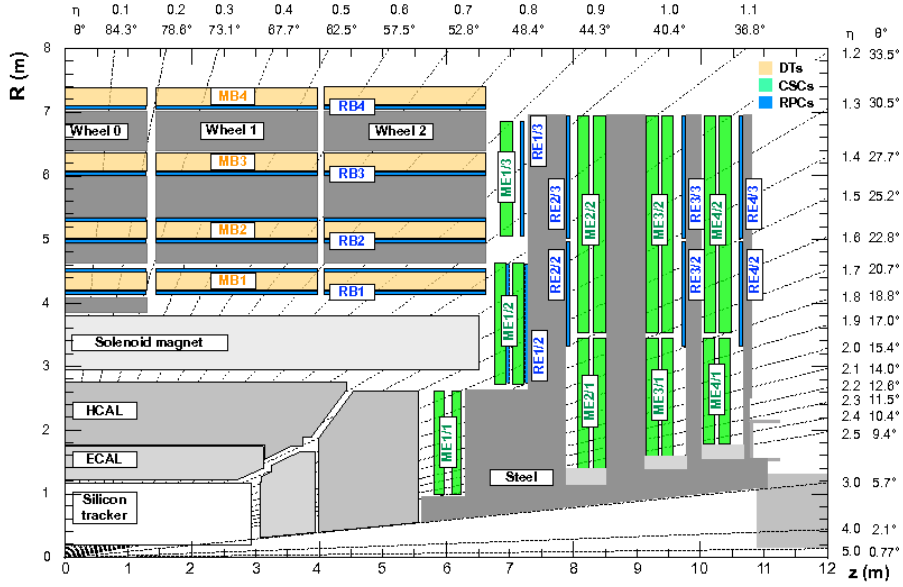


Figure 3.9: Two dimensional CMS inner tracking system layout, phase 0.

Like the HCAL outer detector, the drift tubes contain 5 wheels along the z -axis, 4 stations for each wheel, labeled MB1, MB2, MB3 and MB4. The three inner stations have 60 drift chambers each and the outer chamber has 70. A drift chamber consists of 2 or 3 super layers, each made of by 4 layers of drift cells staggered by half a cell. This honeycomb geometry gives an excellent time-tagging capability, with a time resolution of a few nanoseconds. The full DT provides a pseudo-rapidity coverage of $0 < |\eta| < 1.2$.

3.2.1.3.2 Cathode strip chambers The cathode strip chambers (CSC) are the CMS muon detectors located in the endcap region. The CSC is also a type of wire chamber, but differently from the DT, the CSC measures the location (to be specific, phi coordinates) instead of drift time. The CSC consists of arrays of positively charged anode wires crossed with negatively charged copper cathode strip within a gas volume. The muon knocks

electrons off from the gas atom. Then electrons move to the anode wire to create an avalanche. Positive ions also move to the cathode and trigger a charged pulse.

The CMS endcap muon system consists of 468 cathode strip chambers in the following arrangement: 72 ME1/1, 72 ME1/2, 72 ME1/3, 36 ME2/1, 72 ME2/2, 36 ME3/1, 72 ME3/2, and 36 ME4/1. The full CSC provides a pseudo-rapidity coverage of $1.2 < |\eta| < 2.4$.

The RPC is designed as a fast response provider for the trigger. However, for the endcap region, the CSCs by themselves are sufficient for the trigger requirements for the current instantaneous luminosity of the LHC. Moreover, the CSCs have a better spatial resolution and a larger pseudo-rapidity coverage, providing better precision for the measurements of endcap muons.

3.2.1.3.3 Resistive plate chambers Resistive plate chambers (RPC) are fast gaseous detectors that provide a muon trigger system in parallel with the DT and CSC. The RPC consists of a two high resistively plastic parallel plates with gas between them. One of the plates is a positively charged anode and the other is a negatively charged cathode. Like the CSC, electrons of the gas molecules will be knocked off when a muon passes through, and trigger an avalanche. The pattern of hits from the cathodes provides a quick measurement of the muon momentum, which is sent to the trigger for decision-making. In the CMS RPC, the double-gap modules are used instead of the single-gap modules. This allows a lower bias voltage for each single-gap and higher detector efficiency. The RPC has both good spatial resolution and time resolution (one nanosecond).

The CMS RPCs are distributed in both the barrel and endcap regions. There are 96 RPCs in each wheel in the barrel region. More details on the distribution are given in Table 3.3. In the endcap region, there are 3 RPC stations in the phase 0 design. In order to retain high muon reconstruction efficiency with Run 2 condition, a fourth disk was installed during the long shutdown 1 as part of the phase 1 upgrade. The full RPC has the same pseudorapidity coverage as the DT in barrel, and a smaller coverage ($1.2 < |\eta| < 2.1$) than the CSC in the endcap.

Table 3.3: Numbers of RPCs for different wheels.

RPC	W+2	W+1	W0	W-1	W-2	Total
RB1(in)	12	12	12	12	12	60
RB1(out)	12	12	12	12	12	60
RB2/2(in)	12	-	-	-	12	24
RB2/2(out)	-	12	12	12	-	36
RB2/3(in)	-	12	12	12	-	36
RB2/3(out)	12	-	-	-	12	24
RB3	24	24	24	24	24	120
RB4	24	24	24	24	24	120
Total	96	96	96	96	96	480

3.2.1.4 Trigger system

The LHC has a bunch crossing every 25 ns, which means an intrinsic data rate of 40 MHz. It is impossible for the data acquisition and storage system to deal with such a high rate, and it is not necessary either, since not all the events are interesting for physics analysis. Therefore, we need a system to reduce the data rate and select the events that we are interested in. The system is called trigger system — only triggered data will be read and stored from the LHC collision.

The CMS trigger system is a 2-level trigger system: the FPGA based L1T (Level 1 trigger) and PCs farm based HLT (High level trigger). In the old days (e.g. Tevatron, CDF trigger) there was a custom hardware L2T layer between the L1T and HLT to ease the computing speed gap. However, with improvement to the computing power of the PCs farms, we can avoid the additional L2T. This gives a more robust system with a less complicated structure.

The CMS L1T system receives part of its information from the sub detectors and performs a rough reconstruction of physics object at the FPGA level. The data rate is reduced from 40 MHz to 100 kHz (85 kHz to 90 kHz in real operation, mainly limited by the data acquisition system) after the L1T. The CMS L1T was upgraded (Phase 1) during the 2015-2016 YETS and 2016-2017 EYETS to account for the large increase in the LHC beam intensity in Run 2. The scheme for the phase 1 L1T is shown in Fig 3.10. The calorimeters send the trigger primitive to the calo trigger layer from where the signals are sent to the global trigger. The muon system sends information to the global muon trigger from where they are sent to the global trigger. The L1A (level 1 acceptance) is generated and sent to the sub detectors through TCDS (timing, control and distribution system). The sub detectors send the information to the data acquisition or keep on taking data for next bunch crossing based on the L1 decision.

The CMS HLT system uses all the information from the detector readout and applies a simplified algorithm to reconstruct physics objects. Promptly reconstructed objects are used in the high-level trigger menu for various physics purposes. However, just like the L1T, the HLT has bandwidth limitations. The data rate after the HLT is around 1000

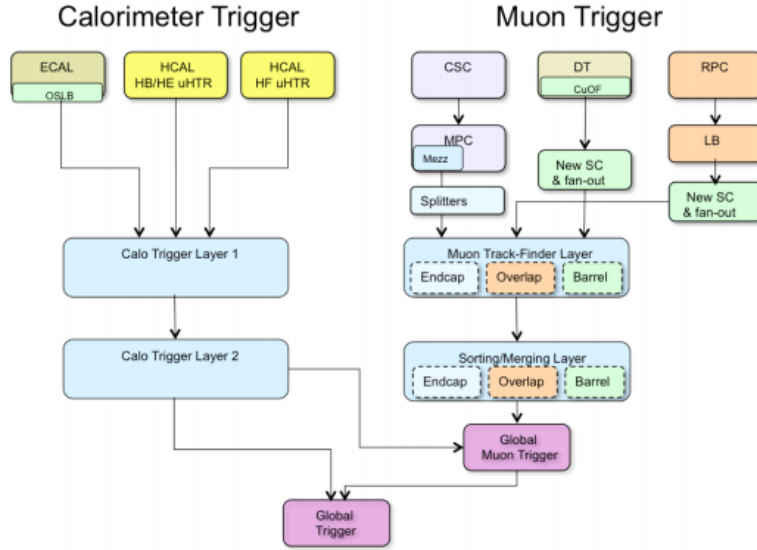


Figure 3.10: Dataflow for the overall phase 1 trigger system.

Hz which is limited mainly by offline storage rather than online computing. HLT is the interface between physics analysis and online operation. On one side, each physics analysis group has a trigger contact person to collect the desired trigger menu from analyzers and make a proposal for a full menu set with rate estimation to the HLT that satisfies the data rate limit. On the other side, each HLT menu takes at least one L1 seed as a starting point of the high-level trigger.

The current (Phase 1) trigger system will continue in operation until the end of LHC Run 3. However, there will be a huge challenge on the trigger system with the HL-LHC. The Track trigger will be necessary to maintain the current physics acceptance with the L1 rate around 1 MHz.

3.2.2 Event reconstruction

The data from the central data acquisition system is selected and built in the high-level trigger farm and then transferred to the primary computing grid at CERN (Tier-0). The data directly from the detector electronics is called DAQ-RAW, which is the input of the online HLT cluster. The data is reformatted and filtered with the high-level trigger. The outcome data are called RAW. Compared to the DAQ-RAW, the RAW contains the HLT information and is ready for offline reconstruction after reformatting.

Before entering the offline reconstruction chain, the data stream is filtered into different primary datasets (PD) for various purposes. Most of the categories are based on physics analysis and the level 1 trigger. For example, in the all-hadronic SUSY search, the common PD for the search region is HTMHT or MET. For the control region, we can select SingleElectron or SingleMuon PD. Then, the RAW with different PD streams is delivered into the offline reconstruction system (CERN Tier-0 or any Tier-1, Fig 3.11).

In the offline reconstruction, the RAW dataset is unpacked from the electronics based digital counts to detector based digital hits. This is the so-called RAW to DIGI step in the reconstruction. After that, the DIGI is converted to reconstructed hits with calibration information. This is called DIGI to RecHits step. The offline database is highly involved in these 2 steps, providing the electronic-detector map, calibration table, pedestal subtraction, radiation damage correction, etc. Then, the physics objects are generated with the reconstruction algorithm. The dataset after the offline reconstruction is called RECO, which combines the RecHits and physics object information together.

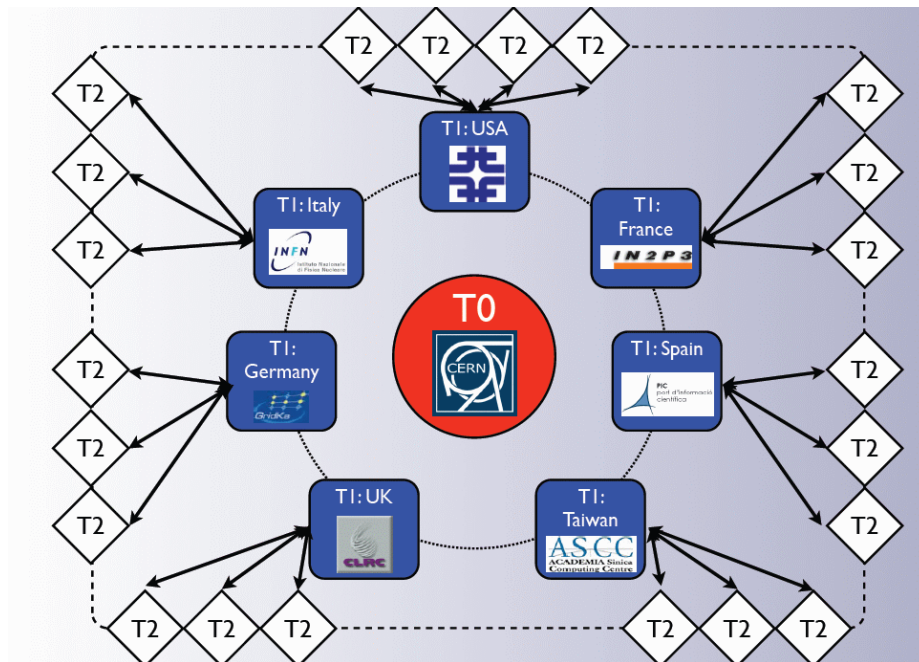


Figure 3.11: CMS offline computing structure.

However, the size of the RECO dataset is about 1.3 MB per event. Further reduction of the event size is necessary for the physics analyses. Therefore, information that is not commonly used for physics analysis is dropped to form a new data format: miniAOD (AOD stands for Analysis Objects Data). It contains the standard physics objects used by most CMS physics analysis groups. More details on those objects are discussed in the following sections.

3.2.2.1 Tracks

Tracks are detected by inner silicon detector signals and form the basis of the event reconstruction. Track reconstruction serves the following purposes:

- Vertex reconstruction: The vertexes in proton-proton collisions can be reconstructed from the intersecting point of tracks. However, there can be multiple vertexes in an event because of the pile up effect.

The reconstructed vertex with the largest value of summed physics-object p_T^2 is taken to be the primary proton-proton interaction vertex. The physics objects are the objects returned by a jet finding algorithm [24, 26] applied to all charged tracks associated with the vertex, plus the corresponding associated missing transverse momentum.

We also reconstruct secondary vertices, i.e., vertices displaced from the beam line, which can indicate the positions of the decay of long-lived particles. Secondary vertex information is very important for b-jet identification.

- Momentum measurement: The CMS inner silicon system has a high resolution in the track momentum measurement thanks to the strong magnetic field (3.8 T). The jet

p_T measurement benefits from the high precision track momentum with the particle flow algorithm. More details are covered in the jet reconstruction section.

- Particle identification: Charged particles can be identified with track information. For example, an electron candidate is found when the energy deposit in the ECAL supercluster can be associated with a track. Muon identification is performed by matching a track from the inner detector with a track segment from the muon detector.

Given the importance of track reconstruction in physics, a reliable algorithm is required. The desirable algorithm must have a near 100% efficiency in track reconstruction together with a relatively small fake rate. The jet energy can be badly mis-measured if there is an unreconstructed or fake track. There are two steps in the track reconstruction:

- Local reconstruction: the signals from the strip and pixel are clustered to evaluate the position and error matrices of hits.
- Global reconstruction: Tracks are reconstructed with several iterations of the combinatorial track finding method [8]. Different seeds are used in different iterations to improve the tracking efficiency for various types of tracks.

The track finding algorithm is very powerful. It can even find tracks with very low energy ($p_T < 2 \text{ GeV}$). Some of the low energy tracks have more hits than the number of detector layers because they spiral in the magnetic field. Other tracks do not reach the preshower detector. Those tracks are kept in the data file for low energy physics studies at CMS.

3.2.2.2 Electrons

Electron reconstruction in CMS relies on the information from the inner tracker and ECAL. The track momentum can be measured with smaller uncertainty in the inner tracker system. On the other hand, the calorimeter has a better resolution for high-energy objects. Therefore, a mixture of “ECAL seed” and “tracker seed” algorithms is designed to optimize the electron identification in both the high and low p_T spectrum.

The ECAL seed electron identification algorithm starts from the supercluster energy deposit in the ECAL. The supercluster is a 5 by 5 cluster combination of lead tungstate crystals in the ECAL. The initial energy and position of the supercluster are used to estimate the electron trajectory in the first tracker layer. In contrast, the tracker-based method begins with the tracks reconstructed by the general algorithm for charged particles. The tracks filtered using a pre-selection criteria and then matched to an ECAL supercluster. This mixture of seed method has been validated with electrons from W boson decays in 2010 [47].

The electron identification group provides several working points to satisfy the needs of various physics analysis groups. For example, we choose the “Veto” working point in the SUSY analysis described in this thesis. The “Veto” has the loosest identification criteria, which helps us to reject background events containing electrons to the maximum extent.

3.2.2.3 Jets and particle flow

A jet is a collimated flux of stable hadrons that originates from a quark or a gluon following hadronization. A jet algorithm is a method to reconstruct the jets. There are several different jet algorithm approaches, but the inputs of the jet algorithm are always the clustered energy deposits in 2-dimensional plane.

There are two major requirements for a jet algorithm: that it be collinear safe and infrared safe. Collinear safe means that the results of the jet algorithm are invariant under collinear splitting of any input momentum. Infrared safe means that a soft emission does not change the results of jet algorithm. Infrared safe is required not only from soft QCD emissions, but also from the collider point of view: the jet algorithm results should not change because of a soft jets from pile-up.

The anti- k_T algorithm [24] is the standard jet algorithm at the LHC. The anti- k_T algorithm is a sequential recombination algorithm. In this approach, we define a special distance between two input observables as in Eq 3.3:

$$d_{ij} = \min(k_{T_i}^{-2}, k_{T_j}^{-2}) \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \quad (3.3)$$

where k_{T_i} is the transverse momentum (or any other response) of the i th input. The numerator is the geometrical distance between the two objects. R is the cone size parameter in the algorithm. The CMS experiment supports $R=0.4$ for the normal jets and $R=0.8$ for fat jets. Fat jets are useful in the analysis of highly boosted objects, as will be described in top tagging section below. The special distance between an input object with respect to the beam direction is defined as: $d_{iB} = k_{T_i}^{-2}$.

The first step is to compute all the d_{ij} and d_{iB} , and find the smallest one. Then the following steps are iterated until all objects are clustered into a jet:

- If the smallest one is a d_{ij} , we combine the two objects i and j , remove objects i and j from the list of objects, and add the combined entity to the list of objects, calculating its d_{ij} and d_{iB} . The combined object is given by the sum of the four-momentum of objects i and j .
- If the smallest one is a d_{iB} , we remove object i and call it a jet.

The anti- k_T algorithm is both collinear and infrared safe. The geometrical distance between two collinear objects is zero. Therefore these two objects will be clustered first. The special distance is infinite for a soft object whose transverse momentum approaches zero. Therefore this new object will be clustered last, contributing nothing to the jet it is associated with since its transverse momentum is zero, or will form a new jet with zero transverse momentum. The jet results for the anti- k_T algorithm remain unchanged in either case.

There are other algorithms that are also both collinear and infrared safe, like the SiSCone [63] and k_T [23] algorithms. However, unlike these latter two algorithms, jets from anti- k_T algorithm always have a nearly circular jet area [25], which makes them easier to correct for detector-related effects.

The performance of jet finding in CMS is improved using the particle-flow (PF) algorithm [2] and by using PF candidates as the input to the anti- k_T jet clustering.

The particle flow algorithm is a particle reconstruction and identification method that uses the information from all sub detectors. The PF algorithm is especially important

for CMS jet and missing transverse momentum reconstruction because of the uncompensated nature of the CMS HCAL. A jet usually has three components from: charged hadrons (π^+ , π^- , etc.), photons (π^0), and neutral hadrons (K_L , etc.). The typical fractions of the three components are listed in the Table 3.4. The inner tracker can provide a precise measurement of the charged hadron fraction. The CMS ECAL provides an excellent measurement of the photon energy. The remaining 10% of the jet energy, due to the neutral hadron fraction will not give a huge uncertainty to the final measurement, on average.

Table 3.4: Average Jet components in GEANT4 Simulation.

Jet Constituent	Energy Fraction	Detection Instrument	Uncertainty
Charged Hadron	65%	Inner Tracker	Negligible
Photon	25%	ECAL	$0.07^2 E_{jet}$
Neutral Hadron	10%	ECAL and HCAL	$0.16^2 E_{jet}$

The key aspect of the PF algorithm is the linking of charged particles between the tracker and calorimeter. First, a track collection is generated, using the iterative track reconstruction method that described in previous section. Then, energy deposits in the calorimeter are clustered to obtain an initial identification of photons, charged hadrons, and neutral hadrons. A linking algorithm [2] is applied to connect the objects from the inner tracker and calorimeters.

The PF jets are still not the collection that is directly used in the physics analyses. The PF jet energies still need to be corrected. For the data, there are three corrections that are applied. The first correction accounts for pile-up and noise. This correction is related to collider and detector effects and is subtracted from the raw energy. The second correction accounts for reconstruction effects, and is evaluated using MC simulation. Finally, the jet

energy is corrected with di-jet data. More details of the jet energy corrections are given in [33]. The jet energy corrections are applied in both data and MC in the physics analysis described in this thesis.

3.2.2.4 Missing transverse momentum

The missing transverse momentum ($p_{\text{T}}^{\text{miss}}$) is the imbalance in the transverse momentum in an event, and is calculated as the negative of the vector sum of all objects in the event. The $p_{\text{T}}^{\text{miss}}$ is an important variable in supersymmetry searches that assume R-parity since the LSP is assumed to escape without detection, leading to potentially significant $p_{\text{T}}^{\text{miss}}$. The $p_{\text{T}}^{\text{miss}}$ is a powerful variable to suppress the standard model background in the search region.

However, on the reconstruction side, $p_{\text{T}}^{\text{miss}}$ is a high-level variable since all visible particles are used in its calculation. Unexpected effects from the detectors or collider issues in object reconstruction can give us the wrong $p_{\text{T}}^{\text{miss}}$. Therefore, we need to establish robust $p_{\text{T}}^{\text{miss}}$ filters before we use the $p_{\text{T}}^{\text{miss}}$, especially for the high $p_{\text{T}}^{\text{miss}}$ case. Source of abnormal $p_{\text{T}}^{\text{miss}}$ can be HCAL noise in the HPDs, beam halo, etc.

After the filters, $p_{\text{T}}^{\text{miss}}$ needs to be corrected, similarly to the jet energies. The usual way to do this in CMS is to propagate the jet energy correction to $p_{\text{T}}^{\text{miss}}$ (so-called Type 1 correction). We apply the type 1 correction in this analysis. More details on the $p_{\text{T}}^{\text{miss}}$ filters and corrections are covered in Ref. [3].

3.2.2.5 Muons

Muons are particles that can penetrate the calorimeter. They are detected when they traverse the muon detectors, outside the solenoid. Muon reconstruction is based on several approaches:

- Global Muon reconstruction: The muon tracks are reconstructed first inside the muon system (DT, RPC, CSC) with a standalone method. Then, the muon-system standalone track is matched with a tracker track by comparing the parameters of the two tracks. Finally, a global muon track is derived by fitting the hits of the two tracks using the Kalman-filter method [41]. This is so-called outside-in method since we start from the muon system, which is outermost element of the CMS detector.
- Tracker Muon reconstruction: This approach begins with the track reconstructed in the inner silicon detectors. Tracks with $p_T < 0.5 \text{ GeV}$ and momentum $p > 2.5 \text{ GeV}$ are selected as the potential muon candidate. Then, the track is extrapolated to the muon system considering the magnet field, average expected energy loss and multiple coulomb scattering effects. The track is considered to be a muon track if at least one muon hit is matched with this track. This is called inside-out method since it begins with the inner tracker system.

The global muon reconstruction method has a better momentum resolution for high p_T muons. The reason is that the inner tracker system covers only a small radial distance, limiting its ability to accurately measure the p_T of high p_T tracks, which exhibit little curvature in the magnetic field. However, the tracker muon reconstruction method

has a better performance for soft muons. This originates from the excellent low p_T track reconstruction provided by the inner tracker system.

A particle-flow based algorithm has been designed to provide good performance for both soft and hard muons. This particle flow muon reconstruction method is a combination of the outside-in and inside-out methods. In this approach, global muon tracks and tracker tracks are gathered together and then selected with special criteria. The selection criteria are adjusted depending on environment of the muon (e.g. isolation) and other detector inputs (e.g. energy deposit in calorimeter). This method is optimized to identify muons within jets. The performances of the three approaches have been studied in Ref. [29].

3.2.3 Event simulation

Event simulation is critical in particle physics analysis. The simulation of the CMS experiment can be categorized into two parts: physics process simulation and detector performance simulation. The former one is the proton-proton collision physics and the latter one is the detector response simulation.

The simulated events are also called MC (Monte Carlo). Monte Carlo is the name of a famous casino house in Monaco. High-energy physicists use MC as a reference to the random numbers in the event-generation processes. This jargon will be used in this thesis.

3.2.3.1 Physics process simulation

The physics process simulation is described in this section. The high-energy proton-proton process calculation is based on the QCD factorization theorem [34]. The QCD factorization theorem declares that the hard-scattering cross section is a convolution

product of a parton distribution function and a perturbative calculable hard scattering. Therefore, the non-perturbative aspects of QCD are incorporated into the parton distribution function. To summarize, the differential cross section of a proton-proton collision can be expressed by Eq 3.4:

$$\frac{\sigma_{pp \rightarrow X}}{d\Phi} = \sum_{\{s_i\}} \int \prod_i \frac{d^3 q_i}{(2\pi^3) 2E_i} \sum_{ab} \int dx_a dx_b (2\pi)^4 \delta^4(x_a p + x_b p' - \sum_i q_i) \quad (3.4)$$

$$f_a^{PDF}(x_a) f_b^{PDF}(x_b) \frac{1}{2\hat{s}} |M|^2(ab \rightarrow \{s_i\}; x_a, x_b, \{q_i\}) D(X(\{x_i, q_i\}; \Phi)),$$

where M is the matrix element, a calculable quantity in perturbative QCD and f^{PDF} is the parton distribution function (PDF) [20]. The PDF cannot be calculated from first principles. Therefore, the PDF needs to be parameterized and determined from data. There are several different PDF collaborations, e.g. CTEQ [51] or NNPDF [16]. In Run 2, CMS uses the CTEQ6L1 PDF set. The D is the fragmentation function, which states the probability for the outgoing parton s_i to produce the final state X through the hadronization process. The fragmentation function, like the PDFs, cannot be derived from first principles but must be determined from data.

In practice, it is not practical to calculate the cross section directly from Eq 3.4 because there are too many combinations of final-state hadrons for the given outgoing partons. Therefore, the calculation is truncated before the hadronization stage and parton shower MC simulations are used to finish the calculation. Collinear parton splitting and soft gluon emission are considered at each order of the strong coupling strength in the parton shower simulations since they are dominant processes. Then the partons from the shower are combined to form hadrons using models, such as the Lund string [13] or the cluster

model, based on the preconfinement property of QCD [12]. More details of the parton shower simulation are given in Ref. [46].

Besides particles created in the hard scattering process, particles in proton-proton collisions can be created in soft processes, such as from the remnants of the protons left after the hard scattering of a quark or gluon from the protons has occurred. A schematic underlying process together with hard process event plot is shown in Fig 3.12. We use minimum bias data collected using only very minimal trigger conditions to study the effects of underlying events. In the simulations, the properties of underlying events are tuned to describe the minimum bias data. More details are given in Ref. [40].

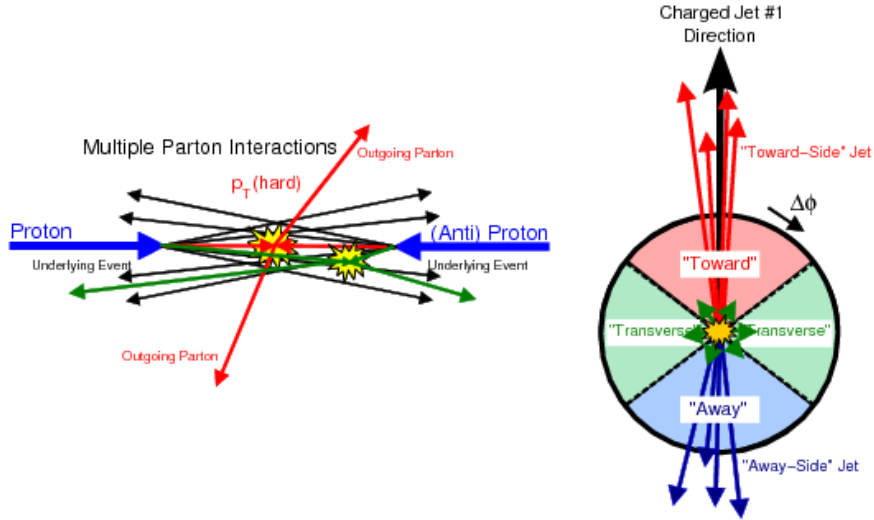


Figure 3.12: Left: Illustration of the underlying event modeled by PYTHIA, together with hard process. Right: Illustration of the correlations in the azimuthal angle relative to the direction of the leading charged jet.

In CMS, two event generators are used to simulate physics events. First, a multi-purpose parton level generator (e.g. MadGraph [10]) is used to calculate the matrix ele-

ments, incorporating the PDF information. Second, the parton shower and hadronization is described using either PYTHIA [67] or HERWIG [18]. All stable and intermediate particles are stored in the MC record for study. The stable particles are passed to detector simulation software. This latter stage is described in the next section.

3.2.3.2 Detector performance simulation

The four-vectors of the final-state particles are processed through a program to simulate the detector response. All the particles and their secondary products from electromagnetic or hadronic showers are traced and simulated from first principle in the full simulation.

The core software for this simulation is GEANT4 [9], which is a software toolkit that simulates the passage of particles through matter. The detector geometry, in this case, the CMS detector geometry [53] is the input to GEANT4. Each subdetector has a geometry input class. The corresponding general GEANT4 geometry class inherits this class. During the simulation, different particles will call various modules when they pass through the CMS detector system, depending on the type of subdetector.

However, the CPU time required for the full simulation is huge due to the complicated CMS geometry and the nature of the first principle calculations. Therefore, CMS also has a fast simulation method [62]. The aim of fast simulation is to reduce the CPU time but keep the event simulation quality at an acceptable level. The simplifications are applied in several parts. For example, the pile-up simulation is simplified, the tracker geometry is roughened, and the showers are simulated with an empirical formula rather than from first principles. The simulation process is greatly shortened because of these simplifications.

The fast simulation method is used for SUSY signal events and for some standard model processes. We can use high level object, like jets, without concern with respect to accuracy in the fast-simulated samples. However, we need to be careful concerning higher-level information (e.g. quark-gluon likelihood) in those simplified samples since the geometry is simplified. For example, in this thesis, the customized top tagger needs a fast-full simulation scale factor before the limit setting, in order to correct for differences between the fast and full detector simulation. More discussion on this is given in the next chapter.

Chapter 4

SUSY Search with a customized top tagger

4.1 Introduction

The physics motivation for this SUSY search, the experimental instruments and the reconstruction of physics objects have been discussed in the previous chapters. In this thesis, a top squark and gluino search in all-hadronic channels has been designed. The target signal models are shown in Fig. 4.1.

The first step of the search is to design a search region with proper selection criteria. Since we are looking into all-hadronic channels, leptons and isolated tracks are vetoed. Isolated tracks are vetoed because they can arise from single-prong tau lepton decays or from electrons or muons that are not identified. Minimum requirements on the number of jets and $p_{\text{T}}^{\text{miss}}$ are needed to suppress the standard model background. A set of $\Delta\phi$ cuts

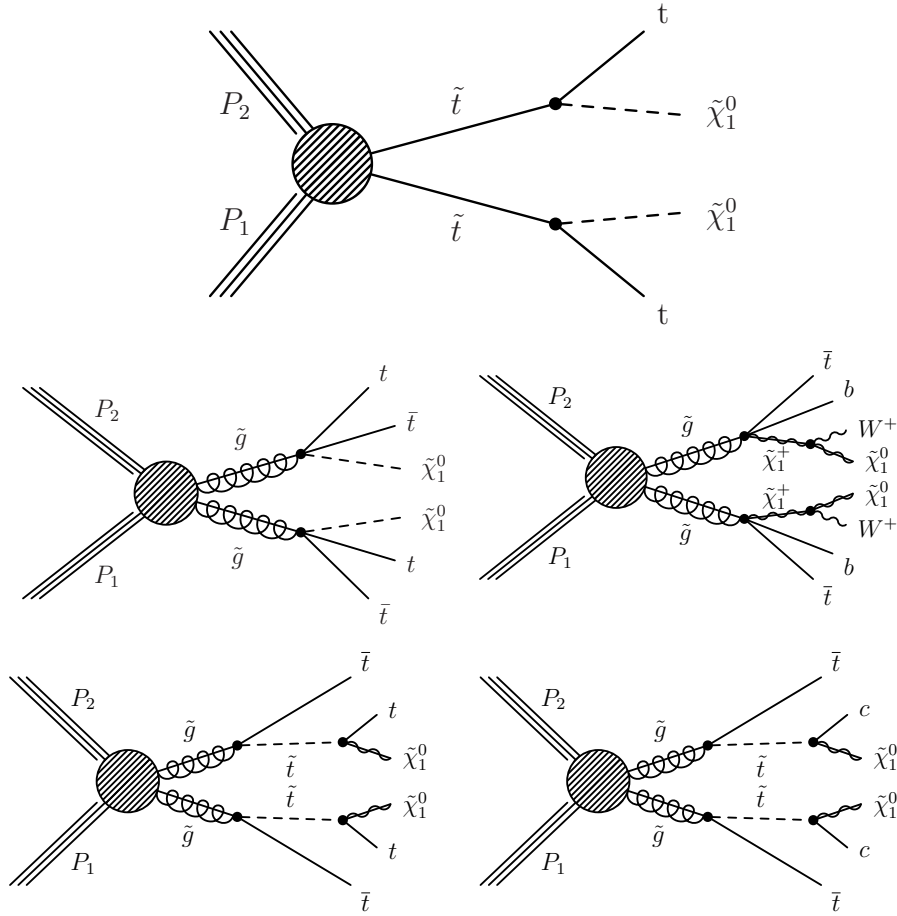


Figure 4.1: Signal models of interest in this search: top squark pair production with the top squark decaying into a top quark and neutralino (top), and top squarks from cascade decays of gluinos (middle and bottom). The SUSY simplified model topology shown at the top is referred to as T2tt, the middle left model as T1tttt, middle right model as T1tttb the bottom left one as T5tttt and the bottom right one as T5ttcc.

between the $p_{\text{T}}^{\text{miss}}$ and several leading jets are applied to reject QCD events. Moreover, since the signal final states have tagged b-jets and top quark candidates, minimum requirements on the numbers of tagged b-jets and reconstructed top quarks are included in the baseline selection. We use the official recommendation for b tagging from the CMS b-tagging working group. However, for top quark candidates, we developed our own algorithms. For the analysis of the 2015 data, we used a cut-based top tagger approach [5]. For the analysis

of the 2016 data, described in this thesis, we designed an improved top tagger, which is described below. Several working points are designed for this top tagger.

The next step is to determine the top tagger working point and optimize the definitions of the search bins. We optimized the search bin definition for both medium and tight top tagger working points, and then chose the tight one because it is more sensitive for some signal models.

Then, we need to estimate the backgrounds for all search bins. The major background is from standard model $t\bar{t}$, W +jets and single top processes. The Z +jets, QCD, TTZ and other rare processes can be important in some search bins.

Finally, we can set limits on the model parameters by using the data yield, background estimation and expected model yields in a likelihood fit. The interpretation of data is based on simplified models [11], shown in Fig 4.1, as was already discussed.

4.2 Customized top tagger

4.2.1 Motivation

A cut-based top jet tagger was applied for the analysis of the 2015 data [4]. The cut-based top jet tagger provided a good reconstruction efficiency over the entire top quark p_T spectrum. However, the mistag rate, i.e., the rate at which an object that is not a top quark is erroneously tagged as such by the top tagging algorithm, is relatively high with this algorithm. Therefore, we designed a new top tagger, to reduce the mistag rate.

new top tagger has two major upgrades: fat jets (AK8 jet) for the mono-jet and di-jet top systems, and a multivariate algorithm for the tri-jet system. We expect a lower mistag rate with a similar top quark tagging efficiency after the upgrade.

The mono-jet top selection is relatively simple. The requirements for the mono-AK8 jet are:

- AK8 jet $p_T \geq 400$ GeV.
- Soft drop mass [52] between 105 and 210 GeV. The soft-drop algorithm reclusters the AK8 jet into subjet using the Cambridge-Aachen algorithm [36]. The soft radiation is removed to avoid bias on the jet mass determination.
- N-subjettiness $\tau_{32} \leq 0.65$. The N-subjettiness is defined in Eq 4.1:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min \Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}, \quad (4.1)$$

where $\Delta R_{N,k}$ is the angular separation between constituent k and candidate subjet N, d_0 is a normalization factor given by: $d_0 = \sum_k p_{T,k} R_0$, R_0 is 0.8 for AK8 clustering, $\tau_{32} = \frac{\tau_3}{\tau_2}$. Rejecting events with τ_{NM} close to 1 selects for jets that are more N-prong like [32]. Therefore, three-prong like AK8 jets are selected with this requirement.

If an AK8 jet is tagged as a top quark for the mono-jet case, it will be removed from the jet collection for the di-jet and tri-jet case.

We require one W-like AK8 jet and b-like AK4 jet in the di-jet case. The W-like AK8 jet is selected with following requirements:

- AK8 jet $p_T \geq 200$ GeV.

- Soft drop mass between 65 and 100 GeV.
- N-subjettiness $\tau_{21} \leq 0.60$. Two-prong like AK8 jets are selected with this requirement.

Then, the W-like AK8 jet is combined with an AK4 jet with $p_T \geq 40$ GeV. The additional requirements on the di-jet system are:

- di-jet system mass is between 100 to 250 GeV.
- the two jets in the di-jet system both lie in a cone of radius $R=1.0$ centered on the direction of their summed p_T vector.
- jet mass ratio between AK8 W-like jet and di-jet system is in the range $[0.85 \frac{m_W}{m_t}, 1.25 \frac{m_W}{m_t}]$, where m_W and m_t are the nominal W boson and top quark masses, respectively.

We design an algorithm to avoid double counting jet energy in the algorithm. An AK4 jet is considered matched if it lies within $\Delta R < 0.4$ of one of the soft-drop subjects of the tagged AK8 jet.

A multivariate analysis is applied to separate signal top quarks from background for candidates in the trijet category. There are two key elements in the multi-variable algorithm design: input variables and the algorithm itself.

To identify a top quark jet, we have two sets of variables for choice. The first set is on the particle candidate level. The p_T, η, ϕ can be used as training variables. However, there are typically about 400 particle-flow candidates in the tri-jet cone ($R=1.5$). The convolutional neural network [50] (CNN) might be a good algorithm to deal with the classification problem with that amount of information for one top quark candidate. However, the CNN training is very time consuming and it is not a commonly used method in

experimental high-energy physics for now. Therefore, we choose the second set: three-jet combination. There are only about 10 features per event, and it is therefore relatively quick to be trained with a decision based algorithm. Several multivariate algorithms are applied to the simulation sample. The random forest algorithm [45] is used as the identification algorithm.

All three-jet combinations are potential top quark candidates. The following variables are considered in the algorithm:

- Top quark candidate properties: mass, p_T , R cone size.
- Constituent jet properties: jet p_T , η , ϕ ; CSV value (b-tag likelihood), quark-gluon discriminator.
- Angular variables between jets: $\Delta\phi$, $\Delta\eta$, ΔR .

All the variables are carefully studied before being used to train the decision tree. We can use the quark gluon discriminator as an example. Because of its larger color charge, gluon jets radiate more, leading to a larger particle multiplicity, a broader p_T spectrum of its emitted particles with respect to the jet axis, and a softer energy spectrum of the emitted particles. The quark gluon discriminator is a likelihood method used to separate quark and gluon jets. A likelihood value near 1 means it is more like a quark jet, and otherwise that it is more like a gluon jet. We expect the top tagger mistag rate to be reduced after incorporating the quark-gluon discriminator into the decision tree since gluon jets, which tend to be fat, can be similar to the fat jets used in the mono-jet and di-jet top quark reconstruction categories. The likelihood is constructed from three variables:

- $p_T D$: jet energy dispersion variable, defined as $p_T D = \frac{\sqrt{\sum_i p_{T,i}^2}}{\sum_i p_{T,i}}$, with a sum over all PF candidates. Quark jets usually have a higher $p_T D$ than gluon jets.
- Multiplicity: the total number of particle-flow candidates in the jet. On average, gluon jets have a higher multiplicity than quark jets.
- Axis2: minor axis RMS in the $\eta - \phi$ plane of the particle flow candidates. This is an angular spread for the jet. Gluon jets are usually more spread out than quark jets.

The multivariate algorithm does a better job to evaluate the correlations between training variables than cut-based method. To determine the performance of the quark gluon discriminator on different jet flavors, p_T and η , we studied the following jet flavors in $t\bar{t}$ simulation samples: light flavor jet, c-jet, b-jet, gluon jet and pile-up jet. The p_T and η schemes are listed in Table 4.1.

Table 4.1: Jet bin for quark-gluon discriminator study.

Jet η Bin	1, HB	2, HBHE	3, HE	4, HEHF	5, HF	6, HF no PDF
Jet η	[0,1.31]	[1.31,1.39]	[1.39,2.65]	[2.65,3.14]	[3.14,4.7]	[4.7,5.19]
Jet p_T Bin	1, no PDF	2	3	4	5	6
Jet p_T	[0,20]	[20,40]	[40,50]	[50,80]	[80,100]	[100,Inf]

The performance in terms of jet η in different jet flavors is shown in Fig 4.3. The discrimination power for low p_T jets is not ideal.

The performance in terms of jet p_T in different jet flavors is shown in Fig 4.4. The discrimination power for HEHF and HF jets is not ideal.

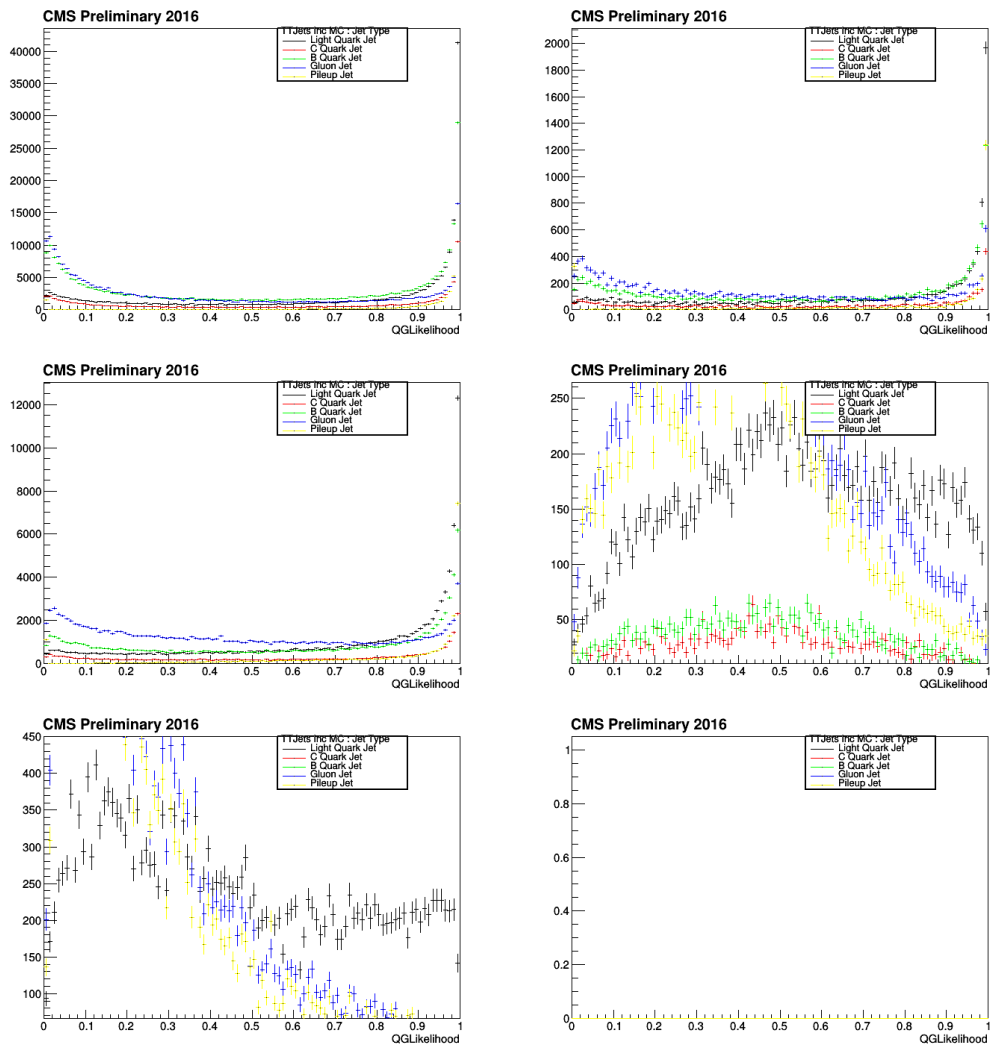


Figure 4.3: Top left: Quark Gluon likelihood for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.

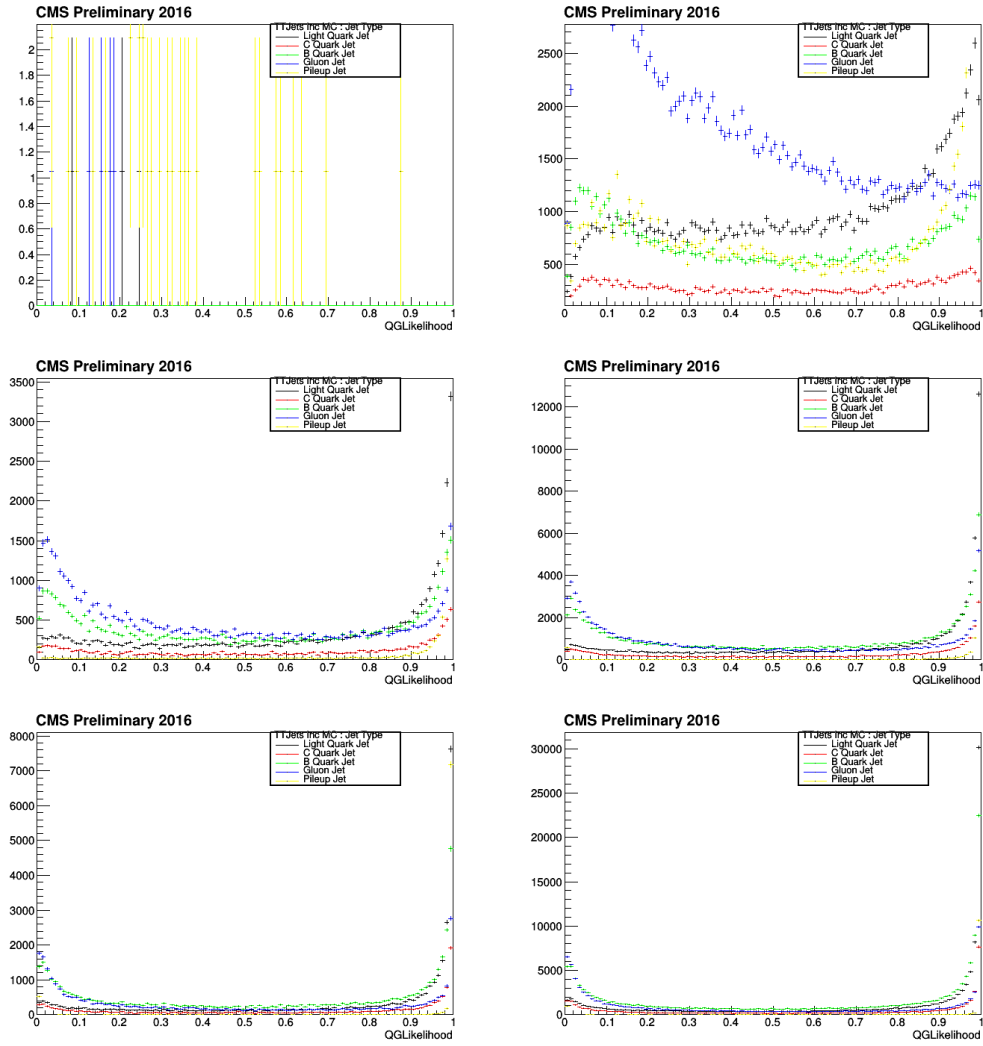


Figure 4.4: Top left: Quark Gluon likelihood for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.

The quark gluon discriminator does not work well for b-jet in all cases. Therefore, for tagged b jets, the likelihood is set equal to 1 (quark jet) before being trained.

All the variables are fed into the random forest [45] training algorithm for training in simulation samples. We designed three points, each with a different balance of reconstruction efficiency versus mistag rate. In the end, we chose the tight working point for the analysis, based on the results of sensitivity studies. The combined top quark reconstruction efficiency (Fig 4.5) is about 60%. The mistag rate (Fig 4.6) is around 20%, which is about a factor of two smaller than for the cut-based top tagger used in the analysis of the 2015 data. Additional studies were performed to study the difference between the performance of the top tagger in the full simulation and data, and also between the full simulation and fast simulation, since we use the fast simulation for limit setting. Scale factors that account for differences between the data and full simulation, or between the fast and full simulation, in the top tagger performance, are binned in the top quark p_T and applied to the data before limit setting. More details are given in Ref. [68].

4.3 Event selection and search bin design

As mentioned at the beginning of chapter 4, the analysis is designed for maximum sensitivity to the SUSY simplified model T2tt, T1tttt, T1ttbb, T5tttt and T5ttcc topology resulting in final states with multiple top quarks, hence multiple jets and b-tagged jets, produced in top squark decay, no leptons, and large p_T^{miss} . The T2tt, T5ttcc and T1tttt final states differ in the number of jets, b-tagged jets and top quarks produced.

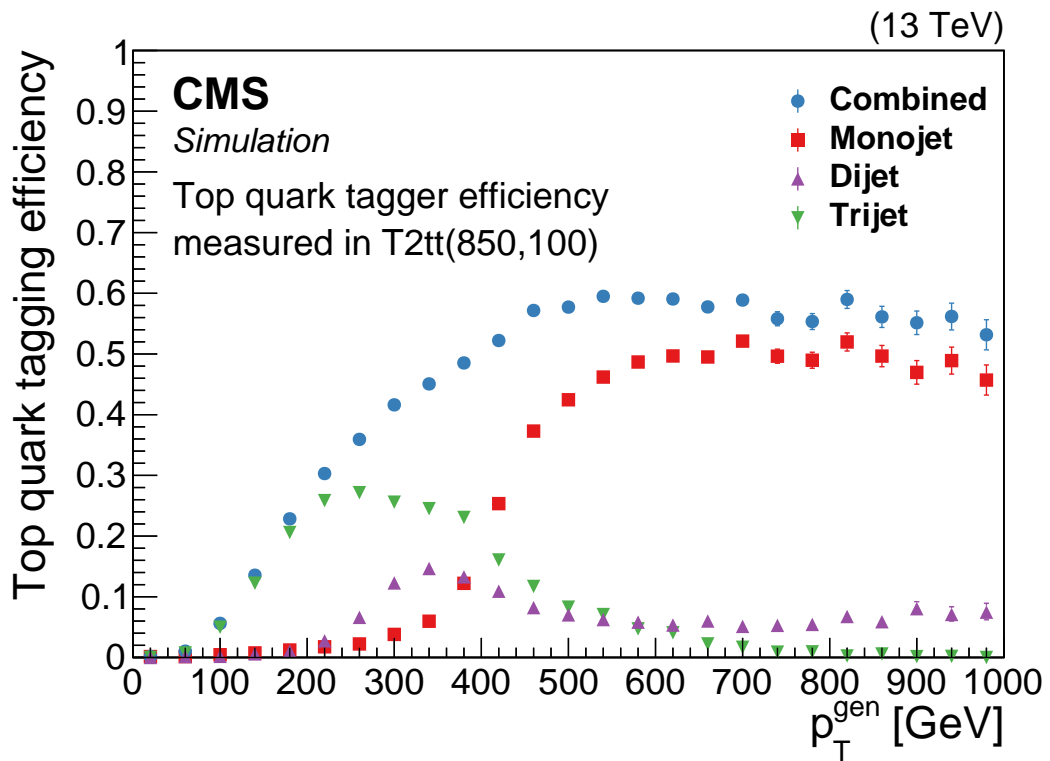


Figure 4.5: Efficiency of the top quark tagger as a function of generator-level top quark p_T for the monojet (red boxes), dijet (magenta upper-triangles), and trijet (green lower-triangles) categories and for their combination (blue circles), as determined using T2tt signal events with a top squark mass of 850 GeV and an LSP mass of 100 GeV. The vertical bars indicate the statistical uncertainties.

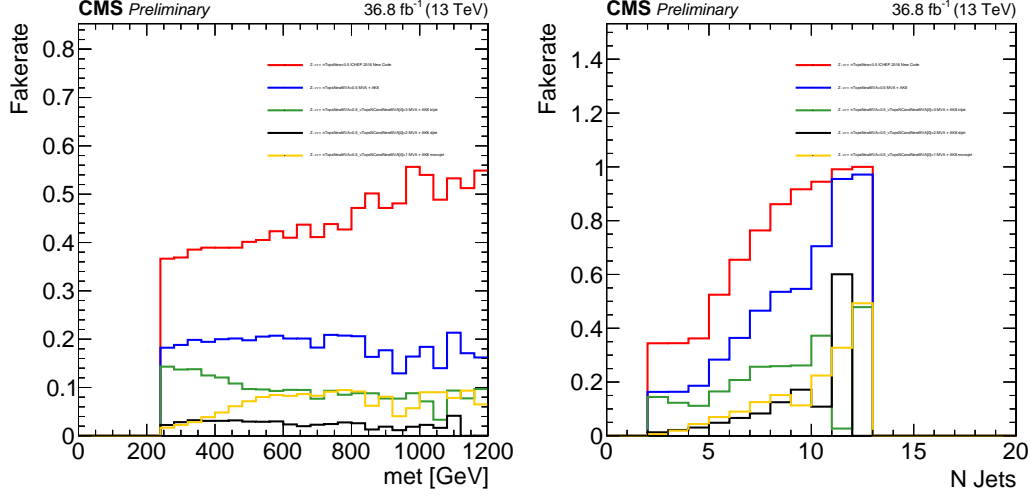


Figure 4.6: Top-jet reconstruction mistag rate, measured in Z +jets simulation samples. Left plot is the distribution in p_T^{miss} , Right plot is in N_{jets} . The red curve is the mistag rate for cut-based top tagger, the blue curve is the mistag rate for multivariate top tagger. The green, black and yellow are the mistag rate for tri-jet, di-jet and mono-jet, respectively.

Targeting the above-mentioned final states, the data are initially selected by requiring a minimum number of jets and b-jets (N_{jets} and $N_{\text{b-jets}}$) and large p_T^{miss} . The search regions are ultimately defined in exclusive bins of N_{tops} , $N_{\text{b-jets}}$, H_T , p_T^{miss} and M_{T2} , where N_{tops} is the number of reconstructed top quarks. SM backgrounds arise from processes such as $t\bar{t}$, Z +jets, W /top+jets and QCD multijets with smaller contributions from rare processes.

The data selection process starts with the triggers and follows with a baseline selection and the definition of the search bins. The top quark reconstruction and identification procedure (top tagging) is described in this section, as well as the MC samples that model signal and backgrounds.

4.3.1 Trigger

The trigger is the first filter applied to the data. We choose p_T^{miss} based triggers in this analysis instead of the HT based triggers used for the analysis of the 2015 data because of HT bit trigger issues in the 2016 run H data, where Run H corresponds to a data-taking period between September 16 2016 and October 28 2016 consisting of 92.5 fb^{-1} of data.

The trigger sets are used in this analysis are: a general trigger, described below, used for the search regions and for the estimation of the QCD multijet background and a single-muon trigger for the Z +jets background evaluation. We measure the search trigger efficiencies with different orthogonal triggers to check differences in the turn-on behavior. The corresponding efficiencies are applied to correct the simulations used in the analysis.

Six search triggers are used to collect events for this analysis. They are seeded by p_T^{miss} in the Level-1 trigger.

- HLT_PFMET100_PFMHT100_IDTight_v*
- HLT_PFMET110_PFMHT110_IDTight_v*
- HLT_PFMET120_PFMHT120_IDTight_v*
- HLT_PFMETNoMu100_PFMHTNoMu100_IDTight_v*
- HLT_PFMETNoMu110_PFMHTNoMu110_IDTight_v*
- HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v*

The probability for these triggers to accept events (trigger efficiency) is measured in a sample of events collected by the single-electron trigger

- HLT_Ele27_WPTight_v*,

which has been commonly used within CMS for the p_T^{miss} trigger efficiency measurement. Events from the single electron dataset are also required to have at least one offline reconstructed electron with $p_T > 30$ GeV. These selection criteria ensure that the single electron trigger is efficient.

To measure the search trigger efficiency, additional requirements to mimic the baseline selection, defined in Sec. 4.3.2, are imposed.

- Satisfy all filters
- Veto reconstructed muon
- $N_{\text{jets}} \geq 4$
- $N_{\text{b-jets}} \geq 1$
- $H_T \geq 300$ GeV
- $\Delta\phi(p_T^{\text{miss}}, j_{1,2,3}) > 0.5, 0.5, 0.3$

The efficiency of the search triggers is measured as a function of the offline p_T^{miss} . Events satisfying the above requirements are defined as the denominator, while events satisfying these criteria that are also selected by the search triggers are defined as the numerator. It has been found that the p_T^{miss} trigger efficiency has a non-trivial dependency on the offline H_T . The search trigger efficiencies are measured in the low H_T ($300 < H_T < 1000$) and high H_T ($H_T > 1000$) region, as shown in Fig. 4.7.

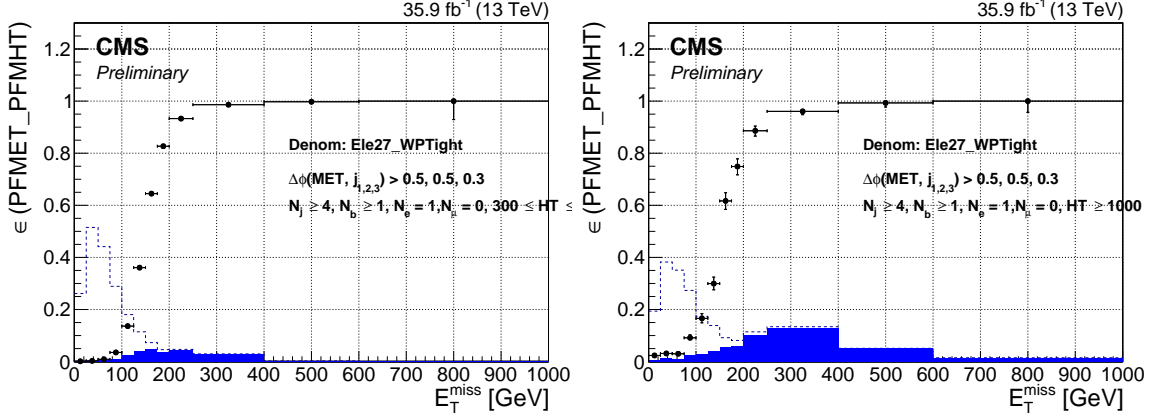


Figure 4.7: The trigger efficiency, denote by the black point, as a function of the offline p_T^{miss} for (left) $300 < H_T < 1000$ and (right) $H_T > 1000$. The error bar indicates the statistical uncertainty of the trigger efficiency. The dash blue line represents the denominator passing the selection, while the solid blue histogram represents the numerator where the denominator events also trigger the search triggers.

Besides measuring the p_T^{miss} trigger efficiency from the single electron dataset, the trigger efficiency can be measured from a single muon and HTMHT dataset. To account for possible bias from different measurements, we take the measurement from the single-electron dataset as the nominal one, and the variation from the measurements from single-muon and HTMHT dataset as the systematic uncertainty in the trigger efficiency, as shown Fig 4.8. For p_T^{miss} above 250 GeV, we observe similar p_T^{miss} trigger efficiencies, with a systematic uncertainty less than 1%.

For the p_T^{miss} trigger efficiency measured from the single-muon dataset, events are collected with the single-muon trigger

- HLT_Mu50_v*,

with the following offline requirements:

- Satisfy all filters

- Leading reconstructed muon has $p_T > 50$ GeV
- Veto reconstructed electron
- $N_{\text{jets}} \geq 4$
- $N_{\text{b-jets}} \geq 1$
- $H_T \geq 300$ GeV
- $\Delta\phi(p_T^{\text{miss}}, j_{1,2,3}) > 0.5, 0.5, 0.3$

For the p_T^{miss} trigger efficiency measured from the HTMHT dataset, events are collected by the HT triggers:

- HLT_PFHT200_v*,
- HLT_PFHT250_v*,
- HLT_PFHT300_v*,
- HLT_PFHT350_v*,
- HLT_PFHT400_v*,
- HLT_PFHT475_v*,
- HLT_PFHT600_v*,
- HLT_PFHT800_v*,
- HLT_PFHT900_v*,
- HLT_CaloJet500_NoJetID_v*,

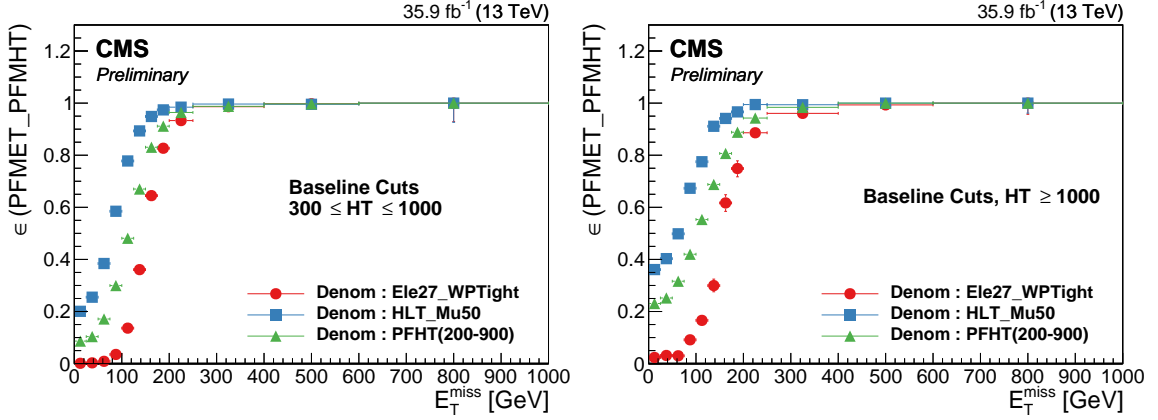


Figure 4.8: The trigger efficiency, denote by the black point, as a function of the offline p_T^{miss} for (left) $300 < H_T < 1000$ and (right) $H_T > 1000$. The error bar indicates the statistical uncertainty of the trigger efficiency. The blue square represents efficiency measured with single-muon dataset. The red point represents efficiency measured with single-electron dataset while the green triangle represents efficiency measured with HT dataset.

in which `HLT_CaloJet500_NoJetID_v*` is recommended at Level-1 to recover the inefficiency of the `L1_HTT` trigger during period H data taking. Events are required to satisfy the following criteria:

- Satisfy all filters
- Veto reconstructed electron
- Veto reconstructed muon
- $N_{\text{jets}} \geq 4$
- $N_{\text{b-jets}} \geq 1$
- $H_T \geq 300$ GeV
- $\Delta\phi(p_T^{\text{miss}}, j_{1,2,3}) > 0.5, 0.5, 0.3$

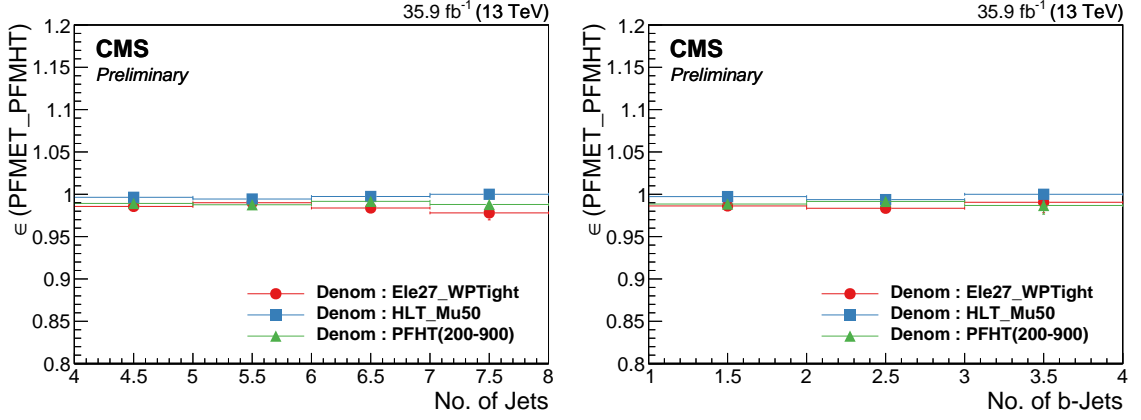


Figure 4.9: The trigger efficiency, denote by the black point, as a function of the offline (left) number of jets with $p_T > 30$ GeV and (right) number of b -tagged jets with $p_T > 30$ GeV. The error bar indicates the statistical uncertainty of the trigger efficiency. The blue square represents efficiency measured with single-muon dataset. The red point represents efficiency measured with single-electron dataset while the green triangle represents efficiency measured with HT dataset.

We also checked the trigger efficiency as a function of number of AK4 jets and b -tagged jets, after requiring $p_T^{\text{miss}} > 250$ GeV, as shown in Fig 4.9. Similar efficiencies are observed from the different measurements.

The QCD multijet background is estimated using the events triggered by the search triggers, but with offline requirements to select a QCD-enriched region. Since there is usually very little genuine p_T^{miss} in QCD events, the p_T^{miss} trigger efficiency in the QCD-enriched region is expected to be different from that of the search region in the low p_T^{miss} region. A measurement of the trigger efficiency in the QCD-enriched region is performed in the HTMHT dataset to avoid bias. Events are required to satisfy similar criteria as for the search trigger efficiency measurement, except an inverted $\Delta\phi(p_T^{\text{miss}}, j_{1,2,3})$ requirement is imposed:

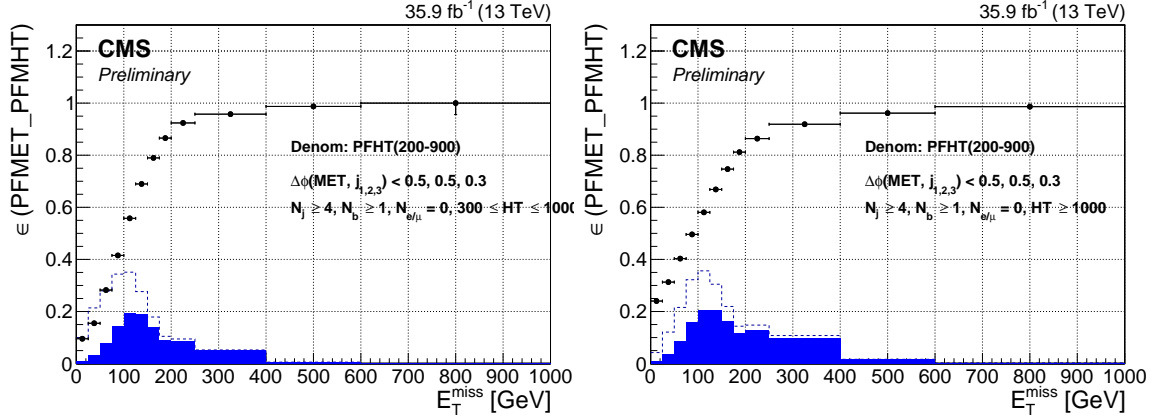


Figure 4.10: The trigger efficiency, denote by the black point, as a function of the offline p_T^{miss} for (left) $300 < H_T < 1000$ and (right) $H_T > 1000$. The error bar indicates the statistical uncertainty of the trigger efficiency. The dash blue line represents the denominator passing the selection, while the solid blue histogram represents the numerator where the denominator events also trigger the search triggers.

- Satisfy all filters
- Veto reconstructed electron
- Veto reconstructed muon
- $N_{\text{jets}} \geq 4$
- $N_{\text{b-jets}} \geq 1$
- $H_T \geq 300$ GeV
- $\Delta\phi(p_T^{\text{miss}}, j_{1,2,3}) < 0.5, 0.5, 0.3$

The trigger efficiency is measured as a function of the offline p_T^{miss} and is shown in Fig 4.10.

Events in the di-muon control sample, which is used for the estimation of the background from events in which a Z boson decays into neutrinos, are collected with the single muon trigger.

- HLT_IsoMu24_eta2p1_v*.
- HLT_IsoTKMu24_eta2p1_v*.
- HLT_Mu50_eta2p1_v*.

The trigger efficiencies are measured in the single-electron sample with the following criteria:

- Satisfy all filters
- Leading reconstructed electron has $p_T > 30$ GeV
- At least one reconstructed muon
- $N_{\text{jets}} \geq 4$
- $N_{\text{b-jets}} \geq 1$
- $H_T \geq 300$ GeV

The measured single muon trigger efficiency as a function of the reconstructed leading muon p_T and η is shown in Fig. 4.11. The muon trigger efficiency is also measured in the HT and MET dataset as cross checks. We also checked the benefits of adding the double muon triggers:

- HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
- HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*.

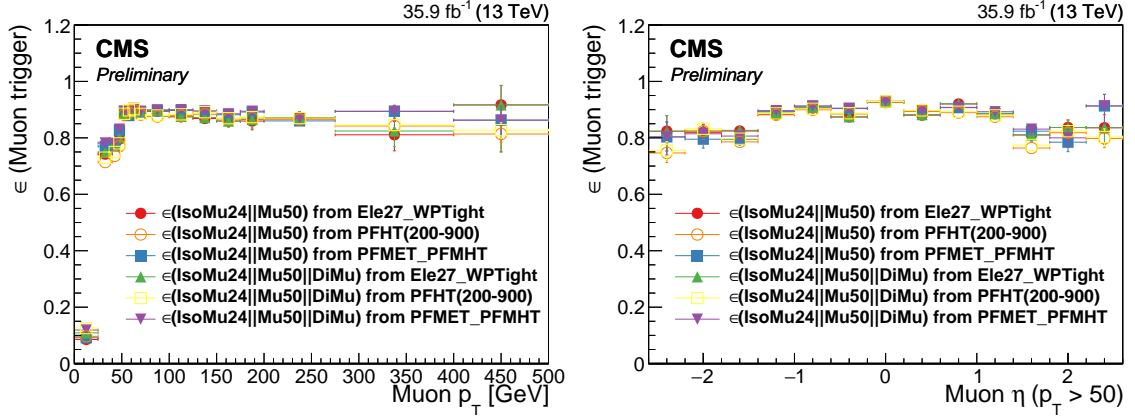


Figure 4.11: The trigger efficiency as a function of the offline leading Muon (left) p_T and (right) η .

We do observe a small gain in the muon trigger efficiencies as shown in Fig 4.11, but no benefit for the $Z \rightarrow \nu\nu$ background method.

We also checked the trigger efficiency as a function of number of AK4 jets and b -tagged jets, after requiring the highest- p_T muon to have $p_T > 50$ GeV, as shown in Fig 4.12.

We observe similar efficiencies from the different measurements.

4.3.2 Baseline selection

The search is based on a sample of all-hadronic events, with b -jets decaying from top quarks, large p_T^{miss} and no leptons. Initially, a loose baseline selection is applied in p_T^{miss} , H_T , N_{jets} and $N_{b\text{-jets}}$. This baseline selection preserves 2-20% of the signal events depending on the signal model. The events satisfying the baseline selection are divided into search regions defined in terms of N_{tops} , $N_{b\text{-jets}}$, p_T^{miss} , H_T and M_{T2} . By having many search regions, the analysis becomes “inclusive”, i.e., sensitive to many different signal topologies.

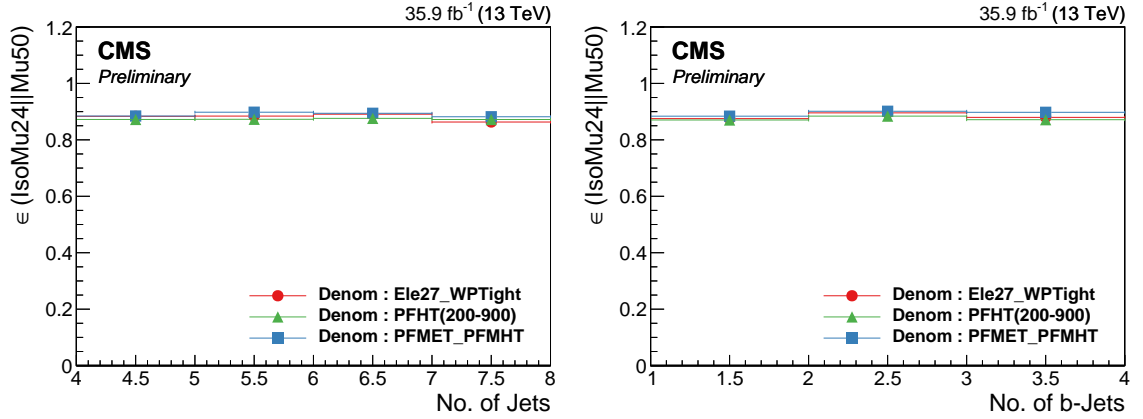


Figure 4.12: The trigger efficiency as a function of the offline (left) number of jets with $p_T > 30$ GeV and (right) number of b -tagged jets with $p_T > 30$ GeV, with leading reconstructed muon p_T above 50 GeV from the single-muon trigger. The red point denote the measurement from single-electron dataset. The green triangle represents measurement from HT dataset, while blue square represents measurement from MET dataset.

The following selection criteria define the baseline selection:

- Satisfies all filters that remove detector- and beam-related noise:
 - HBHE noise filter,
 - HBHEiso noise filter,
 - Ecal dead cell trigger primitive filter,
 - Primary vertex filter,
 - Bad EE super crystal filter,
 - Global tight beam halo filter,
 - Bad muon filter,
 - Bad charged hadron filter,
 - Loose JetID event filter.

- $N_{\text{jets}} \geq 4$:

Since the top squark is produced in pairs and each top squark decays to a top quark and an LSP, the all-hadronic final state will contain at least six jets. Not all the jets satisfy the selection criteria, therefore we only require at least four jets. Jets are reconstructed with the PF technique and clustered with the anti- k_T algorithm with a resolution parameter 0.4 [24] (AK4). Every jet is required to have $p_T > 30$ GeV and $|\eta| < 2.4$. In addition, they must satisfy the loose jet ID criteria for PF jets as recommended by the JetMET group. If any of the jets fail the loose jet ID criteria, the event is rejected (referred to as the loose JetID event filter above).

The two jets with highest p_T are required to have $p_T > 50$ GeV since SUSY predicts centrally produced jets with high p_T .

- $p_T^{\text{miss}} \geq 250$ GeV, where we use the type1-corrected particle-flow p_T^{miss} , with the jet energy corrections (Summer16_23Sep2016V3). The selection threshold is constrained by trigger efficiency requirements.
- $M_{T2} \geq 200$ GeV, which is mainly used to reduce SM background events with low value of M_{T2} . The M_{T2} variable works especially well for the $t\bar{t}$ events, where the M_{T2} shows a kinematic edge around the top quark mass.
- $H_T \geq 300$ GeV, with $H_T = \sum_{\text{jets}} p_T$, the scalar p_T sum over jets. Jets in the H_T calculation must meet the same jet selection criteria defined above.
- $N_{\text{b-jets}} \geq 1$, with b-jets identified using the CSV b-tagging algorithm (CSVm), medium working point.

- Muon veto:

Muon candidates are selected using the “Medium Muon” selection, recommended by the muon Physics Objects Group (POG). Muon candidates must satisfy $p_T > 10$ GeV and $|\eta| < 2.4$. Details of the muon medium ID criteria are listed in the Table 4.2 and Table 4.3. In addition to the official medium selection, we also apply an impact parameter requirement, with details listed in Table 4.4. A PF relative-isolation criteria is applied (mini-isolation), for which the isolation cone shrinks as a function of increasing muon p_T . The p_T in the mini-isolation cone is required to be less than 20% of the muon p_T to eliminate events with an isolated muon.

Table 4.2: Muon Medium ID 2016 HIP Safe.

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 4.3: Muon Medium ID HIP Safe Good Global Muon.

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 4.4: Additional Impact Parameter cut on Muon.

Muon Impact Parameter	
d0 $<$	0.2
dz $<$	0.5

- Electron veto:

Electron candidates are selected using the POG-recommended “Cut Based VETO” selection. Different selection criteria are applied to the barrel and endcap electromagnetic calorimeter regions. The selection criteria are listed in Table 4.5.

Electrons are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. Reconstructed isolated electrons are rejected using PF-based “mini-isolation” criteria, requiring less than 10% of the electron energy in the isolation cone.

Table 4.5: Electron Cut Based Veto 2016 Data in 80X CMSSW offline reconstruction condition.

	ECAL Barrel($ \eta < 1.479$)	ECAL Endcap($ \eta > 1.479$)
full5x5 sigmaIetaIeta <	0.0115	0.037
abs(dEtaInSeed) <	0.00749	0.00895
abs(dPhiIn) <	0.228	0.213
H/E <	0.356	0.211
Rel. comb. PF iso with EA corr <	0.175	0.159
abs(1/E-1/p) <	0.299	0.15
expected missing inner hits <	3	4
pass conversion veto	yes	yes

- Angular criteria:

Selection criteria on the angles between p_T^{miss} and the three jets with highest p_T , $\Delta\phi(p_T^{\text{miss}}, j_{1,2,3}) > 0.5, 0.5, 0.3$, are applied to remove events arising from QCD processes.

- Isolated track veto:

After applying the criteria described above, the residual background comes from $t\bar{t}$, single top, and W +jets events with one $W \rightarrow l\nu$ decay where l can be an electron or muon that is not identified, or a hadron from single-prong hadronic τ lepton decay. To further suppress these backgrounds, we reject events that have one or more isolated tracks. The track isolation is calculated from charged PF candidates consistent with

the reconstructed primary vertex ($|dz(PV)| < 0.1$ cm). The requirements are different for muon, electron and charged hadron tracks. For both electron and muon tracks, the isolated track requirements are: $p_T > 5$ GeV, $|\eta| < 2.5$ and relative isolation less than 0.2. For charged hadron tracks, the p_T requirement is raised to be at least 10 GeV and the relative isolation value to be less than 0.1. To retain more signal, and thus improve signal-to-background event discrimination, events with one isolated track, as defined above, are rejected only if they satisfy

$$m_T(tk, p_T^{\text{miss}}) = \sqrt{2p_T^{tk} p_T^{\text{miss}} (1 - \cos \Delta\phi)} < 100 \text{ GeV} \quad (4.2)$$

where p_T^{tk} is the transverse momentum of the track and $\Delta\phi$ is the azimuthal separation between the track and p_T^{miss} vector.

4.3.3 Search regions

In the analysis of the 2016 data, we bin the search regions in terms of the number of b-tagged jets and top quark candidates. The top quark reconstruction and identification procedure (top-tagging) is described in Sec. 4.2.2.

In order to improve background suppression, in particular the $t\bar{t}$ contribution, the p_T^{miss} , H_T and M_{T2} variables were added to the set that defines the search regions. The variable M_{T2} [54, 17] is an extension of the transverse mass variable that is sensitive to the pair production of heavy particles, each of which decays to a visible particle and an invisible particle. The $p^{3\text{-jet}}$, the p^{Rsys} , and the p_T^{miss} in an event are used to construct M_{T2} assuming the invisible particles are massless. In order to illustrate how M_{T2} is calculated, let us take the process $pp \rightarrow \tilde{t}\tilde{t}^* \rightarrow t\bar{t}\chi_1^0\chi_1^0$ as an example. This process contains two simultaneous

decays of an unseen particle of unknown mass(\tilde{t} or \tilde{t}^*) into an invisible particle (χ_1^0) and visible particle (t or \bar{t}). The variable M_{T2} is defined as:

$$M_{T2} \equiv \min_{\vec{q}_T^{(1)} + \vec{q}_T^{(2)} = \vec{p}_T} [\max\{m_T^2(\vec{p}_T^{t(1)}, \vec{q}_T^{(1)}; m_{\chi_1^0}), m_T^2(\vec{p}_T^{t(2)}, \vec{q}_T^{(2)}; m_{\chi_1^0})\}] \quad (4.3)$$

where the m_T^2 is the transverse mass,

$$m_T^2(\vec{p}_T^{t(1)}, \vec{q}_T^{(1)}; m_{\chi_1^0}) \equiv m_{t(1)}^2 + m_{\chi_1^0}^2 + 2(E_T^{t(1)} E_T^{(1)} - \vec{p}_T^{t(1)} \cdot \vec{q}_T^{(1)}) \quad (4.4)$$

M_{T2} is a minimization of two transverse masses with a constraint that the sum of the transverse momenta of both χ_1^0 's be equal to the missing transverse momentum in the event, i.e., $\vec{q}_T^{(1)} + \vec{q}_T^{(2)} = \vec{p}_T$. M_{T2} has a kinematic upper limit at the \tilde{t} mass. The superscripts (1) and (2) in the equations refer to the individual decays of the \tilde{t} particles. In the specific case of the analysis described in this thesis, we replace the quantities related to superscript (1) by those of associated with the fully reconstructed top quark, i.e., the $p^{3\text{-jet}}$. Similarly, we replace the quantities related to superscript (2) by those of the partially reconstructed top quark, i.e., the p^{Rsys} for $N_{\text{tops}} = 1$ and the fully reconstructed top quark for $N_{\text{tops}} \geq 2$. p_T^{miss} then corresponds to \vec{p}_T in the equation 4.3. Since we assume that the invisible particle is massless, $m_{\chi_1^0}$ is set to zero in Eqs. 4.3- 4.4.

Briefly, we start from the fact that at least one hadronic top quark candidate is required to be present in the search sample. If there are two top quark candidates, M_{T2} is calculated using the pair of top candidates and p_T^{miss} . In case there are more than two top candidates in the sample, we compute M_{T2} for all combinations and choose the M_{T2} with the smallest value. If there is only one top quark candidate identified by the top-tagging algorithm, we reconstruct the other top quark from the remanent of the event using the

b-tagged jet (or the hardest p_T jet if no b-tagged jet found) as a seed and the R-sys jet closest to the seed jet with an invariant mass between 50 GeV and the nominal top quark mass. In case no combination satisfies the invariant mass requirement, we use the seed jet as the only remnant of the other top quark. In the latter case, M_{T2} is calculated from the reconstructed top candidate, the remnant and the p_T^{miss} .

Figs 4.13 and 4.14 show a comparison between total SM backgrounds from simulation and several signal points for the four search bin variables after the baseline selection requirements. We can clearly see that all the variables have good discrimination power. The data are also shown and the total SM backgrounds are scaled to the same yield as the data.

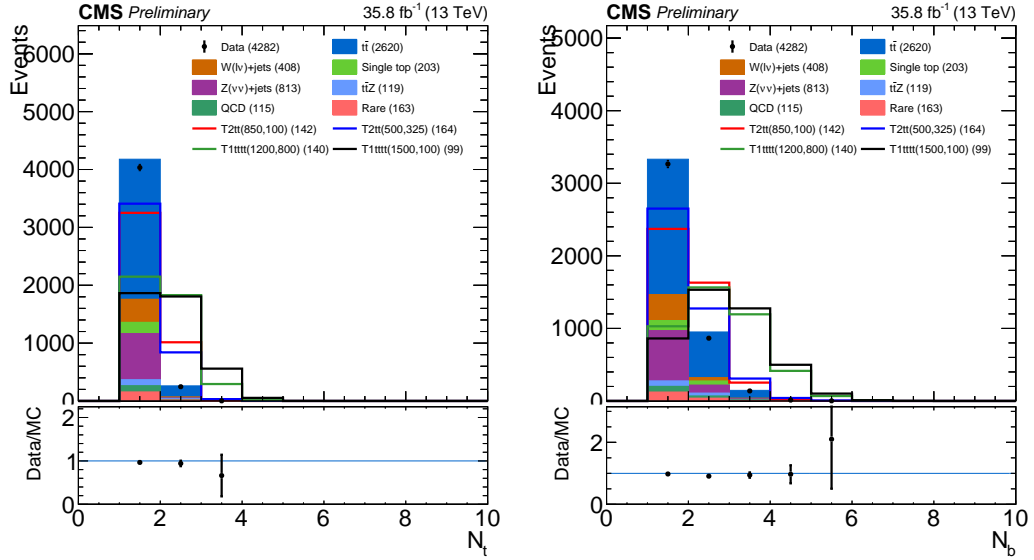


Figure 4.13: Comparison of the distributions between total SM backgrounds from simulation and several signal points for N_{tops} at the left and $N_{\text{b-jets}}$ at the right. Total SM backgrounds and signals are scaled to same data yield for a shape comparison. The yields for the Data and the SM backgrounds are in the legend. The scale is included in the legend for the signal points.

The search bins defined after the baseline selection criteria (in total 84 bins) are illustrated in Fig 4.15. An improvement in the selection efficiency is made by switching from the M_{T2} variable to H_T variable for search bins with $N_{b\text{-jets}} \geq 3$ or $N_{\text{tops}} \geq 3$ since these bins should be sensitive to the T1tttt signals where M_{T2} cannot be as clearly defined as for T2tt events. To accommodate the larger data sample collected in 2016 compared to 2015, and improve the search sensitivity, we use a finer segmentation of the search bins in p_T^{miss} , H_T and M_{T2} than for the analysis of the 2015 data. The search bin optimization is based on a significance scan of each of p_T^{miss} , M_{T2} and H_T dimension. However further adjustment was performed to have a reasonable number of control sample events for the major background predictions. The numbers displayed in the figures are the binning indices that are used throughout the analysis. The bins with $N_{\text{tops}} \geq 3$ are important for T1tttt signal but for T2tt we do not use them for the limit calculation.

4.3.4 MC samples for background and signal Studies

The analysis uses a set of Monte Carlo samples for background estimation method development and predictions as well as SUSY signal samples for the interpretation of the results. All the background samples are generated with the Geant4-based CMS simulation application while all the signal samples for limit setting are generated using the Fast Simulation application.

4.3.4.0.1 Standard Model Samples Monte Carlo samples of SM processes reconstructed with CMSSW release 8.0 (Summer16) are used in this thesis. A complete list of these samples is given in Table 4.6 and Table 4.7. The cross sections listed come from calcu-

lations performed at the next-to-next-to-leading-order (NNLO) unless otherwise noted. All samples use the “Moriond17” pileup scenario, which simulates a pileup distribution with an average of 25 interactions per bunch crossing and a 25 ns interval between bunches.

Table 4.6: Standard model Monte Carlo samples used in the analysis.

Dataset	σ (pb)
QCD MC samples (LO)	
QCD_HT100to200_TuneCUETP8M1_13TeV-madgraph-pythia8	27540000
QCD_HT200to300_TuneCUETP8M1_13TeV-madgraph-pythia8	1735000
QCD_HT300to500_TuneCUETP8M1_13TeV-madgraph-pythia8	366800
QCD_HT500to700_TuneCUETP8M1_13TeV-madgraph-pythia8	29370
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraph-pythia8	6524
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraph-pythia8	1064
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraph-pythia8	121.5
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraph-pythia8	25.42
SM $t\bar{t}$ MC samples	
TTJets_TuneCUETP8M1_13TeV-madgraph-pythia8	816.0
TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraph-pythia8	179.3
TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraph-pythia8	179.3
TTJets_DiLept_TuneCUETP8M1_13TeV-madgraph-pythia8	86.66
TTJets_HT-600to800_TuneCUETP8M1_13TeV-madgraph-pythia8	2.615
TTJets_HT-800to1200_TuneCUETP8M1_13TeV-madgraph-pythia8	1.077
TTJets_HT-1200to2500_TuneCUETP8M1_13TeV-madgraph-pythia8	0.195
TTJets_HT-2500toInf_TuneCUETP8M1_13TeV-madgraph-pythia8	0.002
SM $W \rightarrow l\nu$ MC samples	
WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraph-pythia8	1635
WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraph-pythia8	437.0
WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraph-pythia8	59.50
WJetsToLNu_HT-600ToInf_TuneCUETP8M1_13TeV-madgraph-pythia8	22.80
WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraph-pythia8	15.50
WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraph-pythia8	6.366
WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraph-pythia8	1.614
WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraph-pythia8	0.037

4.3.4.0.2 Signal samples Diagrams associated with the signal simplified models (SMS) used in this search for interpretation of the results are shown in Fig 4.1. The top diagram is

Table 4.7: Standard model Monte Carlo samples used in the analysis.

Dataset	σ (pb)
SM $Z \rightarrow \nu\bar{\nu}$ MC samples	
ZJetsToNuNu_HT-100To200_13TeV-madgraph	345.0
ZJetsToNuNu_HT-200To400_13TeV-madgraph	96.38
ZJetsToNuNu_HT-400To600_13TeV-madgraph	13.46
ZJetsToNuNu_HT-600ToInf_13TeV-madgraph	5.170
ZJetsToNuNu_HT-600To800_13TeV-madgraph	3.146
ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.453
ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.359
ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.0085
SM $Z \rightarrow l^+l^-$ MC samples	
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraph-pythia8	6025
DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraph-pythia8	171.5
DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraph-pythia8	52.58
DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraph-pythia8	6.984
DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraph-pythia8	1.676
DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraph-pythia8	0.831
DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraph-pythia8	0.143
DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraph-pythia8	0.0032
SM single-top MC samples	
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.80
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.80
SM di-boson and other rare process MC samples	
ttHJetTobb_M125_13TeV_amcatnlo-madspin-pythia8	0.293
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.228
TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.530
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnlo-madspin-pythia8	0.204
TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnlo-madspin-pythia8	0.423
WW_TuneCUETP8M1_13TeV_pythia8	115.0
WZ_TuneCUETP8M1_13TeV_pythia8	47.13
ZZ_TuneCUETP8M1_13TeV_pythia8	16.523
WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.165
WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.056
ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.014

often referred to as $T2tt(x, y)$ and represents direct squark-antisquark pair production with x and y the top squark and $\tilde{\chi}_1^0$ masses, respectively [30]. Under the assumption that the SUSY particles that could decay to top squarks are too heavy and beyond the reach of LHC Run 2, this diagram would represent the dominant process for top squark pair production and the target signal process for this analysis.

If the gluino is within the LHC reach in Run 2, gluino-induced processes such as those in the middle and bottom row of Fig 4.1 would become relevant to the analysis. The middle left diagram is called $T1tttt(x, y)$ with x and y the gluino and $\tilde{\chi}_1^0$ masses. In this model, the gluino undergoes a three-body decay into t, \bar{t} and $\tilde{\chi}_1^0$. The event kinematics are similar to the case where $\tilde{g} \rightarrow t\bar{t}, \tilde{t} \rightarrow t\tilde{\chi}_1^0$ [30] as in another model shown on the bottom right, denoted as $T5ttcc(x, y, z)$. The numbers in parentheses refer to the gluino, top squark, and $\tilde{\chi}_1^0$ masses.

Cross sections for a couple of mass points are shown in Table 4.8 for the $T2tt$ and $T1tttt$ SMS. These selected points were generated with full simulation and were used for cut flow studies and search region optimization. Limit setting was performed using fast simulation signal samples.

Table 4.8: Cross sections for a couple of mass points for the $T2tt$ and $T1tttt$ Simplified Models. The selected points were generated with full simulation.

Dataset	σ (pb)
SMS-T2tt_mStop500mLSP325_TuneCUETP8M1_13TeV-madgraph-pythia8	0.5185
SMS-T2tt_mStop850mLSP100_TuneCUETP8M1_13TeV-madgraph-pythia8	0.0190
SMS-T1tttt_mGluino1500mLSP100_TuneCUETP8M1_13TeV-madgraph-pythia8	0.014
SMS-T1tttt_mGluino1200mLSP800_TuneCUETP8M1_13TeV-madgraph-pythia8	0.086

4.4 Background estimation

The standard model processes are suppressed after the baseline selection criteria are applied, and signal sensitivities are optimized through the search bin design. However, there are still standard model backgrounds in the search region. In this analysis, the events from $t\bar{t}$, W +jets and single top are the major background. Z +jets, QCD and other rare processes also make non-negligible contributions in some search bins. The simulated background distributions are shown in the Fig 4.16 in the N_{tops} and $N_{\text{b-jets}}$ categories.

4.4.1 Backgrounds from top and W decays

The background from $t\bar{t}$, W +jets and single-top events that is not removed by the explicit lepton-veto and isolated track veto is the largest background in the analysis. It can contain either a hadronically decaying tau lepton or light leptons (electrons or muons) that are not isolated, not identified/reconstructed or are out of the acceptance region. Both types of background are estimated using the “translation factor method” (described in this section). The classic lost lepton method, which is used to predict specifically the events with light leptons, will be discussed in the following section. It serves as an important cross-check to the translation factor method.

4.4.1.1 Translation factor method

The translation factor is the ratio between the signal region yield and single lepton control sample yield in the simulation samples ($t\bar{t}$, W +jets and single top). The translation factor method is straightforward to apply. First, we calculate the translation factors for each

search bin in the single lepton control sample in simulation events. The simulation events are corrected to account for residual differences with respect to data with ISR-reweighting, b-jet tag efficiencies and lepton reconstruction efficiencies. Then, we apply those factors to the single lepton control samples in data to estimate the backgrounds. The translation factors for misidentified leptons (lost lepton) and decaying tau leptons are shown in Figs 4.17 and 4.18, for the single muon and electron control samples respectively. The final prediction combines the results from the single electron and single muon samples together (Figs 4.19 and 4.20).

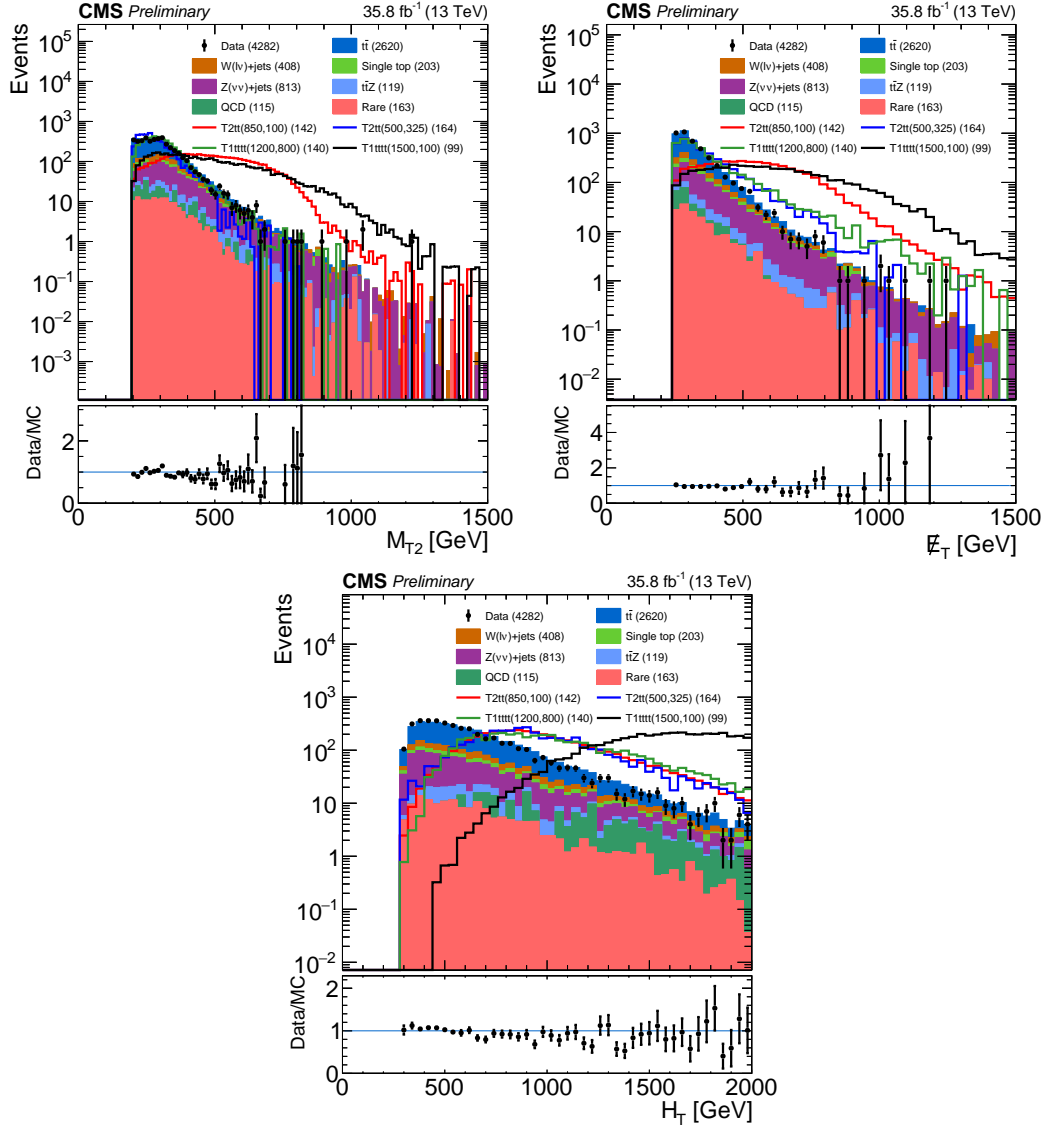


Figure 4.14: Comparison of the distributions between total SM backgrounds from simulation and several signal points for p_T^{miss} at the top left and M_{T2} at the top right with H_T at the bottom. Total SM backgrounds and signals are scaled to same data yield for a shape comparison. The yields for the Data and the SM backgrounds are in the legend. The scale is included in the legend for the signal points.

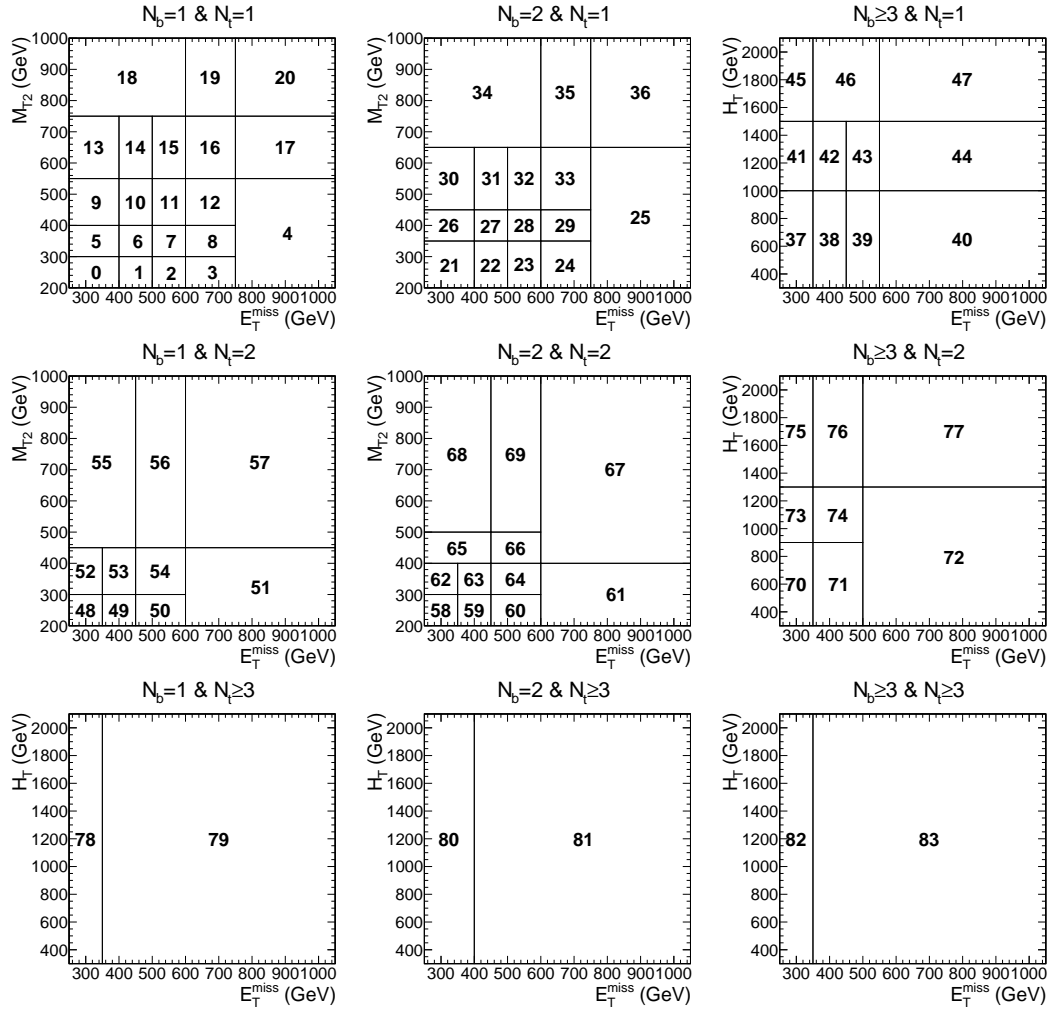


Figure 4.15: Original search bin definitions after baseline cuts (in total 84 search bins).

CMS Simulation Supplementary (13 TeV)

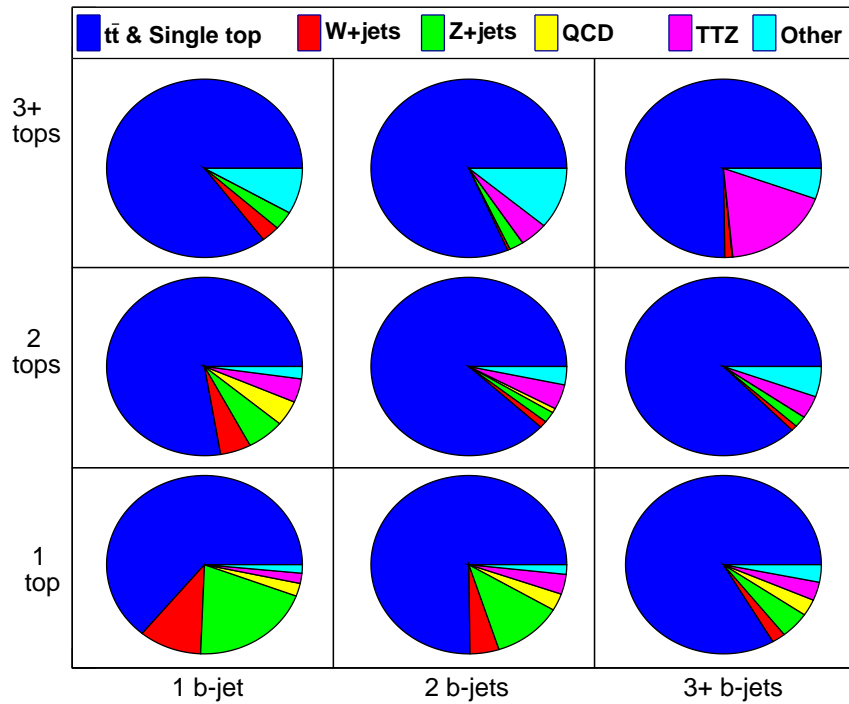


Figure 4.16: Background pie chart, from simulation.

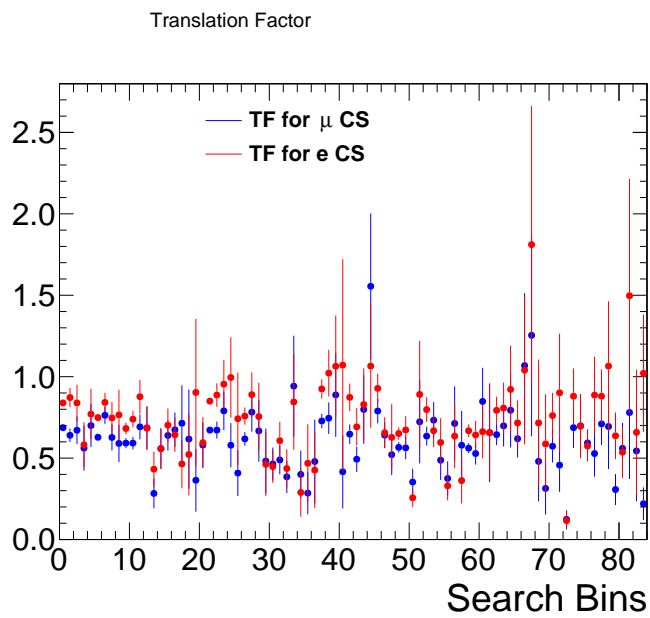


Figure 4.17: Translation factors for the τ_h background prediction with their uncertainties from limited MC statistics for both muon and electron CS.

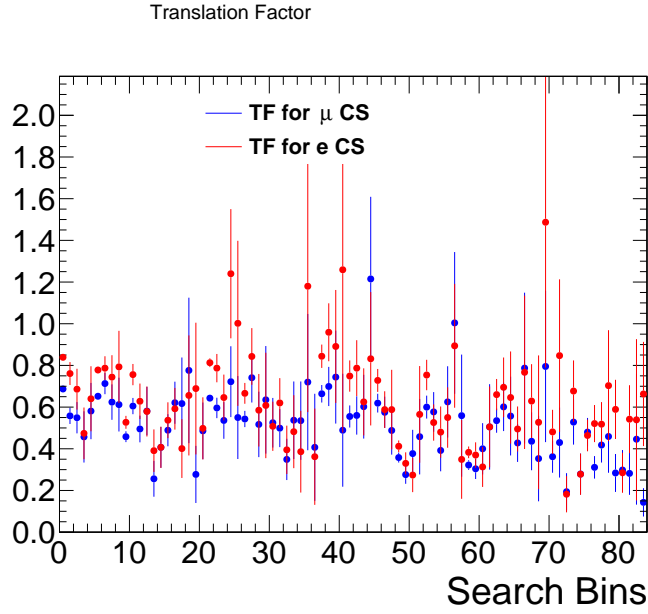


Figure 4.18: Translation factors for the lost lepton background prediction with their uncertainties from limited MC statistics for both muon and electron CS.

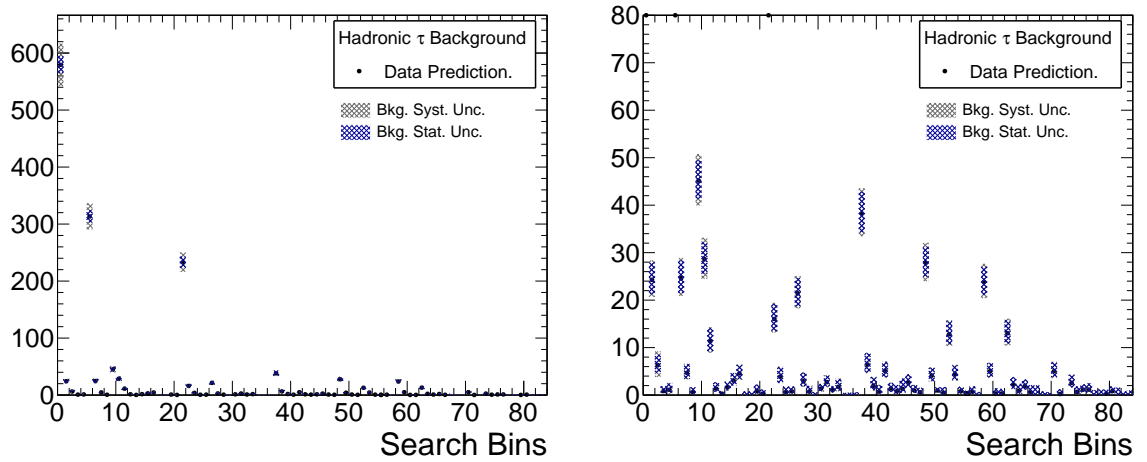


Figure 4.19: Predicted hadronic tau background yield for a 35.9 fb^{-1} data for all the search regions. Right plot is a zoomed version of left plot. Both statistical and total systematic uncertainties are shown.

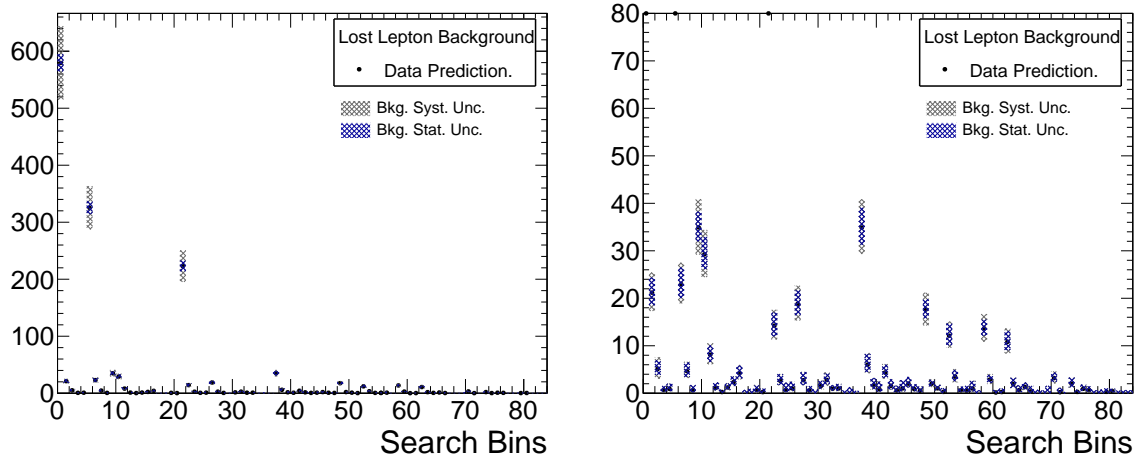


Figure 4.20: Predicted lost lepton background yield for a 35.9 fb^{-1} data for all the search regions. Right plot is a zoomed version of left plot. Both statistical and total systematic uncertainties are shown.

4.4.1.2 Classical lost lepton method

The classical lost lepton method is a validated method that was used in analysis of the 2015 data [5]. Although the analysis of the 2016 data uses the simulation based translation factor method for the background estimation for $t\bar{t}$, single top and W +jets, the classical method is still an important cross check method.

The name (lost lepton) of the background is interesting. On one hand, lost lepton indicates leptons that are lost because they are not in the acceptance, not reconstructed or not isolated. On the other hand, a lost lepton is like a lost boy. The boy becomes crazy and finally is misidentified as a jet. The expected numbers of lost lepton events are categorized in Table 4.9. The numbers of not accepted and not isolated events for muons and electrons are very similar. This is expected because of lepton flavor symmetry among $t\bar{t}$ decay channels. However, the numbers of not reconstructed muon and electron are very different. The difference comes from the nature of detector: the muons detected by muon system have higher reconstruction efficiencies than electrons in ECAL.

Table 4.9: Number of expected lost lepton events (not isolated, not identified/reconstructed and out of the acceptance) in $t\bar{t}$, single top and W +jets simulation events, for a luminosity of 35.9 fb^{-1} .

	iso	id	acc
muons	139.6 ± 16.4	71.3 ± 12.5	707.7 ± 38.2
electrons	148.8 ± 17.8	305.2 ± 25.3	692.8 ± 38.6

4.4.1.2.1 Description of the method The lepton identification chain is shown in Fig 4.21. The event will satisfy the lepton veto if the lepton in the event is out of acceptance, not identified, or not isolated. Therefore, we can calculate the acceptance, identification

and isolation efficiency in the single lepton control sample in simulation. And then, apply those efficiencies to the single lepton data sample with a model to estimate the lost lepton yield in the search region.

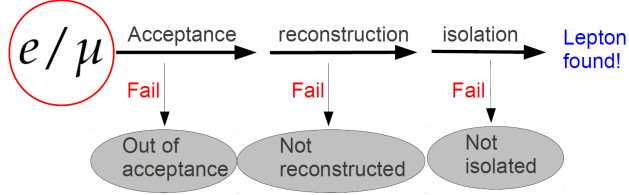


Figure 4.21: Sketch of the requirements electrons and muons from W decays must meet in order to be rejected by the explicit lepton veto.

However, the single lepton control sample can have a large signal contamination for some search bins, i.e., a large contribution of signal events. To reduce the contamination, only events with lepton transverse mass smaller than 100 GeV are considered. The lepton m_T is defined in Eq 4.5:

$$m_T = \sqrt{2p_T(\text{lepton})E_T^{\text{miss}}(1 - \cos(\Delta\Phi))}, \quad (4.5)$$

where $\Delta\Phi$ is the distance in Φ between the muon and the p_T^{miss} . For the electroweak control sample p_T^{miss} originates from the neutrino, m_T represents the transverse W -mass (M_T^W), and therefore the distribution falls sharply above 80 GeV. To compensate for the loss in efficiency due to the m_T selection requirement, a correction factor is added in the lost lepton prediction model.

We also need to consider the events with multiple lost leptons. Here we only estimate a part of the first-order effect: di-muon, di-electron and one electron plus one muon

events. Events with a higher number of leptons are not considered. The simplification is reasonable since the correction factor is very close to 1.

As mentioned in the event selection section, an isolated track veto is used to reject 1-prong hadronic tau events. An overall correction factor is applied to account for this selection requirement.

To summarize, the predicted number of $t\bar{t}$, W +jets and single top events with lost leptons, $N_{LostLepton}$ contributing to the search sample can be calculated by Eq 4.6:

$$N_{LostLepton} = \sum_{CS} \left(\sum_{i=e,\mu} (F_{ISO}^i + F_{ID}^i + F_{Acc}^i) \times F_{dilepton}^i \right) \times E_{Mtw} \times \epsilon_{isotrack}, \quad (4.6)$$

where \sum_{CS} is the sum over the events in the control sample after the baseline selection, F_{ISO}^i , F_{ID}^i and F_{Acc}^i are factors converting the number of events in the control sample to the number of lost lepton events due to respectively isolation, reconstruction or acceptance criteria, $F_{dilepton}^i$ is the correction factor for the di-lepton contribution, E_{Mtw} is the correction factor for the M_T^W selection requirement and $\epsilon_{isotrack}$ is the correction factor to compensate for the isolated track veto. The simulation based factors are described in the remainder of this section.

The control sample is normalized to compensate for the reduction of efficiency because of the $M_T^W < 100$ GeV requirement, as illustrated in Table 4.10, which shows the M_T^W correction factors, obtained from simulation, as a function of muon p_T .

The control sample is weighted according to the lepton isolation efficiency in order to model the non-isolated leptons in the signal region (F_{ISO}^i). The calculation is for muons

Table 4.10: M_T^W correction factors obtained from $t\bar{t}$, single top and W+jets simulation, after baseline selection. Only statistical uncertainties are shown.

Muon p_T [GeV]	10-20	20-30	30-40	40-50	50-70	70-100	> 100
M_T^W factor	1.04 ± 0.01	1.06 ± 0.01	1.07 ± 0.01	1.09 ± 0.01	1.11 ± 0.01	1.19 ± 0.01	1.66 ± 0.02

and electrons depending on the superscript:

$$F_{ISO}^{e/\mu} = \frac{1 - \epsilon_{ISO}^{e/\mu}}{\epsilon_{ISO}^{\mu}} \cdot \frac{\epsilon_{ID}^{e/\mu}}{\epsilon_{ID}^{\mu}} \cdot \frac{\epsilon_{Acc}^{e/\mu}}{\epsilon_{Acc}^{\mu}}. \quad (4.7)$$

To model the sample containing no identified electron or muons in the signal region, the control sample is weighted as follows:

$$F_{ID}^{e/\mu} = \frac{1}{\epsilon_{ISO}^{\mu}} \cdot \frac{1 - \epsilon_{ID}^{e/\mu}}{\epsilon_{ID}^{\mu}} \cdot \frac{\epsilon_{Acc}^{e/\mu}}{\epsilon_{Acc}^{\mu}}. \quad (4.8)$$

The isolation and reconstruction efficiencies as well as the acceptance are obtained from simulated $t\bar{t}$ and W+jets events. They are measured using reconstructed muons and electrons after the baseline selection and parameterized as a function of the muon p_T and the activity around the lepton, defined as the sum of the p_T 's of all PF particle candidates in an annulus outside a standard isolation divided by the p_T of the lepton:

$$A_{\mu/e} = \left(\sum_{PFcands}^{R_{\text{miniIso}} < r < 0.4} p_T \right) / p_T(lepton). \quad (4.9)$$

The lepton isolation and reconstruction efficiencies are analysis independent. Therefore, they are calculated without the baseline selection. They are shown in Figs 4.22, 4.23, 4.24 and 4.25.

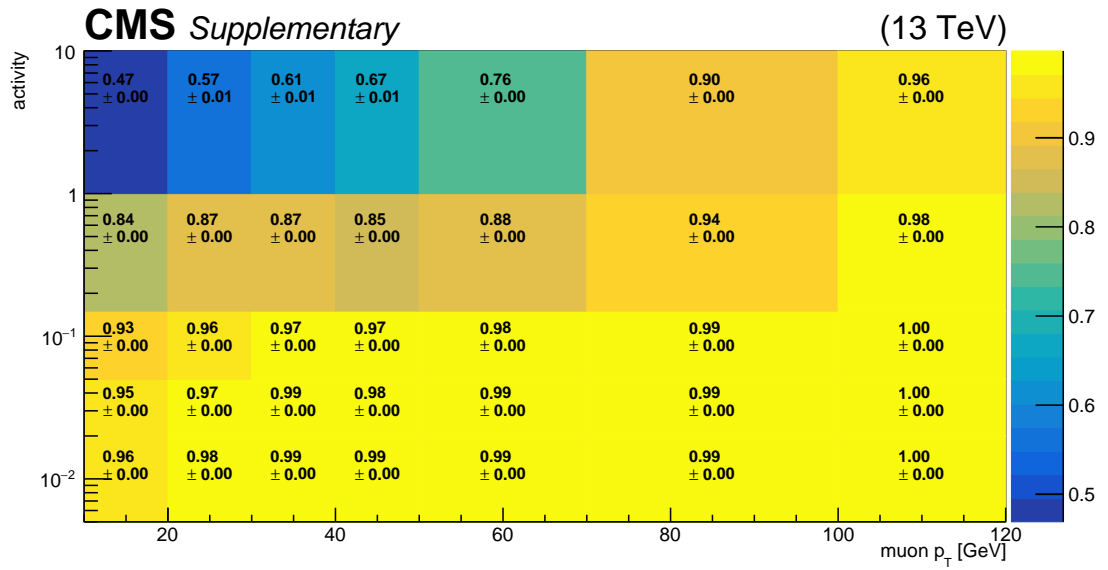


Figure 4.22: Muon isolation efficiencies as a function of the muon p_T and the activity around the muon. Only statistical uncertainties are shown.

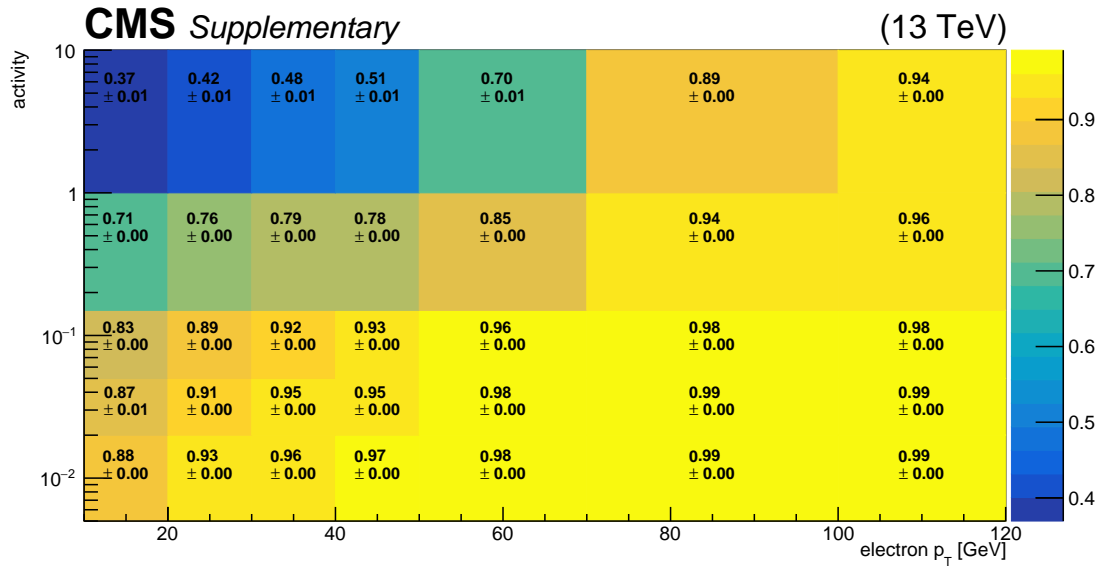


Figure 4.23: Electron isolation efficiencies as a function of the electron p_T and the activity around the electron. Only statistical uncertainties are shown.

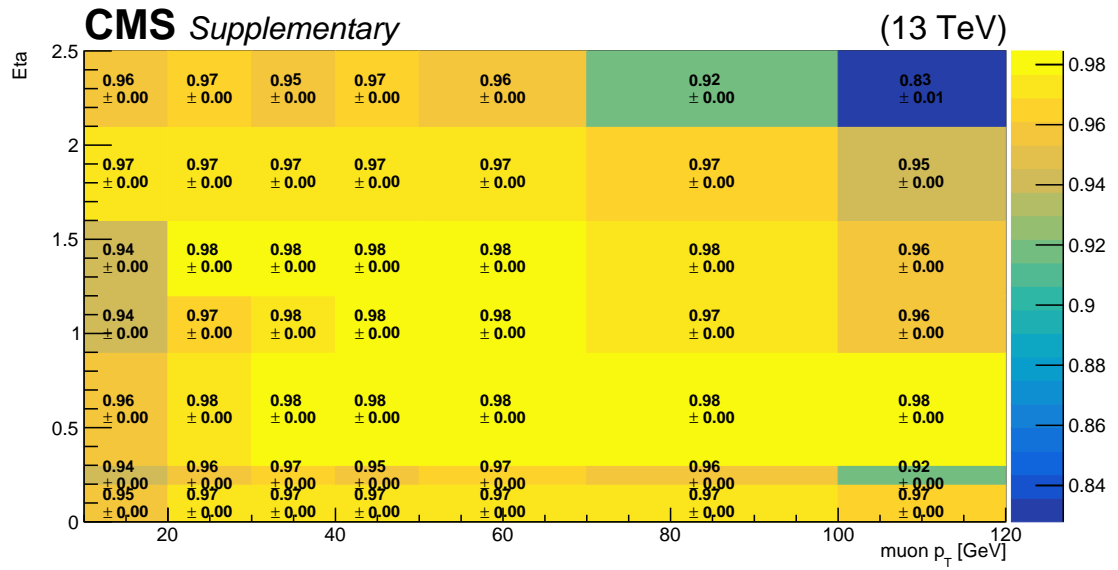


Figure 4.24: Muon reconstruction efficiencies as a function of the muon p_T and Eta. Only statistical uncertainties are shown.

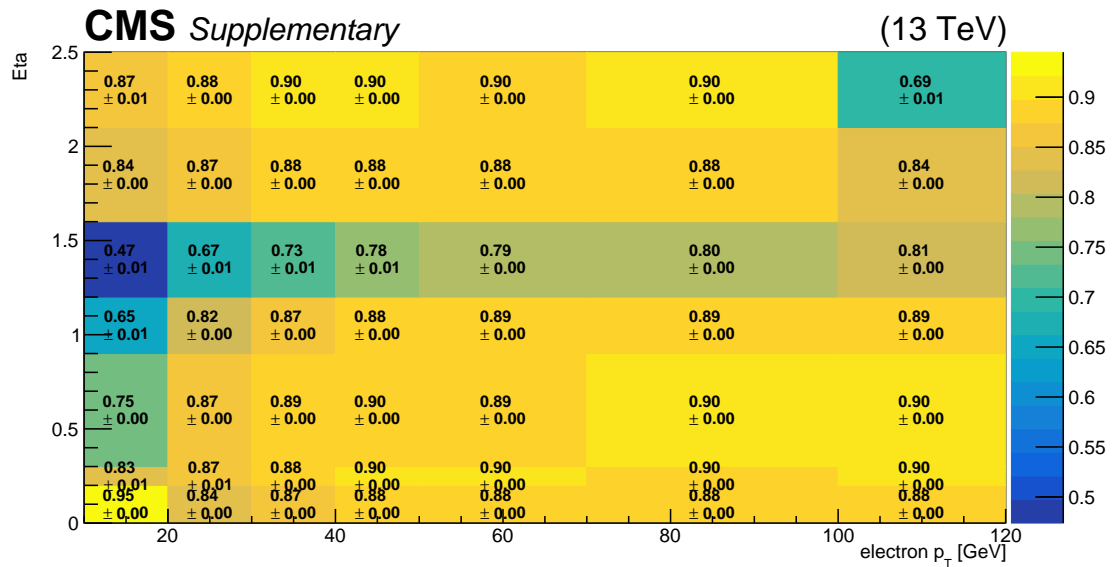


Figure 4.25: Electron reconstruction efficiencies as a function of the electron p_T and Eta. Only statistical uncertainties are shown.

Another multiplicative factor to the number of control region events originates from leptons that are out of acceptance. Such events occur when leptons have a transverse momentum below the lepton veto p_T threshold or when leptons are emitted in the forward region outside the η acceptance. For example, leptons from leptonic-tau decays tend to have low momentum as well as neutrinos that contribute to the event p_T^{miss} . This $F_{\text{Acc}}^{e/\mu}$ factor is modeled according to the following equation:

$$F_{\text{Acc}}^{e/\mu} = \frac{1}{\epsilon_{\text{ISO}}^\mu} \cdot \frac{1}{\epsilon_{\text{ID}}^\mu} \cdot \frac{1 - \epsilon_{\text{Acc}}^{e/\mu}}{\epsilon_{\text{Acc}}^\mu}. \quad (4.10)$$

The acceptance efficiencies are derived for each search bin from $t\bar{t}$ and W+jets simulated events, selected using the baseline criteria. They are shown in Fig 4.26.

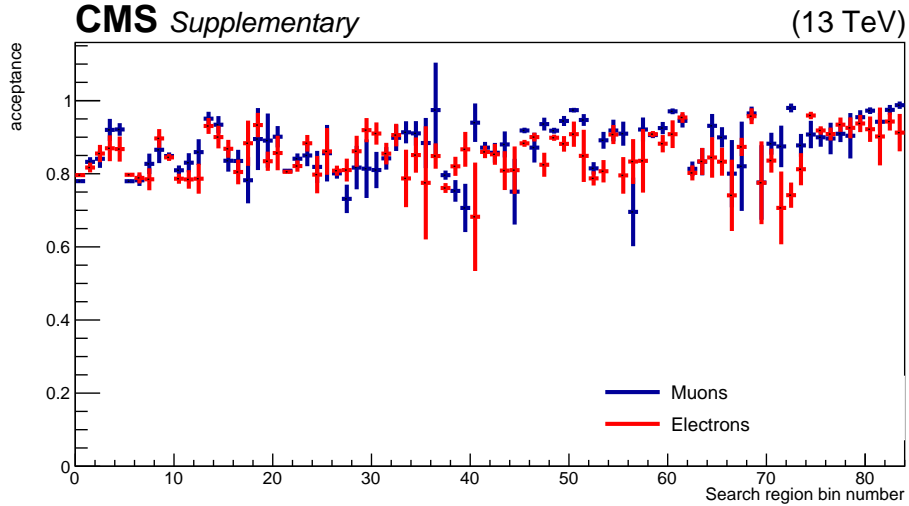


Figure 4.26: Acceptance efficiencies for muons and electrons in each of the search bins. Only the statistical uncertainties are shown.

The $t\bar{t}$ samples include both semi-leptonic/di-leptonic inclusive samples and HT-binned samples. The $t\bar{t}$ inclusive samples have the largest weight when the samples are combined. However, the $t\bar{t}$ inclusive samples have very few events in some bins, which

dramatically influences the efficiencies. We therefore set a criterion in the acceptance and isolated track efficiency calculations that if there are less than five events in a search bin, the acceptance and isolated track efficiencies are calculated for that bin using only the HT-binned samples.

Di-lepton events may also contribute to the background if both leptons are lost. In the muon control sample, there are di-lepton events that contribute when one lepton is lost while the other one is reconstructed and identified as a muon.

If ϵ_μ is the total muon efficiency (reco/identification times acceptance), then the 1-muon control sample contains:

- ϵ_μ events with exactly one identified/isolated muon and no other lost lepton.
- $2\epsilon_\mu(1 - \epsilon_\mu)$ events with one identified/isolated muon and one lost muon.
- $2\epsilon_\mu(1 - \epsilon_e)$ events with one muon and one lost electron.

For these events we apply a $(1 - \epsilon_\mu)/\epsilon_\mu$ factor to predict the number of events with lost muons and $(1 - \epsilon_e)/\epsilon_\mu$ to predict the number of lost electrons. This leads to an overestimate in the number of lost di-leptonic events in the prediction by a factor of two:

- Two lost muons case: prediction is $2(1 - \epsilon_\mu)(1 - \epsilon_\mu)$, expectation is $(1 - \epsilon_\mu)(1 - \epsilon_\mu)$.
- Two lost electrons case: prediction is $2(1 - \epsilon_e)(1 - \epsilon_e)$, expectation is $(1 - \epsilon_e)(1 - \epsilon_e)$.
- One lost muon and one lost electron case: prediction is $4(1 - \epsilon_\mu)(1 - \epsilon_e)$, expectation is $2(1 - \epsilon_\mu)(1 - \epsilon_e)$.

This effect is evaluated in simulated $t\bar{t}$, single top and W+jets events as the ratio between the number of events with one or two lost leptons over the number of events with one lost

lepton plus twice the number of events with two lost lepton. Separate correction factors are applied, $F_{dilepton}^{\mu}=(99.4 \pm 0.7)\%$ for muons and $F_{dilepton}^e=(97.1 \pm 0.9)\%$ for electrons.

The purity of the electron control sample is measured in simulation and found to be 0.96 ± 0.01 . The purity of the muon control sample is assumed to be 1.

Finally, the isolated track veto efficiency factor is applied in Eq 4.6 to get the final number of predicted lost lepton background events. The isolated track veto efficiency is computed from simulated events for each search bin and shown in Fig 4.27.

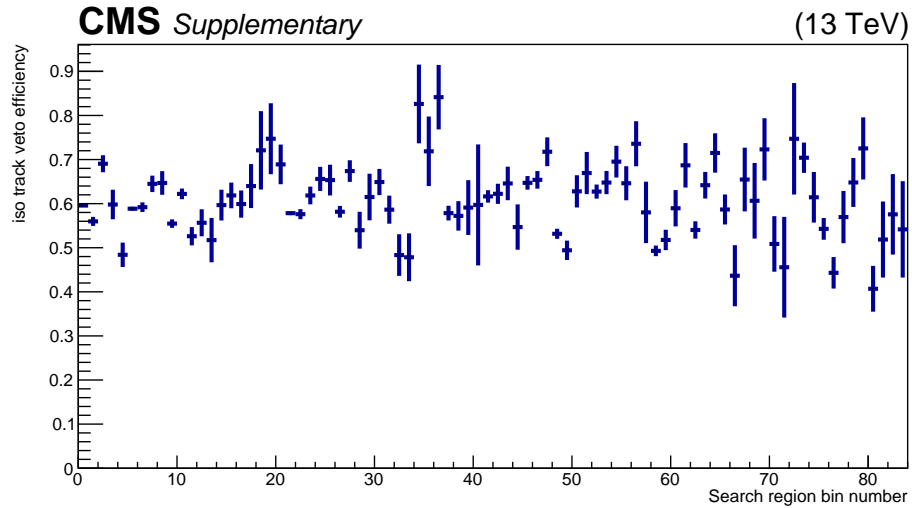


Figure 4.27: Isolated track veto efficiencies for each search bin.

4.4.1.2.2 Closure Test Closure tests have been performed using $t\bar{t}$, single top and W +jets Monte Carlo samples. The prediction, i.e. the number of events obtained using the lost lepton method explained in the previous section is compared with the expectation,

i.e. the true number of Monte Carlo simulated events with one or two lost leptons (coming from the W boson produced in top quark decays).

The results of the closure test are shown in Fig 4.28. The expectation/prediction disagreements are propagated into the systematic uncertainty, and is the major uncertainty in the lost lepton background estimation.

4.4.1.2.3 Systematic uncertainties The following systematic uncertainties are included for the lost lepton background prediction:

- Lepton isolation efficiency:

The muon and electron isolation efficiencies are obtained from simulated events. In order to estimate how well data and simulation agree, Tag-and-Probe efficiencies on the Z resonance from data and simulation are used, which are provided by the SUSY lepton scale-factor group. The maximum between the uncertainty obtained with the Tag-and-Probe comparing Data and simulation and the uncertainty on this value is used. Since no systematic bias is visible, no correction based on data/simulation scaling factors is introduced, and we merely take into account the systematic uncertainty obtained by propagating the SUSY lepton scale-factor group numbers. In addition the statistical uncertainty in the simulation efficiencies are propagated the same way as the Tag-and-Probe uncertainty.

- Lepton reconstruction/ID efficiency:

The muon and electron reconstruction and ID efficiencies are obtained from simulated events. The main uncertainty here arises from possible differences between efficiencies

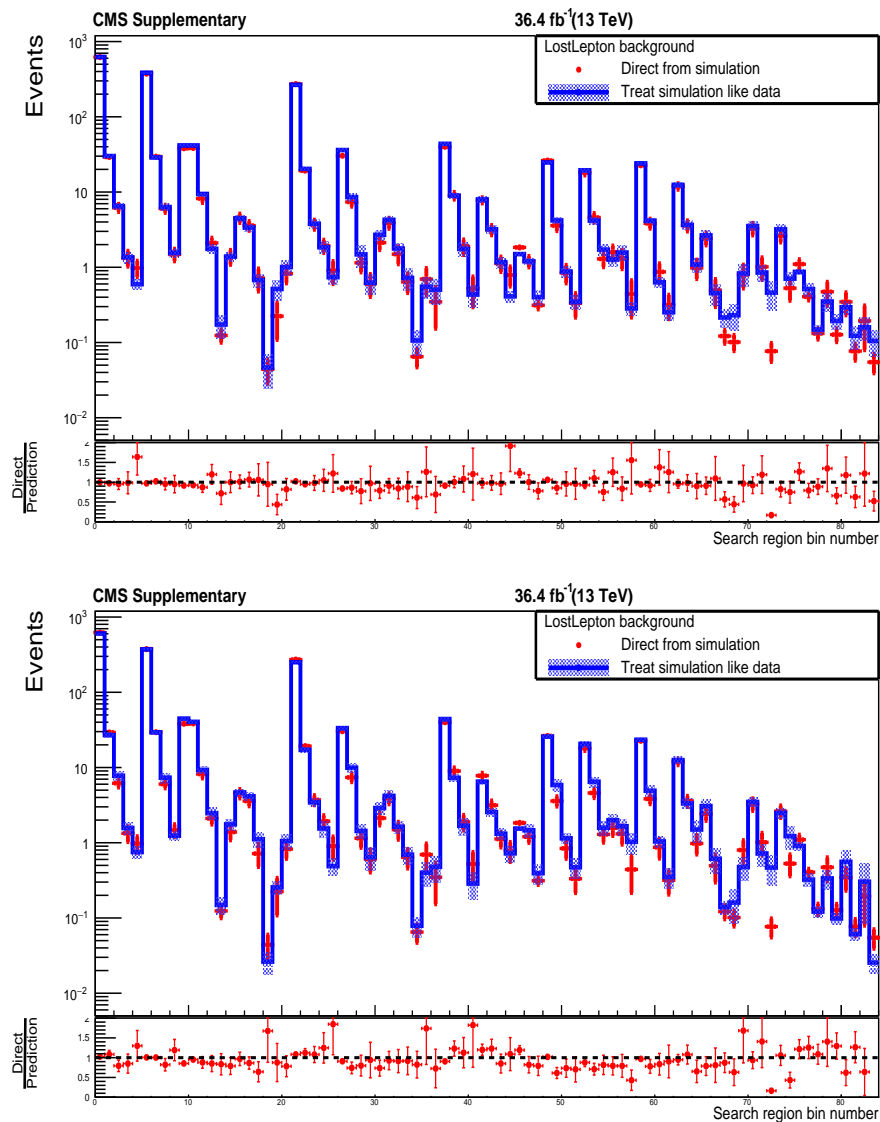


Figure 4.28: The lost lepton background in all the search regions of the analysis as predicted directly from $t\bar{t}$, single top and W +jets simulation (in red) and as predicted by applying the lost lepton background-determination procedure to simulated muon control sample (in black). The lower panel shows the ratio between the true and predicted yields. The top plot shows the prediction computed from the muon control sample. The bottom plot shows the prediction from the electron control sample.

obtained from data and simulation. By examining the reconstruction efficiencies as a function of the search variables it seems to be reasonable that using some inclusive efficiencies also provided by the SUSY lepton scale-factor groups are sufficient for deriving data/simulation uncertainties (same procedure as above). Furthermore, the statistical uncertainties from simulation are propagated.

- Lepton acceptance:

The uncertainty in the acceptance efficiency consists of the uncertainty in the parton distribution functions (PDF), the simulation renormalization/factorization scale and the uncertainty arising from the statistical precision of the efficiency maps. The PDF uncertainties are studied by varying the PDF sets used to produce the simulation samples, according to their uncertainties in the baseline selection.

- Control sample purity:

The purity is expected to be very high ($> 99\%$) so this only leads to a minor systematic uncertainty and a conservative uncertainty of 20% on the impurity is assigned. Furthermore, the statistical uncertainties from simulation are propagated.

- Di-lepton correction:

Both di-leptonic corrections are only minor compared to the remaining ones so a conservative systematic uncertainty is assigned here. An uncertainty of 50% is assigned on the number of di-leptonic events. Furthermore, the statistical uncertainties from simulation are propagated.

- M_T cut efficiency:

The uncertainty associated with the M_T cut consists of two parts: the statistical uncertainty in the efficiency map from simulation and a systematic uncertainty. For the latter, the uncertainty in the jet energy corrections (JEC) is propagated to p_T^{miss} (following the latest recommendations of the MET group) and the efficiency of the M_T cut is recalculated. Following this procedure, a conservative uncertainty of 1% is assigned as a systematic uncertainty.

- Isolated-track vetoes:

The isolated-track vetoes lead to a reduction of about 40% of the lost lepton background. Due to the size of this reduction, it is important to study the validity of the efficiency maps. A detailed study has been performed in the context of the RA2/b analysis [66] comparing Tag-and-Probe efficiencies on the Z resonance from data and simulation. This study has shown that a systematic uncertainty of 10% is a conservative estimate for the systematic uncertainty on the number of events removed by the isolated track veto.

- simulation closure:

Furthermore, an uncertainty in the precision of the simulation closure test is assigned. Since no true non-closure is observed (see Fig 4.28) the larger value of the non-closure or the statistical uncertainty in the non-closure is assigned for each bin.

The source and contribution of the different components of the lost lepton method systematic uncertainties are shown in Table 4.11.

Table 4.11: Source and contribution of the different components of the lost lepton method systematic uncertainties.

Source	Note	Electron control sample	Muon control sample
Muon Iso	Data-MC correction from tag and probe	1% to 3%	1% to 4%
Elec Iso	Data-MC correction from tag and probe	4% to 8%	2% to 5%
Muon ID	Data-MC correction from tag and probe	2% to 7%	4% to 11%
Elec ID	Data-MC correction from tag and probe	6% to 11%	2% to 8%
Acceptance	PDF and MC scale variation	4% to 97%	4% to 97%
Other SM contribution	20% uncertainty on the purity	0%	0%
Di-Muon Correction	Statistical uncertainty + 50%	0%	0%
Di-Electron Correction	Statistical uncertainty + 50%	1%	1%
Transverse Mass Cut	Variation of p_T^{miss} energy scale	0%	0%
Isolated track veto	Data-MC correction on isolated track veto efficiencies	10%	10%
Closure	Non-closure and statistical precision of the closure	2% to 104%	2% to 86%
Total		13% to 143%	12% to 131%

4.4.1.2.4 Lost Lepton background prediction The lost lepton method is applied to data event samples (collected with the search triggers) corresponding to an integrated luminosity of 35.9 fb^{-1} . The final predictions for all search bins are shown in Figs 4.29 and 4.30 (red points) and listed in Tables 4.12-4.17 for both the muon and electron channels.

Applying the procedure indicated by Eq. 4.6, each event in the control sample is weighted by the various efficiencies. A few bins have zero predicted events. In that case the statistical uncertainty is computed as the upper bound of the statistical uncertainty on 0 given by the Garwood interval (1.8) multiplied by the average translation factor to go from a raw number of events to the prediction. This Poisson uncertainty is treated in the Higgs combination tool using a gamma function [31, 15].

Figures 4.29 and 4.30 also show the comparison of the lost lepton background prediction between this method and the TF method. The first figure shows the prediction from the muon CS and the second figure shows the results from electron CS. As seen, both methods provide good agreement within the uncertainties. Thus the lost-lepton method provides a good cross-check of this important background.

4.4.1.3 Classical hadronic tau method

The hadronic decay of τ leptons (τ_h) is one of the largest components of the background from $t\bar{t}$, W +jets and single-top events contributing to the search regions. The

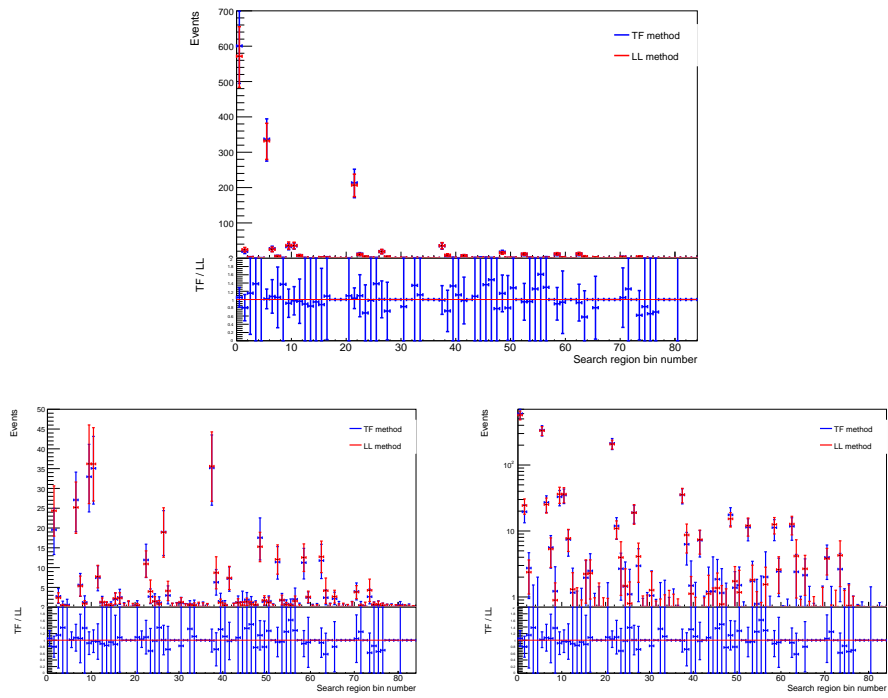


Figure 4.29: Lost lepton background predictions on muon control sample, in red. The blue points are the results obtained with the average TF method. The uncertainties include both the statistical and systematic uncertainties. Bottom left plot is a zoom of the top plot and the bottom right plot is log scale.

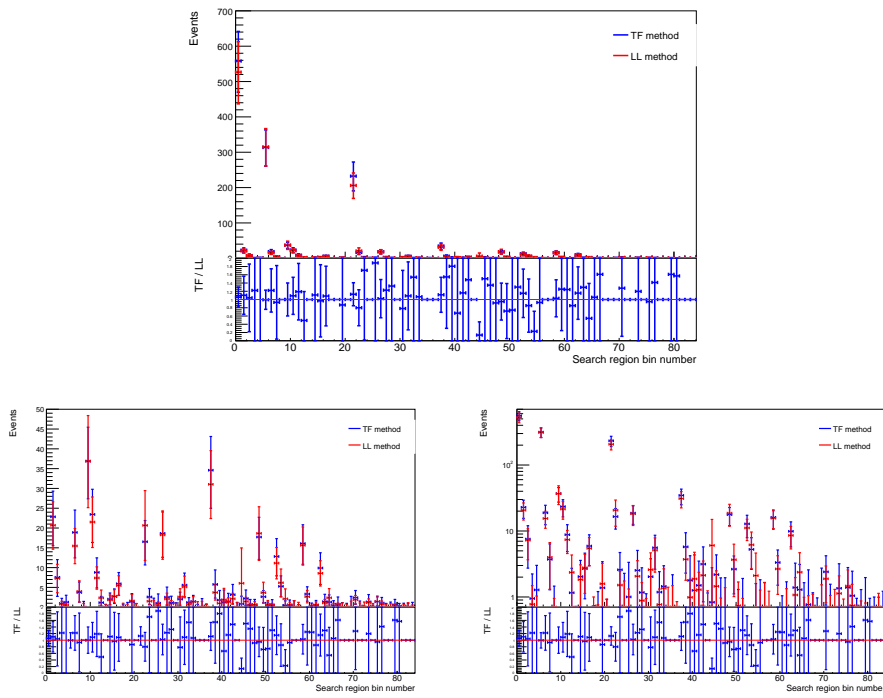


Figure 4.30: Lost lepton background predictions on electron control sample, in red. The blue points are the results obtained with the average TF method. The uncertainties include both the statistical and systematic uncertainties. Bottom left plot is a zoom of the top plot and the bottom right plot is log scale.

Table 4.12: Predicted lost lepton background yield from the muon control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	M_{T2} [GeV]	p_T^{miss} [GeV]	Lost Lepton Prediction
0	1	1	200-300	250-400	571.615 $+21.270$ $+80.824$ -21.237 -85.452
1	1	1	200-300	400-500	24.355 $+4.936$ $+4.015$ -4.837 -4.201
2	1	1	200-300	500-600	2.376 $+1.520$ $+0.669$ -1.070 -0.682
3	1	1	200-300	600-750	0.331 $+0.824$ $+0.182$ -0.331 -0.183
4	1	1	200-550	750+	0.231 $+0.607$ $+0.150$ -0.231 -0.150
5	1	1	300-400	250-400	332.123 $+16.111$ $+47.094$ -16.070 -49.965
6	1	1	300-400	400-500	25.202 $+4.961$ $+4.242$ -4.816 -4.427
7	1	1	300-400	500-600	5.341 $+2.178$ $+1.749$ -1.815 -1.773
8	1	1	300-400	600-750	0.894 $+1.127$ $+0.369$ -0.645 -0.372
9	1	1	400-550	250-400	36.205 $+5.030$ $+8.485$ -4.960 -8.724
10	1	1	400-550	400-500	36.177 $+5.448$ $+7.466$ -5.329 -7.691
11	1	1	400-550	500-600	7.748 $+2.303$ $+1.895$ -2.102 -1.937
12	1	1	400-550	600-750	1.306 $+1.277$ $+0.449$ -0.924 -0.453
13	1	1	550-750	250-400	0.609 $+0.681$ $+0.632$ -0.445 -0.633
14	1	1	550-750	400-500	0.432 $+0.756$ $+0.185$ -0.432 -0.187
15	1	1	550-750	500-600	2.232 $+1.523$ $+0.705$ -1.234 -0.719
16	1	1	550-750	600-750	2.293 $+1.551$ $+0.674$ -1.177 -0.688
17	1	1	550-750	750+	0.000 $+0.990$ $+0.000$ -0.000 -0.000
18	1	1	750+	250-600	0.000 $+1.661$ $+0.000$ -0.000 -0.000
19	1	1	750+	600-750	0.000 $+0.992$ $+0.000$ -0.000 -0.000
20	1	1	750+	750+	0.446 $+1.055$ $+0.243$ -0.446 -0.244
21	1	2	200-350	250-400	207.008 $+12.950$ $+28.139$ -12.904 -29.923
22	1	2	200-350	400-500	10.922 $+2.757$ $+2.101$ -2.558 -2.173
23	1	2	200-350	500-600	3.971 $+2.797$ $+1.238$ -2.627 -1.259
24	1	2	200-350	600-750	1.471 $+1.617$ $+0.595$ -1.043 -0.598
25	1	2	200-650	750+	0.796 $+1.127$ $+0.694$ -0.563 -0.700
26	1	2	350-450	250-400	18.915 $+3.488$ $+5.244$ -3.308 -5.348
27	1	2	350-450	400-500	4.115 $+2.637$ $+1.215$ -2.118 -1.231
28	1	2	350-450	500-600	0.000 $+1.057$ $+0.000$ -0.000 -0.000
29	1	2	350-450	600-750	0.000 $+1.090$ $+0.000$ -0.000 -0.000
30	1	2	450-650	250-400	1.269 $+1.452$ $+0.692$ -0.926 -0.695
31	1	2	450-650	400-500	0.000 $+0.884$ $+0.000$ -0.000 -0.000
32	1	2	450-650	500-600	0.519 $+0.701$ $+0.184$ -0.370 -0.185
33	1	2	450-650	600-750	0.482 $+1.016$ $+0.352$ -0.482 -0.353
34	1	2	650+	250-600	0.000 $+1.282$ $+0.000$ -0.000 -0.000
35	1	2	650+	600-750	0.000 $+1.123$ $+0.000$ -0.000 -0.000
36	1	2	650+	750+	0.000 $+0.953$ $+0.000$ -0.000 -0.000

Table 4.13: Predicted lost lepton background yield from the muon control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.

37	1	3+	300-1000	250-350	35.608	$+5.692$	$+6.655$
						-5.569	-6.875
38	1	3+	300-1000	350-450	8.699	$+3.703$	$+1.994$
						-3.514	-2.035
39	1	3+	300-1000	450-550	1.120	$+1.490$	$+0.472$
						-0.792	-0.476
40	1	3+	300-1000	550+	0.440	$+0.908$	$+0.330$
						-0.440	-0.331
41	1	3+	1000-1500	250-350	7.396	$+2.601$	$+1.654$
						-2.400	-1.697
42	1	3+	1000-1500	350-450	0.000	$+0.992$	$+0.000$
						-0.000	-0.000
43	1	3+	1000-1500	450-550	1.118	$+1.319$	$+0.589$
						-0.827	-0.592
44	1	3+	1000-1500	550+	0.520	$+1.291$	$+0.674$
						-0.520	-0.674
45	1	3+	1500+	250-350	1.366	$+1.192$	$+0.459$
						-0.818	-0.467
46	1	3+	1500+	350-550	0.777	$+1.127$	$+0.449$
						-0.553	-0.451
47	1	3+	1500+	550+	0.626	$+1.199$	$+0.321$
						-0.626	-0.325

Isolated track veto is applied in the baseline selection to reduce the hadronic τ background while sustaining a minimal impact on signal efficiency. After applying the veto, the remnant hadronic τ events are estimated using the method described as following.

When a W boson decays to a neutrino and a hadronically decaying τ lepton, the presence of neutrinos in the final state results in p_T^{miss} , and the event satisfies the lepton veto because the hadronically decaying τ is reconstructed as a jet. This background is estimated from a control sample (CS) of $\mu + \text{jets}$ events selected from data using a $\mu + H_T$ -based trigger, HLT_Mu15_IsoVVVL_PFHT350_v, and requiring exactly one μ with $p_T^\mu > 20 \text{ GeV}$ and $|\eta| < 2.4$. A selection requirement on the transverse mass of the W , $m_T = \sqrt{2p_T^\mu p_T^{\text{miss}}(1 - \cos \Delta\phi)} < 100 \text{ GeV}$, is imposed in order to select events containing a $W \rightarrow \mu\nu$ decay and to suppress possible new physics signal contamination, i.e., signal events present in the $\mu + \text{jets}$ sample. Here, $\Delta\phi$ is the azimuthal angle between the \vec{p}_T^μ and the p_T^{miss} directions. Because the $\mu + \text{jets}$ and $\tau_h + \text{jets}$ events arise from the same physics processes, the hadronic component of the two samples is the same except for the response

Table 4.14: Predicted lost lepton background yield from the muon control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	M_{T2} [GeV]	p_T^{miss} [GeV]	Lost Lepton Prediction
48	2	1	200-300	250-350	15.278 $^{+2.515}_{-2.443}$ $^{+2.699}_{-2.889}$
49	2	1	200-300	350-450	1.733 $^{+1.132}_{-0.989}$ $^{+0.601}_{-0.617}$
50	2	1	200-300	450-600	1.176 $^{+0.884}_{-0.649}$ $^{+1.195}_{-1.197}$
51	2	1	200-450	600+	0.000 $^{+0.705}_{-0.000}$ $^{+0.000}_{-0.000}$
52	2	1	300-450	250-350	12.074 $^{+3.125}_{-2.893}$ $^{+2.258}_{-2.343}$
53	2	1	300-450	350-450	1.818 $^{+1.440}_{-1.083}$ $^{+0.555}_{-0.566}$
54	2	1	300-450	450-600	0.624 $^{+0.939}_{-0.442}$ $^{+0.379}_{-0.381}$
55	2	1	450+	250-450	0.388 $^{+0.992}_{-0.388}$ $^{+0.205}_{-0.206}$
56	2	1	450+	450-600	1.550 $^{+2.135}_{-1.097}$ $^{+0.651}_{-0.655}$
57	2	1	450+	600+	0.000 $^{+0.646}_{-0.000}$ $^{+0.000}_{-0.000}$
58	2	2	200-300	250-350	12.561 $^{+2.605}_{-2.540}$ $^{+2.304}_{-2.466}$
59	2	2	200-300	350-450	2.606 $^{+1.165}_{-1.012}$ $^{+1.057}_{-1.071}$
60	2	2	200-300	450-600	0.000 $^{+0.535}_{-0.000}$ $^{+0.000}_{-0.000}$
61	2	2	200-400	600+	0.000 $^{+0.650}_{-0.000}$ $^{+0.000}_{-0.000}$
62	2	2	300-400	250-350	12.739 $^{+3.172}_{-3.016}$ $^{+2.566}_{-2.647}$
63	2	2	300-400	350-450	4.170 $^{+3.038}_{-2.833}$ $^{+1.481}_{-1.497}$
64	2	2	300-400	450-600	0.000 $^{+0.937}_{-0.000}$ $^{+0.000}_{-0.000}$
65	2	2	400-500	250-450	2.671 $^{+1.609}_{-1.391}$ $^{+0.804}_{-0.815}$
66	2	2	400-500	450-600	0.000 $^{+0.874}_{-0.000}$ $^{+0.000}_{-0.000}$
67	2	2	400+	600+	0.000 $^{+1.032}_{-0.000}$ $^{+0.000}_{-0.000}$
68	2	2	500+	250-450	0.000 $^{+0.710}_{-0.000}$ $^{+0.000}_{-0.000}$
69	2	2	500+	450-600	0.000 $^{+1.422}_{-0.000}$ $^{+0.000}_{-0.000}$
70	2	3+	300-900	250-350	3.810 $^{+1.407}_{-1.187}$ $^{+1.132}_{-1.156}$
71	2	3+	300-900	350-500	0.342 $^{+0.781}_{-0.342}$ $^{+0.232}_{-0.233}$
72	2	3+	300-1300	500+	0.000 $^{+1.289}_{-0.000}$ $^{+0.000}_{-0.000}$
73	2	3+	900-1300	250-350	4.272 $^{+2.567}_{-2.301}$ $^{+1.633}_{-1.658}$
74	2	3+	900-1300	350-500	0.675 $^{+0.772}_{-0.499}$ $^{+0.300}_{-0.305}$
75	2	3+	1300+	250-350	0.732 $^{+1.007}_{-0.732}$ $^{+0.297}_{-0.300}$
76	2	3+	1300+	350-500	0.446 $^{+0.726}_{-0.446}$ $^{+0.367}_{-0.369}$
77	2	3+	1300+	500+	0.000 $^{+0.757}_{-0.000}$ $^{+0.000}_{-0.000}$
78	3+	1	200+	250-350	0.000 $^{+0.673}_{-0.000}$ $^{+0.000}_{-0.000}$
79	3+	1	200+	350+	0.000 $^{+0.661}_{-0.000}$ $^{+0.000}_{-0.000}$
80	3+	2	200+	250-400	0.280 $^{+0.426}_{-0.199}$ $^{+0.166}_{-0.168}$
81	3+	2	200+	400+	0.000 $^{+0.634}_{-0.000}$ $^{+0.000}_{-0.000}$
82	3+	3+	200+	250-350	0.000 $^{+0.538}_{-0.000}$ $^{+0.000}_{-0.000}$
83	3+	3+	200+	350+	0.000 $^{+0.414}_{-0.000}$ $^{+0.000}_{-0.000}$

Table 4.15: Predicted lost lepton background yield from the electron control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	$M_{\text{T}2}$ [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	Lost Lepton Prediction
0	1	1	200-300	250-400	526.622 $^{+28.207}_{-28.173}$ $^{+81.255}_{-85.253}$
1	1	1	200-300	400-500	20.687 $^{+4.571}_{-4.416}$ $^{+3.960}_{-4.112}$
2	1	1	200-300	500-600	7.257 $^{+2.843}_{-2.416}$ $^{+2.633}_{-2.658}$
3	1	1	200-300	600-750	0.777 $^{+1.071}_{-0.572}$ $^{+0.420}_{-0.423}$
4	1	1	200-550	750+	0.542 $^{+0.841}_{-0.401}$ $^{+0.283}_{-0.284}$
5	1	1	300-400	250-400	315.171 $^{+20.698}_{-20.657}$ $^{+47.346}_{-50.007}$
6	1	1	300-400	400-500	15.472 $^{+3.652}_{-3.405}$ $^{+2.807}_{-2.914}$
7	1	1	300-400	500-600	3.992 $^{+2.729}_{-2.221}$ $^{+1.491}_{-1.506}$
8	1	1	300-400	600-750	0.000 $^{+0.968}_{-0.000}$ $^{+0.000}_{-0.000}$
9	1	1	400-550	250-400	36.861 $^{+5.192}_{-5.093}$ $^{+10.358}_{-10.549}$
10	1	1	400-550	400-500	21.512 $^{+4.524}_{-4.309}$ $^{+4.636}_{-4.775}$
11	1	1	400-550	500-600	7.384 $^{+2.451}_{-2.111}$ $^{+1.845}_{-1.884}$
12	1	1	400-550	600-750	2.355 $^{+2.132}_{-1.818}$ $^{+0.770}_{-0.783}$
13	1	1	550-750	250-400	0.000 $^{+0.621}_{-0.000}$ $^{+0.000}_{-0.000}$
14	1	1	550-750	400-500	1.835 $^{+1.175}_{-0.883}$ $^{+0.869}_{-0.875}$
15	1	1	550-750	500-600	2.774 $^{+1.681}_{-1.374}$ $^{+0.887}_{-0.900}$
16	1	1	550-750	600-750	5.459 $^{+2.155}_{-1.844}$ $^{+1.654}_{-1.681}$
17	1	1	550-750	750+	0.000 $^{+1.019}_{-0.000}$ $^{+0.000}_{-0.000}$
18	1	1	750+	250-600	0.000 $^{+0.778}_{-0.000}$ $^{+0.000}_{-0.000}$
19	1	1	750+	600-750	1.586 $^{+1.736}_{-1.191}$ $^{+1.139}_{-1.145}$
20	1	1	750+	750+	0.000 $^{+0.947}_{-0.000}$ $^{+0.000}_{-0.000}$
21	1	2	200-350	250-400	206.079 $^{+15.422}_{-15.369}$ $^{+31.712}_{-33.349}$
22	1	2	200-350	400-500	20.660 $^{+7.708}_{-7.619}$ $^{+4.371}_{-4.510}$
23	1	2	200-350	500-600	1.513 $^{+1.287}_{-0.789}$ $^{+0.492}_{-0.499}$
24	1	2	200-350	600-750	0.000 $^{+1.780}_{-0.000}$ $^{+0.000}_{-0.000}$
25	1	2	200-650	750+	0.531 $^{+1.279}_{-0.553}$ $^{+0.564}_{-0.566}$
26	1	2	350-450	250-400	18.258 $^{+4.157}_{-3.971}$ $^{+4.580}_{-4.690}$
27	1	2	350-450	400-500	2.063 $^{+2.241}_{-1.245}$ $^{+0.814}_{-0.819}$
28	1	2	350-450	500-600	0.882 $^{+1.305}_{-0.652}$ $^{+0.416}_{-0.418}$
29	1	2	350-450	600-750	0.000 $^{+0.967}_{-0.000}$ $^{+0.000}_{-0.000}$
30	1	2	450-650	250-400	2.596 $^{+1.785}_{-1.419}$ $^{+1.502}_{-1.508}$
31	1	2	450-650	400-500	5.126 $^{+2.345}_{-2.103}$ $^{+1.741}_{-1.772}$
32	1	2	450-650	500-600	0.769 $^{+0.771}_{-0.463}$ $^{+0.295}_{-0.297}$
33	1	2	450-650	600-750	1.356 $^{+1.183}_{-0.874}$ $^{+0.984}_{-0.986}$
34	1	2	650+	250-600	0.000 $^{+1.431}_{-0.000}$ $^{+0.000}_{-0.000}$
35	1	2	650+	600-750	0.000 $^{+1.508}_{-0.000}$ $^{+0.000}_{-0.000}$
36	1	2	650+	750+	0.000 $^{+0.867}_{-0.000}$ $^{+0.000}_{-0.000}$

Table 4.16: Predicted lost lepton background yield from the electron control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.

37	1	3+	300-1000	250-350	31.043	+5.945	+6.260
						-5.752	-6.448
38	1	3+	300-1000	350-450	3.717	+2.068	+0.994
						-1.638	-1.009
39	1	3+	300-1000	450-550	0.986	+1.516	+0.471
						-0.727	-0.474
40	1	3+	300-1000	550+	1.879	+2.434	+2.111
						-1.957	-2.113
41	1	3+	1000-1500	250-350	1.291	+1.490	+0.319
						-1.029	-0.328
42	1	3+	1000-1500	350-450	2.129	+1.606	+0.753
						-1.150	-0.764
43	1	3+	1000-1500	450-550	0.000	+1.263	+0.000
						-0.000	-0.000
44	1	3+	1000-1500	550+	6.040	+6.445	+6.336
						-6.292	-6.341
45	1	3+	1500+	250-350	1.452	+1.395	+0.486
						-0.922	-0.493
46	1	3+	1500+	350-550	0.436	+1.303	+0.268
						-0.454	-0.269
47	1	3+	1500+	550+	0.641	+1.405	+0.331
						-0.668	-0.333

of the detector to the muon or the τ_h jet. The trick consists of replacing the muon p_T by a random simulated τ_h jet response, obtained from a "template" function for a hadronically decaying τ lepton. The global variables of the event are recalculated after including this τ_h jet, and the search selections are applied to predict the τ_h background.

As shown in Fig 4.31, the template is measured in four p_T bins to account for the p_T dependence of τ jet response.

The τ_h background prediction is calculated as follows:

$$N_{\tau_h} = \sum_i^{N_{\text{CS}}^\mu} \left(\sum_j^{\text{Template bins}} (P_{\tau_h}^{\text{resp}}) \frac{\epsilon_{\tau \rightarrow \mu}}{\epsilon_{\text{trigger}}^\mu \epsilon_{\text{reco}}^\mu \epsilon_{\text{iso}}^\mu \epsilon_{\text{acc}}^\mu \epsilon_{m_T}^\mu} \frac{\mathcal{B}(W \rightarrow \tau_h)}{\mathcal{B}(W \rightarrow \mu)} \epsilon_{\text{isotrack}} F_{\text{dilepton}} \right), \quad (4.11)$$

where the first summation is over the events in the μ + jets control sample, the second is over the bins of the τ_h response template and $P_{\tau_h}^{\text{resp}}$ is the probability of the τ_h response in each bin.

The classical hadronic tau method is not applied in this analysis because of the issue in the HT bit L1 trigger in 2016 run H data.

Table 4.17: Predicted lost lepton background yield from the electron control sample with statistical and systematic uncertainties for a 35.9 fb^{-1} sample.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	M_{T2} [GeV]	p_T^{miss} [GeV]	Lost Lepton Prediction
48	2	1	200-300	250-350	18.648 $^{+5.730}_{-5.687}$ $^{+3.582}_{-3.780}$
49	2	1	200-300	350-450	3.668 $^{+2.018}_{-1.825}$ $^{+1.912}_{-1.929}$
50	2	1	200-300	450-600	0.737 $^{+0.826}_{-0.573}$ $^{+0.758}_{-0.760}$
51	2	1	200-450	600+	0.434 $^{+1.201}_{-0.452}$ $^{+0.253}_{-0.254}$
52	2	1	300-450	250-350	11.134 $^{+3.300}_{-2.928}$ $^{+2.624}_{-2.683}$
53	2	1	300-450	350-450	6.168 $^{+2.709}_{-2.362}$ $^{+2.553}_{-2.575}$
54	2	1	300-450	450-600	2.109 $^{+2.392}_{-2.197}$ $^{+1.236}_{-1.246}$
55	2	1	450+	250-450	0.593 $^{+1.296}_{-0.617}$ $^{+0.287}_{-0.289}$
56	2	1	450+	450-600	0.000 $^{+1.693}_{-0.000}$ $^{+0.000}_{-0.000}$
57	2	1	450+	600+	0.000 $^{+1.245}_{-0.000}$ $^{+0.000}_{-0.000}$
58	2	2	200-300	250-350	15.625 $^{+3.591}_{-3.533}$ $^{+3.005}_{-3.178}$
59	2	2	200-300	350-450	2.661 $^{+1.242}_{-0.938}$ $^{+1.244}_{-1.255}$
60	2	2	200-300	450-600	0.252 $^{+0.676}_{-0.262}$ $^{+0.116}_{-0.118}$
61	2	2	200-400	600+	0.593 $^{+1.032}_{-0.618}$ $^{+0.409}_{-0.411}$
62	2	2	300-400	250-350	8.563 $^{+2.784}_{-2.533}$ $^{+1.924}_{-1.980}$
63	2	2	300-400	350-450	1.073 $^{+1.422}_{-0.797}$ $^{+0.410}_{-0.416}$
64	2	2	300-400	450-600	2.377 $^{+2.394}_{-1.905}$ $^{+1.315}_{-1.320}$
65	2	2	400-500	250-450	0.469 $^{+1.188}_{-0.489}$ $^{+0.173}_{-0.175}$
66	2	2	400-500	450-600	0.476 $^{+1.333}_{-0.496}$ $^{+0.276}_{-0.278}$
67	2	2	400+	600+	0.000 $^{+1.117}_{-0.000}$ $^{+0.000}_{-0.000}$
68	2	2	500+	250-450	0.000 $^{+0.726}_{-0.000}$ $^{+0.000}_{-0.000}$
69	2	2	500+	450-600	0.000 $^{+1.557}_{-0.000}$ $^{+0.000}_{-0.000}$
70	2	3+	300-900	250-350	1.885 $^{+1.242}_{-0.900}$ $^{+0.570}_{-0.579}$
71	2	3+	300-900	350-500	0.000 $^{+1.072}_{-0.000}$ $^{+0.000}_{-0.000}$
72	2	3+	300-1300	500+	0.000 $^{+1.327}_{-0.000}$ $^{+0.000}_{-0.000}$
73	2	3+	900-1300	250-350	1.131 $^{+1.450}_{-0.839}$ $^{+0.464}_{-0.469}$
74	2	3+	900-1300	350-500	0.000 $^{+0.947}_{-0.000}$ $^{+0.000}_{-0.000}$
75	2	3+	1300+	250-350	1.468 $^{+1.145}_{-0.915}$ $^{+0.611}_{-0.617}$
76	2	3+	1300+	350-500	0.738 $^{+0.859}_{-0.570}$ $^{+0.584}_{-0.586}$
77	2	3+	1300+	500+	0.000 $^{+0.766}_{-0.000}$ $^{+0.000}_{-0.000}$
78	3+	1	200+	250-350	0.000 $^{+0.887}_{-0.000}$ $^{+0.000}_{-0.000}$
79	3+	1	200+	350+	0.366 $^{+0.804}_{-0.381}$ $^{+0.228}_{-0.230}$
80	3+	2	200+	250-400	0.181 $^{+0.634}_{-0.189}$ $^{+0.108}_{-0.109}$
81	3+	2	200+	400+	0.000 $^{+0.645}_{-0.000}$ $^{+0.000}_{-0.000}$
82	3+	3+	200+	250-350	0.000 $^{+1.156}_{-0.000}$ $^{+0.000}_{-0.000}$
83	3+	3+	200+	350+	0.000 $^{+0.571}_{-0.000}$ $^{+0.000}_{-0.000}$

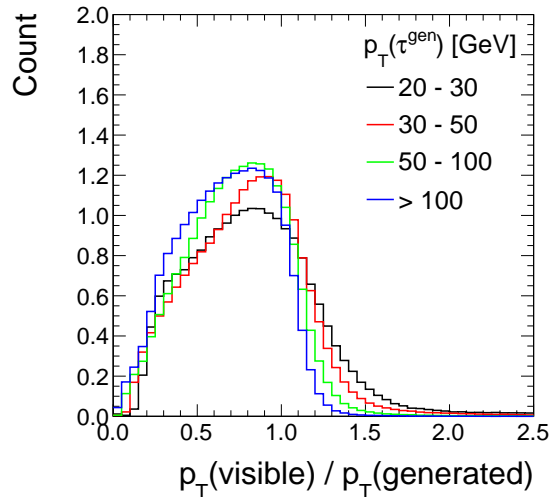


Figure 4.31: Tau jet visible energy fraction templates in tau lepton p_T bins.

4.4.2 Backgrounds from neutrinos in Z decays

An important background for any analysis with jets and large p_T^{miss} is the production of Z bosons in association with jets, where the Z boson decays to neutrinos. In this section the estimation of this background is discussed.

Ideally, we would like to use a data driven approach based on Z +jets events where the Z boson decays to a muon pair. The kinematics of these events are indistinguishable from the kinematics of events where Z decays to neutrinos. The behavior under the search region selection and the characteristics of the distributions of physics observables for both decays, including the minimum H_T and p_T^{miss} requirements or the number of b-tagged jets, would be preserved. This strategy gives a reliable background estimation, however it suffers from having a low number of events, which is due in part to the small branching ratio for

$Z \rightarrow \mu\mu$ as well as the b-tagging and other kinematic requirements placed on the control region events.

To improve on the large statistical uncertainties that would occur when using the above method, a method incorporating data-validated MC is used instead. In this multistage process the final estimate is taken from the $Z \rightarrow \nu\nu$ MC, which is corrected for data/MC differences observed in a control region with loosened selection criteria.

The central value of the $Z \rightarrow \nu\nu$ background prediction for each search bin B can be written as

$$\hat{N}_B = R_{\text{norm}} \cdot \sum_{\text{events} \in B} S_{DY}(N_{\text{jet}})w_{\text{MC}}, \quad (4.12)$$

with \hat{N}_B the predicted number of $Z \rightarrow \nu\nu$ background events in search bin B , and w_{MC} the standard MC event weight including the assumed $Z \rightarrow \nu\nu$ cross section, the data luminosity, the b tag scale factors, and the measured trigger efficiency. Each MC event is corrected using two additional scale factors. The first, R_{norm} , is an overall normalization factor for the $Z \rightarrow \nu\nu$ simulation that is derived in a tight control region in data. This tight control region has the same selection as the search region, apart from the requirement that there be two muons (treated as if they were neutrinos) and that events with any b-tagged jet multiplicity are allowed, so it is a very good proxy for the signal region. The second scale factor, S_{DY} , depends on the number of jets (N_{jet}) in the event and is derived in a loose control region in which the signal region requirements on $p_{\text{T}}^{\text{miss}}$, $M_{\text{T}2}$ and the number of top-tagged jets in the event are relaxed. The scale factor is derived separately for events with 0 and ≥ 1 b-tagged jets. It corrects both the observed mismodelling of the jet multiplicity distribution

in the simulation and the difference in normalization between data and simulation in this loose region.

This corrected MC estimate is further validated, and systematic uncertainties are assigned where appropriate. A first sanity check is to validate the $DY \rightarrow \mu\mu$ MC against the $Z \rightarrow \nu\nu$ MC. Any expected differences due to generator discrepancies, acceptance, and efficiency are corrected for. This is an important step for the overall validation of the method because it relies on the usage of dimuon events in data to do the various cross checks and derive scale factors. These scale factors are eventually applied to $Z \rightarrow \nu\nu$ MC events, hence the need for a good match between $DY \rightarrow \mu\mu$ and $Z \rightarrow \nu\nu$ MC. A second layer of validation is to use the loose control region, for which a reasonable number of events are available in data, to check the shape agreement between data and the simulated distributions. Any disagreements will be incorporated as a systematic uncertainty in the prediction. Finally, we also need to check the data/MC agreement in the loose region, where the shape systematics are assessed, versus the data/MC agreement in the tight region, which is the proxy for the region we want to predict.

The $Z \rightarrow \nu\nu$ background predictions for each search bin, including the statistical and systematic uncertainties, are shown in Fig. 4.32

4.4.3 Backgrounds from QCD multi-jet

The QCD multijet background originates in the measurement process (instrumentation) and is primarily due to mismeasurement of the energy of one or more jets in a SM multijet event. When that happens, large amounts of spurious p_T^{miss} can be present

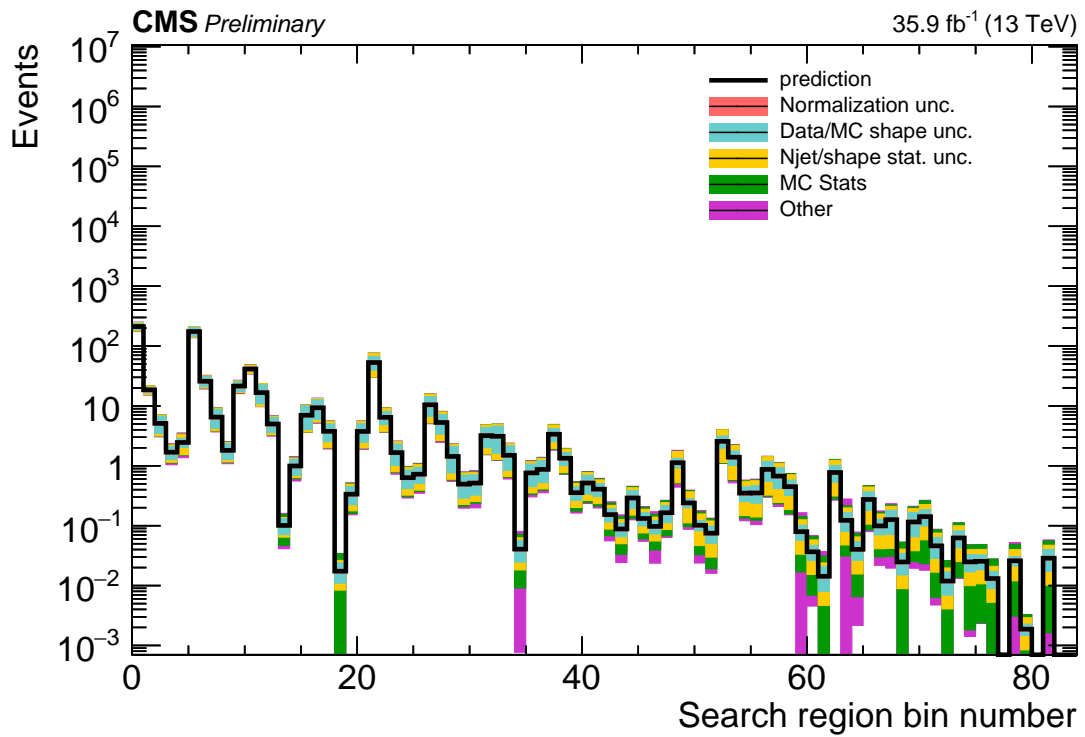


Figure 4.32: $Z \rightarrow \nu\nu$ background prediction for all search bins, including the breakdown of the various uncertainties.

in the reconstructed event, which potentially mimics an all-hadronic SUSY final state that satisfies the search selection. The probability to have such an event, where there are fake b jets and tops, is very low, but the large QCD cross section makes them more likely and therefore their contribution to the signal regions must be estimated. The expectation is for this background to be subdominant in most search bins.

MC simulation tells us a fact that the high p_T^{miss} and $\Delta\phi_{\text{jets}}$ and p_T^{miss} requirements, as well as the requirement that there be multiple jets, remove nearly all the QCD contribution. An unfortunate consequence of the excellent veto power of the selection criteria is that the QCD contribution is also very small in the sideband samples typically used to evaluate the residual QCD background contribution. The fact that these control samples are most frequently leptonic- $t\bar{t}$ dominated makes it difficult to use the more common background estimation techniques that would simply extrapolate QCD dominated distributions from the sidebands into the signal regions.

The procedure developed for the analysis described in this note consists of selecting a signal depleted data control sample rich in QCD events, from which contributions of other SM backgrounds, such as $t\bar{t}$, W +jets, Z +jets, are subtracted. These backgrounds are estimated using the same procedures described in the preceding sections. A translation or transfer factor is determined to extrapolate the number of QCD events from the control sample to the search region. Due to the small number of events in data control regions, we use MC samples to derive the translation factor, although we normalize their values to a data measurement in the $200 \text{ GeV} < p_T^{\text{miss}} < 250 \text{ GeV}$ bin, just below the signal region, where there are enough events to offer a reasonably small statistical uncertainty.

4.4.3.1 Description of the method

The sideband or QCD enriched control sample is defined by applying the full set of baseline selection requirements to the search triggers, except for the $\Delta\phi_{\text{jets}}$ and $p_{\text{T}}^{\text{miss}}$ requirements which are inverted in order to select multijet events. In other words, we require that $\Delta\phi_{1^{\text{st}}\text{jet}}$ and $p_{\text{T}}^{\text{miss}} < 0.5$, $\Delta\phi_{2^{\text{nd}}\text{jet}}$ and $p_{\text{T}}^{\text{miss}} < 0.5$ or $\Delta\phi_{3^{\text{rd}}\text{jet}}$ and $p_{\text{T}}^{\text{miss}} < 0.3$ in order to maximize the number of events with fake $p_{\text{T}}^{\text{miss}}$, which tend to be aligned with one of the highest p_T jets. Fig 4.33 depicts the topology of events passing the $\Delta\phi$ and the inverted $\Delta\phi$ cuts.

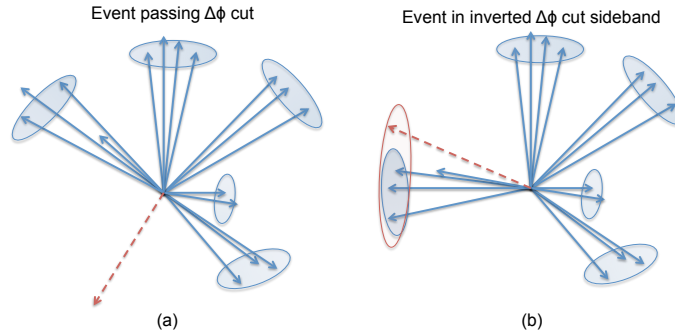


Figure 4.33: (a) Example of an event passing the $\Delta\phi_{\text{jets}}$ and $p_{\text{T}}^{\text{miss}}$ cut. $p_{\text{T}}^{\text{miss}}$ is well separated from the leading three jets. (b) Example of an event failing the $\Delta\phi_{\text{jets}}$ and $p_{\text{T}}^{\text{miss}}$ cut. $p_{\text{T}}^{\text{miss}}$ is well aligned with one of the leading jets and most likely arises from jet mismeasurement.

Although the contribution of QCD multijet events is not negligible in this control sample, it is still far from dominant. MC studies in the Fig 4.34 show the non-QCD background contribution to the control sample.

To assume that the events in the sideband sample are all QCD multijet events would lead to a gross overestimation of the total QCD background, after the translation

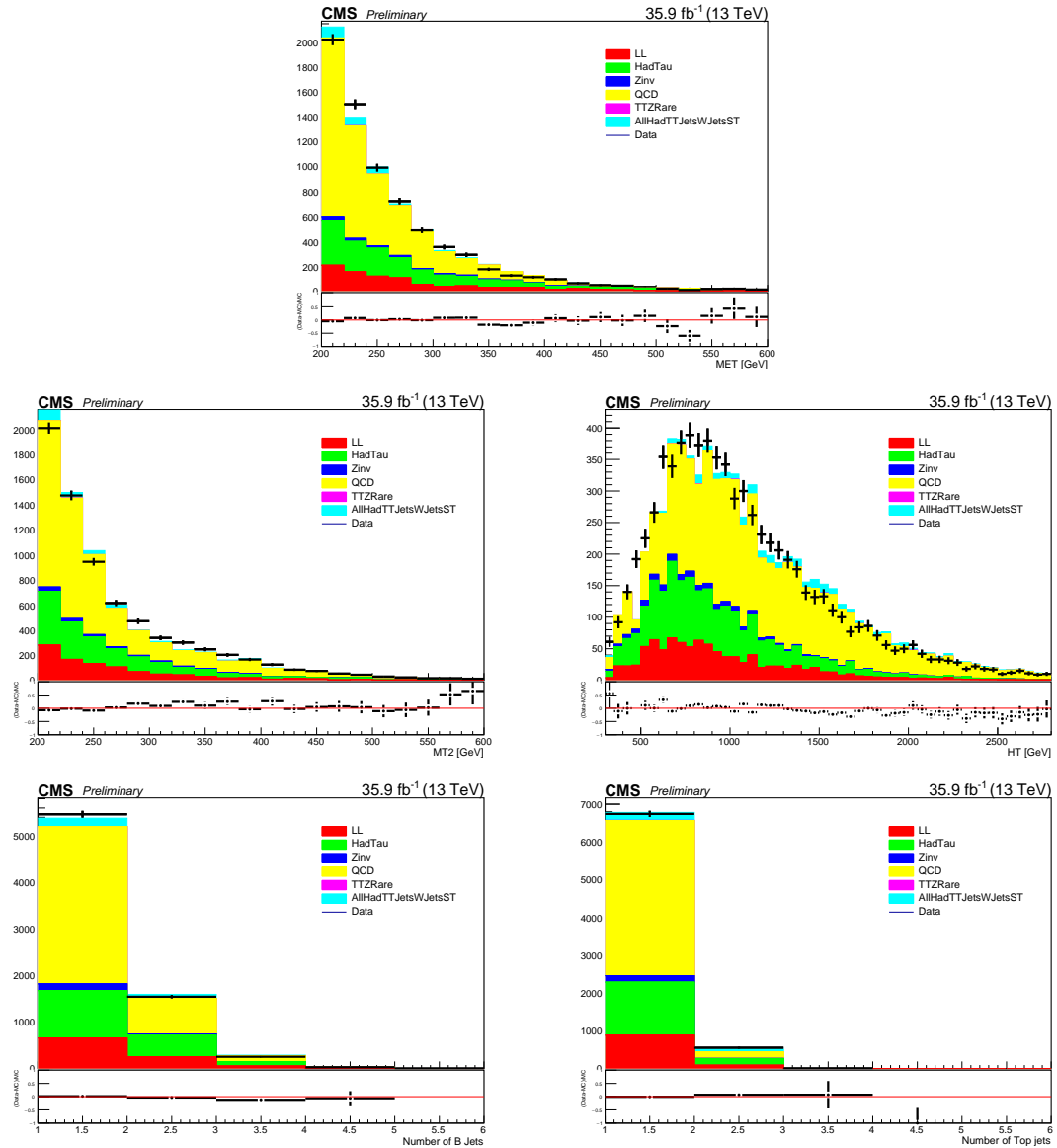


Figure 4.34: Basic kinematic distributions in the QCD control sample. The uncertainty relates to the total statistical uncertainty. The variables are p_T^{miss} on top, M_{T2} (left) and H_T (right) in middle and number of b-jets (left) and number of top candidates (right) on bottom.

factor is applied. Therefore we predict and subtract the contributions from lost leptons (LL), hadronic τ 's (τ_h), and Z +jets ($Z \rightarrow \nu\nu$) processes from the number of data events counted in the inverted $\Delta\phi_{\text{jets}}$ and $p_{\text{T}}^{\text{miss}}$ sample. The remaining events make up the QCD contribution, $N_{QCD}^{\Delta\bar{\phi}}$, is calculated as:

$$N_{QCD}^{\Delta\bar{\phi}} = N_{Data}^{\Delta\bar{\phi}} - N_{LL}^{\Delta\bar{\phi}} - N_{\tau_h}^{\Delta\bar{\phi}} - N_{Z\rightarrow\nu\nu}^{\Delta\bar{\phi}}, \quad (4.13)$$

where $N_X^{\Delta\bar{\phi}}$ is the number of type X events in the inverted $\Delta\phi_{\text{jets}}$ and $p_{\text{T}}^{\text{miss}}$ sideband. The contributions $N_{LL}^{\Delta\bar{\phi}}$, $N_{\tau_h}^{\Delta\bar{\phi}}$ and $N_{Z\rightarrow\nu\nu}^{\Delta\bar{\phi}}$ are estimated using the same methods as described in the previous sections.

The translation factor, T_{QCD}^{MC} , is defined as the ratio of the MC predictions for the $\Delta\phi$ and $\Delta\bar{\phi}$ samples:

$$T_{QCD}^{MC} = \frac{N_{MC-QCD}^{\Delta\phi}}{N_{MC-QCD}^{\Delta\bar{\phi}}}, \quad (4.14)$$

while the final QCD background prediction in the search regions is calculated as:

$$N_{QCD}^{SR} = N_{QCD}^{\Delta\bar{\phi}} \times T_{QCD}^{Scale}, \quad (4.15)$$

where $N_{QCD}^{\Delta\bar{\phi}}$ comes from data (as defined in Eq. 4.13). The T_{QCD}^{Scale} term is the T_{QCD}^{MC} normalized to a translation factor measured in the $200 \text{ GeV} < p_{\text{T}}^{\text{miss}} < 250 \text{ GeV}$ sideband from data. This normalization provides a more accurate estimation of the true translation factors because although we trust (within the assigned uncertainties) the shape of the MC distributions utilized to calculate them, we do not trust their absolute values, which are corrected using the low $200 \text{ GeV} < p_{\text{T}}^{\text{miss}} < 250 \text{ GeV}$ sideband T_{QCD}^{Data} measurement.

Technically speaking, the T_{QCD}^{Scale} is based on MC only in what relates to its dependence on the search variables (shape), and based on data in what relates to its normalization.

The procedure to derive the translation factors is the following:

- Calculate T_{QCD}^{MC} from QCD MC
- Measure T_{QCD}^{Data} from Data in low p_T^{miss} sideband
- Measure T_{QCD}^{Scale} by normalizing the T_{QCD}^{MC} versus p_T^{miss} functions using the sideband T_{QCD}^{Data} factors measured in real data from the $200 \text{ GeV} < p_T^{\text{miss}} < 250 \text{ GeV}$ bin.

To summarize, we have three sets of translation factors:

- **MC** T_{QCD}^{MC} , calculated from QCD MC and used to evaluate the non-closure systematic uncertainty
- **Scaled** T_{QCD} , i.e., T_{QCD}^{Scale} are factors applied to get final QCD background predictions.

The T_{QCD}^{MC} 's are shown in Fig 4.35 and Table 4.18 (QCD HT-binned MC samples).

The error bars are statistical uncertainties only.

Table 4.18: T_{QCD}^{MC} versus p_T^{miss} distributions for N_{tops} or $N_{\text{b-jets}} \geq 2$ requirements.

$p_T^{\text{miss}} [200,250]$	$p_T^{\text{miss}} [250,Inf]$
0.113	0.095

The low p_T^{miss} T_{QCD}^{Data} are measured from data using the following relation:

$$T_{QCD}^{Data} = \frac{N_{Data}^{\Delta\phi} - N_{MC-LL}^{\Delta\phi} - N_{MC-\tau_h}^{\Delta\phi} - N_{MC-Z\rightarrow\nu\nu}^{\Delta\phi} - N_{MC-TTZ,Rare}^{\Delta\phi}}{N_{Data}^{\Delta\bar{\phi}} - N_{MC-LL}^{\Delta\bar{\phi}} - N_{MC-\tau_h}^{\Delta\bar{\phi}} - N_{MC-Z\rightarrow\nu\nu}^{\Delta\bar{\phi}} - N_{MC-TTZ,Rare}^{\Delta\bar{\phi}}} . \quad (4.16)$$

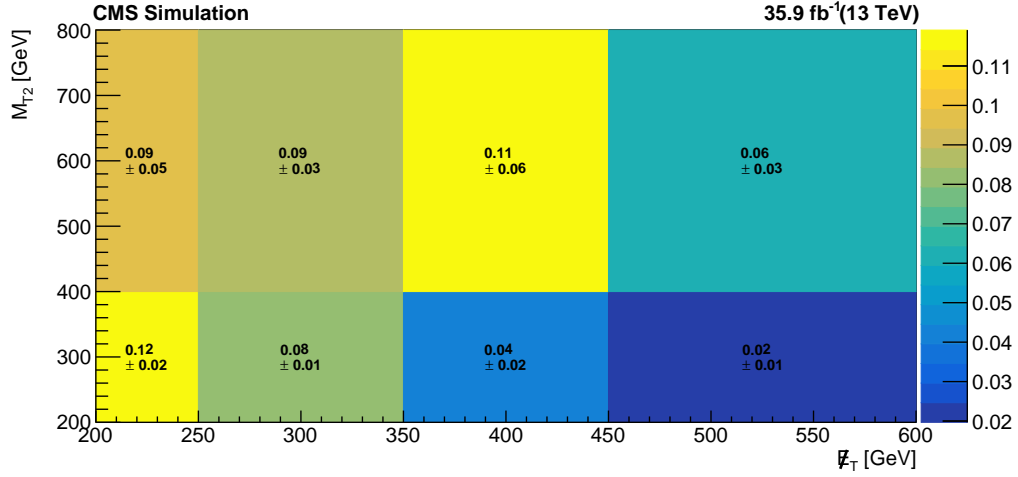


Figure 4.35: T_{QCD}^{MC} versus p_T^{miss} and M_{T2} distributions for N_{tops} and $N_{\text{b-jets}} < 3$ requirements. Both the sideband point, $200 \text{ GeV} < p_T^{\text{miss}} < 250 \text{ GeV}$, and the signal region are included.

The results for T_{QCD}^{Scale} are shown in Fig 4.36 and Table 4.19. The low p_T^{miss} scaled T_{QCD}^{Data} are calculated using data and MC samples. The $t\bar{t}$, W +jets and single top MC events are reweighted with a scale factor, whose value is 84%. The scale factor is derived from the data/MC comparison in the single-muon control samples, with ISR reweighting and the fake b scale factor. The Z+jets event yields are calculated in the same way described in the last section. The TTZ and Rare events are obtained directly from MC samples, with generator-level leptonically decaying W and Z bosons vetoed. The error bars are the total uncertainties, obtained by adding in quadrature the statistical uncertainties and the uncertainty in T_{QCD}^{Data} . Since negative T_{QCD}^{Scale} factors are not physical, we set the background prediction to zero when this is consistent within the uncertainties.

The trigger effects are taken into account when calculating the translation factors. We measure the trigger efficiencies in low/high H_T and $\Delta\phi$ /inverted $\Delta\phi$ data samples, and

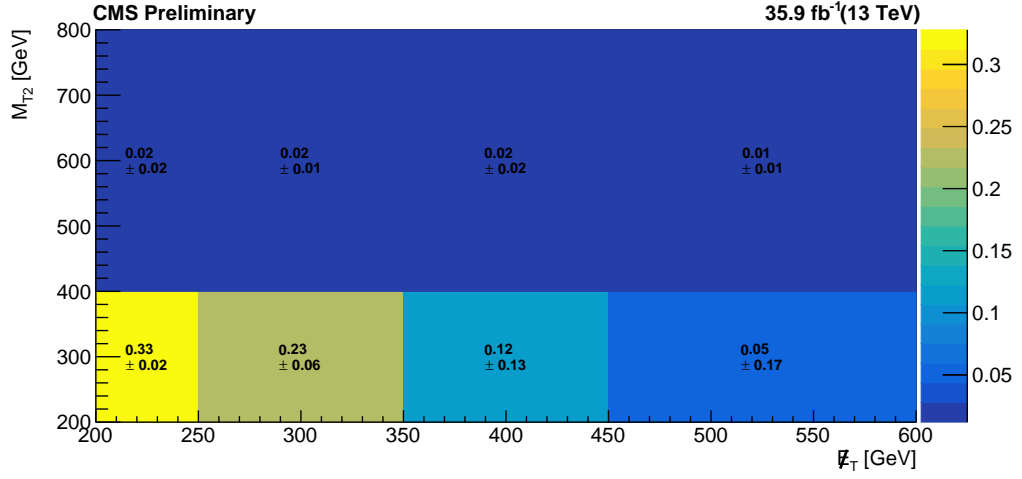


Figure 4.36: T_{QCD}^{Scale} versus p_T^{miss} and M_{T2} distributions Both the sideband point, $200 \text{ GeV} < p_T^{miss} < 250 \text{ GeV}$, and the signal region are included.

Table 4.19: T_{QCD}^{Scale} versus p_T^{miss} distributions for N_{tops} or $N_{\text{b-jets}} \geq 3$ requirements.

p_T^{miss} [200,250]	p_T^{miss} [250,Inf]
0.312	0.261

then apply them into translation factors calculation, both for the MC shape and for the low p_T^{miss} sideband.

4.4.3.2 Closure Test

The closure test is the comparison between the MC expectation after the baseline selection and the T_{QCD}^{MC} prediction. It is not a circular test since we measure the T_{QCD}^{MC} without binning in N_{tops} and $N_{\text{b-jets}}$. The closure test aims to test whether the T_{QCD}^{MC} method works on the QCD MC sample by comparing the distribution of expectation and T_{QCD}^{MC} prediction in QCD MC. As seen in Fig 4.37, the closure is reasonable for all search bin variables. The non-closure uncertainty is evaluated in the following section.

QCD closure tests with T_{QCD}^{MC} applied to all search bins are shown in Fig 4.38. The uncertainties shown in the plot are statistical only.

4.4.3.3 Systematic Uncertainties

The sources of QCD background systematical uncertainty are:

- T_{QCD}^{Scale}
- Non-closure
- Contamination from other backgrounds

4.4.3.3.1 T_{QCD}^{Scale} uncertainties The uncertainty of the T_{QCD}^{Scale} factors are the statistical uncertainties when we calculate T_{QCD}^{MC} and T_{QCD}^{Data} .

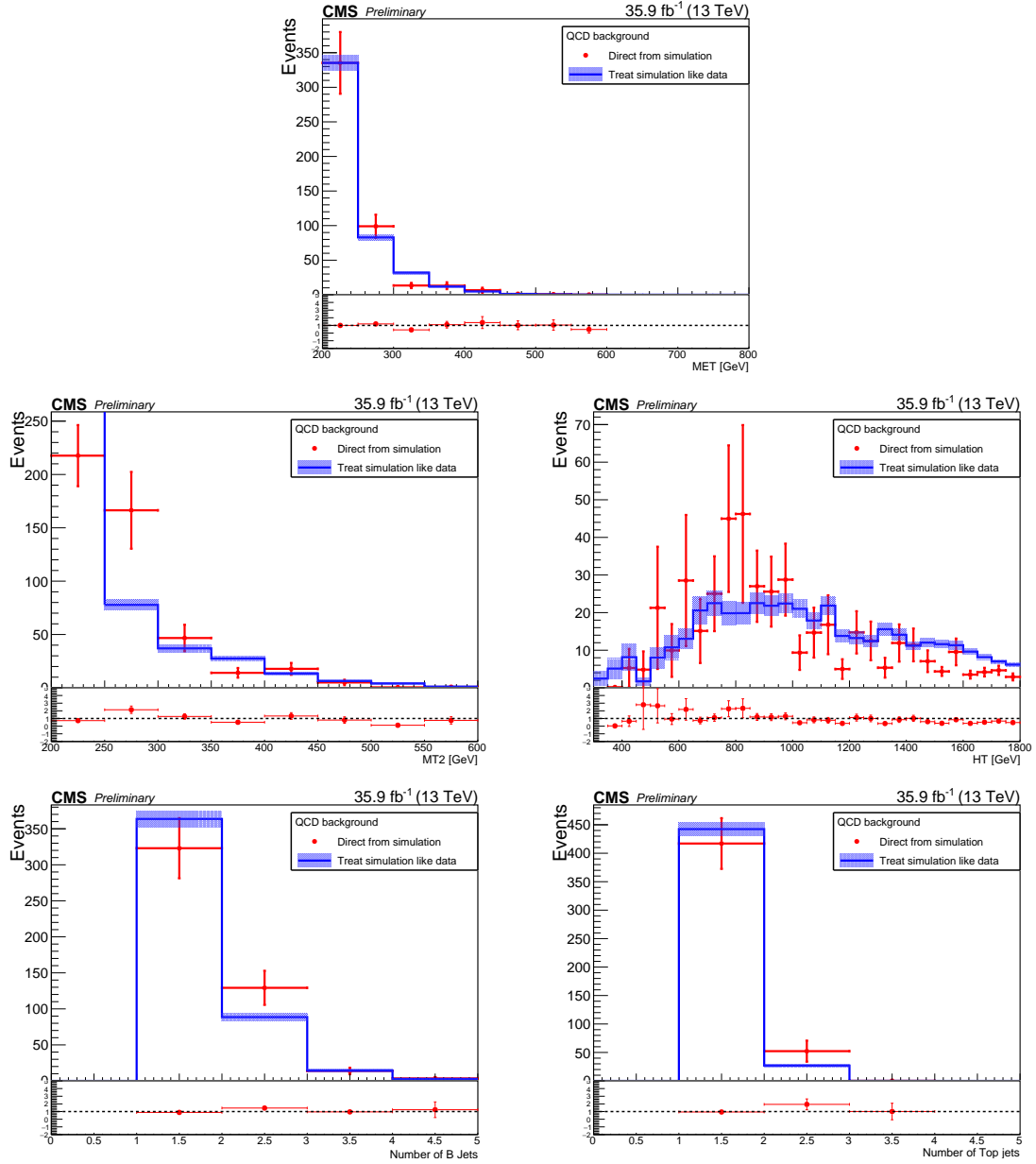


Figure 4.37: Closure test of the method: prediction compared to expectation in H_T binned QCD MC samples. The uncertainty relates to the total statistical uncertainty of the prediction. The shown variables are p_T^{miss} on top, M_{T2} (left) and H_T (right) in middle and number of b jets (left) and number of top jets (right) on bottom.

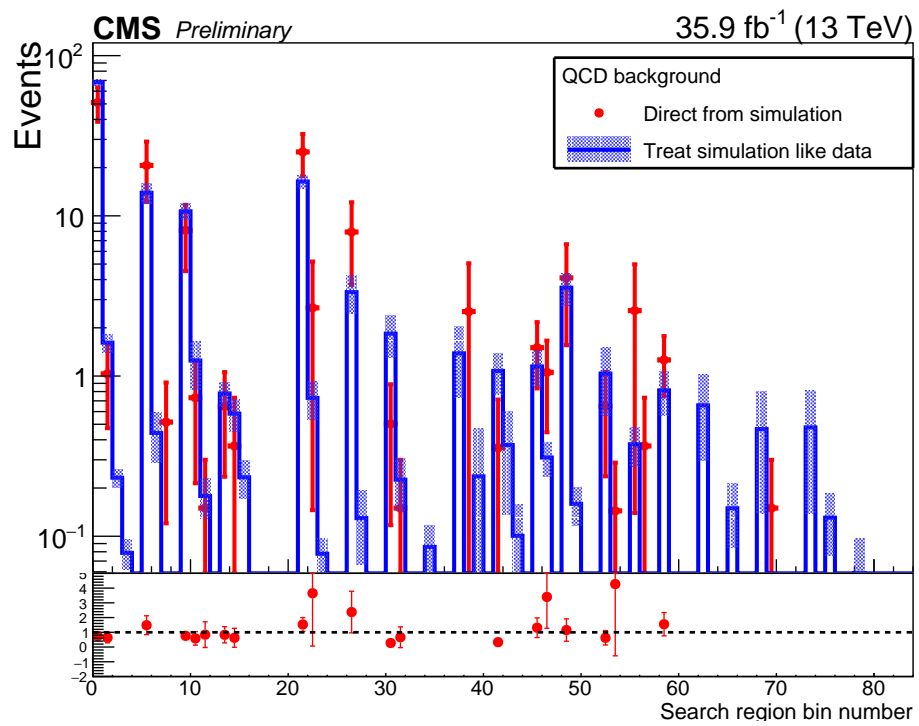


Figure 4.38: QCD background Closure tests in all search bins with T_{QCD}^{MC} .

4.4.3.3.2 Non-Closure uncertainty For the evaluation of the systematic uncertainty, the prediction is derived from the T_{QCD}^{MC} 's. The expectations and predictions are compatible in most of the search bins. In the case of bins without enough (< 4) expected MC events, we take the uncertainties from p_T^{miss} and M_{T2} or p_T^{miss} and H_T 2 dimensional inclusive closure plots, then combine them with 1 dimensional N_{tops} and $N_{\text{b-jets}}$ closure results in quadrature. If there are not enough MC events in p_T^{miss} and M_{T2} or p_T^{miss} and H_T 2 dimensional inclusive closure plots, we will take 4 independent closure results from search bin variables inclusive closure plots and then combine them in quadrature. However, there are no MC events in some bins, such as when p_T^{miss} greater than 700 GeV. In these case, we use the uncertainty from the neighboring bin.

4.4.3.3.3 Contamination uncertainty Since we are using input from other backgrounds to estimate the contamination to the QCD control region, we also incorporate this into the total systematic uncertainty.

Table 4.20: Contributions from different sources of systematic uncertainty to the QCD background prediction.

Process	Source	Effect on QCD Prediction in %
T_{QCD}^{Scale} factors	Statistical uncertainty on T_{QCD}^{MC} and T_{QCD}^{Data}	30 to 330
Closure	Non-closure and statistical precision of the closure	30 to 500
Contamination from other backgrounds	Uncertainties from other background input	2 to 50

4.4.3.4 QCD background prediction

The QCD background predictions for the full 35.9 fb^{-1} dataset in all search bins are shown in Fig 4.39 and listed in Tables 4.21, 4.22 and 4.23. The uncertainties include both statistical and systematic terms.

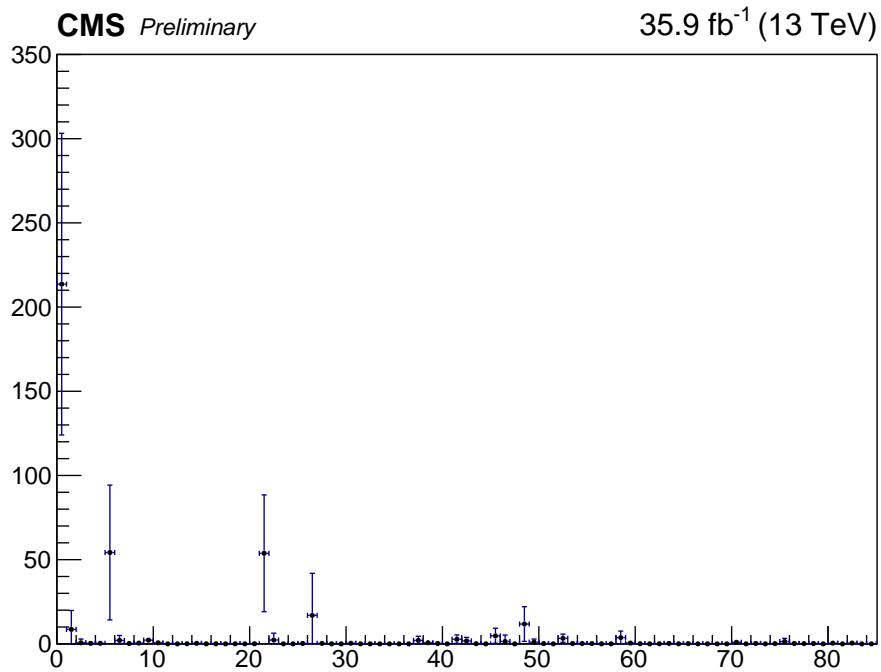


Figure 4.39: QCD background predictions in all search bins.

4.4.4 Backgrounds from TTZ and other standard model rare process

Besides the dominant backgrounds discussed above, other SM backgrounds with small cross sections were also considered and estimated for this analysis. These include the

Table 4.21: Predicted QCD backgrounds corresponding to the full 35.9 fb^{-1} data sample in first 37 search bins.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	$M_{\text{T}2}$ [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	QCD Prediction
0	1	1	200-300	250-400	213.6 $+8.556$ $+89.18$ -8.324 -89.18
1	1	1	200-300	400-500	8.587 $+1.443$ $+11.12$ -1.318 -11.12
2	1	1	200-300	500-600	0.644 $+0.339$ $+2.180$ -0.279 -2.180
3	1	1	200-300	600-750	0.300 $+0.254$ $+1.031$ -0.192 -1.031
4	1	1	200-550	750+	0.327 $+0.239$ $+1.111$ -0.176 -1.111
5	1	1	300-400	250-400	54.22 $+4.673$ $+39.76$ -4.437 -39.76
6	1	1	300-400	400-500	2.029 $+0.767$ $+2.842$ -0.637 -2.842
7	1	1	300-400	500-600	0.212 $+0.222$ $+1.141$ -0.159 -1.141
8	1	1	300-400	600-750	0.442 $+0.239$ $+1.485$ -0.176 -1.485
9	1	1	400-550	250-400	2.214 $+0.250$ $+1.397$ -0.233 -1.397
10	1	1	400-550	400-500	0.621 $+0.170$ $+0.640$ -0.146 -0.640
11	1	1	400-550	500-600	0.000 $+0.027$ $+0.005$ -0.009 -0.005
12	1	1	400-550	600-750	0.007 $+0.031$ $+0.012$ -0.015 -0.012
13	1	1	550-750	250-400	0.061 $+0.060$ $+0.053$ -0.040 -0.053
14	1	1	550-750	400-500	0.319 $+0.121$ $+0.388$ -0.097 -0.388
15	1	1	550-750	500-600	0.024 $+0.047$ $+0.039$ -0.033 -0.039
16	1	1	550-750	600-750	0.036 $+0.042$ $+0.052$ -0.028 -0.052
17	1	1	550-750	750+	0.004 $+0.027$ $+0.007$ -0.009 -0.007
18	1	1	750+	250-600	0.032 $+0.034$ $+0.170$ -0.019 -0.170
19	1	1	750+	600-750	0.001 $+0.027$ $+0.009$ -0.009 -0.009
20	1	1	750+	750+	0.055 $+0.042$ $+0.290$ -0.028 -0.290
21	1	2	200-350	250-400	53.76 $+4.982$ $+34.32$ -4.747 -34.32
22	1	2	200-350	400-500	2.278 $+0.949$ $+3.946$ -0.822 -3.946
23	1	2	200-350	500-600	0.000 $+0.204$ $+0.045$ -0.139 -0.045
24	1	2	200-350	600-750	0.000 $+0.158$ $+0.032$ -0.088 -0.032
25	1	2	200-650	750+	0.232 $+0.194$ $+0.788$ -0.128 -0.788
26	1	2	350-450	250-400	16.92 $+2.498$ $+24.85$ -2.257 -24.85
27	1	2	350-450	400-500	0.283 $+0.402$ $+0.448$ -0.257 -0.448
28	1	2	350-450	500-600	0.014 $+0.031$ $+0.021$ -0.015 -0.021
29	1	2	350-450	600-750	0.000 $+0.021$ $+0.001$ -0.000 -0.001
30	1	2	450-650	250-400	0.367 $+0.107$ $+0.303$ -0.088 -0.303
31	1	2	450-650	400-500	0.006 $+0.074$ $+0.021$ -0.047 -0.021
32	1	2	450-650	500-600	0.025 $+0.037$ $+0.039$ -0.022 -0.039
33	1	2	450-650	600-750	0.007 $+0.027$ $+0.012$ -0.009 -0.012
34	1	2	650+	250-600	0.006 $+0.027$ $+0.035$ -0.009 -0.035
35	1	2	650+	600-750	0.005 $+0.027$ $+0.032$ -0.009 -0.032
36	1	2	650+	750+	0.000 $+0.021$ $+0.003$ -0.000 -0.003

Table 4.22: Predicted QCD backgrounds corresponding to the full 35.9 fb^{-1} data sample in following 11 search bins.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	H_{T} [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	QCD Prediction
37	1	3+	300-1000	250-350	2.024 $^{+1.685}_{-1.399}$ $^{+1.807}_{-1.807}$
38	1	3+	300-1000	350-450	0.668 $^{+0.985}_{-0.674}$ $^{+0.704}_{-0.704}$
39	1	3+	300-1000	450-550	0.212 $^{+0.689}_{-0.337}$ $^{+0.259}_{-0.259}$
40	1	3+	300-1000	550+	0.000 $^{+0.481}_{-0.000}$ $^{+0.018}_{-0.018}$
41	1	3+	1000-1500	250-350	2.718 $^{+1.505}_{-1.216}$ $^{+2.073}_{-2.073}$
42	1	3+	1000-1500	350-450	1.726 $^{+1.154}_{-0.853}$ $^{+1.749}_{-1.749}$
43	1	3+	1000-1500	450-550	0.118 $^{+0.689}_{-0.337}$ $^{+0.186}_{-0.186}$
44	1	3+	1000-1500	550+	0.000 $^{+0.481}_{-0.000}$ $^{+0.077}_{-0.077}$
45	1	3+	1500+	250-350	4.770 $^{+1.505}_{-1.216}$ $^{+4.257}_{-4.257}$
46	1	3+	1500+	350-550	1.426 $^{+1.074}_{-0.769}$ $^{+3.653}_{-3.653}$
47	1	3+	1500+	550+	0.000 $^{+0.601}_{-0.216}$ $^{+0.107}_{-0.107}$

diboson (WW , WZ , ZZ) processes, multiboson (WWW , WWZ , WZZ , ZZZ) processes, the associated production with a top quark pair ($t\bar{t}H$, $t\bar{t}G$, $t\bar{t}W$, and $t\bar{t}Z$), and others (tWZ , WZG , WWG). Due to the requirement that at least one reconstructed top quark candidate be present, the backgrounds associated with top quarks are expected to have higher selection efficiency than those without top quarks. The Feynman diagrams for the dominant $t\bar{t}W$ and $t\bar{t}Z$ production mechanisms in proton-proton collisions are shown in Fig 4.40.

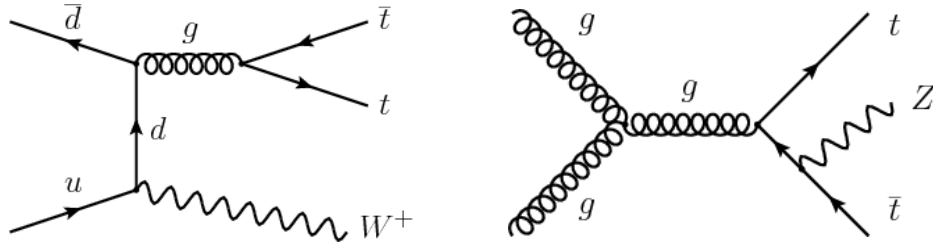


Figure 4.40: Dominant leading order Feynman diagrams for $t\bar{t}W^+$ and $t\bar{t}Z$ production at the LHC. The charge conjugate process of $t\bar{t}W^+$ produces $t\bar{t}W^-$.

The lost lepton and hadronic τ background use a data driven method, which therefore automatically takes into account the leptonically decaying W and Z components

Table 4.23: Predicted QCD backgrounds corresponding to the full 35.9 fb^{-1} data sample in remaining 36 search bins.

Search Bin	N_{tops}	$N_{\text{b-jets}}$	M_{T2} [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	QCD Prediction
48	2	1	200-300	250-350	11.75 $+2.316$ $+10.06$ -2.074 -10.06
49	2	1	200-300	350-450	1.008 $+0.619$ $+1.736$ -0.486 -1.736
50	2	1	200-300	450-600	0.161 $+0.194$ $+0.579$ -0.128 -0.579
51	2	1	200-450	600+	0.000 $+0.124$ $+0.020$ -0.044 -0.020
52	2	1	300-450	250-350	3.269 $+1.346$ $+2.106$ -1.092 -2.106
53	2	1	300-450	350-450	0.201 $+0.376$ $+0.384$ -0.227 -0.384
54	2	1	300-450	450-600	0.223 $+0.194$ $+0.831$ -0.128 -0.831
55	2	1	450+	250-450	0.169 $+0.097$ $+0.243$ -0.071 -0.243
56	2	1	450+	450-600	0.044 $+0.040$ $+0.075$ -0.025 -0.075
57	2	1	450+	600+	0.031 $+0.034$ $+0.052$ -0.019 -0.052
58	2	2	200-300	250-350	3.697 $+1.729$ $+3.479$ -1.481 -3.479
59	2	2	200-300	350-450	0.459 $+0.525$ $+0.854$ -0.388 -0.854
60	2	2	200-300	450-600	0.077 $+0.171$ $+0.290$ -0.103 -0.290
61	2	2	200-400	600+	0.000 $+0.099$ $+0.016$ -0.000 -0.016
62	2	2	300-400	250-350	0.059 $+0.822$ $+0.221$ -0.546 -0.221
63	2	2	300-400	350-450	0.258 $+0.376$ $+0.517$ -0.227 -0.517
64	2	2	300-400	450-600	0.020 $+0.124$ $+0.086$ -0.044 -0.086
65	2	2	400-500	250-450	0.138 $+0.077$ $+0.172$ -0.057 -0.172
66	2	2	400-500	450-600	0.000 $+0.021$ $+0.001$ -0.000 -0.001
67	2	2	400+	600+	0.000 $+0.021$ $+0.001$ -0.000 -0.001
68	2	2	500+	250-450	0.003 $+0.038$ $+0.009$ -0.013 -0.009
69	2	2	500+	450-600	0.004 $+0.027$ $+0.010$ -0.009 -0.010
Search Bin	N_{tops}	$N_{\text{b-jets}}$	H_{T} [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	QCD Prediction
70	2	3+	300-900	250-350	0.929 $+0.884$ $+1.221$ -0.564 -1.221
71	2	3+	300-900	350-500	0.000 $+0.481$ $+0.031$ -0.000 -0.031
72	2	3+	300-1300	500+	0.251 $+0.601$ $+0.307$ -0.216 -0.307
73	2	3+	900-1300	250-350	0.000 $+0.689$ $+0.112$ -0.337 -0.112
74	2	3+	900-1300	350-500	0.147 $+0.689$ $+0.300$ -0.337 -0.300
75	2	3+	1300+	250-350	1.348 $+0.985$ $+1.636$ -0.674 -1.636
76	2	3+	1300+	350-500	0.203 $+0.689$ $+0.316$ -0.337 -0.316
77	2	3+	1300+	500+	0.164 $+0.601$ $+0.241$ -0.216 -0.241
78	3+	1	300+	250-350	0.198 $+0.601$ $+0.273$ -0.216 -0.273
79	3+	1	300+	350+	0.000 $+0.481$ $+0.026$ -0.000 -0.026
80	3+	2	300+	250-400	0.430 $+0.689$ $+0.588$ -0.337 -0.588
81	3+	2	300+	400+	0.000 $+0.481$ $+0.028$ -0.000 -0.028
82	3+	3+	300+	250-350	0.458 $+0.689$ $+0.608$ -0.337 -0.608
83	3+	3+	300+	350+	0.000 $+0.481$ $+0.000$ -0.000 -0.000

of these backgrounds. The invisible Z background method uses simulation only to derive scale factors. So only the hadronically decaying W and Z , $Z \rightarrow \nu\nu$ components of these rare backgrounds need to be explicitly addressed. To remove the overlap between the prediction and the lost lepton and hadronic tau background prediction, a veto on the generator level leptonically decaying W and Z is applied.

4.4.4.1 $t\bar{t}Z$ Background

Similarly to the $Z \rightarrow \nu\nu + \text{jets}$ background, $t\bar{t}Z$ is an irreducible background when the Z decays to $\nu\nu$ and both top quarks decay hadronically. The $t\bar{t}Z$ cross section at 13 TeV is 782.6 pb. The predicted yield of $t\bar{t}Z$ events in the search bins is less than 10% of the total background. Given the small cross section associated with this process, we rely on simulation to generate a prediction, although this estimation is validated using data. A validation study is performed in the three-lepton channel following the $t\bar{t}Z$ normalization exercise within the SUSY group. The three-lepton channel is defined by the following selection:

- Satisfies all filters that remove detector- and beam-related noise:
 - HBHE noise filter,
 - HBHEiso noise filter,
 - ECAL dead cell trigger primitive filter,
 - Primary vertex filter,
 - Bad EE super crystal filter,
 - Global tight beam halo filter,

- Bad muon filter,
 - Bad charged hadron filter,
 - Loose JetID event filter.
- Satisfies the trigger, here we only consider the single muon trigger
 - HLT_Mu50_v*.
 - HLT_IsoMu24_v*.
 - HLT_IsoTKMu24_v*.
 - Has exactly 3 leptons
 - Highest p_T lepton, which is required by the trigger to be a muon, with $p_T > 40$ GeV
 - Leptons with the second and third largest p_T with $p_T > 20$ GeV
 - Has one reconstructed Z boson within the mass window between 81 and 101 GeV
 - At least 4 jets with $p_T > 40$ GeV
 - At least 2 b-tagged jets with $p_T > 30$ GeV
 - $p_T^{\text{miss}} > 30$ GeV

The list of MC samples for $t\bar{t}Z$ and other rare backgrounds is given in Table 4.24. The $t\bar{t}Z$ and other rare backgrounds are normalized to the SM prediction. The measured yield of $t\bar{t}Z$ and other rare backgrounds are compared with single-muon data as shown in Fig 4.41, as a function of the reconstructed Z boson mass. The observed $t\bar{t}Z$ rate is

calculated by subtracting the backgrounds yields from the data yield. The derived $t\bar{t}Z$ scale factor is 1.03 ± 0.31 . The predicted $t\bar{t}Z$ rate is consistent with data within the statistical uncertainty. Thus we don't apply a scale factor for the $t\bar{t}Z$ yield in this analysis. A 30% systematic uncertainty is assigned to the $t\bar{t}Z$ background prediction.

Table 4.24: The list of MC samples for $t\bar{t}Z$ and other rare backgrounds.

Process	Dataset names
$t\bar{t}Z$	TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	ttHToNonbb_M125_13TeV_powheg_pythia8
	VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8
	GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV6_pythia8
	ST_tWll_5f_LO_13TeV-MadGraph-pythia8
	tZq_ll_4f_13TeV-amcatnlo-pythia8_TuneCUETP8M1
	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8
	WZTo3LNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
Rare	ZZTo4L_13TeV_powheg_pythia8
	WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	WZG_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	WWG_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	WWW_4F_TuneCUETP8M1_13TeV-amcatnlo-pythia8
	TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8
NonPrompt	TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	DYJetsToLL_M-50_HT-*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8

Systematic uncertainties come from the choice of the renormalization/factorization scale and PDF in the $t\bar{t}Z$ MC samples are shown in Fig 4.42. Other systematic uncertainties including ISR reweighting, pileup correction, b -tagging scale factor are also considered.

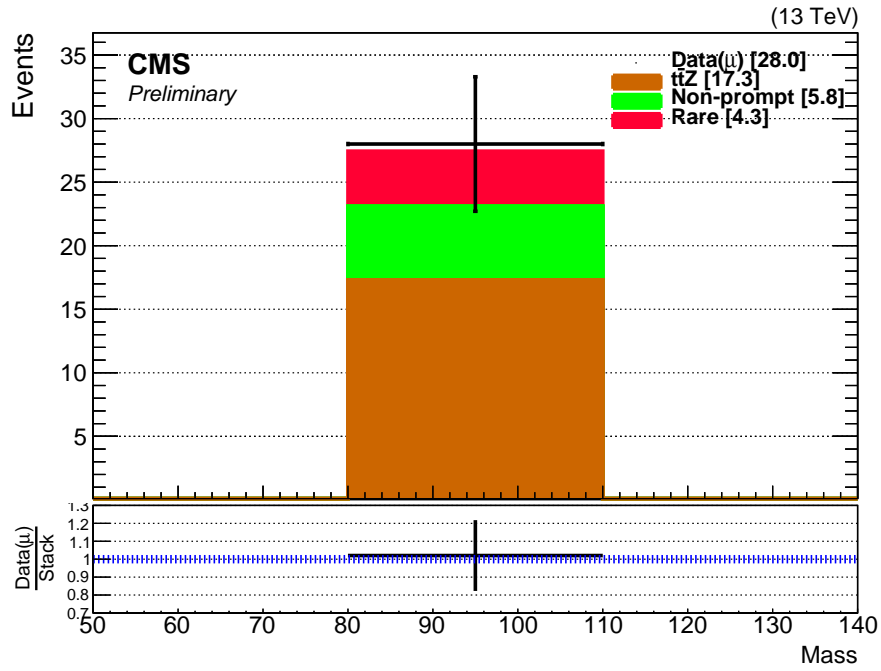


Figure 4.41: $t\bar{t}Z$ validation in the three-leptons channels from the Single Muon data. The error bar denotes the statistical uncertainty.

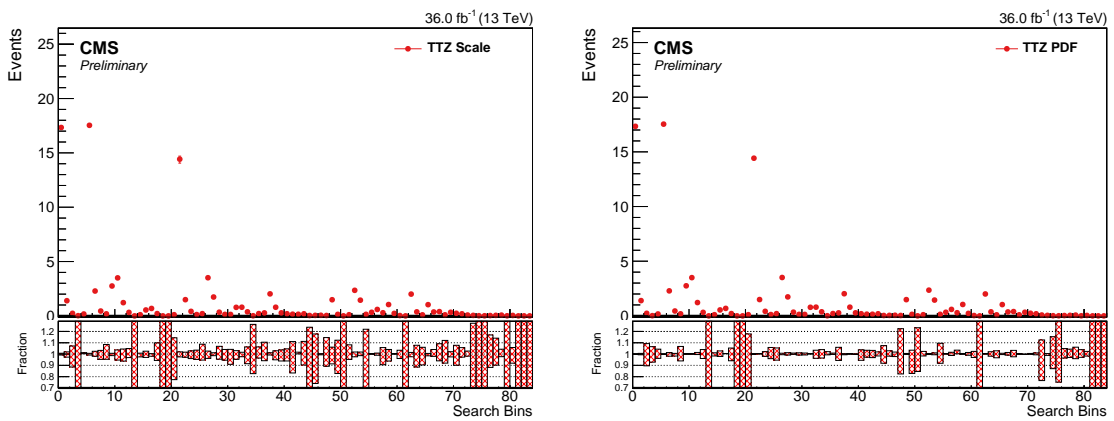


Figure 4.42: Systematic uncertainties from different sources that contribute to the $t\bar{t}Z$ background prediction, normalized to 36 fb^{-1} .

4.4.4.2 Yields of Other SM Backgrounds

The yields of the $t\bar{t}Z$ and other rare backgrounds are derived from MC samples, normalized to the SM predicted cross section. A generator-level veto of the leptonically decaying W and Z bosons is applied to avoid double-counting with the lost-lepton and hadronic tau backgrounds. Except for $t\bar{t}Z$ process, the other backgrounds are combined as the rare background. Their yield with statistical uncertainties is shown in Fig 4.43. Additional systematic uncertainties come from the choice of renormalization/factorization scale and PDF in the rare MC samples, ISR reweighting, pileup correction, and b-tagging scale factors.

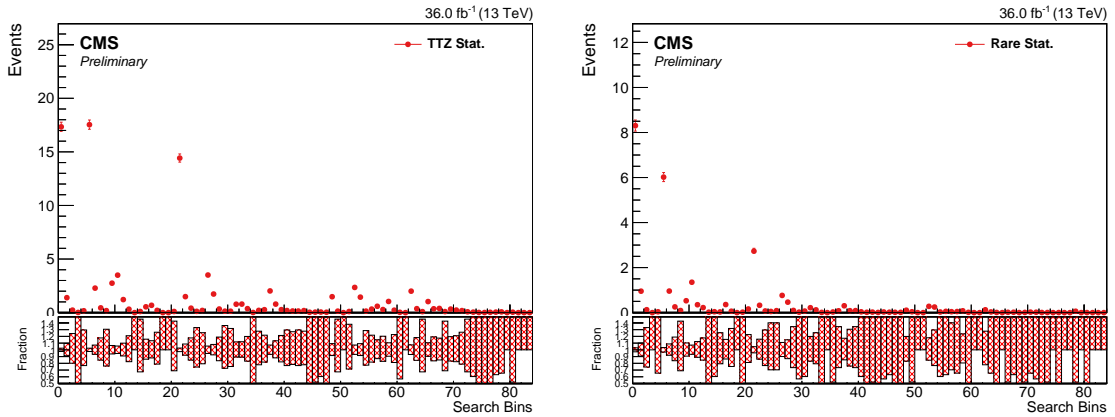


Figure 4.43: Yield of the $t\bar{t}Z$ and rare background prediction normalized to 36 fb^{-1} .

4.5 Results

The data yields and background estimation for all 84 search bins are shown in Fig 4.44. No statistically significant excess in the data above the expectation from the

standard model is observed. The $t\bar{t}$, W +jets and single top events are the largest source of the background. The Z +jets events can be dominant in the high p_T^{miss} search bins. The QCD and rare backgrounds are small in all search bins.

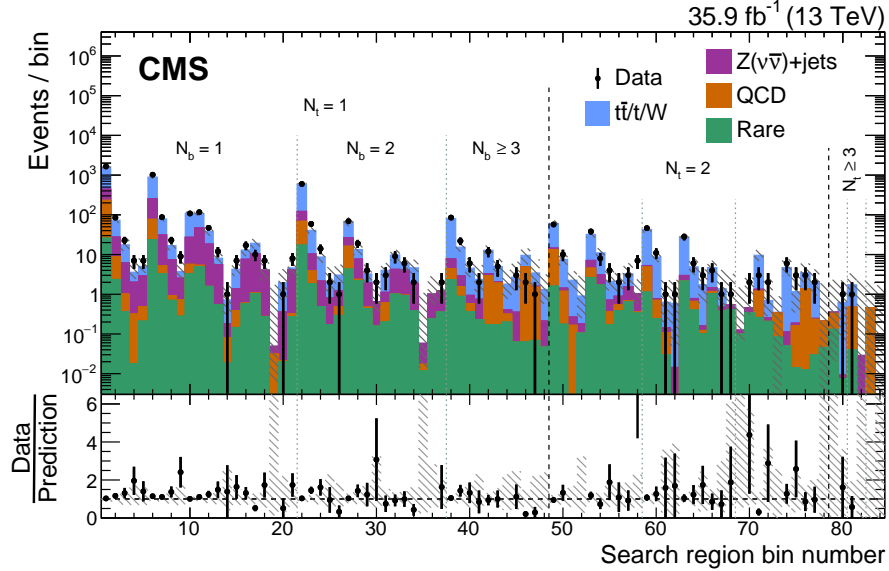


Figure 4.44: Observed event yields in data (black points) and predicted SM background (filled solid area) for the 84 search bins. The lower panel shows the ratio of data over total background prediction in each search bin. For both panels, the error bars show the statistical uncertainty associated with the observed data counts, and the grey (blue) hatched bands indicate the statistical (systematic) uncertainties in the total predicted background.

The analysis is interpreted as the upper limit of the signal model cross section. The upper limit of the signal model cross sections with 95% confidence level are calculated, using a modified frequentist (CL_s) approach [35]. The CL_s is defined in Eq 4.17:

$$CL_s = \frac{CL_{s+b}}{CL_b}, \quad (4.17)$$

where the CL_{s+b} is the confidence level in signal plus background hypothesis, and CL_b is the confidence level with the background only assumption.

In the CL_{s+b} calculation, we assume that the observed number of events n to follows a Poisson distribution with an expected value $E[n] = \mu s + b$. The s term is the mean number of events from a signal model, which is a known number from simulation. The b term is the expected number of background events, also a known number. The μ term is the signal acceptance, which is the value we need to obtain with likelihood.

The b term is the expected number of background events from all sources. We treat the backgrounds as nuisance parameters, whose values are constrained by the control samples (e.g., the single muon control sample for the lost lepton and hadronic tau background, the inverted $\Delta\phi$ control sample for the QCD background, etc.). The control samples are defined as a Poisson distribution with mean value $E[n] = \tau b$. The τ term is a scale factor between the control sample region and signal region, for example, the translation factors in the QCD case. Therefore the likelihood function for μ and b can be expressed as Eq 4.18:

$$L(\mu, b) = \prod_{i=1}^N \frac{(\mu s_i + \sum_{j=1}^M b_{ij})^{n_i}}{n_i!} e^{-(\mu s_i + \sum_{j=1}^M b_{ij})}, \quad (4.18)$$

where M is the number of different background categories. In this analysis, $M=4$: one category for the lost lepton and hadronic tau lepton background, and the other three for the Z+jets, QCD, and TTZ/rare backgrounds. N is the total number of search bins, $N = 84$ in this analysis. Wilks's theorem [70] states that the log of CL_s is asymptotically χ^2 distributed. This asymptotic method is applied in the limit setting to accelerate the process. The likelihood ratio is maximized following the procedure described in Ref [35].

We consider the following uncertainties in the signal yields:

- **MC statistics:** The statistical uncertainties in MC signal samples.

- **Luminosity:** A 2.6% flat contribution is assigned.
- **lepton veto:** Lepton vetoes are categorized as the muon and electron vetoes, the muon and electron track vetoes and the pion track veto. The numbers of signal events vetoed by each of the categories are evaluated. Then the yields are varied by the corresponding veto category uncertainties and propagated to be relative changes to the signal yields in the search bins. Here the amount of vetoed events effectively reflects the various lepton selection efficiencies. The efficiency uncertainties in the lepton selections either come from the lepton scale factor group (for the muon and electron efficiencies) or dedicated studies by either the lost lepton (for the muon and electron track efficiencies) or the hadronic tau method (for the pion track efficiencies). The scale factors from lepton scale factor group are used in this thesis.
- **b-tag efficiency:** The b-tagging and mistagging scale factor uncertainties are applied as a function of the jet p_T and η .
- **b-tag FastSim corrections:** The b-tagging and mistagging performance as derived from the fast simulation is corrected to match the full simulation predictions. Separate correction factors are derived for b-jets, c-jets, and light-flavor-jets, as a function of the jet p_T and η . As with the scale factors above, the correction factors for each type of jet are varied independently within their uncertainties and propagated to the signal bins. The correction factors and uncertainties are derived from an average mixture of $t\bar{t}$ and signal events.

- **Trigger efficiency:** The signal samples are corrected for the trigger inefficiencies. The effect of trigger efficiency uncertainties on the signal samples is at most 2.6% in the lowest p_T^{miss} bins.
- **Pileup acceptance:** The signal acceptance was found not to depend strongly on pileup, and an uncertainty is assigned to quantify this. The relative signal acceptance as a function of n_{vtx} is modeled from Monte Carlo, and is applied to the normalized n_{vtx} distribution in data. The relative signal acceptance is modeled as a linear fit in two bins, $n_{\text{vtx}} < 20$ (low PU) or $n_{\text{vtx}} \geq 20$ (high PU). A confidence band for the linear fit is derived from the uncertainty in the relative acceptance due to the statistical uncertainty in the Monte Carlo event counts. The pileup acceptance uncertainty is found by calculating the expectation value:

$$c = \sum_{n_{\text{vtx}}=0}^{100} f_{\text{MC}}(n_{\text{vtx}})g_{\text{data}}(n_{\text{vtx}}), \quad (4.19)$$

where f_{MC} is taken from the central fit value or the lower or upper limit from the confidence band. The $g_{\text{data}}(n_{\text{vtx}})$ term is measured in a single electron control region (requiring $N_{\text{jets}} \geq 2$, $n_{\text{electron}} = 1$, and `HLT_Elea27_WPTight`). The up and down variations of c are normalized to the value from the central variation of the fit. The magnitude of the uncertainty is found to be 0.2-4.1%

- **Renormalization and factorization scales:** The uncertainty is calculated using the envelope of the weights obtained by varying the renormalization and factorization scales, μ_R and μ_F , by factor of two [22, 27]. The effects on the signal shapes are considered as their uncertainties on the signal cross sections.

- **ISR:** The ISR correction is applied and its uncertainty is propagated following the recommended procedure for Moriond 2017.
- **Jet Energy Corrections:** The jet energy corrections (JEC) are varied within the p_T and η -dependent jet energy scale uncertainties available in the official database. A different set of corrections and uncertainties is used in fast simulation samples. These variations are propagated into the jet-dependent search variables, such as: $N_{\text{b-jets}}$, N_{tops} , p_T^{miss} , M_{T2} , H_T , $\Delta\phi(p_T^{\text{miss}}, j_i)$.
- p_T^{miss} **uncertainty:** To account for uncertainties in p_T^{miss} in the fast simulation, the evaluation of the signal yield is repeated using generator p_T^{miss} . The average of both yields is used as an uncertainty.
- **PDFs:** The LHC4PDF [20] prescription for the uncertainty in the total cross section is included as ± 1 sigma band in the limit plots. As recommended by the SUSY group, no additional PDF uncertainty is included.
- **Full/FastSim scale for top quark reconstruction:** We use the full simulation to fastsim scale factor as measured in Ref [68]. In the fastsim signal events for any of the reconstructed top quark candidates that are matched to a generator-level top quark, we apply the corresponding scale factor measured as a function of top quark p_T . We then propagate the statistical uncertainties on the scale factor to the signal yields in our search bins.
- **Top tagger data/MC difference:** As discussed in Ref [68], the top tagger efficiency agrees well between data and MC (full simulation) within uncertainties. We

apply the measured correction factors for each type of the mono-jet, di-jet and tri-jet reconstructed top quark candidates. We then propagate the measured uncertainties to the signal yields in our search bins.

The analysis interpretations are presented in Fig 4.45. The T2tt interpretation is not given for $|m_{\tilde{t}} - m_{\tilde{\chi}_1^0} - mt| < 25 \text{ GeV}$ and $m_{\tilde{t}} < 275 \text{ GeV}$ because the signal acceptance is difficult to model in these regions, due to the similarity between signal events and the standard model $t\bar{t}$ events in these regions.

The T5tttt interpretation is not presented for $m_{\tilde{\chi}_1^0} < 50 \text{ GeV}$ because of contamination from SM $t\bar{t}$ events with fake p_T^{miss} . 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 [19]. 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.

The expected exclusion limits on the model parameters are summarized in Table 4.25:

Table 4.25: Expected Exclusion summary table.

SMS Model	LSP mass expected exclusion	SUSY Mother mass expected exclusion
T2tt	460 GeV	1130 GeV
T1tttt	1100 GeV	2040 GeV
T1ttbb	1130 GeV	1900 GeV
T5tttt	1080 GeV	1950 GeV
T5ttcc	1120 GeV	1940 GeV

The observed exclusion limits on the model parameters are summarized in Table 4.26:

Table 4.26: Observed Exclusion summary table.

SMS Model	LSP mass observed exclusion	SUSY Mother mass observed exclusion
T2tt	400 GeV	1020 GeV
T1tttt	1100 GeV	2060 GeV
T1ttbb	1150 GeV	2000 GeV
T5tttt	1080 GeV	2000 GeV
T5ttcc	1070 GeV	1850 GeV

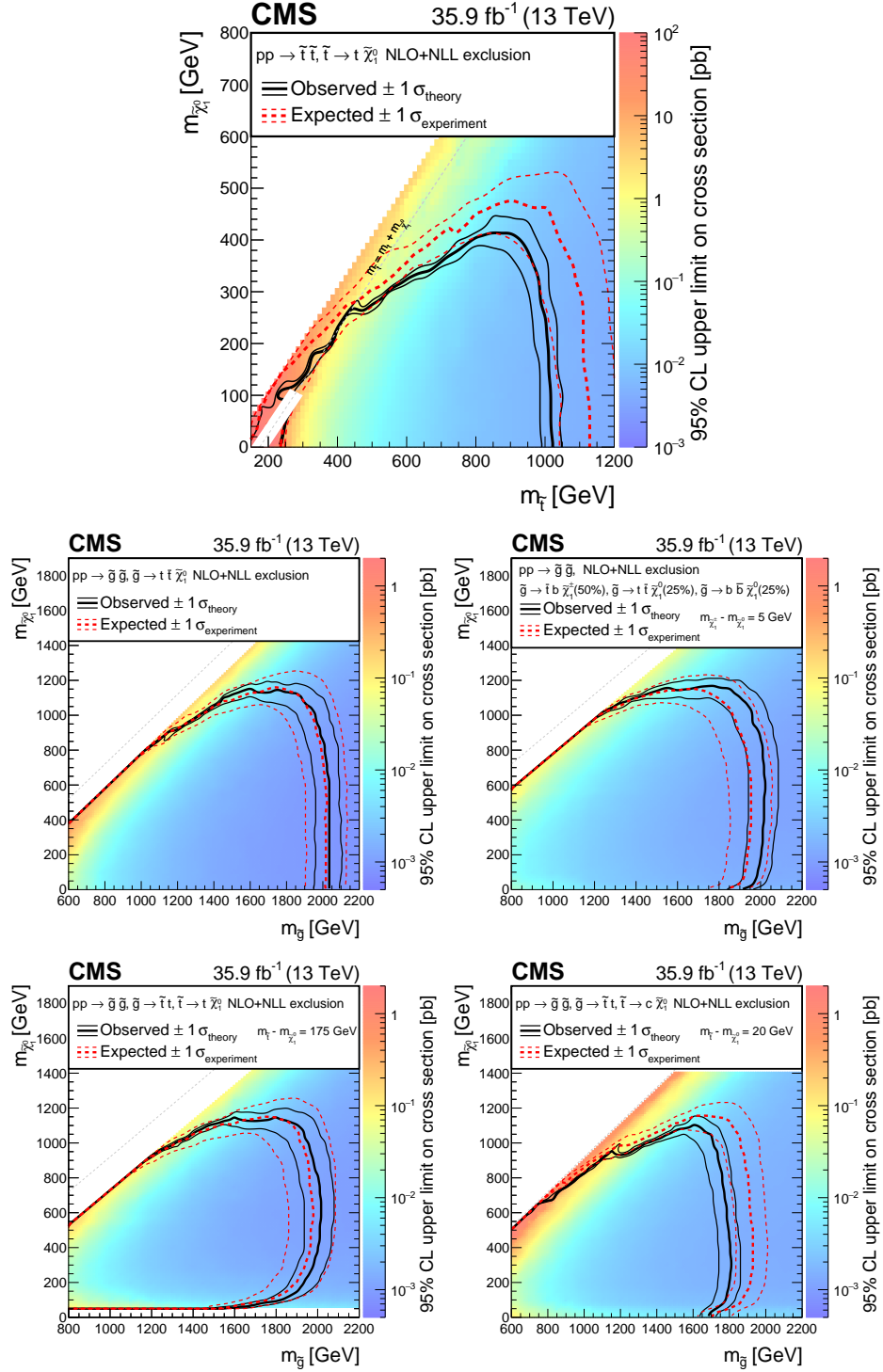


Figure 4.45: Exclusion limit at 95% CL for the signal models in this search: top squark pair production with the top squark decaying into a top quark and neutralino (top), and top squarks from cascade decays of gluinos (middle and bottom). The top is T2tt, the middle left model as T1tttt, middle right model as T1ttbb the bottom left one as T5tttt and the bottom right one as T5ttcc.

Chapter 5

Conclusions

A search for top squarks and gluinos in all-hadronic events produced in proton-proton collisions at center-of-mass energy 13 TeV has been presented. The data sample, collected in 2016 with the CMS detector at the CERN large hadron collider (LHC), corresponds to an integrated luminosity of 35.9 fb^{-1} .

The standard model backgrounds are estimated using data-driven methods as well as MC simulation. No excess of events above the expected standard model background is observed. The result is interpreted in the context of simplified supersymmetric models as 95% confidence level upper limits on the cross section of gluinos and top squarks pair production processes. The T2tt model has been excluded for top squark masses up to 1020 GeV. The corresponding exclusions on the gluino mass are up to 1810-2040 GeV, depending on the type of models. The naturalness of the SUSY MSSM RPC simplified models are under challenge with the large 13 TeV data samples now being collected at the LHC.

Bibliography

- [1] Performance of Jet Algorithms in CMS. Technical Report CMS-PAS-JME-07-003, CERN, Geneva.
- [2] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009.
- [3] Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector. Technical Report CMS-PAS-JME-16-004, CERN, Geneva, 2016.
- [4] Search for supersymmetry in the all-hadronic final state using top quark tagging in pp collisions at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-SUS-16-030, CERN, Geneva, 2016.
- [5] Search for supersymmetry in the all-hadronic final state using top quark tagging in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 96:012004, Jul 2017.
- [6] Search for supersymmetry using hadronic top quark tagging in 13 TeV pp collisions. Technical Report CMS-PAS-SUS-16-050, CERN, Geneva, 2017.
- [7] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.
- [8] W. Adam, B. Mangano, T. Speer, and T. Todorov. Track reconstruction in the CMS tracker. 2005.
- [9] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.
- [10] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 06:128, 2011.
- [11] Johan Alwall, Philip Schuster, and Natalia Toro. Simplified Models for a First Characterization of New Physics at the LHC. *Phys. Rev.*, D79:075020, 2009.
- [12] D. Amati and G. Veneziano. Preconfinement as a Property of Perturbative QCD. *Phys. Lett.*, 83B:87–92, 1979.

- [13] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983.
- [14] R W Assmann. Preliminary Beam-based specifications for the LHC collimators. Technical Report LHC-PROJECT-NOTE-277, CERN, Geneva, Jan 2002.
- [15] ATLAS and CMS Collaborations. Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.
- [16] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [17] Alan Barr, Christopher Lester, and P. Stephens. $m(T2)$: The Truth behind the glamour. *J.Phys.*, G29:2343–2363, 2003.
- [18] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J.*, C76(4):196, 2016.
- [19] Christoph Borschensky, Michael Krmer, Anna Kulesza, Michelangelo Mangano, Sanjay Padhi, Tilman Plehn, and Xavier Portell. Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13, 14, 33$ and 100 TeV. *Eur. Phys. J.*, C74(12):3174, 2014.
- [20] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43:023001, 2016.
- [21] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.
- [22] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi. The $T\bar{T}$ cross-section at 1.8 TeV and 1.96 TeV: A Study of the systematics due to parton densities and scale dependence. *JHEP*, 04:068, 2004.
- [23] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett.*, B641:57–61, 2006.
- [24] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti- $k(t)$ jet clustering algorithm. *JHEP*, 04:063, 2008.
- [25] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008.
- [26] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012.
- [27] Stefano Catani, Daniel de Florian, Massimiliano Grazzini, and Paolo Nason. Soft gluon resummation for Higgs boson production at hadron colliders. *JHEP*, 07:028, 2003.
- [28] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

- [29] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.
- [30] Serguei Chatrchyan et al. Interpretation of Searches for Supersymmetry with simplified Models. *Phys.Rev.*, D88(5):052017, 2013.
- [31] CMS Higgs Combination Group. Documentation of the roostats-based statistics tools for higgs pag. 2015.
- [32] Timothy Cohen, Eder Izaguirre, Mariangela Lisanti, and Hou Keong Lou. Jet Substructure by Accident. *JHEP*, 03:161, 2013.
- [33] The CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. *Journal of Instrumentation*, 6(11):P11002, 2011.
- [34] John C. Collins, Davison E. Soper, and George F. Sterman. Factorization of Hard Processes in QCD. *Adv. Ser. Direct. High Energy Phys.*, 5:1–91, 1989.
- [35] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: *Eur. Phys. J.*C73,2501(2013)].
- [36] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better jet clustering algorithms. *JHEP*, 08:001, 1997.
- [37] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [38] Glennys R. Farrar and Pierre Fayet. Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry. *Phys. Lett.*, 76B:575–579, 1978.
- [39] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, Apr 1998.
- [40] Rick Field. Min-Bias and the Underlying Event at the LHC. *Acta Phys. Pol. B*, 42(arXiv:1110.5530):2631–2656, Oct 2011.
- [41] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987.
- [42] Gavril Giurgiu, D. Fehling, P. Maksimovic, M. Swartz, and V. Chiochia. Pixel Hit Reconstruction with the CMS Detector. 2008.
- [43] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.
- [44] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

- [45] Tin Kam Ho. Random decision forests. In *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1*, ICDAR '95, pages 278–, Washington, DC, USA, 1995. IEEE Computer Society.
- [46] Stefan Hche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*, 2014.
- [47] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $s = 8$ TeV. *JINST*, 10(06):P06005, 2015.
- [48] Katja Klein. The Phase-1 Upgrade of the CMS Pixel Detector. Technical Report CMS-CR-2016-036, CERN, Geneva, Mar 2016.
- [49] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.
- [50] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [51] Hung-Liang Lai, Joey Huston, Zhao Li, Pavel Nadolsky, Jon Pumplin, Daniel Stump, and C. P. Yuan. Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions. *Phys. Rev.*, D82:054021, 2010.
- [52] Andrew J. Larkoski, Simone Marzani, Gregory Soyez, and Jesse Thaler. Soft Drop. *JHEP*, 05:146, 2014.
- [53] Vronique Lefbure, Sudeshna Banerjee, and I. Gonzlez. CMS Simulation Software Using Geant4. 1999.
- [54] C.G. Lester and D.J. Summers. Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders. *Phys.Lett.*, B463:99–103, 1999.
- [55] J. Mans, J. Anderson, B. Dahmes, P. de Barbaro, J. Freeman, T. Grassi, E. Hazen, J. Mans, R. Ruchti, I. Schimdt, et al. CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter. 2012.
- [56] Stephen P. Martin. A Supersymmetry primer. 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [57] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. i. *Phys. Rev.*, 122:345–358, Apr 1961.
- [58] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. ii. *Phys. Rev.*, 124:246–254, Oct 1961.

- [59] Yoichiro Nambu. Axial vector current conservation in weak interactions. *Phys. Rev. Lett.*, 4:380–382, Apr 1960.
- [60] Michele Papucci, Joshua T. Ruderman, and Andreas Weiler. Natural SUSY Endures. *JHEP*, 09:035, 2012.
- [61] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973.
- [62] Rahmat Rahmat, Rob Kroeger, and Andrea Giammanco. The fast simulation of the cms experiment. *Journal of Physics: Conference Series*, 396(6):062016, 2012.
- [63] Gavin P. Salam and Gregory Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 05:086, 2007.
- [64] Martin Schmaltz and David Tucker-Smith. Little Higgs review. *Ann. Rev. Nucl. Part. Sci.*, 55:229–270, 2005.
- [65] Albert M Sirunyan et al. Mechanical stability of the CMS strip tracker measured with a laser alignment system. *JINST*, 12(04):P04023, 2017.
- [66] Albert M Sirunyan et al. Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV. 2017.
- [67] Torbjørn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
- [68] the all-hadronic top squark search team alpha. Tagging algorithm combining resolved and merged top quark decays. *AN-16-461*, 2017.
- [69] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [70] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.*, 9(1):60–62, 1938.
- [71] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954.

Appendix A

Supplementary material for analysis

A.1 Quark-gluon discriminator study

The $p_t D$ in terms of jet η in different jet flavors is shown in Fig [A.1](#).

The $p_t D$ in terms of jet p_T in different jet flavors is shown in Fig [A.2](#).

The multiplicity in terms of jet η in different jet flavors is shown in Fig [A.3](#).

The multiplicity in terms of jet p_T in different jet flavors is shown in Fig [A.4](#).

The Axis2 in terms of jet η in different jet flavors is shown in Fig [A.5](#).

The Axis2 in terms of jet p_T in different jet flavors is shown in Fig [A.6](#).

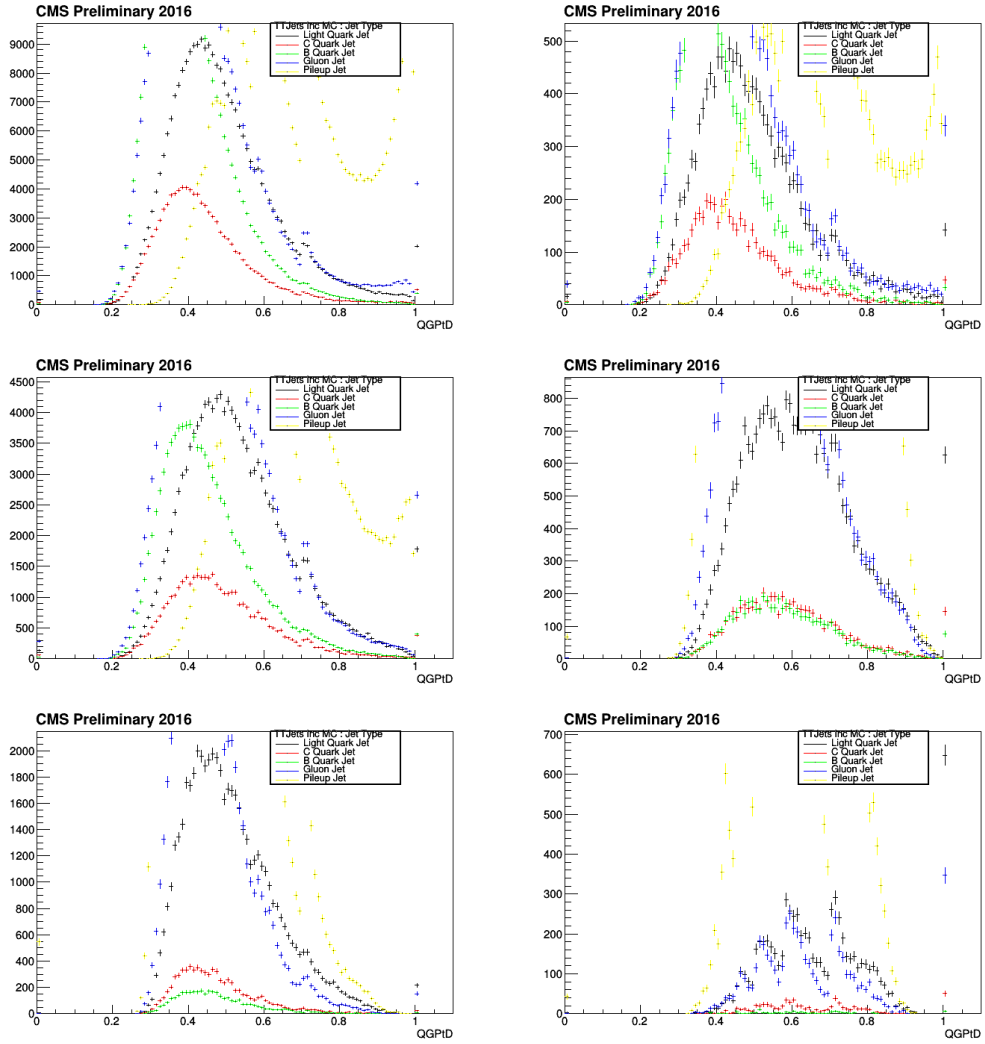


Figure A.1: Top left: Quark Gluon $p_t D$ for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.

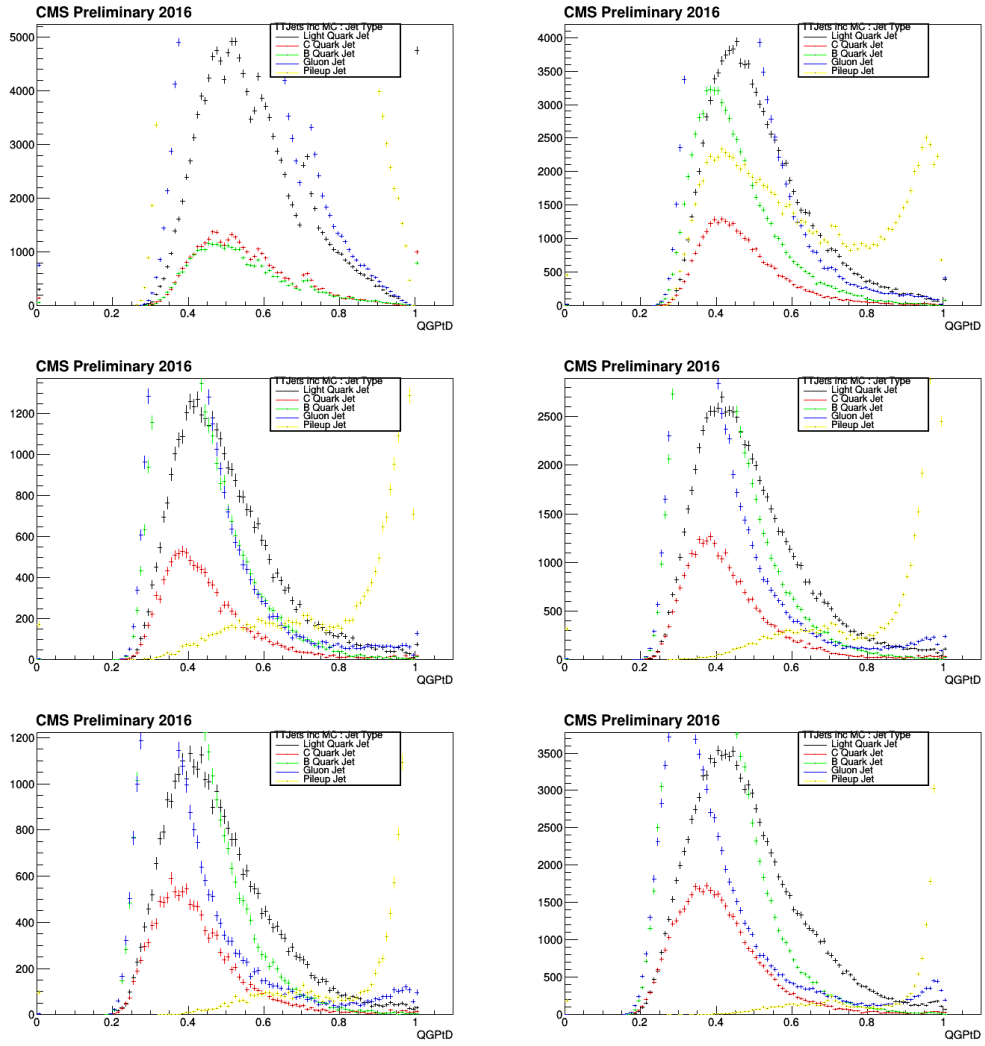


Figure A.2: Top left: Quark Gluon $p_t D$ for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.

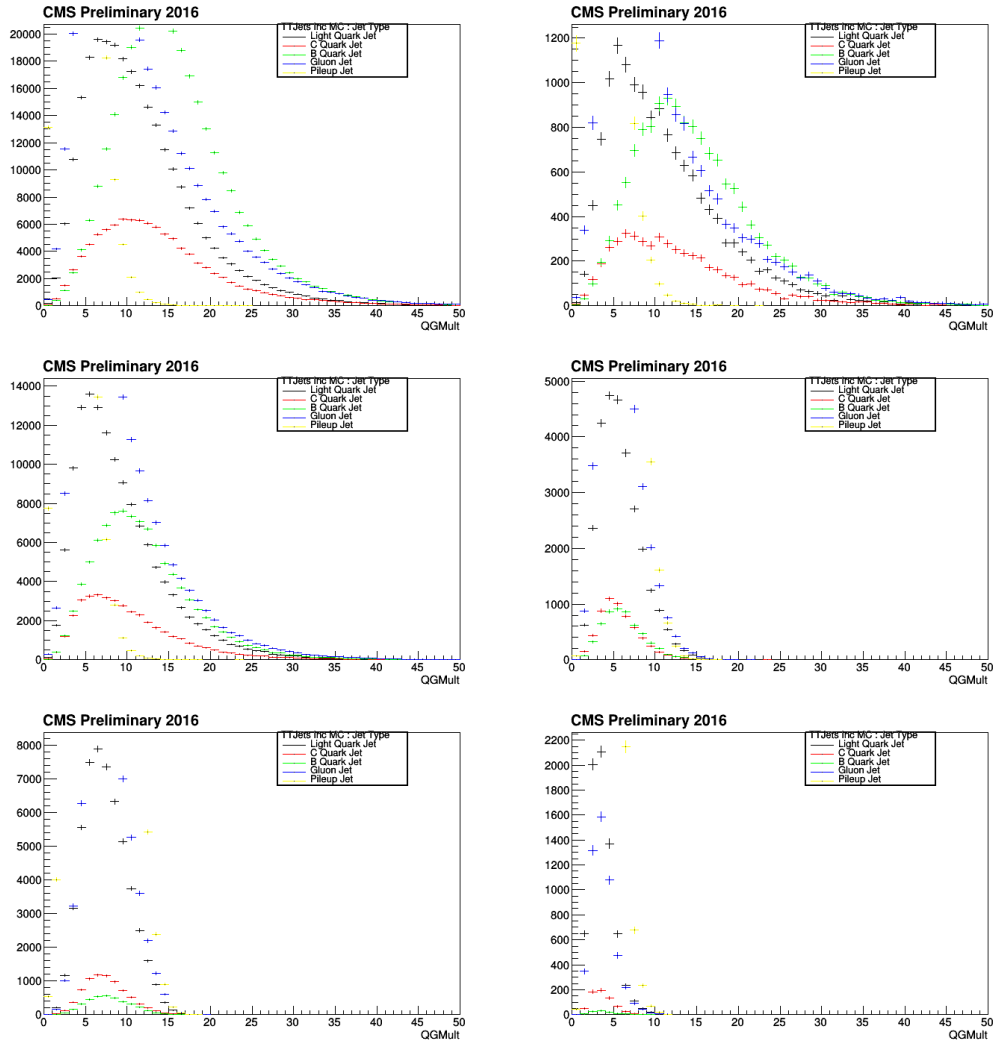


Figure A.3: Top left: Quark Gluon multiplicity for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.

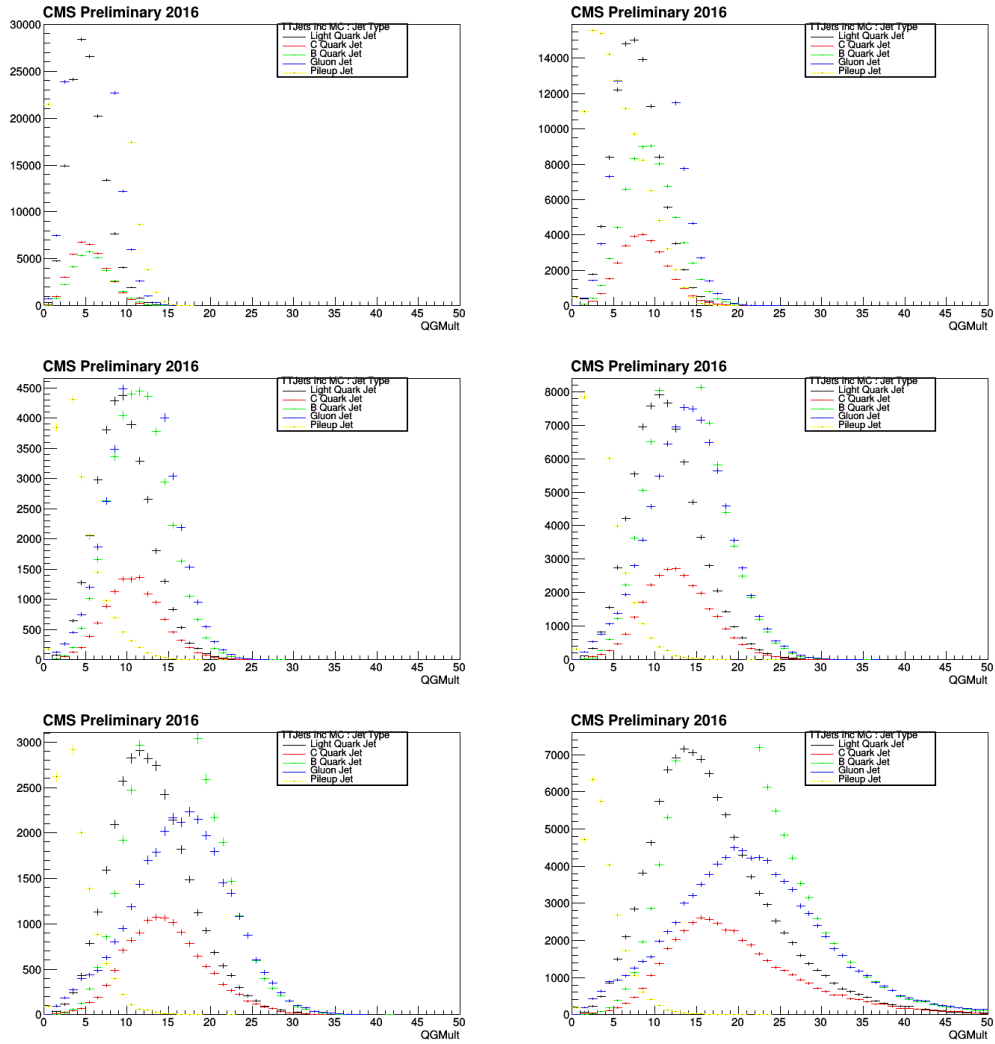


Figure A.4: Top left: Quark Gluon multiplicity for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.

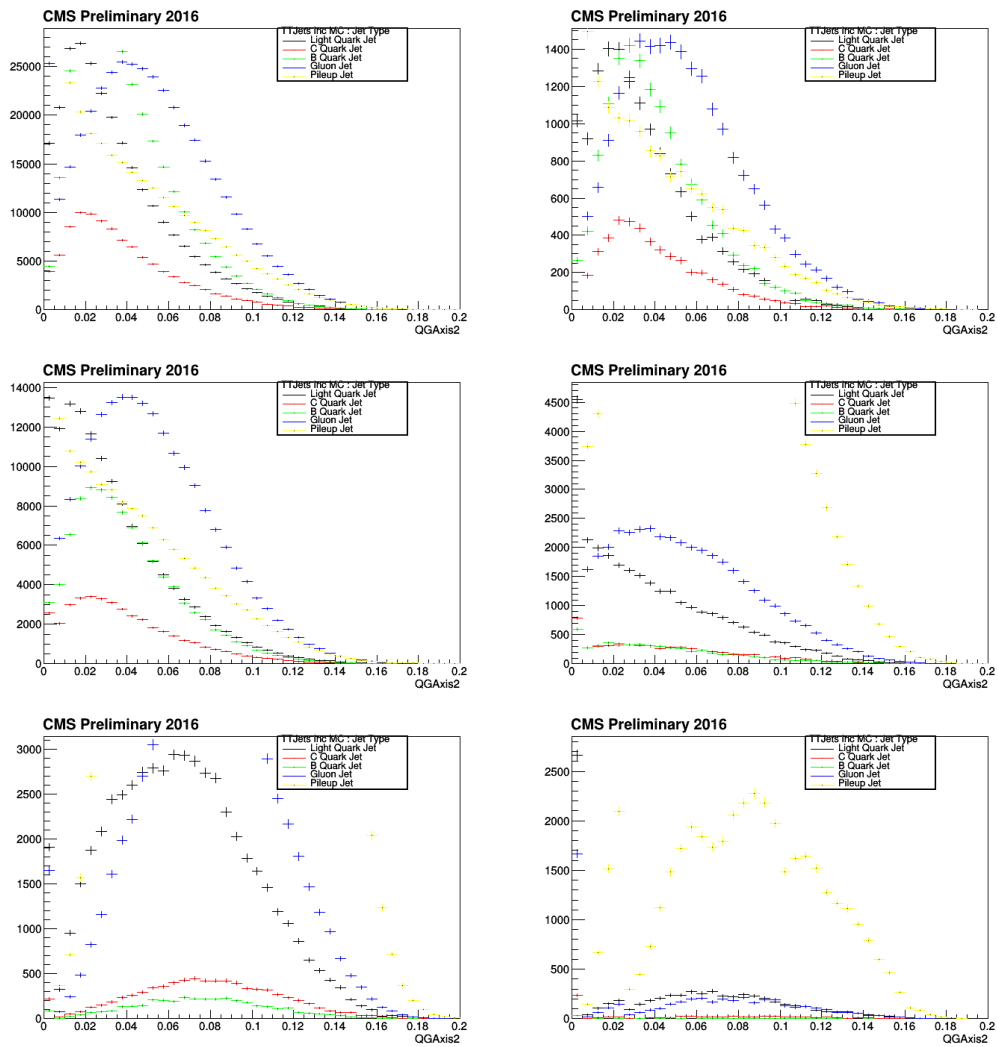


Figure A.5: Top left: Quark Gluon Axis2 for jet η bin 1; Top right: jet η bin 2; Middle left: jet η bin 3; Middle right: jet η bin 4; Middle left: jet η bin 5; Middle right: jet η bin 6.

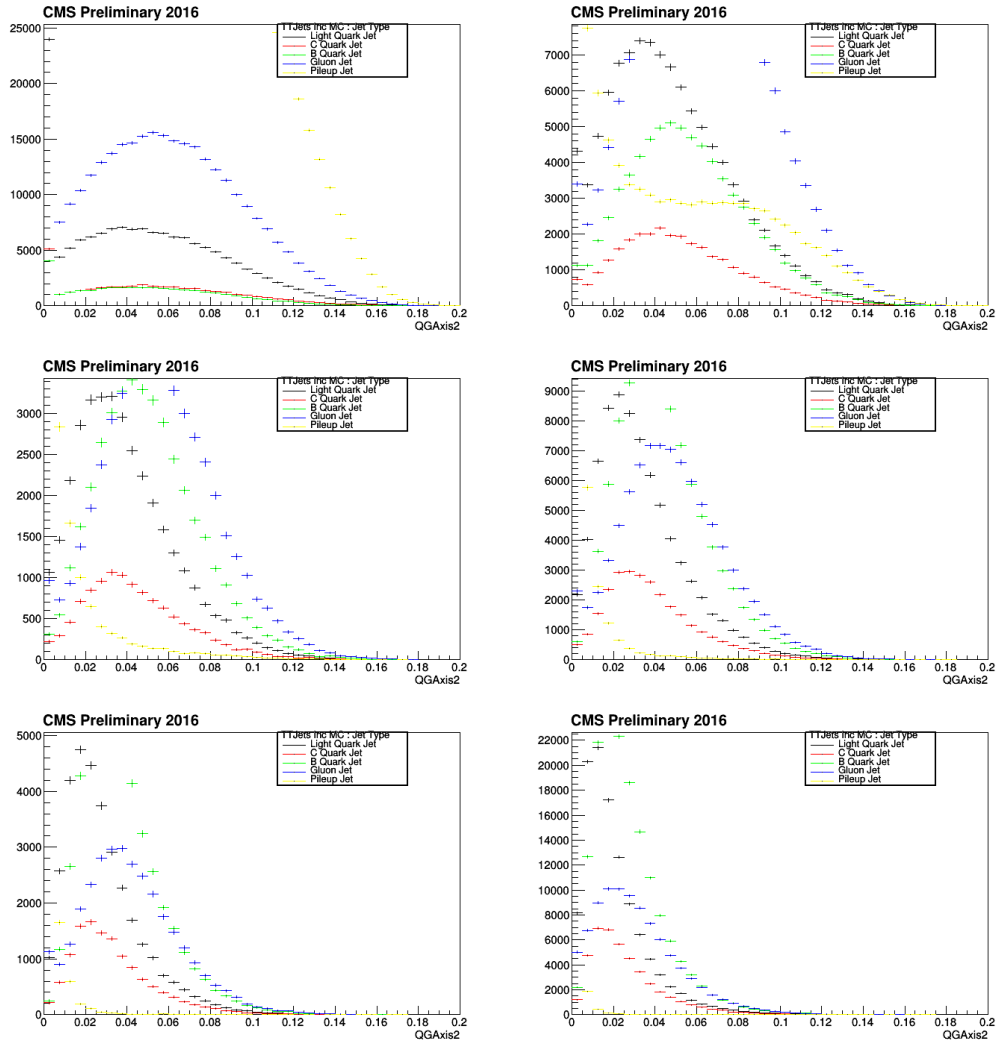


Figure A.6: Top left: Quark Gluon Axis2 for jet p_T bin 1; Top right: jet p_T bin 2; Middle left: jet p_T bin 3; Middle right: jet p_T bin 4; Middle left: jet p_T bin 5; Middle right: jet p_T bin 6.

A.2 Di-Top jets event display

The Event with two reconstructed top jets are demonstrated in Fig A.7. The jet energies and p_T^{miss} shown are without energy correction.

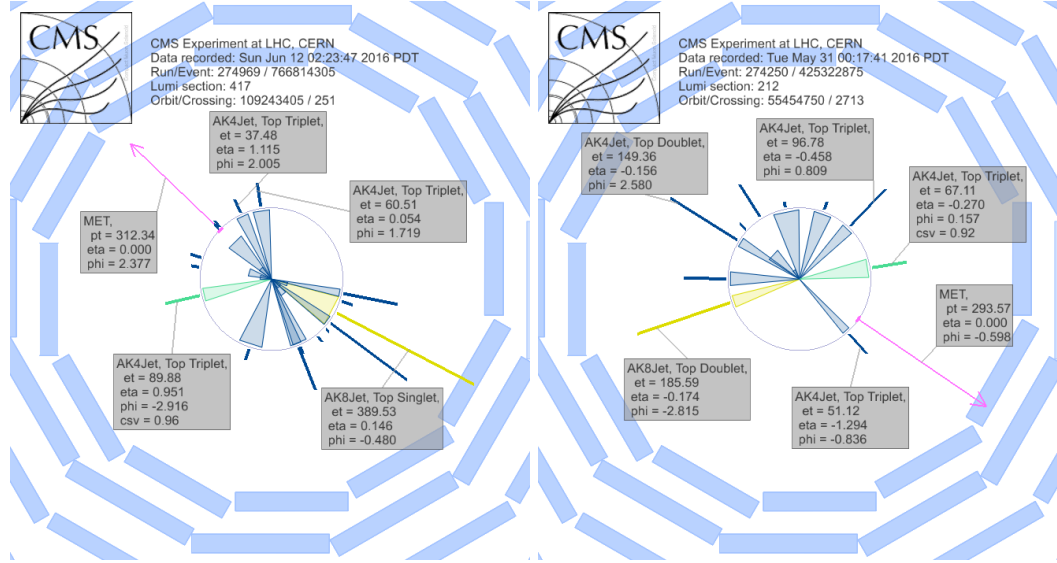


Figure A.7: Event display for di-topjets event in data. Left plot: Event display for a SUSY candidate event (from Run 274250, Event number 425322875) with two reconstructed top-tagged candidates (one is in di-jet category, the other is in tri-jet category) passed the search region selection, in $r-\phi$ view of the CMS detector. The momenta of the AK4 jets that tagged as b-jet in tri-jet category are marked in green, while other AK4 jets are marked in blue. The momentum of the AK8 jets are marked in yellow and missing transverse momentum is marked in purple. Right plot: Event display for a SUSY candidate event (from Run 274969, Event number 766814305) with two reconstructed top-tagged candidates (one is in mono-jet category, the other is in tri-jet category) passed the search region selection, in $r-\phi$ view of the CMS detector. The momenta of the AK4 jets that tagged as b-jet in tri-jet category are marked in green, while other AK4 jets are marked in blue. The momentum of the AK8 jets are marked in yellow and missing transverse momentum is marked in purple.

A.3 Simplified top tagger, aggregate search bin and results

Simplified top tagger and 10 aggregate search bins are defined to simplify the use of our data. We drop the quark gluon discriminator, therefore theorist can repeat our tagging algorithm. As given in Table A.1, the aggregate search bins are not exclusive. The first four aggregate regions represent topologies of general interest. The fifth and sixth are sensitive to direct top squark pair production. The seventh region targets the large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ region of T5ttcc-like models, while the final three target events with a large number of top quarks such as are produced in the T1tttt and T5tttt models.

Table A.1: Definition of the aggregate search regions.

Region	N_{tops}	$N_{\text{b-jets}}$	M_{T2} [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	Motivation
1	≥ 1	≥ 1	≥ 200	≥ 250	Events satisfying selection criteria
2	≥ 2	≥ 2	≥ 200	≥ 250	Events with $N_{\text{tops}} \geq 2$ and $N_{\text{b-jets}} \geq 2$
3	≥ 3	≥ 1	≥ 200	≥ 250	Events with $N_{\text{tops}} \geq 3$ and $N_{\text{b-jets}} \geq 1$
4	≥ 3	≥ 3	≥ 200	≥ 250	T5tttt; small $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$ and $m_{\tilde{\chi}_1^0} < m_t$
5	≥ 2	≥ 1	≥ 200	≥ 400	T2tt; small $\Delta M(\tilde{t}, \tilde{\chi}_1^0)$
6	≥ 1	≥ 2	≥ 600	≥ 400	T2tt; large $\Delta M(\tilde{t}, \tilde{\chi}_1^0)$
Region	N_{tops}	$N_{\text{b-jets}}$	H_{T} [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	Motivation
7	≥ 1	≥ 2	≥ 1400	≥ 500	T1ttbb and T5ttcc; large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$
8	≥ 2	≥ 3	≥ 600	≥ 350	T1tttt; small $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$
9	≥ 2	≥ 3	≥ 300	≥ 500	T1/T5tttt and T1ttbb; intermediate $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$
10	≥ 2	≥ 3	≥ 1300	≥ 500	T1/T5tttt; large $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$

The aggregate search bin results are shown in Fig A.8 and Table A.2, using the same background estimation methods described in this thesis.

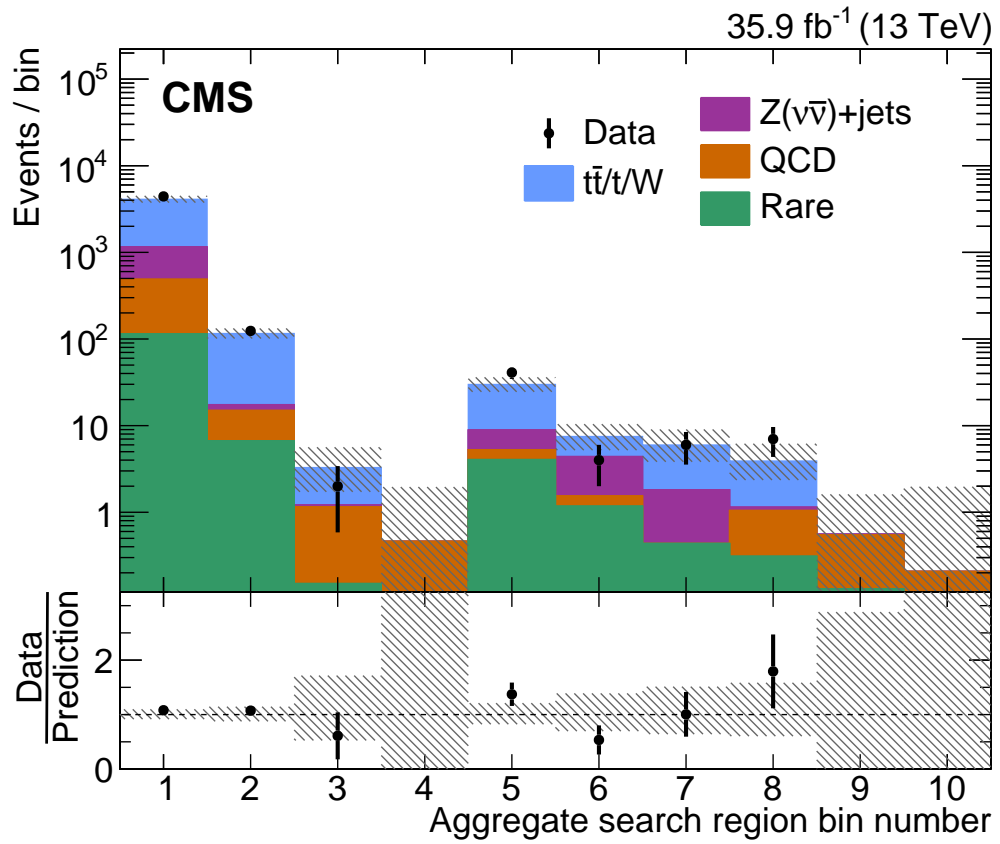


Figure A.8: Observed event yields (black points) and predicted SM background (filled solid areas) for the 10 aggregate search regions. The lower panel shows the ratio of the data to the total background prediction. The hatched bands correspond to the total uncertainty in the background prediction.

Table A.2: The observed number of events and the total background prediction for the aggregate search regions. The first uncertainty in the background prediction is statistical and the second is systematic.

Search region	N_{tops}	$N_{\text{b-jets}}$	$M_{\text{T}2}$ [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	Data	Predicted background
1	≥ 1	≥ 1	≥ 200	≥ 250	4424	$4100 \pm 50_{-340}^{+390}$
2	≥ 2	≥ 2	≥ 200	≥ 250	124	$116 \pm 8_{-12}^{+15}$
3	≥ 3	≥ 3	≥ 200	≥ 250	0	$0.5_{-0.4}^{+1.4} \pm 0.5$
4	≥ 3	≥ 1	≥ 200	≥ 250	2	$3.3_{-1.1}^{+2.0} \pm 1.2_{-1.1}$
5	≥ 2	≥ 1	≥ 200	≥ 400	41	$30_{-3}^{+4} \pm 5_{-4}$
6	≥ 1	≥ 2	≥ 600	≥ 400	4	$7.5_{-1.2}^{+2.1} \pm 2.0_{-1.9}$
Search region	N_{tops}	$N_{\text{b-jets}}$	H_{T} [GeV]	$p_{\text{T}}^{\text{miss}}$ [GeV]	Data	Predicted background
7	≥ 1	≥ 2	≥ 1400	≥ 500	6	$6.0_{-1.5}^{+2.7} \pm 1.5$
8	≥ 2	≥ 3	≥ 600	≥ 350	7	$3.9_{-1.2}^{+2.1} \pm 0.9$
9	≥ 2	≥ 3	≥ 300	≥ 500	0	$0.6_{-0.4}^{+1.0} \pm 0.4$
10	≥ 2	≥ 3	≥ 1300	≥ 500	0	$0.2_{-0.3}^{+1.8} \pm 0.2$

Appendix B

Service work

B.1 HCAL logical map

HCAL Logical map has been described in detail in the HCAL section in chapter 3. The HCAL map validation plots are listed here in different period and various sub-detectors. The FrontEnd electronics variables are the based coordinates that used in the validation plots. The validation plots are separated into 2 categories: BackEnd variables in FrontEnd coordinates and Geometry variables in FrontEnd coordinates. A complete FrontEnd coordinates set has 4 variables, RBX (readout box), RM (readout module), RM fibers and fiber channel for HB, HE and HO. HF is special since it does have RM (readout module). HF FrontEnd can be described by RBX, QIE10 slot, QIE10 fiber and fiber channel. BackEnd and Geometry variables will be described separately in the following sections.

The basic status of HCAL in 2015 and 2016 are listed in Table [B.1](#) and Table [B.2](#) respectively, as bookkeeping. The 2017 HCAL (operating now) status are summarized in

Table B.3. The planned phase 1 and phase 2 HCAL summaries are shown in Table B.4 and Table B.5 respectively.

ADCs that come from same η and ϕ are summed over and delivered to trigger system. This is so-called trigger primitive (TP) in HCAL system. There are 2304 TP channels in HB, 1728 TP channels in HE and 792 TP channels in HF. HO TP is supposed to be delivered to muon trigger, therefore HO TP maintenance is not HCAL groups responsibility.

Table B.1: HCAL Logical Map status in 2015.

Subdet	HB	HE	HF	HO
No. Geo Channels	2592	2592	1728	2160
Photon Dectector	HPD	HPD	PMT	SiPM
Type of FrontEnd	QIE8	QIE8	QIE8	QIE8
No. FE Channels	2592	2592	1728	2304
Type of BackEnd	VME	VME	uTCA	VME
No. BE Channels	2592	2592	1728	2376

Table B.2: HCAL Logical Map status in 2016.

Subdet	HB	HE	HF	HO
No. Geo Channels	2592	2592	1656+144	2160
Photon Dectector	HPD	HPD	PMT	SiPM
Type of FrontEnd	QIE8	QIE8	QIE8+QIE10	QIE8
No. FE Channels	2592	2592	1656+144	2304
Type of BackEnd	uTCA	uTCA	uTCA	VME
No. BE Channels	2592	2592	1656+144	2376

Table B.3: HCAL Logical Map status in 2017.

Subdet	HB	HE	HF	HO
No. Geo Channels	2592	2520+188	3456	2160
Photon Dectector	HPD	HPD+SiPM	PMT	SiPM
Type of FrontEnd	QIE8	QIE8+QIE11	QIE10	QIE8
No. FE Channels	2592	2520+192	3456	2304
Type of BackEnd	uTCA	uTCA	uTCA	VME
No. BE Channels	2592	2520+192	3456	2376

Table B.4: HCAL Logical Map status in 2018.

Subdet	HB	HE	HF	HO
No. Geo Channels	2592	6768	3456	2160
Photon Dectector	HPD	SiPM	PMT	SiPM
Type of FrontEnd	QIE8	QIE11	QIE10	QIE8
No. FE Channels	2592	6912	3456	2304
Type of BackEnd	uTCA	uTCA	uTCA	VME
No. BE Channels	2592	6912	3456	2376

B.1.1 Remapped phase 1 HB in 2018

The 2018 HB is still in HPD+QIE8 FrontEnd+uTCA BackEnd stage. We are in this hardware setting since 2016. However, HE will be upgraded to SiPM+QIE11 FrontEnd+uTCA BackEnd with more readout channels. Then the patch panels need to be reorganized since HB and HE share the same BackEnd electronics. The remapped 2018 HB is followed by the design from Richard Kellogg, coordinated with Jeremy Mans on the requirement of trigger sum algorithm in BackEnd firmware.

The validation plots of remapped HB in 2018 are showed from Fig B.1 to Fig B.12. There are 2592 readout channels in total for remapped HB (same as 2016 HB, before remap). In order to satisfy the trigger latency requirement, the frontend channel permutation is

Table B.5: HCAL Logical Map status in 2023.

Subdet	HB	HE	HF	HO
No. Geo Channels	9072	Retired	3456	2160
Photon Dectector	SiPM	Retired	PMT	SiPM
Type of FrontEnd	QIE11	Retired	QIE10	QIE8
No. FE Channels	9216	Retired	3456	2304
Type of BackEnd	uTCA	Retired	uTCA	VME
No. BE Channels	9216	Retired	3456	2376

applied in this map. HB BackEnd electronics are still in uTCA, 12 uHTR in one crate, 24 out going fibers per uHTR card, with 3 channels per fiber.

HBP Phi in FrontEnd coordiantes

HBP18RM4	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70
HBP18RM3	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69
HBP18RM2	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68
HBP18RM1	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67
HBP17RM4	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66
HBP17RM3	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65
HBP17RM2	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64
HBP17RM1	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63
HBP16RM4	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62
HBP16RM3	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61
HBP16RM2	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60
HBP16RM1	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59
HBP15RM4	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58
HBP15RM3	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57
HBP15RM2	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56
HBP15RM1	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55
HBP14RM4	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54
HBP14RM3	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53
HBP14RM2	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52
HBP14RM1	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51
HBP13RM4	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
HBP13RM3	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49
HBP13RM2	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
HBP13RM1	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47
HBP12RM4	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46
HBP12RM3	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45
HBP12RM2	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44
HBP12RM1	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
HBP11RM4	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42
HBP11RM3	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41
HBP11RM2	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
HBP11RM1	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39
HBP10RM4	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38
HBP10RM3	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
HBP10RM2	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36
HBP10RM1	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
HBP09RM4	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34
HBP09RM3	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
HBP09RM2	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
HBP09RM1	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31
HBP08RM4	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
HBP08RM3	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
HBP08RM2	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
HBP08RM1	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
HBP07RM4	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26
HBP07RM3	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
HBP07RM2	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
HBP07RM1	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23
HBP06RM4	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
HBP06RM3	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
HBP06RM2	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
HBP06RM1	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
HBP05RM4	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
HBP05RM3	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
HBP05RM2	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
HBP05RM1	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
HBP04RM4	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
HBP04RM3	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
HBP04RM2	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
HBP04RM1	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
HBP03RM4	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
HBP03RM3	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
HBP03RM2	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
HBP03RM1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
HBP02RM4	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
HBP02RM3	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
HBP02RM2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
HBP02RM1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
HBP01RM4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
HBP01RM3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HBP01RM2	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71
HBP01RM1	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71

Figure B.2: HCAL (phase 1 HB, plus side) detector ϕ distribution in the frontend electronic coordinates.

HBP Depth in FrontEnd coordiantes

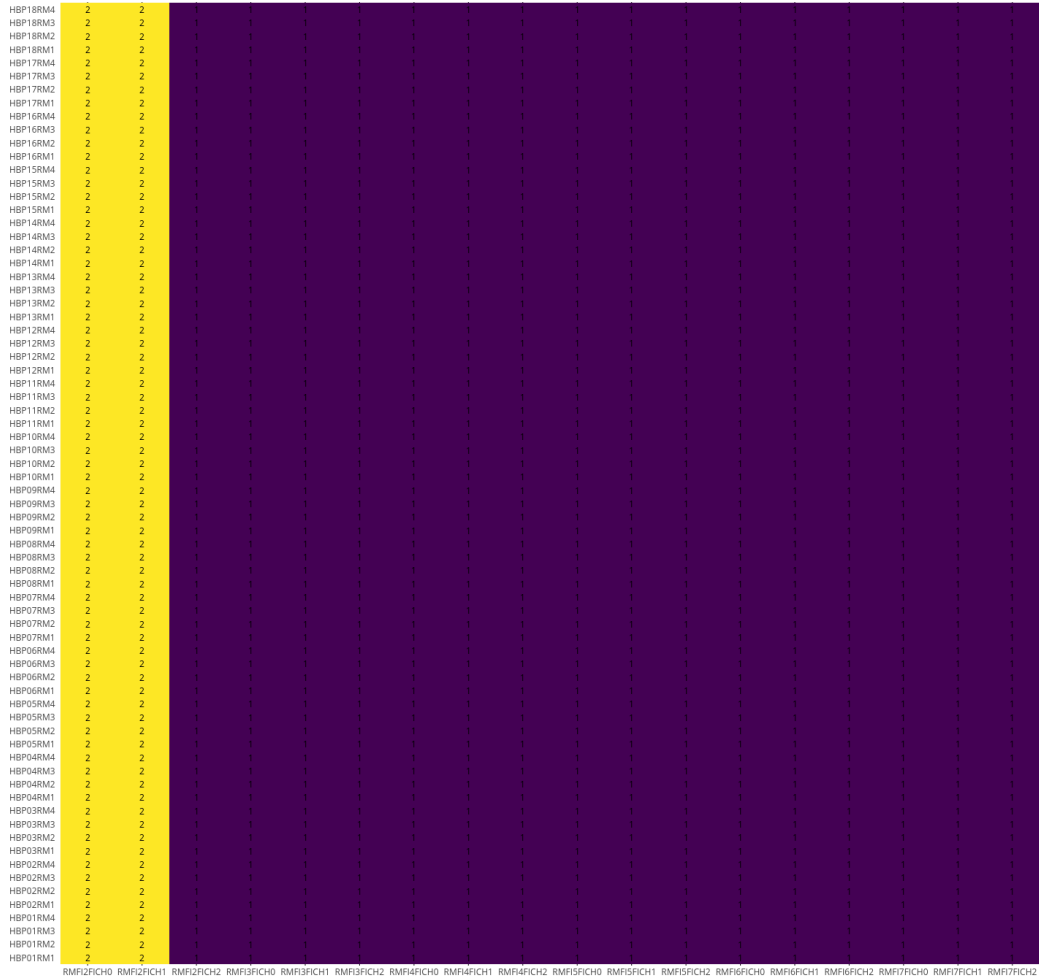


Figure B.3: HCAL (phase 1 HB, plus side) detector depth distribution in the frontend electronic coordinates.

HBM Phi in FrontEnd coordiantes

HBM18R44	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67
HBM18R43	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68
HBM18R42	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69
HBM18R41	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70
HBM17R44	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63
HBM17R43	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64
HBM17R42	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65
HBM17R41	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66
HBM16R44	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59
HBM16R43	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60
HBM16R42	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61
HBM16R41	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62
HBM15R44	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55
HBM15R43	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56
HBM15R42	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57
HBM15R41	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58
HBM14R44	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51
HBM14R43	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52
HBM14R42	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53
HBM14R41	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54
HBM13R44	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47
HBM13R43	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
HBM13R42	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49
HBM13R41	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
HBM12R44	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
HBM12R43	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44
HBM12R42	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45
HBM12R41	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46
HBM11R44	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39
HBM11R43	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
HBM11R42	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41
HBM11R41	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42
HBM10R44	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
HBM10R43	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36
HBM10R42	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
HBM10R41	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38
HBM09R44	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31
HBM09R43	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
HBM09R42	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
HBM09R41	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34
HBM08R44	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
HBM08R43	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
HBM08R42	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
HBM08R41	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
HBM07R44	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23
HBM07R43	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
HBM07R42	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
HBM07R41	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26
HBM06R44	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
HBM06R43	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
HBM06R42	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
HBM06R41	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
HBM05R44	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
HBM05R43	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
HBM05R42	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
HBM05R41	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
HBM04R44	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
HBM04R43	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
HBM04R42	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
HBM04R41	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
HBM03R44	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
HBM03R43	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
HBM03R42	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
HBM03R41	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
HBM02R44	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
HBM02R43	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
HBM02R42	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
HBM02R41	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
HBM01R44	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71
HBM01R43	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72
HBM01R42	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HBM01R41	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

RMF2FCH0 RMF2FCH1 RMF2FCH2 RMF3FCH0 RMF3FCH1 RMF3FCH2 RMF4FCH0 RMF4FCH1 RMF4FCH2 RMF5FCH0 RMF5FCH1 RMF5FCH2 RMF6FCH0 RMF6FCH1 RMF6FCH2 RMF7FCH0 RMF7FCH1 RMF7FCH2

Figure B.5: HCAL (phase 1 HB, minus side) detector ϕ distribution in the frontend electronic coordinates.

HBM Depth in FrontEnd coordiantes

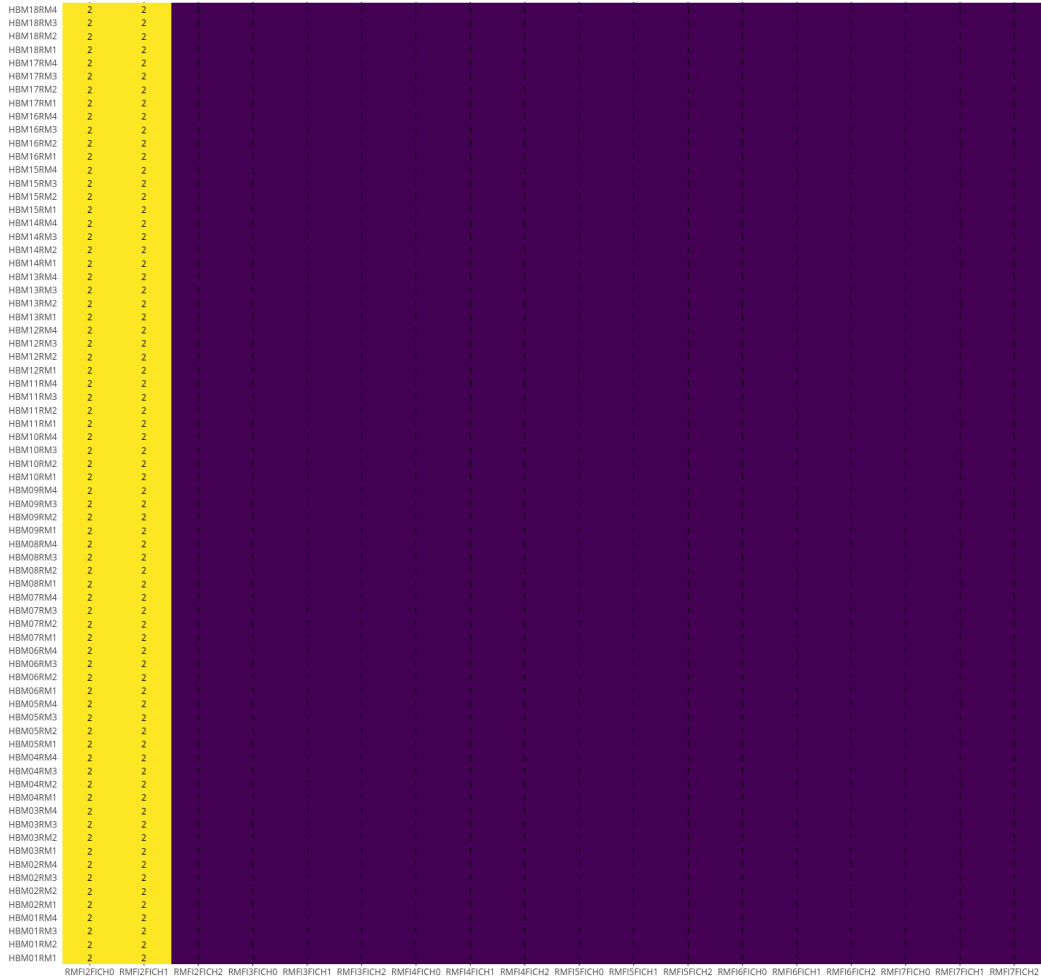


Figure B.6: HCAL (phase 1 HB, minus side) detector depth distribution in the frontend electronic coordinates.

B.1.2 Phase 2 HB in 2020

The 2018 HB will be replaced and move to SiPM+QIE11 FrontEnd+uTCA Back-End stage in 2020, during the long shutdown 2. A complicate swap scheme is applied inside the FrontEnd board to satisfy the trigger latency requirement. Fortunately, thanks to Dicks elegant design, we do not need to unplug HE fibers from patch panel to fit in the new 2020 HB.

The phase 2 2020 HB validation plots are shown from Fig B.13 to Fig B.24. There are 9216 readout channels in phase 2 HB. 9072 channels are physical readout channel, corresponding to actual tower and layer on detector. 144 of them are dummy calibration channels, which are readout by FrontEnd with no corresponding detector components. Those dummy calibration channels are labeled as “HBX” channels. HBX channels are distributed among 144 HB readout modules, but only on two locations: RM fiber 2 fiber channel 2 and RM fiber 3 fiber channel 6. BackEnd electronics are still in 12 uHTR and 24 uHTR fibers, but number of fiber channels per fiber is increased from 3 to 8, to accommodating more readout channels, comparing with 2018 remapped HB.

ngHBP Eta in FrontEnd coordiantes

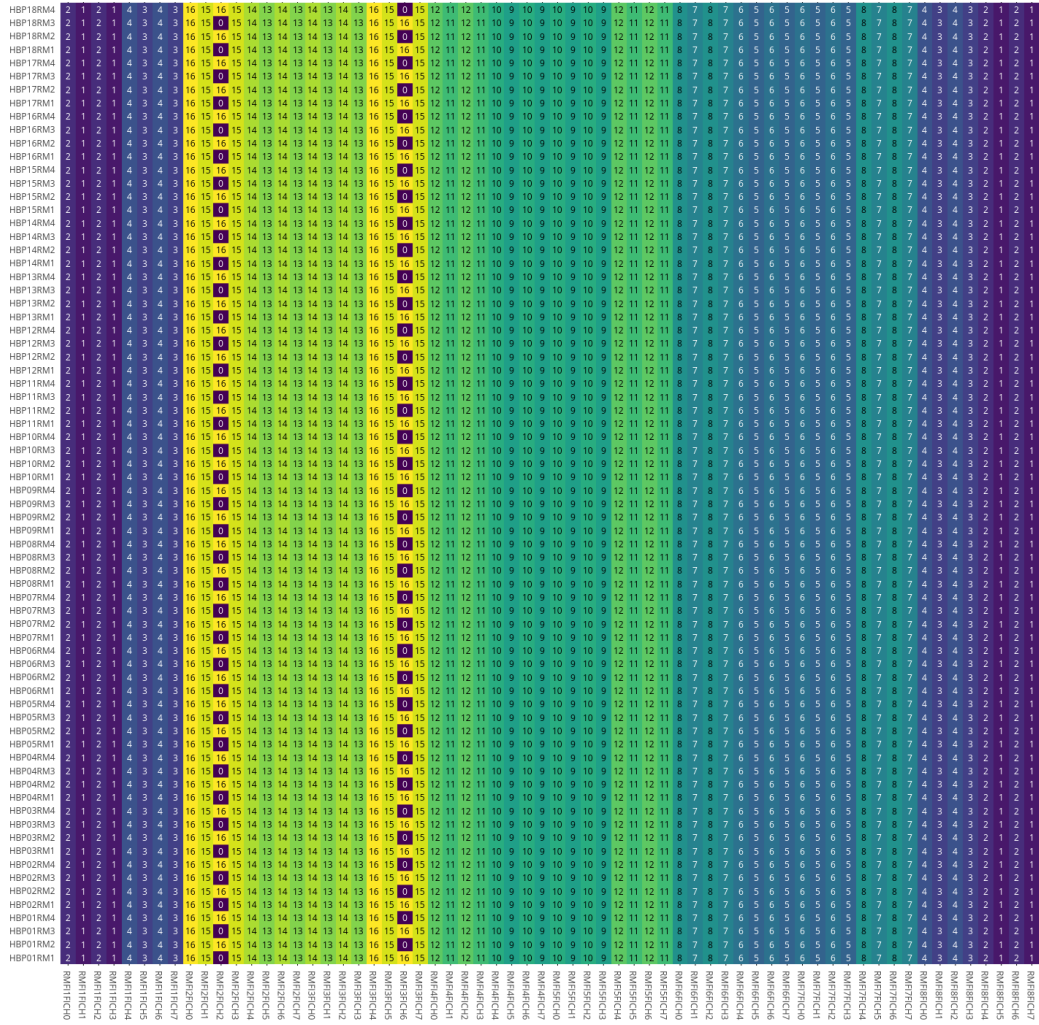


Figure B.13: HCAL (phase 2 HB, plus side) detector η distribution in the frontend electronic coordinates.

ngHBP Depth in FrontEnd coordiantes

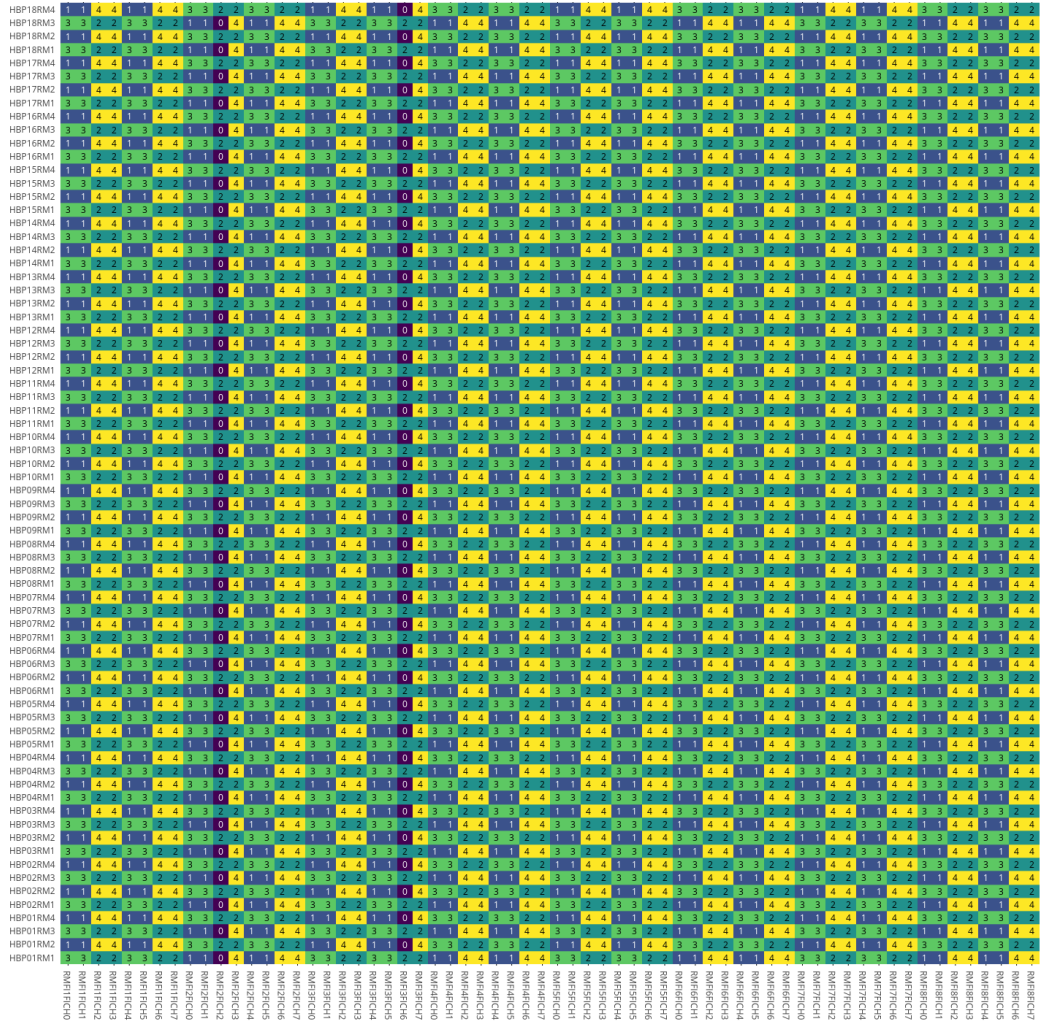


Figure B.15: HCAL (phase 2 HB, plus side) detector depth distribution in the frontend electronic coordinates.

ngHBM Depth in FrontEnd coordiantes

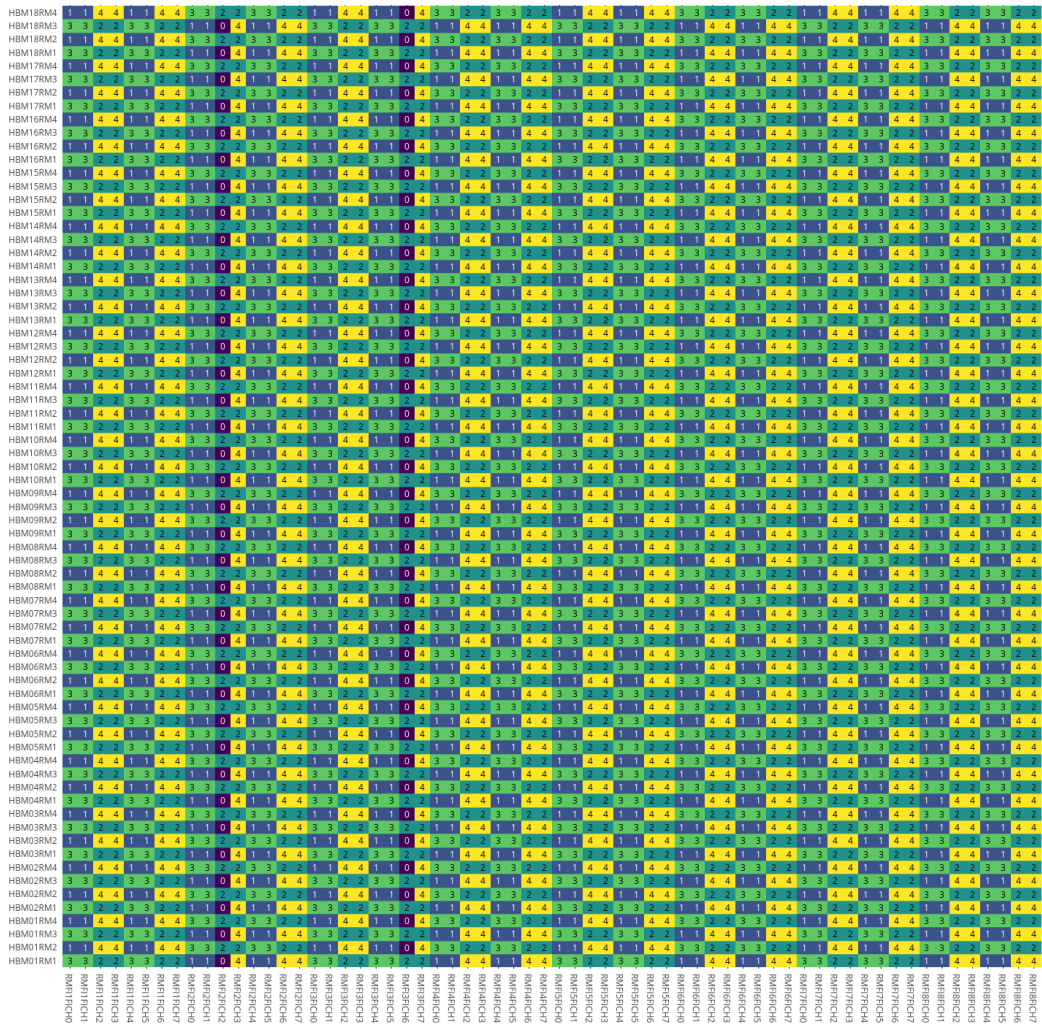


Figure B.18: HCAL (phase 2 HB, minus side) detector depth distribution in the frontend electronic coordinates.

ngHBM uHTR in FrontEnd coordiantes

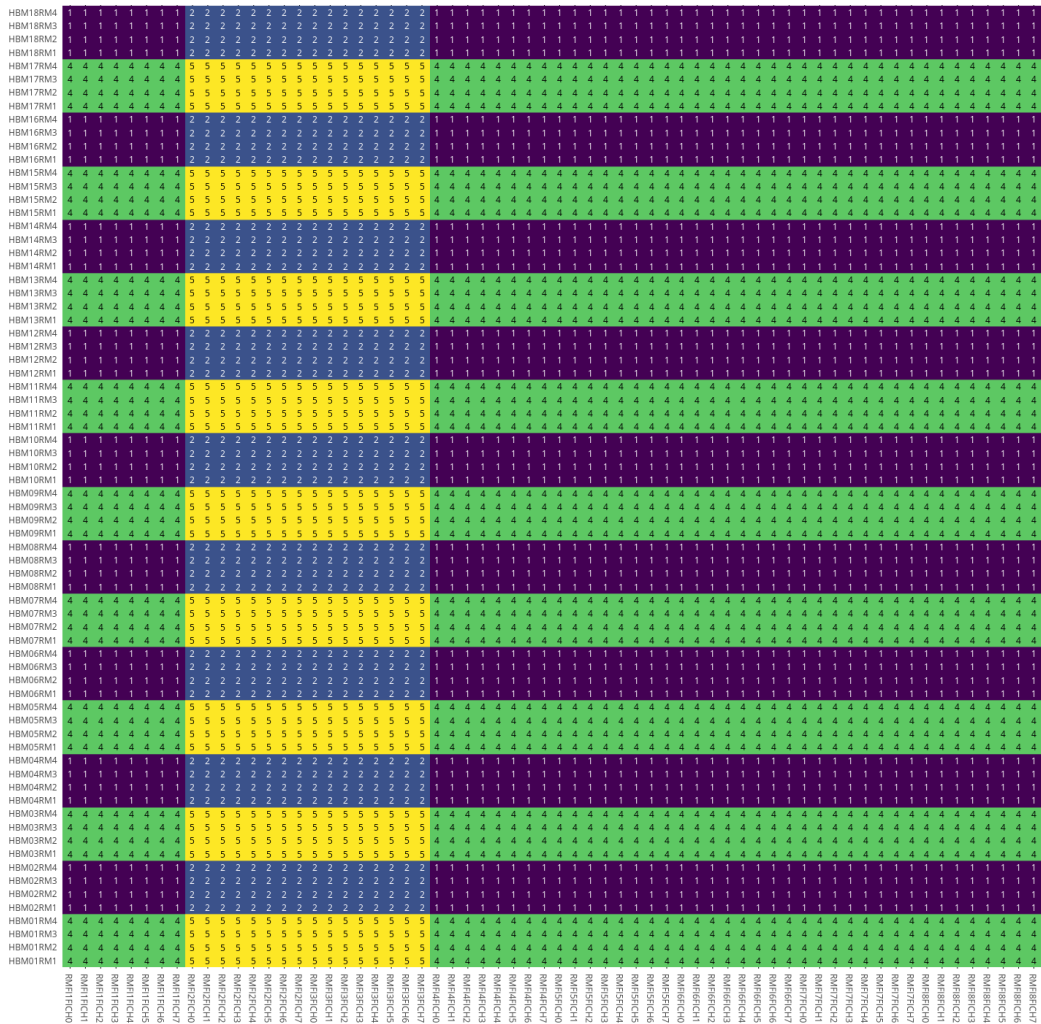


Figure B.23: HCAL (phase 2 HB, minus side) backend electronic coordinate uHTR slot distribution in the front-end electronic coordinates.

B.1.3 Phase 1 HE in 2018

The 2018 HE is in SiPM+QIE11 FrontEnd+uTCA BackEnd stage. HE mapping algorithm is always more complicate than HB due to the tricky geometry structure in detector coordinates. Comparing to 2016 HE with QIE8, no swap trick inside the FrontEnd board is applied in 2018 HE. This indeed increases the difficulty for the firmware expert to pin down the latency within L1 trigger requirement.

The phase 1 2018 HE validation plots are shown from Fig B.25 to Fig B.36. There are 6912 readout channels in phase 1 HE. 6768 channels are physical readout channel, corresponding to actual tower and layer on detector. 144 of them are dummy calibration channels, which are readout by FrontEnd with no corresponding detector components. Those dummy calibration channels are labeled as “HEX” channels. HEX channels are distributed among 144 HE readout modules, on four different locations: RM fiber 2 fiber channel 1, RM fiber 3 fiber channel 6, RM fiber 5 fiber channel 5 and RM fiber 7 fiber channel 1. BackEnd electronics are still in 12 uHTR and 24 uHTR fibers, but number of fiber channels per fiber is increased from 3 to 6, to accommodating more readout channels, comparing with 2016 HE.

ngHEP Depth in FrontEnd coordiantes

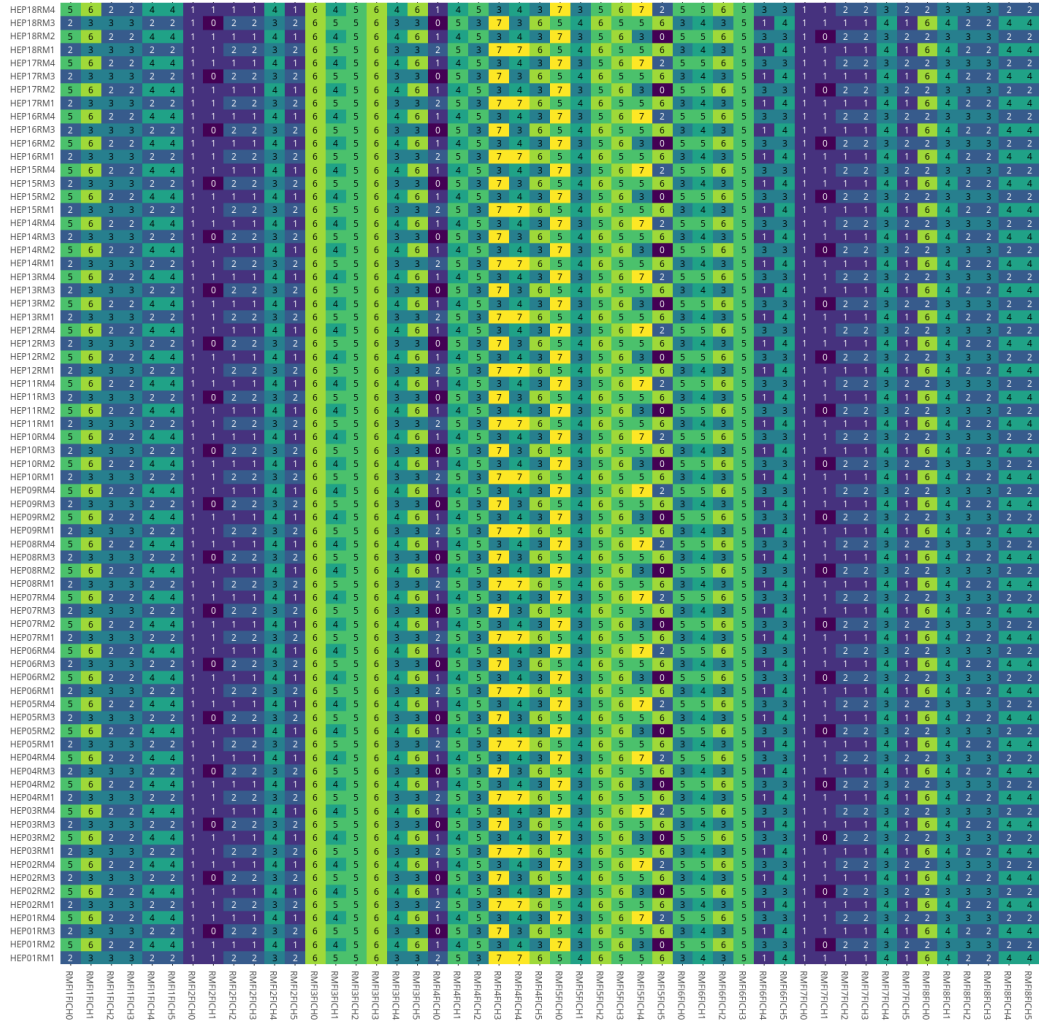


Figure B.27: HCAL (phase 1 HE, plus side) detector depth distribution in the frontend electronic coordinates.

ngHEM Depth in FrontEnd coordiantes

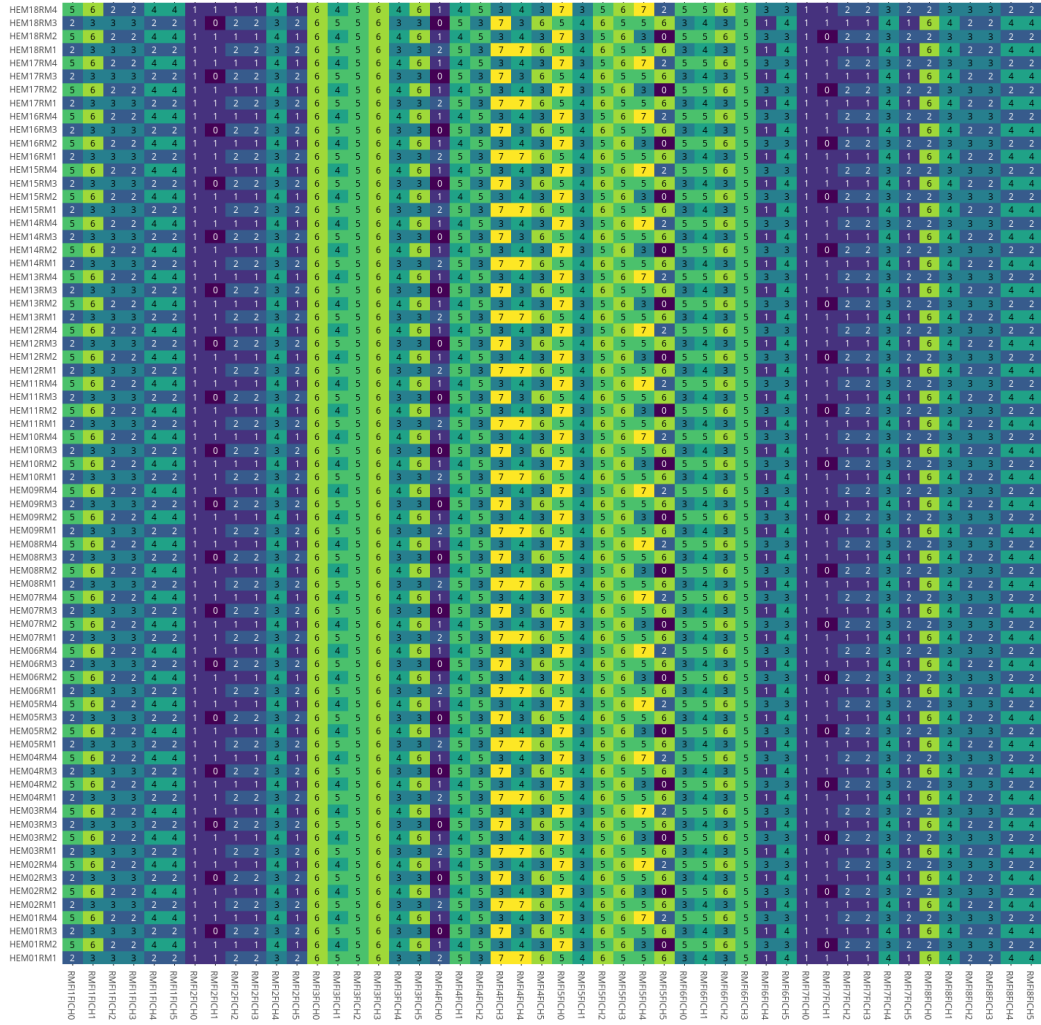


Figure B.30: HCAL (phase 1 HE, minus side) detector depth distribution in the frontend electronic coordinates.

ngHEP uHTR in FrontEnd coordiantes

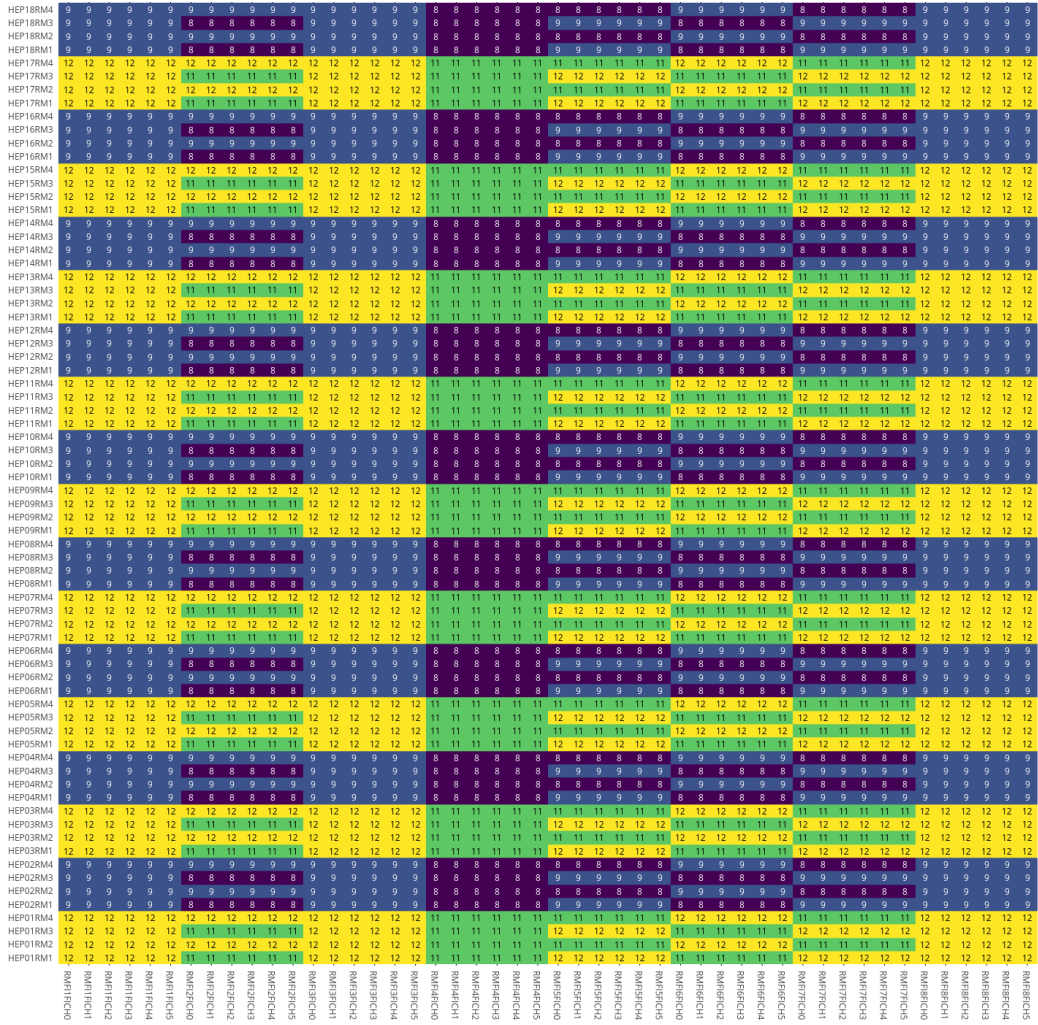


Figure B.32: HCAL (phase 1 HE, plus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.

ngHEM uHTR in FrontEnd coordiantes

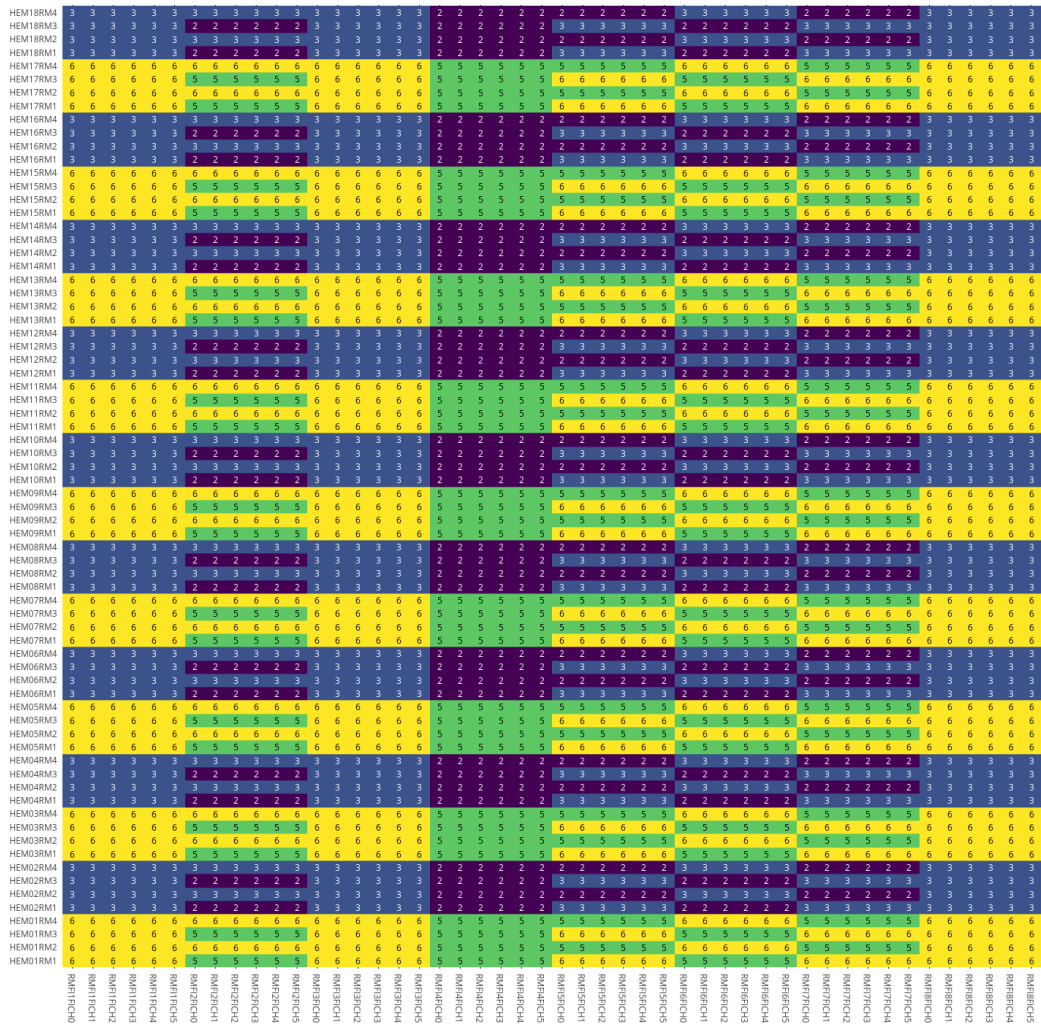


Figure B.35: HCAL (phase 1 HE, minus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.

B.1.4 Phase 1 HF in 2017

The 2017 HF is in PMT+QIE10 FrontEnd+uTCA BackEnd stage. HF BackEnd electronics are upgraded from VME to uTCA during 2015 YETS. HF is the first sub-detector to have uTCA based BackEnd electronics installed and operated in CMS. HF The QIE10 FrontEnd electronics are installed on HF during 2017 YETS.

The phase 1 2017 HF validation plots are shown from Fig [B.37](#) to Fig [B.48](#). There are 3456 readout channels in phase 1 HF. BackEnd electronics are in standard uTCA way: 12 uHTR and 24 uHTR fibers. The number of fiber channels per fiber is increased from 3 to 4 to accommodate dual readout channels, comparing with legacy HF.

ngHFP Eta in FrontEnd coordiantes

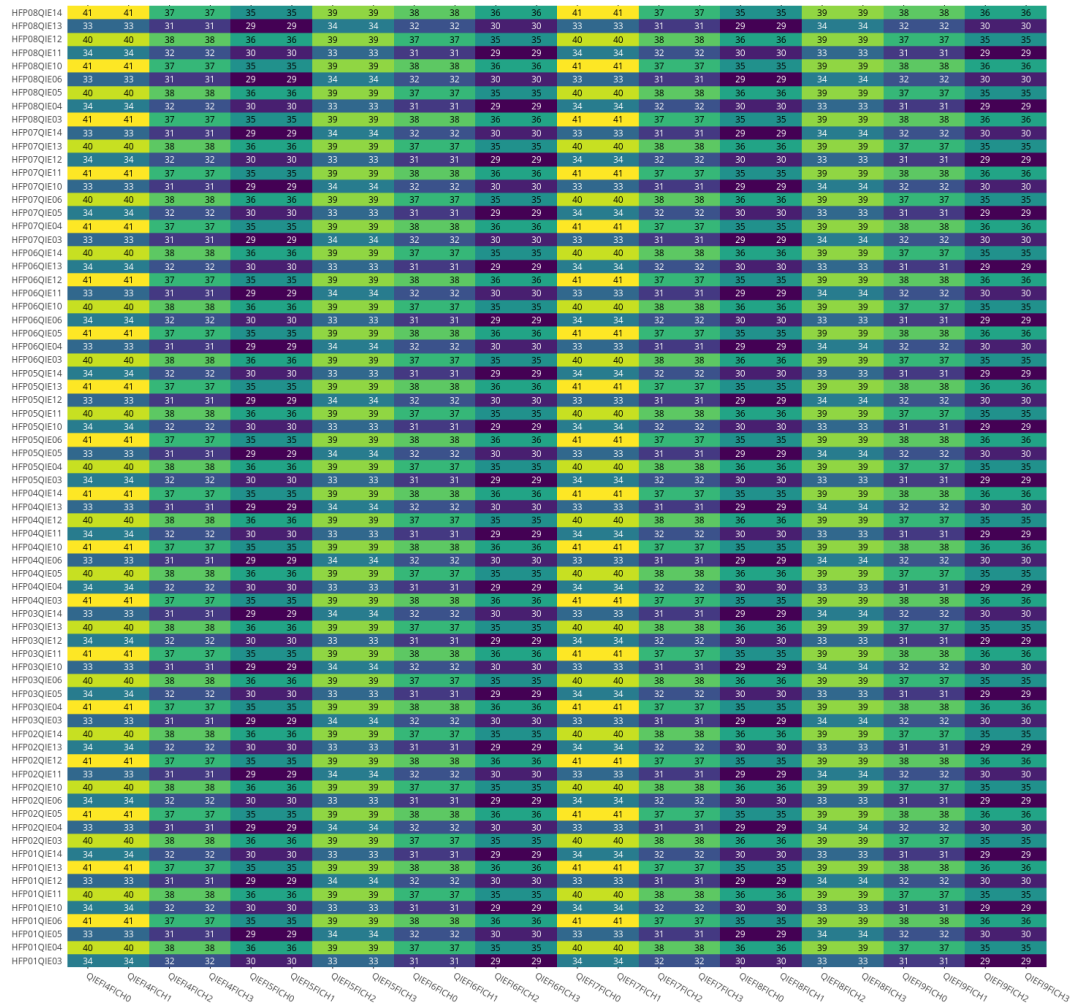


Figure B.37: HCAL (phase 1 HF, plus side) detector η distribution in the frontend electronic coordinates.

ngHFP Phi in FrontEnd coordiantes

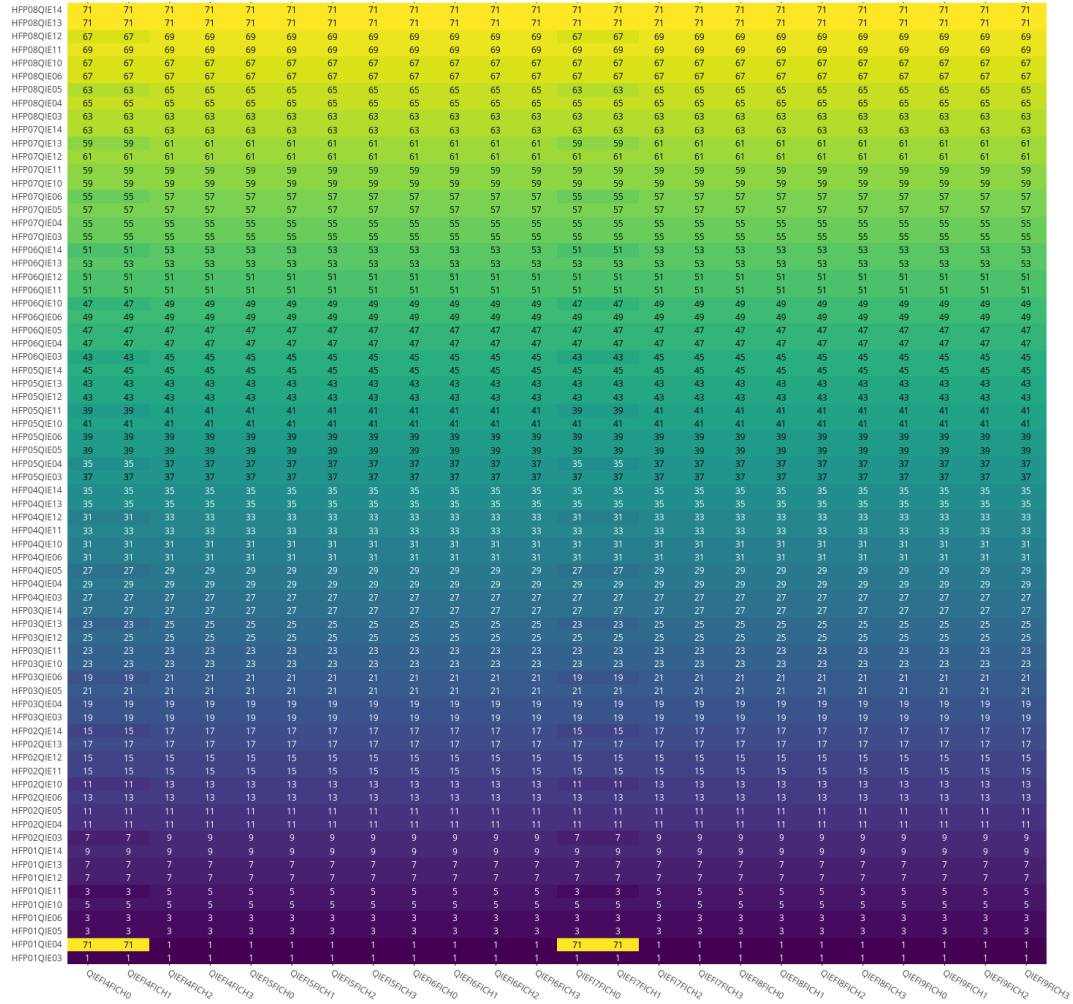


Figure B.38: HCAL (phase 1 HF, plus side) detector ϕ distribution in the frontend electronic coordinates.

ngHFP Depth in FrontEnd coordiantes

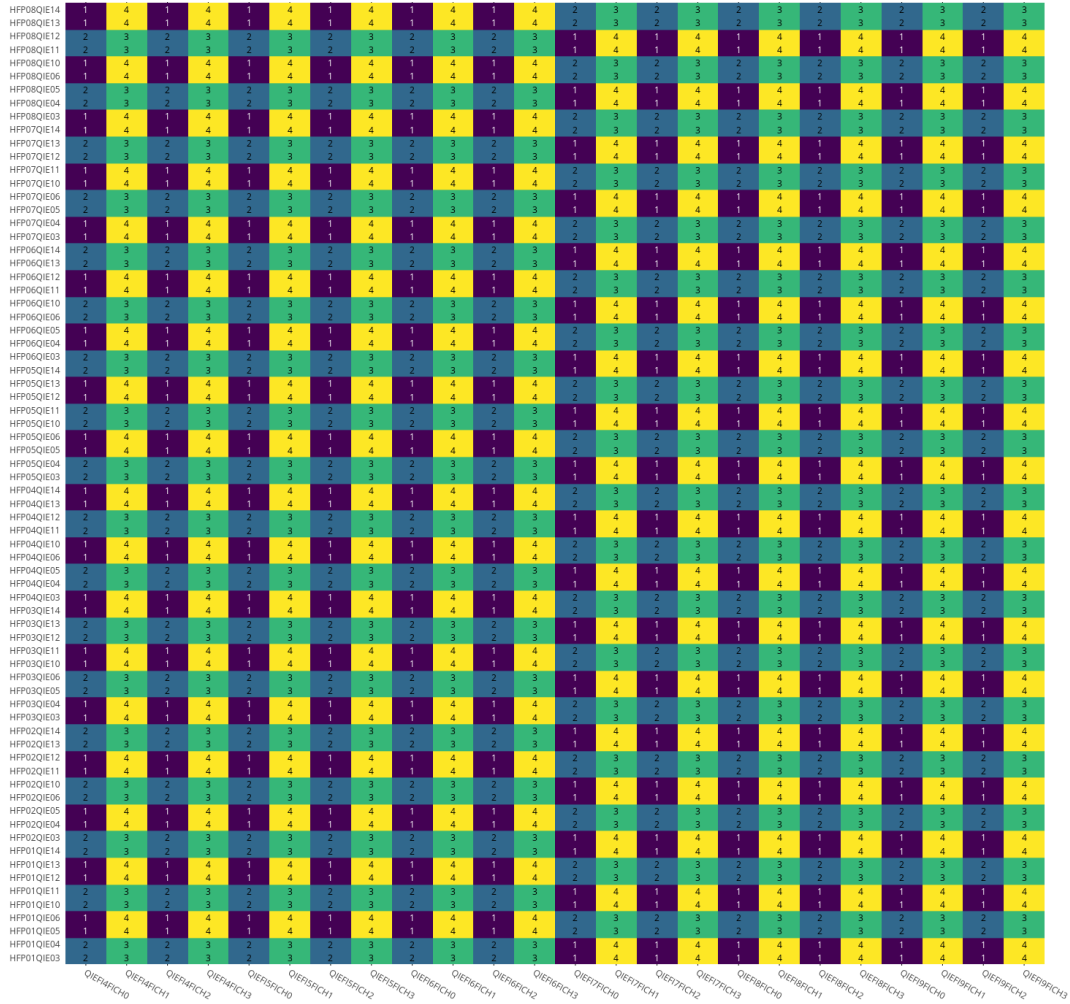


Figure B.39: HCAL (phase 1 HF, plus side) detector depth distribution in the frontend electronic coordinates.

ngHF_M Eta in FrontEnd coordiantes

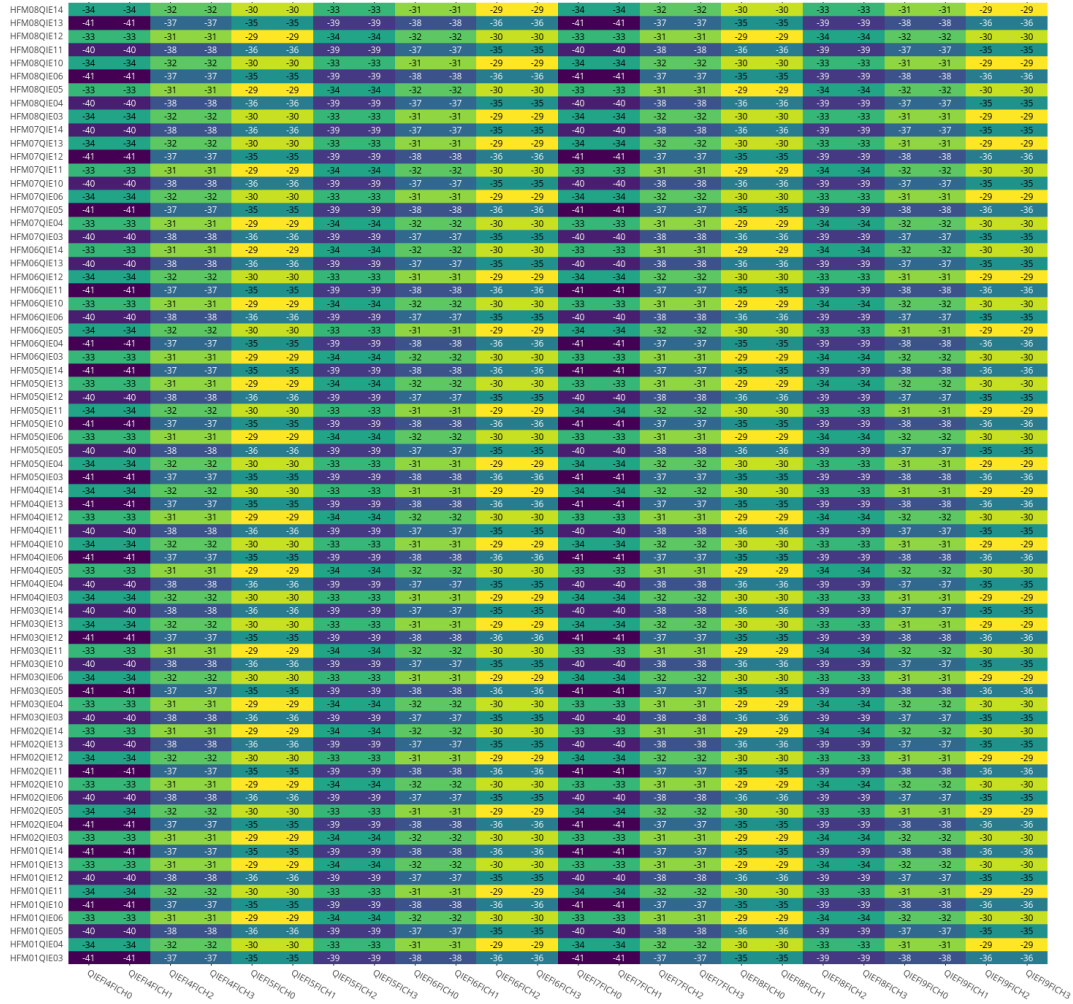


Figure B.40: HCAL (phase 1 HF, minus side) detector η distribution in the frontend electronic coordinates.

ngHFM Phi in FrontEnd coordiantes

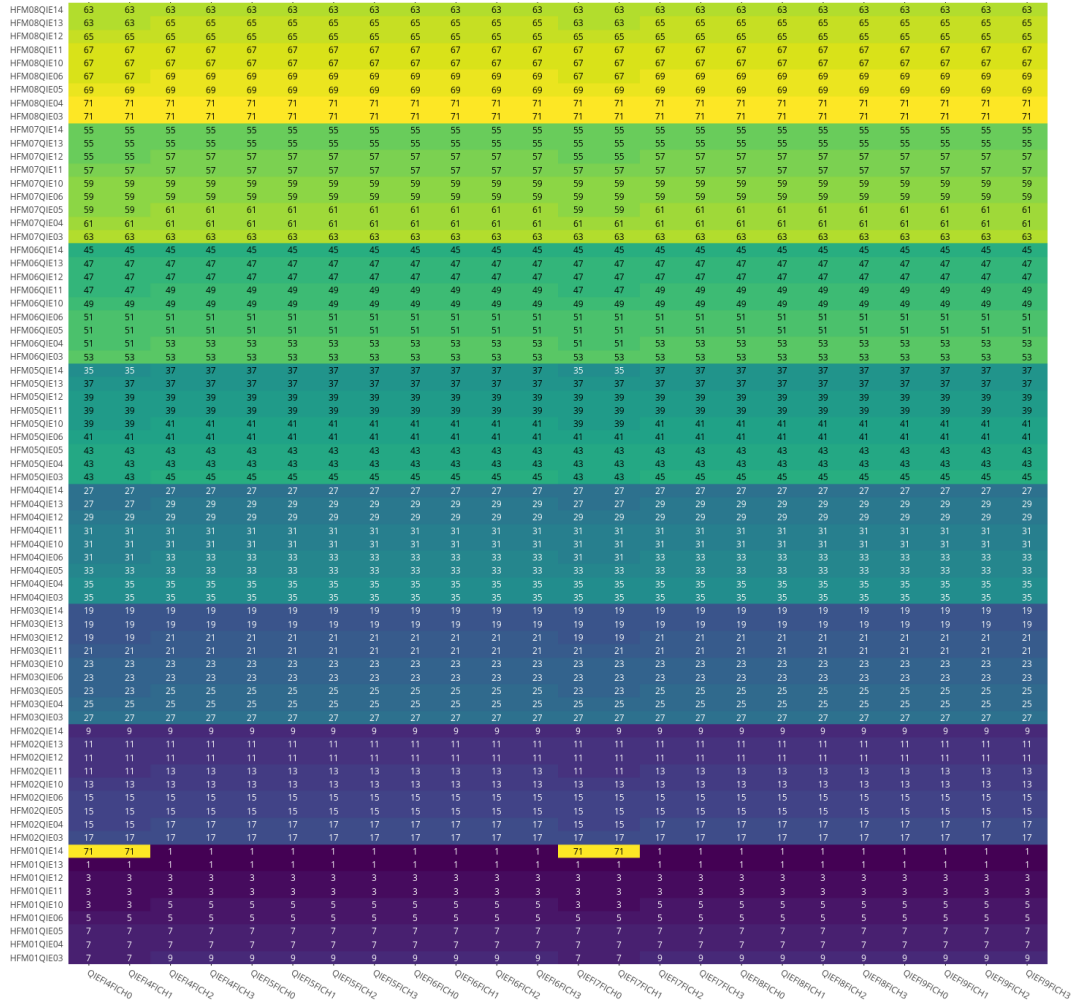


Figure B.41: HCAL (phase 1 HF, minus side) detector ϕ distribution in the frontend electronic coordinates.

ngHFM Depth in FrontEnd coordiantes

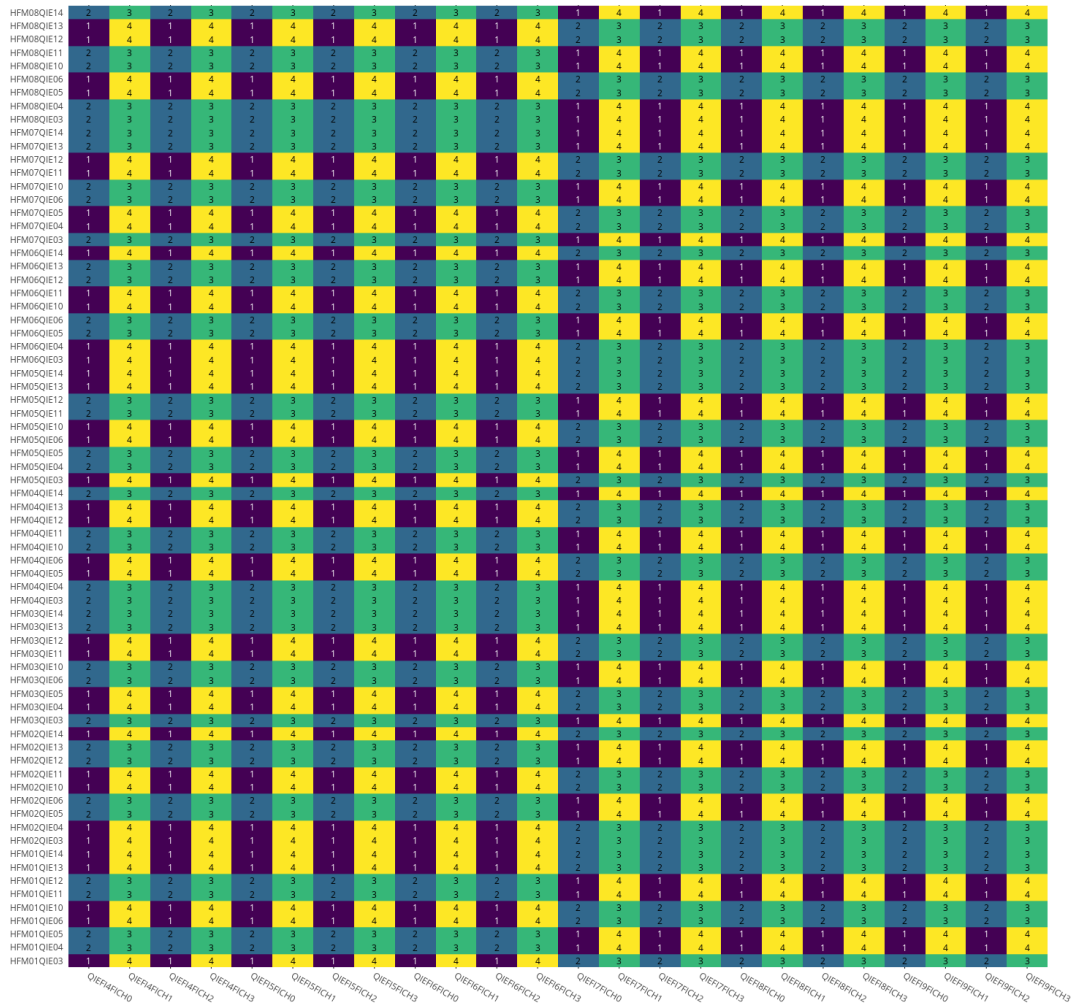


Figure B.42: HCAL (phase 1 HF, minus side) detector depth distribution in the frontend electronic coordinates.

ngHFP Crate in FrontEnd coordiantes

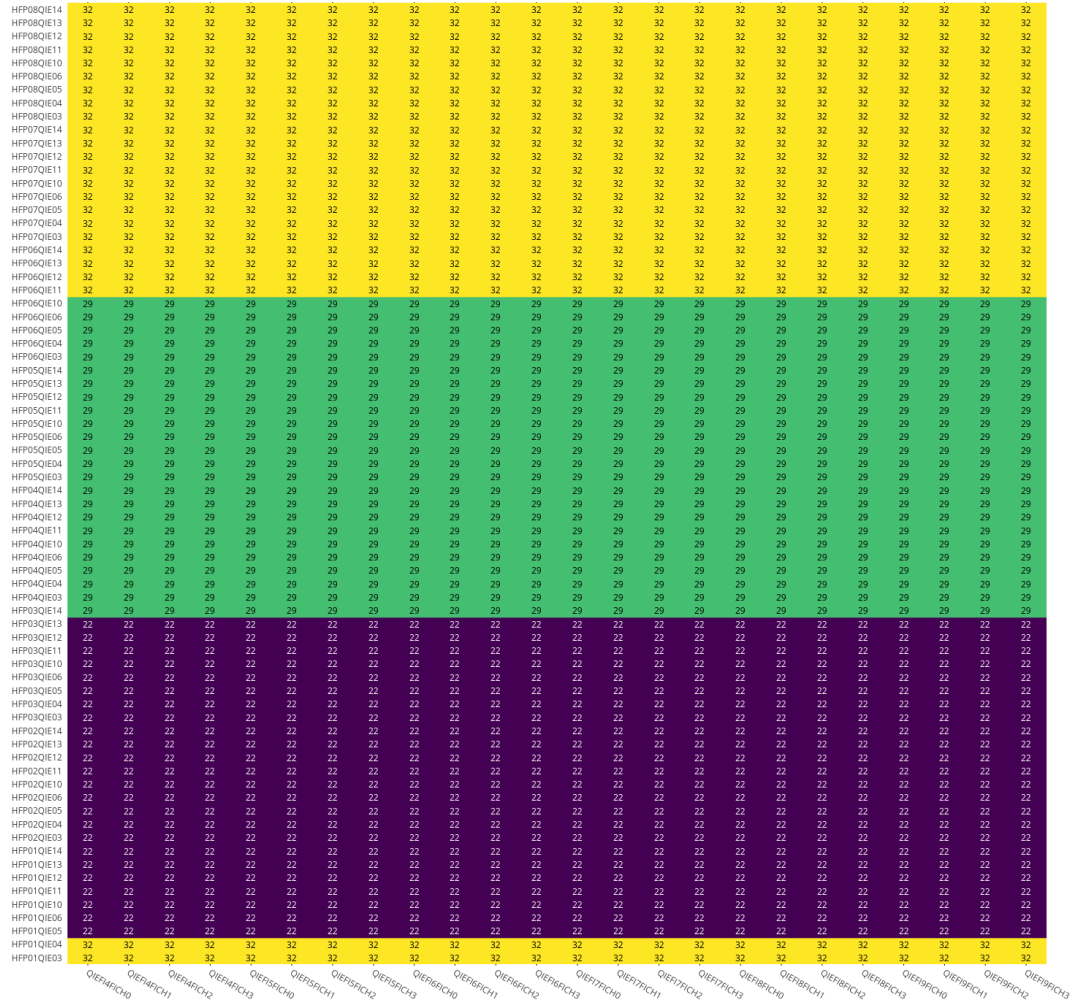


Figure B.43: HCAL (phase 1 HF, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.

ngHFP uHTR in FrontEnd coordiantes

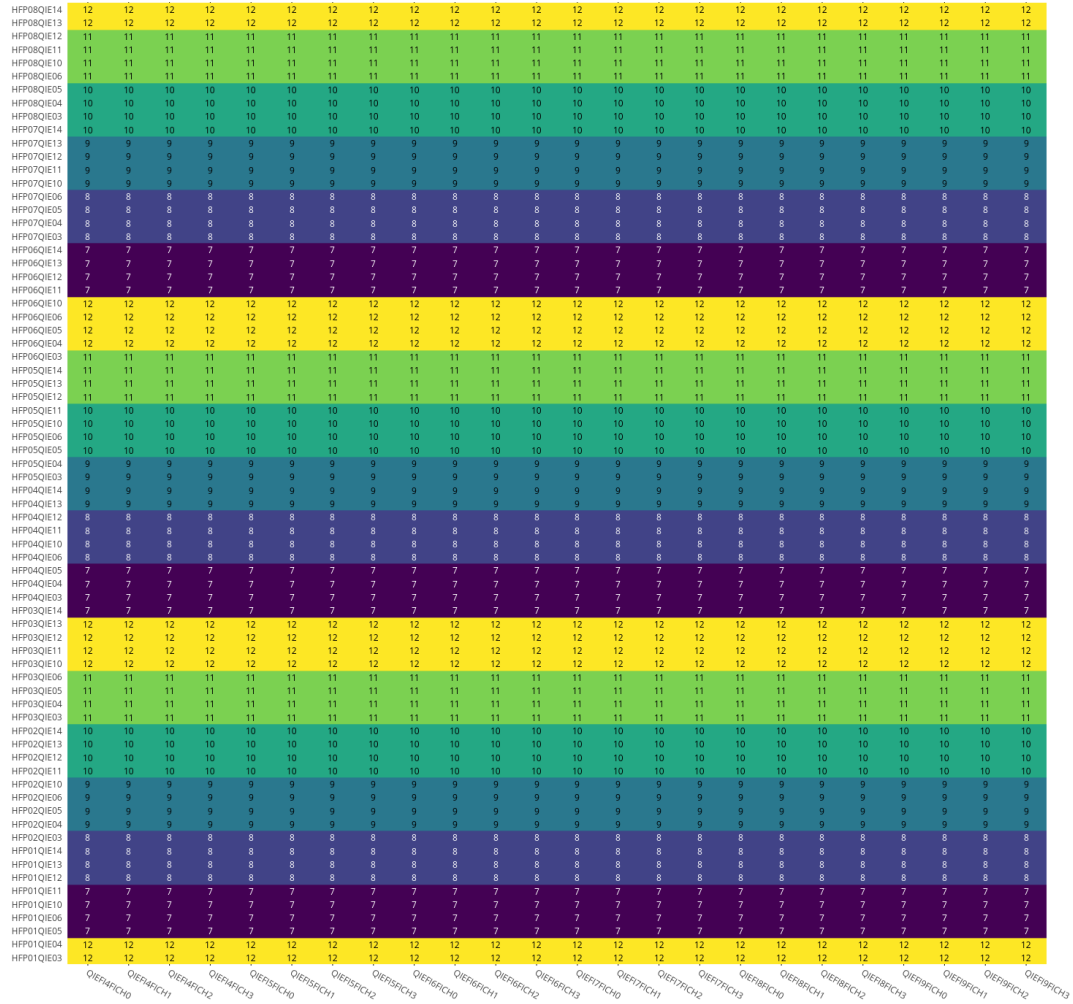


Figure B.44: HCAL (phase 1 HF, plus side) backend electronic coordinate uHTR slot distribution in the frontend electronic coordinates.

ngHFM uHTR in FrontEnd coordiantes

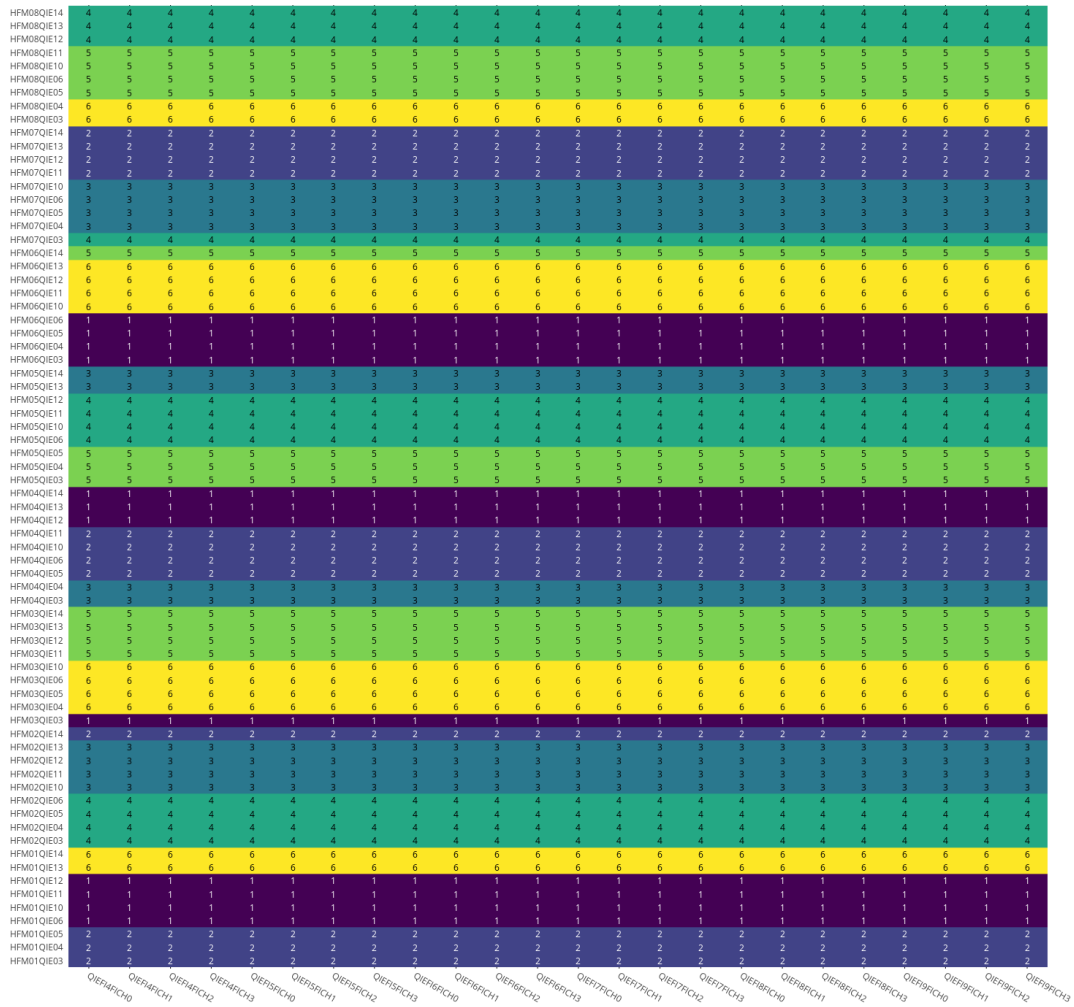


Figure B.47: HCAL (phase 1 HF, minus side) backend electronic coordinate uHTR slot distribution in the front-end electronic coordinates.

B.1.5 HO since 2015

HO has been in SiPM+QIE8 FrontEnd+VME BackEnd stage since long shutdown

1. Unlike HB, HE and HF, HO geometry scheme is more similar to muon system. There are 5 wheels in HO, labeled from -2 to +2. The central wheel (wheel 0) has 12 readout boxes per wheel, 3 readout modules per box. The other wheels have 6 readout boxes per wheel, 4 readout modules per box. HO patch panels are organized in 2015 on requirements of the trigger bits delivered to muon system (TwinMUX).

The HO validation plots are shown from Fig B.49 to Fig B.78. There are 2376 readout channels in HO. 2160 channels are physical readout channel. 216 channels are calibration channels, labeled as “HOX”. 144 of them are normal calibration channels, which connected to the patch panel and BackEnd electronics. The rest 72 HOX channels are even not connected to the patch panel. HO is the only HCAL sub-system that still use VME based BackEnd electronics. There are 4 VME crates for HO, 12 HTR per crate. Each HTR card has two FPGA, labeled as “Top” or “Bottom”. There are 8 fibers comes out from each HTR card. Each fiber has 3 channels.

H00 Eta in FrontEnd coordiantes

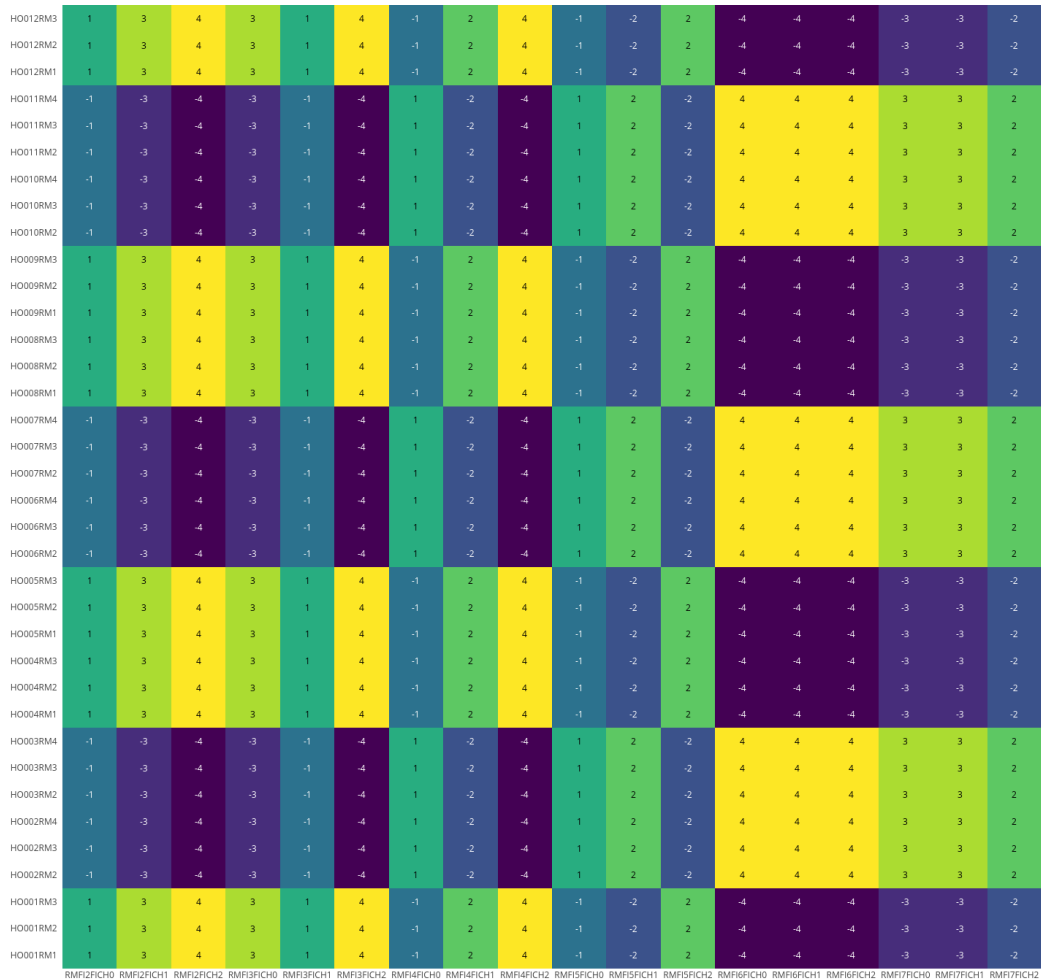


Figure B.49: HCAL (phase 1 HO, sector 0) detector η distribution in the frontend electronic coordinates.

H00 Phi in FrontEnd coordiantes

HO012RM3	70	70	70	69	69	69	70	70	69	69	69	70	69	69	70	69	70
HO012RM2	68	68	68	67	67	67	68	68	67	67	67	68	67	67	68	67	68
HO012RM1	66	66	66	65	65	65	66	66	65	65	65	66	65	65	66	65	66
HO011RM4	59	59	59	60	60	60	59	59	60	60	60	59	60	60	59	60	59
HO011RM3	61	61	61	62	62	62	61	61	62	62	62	61	62	62	61	62	61
HO011RM2	63	63	63	64	64	64	63	63	64	64	64	63	64	64	63	64	63
HO010RM4	53	53	53	54	54	54	53	53	54	54	54	53	54	54	53	54	53
HO010RM3	55	55	55	56	56	56	55	55	56	56	56	55	56	56	55	56	55
HO010RM2	57	57	57	58	58	58	57	57	58	58	58	57	58	58	57	58	57
HO009RM3	52	52	52	51	51	51	52	52	51	51	51	52	51	51	52	51	52
HO009RM2	50	50	50	49	49	49	50	50	49	49	49	50	49	49	50	49	50
HO009RM1	48	48	48	47	47	47	48	48	47	47	47	48	47	47	48	47	48
HO008RM3	46	46	46	45	45	45	46	46	45	45	45	46	45	45	46	45	46
HO008RM2	44	44	44	43	43	43	44	44	43	43	43	44	43	43	44	43	44
HO008RM1	42	42	42	41	41	41	42	42	41	41	41	42	41	41	42	41	42
HO007RM4	35	35	35	36	36	36	35	35	36	36	36	35	36	36	35	36	35
HO007RM3	37	37	37	38	38	38	37	37	38	38	38	37	38	38	37	38	37
HO007RM2	39	39	39	40	40	40	39	39	40	40	40	39	40	40	39	40	39
HO006RM4	29	29	29	30	30	30	29	29	30	30	30	29	30	30	29	30	29
HO006RM3	31	31	31	32	32	32	31	31	32	32	32	31	32	32	31	32	31
HO006RM2	33	33	33	34	34	34	33	33	34	34	34	33	34	34	33	34	33
HO005RM3	28	28	28	27	27	27	28	28	27	27	27	28	27	27	28	27	28
HO005RM2	26	26	26	25	25	25	26	26	25	25	25	26	25	25	26	25	26
HO005RM1	24	24	24	23	23	23	24	24	23	23	23	24	23	23	24	23	24
HO004RM3	22	22	22	21	21	21	22	22	21	21	21	22	21	21	22	21	22
HO004RM2	20	20	20	19	19	19	20	20	19	19	19	20	19	19	20	19	20
HO004RM1	18	18	18	17	17	17	18	18	17	17	17	18	17	17	18	17	18
HO003RM4	11	11	11	12	12	12	11	11	12	12	12	11	12	12	11	12	11
HO003RM3	13	13	13	14	14	14	13	13	14	14	14	13	14	14	13	14	13
HO003RM2	15	15	15	16	16	16	15	15	16	16	16	15	16	16	15	16	15
HO002RM4	5	5	5	6	6	6	5	5	6	6	6	5	6	6	5	6	5
HO002RM3	7	7	7	8	8	8	7	7	8	8	8	7	8	8	7	8	7
HO002RM2	9	9	9	10	10	10	9	9	10	10	10	9	10	10	9	10	9
HO001RM3	4	4	4	3	3	3	4	4	3	3	3	4	3	3	4	3	4
HO001RM2	2	2	2	1	1	1	2	2	1	1	1	2	1	1	2	1	2
HO001RM1	72	72	72	71	71	71	72	72	71	71	71	72	71	71	72	71	72

RMF2FCH0 RMF2FCH1 RMF2FCH2 RMF3FCH0 RMF3FCH1 RMF3FCH2 RMF4FCH0 RMF4FCH1 RMF4FCH2 RMF5FCH0 RMF5FCH1 RMF5FCH2 RMF6FCH0 RMF6FCH1 RMF6FCH2 RMF7FCH0 RMF7FCH1 RMF7FCH2

Figure B.50: HCAL (phase 1 HO, sector 0) detector ϕ distribution in the frontend electronic coordinates.

HO1P Eta in FrontEnd coordiantes

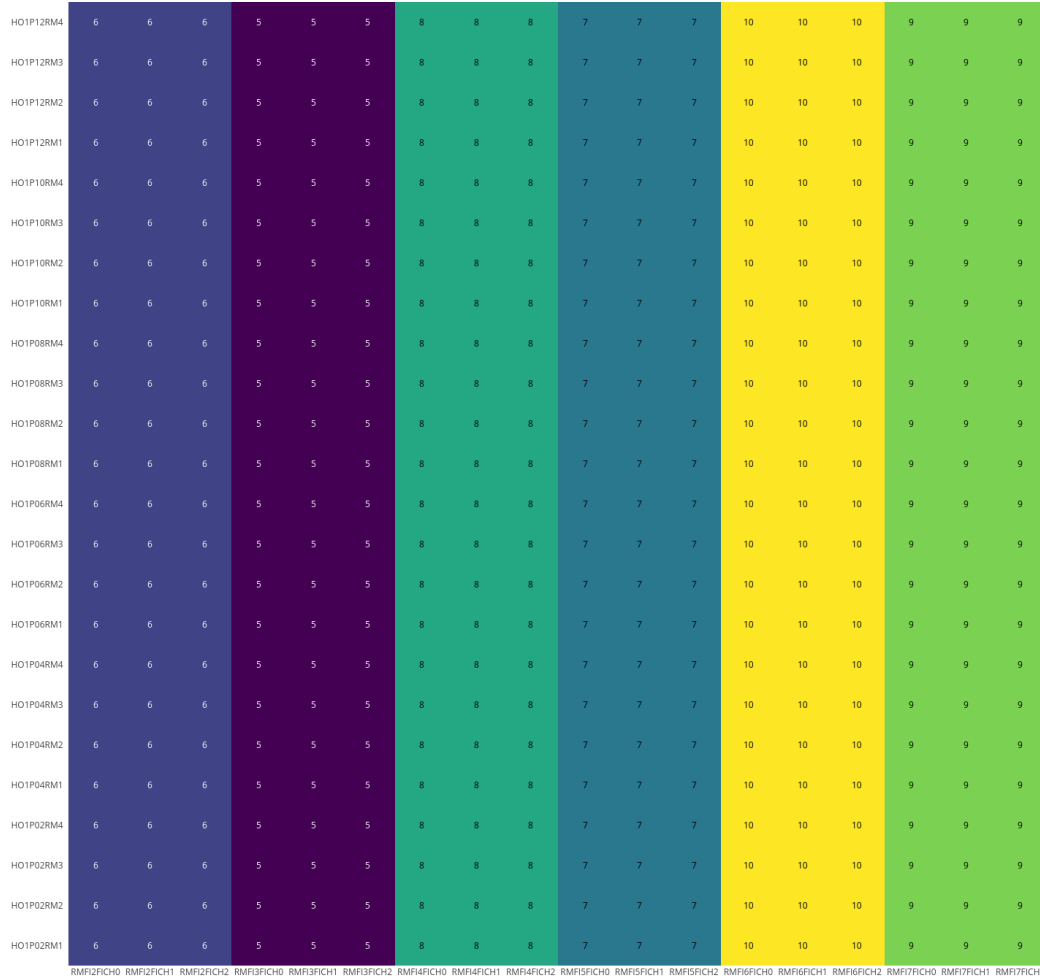


Figure B.51: HCAL (phase 1 HO, sector 1, plus side) detector η distribution in the frontend electronic coordinates.

HO1P Phi in FrontEnd coordiantes

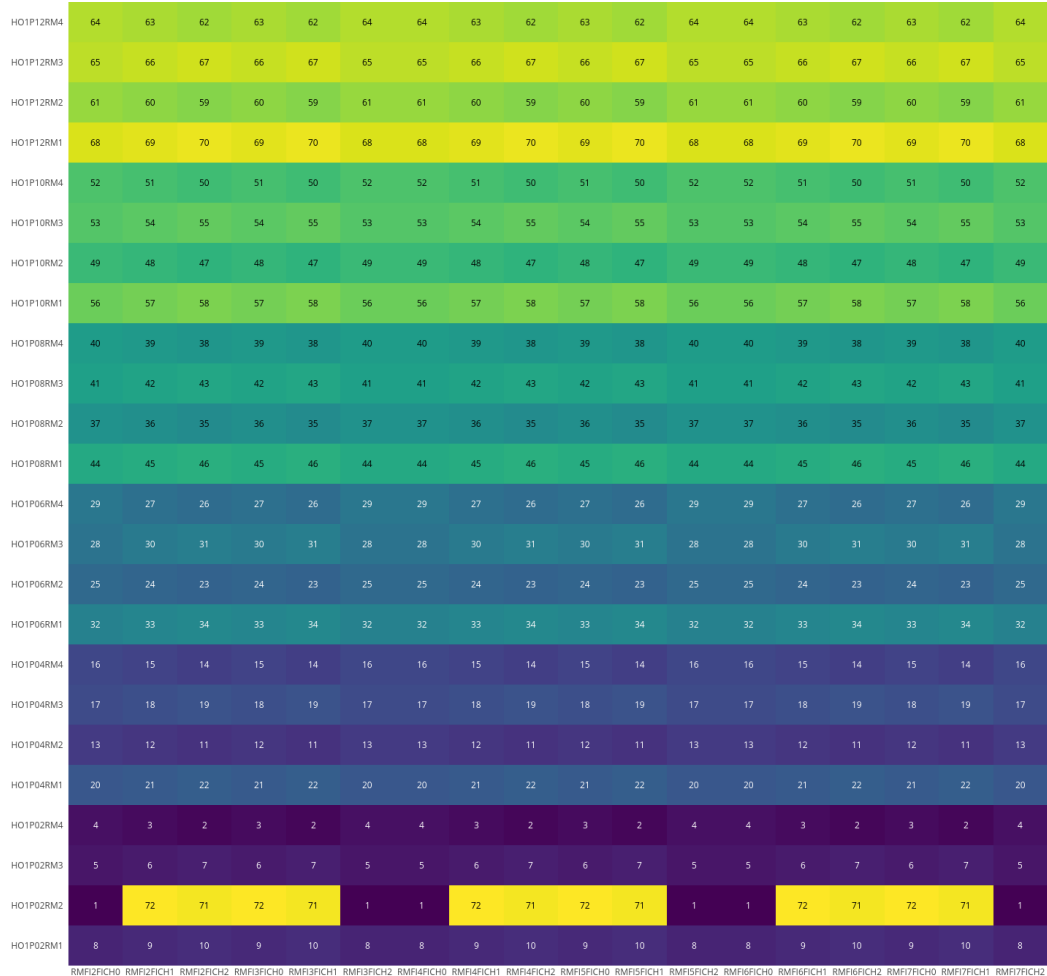


Figure B.52: HCAL (phase 1 HO, sector 1, plus side) detector ϕ distribution in the frontend electronic coordinates.

HO1M Eta in FrontEnd coordiantes

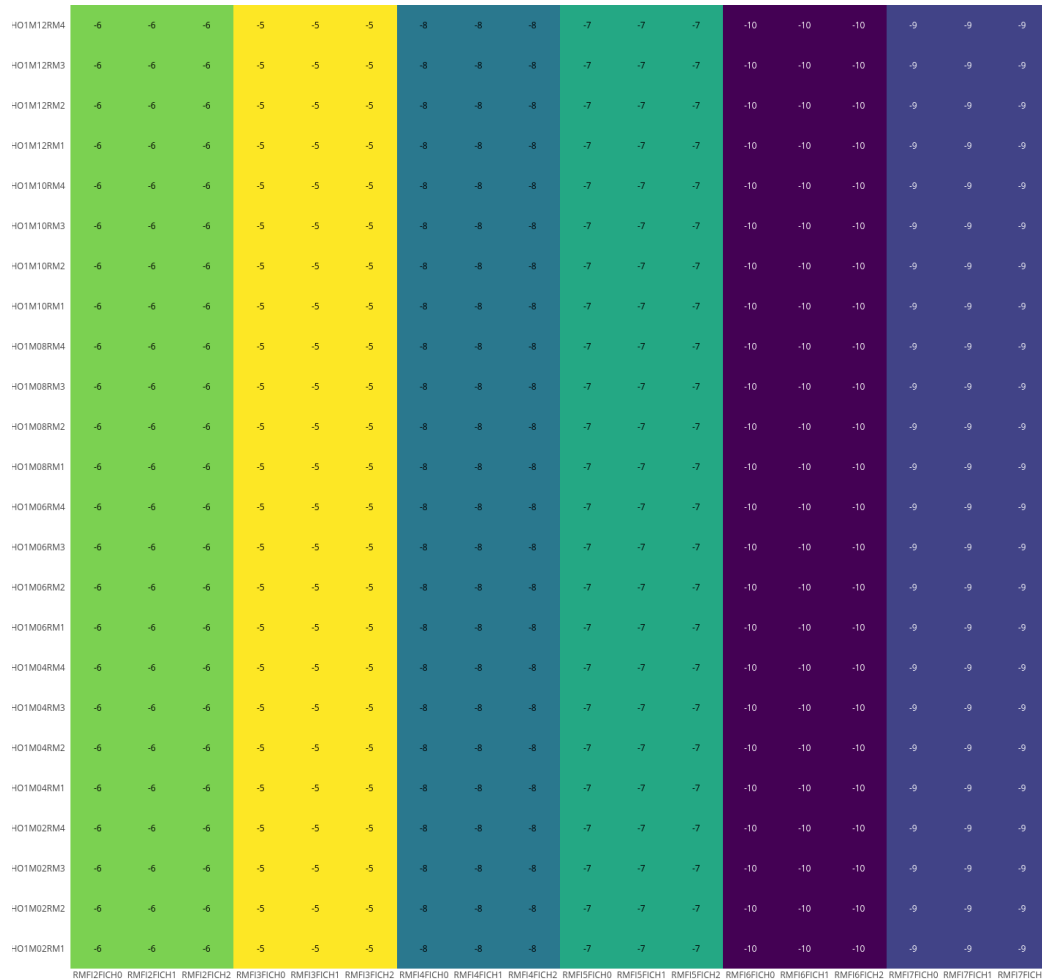


Figure B.53: HCAL (phase 1 HO, sector 1, minus side) detector η distribution in the frontend electronic coordinates.

HO1M Phi in FrontEnd coordiantes

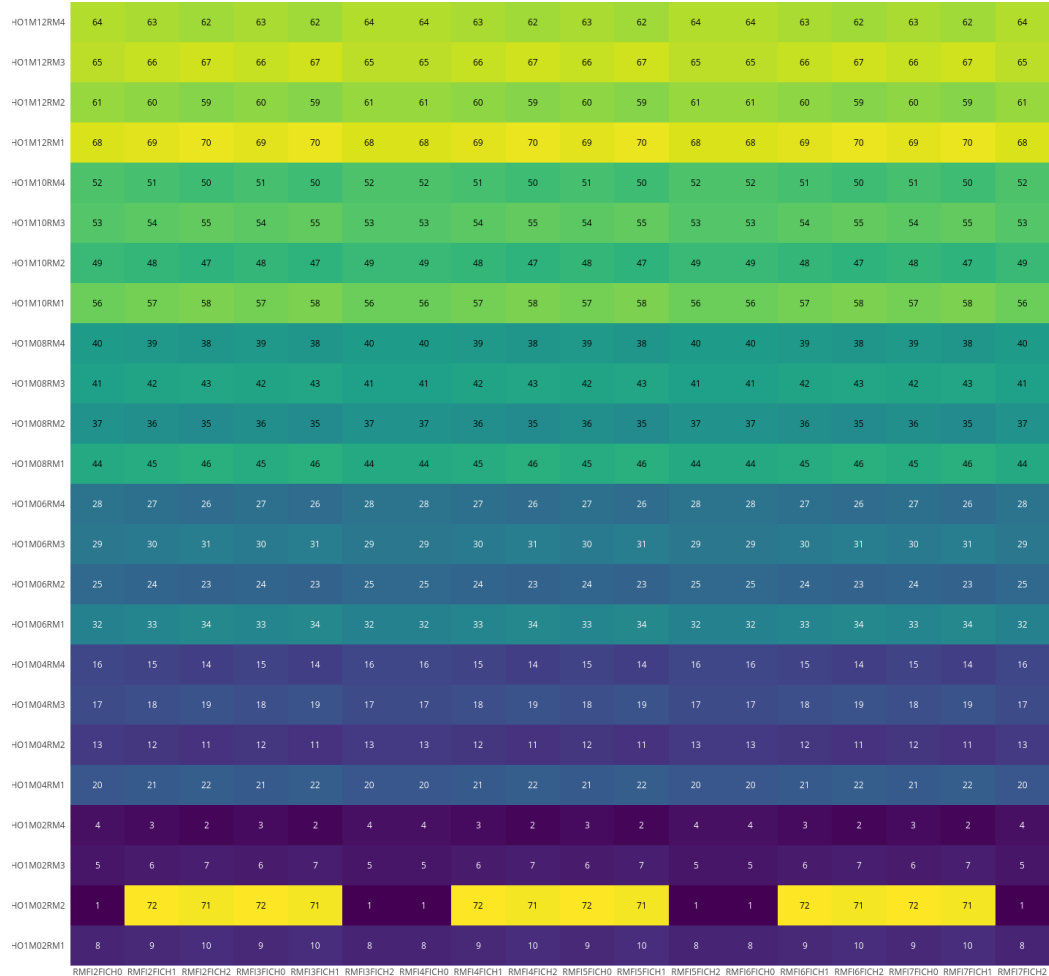


Figure B.54: HCAL (phase 1 HO, sector 1, minus side) detector ϕ distribution in the front-end electronic coordinates.

HO2P Eta in FrontEnd coordiantes

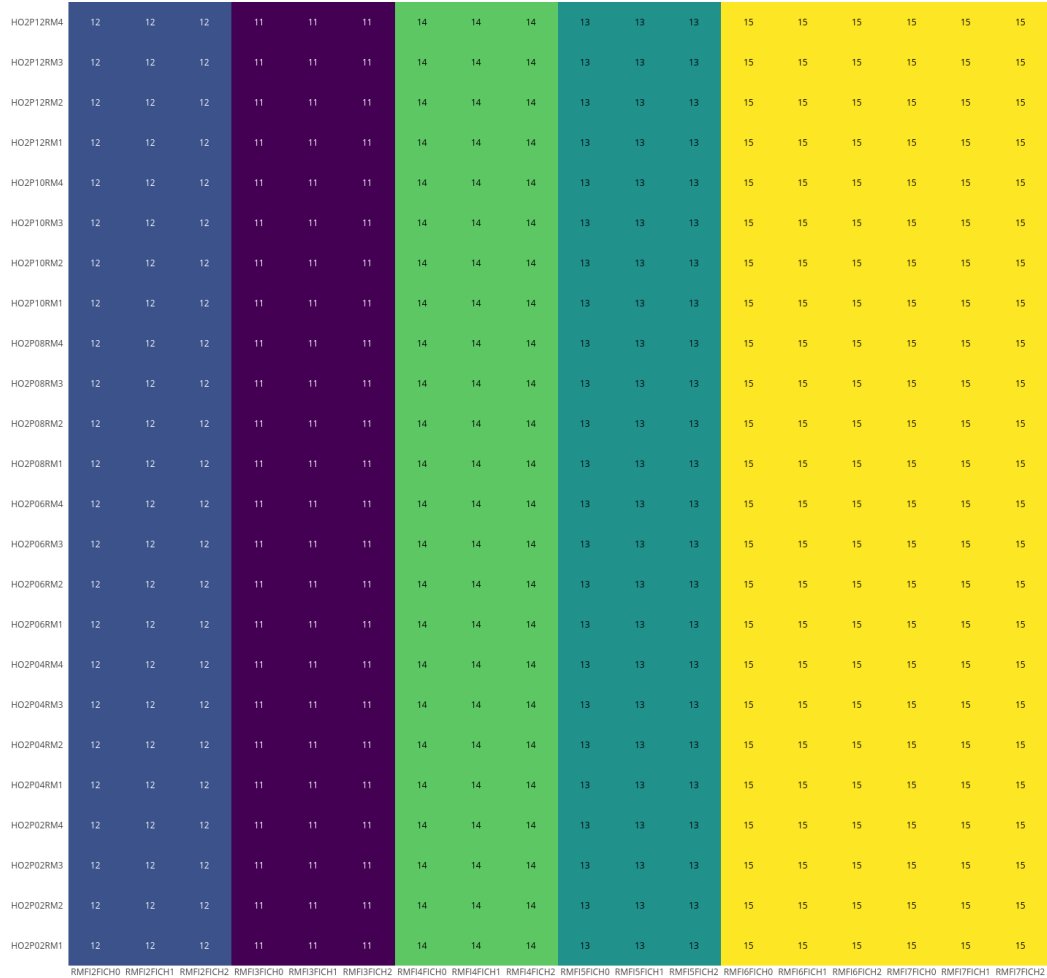


Figure B.55: HCAL (phase 1 HO, sector 2, plus side) detector η distribution in the frontend electronic coordinates.

HO2P Phi in FrontEnd coordiantes

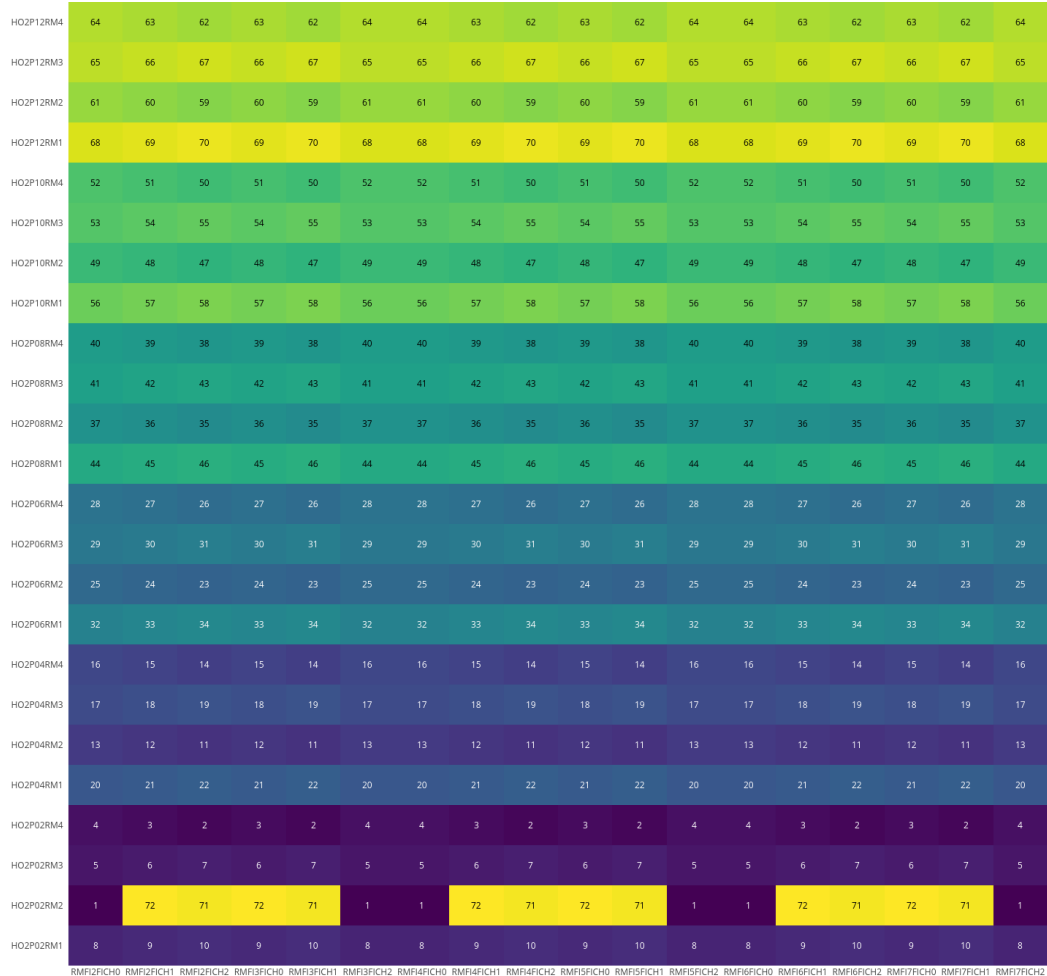


Figure B.56: HCAL (phase 1 HO, sector 2, plus side) detector ϕ distribution in the frontend electronic coordinates.

HO2M Eta in FrontEnd coordiantes

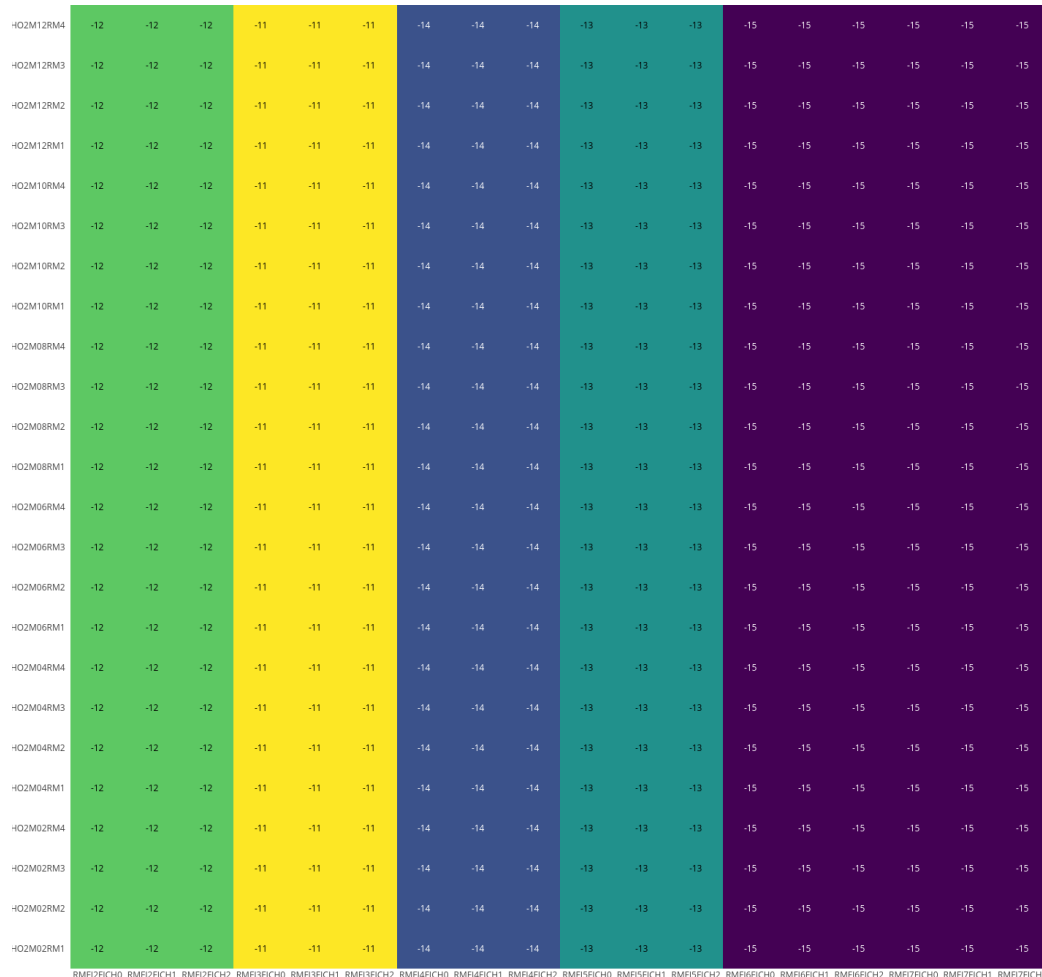


Figure B.57: HCAL (phase 1 HO, sector 2, minus side) detector η distribution in the frontend electronic coordinates.

HO1M Phi in FrontEnd coordiantes

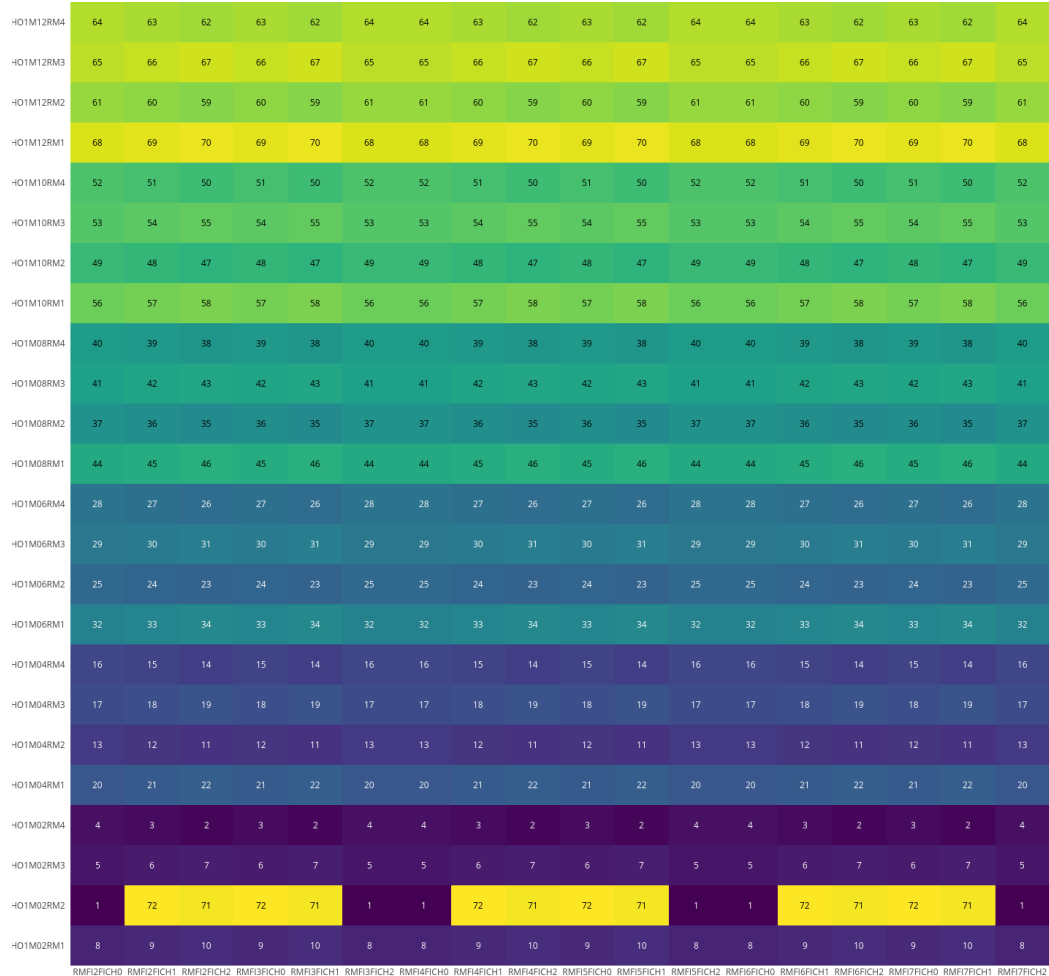


Figure B.58: HCAL (phase 1 HO, sector 2, minus side) detector ϕ distribution in the front-end electronic coordinates.

H00 Crate in FrontEnd coordiantes

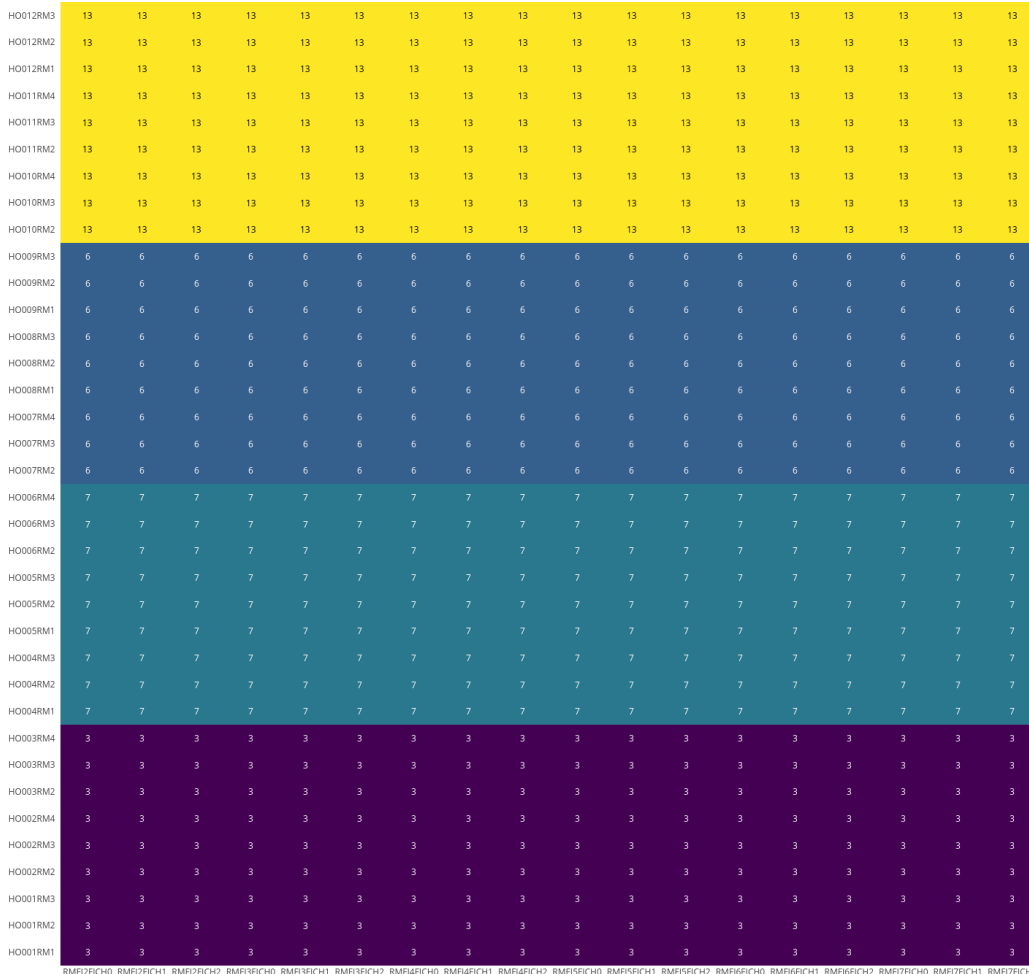


Figure B.59: HCAL (phase 1 HO, sector 0) backend electronic coordinate crate distribution in the frontend electronic coordinates.

HO0 HTR in FrontEnd coordiantes

HO012RM3	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO012RM2	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO012RM1	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO011RM4	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO011RM3	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO011RM2	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO010RM4	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO010RM3	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO010RM2	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO009RM3	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO009RM2	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO009RM1	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO008RM3	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO008RM2	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO008RM1	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO007RM4	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO007RM3	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO007RM2	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO006RM4	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO006RM3	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO006RM2	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO005RM3	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO005RM2	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO005RM1	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO004RM3	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO004RM2	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO004RM1	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO003RM4	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO003RM3	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO003RM2	17	17	17	17	17	17	16	16	16	16	16	16	17	17	17	16	16	16
HO002RM4	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO002RM3	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO002RM2	13	13	13	13	13	13	7	7	7	7	7	7	13	13	13	7	7	7
HO001RM3	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO001RM2	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3
HO001RM1	4	4	4	4	4	4	3	3	3	3	3	3	4	4	4	3	3	3

RMF12FICH0 RMF12FICH1 RMF12FICH2 RMF13FICH0 RMF13FICH1 RMF13FICH2 RMF14FICH0 RMF14FICH1 RMF14FICH2 RMF15FICH0 RMF15FICH1 RMF15FICH2 RMF16FICH0 RMF16FICH1 RMF16FICH2 RMF17FICH0 RMF17FICH1 RMF17FICH2

Figure B.60: HCAL (phase 1 HO, sector 0) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.

H00 HTR_TB in FrontEnd coordiantes

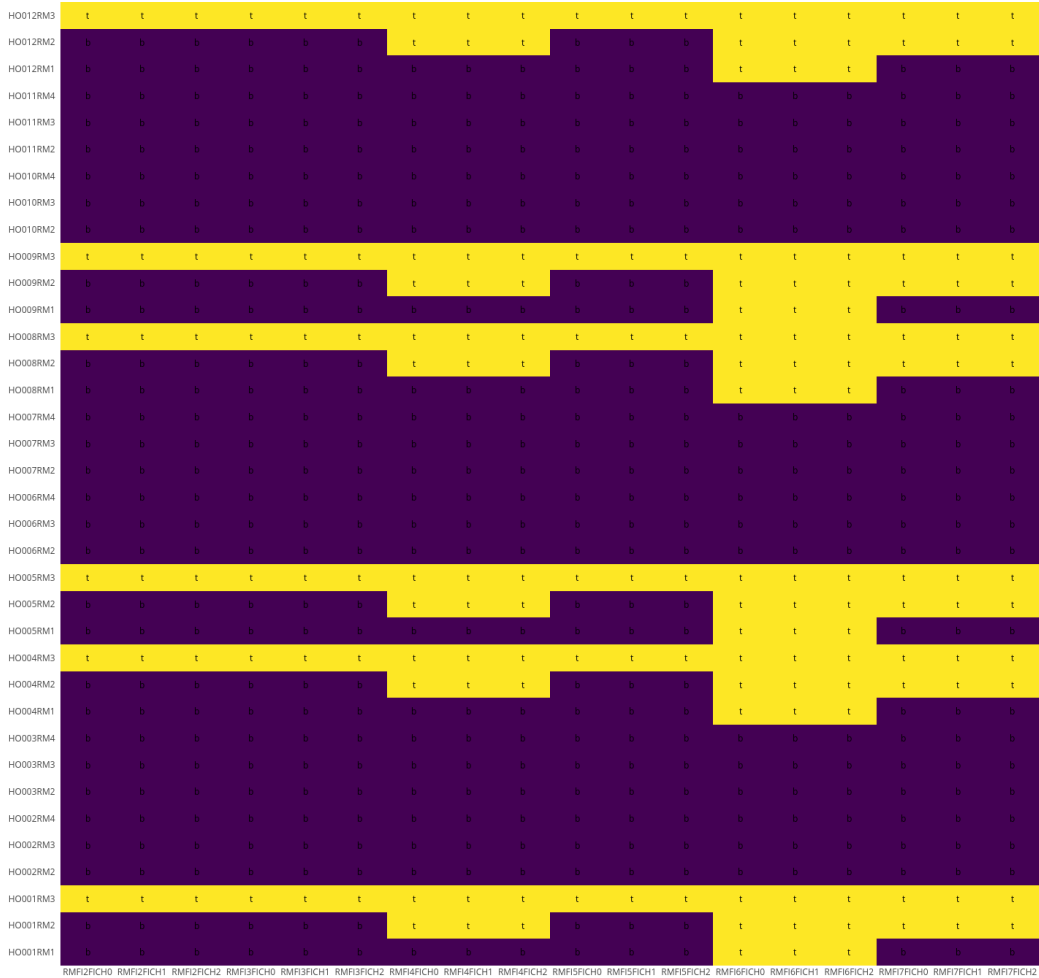


Figure B.61: HCAL (phase 1 HO, sector 0) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.

H00 HTR_FI in FrontEnd coordiantes

HO012RM3	7	7	7	8	8	8	5	5	5	6	6	6	4	4	4	4	4	4
HO012RM2	4	4	4	5	5	5	8	8	8	4	4	4	5	5	5	7	7	7
HO012RM1	6	6	6	7	7	7	6	6	6	7	7	7	6	6	6	5	5	5
HO011RM4	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6	5	5	5
HO011RM3	5	5	5	4	4	4	4	4	4	8	8	8	5	5	5	7	7	7
HO011RM2	8	8	8	7	7	7	6	6	6	5	5	5	4	4	4	4	4	4
HO010RM4	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6	5	5	5
HO010RM3	5	5	5	4	4	4	4	4	4	8	8	8	5	5	5	7	7	7
HO010RM2	8	8	8	7	7	7	6	6	6	5	5	5	4	4	4	4	4	4
HO009RM3	7	7	7	8	8	8	5	5	5	6	6	6	4	4	4	4	4	4
HO009RM2	4	4	4	5	5	5	8	8	8	4	4	4	5	5	5	7	7	7
HO009RM1	6	6	6	7	7	7	6	6	6	7	7	7	6	6	6	5	5	5
HO008RM3	7	7	7	8	8	8	5	5	5	6	6	6	4	4	4	4	4	4
HO008RM2	4	4	4	5	5	5	8	8	8	4	4	4	5	5	5	7	7	7
HO008RM1	6	6	6	7	7	7	6	6	6	7	7	7	6	6	6	5	5	5
HO007RM4	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6	5	5	5
HO007RM3	5	5	5	4	4	4	4	4	4	8	8	8	5	5	5	7	7	7
HO007RM2	8	8	8	7	7	7	6	6	6	5	5	5	4	4	4	4	4	4
HO006RM4	6	6	6	7	7	7	7	7	7	6	6	6	6	6	6	5	5	5
HO006RM3	4	4	4	5	5	5	4	4	4	8	8	8	5	5	5	7	7	7
HO006RM2	8	8	8	7	7	7	6	6	6	5	5	5	4	4	4	4	4	4
HO005RM3	7	7	7	8	8	8	5	5	5	6	6	6	4	4	4	4	4	4
HO005RM2	4	4	4	5	5	5	8	8	8	4	4	4	5	5	5	7	7	7
HO005RM1	6	6	6	7	7	7	6	6	6	7	7	7	6	6	6	5	5	5
HO004RM3	7	7	7	8	8	8	5	5	5	6	6	6	4	4	4	4	4	4
HO004RM2	4	4	4	5	5	5	8	8	8	4	4	4	5	5	5	7	7	7
HO004RM1	6	6	6	7	7	7	6	6	6	7	7	7	6	6	6	5	5	5
HO003RM4	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6	5	5	5
HO003RM3	5	5	5	4	4	4	4	4	4	8	8	8	5	5	5	7	7	7
HO003RM2	8	8	8	7	7	7	6	6	6	5	5	5	4	4	4	4	4	4
HO002RM4	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6	5	5	5
HO002RM3	5	5	5	4	4	4	4	4	4	8	8	8	5	5	5	7	7	7
HO002RM2	8	8	8	7	7	7	6	6	6	5	5	5	4	4	4	4	4	4
HO001RM3	7	7	7	8	8	8	5	5	5	6	6	6	4	4	4	4	4	4
HO001RM2	4	4	4	5	5	5	8	8	8	4	4	4	5	5	5	7	7	7
HO001RM1	6	6	6	7	7	7	6	6	6	7	7	7	6	6	6	5	5	5

Figure B.62: HCAL (phase 1 HO, sector 0) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.

HO1P Crate in FrontEnd coordiantes

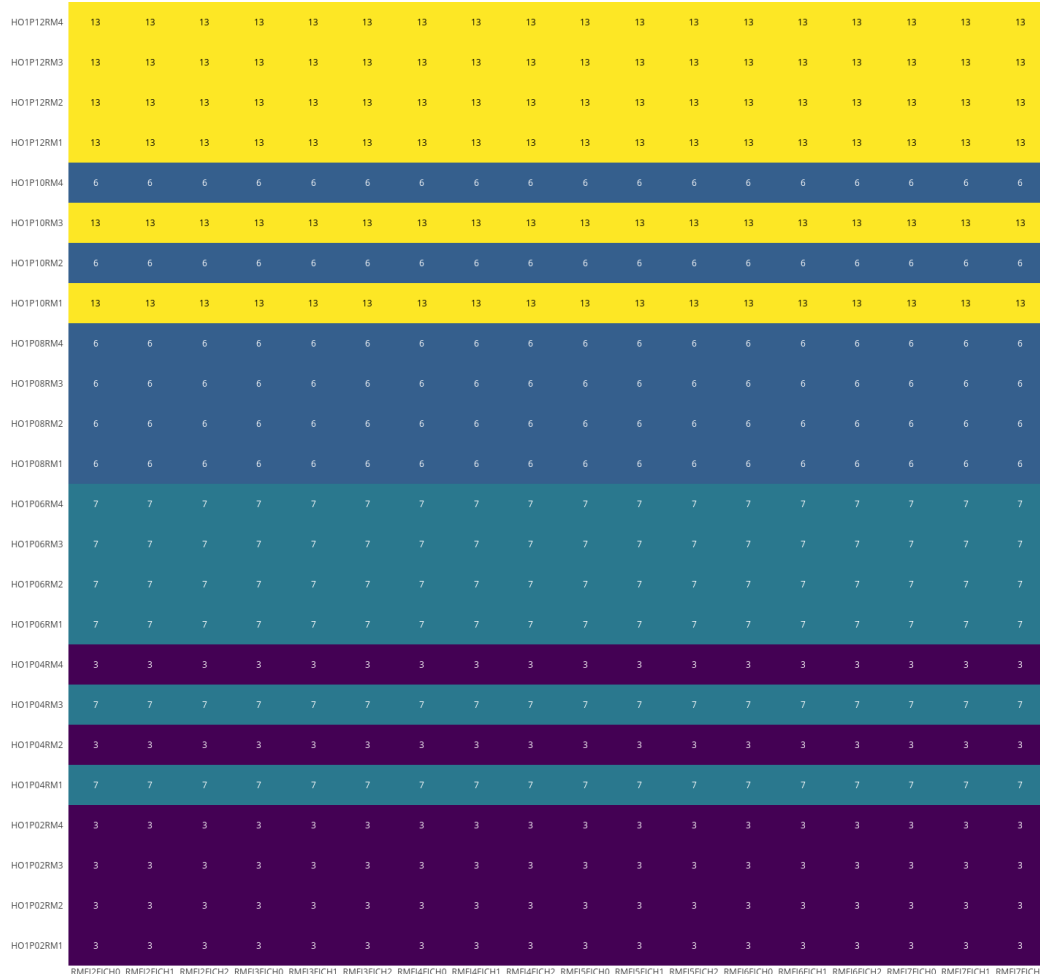


Figure B.63: HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.

HO1P HTR in FrontEnd coordiantes

HO1P12RM4	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
HO1P12RM3	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
HO1P12RM2	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
HO1P12RM1	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
HO1P10RM4	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
HO1P10RM3	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
HO1P10RM2	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
HO1P10RM1	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
HO1P08RM4	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
HO1P08RM3	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
HO1P08RM2	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
HO1P08RM1	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
HO1P06RM4	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
HO1P06RM3	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
HO1P06RM2	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
HO1P06RM1	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
HO1P04RM4	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
HO1P04RM3	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
HO1P04RM2	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
HO1P04RM1	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
HO1P02RM4	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
HO1P02RM3	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
HO1P02RM2	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
HO1P02RM1	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14

RMF2FCH0 RMF2FCH1 RMF2FCH2 RMF3FCH0 RMF3FCH1 RMF3FCH2 RMF4FCH0 RMF4FCH1 RMF4FCH2 RMF5FCH0 RMF5FCH1 RMF5FCH2 RMF6FCH0 RMF6FCH1 RMF6FCH2 RMF7FCH0 RMF7FCH1 RMF7FCH2

Figure B.64: HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.

HO1P HTR_TB in FrontEnd coordiantes

HO1P12RM4	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P12RM3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P12RM2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P12RM1	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P10RM4	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P10RM3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P10RM2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P10RM1	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P08RM4	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P08RM3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P08RM2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P08RM1	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P06RM4	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P06RM3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P06RM2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P06RM1	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P04RM4	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P04RM3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P04RM2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P04RM1	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P02RM4	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
HO1P02RM3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P02RM2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
HO1P02RM1	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t

RMF2FCH0 RMF2FCH1 RMF2FCH2 RMF3FCH0 RMF3FCH1 RMF3FCH2 RMF4FCH0 RMF4FCH1 RMF4FCH2 RMF5FCH0 RMF5FCH1 RMF5FCH2 RMF6FCH0 RMF6FCH1 RMF6FCH2 RMF7FCH0 RMF7FCH1 RMF7FCH2

Figure B.65: HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.

HO1P HTR_FI in FrontEnd coordiantes

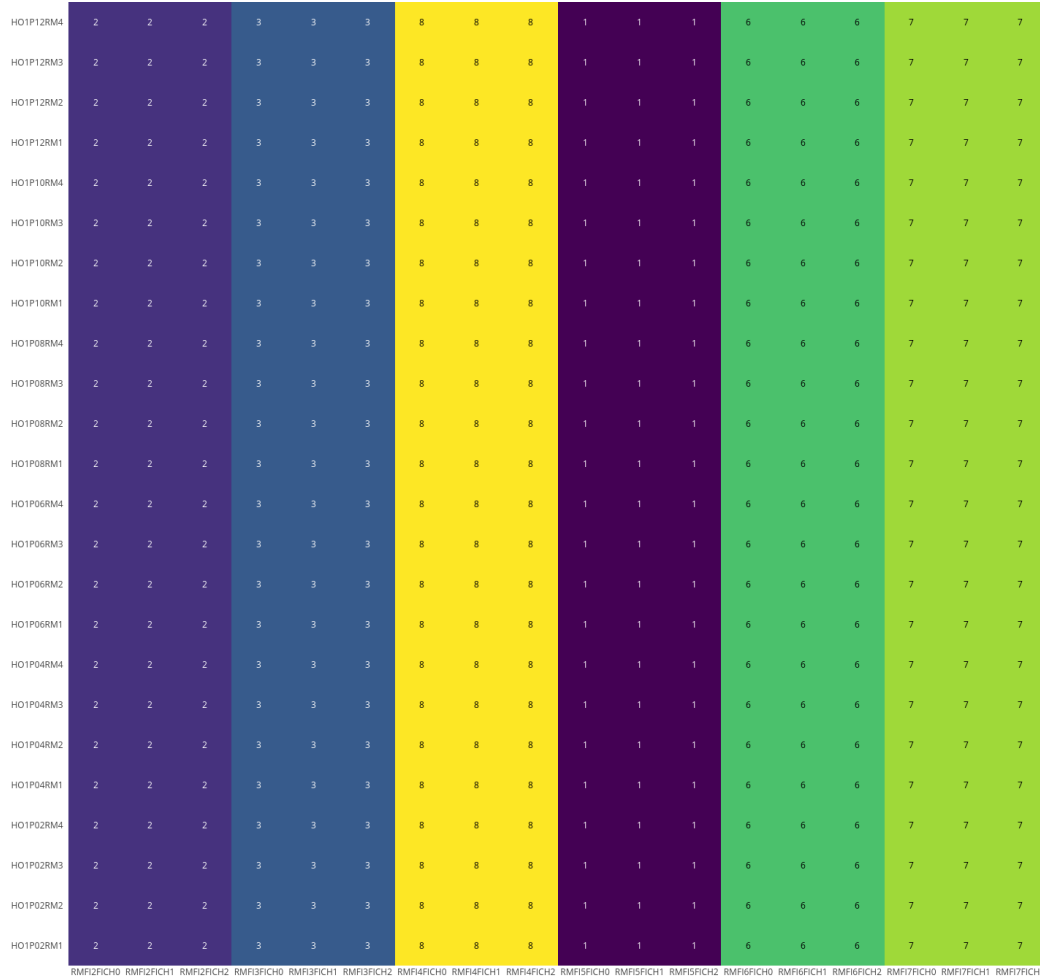


Figure B.66: HCAL (phase 1 HO, sector 1, plus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.

HO1M Crate in FrontEnd coordiantes

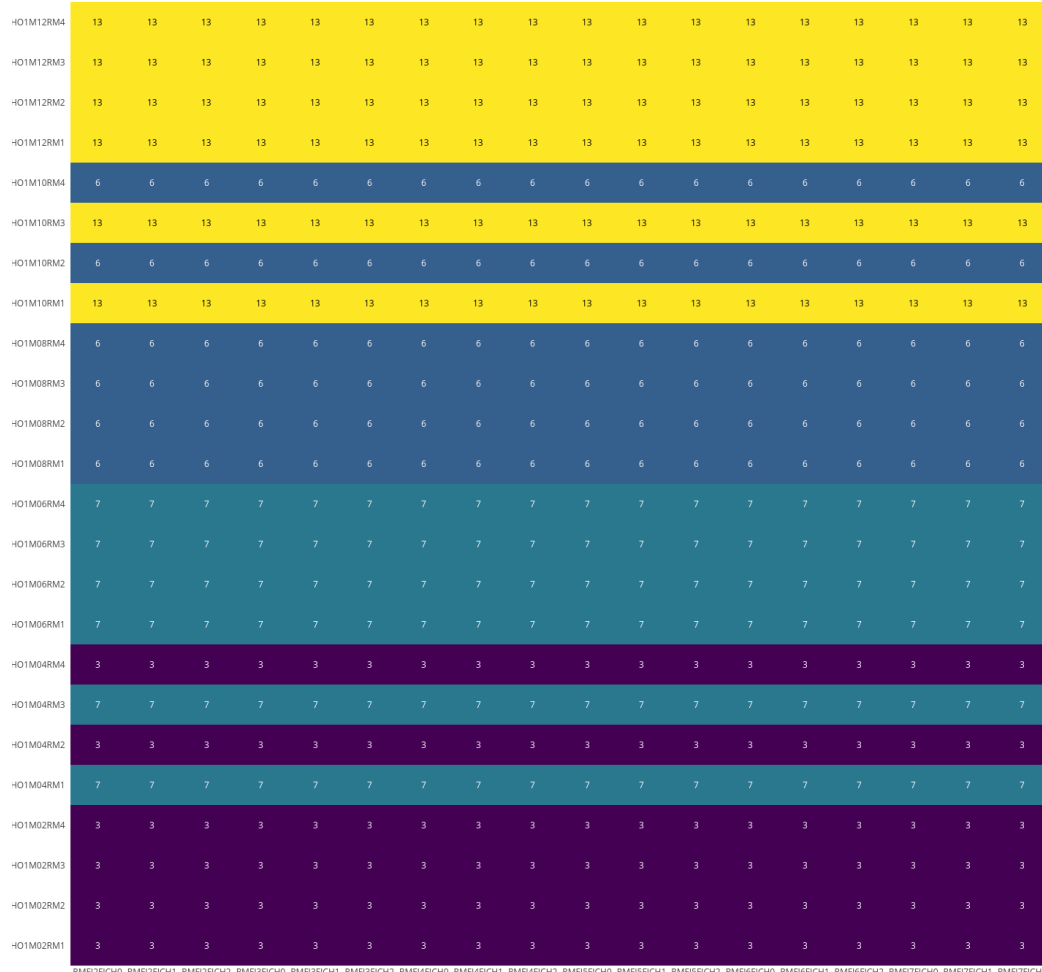


Figure B.67: HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.

HO1M HTR in FrontEnd coordiantes

-HO1M12RM4	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
-HO1M12RM3	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
-HO1M12RM2	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
-HO1M12RM1	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
-HO1M10RM4	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
-HO1M10RM3	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
-HO1M10RM2	17	17	17	17	17	17	18	18	18	17	17	17	18	18	18	18	18	18
-HO1M10RM1	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
-HO1M08RM4	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
-HO1M08RM3	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
-HO1M08RM2	3	3	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
-HO1M08RM1	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
-HO1M06RM4	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
-HO1M06RM3	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
-HO1M06RM2	13	13	13	13	13	13	14	14	14	13	13	13	14	14	14	14	14	14
-HO1M06RM1	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
-HO1M04RM4	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
-HO1M04RM3	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
-HO1M04RM2	16	16	16	16	16	16	15	15	15	16	16	16	15	15	15	15	15	15
-HO1M04RM1	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
-HO1M02RM4	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
-HO1M02RM3	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6
-HO1M02RM2	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5	5	5
-HO1M02RM1	7	7	7	7	7	7	6	6	6	7	7	7	6	6	6	6	6	6

RMF12FCH0 RMF12FCH1 RMF12FCH2 RMF13FCH0 RMF13FCH1 RMF13FCH2 RMF14FCH0 RMF14FCH1 RMF14FCH2 RMF15FCH0 RMF15FCH1 RMF15FCH2 RMF16FCH0 RMF16FCH1 RMF16FCH2 RMF17FCH0 RMF17FCH1 RMF17FCH2

Figure B.68: HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.

HO1M HTR_TB in FrontEnd coordiantes

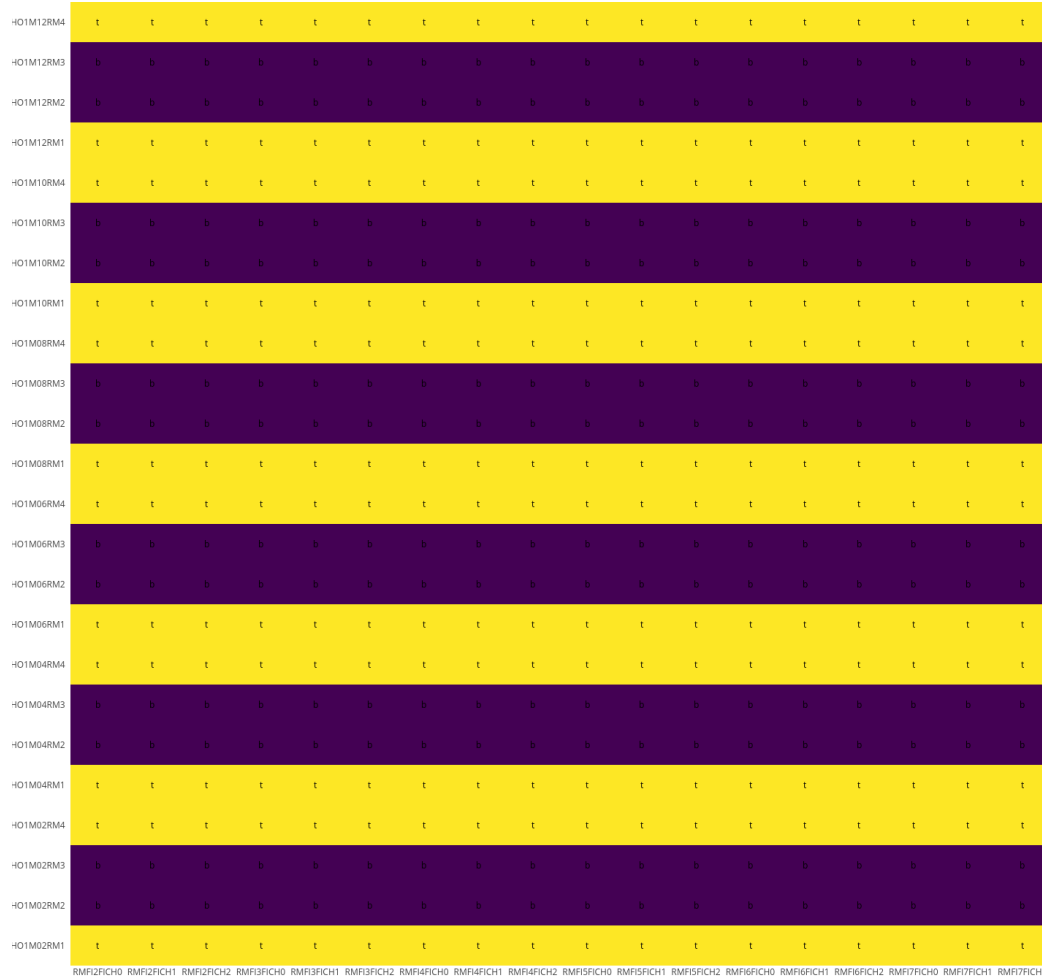


Figure B.69: HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.

HO1M HTR_FI in FrontEnd coordiantes

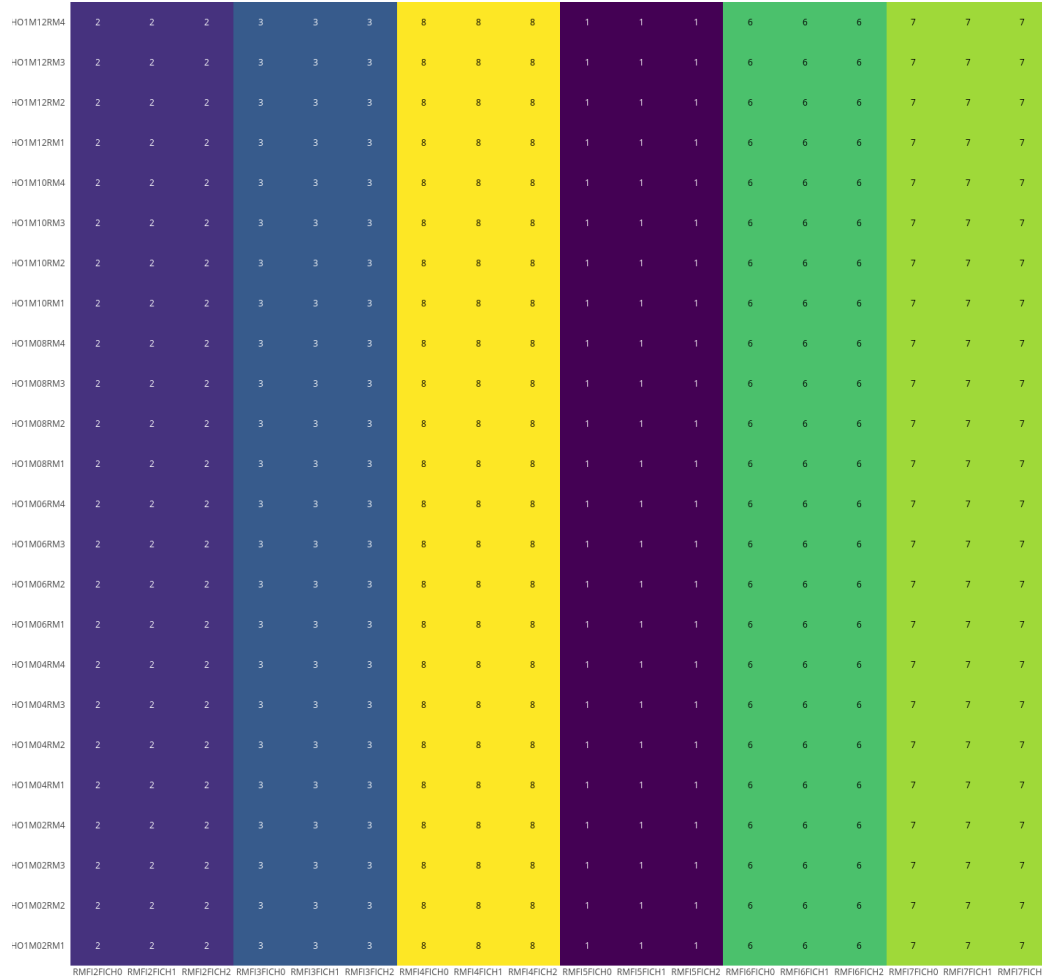


Figure B.70: HCAL (phase 1 HO, sector 1, minus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.

HO2P Crate in FrontEnd coordiantes

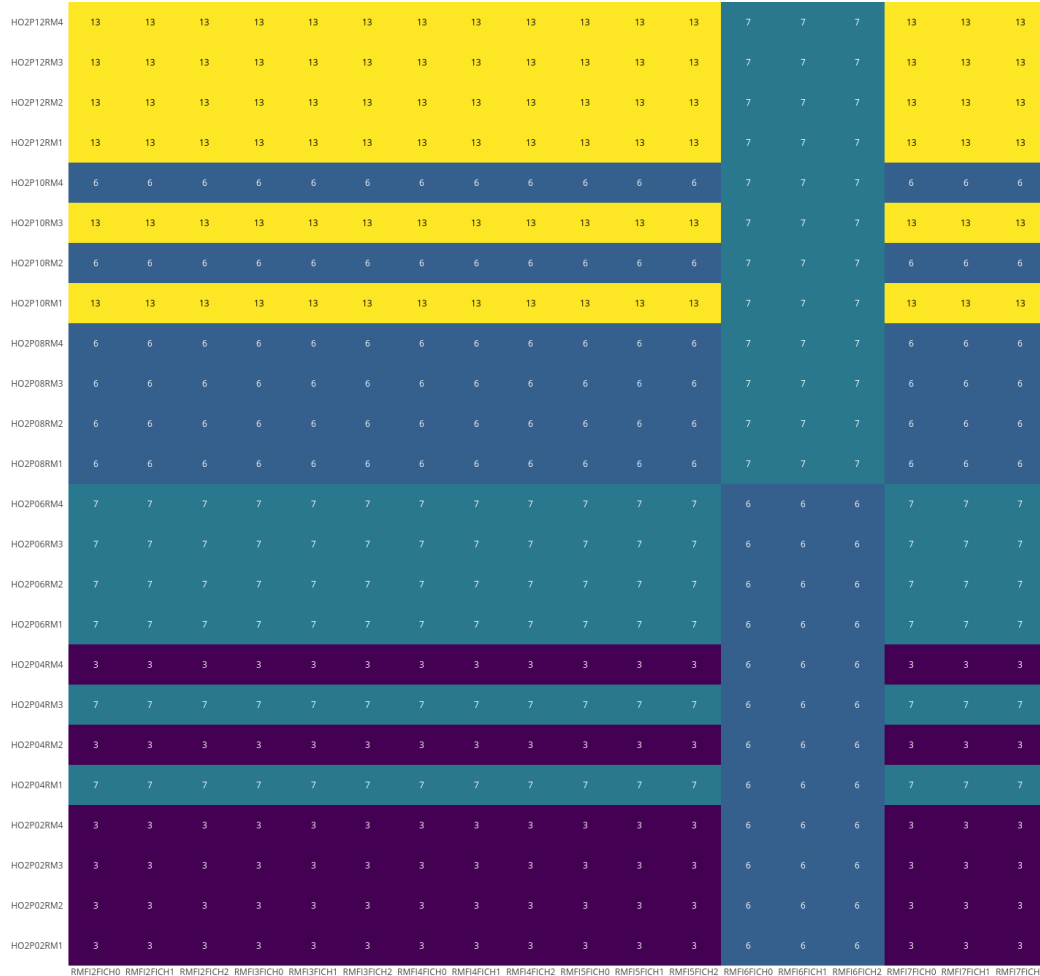


Figure B.71: HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.

HO2P HTR in FrontEnd coordiantes

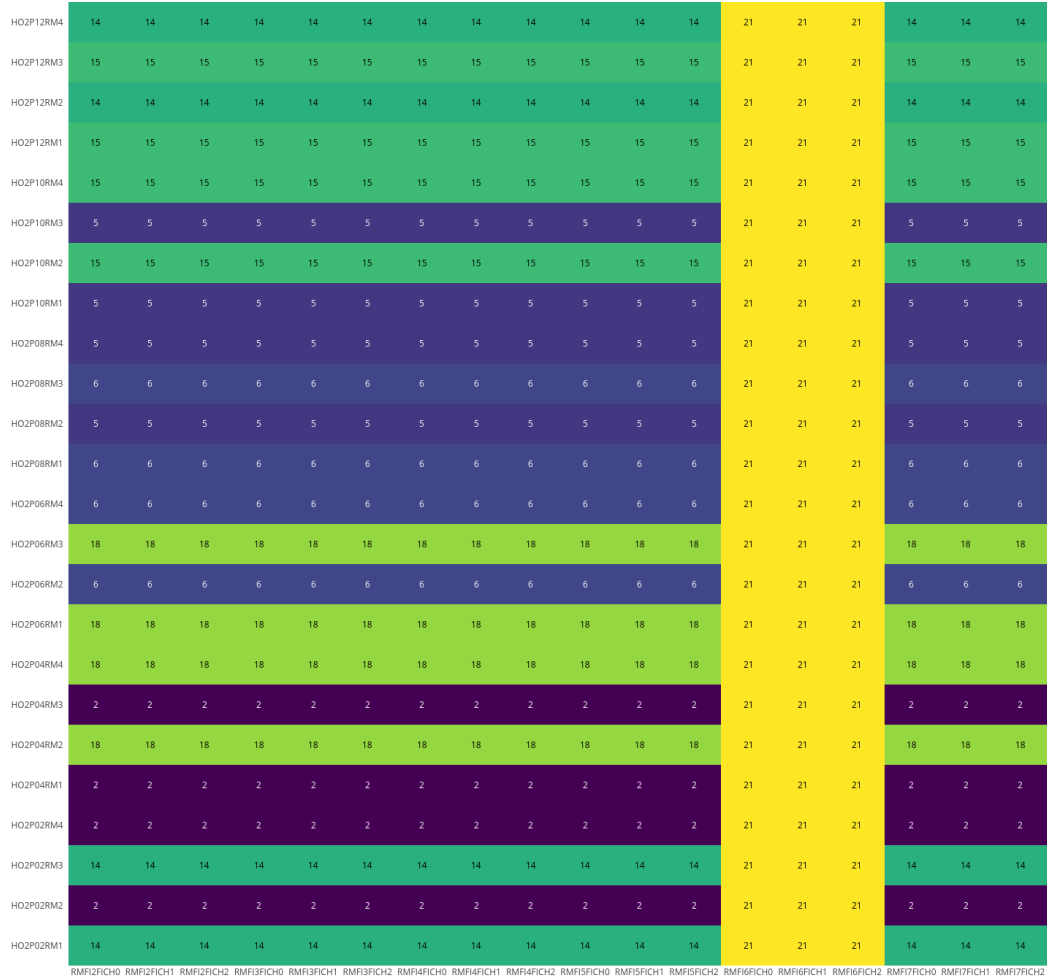


Figure B.72: HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate HTR slot distribution in the frontend electronic coordinates.

HO2P HTR_TB in FrontEnd coordiantes

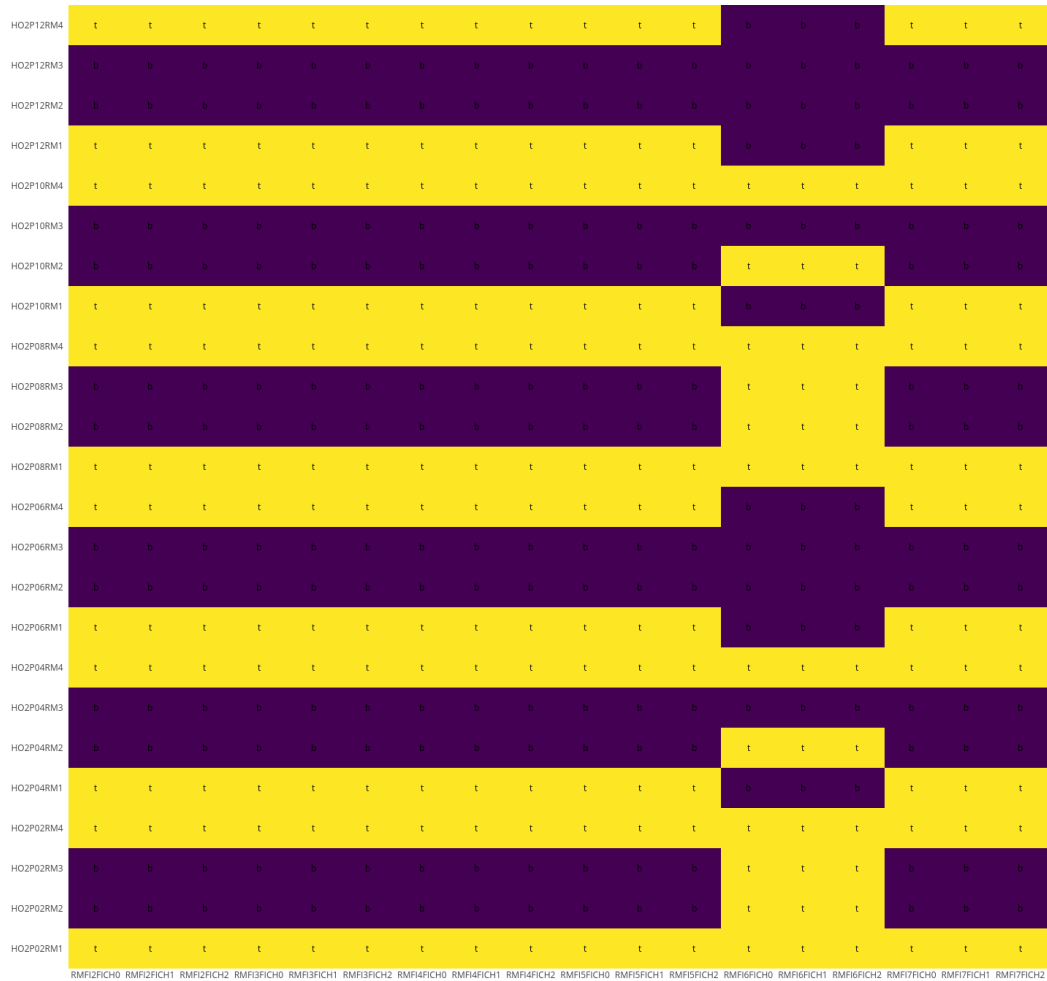


Figure B.73: HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.

HO2P HTR_FI in FrontEnd coordiantes

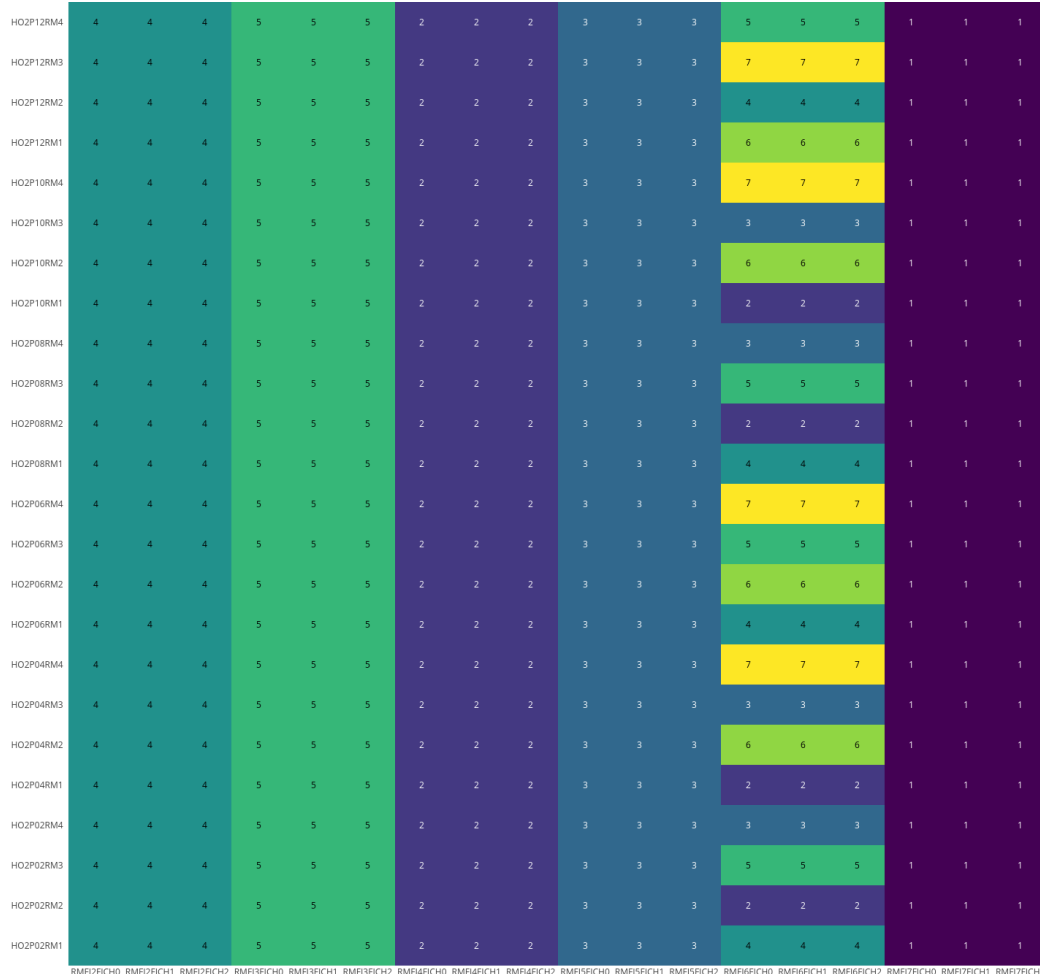


Figure B.74: HCAL (phase 1 HO, sector 2, plus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.

HO2M Crate in FrontEnd coordiantes

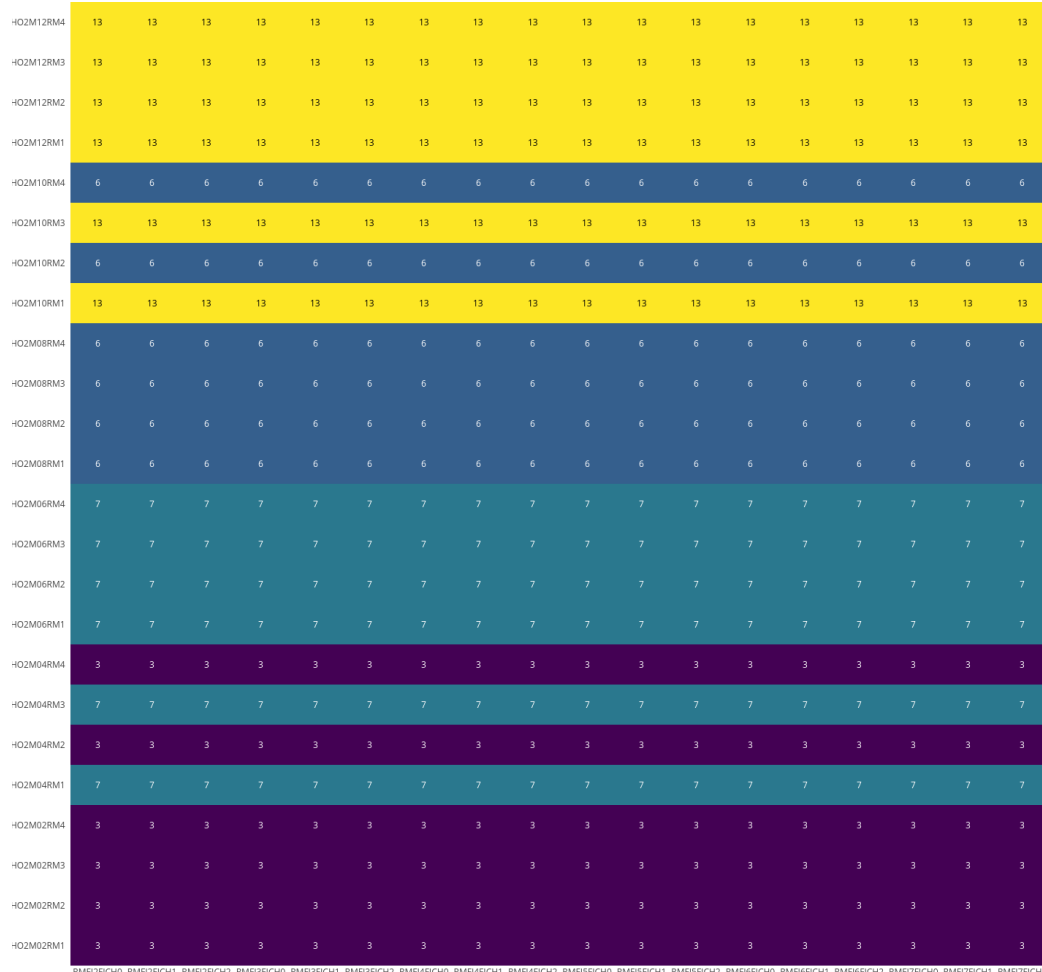


Figure B.75: HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate crate distribution in the frontend electronic coordinates.

HO2M HTR_TB in FrontEnd coordiantes

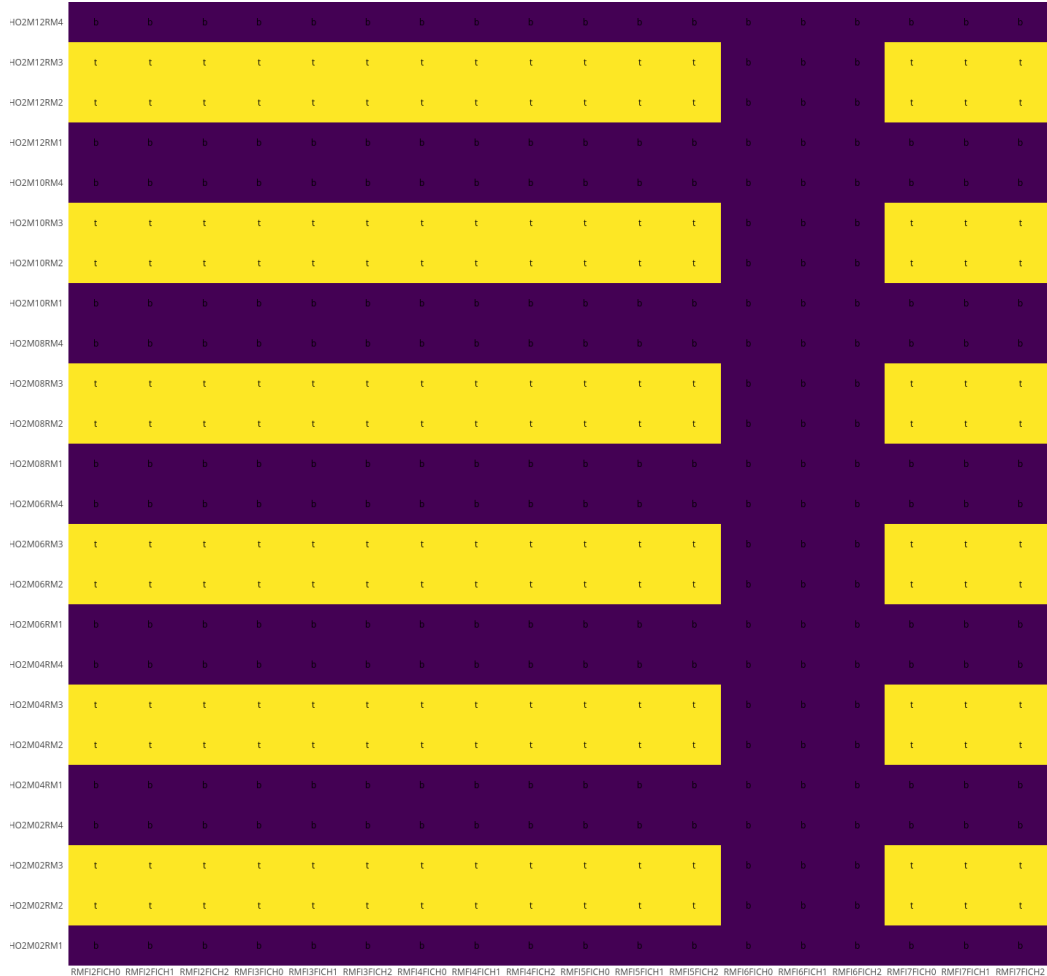


Figure B.77: HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate HTR fpga distribution in the frontend electronic coordinates.

HO2M HTR_FI in FrontEnd coordiantes

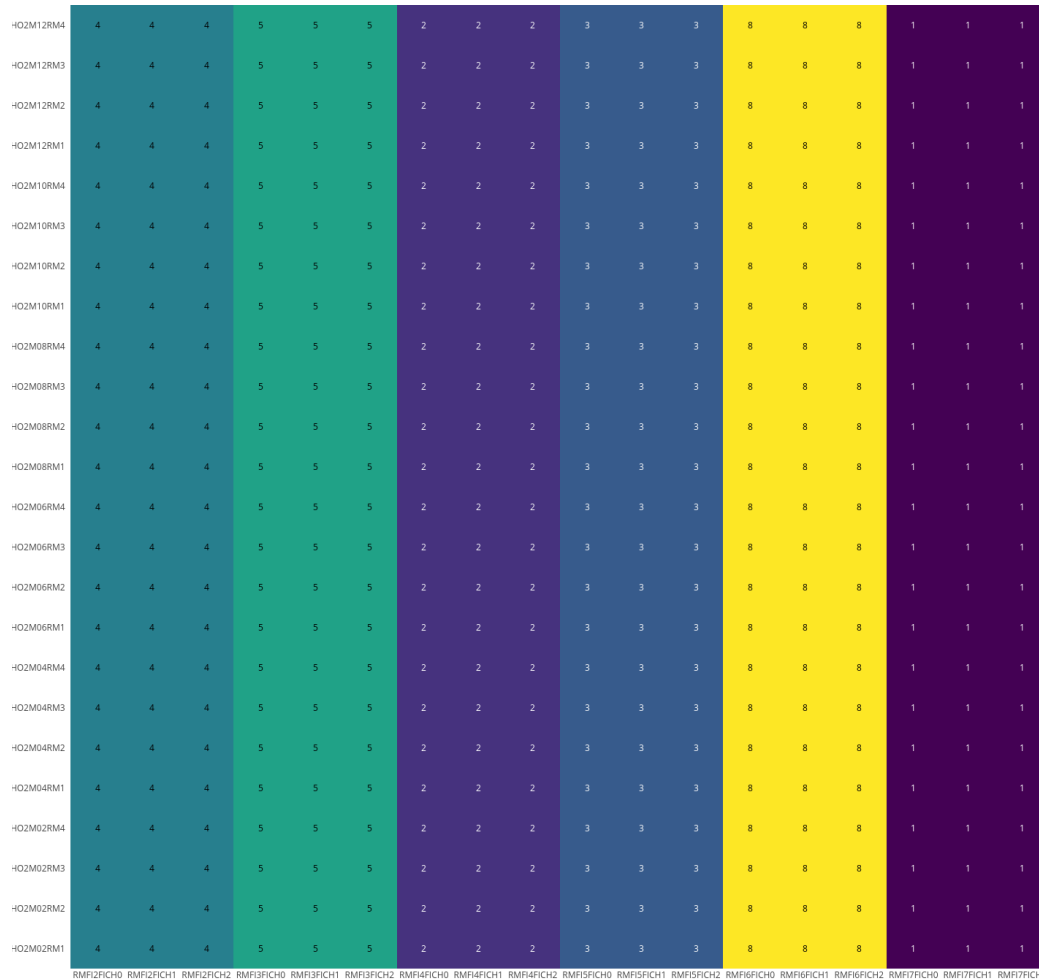


Figure B.78: HCAL (phase 1 HO, sector 2, minus side) backend electronic coordinate HTR fiber distribution in the frontend electronic coordinates.