ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

Ph.D. THESIS

Semra TÜRKÇAPAR

SEARCH FOR STEALTH SY$\bar{Y}$ TOP SQUARK DECAYS WITH AN AUTOMATED ABCD METHOD USING A DOUBLE DISCO NEURAL NETWORK IN PROTON-PROTON COLLISIONS AT $\sqrt{s}$ = 13 TeV WITH THE CMS EXPERIMENT

DEPARTMENT OF PHYSICS

ADANA, 2023

ÇUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

SEARCH FOR STEALTH SY$\bar{Y}$ TOP SQUARK DECAYS WITH AN AUTOMATED ABCD METHOD USING A DOUBLE DISCO NEURAL NETWORK IN PROTON-PROTON COLLISIONS AT $\sqrt{s}$ = 13 TeV WITH THE CMS EXPERIMENT

Semra TÜRKÇAPAR

Ph.D. THESIS

DEPARTMENT OF PHYSICS

We certify that the thesis titled above was reviewed and approved for the award of degree of the Doctor of Physics by the board of jury on 17/03/2023.

....	....	....
Prof. Dr. Ayşe POLATÖZ SUPERVISOR	Prof. Dr. Eda EŞKUT MEMBER	Prof. Dr. İsa DUMANOĞLU MEMBER
....	....	
Prof. Dr. Kenan SÖĞÜT MEMBER	Asst. Prof. Dr. Yalçın GÜLER MEMBER	

This Ph.D.Thesis is written at the Department of Physics of Institute of Natural and Applied Sciences of Çukurova University.
Registration Number:

Prof. Dr. Sadık DİNÇER
Director
Institute of Natural and Applied Sciences

This study was supported by Çukurova University Scientific Research Fund.
Project No: FDK-2017-9644

Note: The usage of the presented specific declarations, tables, figures and photographs either in this thesis or in any other reference without citation is subject to "The Law of Arts and Intellectual Products" numbered 5846 of Turkish Republic.

This thesis is dedicated to the ”women” who are left alone in their struggles—does not matter which kind of struggles! They always make me feel stronger.

ABSTRACT

Ph.D THESIS

SEARCH FOR STEALTH SY$\bar{Y}$ TOP SQUARK DECAYS WITH AN AUTOMATED ABCD METHOD USING A DOUBLE DISCO NEURAL NETWORK IN PROTON-PROTON COLLISIONS AT $\sqrt{s}$ = 13 TeV WITH THE CMS EXPERIMENT

Semra TÜRKÇAPAR

DEPARTMENT OF PHYSICS
INSTITUTE OF NATURAL AND APPLIED SCIENCES
UNIVERSITY OF ÇUKUROVA

Supervisor : Prof. Dr. Ayşe POLATÖZ
Year : 2023, Pages : 286
Jury : Prof. Dr. Ayşe POLATÖZ
Prof. Dr. Eda EŞKUT
Prof. Dr. İsa DUMANOĞLU
Prof. Dr. Kenan SÖĞÜT
Asst. Prof. Dr. Yalçın GÜLER

In this thesis, a search is presented for new physics beyond the standard model which includes a version of Supersymmetry (SUSY) characterized by Stealth SY$\bar{Y}$ SUSY. This search considers events with two top quarks, no extra transverse momentum, and many light flavor jets as pair-produced top squark decays. The Run2 data used were collected with the CMS detector at the LHC from 2016 to 2018, and correspond to a total integrated luminosity of 138 fb^{-1}. An automated ABCD method using a Double DisCo Neural Network is implemented to estimate the main irreducible background ($t\bar{t}+$jets) in data. The analysis considers the three orthogonal decay channels: fully-hadronic, semi-leptonic and fully leptonic and the results of these channels are combined for best signal sensitivity.

Key Words: CMS, Physics, Supersymmetry, NN

ÖZ

DOKTORA TEZİ

CMS DENEYİNDE $\sqrt{s}$ = 13 TeV'LİK PROTON-PROTON ÇARPIŞMALARINDA STEALTH SY$\bar{Y}$ TOP SKUARK BOZUNUMLARININ DOUBLE DISCO NN KULLANAN OTOMATİKLEŞTİRİLMİŞ ABCD METODU İLE ARAŞTIRILMASI

Semra TÜRKÇAPAR

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
FiZiK ANABİLİM DALI

Danışman : Prof. Dr. Ayşe POLATÖZ
Yıl : 2023, Sayfa : 286
Jüri : Prof. Dr. Ayşe POLATÖZ
Prof. Dr. Eda EŞKUT
Prof. Dr. İsa DUMANOĞLU
Prof. Dr. Kenan SÖĞÜT
Dr. Öğr. Üyesi Yalçın GÜLER

Bu tezde, standart modelin ötesinde "Stealth SY$\bar{Y}$" SUSY ile karakterize edilen yeni fizik araştırması sunulmuştur. Bu araştırmada iki top kuark içeren, ekstra dik momentumun olmadığı ve birçok hafif çeşnili jet içeren, çiftler halinde üretilen süpersimetrik top kuarkların bozunumları ele alınmıştır. Kullanılan Run2 verileri, 2016'dan 2018'e kadar LHC'de CMS dedektörü ile toplanmıştır ve 138 fb^{-1} toplam ışıklılığa karşılık gelmektedir. Verilerdeki ana indirgenemez arka planı ($t\bar{t}+$jets) tahmin etmek için, Double DisCo NN kullanarak otomatikleştirilmiş bir ABCD methodu uygulanmıştır. Analiz üç ortogonal bozunma kanalını göz önünde bulundurmaktadır: tamamen-hadronik, yarı-leptonik ve tamamen-leptonik ve bu kanalların sonuçları en iyi sinyal hassasiyeti için birleştirilmiştir.

Anahtar Kelimeler: CMS, Fizik, Süpersimetri, NN

EXTENDED SUMMARY

With the discovery of the Higgs boson by the Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS) experiments at the Large Hadron Collider (LHC), the Standard Model (SM) was considered complete and it describes all observed particles and their interactions except the gravitational interactions. There is a wide range of experimental evidence from the LHC demonstrating that the SM has been the most successful theory to date. Even if it is a well-tested theory supported by experimental results, there are still unanswered questions that cannot be described by the mathematical framework of the SM.

One of the big remaining questions is why the SM cannot include gravity. Of course, there are also still no answers from the SM for what is the dark matter candidate, how neutrinos get their mass, the disparity between visible matter and antimatter, but the Hierarchy problem must be the biggest one. The observed Higgs boson mass is on the electroweak energy scale, since quantum corrections to the Higgs mass are proportional to the highest applicable energy scale. The SM does not provide a mechanism or symmetry to protect the Higgs mass against these very high energy scales and some fine tuning/balancing must take place. In order to answer all these questions or provide a more complete theory, there have been a lot of proposed models for physics Beyond the SM (BSM). Some of those models not only help understand the nature of the Higgs mechanism, but also provide new mechanisms to find a solution for the Hierarchy problem.

Supersymmetry (SUSY) is the most promising model, which can provide a solution to the Hierarchy problem. It introduces a new symmetry that relates fermions to bosons—each fermion in the SM gets a bosonic superpartner, and each boson gets a fermionic superpartner. By introducing such a symmetry, the quantum corrections

to the Higgs mass are automatically cancelled without any fine tuning.

In any SUSY search, supersymmetric particles decay through SM and undetected lightest SUSY particles (LSPs). Most of the traditional SUSY searches at the LHC look for an excess of events with large missing transverse momentum originating from the undetected LSPs. In this case, this large missing transverse momentum carried by the LSP may be a dark matter candidate. In this thesis, unexplored top squark decays ending with ”low missing transverse momentum”, which is called ”Stealth” SUSY, are presented. The Stealth SUSY supposes a ”hidden sector” where SUSY is a conserved symmetry. This hidden sector contains an sfermion and its scaler partner. In this model, the sfermion singlino ($\tilde{S}$) decays through hidden sector to close-in-mass its scaler partner singlet (S) and a gravitino ($\tilde{G}$), which is the LSP, and whose mass is usually assumed to be around 1 GeV. In this case, the top squark could decay through this hidden sector and $\tilde{G}$ would carry low missing transverse momentum.

The thesis analysis uses Run2 data collected during the years 2016, 2017 and 2018 in the proton-proton collisions by the CMS detector. The CMS detector is one of the general purpose detectors at the LHC, which is also suitable for finding evidence for SUSY. It allows good muon identification, high momentum and energy resolution. The CMS also provides excellent jet identification and jet resolution via the calorimeter systems.

In this thesis, a dedicated analysis searching for Stealth top squark decays with an automated ABCD method using Double DisCo Neural Network is presented. The automated ABCD method is applied to estimate the $t\bar{t}+$jets in data. Double DisCo is a novel neural network (NN), consisting of two binary classifiers and it is customized for this analysis with a mass regression model and some additional loss terms in the total loss function. Additionally, another NN technique is used for an object tagging. Also, all fully-hadronic, semi-leptonic and fully-leptonic final states of top squark decays are included for better signal sensitivity.

GENİŞLETİLMİŞ ÖZET

Higgs Bozununun Büyük Hadron Çarpıştırıcısı'nda (BHÇ) CMS ve ATLAS deneyleri tarafından keşfedilmesiyle, Standart Model (SM) tamamlanmış kabul edildi ve böylece SM'in parçacıkları ve yerçekimi kuvveti dışında diğer tüm etkileşimleri gözlemlendi. BHÇ'de SM'nin bugüne kadarki en başarılı teori olduğunu gösteren çok çeşitli deneysel kanıtlar vardır. Halen SM deneysel sonuçlarla desteklenen iyi test edilmiş bir teori olmasına rağmen, matematiksel çerçevesiyle cevaplayamadığı sorular vardır.

En büyük sorulardan biri, SM'in neden yerçekimi kuvvetini içermediğidir. Kuşkusuz ki, SM'in karanlık madde adayı, nötrino kütlesi, görünür madde ve anti-madde arasındaki fark gibi açıklayamadığı soruları da var, fakat en önemli sorun Hiyerarşi problemidir. Gözlenen Higgs bozon kütlesi elektrozayıf enerji ölçeğindedir, ama Higgs kütlesine yapılan kuantum düzeltmeleri uygulanabilir en yüksek enerji ölçeğiyle orantılıdır. SM, Higgs kütlesini bu çok yüksek enerji ölçeklerine karşı korumak için bir mekanizma veya simetri sağlamaz ve bazı ince ayar ya da dengelemelerin yapılması gerekir. Tüm bu soruları yanıtlamak veya daha eksiksiz bir teori sağlamak için SM'nin ötesinde önerilen birçok yeni fizik modeli vardır. Bu modellerden bazıları yalnızca Higgs mekanizmasının doğasını anlamaya yardımcı olmakla kalmaz, aynı zamanda Hiyerarşi sorununa da çözüm bulmak için yeni mekanizmalar sağlar.

Süpersimetri (SUSY), Hiyerarşi sorununa çözüm sağlayabilecek en umut verici modeldir. SUSY fermiyonları bozonlarla ilişkilendiren yeni bir simetri sunar—SM'deki her fermiyon bir skaler süpereş ve her bozon da bir fermiyonik süpereş alır. Böyle bir simetri getirilerek, Higgs kütlesine yapılan kuantum düzeltmeleri, herhangi bir ince ayar gerekmeden otomatik olarak iptal edilir.

SUSY araştırmasında, süpersimetrik parçacıklar SM ve tespit edilmemiş en hafif SUSY parçacıklarına (LSP) bozunur. BHÇ'deki geleneksel SUSY araştırmalarının çoğu LSP'den gelen geniş kayıp dik momentum içeren olaylara kanıt arar. Bu durumda, LSP tarafından taşınan geniş kayıp dik momentum karanlık madde adayı olabilir. Bu tezde, "Stealth" SUSY olarak adlandırılan ve düşük kayıp dik momentum ile sonlanan keşfedilmemiş süpersimetrik top kuarkların bozunumları sunulmuştur. Stealth SUSY, süpersimetrinin korunduğu bir "gizli sektör" olduğunu varsayar. Bu gizli sektör, bir süpersimetrik fermiyon ve onun skaler eşini içerir. Bu modelde, süpersimetrik fermiyon "singlino" gizli sektör aracılığıyla skaler eşi "singlet" ve bir hafif SUSY parçacığı olan ve kütlesinin genellikle 1 GeV civarında olduğu varsayılan "gravitino" ya bozunur. Bu durumda, süpersimetrik top kuark bu gizli sektör aracılığıyla bozunabilir ve gravitino düşük kayıp dik momentum taşır.

Bu tezde sunulan analizde, CMS dedektörü tarafından proton-proton çarpışmalarında 2016, 2017 ve 2018 yıllarında toplanan Run2 verileri kullanılmıştır. CMS dedektörü, SUSY için kanıt bulmak için de uygun olan BHÇ'deki genel amaçlı dedektörlerden biri olup iyi bir müon tanımlaması, yüksek momentum ve enerji çözünürlüğü sağlar. CMS dedektörü ayrıca kalorimetre sistemleri aracılığıyla mükemmel jet tanımlama ve jet çözünürlüğü sağlar.

Bu tezde, Double DisCo NN kullanarak otomatikleştirilmiş bir ABCD yöntemiyle "Stealth" süpersimetrik top kuark bozunumlarını arayan özel bir analiz sunulmaktadır. Bu otamatikleştirilmiş ABCD metodu sinyal bölgesindeki $t\bar{t}+$jets ardalanlarını tespit etmek için kullanılmaktadır. Double DisCo iki tane diskriminatörden oluşan orijinal bir NN metodudur ve "mass regression" modeli ve toplam "loss" fonksiyonundaki bazı ek terimlerle bu analiz için özelleştirilmiştir. Buna ek olarak, nesne etiketleme için başka bir NN tekniği de kullanılmaktadır. Ayrıca, daha iyi sinyal duyarlılığı için süpersimetrik top kuark bozunumlarının tam hadronik, yarı leptonik ve tam leptonik son durumlarının hepsi dahil edilmiştir.

ACKNOWLEDGEMENTS

”The true sign of intelligence is not knowledge but imagination!” – Albert Einstein. I have started my long way with Albert Einstein! As always inspired by him, I have never given up to imagine for being wherever I am now! It has been a long way to getting *”my Ph.D.”*, which I had been working hard and struggling a lot to make it work—sometimes it looks like a miracle...sometimes it seems like a loss...but I always feel lucky because of many notable people passing through my life during my Ph.D. term! With inspiration and support from all those, this Ph.D. degree is mine to the end!

I want to express my endless thanks to Prof. Dr. Gülsen Önengüt, who encouraged me to step into a door instead of rejecting me even if there were a lot of students on her team. I had another valuable encouragement from my advisor Prof. Dr. Ayşe Polatöz, who believed in me and let me join this team—otherwise, I could not be wherever I am! I appreciate that my professor never rejected me whenever I wanted to work on new things during my entire PhD—this means a lot! Also, she was very supportive and kept me motivated. I want to dedicate my special thanks to her with this *”Ph.D. degree.”*

I would like to express my special thanks to my thesis committee, Prof. Dr. Eda Eşkut, Prof. Dr. İsa Dumanoğlu, Prof. Dr. Kenan Söğüt and Dr. Öğr. Üyesi Yalçın Güler for reading this thesis and giving their significant comments. Also, I thank Prof. Dr. Aysel Kayış Topaksu for her direct and indirect contributions to being successful on my way.

I would like to thank Assoc. Prof. Dr. Alberto Belloni, who made me realize that I am really interested in anything related to electronics. Moreover, I should be lucky to meet Prof. Dr. Drew Baden, who taught me that I have to keep asking my questions without thinking if my questions are silly, if people judge me because of my physics knowledge. Also, I always remember that Prof. Dr. Jay Dittmann told me that it does not matter coming from the same educational background as others, it only matters if I am curious enough, working hard, and being responsible. I think my most favorite is Research Scientist Dr. Richard Kellogg, who must be the most expert one and, even at his older age, his curiosity never goes away—this still keeps me motivated! Besides, every single day he had been asking “would you like to drink coffee”—lots of discussion all together by drinking morning coffees in his office. Recently, it is a crossed corner in my life, that a brand new door was opened by Scientist Dr. Jim Hirschauer. I was invited to Fermilab as a visitor PhD student by him—he provided me to start a new path in my career. Besides, I can not discard the mesmerizing effect from a *”woman”*, Asst. Prof. Dr. Nadja Strobbe, who always inspires me as a young, strong woman in science! Also, Senior Scientist Dr. Pawel De Barbaro and Prof. Dr. Paolo Rumerio gave me good opportunities as project manager. I am very grateful that all these amazing people touched my life in many respects and supported my struggle on my way to getting this Ph.D. degree!

I also thank my *”several crews”* for learning together, handling the electronics and physics analyses together, laughing together and many other actions together during both my CERN and Fermilab times. Starting with my *”HCAL crew”*, many thanks to Danny Noonan, Fábio Alves, Josh Kunkle, Ayşe Bat, Edward Laird, Caleb Smith—always able to discuss any detector and analysis topics, have lots of fun together. My *”HB CU crew”* Doyeong Kim, Brooks McMaster, Zuhal Şeyma Demiroğlu, special thanks to being a great team all together and always open to learn together—we grew each other a lot in many respects. Last, my *”SUSY crew”*, I am very grateful to meet all them, who are Kelvin Mei, Chris Madrid, Alex Mercaldi, Bryan Crossman, Charlie Kapsiak, Devin Mahon. They helped me a lot to do this thesis analysis and especially some of them were more close to me—more than just being a work friend!

Besides all these smart and fun crews, I would like to express my very special thanks to *”Windfall crew”* who open new windows in both my personal and professional lives. Many thanks to Tote Hughes, who showed me how I can have fun when I work. Also, he always encouraged me when I spoke in English—made me keep speaking even if there were lots of mistakes. I think one of the most important ones to touch my life magically, Joseph Mariano, who turned a mirror towards me for seeing myself in many aspects—for having my own path and career as a *”woman”*, he showed how strong I am, how enthusiastic I am, how skillful I am, how much potential I have. I would like to dedicate this *thesis–Ph.D. degree* to show him how much I appreciate his endless support! I would like to express special thanks to another, most important one, as my favourite working friend ever, Joshua Hiltbrand, who has changed my life view a lot and shown me a better way to look at life for being more happy, more calm and less stressful. Also, many tasty moments with him while deeply discussing SUSY, making fancy codes, and learning more about the detectors!

I must be lucky to have parents, sister and brother who always encourage me to never give up. In particular, my mom is my role model as a strong woman, who encourages me to fight against difficulties in any situation. I am very grateful to have her in my life. Additionally, I am glad that my sister also supported me as another strong woman to choose this path in my life—it means a lot. Of course, I cannot discard that I also have ”man” power from my dad and brother in different aspects!

Lastly, I also would like to thank TAEK (Turkish Atomic Energy Agency) for financially supporting my work at CERN. Also special thanks to the U.S. CMS group for funding me a couple of times during my CERN time. Last, thanks to the LHC Physics Center (LPC) Guest & Visitor Program for funding me during my Fermilab time.

CONTENTS **PAGE**

ABSTRACT I
ÖZ II
EXTENDED SUMMARY II
GENİŞLETİLMİŞ ÖZET VII
ACKNOWLEDGEMENTS XI
CONTENTS XIV
LIST OF TABLES XVI
LIST OF FIGURES XVIII
ABBREVIATIONS XX
1. INTRODUCTION 1
2. THEORETICAL FRAMEWORK 5
2.1. Standard Model (SM) 5
2.1.1. Quantum Electrodynamics (QED) 9
2.1.2. Quantum Chromodynamics (QCD) 10
2.1.3. Electroweak Theory 12
2.1.4. Brout-Englert-Higgs Mechanism 13
2.2. Shortcomings of the Standard Model 16
2.3. Supersymmetry (SUSY) 20
2.3.1. R-Parity Violating (RPV) SUSY 24
2.3.2. Stealth SUSY 25
3. THE LHC AND CMS EXPERIMENT 27
3.1. CERN Accelerator Complex 28
3.2. Proton-Proton Collisions at the LHC 29
3.3. Large Hadron Collider (LHC) 31

3.4. CMS Detector 32
3.4.1. Tracker System 35
3.4.2. Electromagnetic Calorimeter 37
3.4.3. Hadronic Calorimeter 40
3.4.4. Magnet 48
3.4.5. Muon System 50
3.4.6. BRIL System 52
3.5. CMS Trigger System 53
3.5.1. L1 Trigger 53
3.5.2. HLT 55
3.6. CMS Event Recontruction 55
4. SEARCH STRATEGY FOR RPV AND STEALTH SUSY 59
4.1. General Search Strategy for RPV and Stealth SUSY 59
4.2. Motivation from Legacy Analysis 62
4.3. Search Strategy for Current Analysis 66
5. EVENT RECONSTRUCTION AND SELECTION 71
5.1. Data Sets and Simulation Samples 71
5.1.1. Data 71
5.1.2. Simulation 72
5.2. Reconstructed Physics Objects 74
5.2.1. Electrons 74
5.2.2. Muons 75
5.2.3. Jets 76
5.2.4. b-tagged Jets 77
5.2.5. top-tagged Jets 78
5.3. Scale Factors 79

5.3.1. Trigger Scale Factors 79
5.3.2. top-tagging Scale Factors 91
5.4. Event Selections 97
5.4.1. Signal Region (SR) Selections 97
5.4.2. QCD Control Region (QCD CR) Selections 106
6. HADRONIC TOP TAGGING 109
6.1. Hadronic Top Tagger 110
6.1.1. Tagging Algorithms 111
6.1.2. Working Point (WP) and ROC Curve for Performance 113
6.2. Working Point Optimization for DeepResolved Tagger 115
7. TOP SQUARK RECONSTRUCTION 121
8. $t\bar{t}+$jets ESTIMATION AND SYSTEMATICS 129
8.1. ABCD Method 129
8.2. Double DisCo Neural Network Model 131
8.2.1. Mass Regression Model 132
8.2.2. Binary Classifier Models 134
8.2.3. Total Loss Function 134
8.2.4. Hyperparameters 138
8.3. Double DisCo Trainings 139
8.3.1. Input Variables 139
8.3.2. Training Samples 140
8.3.3. Training Performance 142
8.4. Validating the ABCD Method 166
8.4.1. Defining the ABCD Regions 166
8.4.2. Defining the Validation Regions 182
8.4.3. Verifying the Validation Regions 186

8.5. $t\bar{t}+$jets Closure Correction 193
8.6. $t\bar{t}+$jets Systematics 197
8.6.1. Simulation-Based Systematics 197
8.6.2. Data-Based Systematics 200
9. OTHER ESTIMATIONS AND SYSTEMATICS 211
9.1. QCD multi-jet Estimation and Systematics 211
9.1.1. Estimation Method 215
9.1.2. Systematics 216
9.2. Minor Background Estimations and Systematics 217
10. FIT PROCEDURE AND VALIDATION 219
10.1. Fit Procedure 220
10.2. Fit Validation 226
11. PRELIMINARY RESULTS AND INTERPRETATION 235
12. CONCLUSION 239
REFERENCES 241
CURRICULUM VITAE 241
Appendix A Hadronic Top Tagging 244
A1. Number of Tops Populations 244
Appendix B Event Reconstruction and Selection 246
B1. Simulated Background and Signal Samples 246
Appendix C Double DisCo Neural Network 248
C1. Choosing the Input Variables 248
Appendix D Validating the ABCD Method 251
D1. Defining the ABCD Regions 251
D1.1. Initial Study for Optimizing the ABCD Bin Edges 251

D1.2. Optimizing the ABCD Bin Edges 257
D2. Verifying the Validation Regions 269
D3. $t\bar{t}+$jets Data-Based Systematics 272
Appendix E Fit Procedure and Validation 275

LIST OF TABLES **PAGE**

Table 2.1. The bosons of the SM subdivided by their quantum numbers: electric charge Q, third component of weak isospin T_3, weak hypercharge Y, and color charge C. * There are eight linear color-anticolor combinations corresponding to the eight different gluons 8

Table 3.1. Properties of the CMS solenoid magnet. 48

Table 5.1. Run2UL data sets used in this analysis. Note data 2018 SingleElectron data set is renamed as EGamma, as shown in (CMS Collaboration, 2022c) 72

Table 5.2. The minor background processes used in this analysis are listed. These processes are grouped into "$t\bar{t}+\mathrm{X}$ " and "Other" by the analysis. Each individual minor background listed is about less than 1% of the total background 74

Table 5.3. Good electron selection criteria for the signal region 75

Table 5.4. Good muon selection criteria for the signal regions. Nonisolated muon selection criteria for the QCD control region. For the control region, only SingleMu50 type triggers are used, and thus the offline p_{T} selection is increased to 55 GeV 76

Table 5.5. Working points (WPs) for tagging b jets 78

Table 5.6. WPs for each algorithm corresponding to each year. The top WP given for DeepResolved tagger is the optimized one. DeepResolved tagger also requires one of the decay products of top quark as b-jet by implementing "medium" DeepFalvor WP 79
Table 5.7. Each jet, electron and muon trigger used by year, which are OR'ed together for a given year. Both CSV and DeepCSV triggers are used to cover the full run range for a given year. 2016 jet/electron/muon triggers are applied for both 2016preVFP and 2016postVFP years 81
Table 5.8. The preselections for measuring the jet trigger efficiencies and jet trigger scale factors 82
Table 5.9. The preselections for measuring the electron trigger efficiencies and electron trigger scale factors 85
Table 5.10. The preselections for measuring the muon trigger efficiencies and muon trigger scale factors 88
Table 5.11. The selections for defining the semi-leptonic $t\bar{t}+$jets control region (left) and QCD multi-jet control region (right). Definitions of objects and triggers follow from Section 5.2. and Section 5.3.1., respectively. Requirements involving a "loose b jet" refer to some other b-tagged jet (the required medium b jet trivially passes the loose WP) in the event that passes the loose Deep Flavour WP 92
Table 5.12. The list of the SR selections for the fully-hadronic channel 98
Table 5.13. The list of the SR selections for the semi-leptonic channel 101
Table 5.14. The list of the SR selections for the fully-leptonic channel 104

Table 5.15. The two QCD CR selections used for the fully-hadronic and semi-leptonic QCD multi-jet background estimation 107

Table 6.1. WPs for each algorithm corresponding to each year. The top WP for DeepResolved tagger is threshold value based on the original version of the tagger. DeepResolved tagger also requires one of the decay products of top quark as b-jet by implementing "medium" DeepCSV WP 113

Table 6.2. The selection criterias to measure efficiency, fake rate, TPR and FPR 116

Table 8.1. The list of hyper parameters for each of the fully-hadronic, semi-leptonic and fully-leptonic channel. Note that the given LR for the whole network for a given channel 139

Table 8.2. The input variables of the network for each of the fully-hadronic, semi-leptonic and fully-leptonic channel. The four-vector of top squark with "Old Seeding Method" is used for trainings of the semi-leptonic and fully-leptonic channels. The four-vector of top squark with "Top Seeding Method" is used for training of the fully-hadronic channel 141

Table 8.3. For the fully-hadronic channel, top four choices for the ABCD bin edges, ranked by significance. Also, the choice of non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance. For each ABCD bin edges choice, the values of the metric elements are also listed for each N_{jets} bin. Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV is applied while optimizing the edges 174

Table 8.4. For the semi-leptonic channel, top five choices for the ABCD bin edges, ranked by significance. Also, the choice of non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance. For each ABCD bin edges choice, the values of the metric elements are also listed for each N_{jets} bin. Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV is applied while optimizing the edges 175
Table 8.5. For the fully-leptonic channel, top five choices for the ABCD bin edges, ranked by significance. Also, the choice of non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance. For each ABCD bin edges choice, the values of the metric elements are also listed for each N_{jets} bin. Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV is applied while optimizing the edges 176
Table 8.6. For the fully-hadronic channel, the values of the maximum corrected data closures and data-based systematics for all N_{jets} . 205
Table 8.7. For the semi-leptonic channel, the values of the maximum corrected data closures and data-based systematics for all N_{jets} 205
Table 8.8. For the fully-leptonic channel, the values of the maximum corrected data closures and data-based systematics for all N_{jets} 205

Table 9.1. For the fully-hadronic and semi-leptonic channels, TF values used for each of **A**, **B**, **C**, **D** NN bins for the QCD multi-jet background estimation. Note that **A**, **B**, **C**, **D** bins are separated by the NN trainings, targeting the Stealth SY$\bar{Y}$ signal 216

Table 10.1. The systematic uncertainties in the final fit for all backgrounds processes and Stealth SY$\bar{Y}$ signal model. The first group for all background and signal processes excepting that lumi is not for $t\bar{t}+$jets and QCD multi-jet. The second groups for minor backgrounds and signal processes, the third group for only minor background processes, the forth group for signal and all background processes. The last the two groups for $t\bar{t}+$jets and QCD multi-jet background processes, respectively. Abbreviated names for each source of uncertainty are explained in the text, denoting the related N_{jets} and ABCD bins 222

Table B1. Simulated signal samples and related cross section for each sample are listed for this analysis. Cross sections are the same for all 2016preVFP, 2016postVFP, 2017 and 2018 246

Table B2. Simulated background samples and the production cross section for each sample are listed for this analysis. Due to identical physics simulation settings for each year, the cross section for a given sample is the same for 2016preVFP, 2016postVFP, 2017 and 2018 247

XXV

LIST OF FIGURES **PAGE**

Figure 1.1. Top squark decay mode considered in this analysis. The Stealth SY$\bar{Y}$ model. Note that a vertex in the diagram is effective couplings, like in the singlet decay 2

Figure 2.1. The elementary particles of the SM. The model includes three generations of matter particles (leptons and quarks) as well as the gauge and Higgs bosons. There are given properties of the particles, including the particle name, symbol, mass, spin, electric charge, and color charge 6

Figure 2.2. The allowed particle interactions via the QCD interaction arising from the requirement of $SU(3)$ local gauge invariance. The form of qqg interaction vertex (left), triple gluon interaction vertex (middle), quartic gluon interaction vertex (right) (Thomson, M., 2013) 11

Figure 2.3. The one-loop quantum corrections to the Higgs mass, arising from the coupling of the Higgs boson to itself, the coupling of the Higgs boson to the fermions, the coupling of the Higgs boson to the vector bosons (left to right) (Grojean, 2015) 20

Figure 2.4. The SM particles and their superpartners within a MSSM. Each fermion is transformed into bosonic superpartner and vice versa. Superpartners differ in spin by half unit, but the remaining quantum numbers are the same. If SUSY is not broken, then each SM particle and its superpartner have also the same mass 22

Figure 2.5. At the LHC, SUSY cross sections as a function of particle mass. The red curve corresponds to the top squark pair production at 13 TeV (CMS Collaboration, 2021) 24

Figure 2.6. The “LLE-coupling”, “LQD-coupling” and “UDD-coupling” for R-parity violating interactions in the MSSM. λ is coupling strength for LLE, λ' is coupling strength for LQD, λ'' is coupling strength for UDD (Barbier et al., 2005) 25

Figure 2.7. The scalar particle Singlet (S) decays via $Y - \bar{Y}$ vector-like messenger fields in the “stealth sector” to the SM particles (Fan, Reece, and Ruderman, 2012) 26

Figure 3.1. The whole CERN accelerator complex, including the LHC. The four main experiments, ALICE, ATLAS, CMS, LHCb, are also shown (Lopienska, 2022) 29

Figure 3.2. Cumulative delivered and recorded luminosity versus time for 2015-2018 for pp data (left). CMS integrated luminosity delivered, pp collison at $\sqrt{s} = 13$ TeV (right) (“Public CMS Luminosity Information” 2022) 31

Figure 3.3. Overview of the CMS detector with the sub-dectetors is indicated on the figure. Figure taken from (Sakuma, 2016) 33

Figure 3.4. The CMS coordinate system, including the LHC and a compass (the z axis points to the Jura). Figure taken from (Neutelings, 2017) 35

Figure 3.5. Schematic cross section through the CMS tracker system. In the center is the pixel detector, divided into the barrel and endcap regions, and right on the outside is the silicon strip detector system. Figure taken from (T. C. Collaboration, 2014) 36

Figure 3.6. A cross sectional view of the pixel detector where top is the new Phase 1 pixel detector and the bottom is the legacy pixel detector. Diagram taken from (The Tracker Group Of The CMS Collaboration, 2020) 37

Figure 3.7. Schematic overview of the ECAL, including the barrel, the end-cap and the preshower (Siddireddy, 2018) 38

Figure 3.8. Image of a lead tungstate ($PbWO_4$) crystal with an electromagnetic shower shown inside. Figure taken from (Wadud, 2019) 40

Figure 3.9. A schematic view of one quarter of the CMS HCAL during 2016 LHC operation, showing the positions of HB, HE, HO and HF calorimeters. The layers marked in blue are grouped together as depth = 1, while the ones in yellow, green, and magenta are combined as depths 2, 3, and 4, respectively (Sirunyan, 2020) 43

Figure 3.10. Photos from Phase-1 upgrades for the CMS HCAL. The HF burn-in test stand (top), including PMTs, FE and BE electronics. One of the HF cavern (bottom left). The bottom right photo from Installation and Commissioning of Front-End Electronics of the HF calorimeter 47

Figure 3.11. Picture of the solenoid with steal return yoke during installation (top). Magnetic field and field lines predicted on a longitudinal section of the CMS detector (bottom), for the underground model at a central magnetic flux density of 3.8 T. Each field line represents a magnetic flux increment of 6 Wb (C. Collaboration, 2010). 49

Figure 3.12. Schematic view of one quarter of the CMS muon system with DTs, CSCs and RPCs (left) (Seo, 2013). The DTs are labeled ”MB” (Muon Barrel) whereas CSCs are labeled ”ME” (Muon Endcap). Muon chambers interspersed in the magnet return yoke (right) 51
Figure 3.13. Position of the various BRIL systems inside CMS (Guthoff, 2016) 52
Figure 3.14. The CMS L1 trigger system (Jeitler, 2013)........ 54
Figure 3.15. A sketch of the signatures of different types of particles in the CMS detector 56

Figure 4.1. The Feynman diagrams for top squark pair production where the top squarks decay via the benchmark RPV (left) and Stealth SY$\bar{Y}$ (right) models. Considering a final state of each model, there is high jet multiplicity (N_{jets}) and minimal additional p_T^{miss} 60
Figure 4.2. The signal region shown divided up by N_{jets} and S_{NN} bins. Low N_{jets} and S_{NN} bin is background enriched and high N_{jets} and S_{NN} bin is signal enriched 64
Figure 4.3. Expected and observed 95% C.L upper limit on the top squark pair production cross section as a function of the top squark mass for the RPV (top) and Stealth $SY\bar{Y}$ (bottom) SUSY models. Particle masses and branching fractions assumed for each model are included on each plot. The expected cross section computed at NNLO+NNLL accuracy is shown in the red curve (Sirunyan, 2021b) 66

Figure 4.4. A toy example for including each of A, B, C, D regions with some event yields. Each of the two dicriminator values of the network corresponds to each of the two independent variables. The number of predicted background events in **A** region can be computed via $N_A = (N_B N_C)/N_D$ 69

Figure 5.1. For the fully-hadronic channel, the jet trigger efficiency measured as a function of jet H_{T} (left) and p_{T} of the 6^{th} jet (right) in $t\bar{t}+$jets simulation and in SingleMuon data for 2016preVFP (top), 2016postVFP (middle top), 2017 (middle bottom), and 2018 (bottom) of full Run2UL 83

Figure 5.2. For the fully-hadronic channel, the jet trigger efficiency scale factor as a function of jet H_{T} and p_{T} of 6^{th} jet for 2016preVFP, 2016postVFP, 2017, 2018 (top to bottom) of full Run2UL and $N_{\mathrm{bjets}} = 1, \geq 2$ cuts (left to right) 84

Figure 5.3. For the semi-leptonic and fully-leptonic channels, the electron trigger efficiency as a function of electron p_{T} (left) and η (right) in $t\bar{t}+$jets simulation and in SingleMuon data for 2016preVFP (top), 2016postVFP (middle top), 2017 (middle bottom), and 2018 (bottom) of full Run2UL 86

Figure 5.4. For the semi-leptonic and fully-leptonic channels, the electron trigger efficiency scale factor as a function of electron p_{T} and η for 2016preVFP (top left), 2016postVFP (top right), 2017 (bottom left), and 2018 (bottom right) of full Run2UL 87

Figure 5.5. For the semi-leptonic and fully-leptonic channels, the muon trigger efficiency as a function of muon p_{T} (left) and η (right) in $t\bar{t}+$jets simulation and in SingleElectron data for 2016preVFP (top), 2016postVFP (middle top), 2017 (middle bottom), and 2018 (bottom) of full Run2UL 89

Figure 5.6. For the semi-leptonic and fully-leptonic channels, the muon trigger efficiency scale factor as a function of muon p_{T} and η for 2016preVFP (top left), 2016postVFP (top right), 2017 (bottom left), and 2018 (bottom right) of full Run2UL 90

Figure 5.7. The top tagger tagging efficiency in data (black) and simulation (red), and accompanying scale factor (blue) for resolved (left) and merged (right) top candidates, as a function of the candidates transverse momentum. Results are for 2016preVFP (top), 2016postVFP (middle-top), 2017 (middle-bottom), and 2018 (bottom) of full RunUL 95

Figure 5.8. The top tagger mistag rate in data (black) and simulation (red), and accompanying scale factor (blue) for resolved (left) and merged (right) top candidates, as a function of the candidates transverse momentum. Results are for 2016preVFP (top), 2016postVFP (middle-top), 2017 (middle-bottom), and 2018 (bottom) of full RunUL 96

Figure 5.9. For the fully-hadronic channel, N_{jets} (top left), N_{bjets} (top right), N_{tops} (middle left), top mass (middle right), $p_{\mathrm{T}}/ H_{\mathrm{T}}$ (bottom left) and ΔR_{bjets} (bottom right) distributions for Run2UL. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield 99

Figure 5.10. For the semi-leptonic channel, N_{jets} (top left), N_{bjets} (top right), $p_{\mathrm{T}}/ H_{\mathrm{T}}$ (bottom left) and $M_{\mathrm{b,l}}$ (bottom right) distributions for Run2UL. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield 102

Figure 5.11. For the fully-leptonic channel, N_{jets} (top left), N_{bjets} (top right), H_{T} (bottom left) and $M_{\mathrm{l,l}}$ (bottom right) distributions for Run2UL. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield 105

Figure 5.12. For the fully-hadronic (left) and semi-leptonic (right) channels, data-simulation comparison for QCD CR selections. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield 107

Figure 6.1. Top quark decays, $t \to bW$. The hadronic decays of top quarks where $W^{\pm} \to q\bar{q}$. The leptonic decays of top quarks where $W^{\pm} \to l\nu$ 110

Figure 6.2. Resolved and merged top decays based on the boosting regimes. Resolved top consists of three individual AK4 jets. Merged top consists of three jets which are clustered into one AK8 jet.......... 111
Figure 6.3. For 2016preVFP, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_{T} from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events 117
Figure 6.4. For 2016postVFP, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_{T} from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events 118
Figure 6.5. For 2017, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_{T} from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events 119

Figure 6.6. For 2018, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_{T} from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events 120

Figure 7.1. For the fully-hadronic channel, four-vector of clustered top squark 1 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The top seeding method is used to cluster top squarks. Top squark 1 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL 123

Figure 7.2. For the fully-hadronic channel, four-vector of clustered top squark 2 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The top seeding method is used to cluster top squarks. Top squark 2 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL 124

Figure 7.3. For the semi-leptonic channel, four-vector of clustered top squark 1 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 1 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL 125

Figure 7.4. For the semi-leptonic channel, four-vector of clustered top squark 2 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 2 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL 126

Figure 7.5. For the fully-leptonic channel, four-vector of clustered top squark 1 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 1 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL 127

Figure 7.6. For the fully-leptonic channel, four-vector of clustered top squark 2 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 2 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL 128

Figure 8.1. A dramatization of the ABCD method for two arbitrary variables. The number of background events in region A, the signal region, can be calculated as $N_{\mathrm{pred.},A} = (N_B N_C)/N_D$ 131

Figure 8.2. The "customized" Double DisCo neural network is shown with its three models. Mass regression and two binary clasifier models are shown with their input, hidden and output layers. Output of mass regression model is an input for each binary classifier 133

Figure 8.3. For the fully-hadronic channel, comparisons of loss as a function of epoch for the training (blue) and validation (yellow) data sets. The comparisons are shown for the mass regression component (top left), the classifier component (top right), the distance correlation component (bottom left) and ABCD closure component (bottom right) 143
Figure 8.4. For the semi-leptonic channel, comparisons of loss as a function of epoch for the training (blue) and validation (yellow) data sets. The comparisons are shown for the mass regression component (top left), the classifier component (top right), the distance correlation component (bottom left) and ABCD closure component (bottom right) 144
Figure 8.5. For the fully-leptonic channel, comparisons of loss as a function of epoch for the training (blue) and validation (yellow) data sets. The comparisons are shown for the mass regression component (top left), the classifier component (top right), the distance correlation component (bottom left) and ABCD closure component (bottom right) 145

Figure 8.6. For the fully-hadronic channel, the 2D background discriminant output used for training of the network (top left), the 2D background discriminant output used for validating of the network (top right) to check for over training. The mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for training of the network (bottom left) and the mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for validating of the network (bottom right). The truth label in the mass regression was chosen as the GEN-matched reconstructed top mass for background and the GEN-matched reconstructed top squark mass for signal 146
Figure 8.7. For the semi-leptonic channel, the 2D background discriminant output used for training of the network (top left), the 2D background discriminant output used for validating of the network (top right) to check for over training. The mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for training of the network (bottom left) and the mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for validating of the network (bottom right). The truth label in the mass regression was chosen as the GEN-matched reconstructed top mass for background and the GEN-matched reconstructed top squark mass for signal 147

Figure 8.8. For the fully-leptonic channel, the 2D background discriminant output used for training of the network (top left), the 2D background discriminant output used for validating of the network (top right) to check for over training. The mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for training of the network (bottom left) and the mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for validating of the network (bottom right). The truth label in the mass regression was chosen as the GEN-matched reconstructed top mass for background and the GEN-matched reconstructed top squark mass for signal 148
Figure 8.9. For the fully-hadronic channel, mass regression output (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 8,9,10,11,12$\geq$ 13 (left to right, top to bottom) 150
Figure 8.10. For the semi-leptonic channel, mass regression output (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 7,8,9,10,11,$\geq$ 12 (left to right, top to bottom) 151
Figure 8.11. For the fully-leptonic channel, mass regression output (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 6,78,9,10,$\geq$ 11 (left to right, top to bottom) 152
Figure 8.12. For the fully-hadronic channel, Discriminant 1 and 2 outputs (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 8,9,10,11,12,$\geq$ 13 (left to right, top to bottom) 154
Figure 8.13. For the semi-leptonic channel, Discriminant 1 and 2 outputs (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 7,8,9,10,11,$\geq$ 12 (left to right, top to bottom) 155

Figure 8.14. For the fully-leptonic channel, Discriminant 1 and 2 outputs (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 6,7,8,9,10,$\geq$ 11 (left to right, top to bottom) 156
Figure 8.15. For the fully-hadronic (top), semi-leptonic (middle), and fully-leptonic (bottom) channels, the ROC plot for discriminant 1 (left) and discriminant 2 (right), for each N_{jets} separately assuming Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ GeV as the signal. Performance is similar for the different N_{jets} categories and between the two discriminants 158
Figure 8.16. For the fully-hadronic channel, the ROC for discriminant 1 (left) and discriminant 2 (right) for N_{jets} = 8,10,$\geq$ 13 (top to bottom), showing results for several Stealth SY$\bar{Y}$ signal mass points 159
Figure 8.17. For the semi-leptonic channel, the ROC for discriminant 1 (left) and discriminant 2 (right) for N_{jets} = 7,9,$\geq$ 12 (top to bottom), showing results for several Stealth SY$\bar{Y}$ signal mass points 160
Figure 8.18. For the fully-leptonic channel, the ROC for discriminant 1 (left) and discriminant 2 (right) for N_{jets} = 6,8,$\geq$ 11 (top to bottom), showing results for several Stealth SY$\bar{Y}$ signal mass points 161
Figure 8.19. For the fully-hadronic channel, the 2D discriminant plane for $t\bar{t}+$ jets (top two rows) and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ (bottom two rows) separated in N_{jets} = 8,9,10,11,12,$\geq$ 13 (left to right, top to bottom) 163
Figure 8.20. For the semi-leptonic channel, the 2D discriminant plane for $t\bar{t}+$ jets (top two rows) and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ (bottom two rows) separated in N_{jets} = 7,8,9,10,11,$\geq$ 12 (left to right, top to bottom) . 164

Figure 8.21. For the fully-leptonic channel, the 2D discriminant plane for $t\bar{t}+$ jets (top two rows) and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ (bottom two rows) separated in N_{jets} = 6,7,8,9,10,$\geq$ 11 (left to right, top to bottom) ... 165
Figure 8.22. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the non-closure (left) and pull (right) as a function of "non-optimized" ABCD bin edges for $N_{\text{jets}} = 8$ (for fully-hadronic channel), $N_{\text{jets}} = 7$ (for semi-leptonic channel) and $N_{\text{jets}} = 6$ (for fully-leptonic channel). Overall non-closure and pull are about ~ 0 across a majority of the 2D plane of ABCD edge choices 168
Figure 8.23. For the fully-hadronic channel, significance as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 8,9,10,11,12, $\geq$ 13 bins (left to right, top to bottom) 170
Figure 8.24. For the semi-leptonic channel, significance as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 7,8,9,10,11,$\geq$ 12 bins (left to right, top to bottom) 171
Figure 8.25. For the fully-leptonic channel, significance as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 6,7,8,9,10,$\geq$ 11 bins (left to right, top to bottom) 172
Figure 8.26. For the fully-hadronic (top), semi-leptonic (middle), fully-leptonic (bottom) channels, significance as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,$\geq$13 bins (left to right). N_{jets} = 7,9,$\geq$12 bins (left to right). N_{jets} = 6,8,$\geq$11 bins (left to right). According to signal fraction, the best significance is observed in the highest N_{jets} bin for each channel where the Disc. 1 and Disc. 2 edges are crossed 177

Figure 8.27. For the fully-hadronic (top), semi-leptonic (middle), fully-leptonic (bottom) channels, non-closure as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,$\geq$13 bins (left to right). N_{jets} = 7,9,$\geq$12 bins (left to right). N_{jets} = 6,8,$\geq$11 bins (left to right). As expected for lower N_{jets} bins, non-closure is less then $\sim 5\%$, where the Disc. 1 and Disc. 2 edges are crossed .. 178

Figure 8.28. For the fully-hadronic (top), semi-leptonic (middle), fully-leptonic (bottom) channels, pull as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,$\geq$13 bins (left to right). N_{jets} = 7,9,$\geq$12 bins (left to right). N_{jets} = 6,8,$\geq$11 bins (left to right). In general, the pull is about zero as expected where the Disc. 1 and Disc. 2 edges are crossed for a given channel 179

Figure 8.29. For the fully-hadronic channel, signal fractions in each B, C, D region (left to right) as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,$\geq$13 bins (top to bottom). The signal fractions for **B**, **C**, and **D** regions are about zero for lower N_{jets} bins for a given channel 180

Figure 8.30. For the semi-leptonic channel, signal fractions in each B, C, D region (left to right) as a function of "optimized" ABCD bin edges choice. N_{jets} = 7,9,$\geq$12 bins (top to bottom) 181

Figure 8.31. For the fully-leptonic channel, signal fractions in each B, C, D region (left to right) as a function of "optimized" ABCD bin edges choice. N_{jets} = 6,8,$\geq$11 bins (top to bottom) 182

Figure 8.32. Validation regions are defined based on the sub-division of each "BD" regions, "CD" regions and "D" region 183

Figure 8.33. Validation region I with ABCD bin edges and right-most boundary as sub-division of BD regions (top). Validation region II with ABCD bin edges and top-most boundary as sub-division of CD regions (middle). Validation region III with ABCD bin edges and right-most and top-most boundaries as sub-division of D region (bottom) 185
Figure 8.34. For the fully-hadronic channel, non-closure as a function of validation I edges, validation II edges, and validation III edges (left to right). Top to bottom $N_{\text{jets}} = 8, 10, \geq 13$ bins. The large values of the non-closure are observed for the higher N_{jets} bins due to the statistical uncertainty on the closure 187
Figure 8.35. For the semi-leptonic channel, non-closures as a function of validation I edges, validation II edges, and validation III edges (left to right). Top to bottom $N_{\text{jets}} = 7, 9, \geq 12$ bins. The large values of the non-closure are observed for the higher N_{jets} bins due to the statistical uncertainty on the closure 188
Figure 8.36. For the fully-leptonic channel, non-closures as a function of validation I edges, validation II edges, and validation III edges (left to right). Top to bottom $N_{\text{jets}} = 6, 8, \geq 11$ bins. The large values of the non-closure are observed for the higher N_{jets} bins due to the statistical uncertainty on the closure 189
Figure 8.37. For the fully-hadronic channel, the closure per-N_{jets} for $t\bar{t}$ + jets in the ABCD region (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) 190

Figure 8.38. For the semi-leptonic channel, the closure per-N_{jets} for $t\bar{t}+$jets in the ABCD region (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) In ABCD, Validation CD and Validation D regions, the non-closure values are still reasonable even if pull values are far from 0 191
Figure 8.39. For the fully-leptonic channel, the closure per-N_{jets} for $t\bar{t}+$jets in the ABCD region (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) 192
Figure 8.40. For the fully-hadronic channel, the closure corrections as a function of boundary value of the validation regions. The closure corrections for N_{jets} = 8, 9, 10, 11, 12, ≥ 13 (left to right, top to bottom) and calculated from $t\bar{t}+$jets simulation 194
Figure 8.41. For the semi-leptonic channel, the closure corrections as a function of boundary value of the validation regions. The closure corrections for N_{jets} = 7, 8, 9, 10, 11, ≥ 12 (left to right, top to bottom) and calculated from $t\bar{t}+$jets simulation 195
Figure 8.42. For the fully-leptonic channel, the closure corrections as a function of boundary value of the validation regions. The closure corrections for N_{jets} = 6, 7, 8, 9, 10, ≥ 11 (left to right, top to bottom) and calculated from $t\bar{t}+$jets simulation 196

Figure 8.43. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the closure correction ratios for all variations (left) and the gray band given on this plot is the pure statistical uncertainty on the nominal closure correction. A single example of closure correction for the modeling variation that affects "color reconnection" (right) and the top panel compares it with the nominal closure correction derived per N_{jets}, while the bottom panel then shows the closure correction ratio between them. Note that there is an offset to be applied for numbering the N_{jets} bins on the x-axis of all plots 199

Figure 8.44. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels; the closures corrections (left), data closures (middle) and corrected data closures (right) as a function of boundary value for $N_{\mathrm{jets}} = 8$ (for fully-hadronic channel), $N_{\mathrm{jets}} = 7$ (for semi-leptonic channel), $N_{\mathrm{jets}} = 6$ (for fully-leptonic channel) 201

Figure 8.45. For the fully-hadronic channel, the closure correction per-N_{jets} for ABCD (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) 202

Figure 8.46. For the semi-leptonic channel, the closure correction per-N_{jets} for ABCD (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) 203

Figure 8.47. For the fully-leptonic channel, the closure correction per-N_{jets} for ABCD (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) 204

Figure 8.48. For the fully-hadronic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 8$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom) 207

Figure 8.49. For the semi-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 7$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom) 208

Figure 8.50. For the fully-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 6$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom) 209

Figure 9.1. For the fully-hadronic (left) and semi-leptonic (right) channels, data *vs* simulation comparisons of the A, B, C, D NN bins in the two QCD CRs for the Stealth SY$\bar{Y}$ trainings 212

Figure 9.2. For the fully-hadronic (left) and semi-leptonic (right) channels, comparisons of the N_{jets} (top), Double DisCo NN Disc.1 (middle) and Disc.2 (bottom) distributions of the QCD multi-jet simulation between QCD CR and SR. Good agrement of the N_{jets} distributions for QCD CR and SR. Double DisCo NN trained on the Stealth SY$\bar{Y}$ model. Comparisons of Disc. 1 and Disc. 2 give further confidence that QCD CR selection is matched to SR selection 214

Figure 9.3. For the fully-hadronic (left) and semi-leptonic (right) channels, comparison between the QCD multi-jet background prediction and the QCD multi-jet background simulation in the SR 216

Figure 10.1. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the “Background Only” fit to pseudaData. In each plot, the bottom panel shows the goodness-of-fit. The signal is overlaid for illustration purposes here, it is not included in the fit 228

Figure 10.2. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the “Background Only” fit to pseuodaDataS. In each plot, the bottom panel shows the goodness of fit. Note that the pseuodaDataS is with Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV 230

Figure 10.3. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the “Background + Signal” fit to pseuodaDataS. In each plot, the bottom panel shows the goodness of fit. Note that the pseuodaDataS is with Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV 232

Figure 10.4. For the semi-leptonic channel, “Asimov style” impact plots for Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV with pseuodaData (top) and pseuodaDataS (bottom). $\hat{r}$ = 0 with ±11% uncerstainty for pseuodaData and $\hat{r}$ = 1 with ±14% uncerstainty for pseuodaDataS 234

Figure 11.1. For the fully-hadronic (top left), semi-leptonic (top right), fully-leptonic (bottom left) and combination of all channels (bottom right), the observed limits for Stealth SY$\bar{Y}$ model. The analysis is sensitive to the Stealth SY$\bar{Y}$ model up to $m_{\tilde{t}}$ = 900 GeV. Note that limits are with “Asimov style” 235

Figure 11.2. The local p-values as a function of top squark mass for the Stealth SY$\bar{Y}$ model with Run2 simulated events. Results are shown for separate fits of the fully-hadronic (green dashed line), semi-leptonic (blue dashed line), fully-leptonic (red dashed line) channels, and combination of all three channels (black line). The bottom panel shows the signal strength for combination of the all channels with uncertainty denoted by the green band. Note that limits are with “Asimov style” 237

Figure A1. Plots of the number of resolved top tags versus number of merged top tags for $t\bar{t}+$jets (top-left), QCD multi-jet (top-right), and Stealth SY$\bar{Y}$ signal events that pass the fully-hadronic event selection. The selected top squark mass points are 350 (middle left), 550 (middle right), 850 (bottom left), and 1150 (bottom right) GeV. Each colored square represents a possible combination of resolved and merged top tags, while the number reports the fraction of events for the specific top squark mass 245

Figure C1. For the fully-hadronic channel, a sample of input variables used in the final training of the neural network. The top row is the 4-vector of the leading p_T jet. N.B. This jet is rotated to $\phi = 0$ and all other objects in an event rotated accordingly. To avoid passing a delta function to the network, the jet 1 ϕ is not passed as an input. The top middle row is the 4-vector of the leading p_T top squark object. The bottom middle row shows Fox-Wolfram moments 2 through 5. The bottom row shows the energy-momentum tensor eigenvalues 248

Figure C2. For the semi-leptonic channel, a sample of input variables used in the final training of the neural network. The top row is the 4-vector of the leading p_T jet. N.B. This jet is rotated to $\phi = 0$ and all other objects in an event rotated accordingly. To avoid passing a delta function to the network, the jet 1 ϕ is not passed as an input. The top middle row is the 4-vector of the leading p_T top squark object. The bottom middle row shows Fox-Wolfram moments 2 through 5. The bottom row shows the energy-momentum tensor eigenvalues 249

Figure C3. For the fully-leptonic channel, a sample of input variables used in the final training of the neural network. The top row is the 4-vector of the leading p_{T} jet. N.B. This jet is rotated to $\phi = 0$ and all other objects in an event rotated accordingly. To avoid passing a delta function to the network, the jet 1 ϕ is not passed as an input. The top middle row is the 4-vector of the leading p_{T} top squark object. The bottom middle row shows Fox-Wolfram moments 2 through 5. The bottom row shows the energy-momentum tensor eigenvalues 250

Figure D1. For the fully-hadronic channel, the non-closure as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 8,9,10,11,12,$\geq$13 bins (left to right, top to bottom). Overall non-closure is about ~ 0 as expected by the sense of perfect ABCD closure in the lower N_{jets} bins 251

Figure D2. For the semi-leptonic channel, the non-closure as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 7,8,9,10,11,$\geq$12 bins (left to right, top to bottom). Overall non-closure is about ~ 0 as expected by the sense of perfect ABCD closure in the lower N_{jets} bins 252

Figure D3. For the fully-leptonic channel, the non-closure as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 6,7,8,9,10,$\geq$11 bins (left to right, top to bottom). Overall non-closure is about ~ 0 as expected by the sense of perfect ABCD closure in the lower N_{jets} bins 253

Figure D4. For the fully-hadronic channel, pull as a function of "non-optimized" ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 8,9,10,11,12,≥13 bins (left to righti, top to bottom) 254
Figure D5. For the semi-leptonic channel, pull as a function of "non-optimized" ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 7,8,9,10,11,≥12 bins (left to right, top to bottom) 255
Figure D6. For the fully-leptonic channel, pull as a function of "non-optimized" ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 6,7,8,9,10,≥11 bins (left to right, top to bottom) 256
Figure D7. For the fully-hadronic channel, the significance as a function of "optimized" ABCD bin edges. N_{jets} = 8,9,10,11,12≥13 bins (left to right, top to bottom) 257
Figure D8. For the semi-leptonic channel, the significance as a function of "optimized" ABCD bin edges. N_{jets} = 7,8,9,10,11,≥12 bins (left to right, top to bottom) 258
Figure D9. For the fully-leptonic channel, the significance as a function of "optimized" ABCD bin edges. N_{jets} = 6,7,8,9,10,≥11 bins (left to right, top to bottom) 259
Figure D10. For the fully-hadronic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 8,9,10,11,12≥13 bins (left to right, top to bottom) 260
Figure D11. For the semi-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 7,8,9,10,11,≥12 bins (left to right, top to bottom) 261

Figure D12. For the fully-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 6,7,8,9,10,$\geq$11 bins (lest to right, top to bottom) ... 262
Figure D13. For the fully-hadronic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 8,9,10,11,12$\geq$13 bins (left to right, top to bottom) ... 263
Figure D14. For the semi-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 7,8,9,10,11,$\geq$12 bins (left to right, top to bottom) ... 264
Figure D15. For the fully-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 6,7,8,9,10,$\geq$11 bins (lest to right, top to bottom) ... 265
Figure D16. For the fully-hadronic channel, the signal fractions in each **B**, **C**, **D** region (left to right) as a function of "optimized" ABCD bin edges. N_{jets} = 8,9,10,11,12$\geq$13 bins (top to bottom) ... 266
Figure D17. For the semi-leptonic channel, the signal fractions in each **B**, **C**, **D** region (left to right) as a function of "optimized" ABCD bin edges. N_{jets} = 7,8,9,10,11,$\geq$12 bins (top to bottom) ... 267
Figure D18. For the fully-leptonic channel, the signal fractions in each **B**, **C**, **D** region (left to right) as a function of "optimized" ABCD bin edges. N_{jets} = 6,7,8,9,10,$\geq$11 bins (top to bottom) ... 268
Figure D19. For the fully-hadronic channel, the non-closure as a function of validation I (top) edges, validation II (middle) edges, and validation III (bottom) edges. Left to right N_{jets} = 9,11,12 bins. Note that these are for the Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV model ... 269

Figure D20. For the semi-leptonic channel, the non-closure as a function of validation I (top) edges, validation II (middle) edges, and validation III (bottom) edges. Left to right N_{jets} = 8,10,11 bins. Note that these are for the Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV model 270

Figure D21. For the fully-leptonic channel, the non-closure as a function of validation I (top) edges, validation II (middle) edges, and validation III (bottom) edges. Left to right N_{jets} = 7,9,10 bins. Note that these are for the Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV model 271

Figure D22. For the fully-hadronic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 9$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom) 272

Figure D23. For the semi-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 8$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom) 273

Figure D24. For the fully-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 7$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom) 274

Figure E1. For the fully-hadronic channel, asimov style impact plots for Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV with pseuodaData (top) and pseuodaDataS (bottom) 275

Figure E2. For the fully-leptonic channel, asimov style impact plots for Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV with pseuodaData (top) and pseuodaDataS (bottom) 276

ABBREVATIONS

ADC	:	Analog-to-Digital Converter
ALICE	:	A Large Ion Collider Experiment
AMC13	:	Advanced Mezzannine Card 13
APD	:	Avalanche Photo Diode
ASIC	:	Application-Specific Integrated Circuit
ATLAS	:	A Toroidal LHC ApparatuS
BCM1F	:	Fast Beam Conditions Monitor
BCML	:	Beam Condition Monitor for Loss monitoring
BE	:	Back-End
BHM	:	Beam Halo Monitor
BPTX	:	Beam Position for Timing at the eXperiments
BRIL	:	Beam Radiation Instrumentation and Lumi- nosity
BSM	:	Beyond the Standard Model
cDAQ	:	central Data Acquisition
CERN	:	Conceil Europeenne pour la Recherche Nucleaire
CMS	:	Compact Muon Selenoid
CSCs	:	Cathode Strip Chambers
CU	:	Calibration Unit
DAQ	:	Data Acquisition
DTs	:	Drift Tubes
ECAL	:	Electromagnetic Calorimeter
ES	:	Endcap preShower detector
EYTS	:	Extended Year End Technical Stop
FE	:	Front-End
FPGAs	:	Field Programmable Gate Arrays

FSR	: Final State Radiation
GCT	: Global Calorimeter Trigger
GMT	: Global Muon Trigger
GT	: Global Trigger
HB	: Hadronic Barrel
HCAL	: Hadronic Calorimeter
HE	: Hadronic Endcap
HF	: Hadronic Forward
HFlumi	: Hadron Forward calorimeter for luminosity
HLT	: High-Level Trigger
HO	: Hadronic Outer
HPD	: Hybrid Photo Diode
IP	: Interaction Point
ISR	: Initial State Radiation
L1	: Level 1
L1A	: Level 1 Accept
LED	: Light Emitting Diode
LHC	: Large Hadronic Collider
LHCb	: LHC-beauty
LINAC2	: Linear Accelerator 2
LS1	: Long Shutdown 1
LSP	: Lightest Supersymmetric Particle
MCH	: MicroTCA Carrier Hub
MIB	: Machine Induced Background
ngCCM	: new generation Clock Control Module
PF	: Particle Flow
PLT	: Pixel Luminosity Telescope

PMT : Photo Multiplier Tube

pp : proton-proton

PS : Proton Synchrotron

PSB : Proton Synchrotron Booster

QCD : Quantum Chromodynamics

QED : Quantum Electrodynamics

QFT : Quantum Field Theory

QGP : Quark-Gluon Plasma

QIE : Charge Integrated and Endcoding

RCT : Regional Calorimeter Trigger

RPCs : Resistive Plate Chambers

RPV : R-Parity Violating

SiPM : Silicon Photo Multiplier

SM : Standard Model

SPS : Super Proton Synchrotron

SR : Search Region

SUSY : Supersymmetry

TDC : Time-to-Digital Converter

TPs : Trigger Primitives

uHTR : micro HCAL Trigger Readout

uTCA : Micro Telecommunication and Computing Architecture

VEV : Vacuum Expectation Value

VME : Versa Module Eurocard

VPT : Vacuum Photo Triode

VTTx : Versatile Twin Transmitter

WLS : Wave-Length Shifting

LVII

1. INTRODUCTION

Particle physics is a branch of physics that studies the basic building blocks of matter and the interactions between them. Most of the elementary particles are not found in nature under normal conditions. But in particle accelerators, they are produced during high-energy collisions of other particles. The most important and widely accepted theory of Particle Physics, the Standard Model (SM), was developed over the last century to describe the known fundamental particles and forces. The full set of fundamental particles in the SM has been successfully completed with the discovery of the Higgs boson at the European Organization for Nuclear Research (CERN) by A Toroidal LHC ApparatuS (ATLAS) and Compact Muon Selenoid (CMS) in 2012. Although the SM is theoretically self-consistent, it leaves some phenomena unexplained. This motivates particle physicists to search for evidence of new particles beyond the standard model (BSM). There are many well motivated and previously unexplored physics models beyond the SM, such as Supersymmetry (SUSY).

Many searches for SUSY have been done at the LHC, and strong limits on top squark and gluino masses have been set for many simplified models. Supersymmetric particles decay through SM and undetected lightest SUSY particles (LSPs) in almost all traditional SUSY searches. Most of the SUSY searches look for an excess of events with large missing transverse momentum (p_{T}^{miss}) originating from the undetected LSPs, usually the lightest neutralino. Therefore, these searches would not be sensitive to well-motivated models of new physics that do not produce large p_{T}^{miss}, such as Stealth SUSY (Fan, Krall, et al., 2016, Fan, Reece, and Ruderman, 2012).

In the Stealth SUSY model, a "hidden sector" is introduced where SUSY is a nearly conserved symmetry. This hidden sector contains a supersymmetric fermion and its scaler partner. The supersymmetric fermion singlino ($\bar{S}$) decays through the hidden sector to very close-in-mass its scaler partner singlet (S) and a gravitino ($\bar{G}$), as given in Figure 1.1.. In this case, gravitino is the LSP and its mass is usually assumed to be around 1 GeV. Then, the top squark could decay through this hidden sector and $\bar{G}$ would carry low-p_{T}^{miss}. The singlet can decay back to SM particles via the "SY$\bar{Y}$" coupling, which is allowed by the hidden sector and the singlet decay results in $S \rightarrow gg$. Considering the Stealth SUSY with SY$\bar{Y}$ coupling, there are two top quarks, many light-flavor gluon jets and low-p_{T}^{miss} in the final state of the top squark decays.

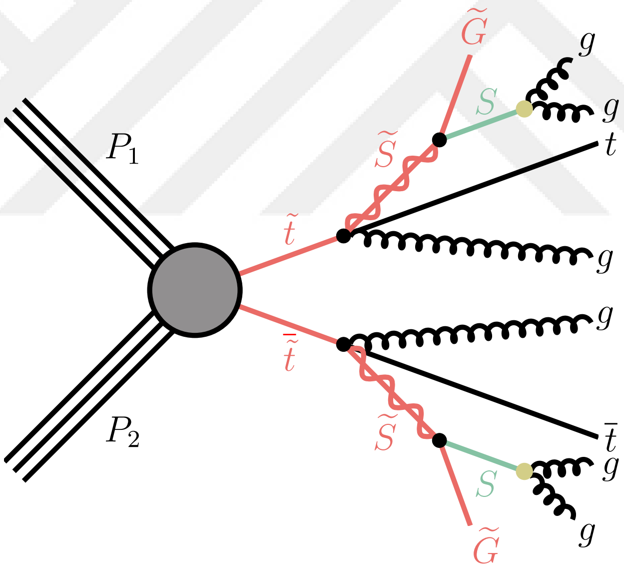


Figure 1.1. Top squark decay mode considered in this analysis. The Stealth SY$\bar{Y}$ model. Note that a vertex in the diagram is effective couplings, like in the singlet decay

In this thesis, a novel search for top squark pair production followed by the decay of each top squark into a top quark and three light-flavor gluon jets via the benchmark Stealth SY$\bar{Y}$ model is presented. The fully-hadronic, semi-leptonic and fully-leptonic final states of top squark decays are included. Additionally, a novel machine learning technique (Double DisCo NN) is used as a part of the $t\bar{t}+$jets background estimation. The Double DisCo NN has not been used in any analysis in the CMS yet.

Chapter 2 is dedicated to summarizing the SM as well as the shortcomings of the SM. Additionally, more details about the SUSY models are mentioned in this chapter. Chapter 3 contains general information about the CMS detector, where data used in this analysis is taken. Chapter 4 includes information about analysis motivation and strategy. A brief of the legacy version of this analysis is also mentioned in this chapter. Chapter 5 goes through the data and simulation samples, trigger performance, object and event selection criteria. Chapter 6 focuses on the hadronic top tagging for the fully-hadronic final state of the analysis. In addition to the tagging algorithm, a little study is dedicated to improving the performance of this tagging algorithm. Chapter 7 also focuses on a specific study, which is loose reconstruction of the top squark four-vector. The goal of this study is to get another set of variables related to the signal topologies to feed into the Double DisCo NN. Chapter 8 is dedicated to the estimation of the $t\bar{t}+$jets background and calculation of systematic uncertainties of it. The QCD multi-jet background and other minor background estimations and related systematics are given in Chapter 9. After calculating all systematic uncertainties, fit calculations are performed using the Higgs Combine tool and the fit results are presented in Chapter 10. Right after fit results, the preliminary limits are given in Chapter 11. Last, Chapter 12 presents the conclusion.

2. THEORETICAL FRAMEWORK

The theoretical framework of the SM has remarkable success in that there is a wide range of precise experimental measurements at the LHC in the last couple of decades, including the discovery of the Higgs boson (S. e. a. Chatrchyan, 2012). The main ingredients of the SM are the *Dirac equation* of relativistic quantum mechanics describing the dynamics of the fermions, *Quantum Field Theory* providing a fundamental description of the particles and their interactions, the *local gauge principle* determining the exact nature of these interactions, and the *Higgs mechanism* of electroweak symmetry breaking generating particle masses. Although the SM is one of the most well-tested theories in particle physics with the discovery of Higgs boson, there are a lot of unexplained mysteries that can not be understood within the SM's mathematical framework or lack of experimental evidence. This opens a new window to Beyond the Standard Model (BSM).

In this chapter, a brief overview of the SM is discussed in Section 2.1.. Next, in Section 2.2., the discussions about shortcomings of the SM are presented. Last, SUSY is introduced as an example of the BSM physics in Section 2.3.. Additionally, two of the supersymmetric models, which are based on the analysis in this thesis, are mentioned in the last section.

2.1. Standard Model (SM)

The SM is a relativistic Quantum Field Theory (QFT), which combines gauge field theories of particle physics into a single framework to describe all interactions of subatomic particles, except those due to gravity.

The elementary particles of the SM are grouped into two sectors, which are fermions and bosons. The elementary particles of the SM are shown in Figure 2.1., which is a table of all the fundamental particles currently discovered along with a few of their fundamental properties such as spin, charge, mass.

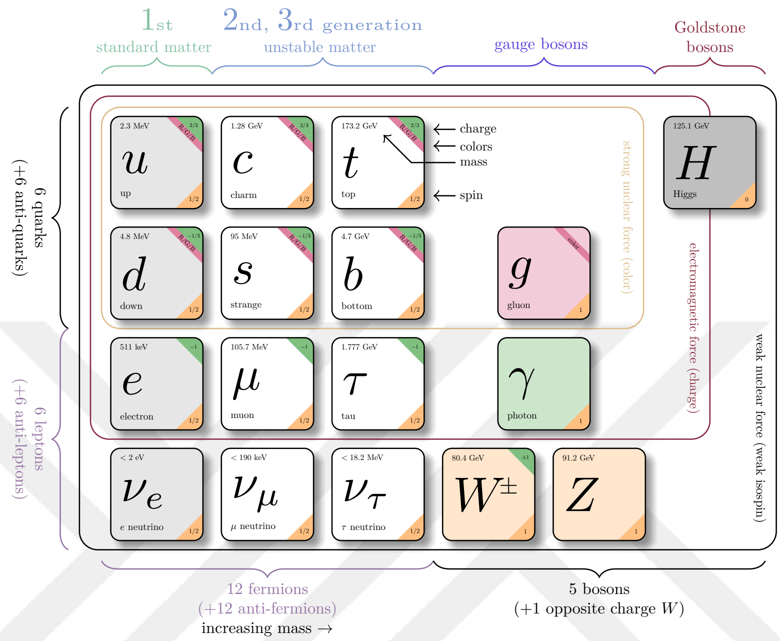


Figure 2.1. The elementary particles of the SM. The model includes three generations of matter particles (leptons and quarks) as well as the gauge and Higgs bosons. There are given properties of the particles, including the particle name, symbol, mass, spin, electric charge, and color charge

The fermions are the particles making up all visible matter in the universe. They follow the Pauli's exclusion principle and have a spin $\frac{1}{2}$. Each fermion has a corresponding anti-particle that has the same mass, but with the opposite quantum number and charge. The fermions are divided into two categories: *quarks* and *leptons*. There are six quarks (u, d, c, s, t, b) and six leptons $(e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau)$. The quarks interact via the electromagnetic and weak forces, but unlike the lepton, quarks can also interact via the strong force. Based on the interaction of the strong force, the quarks also have color charge, which can have three values—referred to as *"red"*, *"green"*, and *"blue"*. Also, antiquarks contain charges of *"anti-red"*, *"anti-green"*, *"anti-blue"*. Interacting via the strong force binds the quarks together inside hadrons, where mesons consist of $q\bar{q}$ and baryons consist of qqq. The fermions come in three generations with identical quantum numbers but different masses. Each generation comprises a weakly charged doublet of quarks with three different color charges, a neutrino and an electrically charged lepton (Altarelli, 2005).

The bosons are force carrier particles in the SM and they are in charge of the interactions between the fermions. They have integer spin and they are divided into two subcategories, which are the *gauge (vector)* and the *Higgs (scalar)* bosons. The gauge bosons have a spin of 1 while the Higgs boson has a spin of 0, as shown in Table 2.1.. The gauge bosons are defined as force carriers that mediate the strong, the weak, and the electromagnetic interactions. The eight gluons mediate the strong interactions between color charged particles (the quarks). They are massless. The eightfold multiplicity of gluons is labeled by a combination of color and anticolor charge. Because gluons have an effective color charge, they can also interact among themselves. The photons mediate the electromagnetic force between electrically charged particles. Like gluons, the photon is also massless. The $W^{\pm}$ and Z^{0} bosons mediate the weak interactions between particles of different flavours (all quarks and leptons). They are massive and they, along with the photon, are carriers of the electroweak force. The Higgs boson is a scalar boson with a spin of 0, with no electric or color charge. It plays a unique role in the SM with two main functions. The first function is to provide the masses of $W^{\pm}$ and Z^{0} bosons through spontaneous symmetry breaking. The second one is to provide the masses for all the fermions through Yukawa interactions. More details about the Higgs boson are given in Section 2.1.4..

Table 2.1. The bosons of the SM subdivided by their quantum numbers: electric charge Q, third component of weak isospin T_3, weak hypercharge Y, and color charge C. * There are eight linear color-anticolor combinations corresponding to the eight different gluons

Bosons	**Mass**	**Spin**	Q	T_3	Y	C
Gluon	0	1	0	0	0	(r,g,b)+$(\bar{r},\bar{g},\bar{b})^*$
Photon	0	1	0	0	0	-
$W^{\pm}$	80.4	1	±1	±1	0	-
Z^0	91.2	1	0	0	0	-
Higgs	125.2	0	0	$-\frac{1}{2}$	1	-

The interactions in the SM are mediated by gauge bosons associated with the symmetry groups $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. This symmetry group describes three of the four known fundamental interactions in nature, excluding gravity. The symmetry group $SU(3)_C$ describes strong interactions via the Quantum Chromodynamics (QCD), discussed in Section 2.1.2., while the group $U(1)_Y$ describes electromagnetic interactions via the Quantum Electrodynamics (QED), discussed in Section 2.1.1.. The group $SU(2)_L \otimes U(1)_Y$ describes unification of electromagnetic and weak interactions via the electroweak theory, discussed in Section 2.1.3.. The subscript C refers to color charge, whereas L refers to coupling to left-handed fermions. The subscript Y refers to the "weak hypercharge" and it is determined from the electric charge and 3^{rd} component of isospin. The electric charge, Q, is given as the relation among them.

$$Q = T_3 + \frac{1}{2}Y \tag{2.1}$$

2.1.1. Quantum Electrodynamics (QED)

Quantum Electrodynamics (QED) (Aitchison, I. and Hey, A., 2013) is a quantum field theory (QFT) describing how the leptons interact with each other through the electrodynamic force mediated by a massless photon. It was the first theory where full agreement between quantum mechanics and special relativity was achieved and was developed by Tomonaga, Schwinger, and Feynman. They won the Nobel Prize in Physics in 1965 for their contributions to the theory.

In a QFT, particles are represented by fields, which are in turn represented mathematically by Lagrangian densities $\mathcal{L}$. On the mathematical level, the QED Lagrangian, Eq.2.2, models an abelian gauge theory under the $U(1)$ local gauge symmetry. Electromagnetic interaction can be described by the QED Lagrangian.

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^{\mu}D\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - q\bar{\psi}\gamma^{\mu}A_{\mu}\psi \tag{2.2}$$

where ψ is the bispinor field of the spin $\frac{1}{2}$ particle, γ is the photon as the mediator of electromagnetic interactions, D_{μ} is the gauge covariant derivative, m and q are the charge and mass of the particle, A_{μ} is the covariant four-potential of the electromagnetic field of the particle, $F_{\mu\nu}$ is the electromagnetic field tensor. The kinetic nature of the fermion and photon are described by the first two terms. The last term in the equation describes photon and fermion interaction.

Based on the Feynman rules, there is no QED vertex connecting more than three particles, because the electromagnetic interaction is between a single photon and two spin-half fermions. For this very reason, all valid QED processes are described by Feynman diagrams formed from the basic three-particle QED vertex.

2.1.2. Quantum Chromodynamics (QCD)

Quantum Chromodynamics (QCD) is a non-abelian gauge field theory with the $SU(3)$ local gauge symmetry (Aitchison, I. J. R. and Hey, A. J. G., 2013), describing how quarks interact with each other through the strong force mediated by eight massless gluons. Akin to electric charge in QED, color charge is the conserved quantity in QCD. There are three varieties of color charge labelled as red (R), blue (B), green (G), where color is simply a label for the orthogonal states in the $SU(3)$ color space. Both quarks and gluons carry color charges. Only particles having non-zero color charge undergo strong interactions. Then, gluons can also interact with each other, which is referred to as self-interaction.

In order to describe the strong interaction between quarks and gluons, the Lagrangian of QCD can be described as in Equation 2.3, where ψ_i is the quark field, D_μ is the gauge covariant derivative, $F^a_{\mu\nu}$ is the gluon field tensor. The first term describes the dynamics of the quarks and the second term describes the dynamical properties of the gluon.

$$\mathcal{L}_{QCD} = \bar{\psi}_i(i\gamma^\mu(D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a \tag{2.3}$$

The gluon field tensor is given as in Equation 2.4, where g_s is the coupling constant of the strong force, f_{abc} is the structure constants of SU(3). The last term of the gluon field tensor gives rise to gluon self-interactions (Thomson, M., 2013) based on the non-abelian gauge field. The mathematical forms of these triple and quartic gluon vertices, shown in Figure 2.2., are specified by $SU(3)$ gauge symmetry. The gluon self-interaction causes the strength of the strong interaction and leads to two remarkable features: “asymptotic freedom” and “color confinement”.

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f^{abc} A^b_\mu A^c_\nu \tag{2.4}$$

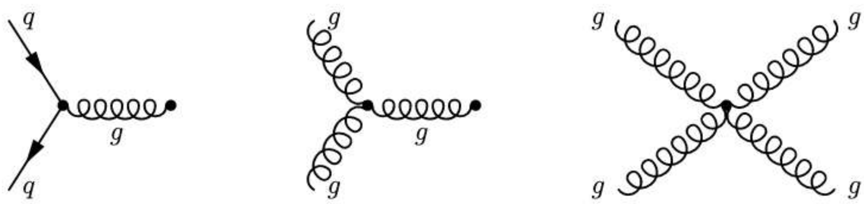


Figure 2.2. The allowed particle interactions via the QCD interaction arising from the requirement of $SU(3)$ local gauge invariance. The form of qqg interaction vertex (left), triple gluon interaction vertex (middle), quartic gluon interaction vertex (right) (Thomson, M., 2013)

Asymptotic freedom and color confinement are the most prominent properties of QCD. According to theory (Gross and Wilczek, 1973), asymptotic freedom originates in QCD running coupling constant, α_s, which determines the strength of an interaction. At low energy or large distances, α_s becomes larger and this makes the traditional perturbative approach to field theory impossible. However, at high energy or short distances, α_s approaches zero which means that the theory is asymptotically free and it models the strong interaction. Another prominent property of QCD is color confinement. At the higher energies, α_s becomes weaker, which explains why free quarks are never observed in nature. This is because removing a free quark from a bound state needs more energy than creating a quark anti-quark pair from the vacuum. So, quarks and gluons are not seen directly in experiments. Instead of this, they turn into hadrons, which are observed in the detectors. This process is called hadronization (Thomson, M., 2013), which is a non-perturbative process following parton showering and yields to the production of collimated beams of particles called "jets".

2.1.3. Electroweak Theory

The electromagnetic and weak nuclear forces are unified as the electroweak force with contributions of Glashow (Glashow, 1961), Salam (A. Salam and Ward, 1964), Weinberg (Weinberg, 1967)—awarded the 1979 Nobel Prize in Physics. This sector can be described as a single non-abelian gauge theory with $SU(2)_L \otimes U(1)_Y$ symmetry group, where two additional quantum numbers are introduced: weak isospin (T) and weak hypercharge (Y).

Based on the Goldstone theorem (Goldstone, Abdus Salam, and Weinberg, 1962), there are four massless mediators. The gauge bosons W_1, W_2, W_3 given by $SU(2)_L$ symmetry group form a weak isospin triplet, and gauge boson B given by $U(1)_Y$ symmetry group is a weak isospin singlet. A spontaneous symmetry breaking associated with Higgs mechanism, more details given in Section 2.1.4., explains that $W^{\pm} and\ Z^0$ are massive although the photon is massless.

The electroweak theory is just the Lie algebra with the additional requirement that the $SU(2)$ symmetry only acts on the left-handed particles—$SU(2)_L$. This means that left-handed fermions are doublets with weak isospin $\pm\frac{1}{2}$ and right-handed fermions are singlets with weak isospin 0 under $S(U)_L$ symmetry.

The Lagrangian of electroweak theory "before symmetry breaking" is given as:

$$
\begin{aligned}
\mathcal{L}_{EW} = & -\frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\
& + \bar{Q}_j i \not{D} Q_j + \bar{u}_j i \not{D} u_j + \bar{d}_j i \not{D} d_j + \bar{L}_j i \not{D} L_j + \bar{e}_j i \not{D} e_j
\end{aligned}
\tag{2.5}
$$

The first line of the Equation 2.5 describes the interaction between the three W vector bosons and the B vector boson, where the $W^a_{\mu\nu}$ (a = 1, 2, 3) and $B_{\mu\nu}$ are the field strength tensors for the weak isospin and weak hypercharge gauge fields. The second line of the Equation 2.5 describes the interaction of the gauge bosons and the fermions are through the gauge covariant derivative, where the subscript j sums over the three generations of fermions; Q, u, and d are the left-handed doublet, right-handed singlet up, and right handed singlet down quark fields; and L and e are the left-handed doublet and right-handed singlet electron fields; $\not{D}$ means the contraction of the 4-gradient with the Dirac matrices.

2.1.4. Brout-Englert-Higgs Mechanism

The Higgs mechanism relies on the Goldstone theorem (ibid.), which predicts the appearance of a massless spinless particle called Goldstone boson when the ”global” symmetry of a scalar field is spontaneously broken. This spontaneous symmetry breaking also can be described for ”local” gauge symmetries by the Higgs mechanism, proposed by Brout, Englert and Higgs (Englert and Brout, 1964, Higgs, 1964).

The Higgs mechanism gives rise to the Higgs field together with the Higgs boson, which refers to the generation of masses for the $W^{\pm}$and Z^0 gauge bosons through spontaneous symmetry breaking. The $SU(2)_L \times U(1)_Y$ symmetry must be broken to $U_{EM}(1)$ by scalar Higgs field.

For three gauge bosons to acquire mass, the scalar fields need to contain at least three degrees of freedom for the mechanism to work. The easy way to do this is a complex scalar field in the spinor representation of $SU(2)_L$, Φ, and it is introduced:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.6}$$

On the mathematical side, to simplify expressions, the Lagrangian of electroweak interaction with only the electron doublet is given by Equation 2.7, where i represents each of three generations, a represents number of generators in the gauge group, μ and ν are Lorentz indices.

$$\mathcal{L}_{EW} = -\frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \bar{L}_i(iD_\mu\gamma^\mu)L_i + \bar{e_{R,i}}(iD_\mu\gamma^\mu)e_{R,i} \tag{2.7}$$

The electroweak field strengths are given by Equation 2.8 and the covariant derivatives for left- and right-handed leptons are given by Equation 2.9, where T_a are the generators of the $SU(2)_L$ gauge group and g_1, g_2 are the coupling constants for the electroweak interaction.

$$\begin{aligned} W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g_2 \in^{abc} W^b_\mu W^c_\nu \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{2.8}$$

$$\begin{aligned} D_\mu L_L &= (\partial_\mu - ig_2 T_a W^a_\nu - ig_1 Y B_\mu)L_L \\ D_\mu e_R &= (\partial_\mu - ig_1 Y B_\mu)e_R \end{aligned} \tag{2.9}$$

Having introduced the scalar doublet Φ, an additional scalar part must be added to the Lagrangian of electroweak interaction, as given by Equation 2.10, where the first term is associated with the kinetic energy of scalar particle and the second term is the scalar potential, which is often called the ”Mexican Hat” potential.

$$\mathcal{L}_{Scalar} = (D^\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi) \tag{2.10}$$

The form of the scalar potential is not known from first principles and it is an assumption in the SM taking the simplest form possible, which has the desired properties of spontaneous symmetry breaking and the ability to be renormalized. The first term of $V(\Phi)$ represents mass of the particle and the second term is identified as self-interaction of the scalar field.

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \tag{2.11}$$

The value of λ must be positive in order for the vacuum to be stable. Depending on one of the cases of μ^2, the spontaneous symmetry breaking can occur (Thomson, M., 2013). In the case $\mu^2 > 0$, the $V(\Phi)$ potential is always positive, which produces a trivial minimum—no spontaneous symmetry breaking takes place.

$$\langle 0|\Phi|0\rangle \equiv \Phi_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.12}$$

In the case $\mu^2 < 0$, then the neutral component of the scalar field develops a non-zero vacuum expectation value (VEV), v, which breaks the symmetry. Then, Φ is expanded in terms of the v, where a new scalar field H(x), Higgs field, is introduced.

$$\langle 0|\Phi|0\rangle \equiv \Phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \ v = \sqrt{-\frac{\mu^2}{\lambda}} \tag{2.13}$$

After inserting this expanded kinetic part of $\mathcal{L}_{Scalar}$ and redefining the gauge fields, the $\mathcal{L}_{Scalar}$ is replaced with a new Lagrangian as given in Equation 2.14.

$$\mathcal{L}_{Broken} = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}g_2^2(v+H)^2 W_\mu^+ W^{\mu-} + \frac{1}{8}(v+H)^2(g_1^2+g_2^2)Z_\mu Z^\mu \tag{2.14}$$

From this Lagrangian, it is seen that mass terms for the W and Z bosons take the general forms $M_W^2 W_\mu W^\mu$ and $\frac{1}{2}M_Z^2 Z_\mu Z^\mu$ and photon A_μ remains massless. Thus the masses of the electroweak gauge bosons are

$$\begin{aligned} M_W &= \frac{1}{2}vg_2 \\ M_Z &= \frac{1}{2}v\sqrt{g_1^2+g_2^2} \\ M_A &= 0 \end{aligned} \tag{2.15}$$

The mass of the Higgs boson is another measurable quantity from the Higgs mechanism. The Lagrangian containing $H(x)$ term is given as

$$\mathcal{L}_H = \frac{1}{2}(\partial_\mu H)(\partial^\mu H) - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4 \tag{2.16}$$

Scalar masses have the general form $\frac{1}{2}m\phi$; the Higgs boson mass is thus given as in Equation 2.17. It needs to be determined experimentally since there is no other way to access the parameter λ.

$$m_H = 2\lambda v^2 = -2\mu^2 \tag{2.17}$$

Besides having massive bosons, the Higgs field can also provide the masses for all the fermions through Yukawa terms added to the Lagrangian. Therefore, mass of a fermion f can be written in terms of its Yukawa coupling as in Equation 2.18.

$$m_f = y_f \frac{v}{\sqrt{2}} \tag{2.18}$$

2.2. Shortcomings of the Standard Model

Despite being the most successful theory of particle physics to date, the SM does not provide an ultimate theory and a complete picture of the universe describing fundamental particles and all interactions between them. For this reason, the SM is inherently an incomplete theory with some questions about the SM itself as well as phenomena that it is unable to explain. All these unanswered questions provide that a large share of the published output by theoretical physicists proposing various forms of the BSM and many new experiments for finding answers for all those questions. Some of the important questions are listed below.

Gravity

One of the important question is how to include gravity in the SM since gravity is the most familiar force in our everyday lives. Although the SM describes the three fundamental forces at the subatomic scale, gravity is not included because it is absurdly weak and does not seem to have any impact on the subatomic interactions. However even if fitting gravity comfortably into the SM framework is hard, there are still some theories such as supergravity (Nastase, 2011) including a particle called the ”graviton” which might transmit gravity like the same way photons carrying the electromagnetic force.

Dark Matter and Dark Energy

The visible universe is made of protons, neutrons, and electrons. One of the interesting discoveries of the 20th century must be that this baryonic matter makes up less than 5% of the mass of the universe. This means that the SM leaves over 90% of the universe unexplained. The unexplained part of the universe appears to be made of invisible substance called ”dark matter” and ”dark energy” based on cosmological and astrophysical measurements. Based on the latest Planck results (Ade, 2016), it is about 26% of the energy density of the universe is covered by dark matter and there are very strong indications that the remaining 69% of the universe is dark energy. Even if there is remarkable evidence for dark matter based on the cosmological level, currently the SM does not contain any particle that could be a dark matter candidate. Although it does not interact with baryonic matter and it is completely invisible, scientists are still confident enough due to the gravitational effects appearing to have on galaxies and galaxy clusters.

Matter-Antimatter Asymmetry

Cosmology assumes that the universe was created in the so-called “Big Bang” from pure energy, and is now composed of about 5% baryonic matter, 26% dark matter and 69% dark energy. The Big Bang should have created equal amounts of matter and antimatter in the early universe. Following the 5% of baryonic matter observed in the visible universe, another question arises as to why is our universe made of matter and not antimatter. In other words, the visible universe is filled with matter, but contains very little amount of anti-matter. So, matter and antimatter is unbalanced and the SM can also not explain this ”matter-antimatter asymmetry”. As proposed by Sakharov (Sakharov, 1967) in 1967, there are necessary conditions for this asymmetry: (1)

Baryon number (B) violation, (2) Charge (C) and Charge-Parity (CP) violation, (3) Interactions out of thermal equilibrium. Regarding to these conditions, only the CP violation in rare meson decays is observed in the SM.

Neutrino Oscillations

One of the most abundant particles in the universe are neutrinos, which pass through most matter unnoticed and also pass harmlessly through human body. For many years, scientists have spent a lot of time exploring the properties of these elusive particles. One of the experimental observations is ”neutrino oscillations” which was observed in several experiments—the OPERA experiment (OPERA Collaboration, 2014), the KamLAND experiment (The KamLAND Collaboration, 2008) and the T2K experiment (T2K Collaboration, 2014). This observation implies that at least two of the three neutrinos must have mass. But, no measurements of the absolute masses have been made so far that neutrinos are massless particles and the SM does not predict neutrino mass. Neutrino masses can be included in the SM (Klinkhamer, 2013), but still there is a question whether they are normal Dirac fermions, or Majorana fermions.

Hierarchy Problem

There are different energy scales in the universe. The SM works well at the magnitude of the electroweak scale (100 GeV). However, there is an upper energy scale in the high energy physics at the magnitude of Planck scale ($\sim 10^{19}$ GeV), where quantum gravitional effects become important—beyond which the SM is no longer valid. While the experimental frontier has been reaching out the TeV scale, another unanswered question of the SM is why the electroweak scale is so much smaller than the Planck scale May be this answer can be evidence for physics BSM (Virdee, 2016). This large difference between electroweak and Planck scales is about $\sim 10^{17}$ GeV and this can have impact when calculating the quantum corrections for different quantities such as mass for scalar particles. With the experimental observation of the Higgs mass at 125 GeV, there is a disparity of expected and observed Higgs masses which is called as ”the Hierarchy problem” and it implies some amount of ”fine-tuning” (Djouadi, 2008) in terms of Higgs vacuum expectation value (v) and cut-off energy scale (Λ).

Observed particle masses are a combination of the ”bare” mass at tree level and the corrections from loop diagrams. The Higgs mass m_H can be written as in terms of the bare mass parameter and the radiative corrections (Biswas, Kundu, and Mondal, 2020).

$$m_H^2 = m_{Bare}^2 + \Delta m_H^2 \tag{2.19}$$

The hierarchy problem can be seen due to the lack of a symmetry, which protects the Higgs mass against to very high energy scales. The chiral symmetry protects against large radiative corrections to fermion masses while the local gauge symmetry protects the gauge boson masses. In the case of Higgs mass, there is no such a symmetry and the problem comes from loop corrections to the Higgs mass due to a quadratic dependence from the correction term. The one loop corrections to the Higgs mass are shown in Figure 2.3., arising from the coupling of the Higgs boson to the fermions, the coupling of the Higgs boson to the vector bosons, the coupling of the Higgs boson to itself.

The one-loop contributions to the Higgs mass are considered as in Equation 2.20, where λ is the quartic coupling of the scalar potential for the Higgs field, y_f are the Yukawa couplings, g_1 and g_2 are the weak gauge boson couplings (Grojean, 2015).

$$\Delta m_H^2 = \left[6\lambda - 6y_f^2 + \frac{1}{4}(9g_1^2 + 3g_2^2)\right] \tag{2.20}$$

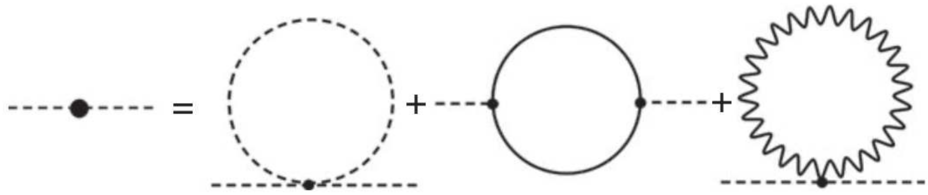


Figure 2.3. The one-loop quantum corrections to the Higgs mass, arising from the coupling of the Higgs boson to itself, the coupling of the Higgs boson to the fermions, the coupling of the Higgs boson to the vector bosons (left to right) (Grojean, 2015)

As a consequence, to end up with the Higgs boson at the mass of electroweak scale, some cancellations are needed for the balance between bare mass and large quantum corrections. This cancellations are highly unnatural. Regarding to the Equation 2.20, there is the relative minus sign between fermion loop and boson loop contributions to Δm_H^2. Thus, one way to solve the hierarchy problem is that a new symmetry ought to relate fermions and bosons, which is called “Supersymmetry”.

2.3. Supersymmetry (SUSY)

SUSY is a popular extension of the SM that posits a space-time symmetry relating fermions with bosons.Under this symmetry, each particle in the SM has a “superpartner”, whose spin differs by $\frac{1}{2}$, but whose remaining quantum numbers are the same. In other words, every fermion would have a bosonic superpartner and every boson would have a fermionic superpartner (Martin, 1998).

Based on SUSY algebra, the SM particles and their superpartners constitute “supermultiplets” where both have the same mass. If there were experimental evidence to observe that SM particles have exactly the same mass as their superpartners, then SUSY would be an exact symmetry and this could solved the hierarchy problem. But, there is no experimental evidence yet from the LHC to observe the superpartners. This indicates that SUSY must be a broken symmetry or superpartners could be discovered at high energies. However, if the superpartners have somewhat similar masses to their SM partners, there will still be a large cancellation and there can be balance between the bare and observed masses of the Higgs.

A Minimal Supersymmetric Standard Model (MSSM), where the superpartner masses lie below a few TeV (Workman, 2022), is introduced as the simplest possible SUSY extension of the SM with the minimal particle content. Each SM particle and its superpartner constitute a chiral or gauge supermultiplet. The leptons and their superpartners "sleptons" constitute chiral supermultiplets, because left- and right-handed fermions transform differently under the electroweak gauge symmetry. The quarks and their superpartners "squarks" also constitute chiral supermultiplets. Note that sleptons and squarks refer to "scalar leptons" and "scalar quarks", respectively. The Higgs boson is spin-0 and it must reside in a chiral supermultiplet. Actually it turns out that SUSY needs two Higgs doublets to ensure anomaly cancellation. All the gauge bosons then have superpartners "gauginos", including higgsinos, winos, a bino and a gluino.

The supermultiplets are shown in Figure 2.4.. There can be mixing between left- and right-chiral sfermion states, which give rise to the physical mass eigenstates. But, assumptions for MSSM are made such that only third generation sfermions have non-trival mixing. In addition, there is no mixing for sneutrinos, because they have only left-chiral states. Thus, mass states of the first and second generation squarks are denoted with subscript "L,R", whereas mass states of the third generation squark are denoted with subscript "1,2". After electroweak symmetry breaking, all gauginos excepting the gluino mix together to form "neutralinos" such as $\tilde{\chi}_1^0$ and "charginos" such as $\tilde{\chi}_1^{\pm}$.

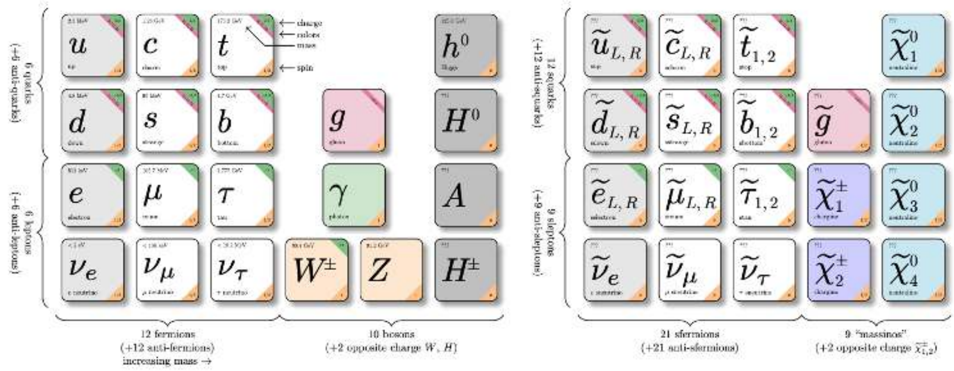


Figure 2.4. The SM particles and their superpartners within a MSSM. Each fermion is transformed into bosonic superpartner and vice versa. Superpartners differ in spin by half unit, but the remaining quantum numbers are the same. If SUSY is not broken, then each SM particle and its superpartner have also the same mass

The MSSM superpotential is given as in Equation 2.21, where Q is a squark doublet; U, D are squark singlets; L are slepton doublets; E is slepton singlet; H are Higgs doublets and i, j, k are generation indices; λ and μ are uniquie coupling strengths for each type of interactions. This shows coupling between the different gauge fields, which manifests as allowed interactions between the SM and supersymmetric particles. In the second line of the Equation 2.21, all given terms correspond to possible particle interactions that violate baryon and/or lepton number. Such interactions give the proton a way of decaying quickly.As proton decay has not been observed experimentally, “R-parity” is introduced as an additional symmetry to supress all those terms.

$$\begin{aligned} W = {} & \lambda^u_{ij} Q_i U_j H_u + \lambda^d_{ij} Q_i D_j H_d + \lambda^e_{ij} L_i E^c_j H_d + \mu H_u H_d \\ & + \frac{1}{2}\lambda_{ijk} L_i L_j E^c_k + \lambda'_{ijk} L_i Q_j D^c_k + \frac{1}{2}\lambda''_{ijk} U^c_i D^c_j D^c_k + \mu_i L_i H_u \end{aligned} \tag{2.21}$$

R-parity is an additional quantum number given in Equation 2.22, where B is baryon number, L is lepton number and s is spin. The SM particles carry an R-parity of +1 while SUSY particles have $R = -1$. Any interactions between the SM particles and SUSY particles must conserve R-parity. In other words, R-parity conservation forces the baryon and lepton number violating couplings to be zero, then the proton is stable.

$$R = (-1)^{3B+L+2s} \tag{2.22}$$

Introducing the R-parity has important consequences for the phenomenology of the MSSM. Assuming the R-parity conservation, SUSY particles can only be produced in pairs at particle colliders which collide the SM particles. Likewise, a SUSY particle would always decay into an odd number of other SUSY particles. The lightest supersymmetric particle (LSP) must therefore be stable and if it is electrically neutral then it would be an excellent dark matter candidate.

None of SUSY particles have been observed yet at the LHC to prove that the MSSM exists or that R-parity is conserved. Also, this means there is no evidence for why SUSY is broken symmetry. So, perhaps SUSY is not as easily observable, and is described by more complicated models. The next two sections present two such models, which are R-Parity Violating (RPV) and Stealth (SY$\bar{Y}$) SUSY, if there are some possible ways to find evidences for all those above.

At the LHC, the cross section for production of SUSY particles expected are shown in Figure 2.5.. In general, SUSY groups have looked for the higher cross section options such as top squark production (CMS Collaboration, 2021). In this thesis, top squark (stop) pair production is presented with the RPV and Stealth SY$\bar{Y}$ SUSY models, further details in Section 4..

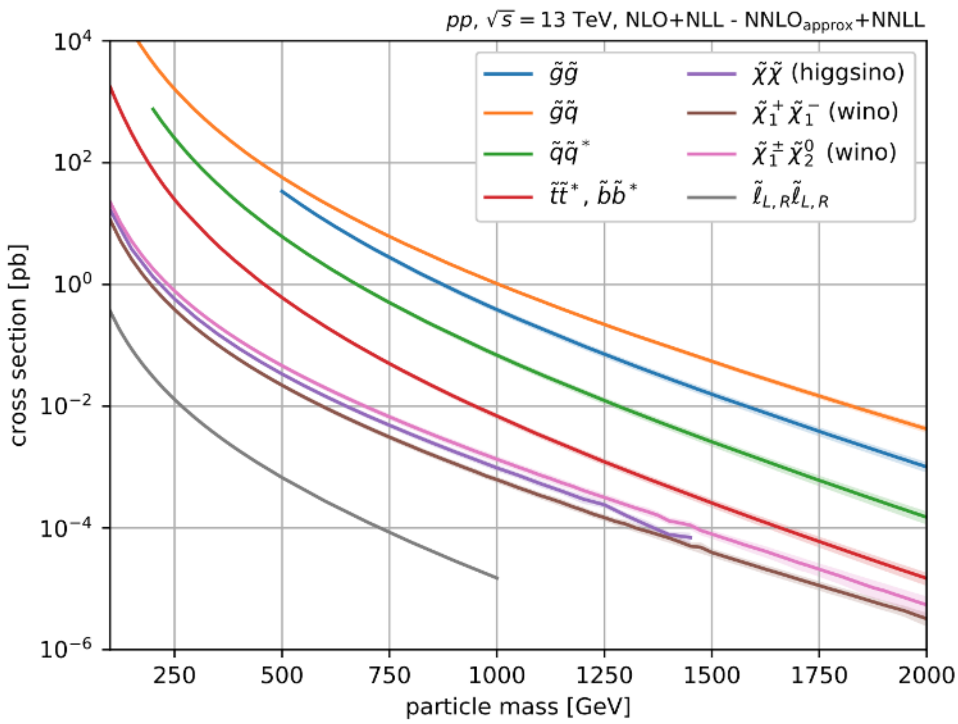


Figure 2.5. At the LHC, SUSY cross sections as a function of particle mass. The red curve corresponds to the top squark pair production at 13 TeV (CMS Collaboration, 2021)

2.3.1. R-Parity Violating (RPV) SUSY

The R-Parity Violating (RPV) SUSY (Evans and Kats, 2013), where R-Parity is not conserved, is introduced as one can consider adding back those problematic terms in the superpotential. Those terms are given in Equation 2.23. To keep the proton stable, only one of the terms allowing lepton or baryon violation dominates the interaction, whereas the other terms are supressed. This then allows RPV phenomenology of the MSSM without obvious proton decay problems.

$$W_{\text{RPV}} = \frac{1}{2}\lambda_{ijk}L_iL_jE^c_k + \lambda'_{ijk}L_iQ_jD^c_k + \frac{1}{2}\lambda''_{ijk}U^c_iD^c_jD^c_k + \mu_iL_iH_u \tag{2.23}$$

In Equation 2.23, the first and second terms are “LLE-coupling” and “LQD-coupling" allowing lepton number violation, whereas the third term is "UDD-coupling" allowing baryon number violation. All these couplings give rise to a SUSY particle decays through the SM particles. These couplings are shown with the λ coupling constants in Figure 2.6..

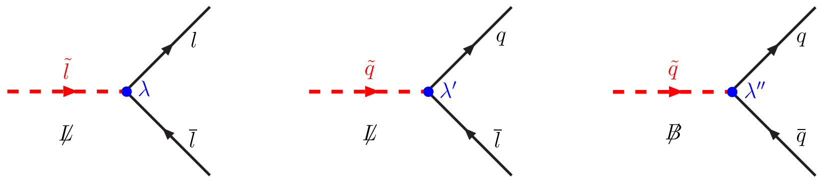


Figure 2.6. The "LLE-coupling", "LQD-coupling" and "UDD-coupling" for R-parity violating interactions in the MSSM. λ is coupling strength for LLE, λ' is coupling strength for LQD, λ'' is coupling strength for UDD (Barbier et al., 2005)

In contrast, most of the SUSY searches at the LHC look for an excess of events with large-E_T^{miss} originating from the undetected LSPs, the RPV model produces low-p_T^{miss} signatures by providing a mechanism for the LSP to decay. Lepton or baryon number violation allows an additional interaction via the related couplings. Then, the LSP (typically $\tilde{\chi}_1^0$) is not stable and decays to purely SM particles.

2.3.2. Stealth SUSY

Stealth SUSY models are extensions to the MSSM, which would make detecting SUSY more difficult and explaining the current absence of evidence. These Stealth SUSY models posit a new hidden sector called as "stealth sector" of light particles with small or absent couplings to the SUSY breaking sector and finite couplings to the visible sector.

In this sector, SUSY is approximately conserved symmetry due to the weak connection to the SUSY breaking sector, which results in stealth particles that are close-in-mass with their superpartners. A collection of "portals", which includes mediation by Higgs boson or vector-like matter as messenger fields, allows production and decay of stealth particles via interections with the MSSM.

The Stealth SY$\bar{Y}$ model (Fan, Krall, et al., 2016) assumes a stealth sector, which contains an supersymmetric fermion singlino ($\tilde{S}$) and its scalar partner singlet (S). Additionally, there is an assumed vector portal which implies the existence of vector-like messenger fields $Y - \bar{Y}$. The scalar particle singlet (S) decays via the $Y - \bar{Y}$ vector-like messenger fields in the "stealth sector" to the SM particles as shown in Figure 2.7.. The Stealth SY$\bar{Y}$ sector superpotential then is given in Equation 2.24, where S is the chiral superfield containing singlet and singlino; Y and $\bar{Y}$ are two additional chiral supermultiplets added as vector-like messenger fields; m_Y and m are supersymmetric masses with $m_Y \sim$ TeV and $m \sim 100$ GeV.

$$W = \frac{m^2}{2} S^2 + \lambda S Y \bar{Y} + m_Y Y \bar{Y} \tag{2.24}$$

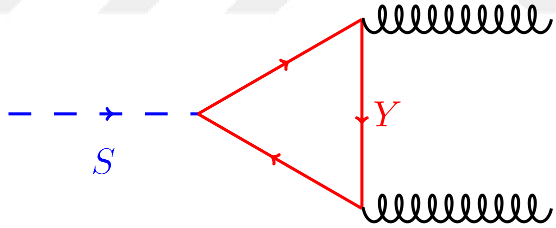


Figure 2.7. The scalar particle Singlet (S) decays via $Y - \bar{Y}$ vector-like messenger fields in the "stealth sector" to the SM particles (Fan, Reece, and Ruderman, 2012)

In this model, $\tilde{S}$ decays through the hidden sector to a close-in-mass S and a Gravitino ($\tilde{G}$). The $\tilde{G}$ is the LSP and it is very light—it is usually assumed to be around 1 GeV. Also, the mass difference between $\tilde{S}$ and S is naturally very small in this model. So, the Stealth SY$\bar{Y}$ model produces low-p_{T}^{miss} signatures.

3. THE LHC AND CMS EXPERIMENT

Scientists from all around the world have been working on lots of research to understand how the universe exists, what the fundamental particles are. Especially during this century, there are lots of laboratories to search for new physics and develop new technology. Not only new particles have been discovered, but also these developed technologies have been used in different areas such as medicine, chemistry, etc.

The European Organization for Nuclear Research (CERN) is the largest research laboratory in the world which has collaborators all around the world and also offers a perfect opportunity for the future's scientists. Besides, CERN hosts an accelerator complex which consists of different types of accelerators. This complex allows the scientists to explore on different resereaches. The Large Hadron Collider (LHC), is the last chain of the CERN accelerator complex and opens a new frontier in particle physics. After several achievements for its performance, the LHC allows us to significantly increase the discovery reach for SUSY, extra dimensions of space and time, and physics BSM with Run3 data taking.

The details of the CERN ccelerator complex are given in Section 3.1. and Section 3.2. presents what is an "event" in the proton-proton (pp) collisions at the LHC. The main analysis in this thesis uses the collected Run2 data from pp collisions at the LHC and measured in the CMS detector. An overview of LHC is presented in Section 3.3. and the CMS detector is detailed in Section 3.4.. Then, the trigger system is described in Section 3.5.. Lastly, object definitions in the CMS detector are described in Section 3.6., which includes the objects used in this analysis.

3.1. CERN Accelerator Complex

The CERN accelerator complex consists of a series of machines to accelerate particles to increasingly higher energies. Each accelerator boosts the energy of a beam of particles before injecting it into the next machine in the queue. The LHC is the last chain of the accelerator complex. Proton bunches are not simply injected into the LHC itself—without passing through each machine in the queue.

The journey of particles starts with a bottle of hydrogen gas by stripping hydrogen atoms to electrons for generating the protons. These protons are accelerated by a linear accelerator, Linear Accelerator 2 (LINAC2), to the energy of 50 MeV. Then the beam is injected in the Proton Synchrotron Booster (PSB), where they are accelerated to 1.4 GeV. From there, the protons move to larger accelerators which are the Proton Synchrotron (PS) and the Super Proton Synchroton (SPS). The PS reaches the energies of 25 GeV and starts ”bunching” the protons together. After then, proton bunches are passed into the SPS, which accelerates the proton bunches to about 450 GeV. Finally, the proton bunches are transferred to the last chain of the accelerator complex, the LHC, where each proton bunch is accelerated to 6.5 TeV of energy while circulating around the ring many times. The figure 3.1. shows this accelerator complex. When these proton bunches collide, it is called a ”bunch crossing”. There is 25 ns between each bunch crossing which implies the collision frequency of 40 MHz.

In addition to accelerating of protons, the accelerator complex is also used for accelerating lead ions produced from a highly purified lead sample.

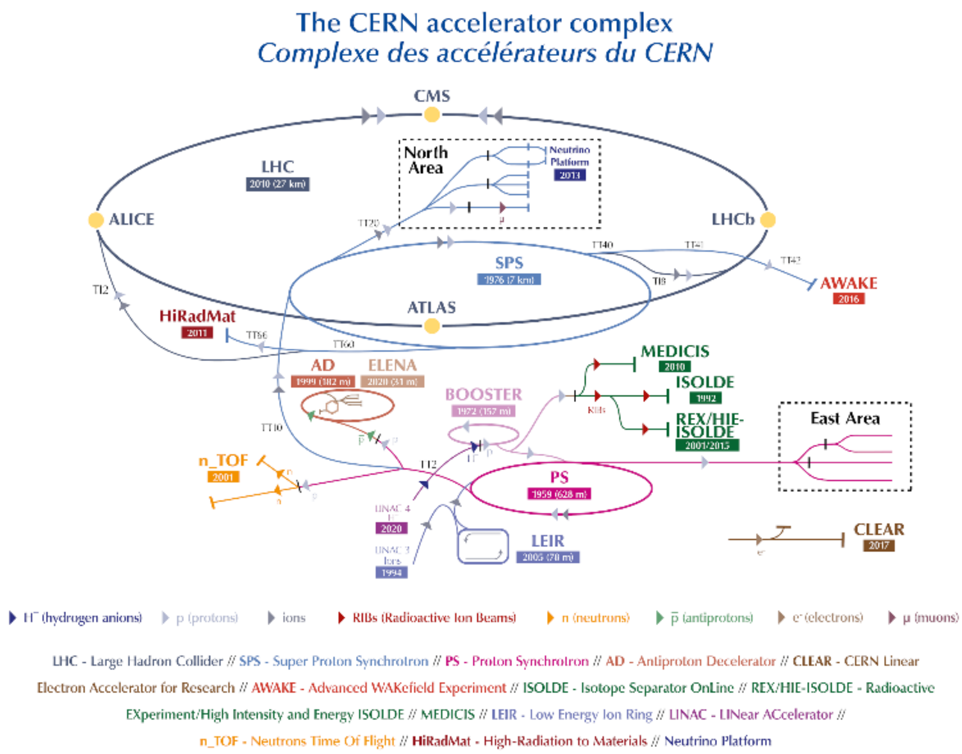


Figure 3.1. The whole CERN accelerator complex, including the LHC. The four main experiments, ALICE, ATLAS, CMS, LHCb, are also shown (Lopienska, 2022)

3.2. Proton-Proton Collisions at the LHC

At the LHC, pp collisions provide lots of information about the secrets of particle physics. The cross section is used to describe the probability when two particles collide and interact in a certain way. In each bunch crossing, many different processes can occur. The cross section of a particular process depends on the type and energy of the colliding particles. For instance, W boson production has large cross section, so it is observed more often than the Higgs boson production. Since Higgs boson production has a much lower cross section and it is more difficult to produce.

An ”event” is defined as everything that happens in a proton bunch crossing. But, only a fraction of the events have something interesting—rest of the events produced by the LHC are never saved for processing later. For any particular physics process in the LHC, the number of events per second for this process can be calculated as in Equation 3.1, where $\mathcal{L}$ is the instantaneous luminosity and σ_{events} is the cross section for a particular process.

$$\frac{dN_{event}}{dt} = \mathcal{L}\sigma_{events} \tag{3.1}$$

The LHC is designed to maximize the number of interesting events by integrating the instantaneous luminosity ($\mathcal{L}$), which measures how tightly particles are packed into a proton beam. It is defined as in Equation 3.2, where N_p is the number of particles per bunch, n_b is the number of bunches per proton beam, f_{rev} is revolution frequency, γ_r is relativistic gamma factor, ε_n is the normalized transverse proton beam emittance, β^* is the beta function at the interaction point, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point.

$$\mathcal{L} = \frac{N_p^2 n_b f_{rev} \gamma_r}{4\pi\varepsilon_n\beta^*}F \tag{3.2}$$

The total integrated luminosity, which considers the total number of events during a period of data-taking, can be calculated as in Equation 3.3. During the Run2 data taking at the LHC, an energy of 13 TeV was achieved with a total bunch crossing rate of 40 MHz, and a peak luminosity of $\mathcal{L} = 10^{34} cm^{-2} s^{-1}$. The figure 3.2. shows the integrated luminosity delivered at the LHC and recorded by the CMS during the Run2 data taking.

$$L = \int \mathcal{L} dt \tag{3.3}$$

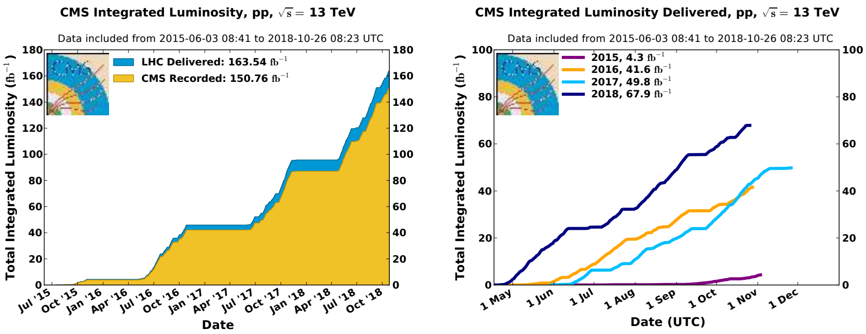


Figure 3.2. Cumulative delivered and recorded luminosity versus time for 2015-2018 for pp data (left). CMS integrated luminosity delivered, pp collison at $\sqrt{s} = 13$ TeV (right) ("Public CMS Luminosity Information" 2022)

3.3. Large Hadron Collider (LHC)

The LHC is world's largest and highest-energy particle accelerator and collider, which reaches energies for its center-of-mass collisions at around 14 TeV and provides around 600 million pp collisions per second with a luminosity ($\mathcal{L}$) on the order of $10^{34} cm^{-2} s^{-1}$. It is located on the border of Switzerland and France. As the last chain of the accelerator complex, the LHC consists of a 26.7 km ring of superconducting magnets with a number of accelerating structures to boost the particles. It is constructed to collide two proton bunches head-to-head around the ring by crossing at four different points, which contain a separate detector for each: A Large Ion Collider Experiment (ALICE), A Toroidal LHC ApparatuS (ATLAS), Compact Muon Solenoid (CMS) and LHC-beauty (LHCb):

- **ALICE:** It is designed to characterize the physical properties of the Quark-Gluon Plasma (QGP), a state of matter created under the extreme conditions of temperature and energy density created in nuclear collisions which head-on collisions between massive ions, such as gold or lead nuclei.
- **ATLAS:** It is one of the general purpose detectors at the LHC searching for SM, Higgs physics and all possible new particles. Also, it is the largest volume

particle detector ever designed.

- **CMS:** It is another general purpose detector like ATLAS with a similar physics programme. The huge solenoid magnet, operating at a magnetic field strength of 3.8 T, is its main feature. More details will be given in this chapter.
- **LHCb:** It is a more specific detector to search for matter-anti matter symmetry by observing the B meson decays. Instead of the cylindrical structure as CMS and ATLAS detectors, it uses a series of subdetectors to detect mainly forward particles – those thrown forwards by the collision in one direction.

3.4. CMS Detector

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors at the LHC and performs a wide range of tasks from measurements of SM physics processes to searches for physics BSM. In order to satisfy the goals of the LHC physics program, the CMS detector meets certain detector requirements which are (1) Good muon identification and momentum resolution over a wide range of momentum, (2) Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker, (3) Good electromagnetic energy resolution, (4) Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters.

It yields head-on collisions of two proton (ion) beams of 7 TeV (2.75 TeV per nucleon) each, with a design luminosity of $10^{34} cm^{-2}s^{-1}$ ($10^{27} cm^{-2}s^{-1}$). The CMS detector is built around a large solenoid magnet, which operates at a magnetic field strength of 3.8 T. The entire detector is placed in a cave about 100 m below the ground and it is 28.7 m long, 15 m high, 15 m wide and weighs 14000 tons (S. Chatrchyan et al., 2008).

The CMS consists of a series of sub-detectors and a superconducting solenoid magnet. Each sub-detector is designed to observe unique information about the particles while the solenoid magnet provides a large bending force to charged particles for measuring their momentum. The bore of the magnet coil is large enough to accommodate the tracker system, electromagnetic and hadronic calorimeters. The muon detector is installed in between the return yoke layers and surrounds the magnet, as illustrated in Figure 3.3.. The tracker system is the innermost one and closest to the beam pipe to measure the charged particle's momentum. The tracker is surrounded by the Electromagnetic Calorimeter (ECAL), which measures the energies of electromagnetically interacting particles. The last system inside the bore of magnet coil is the Hadronic Calorimeter (HCAL), which measures the energy of hadrons made of quarks and gluons. In addition to these sub-systems, there is also a system to measure the luminosity during the data taking at the CMS, which is called the Beam Radiation Instrumentation and Luminosity (BRIL).

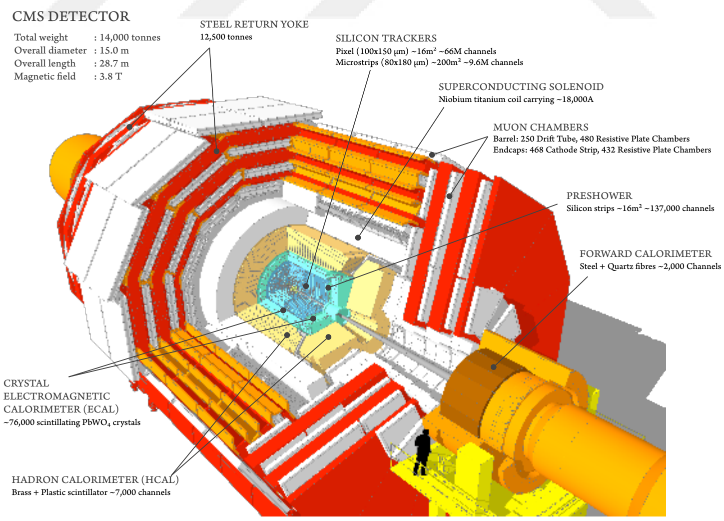


Figure 3.3. Overview of the CMS detector with the sub-dectetors is indicated on the figure. Figure taken from (Sakuma, 2016)

Before describing the sub-detectors of the CMS, the CMS coordinate system is mentioned here. The CMS uses a right-handed coordinate system, shown in Figure 3.4., to define the detector geometry and it is centered at the nominal pp interaction point (IP). The x-axis radially points to the center of the LHC. The y-axis points vertically upward. The z-axis is set along the counter-clockwise beam direction.

Instead of the x-y-z coordinates, two angles are commonly used to describe particle track and momentum positions. First, the azimuthal angle ϕ in the transverse (x-y) plane is measured originating from x-axis and defined as in Equation 3.4. The momentum and energy in the transverse plane is very important to provide some information about the particles. Many quantities are defined in this plane such as the transverse momentum (p_{T}), the transverse energy (E_{T}), and scalar sum of the transverse momentum (H_{T}).

$$tan\phi = \frac{y}{x} \tag{3.4}$$

The another angle, the polar angle θ is measured from the z-axis and it can be mapped to a widely used variable known as the pseudorapidity (η), defined as in Equation 3.5.

$$\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right] \tag{3.5}$$

Also, we can define useful variables in terms of η such as ΔR which is the angular separation between two particles, defined as in Equation 3.6.

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.6}$$

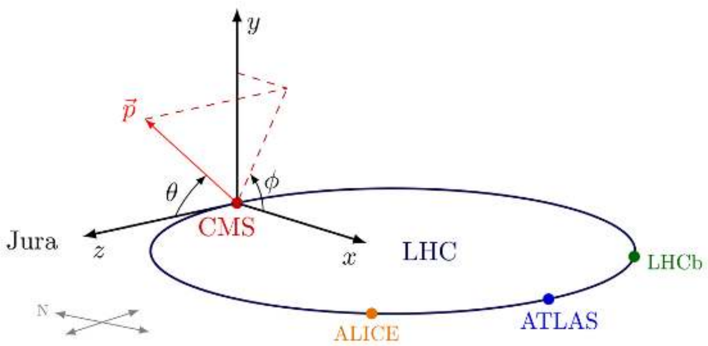


Figure 3.4. The CMS coordinate system, including the LHC and a compass (the z axis points to the Jura). Figure taken from (Neutelings, 2017)

3.4.1. Tracker System

The tracker system is mainly designed to measure precisely the momentum of charged particle tracks with energies as low as 1 GeV up to pseudorapidities $|\eta| < 2.5$ (Karimäki et al., 1997). In addition to its main property, its other roles are to determine the primary vertex and aid in the reconstruction of secondary vertices for heavy-flavor signatures. It is located close to the interaction point for making the most accurate measurements and made with minimal material in order to avoid secondary interactions, such as multiple scattering, bremsstrahlung, photon conversion, and other nuclear interactions.

Due to its location, the tracker system is exposed to large particle fluxes, which can damage its electronics. For this very reason, it is made up of many thin layers of silicon, which are radiation-hard materials. It is the largest detector of its kind with over 200 m^2 of active silicon. Moreover, in order to reduce radiation effects and extend the lifetime of the detector modules, the tracking detectors are designed to run at subzero temperatures, which includes a special cooling system to keep temperature around -20°C.

The CMS tracker system consists of two sub-detectors with independent cooling, powering, and read-out systems: the pixel detector and the silicon strip detector (Veszpremi, 2014). The layout of the tracker system is shown in Figure 3.5.. The pixel detector has a total of 1440 modules whereas the silicon strip detector has a total of 15,148 modules. Each module has its own set of front-end electronics to record the particles passing through the tracker system. The tracker has about 75 million read-out channels, providing measurement of position of a particle within 10 μm and measurement of the p_T of a particle with a 1-2% resolution for a particle with 100 GeV momentum.

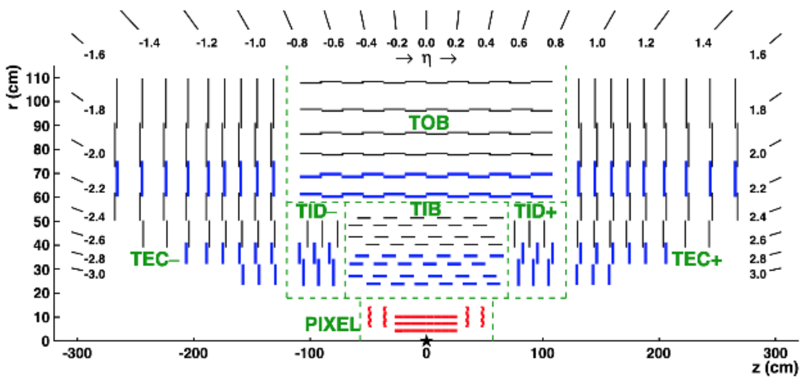


Figure 3.5. Schematic cross section through the CMS tracker system. In the center is the pixel detector, divided into the barrel and endcap regions, and right on the outside is the silicon strip detector system. Figure taken from (T. C. Collaboration, 2014)

The pixel detector is originally comprised of three cylindrical layers in the barrel region (BPIX) at radii of 4.3 cm, 7.2 cm, and 11 cm, respectively. On each endcap, there are two silicon disks at 34.5 cm and 46.5 cm from the interaction point and the combination of the four disks is called FPIX. It has a cooling system to keep the temperature from rising too much above the operating temperature of -10°C. As a part of the Phase-1 upgrades (2016-2017), there were upgrades in both BPIX and FPIX. An additional layer to the BPIX was added and innermost layer becomes closer to the beam pipe. For the FPIX, an extra ring was added to each endcap and each ring is divided into two half rings, as shown in Figure 3.6..

The silicon strip detector surrounds the pixel detector. The detector has 10 tracking layers in the barrel region, which are 4 inner barrel layers (TIB) with 2 inner endcaps (TID) and 6 outer barrel layers (TOB) with two endcaps (TEC) closing off the tracker. The strip detector has also a cooling system to keep the temperature around -20°C.

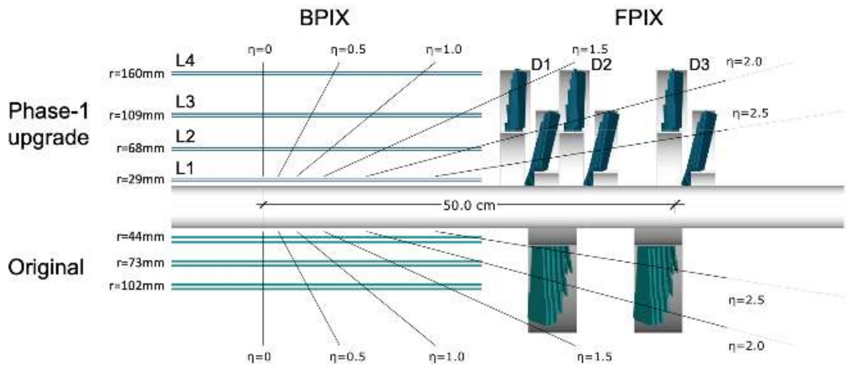


Figure 3.6. A cross sectional view of the pixel detector where top is the new Phase 1 pixel detector and the bottom is the legacy pixel detector. Diagram taken from (The Tracker Group Of The CMS Collaboration, 2020)

3.4.2. Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is designed to measure the energy and position of any electromagnetically interacting particles such as electrons and photons. The main component of the ECAL, lead tungstate ($PbWO_4$) crystals, provides a very good energy resolution, which played an important role in the discovery of Higgs boson during the data taking of Run1 (S. e. a. Chatrchyan, 2012). Moreover, the ECAL plays an important role in discovering a wide range of SM and other new physics processes.

The ECAL is a hermetic, homogeneous, fine grained lead tungstate ($PbWO_4$) crystal calorimeter and located in between the tracker and the HCAL. It consists of a barrel (EB) with pseudorapidity coverage up to $|\eta| = 1.48$, where it is closed by two endcaps (EE) with the coverage of up to $|\eta| = 3.0$, and a preshower detector (ES) is placed in front of the endcaps at $1.65 < |\eta| < 2.6$. The schematic overview of the ECAL is shown in Figure 3.7..

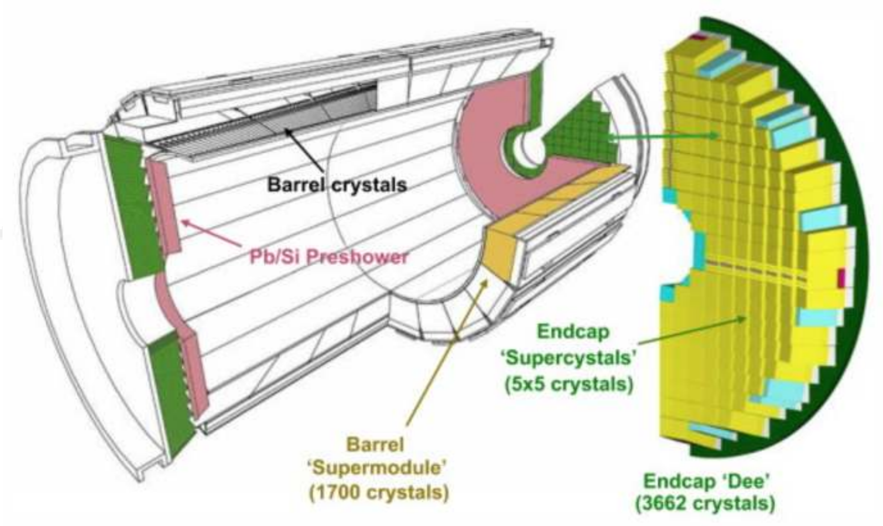


Figure 3.7. Schematic overview of the ECAL, including the barrel, the endcap and the preshower (Siddireddy, 2018)

A total of 75,848 lead tungstate ($PbWO_4$) crystals are arranged in the ECAL, where 61,200 of them are in the barrel and 14,648 of them in the endcaps. The main features of these scintillating crystals are high density, extremely short radiation length and small Moliére radius, which are providing the realization of a homogeneous compact calorimeter with high granularity. Also, the crystals produce fast signals, where 80% of the light is emitted in 25 ns, which is the time between two bunch crossings at the LHC. All these properties allow the ECAL to offer the best performance for energy resolution (Biino, 2015).

When electrons or photons pass through the ECAL, they interact with the $PbWO_4$ crystals. Secondary particles coming from each interaction, which are electron bremsstrahlung and photon conversions, are collected as electromagnetic shower and results in the production of scintillation light, which is used for inferring energy deposited. Then, the light from each crystal is read out by photo-detectors, which are avalanche photodiodes (APDs) in the EB and vacuum phototriodes (VPTs) in the EE. Figure 3.8. shows an image of a $PbWO_4$ crystal with an electromagnetic shower shown inside. The light yield of the crystals and the gain of the photodetectors strongly depend on temperature, therefore ECAL is operated at a controlled temperature, which is maintained by a dedicated water cooling system and temperature monitoring sensors.

The preshower detector is a sampling calorimeter, consisting of two planes of lead radiators followed by detector planes of silicon strips to measure the energy and transverse profile of showers as well as to distinguish between single high-energy photons and low-energy photons coming from the decay of neutral pions.

The energy resolution for the ECAL, given in Equation 3.7, can be characterized as a function of the energy in terms of a few experimentally measured constants, where S is the stochastic term from photon statistics, N is the electronic noise, and C is a constant offset.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{3.7}$$

Figure 3.8. Image of a lead tungstate ($PbWO_4$) crystal with an electromagnetic shower shown inside. Figure taken from (Wadud, 2019)

3.4.3. Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is the outermost detector layer within the CMS magnet (Mans et al., 2012) and is designed to measure the energy of hadrons made of quarks and gluons. Additionally, the HCAL provides indirect measurement of the presence of non-interacting, uncharged particles such as neutrinos. For this reason the HCAL is built to be as hermetic as possible to capture every particle emerging from the collisions. This hermeticity allows for a measurement of missing transverse energy, which allows for the observation of very weakly interacting particles.

The HCAL is a sampling calorimeter, which uses alternating layers of brass absorber and fluorescent scintillator materials as detector active materials. Hadrons interact with the absorbers and this interaction produces a cascade of new particles referred to as a hadronic shower. The measurements of hadronic showers help to deduce particles' positions, energies and arrival times. The charged particles in a hadronic shower interact with the scintillators and this interaction rapidly produces

”a light pulse”. This light pulse passes through photo-detectors, Front-End (FE) and Back-End (BE) electronics and it is processed by photo-detectors, Charge Integrated and Encoding (QIE) cards and micro HCAL Trigger Readout (μHTR) cards with the supports of other components of the FE and BE electronics to produce triggered data. Last, this digital data is transfered to Data Acquisition (DAQ) system for event reconstruction.

- **Photo Detectors:** A special optical fiber collects the light pulse and feeds it into the readout boxes where a photo-detector amplifies the signal. The signal pulse then is converted to the electrical signal by the photo-detector and this electrical signal is transported towards to the QIE card. The HCAL currently uses two types of photo-detectors which are Silicon Photo Multipliers (SiPMs) and 4-anode Photo Multiplier Tubes (PMTs).
- **QIE Cards:** They are main components of the FE electronics. These cards include QIE chips, which consist of analog-to-digital converters (ADCs) for digitizing the pulse heights and time-to-digital converters (TDCs) for measuring the arrival time of the pulse. When this electrical signal is arrived to the QIE card, the QIE chip integrates and digitizes it. Then this dizitized data is transported to the μHTR card via the optical fibers. The HCAL currently uses two types of QIE chips which are QIE10 and QIE11 Application-Specific Integrated Circuit (ASIC) chips. For each sub-detector, there are different amount of the QIE chips and these chips refer to readout channels.
- **μHTR Cards:** They are main components of the BE electronics. These cards receive the digital data from the QIE cards and trigger this digital data. Then, the triggered data is transported to the DAQ system for event reconstruction. The HCAL uses the same μHTR cards for all sub-detectors, but each sub-detector has different amount of them based on the number of readout channels.

The HCAL consists of four sub-detectors which are the HCAL Barrel (HB), Encap (HE), Outher (HO) and Forward (HF), shown in Figure 3.9.. Each sub detector has own photo-detectors, front-end (FE) and back-end (BE) electronics. Currently the HB, HE, HO use the SiPMs whereas the HF uses the 4-anode PMTs. Similarly the HB and HE use the QIE11 chips whereas the HF uses QIE10 chips and HO still uses the phase-0 QIE chips.

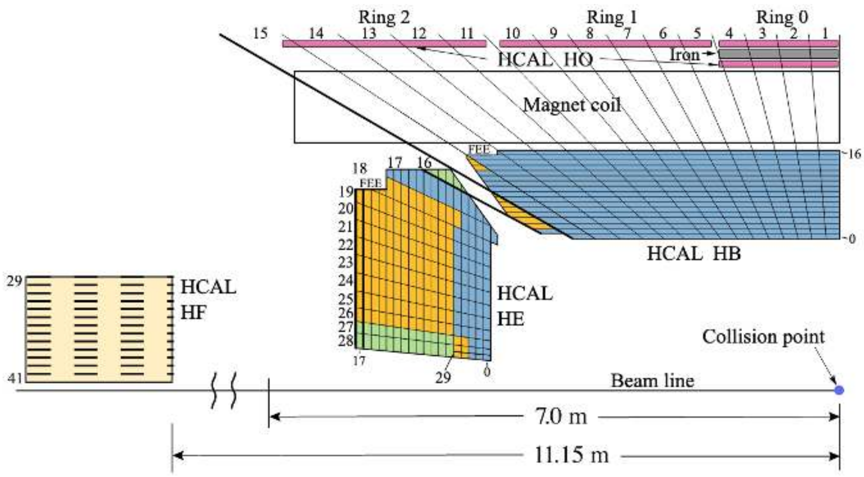


Figure 3.9. A schematic view of one quarter of the CMS HCAL during 2016 LHC operation, showing the positions of HB, HE, HO and HF calorimeters. The layers marked in blue are grouped together as depth = 1, while the ones in yellow, green, and magenta are combined as depths 2, 3, and 4, respectively (Sirunyan, 2020)

The HB and HE are sampling calorimeters, using brass absorbers and plastic scintillator materials. The light coming from the scintillators is transported to SiPMs via Wave-Length Shifting (WLS) fibers. The HB covers the pseudorapidity ranges $|\eta| < 1.3$. The HE covers an important portion of the rapidity range, $1.3 < |\eta| < 3.0$, which captures about 34% of the particles produced in the final state. As a central part of the HCAL, the HB and HE provide excellent jet identification, moderate single particle, and jet resolution.

The HO is placed outside of the solenoid with pseudorapidity coverage $|\eta| < 1.3$. Outside of the solenoid, the magnetic field is returned through an iron yoke, which is designed in the form of five 2.536 m wide (along z-axis) rings. The HO is placed as the first sensitive layer in each of these five rings. It uses the same active materials and WLS fibers like HB and HE, but it uses the steel return yoke and magnet material as an absorber. In the central pseudorapidity region, HB does not provide sufficient containment for hadron showers and the HCAL is extended outside the solenoid with HO acting as a ”tail catcher”. It is an additional absorber and its main goal is to identify late starting showers and to measure the shower energy deposited after HB. Also, it is useful for muon identification.

The HF consists of two modules, HF+ and HF- located symmetrically on each side at about 11 m from the interaction point along the beam axis. It is a Cherenkov calorimeter, consisting of a steel absorber and quartz fibers. These fibers also transport the light pulse to the 4-anode PMTs. There are long (1.65 m) and short (1.43 m) quartz fibers, which help to discriminate between electromagnetic particles and hadronic particles. These quartz fibers are radiation hard because of the high flux of charged particles in its pseudorapidity coverage of $3.0 < |\eta| < 5.0$. Measurements made by HF play an important role in identifying jets, and calculating missing transverse energy. Also, its other role is being part of the BRIL system as a luminositymeter to measure the luminosity during the data taking at the CMS.

Phase-1 Upgrades

During the RunI data taking with "Phase-0" electronics, the Hybrid Photo Diodes (HPDs) were used as photo-detectors for the HB, HE, HO calorimeters and there had been found to produce large noise pulses in the presence of magnetic field. In order to eliminate this problem, the HPDs were replaced with SiPMs. The SiPMs are driven by several beneficial properties. They are inherently magnetic field insensitive. They have considerably high photon-detection efficiency and high gain. The size of the SiPM also facilitates increased depth segmentation because of the higher channel density that is possible. This allows for depth dependent calibrations to better adapt to layer dependent radiation damage.

Likewise, the HF calorimeter was using the 1-anode PMTs as photo-detector. But, these PMTs generated a large, fake signal and this reduced the data quality and physics events. So, 1-anode PMTs were replaced with the 4-anode PMTs. During the Long Shutdown 1 (LS1), the signals from 4-anodes were combined as one channel and these 1-channel readout mode PMTs were used to take data until end of 2015. During the Extended Year End Technical Stop (EYTS) 2016-2017, the PMT boxes were reworked to implement 2-channel readout mode PMTs. These new PMTs not only reduce the intrinsic level of background, but also enable tagging of background events and recovering the underlying signal event (if any) by using the multi-anode features.

There was another issue caused by QIE chips. During the RunI data taking, QIE8s were used by all HB, HE, HO, HF calorimeters. These QIEs caused an optimization challenge. So, they were replaced with QIE10/QIE11 (for HF/HB, HE) chips taking place on the QIE cards. Since the number of bits required to span the dynamic range is significantly larger than is needed to satisfy the precision at the highest energies. Also, this replacement was needed to support the replecements of photo-detectors. The QIE10 and the QIE11 are identical in design, except the QIE11 includes a programmable current shunt . The QIE10/QIE11 chips provide: Deadtime-less operation at 40 MHz (25 ns); Large dynamic range – maximum charge capacity; Timing information (TDC); Programmable, internal calibration; Low power draw.

The FE and BE electronics were also upgraded. The new FE electronics provide faster data link and more channels, and better radiation hardness. Whereas the new BE electronics provide high speed data transfer rates, and handle more data bandwith.

Based on these problems and replacements, the Phase-1 upgrades of the HCAL detector improved performance with a large number of pileup events, providing new capabilities for anomalous background rejection and smaller radiation damage effects. On the other hand, these upgrades provided strong physics performance and mitigate radiation damage effects.

During the Phase-1 upgrades, I had direct involvement in various aspects of the all CMS HCAL sub-detectors. My main contributions were on the HF calorimeter. I involved with (1) Assembling and testing the QIE cards individually, (2) Testing the QIE cards on the burn-in test stand including PMTs, FE and BE electronics, (3) Participating the installation and commissioning of the FE Electronics. Also, as a part of commissioning, I participated in the calibration test for the FE electronics with a radioactive source. Figure 3.10. shows some photos from Installation and Commissioning of Front-End Electronics of the HF calorimeter. Additionally, as another main contribution, I led the testing and integration of Calibration Units (CUs) for the HB calorimeter.

Figure 3.10. Photos from Phase-1 upgrades for the CMS HCAL. The HF burn-in test stand (top), including PMTs, FE and BE electronics. One of the HF cavern (bottom left). The bottom right photo from Installation and Commissioning of Front-End Electronics of the HF calorimeter

3.4.4. Magnet

The CMS detector has the largest solenoid superconducting magnet ever built. The Figure 3.11. shows picture of the solenoid magnet during installation. The main purpose of the magnet is to bend trajectories of charged particles in order to make a precise momentum measurement with the tracking detectors.

The CMS magnet system consists of a superconducting solenoid that can produce a uniform magnetic field of 3.8 T, which is 100,000 times stronger than the Earth's. It is enclosed in a 12,000-tonne steel return yoke, which acts as a filter to allow through only muons and weakly interacting particles such as neutrinos. The properties of the CMS solenoid magnet are shown in Table 3.1.

Table 3.1. Properties of the CMS solenoid magnet.

Property	**Value**
Field	3.8 T
Inner Diameter	5.9 m
Length	12.9 m
Weight	12,000 tonnes
Number of Turns	2168
Current	19.5 kA
Stored Energy	2.7 GJ
Hoop Stress	64 atm

During detector operation, the solenoid is cooled to 4.5 K and magnetic field results in a momentum resolution of $\Delta p/p \approx 10\%$ at a momentum of 1 TeV, which is sufficient to determine the sign of a muon up to $p_{\mathrm{T}} \approx 1$ TeV.

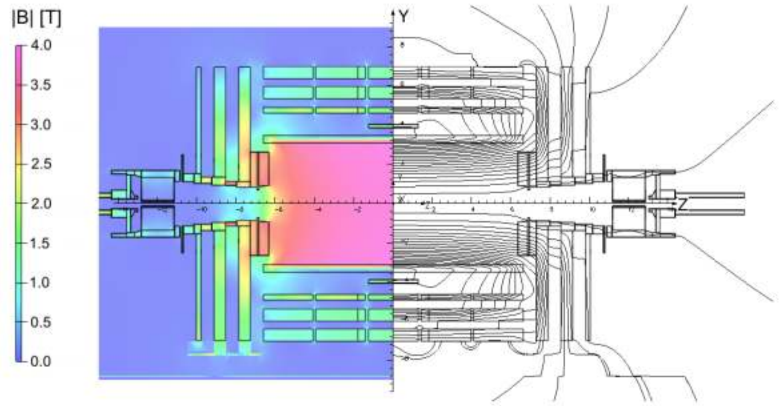


Figure 3.11. Picture of the solenoid with steal return yoke during installation (top). Magnetic field and field lines predicted on a longitudinal section of the CMS detector (bottom), for the underground model at a central magnetic flux density of 3.8 T. Each field line represents a magnetic flux increment of 6 Wb (C. Collaboration, 2010)

3.4.5. Muon System

Detecting muons is one of the most important tasks in the CMS, the presence of muons in the final state of a pp collision is almost always a sign of interesting physics such as Higgs boson, supersymmetric particles, electroweak precision measurements for the b-physics and top physics, etc. For these reasons the muon system is built to have excellent trigger, identification and reconstruction performances.

Muons are charged particles like electrons, but they are heavier than electrons. So, they do not interact with the calorimeters as much as electrons and this allows muons to pass through the calorimeters unstopped. Therefore, the muon system is located outside the solenoid, where muons are the only particles that would register a signal.

The CMS muon system covers the range of $|\eta| < 2.4$ and is divided into barrel and endcap regions. The basic detector process utilized in the CMS muon systems is gas ionization. So, the muon system is composed of gaseous detectors, which are Drift Tubes (DTs), Cathode Strip Chambers (CSCs) and Resistive Plate Chambers (RPCs). These gaseous detectors are sandwiched among the layers of the steel return yoke to allow a traversing muon to be detected at multiple points along the track path (Seo, 2013). Both DTs and CSCs are divided into four stations, where DTs are labeled MB1 through MB4 (Muon Barrel) and the CSCs labeled ME1 through ME4 (Muon Endcap). Layout of one quadrant of the CMS detector with the muon system is shown in Figure 3.12., along with a picture from the muon chambers in the return yoke taken from (Hoch, 2007).

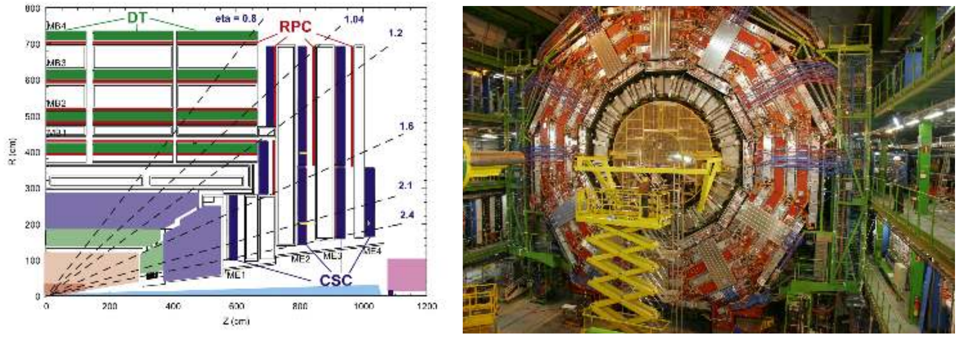


Figure 3.12. Schematic view of one quarter of the CMS muon system with DTs, CSCs and RPCs (left) (Seo, 2013). The DTs are labeled ”MB” (Muon Barrel) whereas CSCs are labeled ”ME” (Muon Endcap). Muon chambers interspersed in the magnet return yoke (right)

The DTs are designed to perform accurate positional information and additional trigger information for muons in the central region of the barrel covering the pseudorapidity range $|\eta| < 1.2$, where the track density and the residual magnetic field are low. In total, there are 250 DTs in these stations inside the magnet yoke. The CSCs also have the same purpose as the DTs in the endcap region, $0.9 < |\eta| < 2.4$, where muon rate is higher and a large residual magnetic field is present. The RPCs are placed in both barrel and endcap regions and they can provide a fast, independent trigger with a looser p_{T} threshold over a large portion of the pseudorapidity range $|\eta| < 1.6$ (Sirunyan, 2018). There are 468 CSCs intermixed with the RPCs to cover the endcap regions.

The muon system forms the CMS central trigger system with the HCAL and ECAL. As a part of the CMS phase1 upgrades, during the Long Shutdown 1 (2013-2014) and the last year-end technical stop (2015-2016), significant consolidation and upgrades have been carried out on the muon detectors and on the Level-1 (L1) muon trigger. The algorithms for muon reconstruction and identification have also been improved for both the High-Level Trigger (HLT) and the offline reconstruction (Cabrera, 2017).

3.4.6. BRIL System

Measuring the luminosity precisely is necessary for all measurements of cross-sections at the CMS. The Beam Radiation Instrumentation and Luminosity (BRIL) system is generally designed to measure the luminosity and additionally monitor the beam conditions during the data taking at the CMS (Guthoff, 2017) with all its sub-systems. It consists of several sub-systems. There are three systems for luminosity measurements, which are the Pixel Luminosity Telescope (PLT), the Fast Beam Conditions Monitor (BCM1F) and the Hadron Forward calorimeter for luminosity (HFlumi). In addition to the luminometers, there are also monitoring systems, which are the Beam Halo Monitor (BHM), the Beam Condition Monitor for Loss monitoring (BCML) and the Beam Position for Timing at the eXperiments (BPTX). The locations of the BRIL sub-systems are shown in Figure 3.13..

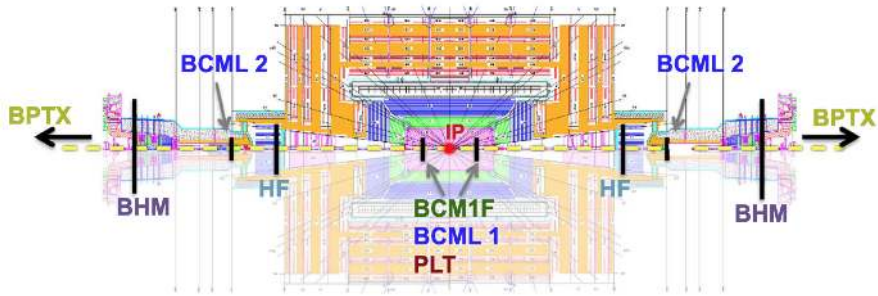


Figure 3.13. Position of the various BRIL systems inside CMS (Guthoff, 2016)

Also, the BRIL system provides to (Guthoff, 2016): (1) Measuring machine induced background (MIB) to quantify operational conditions and safe operation for the silicon tracker system, (2) Providing active protection in case of intense beam loss events as part of the LHC beam loss monitoring system, (3) Measuring the beam timing: The precise time of the incoming bunches with respect to the LHC orbit clock and in comparison of both beams.

3.5. CMS Trigger System

The nominal bunch crossing of 25 ns (40 MHz) at the LHC corresponds to $\sim 10^9$ interactions per second for the design luminosity (Sirunyan, 2017b). Only a small fraction of these collisions contain events of interest to the CMS physics program, which means that the rest of the events in each collision are never stored based on the physical and technological capability. For these reasons, the CMS trigger system has to make a decision to keep the largest possible number of interesting events for analyses in each LHC bunch crossing. Thus, the CMS trigger system effectively decrease the number of events from 40 MHz to 1 kHz in two stages. The hardware-based Level 1 (L1) Trigger system receives information coming from both ECAL and HCAL as well as the muon chambers and reduces the event rate from 40 MHz to 100 kHz in about 4 μs. The software-based High Level Trigger (HLT) reduces the rate from 100 kHz to 1kHz. Differently from the L1 trigger, the HLT has access to the full detector information.

3.5.1. L1 Trigger

The L1 trigger is a hardware-based system and performs the first level of online event selection with the raw information from the calorimeters and the muon system in order to make a decision within 4μs of the collision. The output rate of the L1 trigger is limited due to front-end detector limitations and computing power in the HLT. Because of the limited time for processing the detector data, the L1 trigger uses simple algorithms to process the incoming information on the custom hardware, called Field Programmable Gate Arrays (FPGAs) and ASICs to construct the physics objects for the triggering decision. These trigger objects, referred to as Trigger Primitives (TPs) are generated based on the energy deposits in the calorimeters and hits representing location and timing information of the muons in the muon chambers. The Regional Calorimeter Trigger (RCT) ranks and sorts the trigger objects based on the energy or momentum of the object and the quality of the reconstructed TPs. Then, the ranked trigger objects are evaluated by the Global Calorimeter Trigger (GCT) and Global Muon Trigger (GMT). Last, they are transfered to the Global Trigger (GT), where the final decision is then determined. The schematic of the hardware-based Level-1 Trigger is shown in Figure 3.14.. If the collision event is considered interesting by the GT, it distributes a Level 1 Accept (L1A) decision to the individual sub-detectors in order to begin the DAQ read-out of the event.

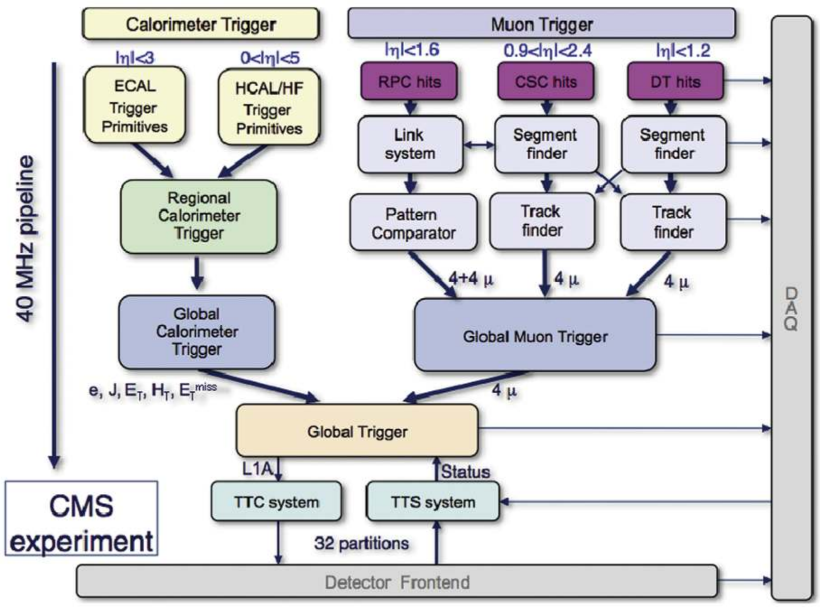


Figure 3.14. The CMS L1 trigger system (Jeitler, 2013).

3.5.2. HLT

The HLT is a software-based system to filter the events (Trocino, 2014). It is a large array of computers, processing the events that pass the L1A trigger. The HLT is a CPU farm made from over 13,000 processors. Its goal is to reduce event rate from 100 kHz to 1kHz, which is a sustainable level for storage. Its software package consists of a collection of C++ sequences for running event reconstruction and filters. Each HLT trigger path has its own modules that includes has some selection algorithms. If an event passes these modules in an HLT path, it means that the event passes that HLT trigger. Then, this event is fully reconstructed and saved to a primary dataset.

3.6. CMS Event Recontruction

The raw data collected by the CMS detector is not suitable to be used directly in physics analyses. There is a set of advanced algorithms to identify the particle content of an event and to recontruct the properties of each identified particle. It is very often that a particle leaves traces in more than one sub-detectors, as shown in Figure 3.15..

In the CMS, the *Particle Flow (PF) algorithm* (Sirunyan, 2017a) aims to reconstruct all stable particles in an event. This algorithm provides a global event description by combining information from the tracker system, the ECAL, the HCAL and the muon system to reconstruct the PF candidates – electrons, muons, photons, neutral hadrons and charged hadrons – by determining their tracks, directions and energies.

The PF algorithm starts with reconstructing charged-particle tracks. All the tracks in the tracker system are reconstructed by using the *iterative tracking algorithm*, which optimizes track efficiency and minimizes the fake tracks. Then, energy deposits in the preshower, ECAL barrel and endcap, HCAL barrel and endcap are individually clustered based on a certain threshold of energy deposit by *clustering algorithm*. Last, information from these algorithms is linked together by using *link-*

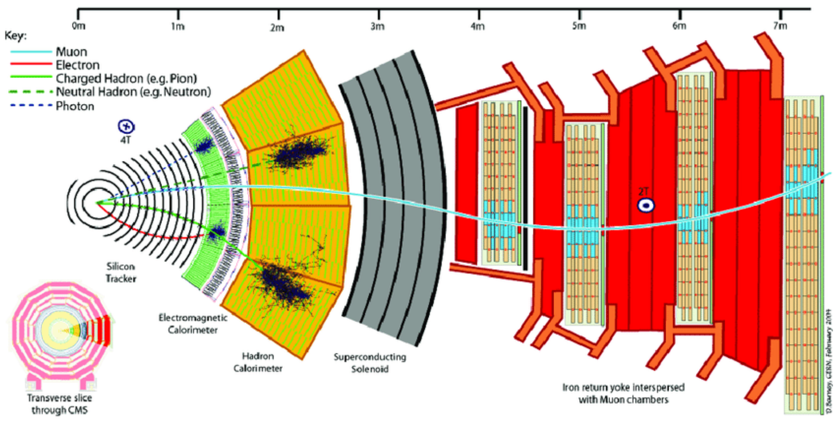


Figure 3.15. A sketch of the signatures of different types of particles in the CMS detector

ing algorithm to fully reconstruct each single particle, while removing any possible double counting from different detectors. The linking algorithm produces ”blocks” of elements linked directly or indirectly. Then, the PF algorithm reconstructs and identifies PF candidates, which is a set of particles from each block of elements (Dordevic, 2018):

- **PF electrons** are reconstructed from linked tracks and the cluster in the ECAL without any signal in the HCAL.
- **PF muons** are reconstructed from a track in the tracker system, which is linked with a corresponding track in the muon system.
- **PF photons and neutral hadrons** are reconstructed from calorimeter clusters, which are not linked to the tracks.
- **PF charged hadrons** are reconstructed by linking the track to one or two calorimeter clusters.

Then, these PF candidates are used to form additional objects – such as jets, b-tagged jets, top-tagged jets, etc. – for physics analyses. The objects used in this thesis are defined in Section 5.2.

4. SEARCH STRATEGY FOR RPV AND STEALTH SUSY

Motivated by many searches at the LHC searching for evidence of SUSY particles, we are interested in the pair production of top squarks. Most searches for the top squark look for an excess of events with large p_T^{miss} originating from the undetected lightest SUSY particle (LSP) produced in the top squark decays. But, there is no evidence yet for SUSY particle decays yielding a large p_T^{miss} signature in the detector. So, our motivation is to search for SUSY particles whose decays lead to low p_T^{miss}.

In this thesis, RPV and Stealth SY$\bar{Y}$ SUSY are introduced with two dedicated analyses. One is the legacy analysis and another is the current analysis—it is still ongoing. This thesis is based on the full status of the current analysis, including only Stealth SY$\bar{Y}$ model. General search strategy for these SUSY models are given in Section 4.1.. Then, with the motivation from the legacy analysis—given a short brief in Section 4.2., the current analysis strategy is defined in Section 4.3.

4.1. General Search Strategy for RPV and Stealth SUSY

The RPV model produces low-p_T^{miss} signature by providing a mechanism for the LSP "Neutralino" ($\tilde{X}_1^0$) to decay via the UDD-coupling. Final states of this model includes 2 top quarks and 6 light-flavor (quark) jets. The Stealth SY$\bar{Y}$ model supposes a "hidden sector" containing an supersymmetric fermion and its scalar partner. The supersymmetric fermion "Singlino" ($\tilde{S}$) decays through the hidden sector to a close-in-mass "Singlet"(S) and a Gravitino ($\tilde{G}$). The $\tilde{G}$ is the LSP and does not contribute much additional p_T^{miss} in the event. In this case, the top squark could decay through this hidden sector and ends in low-p_T^{miss}. The final state of this model includes 2 top quarks and 6 gluon jets. The decay chain of the top squark for each model is shown in Figure 4.1.. Considering a final state of $t\bar{t}+$jets with two top quarks, many extra light-flavor jets, and without additional p_T^{miss}, there will be lots of jets—signal events yield 12 jets. So, the primary topology feature of the signal is high jet multiplicity. To attain the best signal sensitivity, events are categorized by the number of jets (N_{jets}), to make use of the signal's high jet multiplicity.

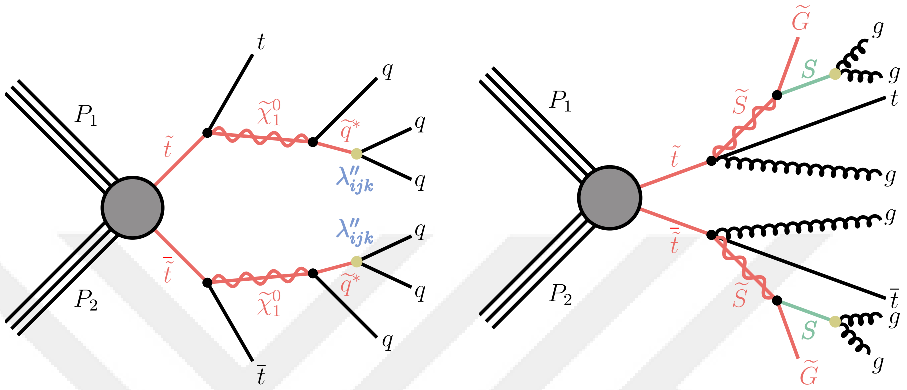


Figure 4.1. The Feynman diagrams for top squark pair production where the top squarks decay via the benchmark RPV (left) and Stealth SY$\bar{Y}$ (right) models. Considering a final state of each model, there is high jet multiplicity (N_{jets}) and minimal additional p_T^{miss}

Based on the final state of each signal model process, there is also pair production process of the top quark. The top quark decays through b quark and $W^{\pm}$ boson. The hadronic decay of the top quark is where $W^{\pm} \rightarrow q\bar{q}$ and leptonic decay of the top quark is where $W^{\pm} \rightarrow l\nu$. Considering the decay mode of the top quark, there are three channels: (1) fully-hadronic channel where both top quarks decay hadronically occurring about 44% of the time, (2) semi-leptonic channel where one of the top quark decays leptonically occuring about 42% of the time, (3) fully-leptonic channel where both top quarks decay leptonically occuring about 5% of the time. In this search, interest of semi-leptonic decay is (e or μ) and interest of fully-leptonic decay is (ee / $\mu\mu$ / $e\mu$).

The main goal of this search is to distinguish the SUSY processes (as signal) from the SM processes (as background). Based on the signal topology, only a few SM processes produce as many jets as the signal models and have large cross-section and includes additional jets from initial- and final-state radiation (ISR and FSR). In the SM processes, $t\bar{t}+$jets and QCD multi-jet are prominent backgrounds based on the large cross-section at the LHC. In this search, $t\bar{t}+$jets is the main background for both signal models and it is very similar to the signal models based on the final event topology. But, still the QCD multi-jet process is challenging, specially for the fully-hadronic channel, which can contain gluon splitting to pairs of b quarks. Then, most of the simulated events in the signal region are expected to be $t\bar{t}+$jets, with contributions of less than 10% from QCD multi-jet and a few percent from the remaining minor backgrounds including ”$t\bar{t}$ + X”, where X is one or more W/Z/H bosons and ”Other”, which includes DY + jets, diboson, triboson, SingleTop + jets and $W+$jets.

The first step of the search strategy is to develop a signal region selection. Then, a data-driven background estimation method using a neural network is used to determine the $t\bar{t}+$jets background. Using a neural network provides an additional performance due to the similarity between signal and background. Background from QCD multi-jet production is estimated in data in a dedicated control region. The estimations of the $t\bar{t}$ + X and Other backgrounds are taken directly from simulation. Then all these estimated backgrounds events are fit components with systematic uncertainties. For the fit procedure, the Higgs combine tool (Higgs Combine Tool, 2021) is implemented. The full Run2 data, corresponding to a total integrated luminosity of 138 fb^{-1}, collected with the CMS detector at the LHC in 2016 to 2018 are used.

There are two rounds of the analysis: *legacy* and *current* analyses. The legacy analysis searches for the pair production of top squark in the semi-leptonic final state via the benchmark RPV and Stealth $SY\bar{Y}$ models—this is the first search of its kind at the LHC (Sirunyan, 2021b). The main distinguishing feature of the signals is high jet multiplicity (N_{jets}). Based on all backgrounds mentioned above, there are additional jets from ISR and FSR. Based on the semi-leptonic channel, requiring exactly one charged lepton (e/μ) reduces the QCD multi-jet background. But, still it is the second main background and implements a deticated control region for estimation. Then, the main background $t\bar{t}+$jets are distinguish from signal by means of a neural network (NN). Events are divided into NN score regions. A custom fit function is designed to describe the N_{jets} distribution in each of the NN score region. A short brief of this round is given in Section 4.2..

In the current analysis, there are couple of differences from the legacy version of the analysis: (1) Including all fully-hadronic, semi-leptonic and fully-leptonic channels can provide better signal sensitivity; (2) Implementing an object tagger, which is a NN, reduces the QCD multi-jet events for the fully-hadronic channel; (3) Applying a more standard background estimation technique for $t\bar{t}+$jets—the ABCD method using a Double DisCo NN. Then, the same QCD control region from the legacy analysis is also implemented for this round. More details are given in Section 4.3.. This thesis is based on the current analysis.

4.2. Motivation from Legacy Analysis

This round is a search for pair production of top squark with the *semi-leptonic* final state via the benchmark RPV and Stealth $SY\bar{Y}$ models. This is the first search of its kind at the LHC (ibid.). A short brief of this analysis is presented in this section.

$t\bar{t}$ **Background Estimation**

First, events are required to have at least seven jets, at least one b-tagged jet, and exactly one lepton as signal region. Based on the primary topology feature of the signal models, jet multiplicity is hard to model at high multiplicity, so we wanted to rely on data and fit the number of jets (N_{jets}) distribution. By making the N_{jets} selection, the higher N_{jets} multiplicity bins (nominally around ten jets) are dominated by signal and the lower N_{jets} multiplicity bins are dominated by $t\bar{t}+$jets background, and are used as a background dominated region to constrain the normalization of the data-driven background fit shape. The N_{jets} shape of the $t\bar{t}+$jets background is then fit in the signal region.

Because of the similarity between final state of the signal models and $t\bar{t}+$jets background, a $t\bar{t}+$jets dominated region may have high signal contamination. Thus, a Neural Network (NN) is designed to distinguish between signal and $t\bar{t}+$jets events to better separate between them. The NN is trained to enhance the discrimination between signal and $t\bar{t}+$jets background. The gradient reversal technique (Ganin and Lempitsky, 2014) is used to minimize N_{jets} dependence in the NN discriminator score (S_{NN}). The jet four-momenta, the lepton four-momentum and additional event wide variables are chosen as NN input variables which are computed in an approximate center-of-mass frame defined by all jets in the event with $p_{\text{T}} > 30$ GeV and $\eta < 5$. For the NN training, simulated $t\bar{t}$ events are used for the background sample, and a mixture of RPV and Stealth $SY\bar{Y}$ simulated events with $m_{\tilde{t}}$ from 350–850 GeV is used as the signal sample.

After giving a NN score to events, the signal region is divided into four regions based on NN score - four S_{NN} bins. The $S_{NN,1}$ is highly dominated by the $t\bar{t}+$jets background and $S_{NN,4}$ has the highest signal sensitivity, like in Figure 4.2.. In each of the S_{NN} bins, the N_{jets} shape for the $t\bar{t}+$jets background is provided to be the same: (1)

the N_{jets} shape is designed to be independent of the S_{NN} bin; (2) the definition of the S_{NN} bins is made per N_{jets} to remove residual differences in N_{jets} shape; (3) systematic uncertainties give the fit flexibility to accommodate differences between data and simulation. Based on given procedure above, the $t\bar{t}$ + jets shape is simultaneously fit in all four S_{NN} bins, with an assumption that the $t\bar{t}$ + jets shape is the same in each S_{NN} bin.

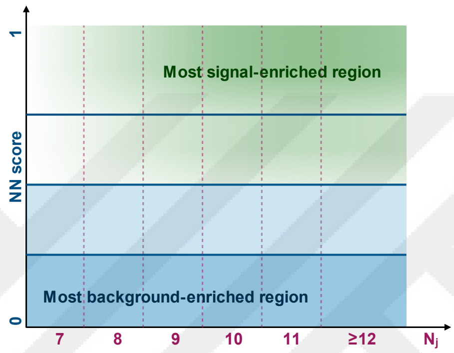


Figure 4.2. The signal region shown divided up by N_{jets} and S_{NN} bins. Low N_{jets} and S_{NN} bin is background enriched and high N_{jets} and S_{NN} bin is signal enriched

For fitting the N_{jets} shape of the $t\bar{t}$ + jets background, a custom fit function is designed based on the theoretical motivation - QCD multi-jet scaling patterns (Gerwick et al., 2012) - in which the ratio $R(i) = M_{i+1}\ /\ M_i$, where M_i is the number of events with N_{jets} = i, can be described by a falling "Poisson" component at low N_{jets} and a constant "staircase" component at high N_{jets}. Then, this $R(i)$ is used as a component of the custom fit function. The custom fit function is used recursively to predict the number of $t\bar{t}$ events with a certain N_{jets}.

QCD multi-jet Background Estimation

QCD multi-jet production is another prominent background. At the high N_{jets} bins, there are limited number of simulated events with large event weights. The large uncertainties from the low event counts and large event weights give the fit too much freedom and could absorb a signal. Then, a control region in data is used to estimate QCD multi-jet background in the signal region.

First, the QCD control region (QCD CR), which is highly enriched with QCD multi-jet events, is defined. The events are required to have at least seven jets and there is no b-tagged jet requirement. The QCD CR is orthogonal to the search region, there is a requirement that there is exactly one non-isolated muon and no good muons and no good electrons. In order to be able to trigger on such a muon, the transverse momentum requirement on the muon is $p_{\mathrm{T}} > 55$ GeV.

The N_{jets} shape is taken from background subtracted data where non-QCD multi-jet background predictions from simulation are subtracted out. Then, the N_{jets} shape in the data is normalized by using a transfer factor (TF), which is the ratio between the number of weighted QCD multi-jet events in the signal region (N_{SR}) and the number of weighted QCD multi-jet events in the QCD CR. Note that everything is based on simulation for the TF.

Fit Procedure and Results

The fit components – which are $t\bar{t}+$jets and QCD multi-jet background estimation, minor background estimations coming from simulation, and systematic uncertainties based on $t\bar{t}$ modeling, jet energy correction (JEC), jet energy resolution (JER), scale factors and modeling of N_{jets}-S_{NN} correlation test using QCD CR – are fed into the Higgs Combine Tool.

The result of this analysis is a maximum observed local significance of $\sim$2.8σ corresponding to the RPV $m_{\tilde{t}}$ = 400 GeV mass point. The observed exclusion limits are set at 675 GeV and 900 GeV for RPV and Stealth SUSY, respectively. Figure 4.3. shows the expected and observed cross section limits as a function of $m_{\tilde{t}}$ for the benchmark RPV and Stealth $SY\bar{Y}$ signal models.

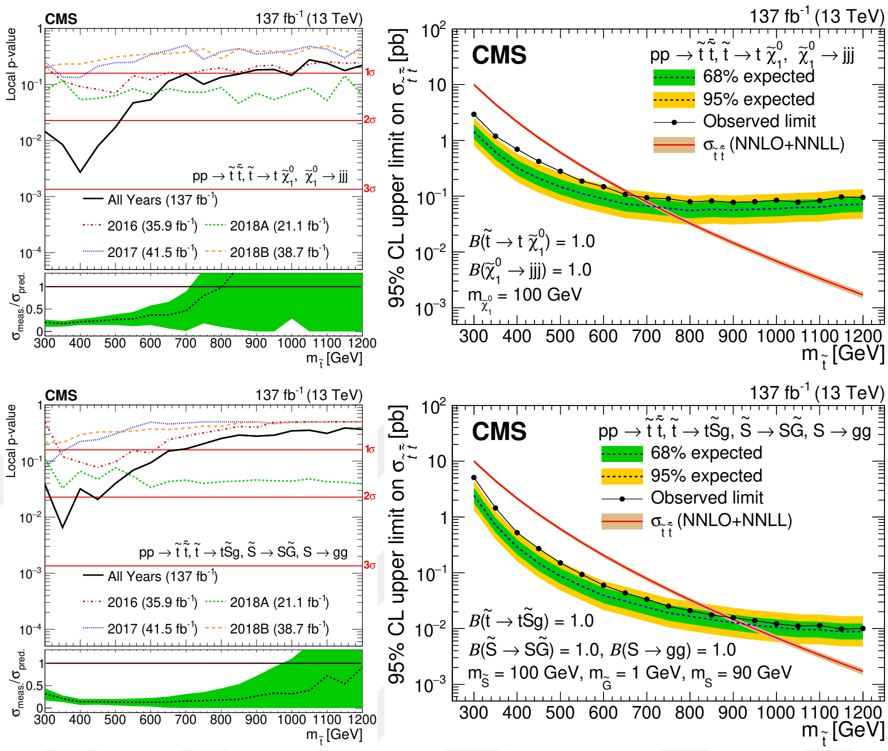


Figure 4.3. Expected and observed 95% C.L upper limit on the top squark pair production cross section as a function of the top squark mass for the RPV (top) and Stealth $SY\bar{Y}$ (bottom) SUSY models. Particle masses and branching fractions assumed for each model are included on each plot. The expected cross section computed at NNLO+NNLL accuracy is shown in the red curve (Sirunyan, 2021b)

4.3. Search Strategy for Current Analysis

Motivated by the legacy analysis, the current analysis searches for pair production of the top squarks and subsequent decays in the fully-hadronic, semi-leptonic and fully-leptonic channels. The final results for these three channels are to be combined in the end for the best possible signal sensitivity. Additionally, with respect to the legacy version of the analysis, the addition of the fully-hadronic and fully-leptonic channels will help to make a stronger statement about the past results. This current round of the analysis is mature, but still ongoing. In this thesis, the search for the Stealth SY$\bar{Y}$ model is presented, including all necessary analysis-level decisions and studies that are necessary before unblinding. This includes a full description of the main background estimation method and preliminary, expected results from the final fit procedure. All the details are presented in rest of the thesis chapters. In the first step of the analysis, the signal region for each channel is developed. In addition to the object selection criteria, additional requirements for fully-hadronic channel may come from top quark pair production of the signal topology. By considering each channel, most of the background comes from the fully-hadronic channel based on the decay branching fraction of the top quarks in the collision. There are also extra challenges for this channel due to QCD multi-jet events, which contain gluon splitting to pairs of b quarks. So, the “Hadronic Top Tagger” (detailed in Section 6.) is implemented as a part of the signal region selection for the fully-hadronic channel to reduce the number of QCD multi-jet events. The hadronic top tagger is a neural network, including a combination of two approaches to tag the hadronically decaying top quarks in all p_{T} ranges. This object tagger allows tagging of hadronically decaying top quarks by eliminating the fake top objects.

The ABCD method using a Double DisCo Neural Network is applied for the $t\bar{t}+$jets background estimation to have less challenges on the N_{jets} dependency and systematic uncertainties. The ABCD method is one of the common data-driven background estimation techniques in high energy physics, which extrapolates or interpolates from some control regions into a signal region of interest by choosing two independent variables. To define these regions, the Double DisCo NN is used to produce two discriminator values from 2 binary classifiers. These discriminator values represent the two independent variables. The method proceeds by forming a two dimensional plane, where **A** is the signal region and **B**, **C**, **D** are the control regions. A toy example including each region with some events is shown in Figure 4.4.. Then the background in the signal region (**A**) can be predicted by using the control regions (**B**, **C**, **D**) via the relation $N_A = (N_B N_C)/N_D$, where N_A is the number of events in **A** region, N_B is the number of events in **B** region, N_C is the number of events in **C** region, and N_D is the number of events in **D** region. If the network performance is good and the method works well, then background can be predicted in the data. Lastly a correction factor is defined as part of calculating systematic uncertainties. All details are presented in dedicated chapter of thesis, Section 8.. Also, as a part of the Double DisCo NN trainings, an extra set of variables – clustered hemisphere variables given in Section 7. – is fed into to the network for getting better discrimination between signal and background events. With the Double DisCo NN trained on simulated events for all years of Run2 for each of the three analysis channels individually, the ABCD method can be performed. To attain the best signal sensitivity, events are categorized by the number of jets (N_{jets}), to make use of the signal's high jet multiplicity. Then, the ABCD method can be performed simultaneously for each N_{jets} bin where sensitivity is maximized in the high N_{jets} bins.

The second largest background yield comes from QCD multi-jet events. At high N_{jets}, the QCD multi-jet background has low simulation event counts and large event weights. This can cause large statistical uncertainties. These large statistical uncertainties can hide a signal, because they give the fit too much freedom, which decreases the overall sensitivity of the analysis. Therefore, to avoid factoring in these large statistical fluctuations, the QCD multi-jet background is estimated using collision data in a QCD-dominated control region (detailed in Section 5.4.2.). For the fully-leptonic channel, total number of QCD multi-jet events is negligible based on the simulation, and thus it is not considered in the final fit. Thus, the procedure defined in Section 9.1.1. only applies to the fully-hadronic and semi-leptonic analyses.

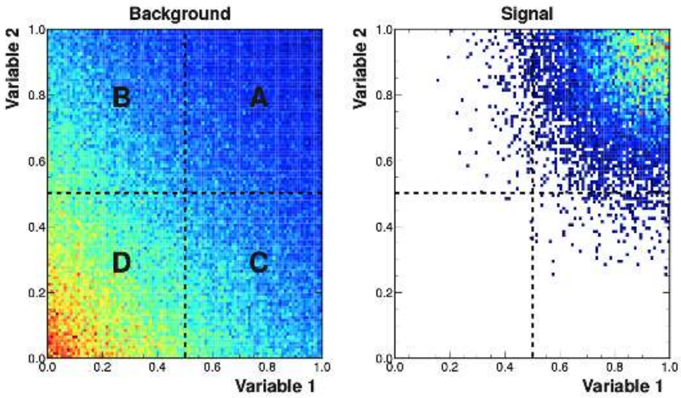


Figure 4.4. A toy example for including each of A, B, C, D regions with some event yields. Each of the two dicriminator values of the network corresponds to each of the two independent variables. The number of predicted background events in **A** region can be computed via $N_A = (N_B N_C)/N_D$

For the fit procedure, the Higgs combine tool (Higgs Combine Tool, 2021) is implemented with the components including predicted events of each background with a related systematic uncertainty. The fit procedure and preliminary fit results are presented in Section 10.. Note that the $t\bar{t}$ + jets background estimation method is the same for all three channels. But, there is no QCD multi-jet background estimation for fully-leptonic channel.

5. EVENT RECONSTRUCTION AND SELECTION

In this chapter, data sets and simulations are mentioned in Section 5.1.. The reconstructed physics objects used in this analysis are presented in Section 5.2.. The trigger scale factors and top tagging scale factors are computed locally for this analysis and dedicated studies are shown in Section 5.3.. Last, the event selection requirements are listed in Section 5.4. and some comparisons between simulation and data are shown.

5.1. Data Sets and Simulation Samples

5.1.1. Data

The analysis presented in this thesis uses the Run2 data collected during the operation periods in 2016, 2017, and 2018 by the CMS detector. Overall this period of ultra legacy (UL) version of Run2 (CMS Collaboration, 2022d) data collection, pp collision occured at a center-of-mass energy of $\sqrt{s} = 13$ TeV with corresponding luminosity 138 fb^{-1}. Summary of the UL version of the Run2 (Run2UL) data collection is shown in Figure 3.2..

For the online event selections, a set of triggers is used for each of fully-hadronic, semi-leptonic and fully-leptonic channel, more details are given in Table 5.7.. The JetHT dataset is used for the fully-hadronic channel and the Single-Electron and SingleMuon datasets are used for the semi-leptonic and fully-leptonic channels, along with a complete set of simulated background samples for all channels. Data sets are listed in Table 5.1.. Also, event veto is applied for 2018 data, which there was an issue corresponding to the HCAL HE readout electronics during the data taking. So, data events with an electron in the region of HEM15 and HEM16 sectors were vetod.

Table 5.1. Run2UL data sets used in this analysis. Note data 2018 SingleElectron data set is renamed as EGamma, as shown in (CMS Collaboration, 2022c)

JetHT	**Integrated luminosity (pb^{-1})**
2016preVFP	19517.540
2016postVFP	16810.813
2017	41471.589
2018	59817.406
SingleElectron	**Integrated luminosity (pb^{-1})**
2016preVFP	19512.351
2016postVFP	16809.957
2017	41476.022
2018	59816.229
SingleMuon	**Integrated luminosity (pb^{-1})**
2016preVFP	19514.702
2016postVFP	16810.812
2017	41475.261
2018	59781.960

5.1.2. Simulation

After collecting the data by the CMS detector, it is important to do useful comparison about what is observed in data and what is expected from theoretical predictions. The SM backgrounds and particular BSM signal processes used in this analysis are simulated by using Monte Carlo (MC) event generation method. This method is broadly split into 2 parts. The first part is corresponding to the event generation and the parton shower (PS) calculation, which implements the event generators such as Pythia. The second part is detector simulation, which implements the Geant4. Although the second part is based on each subdetector feature and PS calculation, the first part is detector-independent.

Signal samples are generated in two steps: (1) MadGraph5_amc@nlo is used at leading order (LO) for the pair production of top squarks, including up to two additional partons; (2) pythia is then used for the decay of the top squarks, as well as the hadronization and parton shower. For each of the two considered signal models, top squark masses between 300 and 1400 GeV are generated, in steps of 50 GeV. In addition to the most dominant backgrounds samples $t\bar{t}+$jets and QCD multi-jet simulated events, there are also minor background samples listed in Table 5.2.. Also, more information of the production cross sections for each set of simulated background and signal samples are given in Appendix B1..

When using simulated events, a suite of scale factors are applied to the simulation in order to more accurately compare with data. These scale factors include b-tag efficiency, trigger efficiency, pileup correction factor, and weighting for prefiring in the detector for all years. In addition to these scale factors there are also lepton identification and isolation efficiency for the semi-leptonic and fully-leptonic analyses, while a top-tagging efficiency scale factor is applied for the fully-hadronic analysis. Only the trigger scale factor and top-tagging efficiency scale factor are uniquely derived for the analyses, documented in Sec.5.3.2... All other scale factors and weights are provided by the respective POGs, based on their Run2UL recommendations (CMS Collaboration, 2023). Last, the simulated events for each process must be weighted (ω_i) in terms of the cross section (σ) and integrated luminosity (L), given in Equation 5.1. Then, this is multiplied by all other scale factors for each analysis to get absolute final event weight.

$$\omega_i = \frac{\sigma * L}{N_{gen}} \tag{5.1}$$

Table 5.2. The minor background processes used in this analysis are listed. These processes are grouped into ”$t\bar{t}+$X ” and ”Other” by the analysis. Each individual minor background listed is about less than 1% of the total background

Category	$t\bar{t}+$**X**		**Other**
Processes	$t\bar{t}Z \to \ell\bar{\ell}\nu\bar{\nu}$	$t\bar{t}+t\bar{t}$	DY + jets
	$t\bar{t}Z \to q\bar{q}$	$t\bar{t}+hh$	$\bar{t}Zq$
	$t\bar{t}W+$jets $\to \ell\nu$	$t\bar{t}+tW$	WW
	$t\bar{t}W+$jets $\to q\bar{q}$	$t\bar{t}+Wh$	WZ
	$t\bar{t}\gamma+$jets	$t\bar{t}+WW$	ZZ
	$t\bar{t}h+$jets	$t\bar{t}+WZ$	WWW
		$t\bar{t}+Zh$	WWZ
		$t\bar{t}+ZZ$	WZZ
		$t\bar{t}t+$jets	ZZZ
			Single t
			$W+$jets

5.2. Reconstructed Physics Objects

5.2.1. Electrons

The electron definition starts with the collection of electrons determined through the Gaussian Sum Filter algorithm. The “tight” working point (WP) is imposed for the cut-based identification for all three years. There are related variables for the identification (CMS Collaboration, 2020a) is applied in order to increase the purity of the sample of electrons. Also, the electrons are further selected based on their transverse momentum and pseudorapidity based on trigger threshold and the highest efficiency and resolution. Besides these criteria, the electron isolation is calculated using the mini-isolation (miniIso) algorithm (CMS Collaboration, 2020b) (rather than the default built-in isolation requirement in the ID itself) and electrons are required to pass miniIso cut that the analysis search region is expected to be high occupancy. This algorithm calculates the isolation by implementing the cone size (cone around the electron) and comparing the transverse momentum of the electron to other objects in the cone. Last, impact parameter cuts are applied to enforce that the electron arises from the associated vertex with using two dimensions—the transverse (d_0) and longitudinal (dz) distances from the primary vertex as given for both barrel and endcap regions. The relevant selections for ”good” electrons is shown in table Table 5.3..

Table 5.3. Good electron selection criteria for the signal region

Good Electron Selections
”Tight” cut-based electron ID
$p_{\mathrm{T}} > 30(37)$ GeV for 2016 (2017/2018)
$\eta < 2.4$
MiniIso < 0.1
$\lvert d_0 \rvert < 0.05$ (0.10) cm for barrel (endcap)
$\lvert d_z \rvert < 0.10$ (0.20) cm for barrel (endcap)

5.2.2. Muons

The muon definition starts with the collection of muon determined through the PF algorithm. The “medium” WP is imposed for the cut-based identification and a set of parameters are used (CMS Collaboration, 2022a). Also, the muon are further selected based on their transverse momentum and pseudorapidity given as $p_T > 30$

GeV and $|\eta| < 2.4$ for all three years. Mini-isolation for muons (CMS Collaboration, 2020c) is defined like that for electrons, with the miniIso cut requirement on muons. Finally, there are impact parameter cuts, as given in Table 5.4., which are based on transverse (d_B) and longitudinal (d_z) distances from primary vertex.

To create a control region for QCD multi-jet background estimation, determined in Section 5.4.2., non-isolated muons are specifically defined for the analysis. Non-isolated muon selection criteria with the other selections above are listed in Table 5.4..

Table 5.4. Good muon selection criteria for the signal regions. Non-isolated muon selection criteria for the QCD control region. For the control region, only SingleMu50 type triggers are used, and thus the offline p_T selection is increased to 55 GeV

Good Muon Selections	**Non-isolated Muon Selections**
”Medium” cut-based muon ID	”Medium” cut-based muon ID
$p_T > 30$ GeV	$p_T > 55$ GeV
$\eta < 2.4$	$\eta < 2.4$
MiniIso < 0.2	MiniIso > 0.2
$\|d_B\| < 0.2$ cm	$\|d_B\| < 0.2$ cm
$\|d_z\| < 0.5$ cm	$\|d_z\| < 0.5$ cm

5.2.3. Jets

There is an additional object type that plays an important role in many searches for SM and BSM due to their large production rate at the LHC, which is called ”jet”. Jets are complex objects that quark and gluons coming from particle showers and hadronization are grouped together into a single 4-vector. So, there are some algorithms to define the jets, which select final-state particles and combine them.

The jets used in this analysis are AK4 CHS jets, that PF candidates are clustered with anti-k_T algorithm (Cacciari, G. P. Salam, and Soyez, 2008) and additional step are applied to subtruct charge hadrons. The anti-k_T algorithm is a sequential clustering algorithm and defines distances d_{ij} between particles i and j. The distance measures are given by

$$d_{ij} = min\left(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2}\right)\frac{\Delta_{ij}^2}{R^2} \tag{5.2}$$

where k_{ti} is the transverse momentum of the ith candidate, Δ_{ij} is the spatial separation of the candidate pair in the coordinate plane of $\eta - \phi$ and R is a tunable parameter to set the scale of the size of the clustered jets, standard choices are $R = 0.4$ and $R = 0.8$ in the CMS.

The jets used by the analysis presented in this thesis are clustered with the anti-k_T algorithm using a radius of 0.4, called AK4 jets. Exact p_{T} and η requirements on jets for the individual channel event selections can be found in Section 5.4.1. Each jet passes ”JetID”(*Jet algorithms performance in 13 TeV data* 2017) requirements. To avoid overlap between jets and leptons, a simple cleaning procedure is used. Any jet that is within a ΔR of 0.4 of an identified and isolated lepton and whose transverse momentum is within 100% of the p_T of the lepton, is removed from consideration.

5.2.4. b-tagged Jets

Some of the jets should be identified as a b jet. For determining if a jet is considered b jet, deep flavor algorithm (Bols et al., 2020) is used to clasify jets originating from the hadronization of b quarks. When using a neural network (NN) to tag objects, there is a particular threshold of discriminator value called ”working point (WP)” where gives fewer mistag rate and higher efficiency. The ”Medium” WP with a 1% mistag rate for the b tag discriminator value is used. The working points depending on each year are given in Table 5.5..

Table 5.5. Working points (WPs) for tagging b jets

	2016preVFP	2016postVFP	2017	2018
WP	0.2598	0.2489	0.3040	0.2783

5.2.5. top-tagged Jets

Considering the final state of the signal process, $t\bar{t}+$jets, the two top quark decays represent non-trivial resonances that can be identified and used to reject many-jet background events such as those originating from QCD multi-jet processes. For the fully-hadronic channel, QCD multi-jet production can easily be a significant background. To help reducing the QCD multi-jet background for the fully-hadronic analysis, the **Hadronic Top Tagger**, which was developed for a SUSY analysis (Sirunyan, 2021a), is implemented as a part of the signal region selection to tag hadronically decaying top quarks.

Regarding to the fully-hadronic channel, the leptonic decays of top quarks are ignored in this tagger. Two deep NN algorithms, which are DeepResolved and DeepAK8, are combined as the Hadronic top tagger to tag top quarks in all p_{T} ranges. More details of the Hadronic Top Tagger and a dedicated study to optimize the WP for DeepResolved tagger are given in Section 6.. The WPs depending on both algorithms are given in Table 5.6.. Currently, for the DeepResolved tagger, the top WP is the optimized based on the dedicated study–given in Section 6.2.. Also, DeepResolved tagger requires one of the decay products of top quark as b-jet by implementing b-jet WP. Currently, “medium” DeepFalvor WP is used for b-tagging. Note that the original version of the DeepResolved tagger uses the WPs given in Section 6.1.2..

Table 5.6. WPs for each algorithm corresponding to each year. The top WP given for DeepResolved tagger is the optimized one. DeepResolved tagger also requires one of the decay products of top quark as b-jet by implementing “medium” DeepFalvor WP

	2016preVFP	2016postVFP	2017	2018
DeepResolved	WP	WP	WP	WP
top	0.95	0.95	0.95	0.95
b-jet	0.2598	0.2489	0.3040	0.2783
DeepAK8	WP	WP	WP	WP
top	0.937	0.937	0.895	0.895

5.3. Scale Factors

5.3.1. Trigger Scale Factors

There are a lot of trigger paths related to different physics analyses in the CMS. Triggers are applied as a part of the online event selections. In this analysis, the jet triggers are applied in the fully-hadronic channel whereas electron and muon triggers are applied in the semi-leptonic and fully-leptonic channels. Each trigger is applied independently per event, and each event is required to satisfy at least one of the trigger requirements. All these triggers per year that are applied in the analysis are listed in Table 5.7.

For the fully-hadronic channel, jet triggers are chosen to target events with a large amount of hadronic activity that is characteristic of the final states of interest with top quarks decays with many jets. Accordingly, the lowest, unprescaled triggers requiring high p_{T} jets, high H_{T}, or large numbers of jets with DeepCSV for b-tagging are chosen. These are recommended by the TOP PAG for triggering on SM top decays, which closely resemble our target signal.

For semi-leptonic and fully-leptonic channels, electron and muon triggers are applied. For events containing exactly one "good" electron, the standard SUSY-recommended single-electron triggers are applied: HLT_Ele27_WPTight_Gsf for 2016 and HLT_Ele35_WPTight for 2017 and 2018. As recommended by the SUSY PAG, two additional triggers are added, which are the lowest p_T, unprescaled, non-isolated electron triggers (to recover efficiency at high p_T, where inefficiencies arise from mismatches between the isolation used online and mini-isolation used offline) and a high p_T photon trigger. For events containing exactly one "good" muon, the standard, lowest unprescaled, isolated and non-isolated single muon triggers are applied.

Trigger efficiency must be different for data and simulation, because it is not possible to send the simulation events through the hardware trigger. Instead, an emulation of the trigger is applied to simulation. Differences between the trigger efficiencies of the data and simulation after their respective treatments of the trigger are used to derive a trigger scale factor, which is then applied to simulation to match the efficiencies observed in data. This difference needs to be considered and the way to do this is that derives a scale factor.

First, the efficiency of the combination of triggers is needed to be measured for each data set. For measuring the trigger efficiency, a preselection region is defined as "denominator", which is very similar to the signal region. Then, "numerator" is defined as the number of events passing the the set of triggers in addition to the preselections used for the denominator. The efficiency can be calculated as ratio between numerator and denominator. The trigger scale factor can be derived as a ratio between simulation efficieny and data efficiency.

Table 5.7. Each jet, electron and muon trigger used by year, which are OR'ed together for a given year. Both CSV and DeepCSV triggers are used to cover the full run range for a given year. 2016 jet/electron/muon triggers are applied for both 2016preVFP and 2016postVFP years

2016 Jet Triggers	**2016 Electron & Muon Triggers**	
HLT_PFJet450	HLT_Ele27_WPTight_Gsf	HLT_IsoMu24
HLT_PFHT900	HLT_Ele115_CaloIdVT_GsfTrkIdT	HLT_IsoTkMu24
HLT_PFHT450_SixJet40_BTagCSV_p056	HLT_Photon175	HLT_Mu50
HLT_PFHT400_SixJet30_DoubleBTagCSV_p056		HLT_TkMu50
2017 Jet Triggers	**2017 Electron & Muon Triggers**	
HLT_PFJet500	HLT_Ele35_WPTight	HLT_IsoMu24
HLT_PFHT1050	HLT_Ele115_CaloIdVT_GsfTrkIdT	HLT_IsoMu27
HLT_PFHT380_SixPFJet32_DoublePFBTagDeepCSV_2p2	HLT_Photon200	HLT_Mu50
HLT_PFHT380_SixPFJet32_DoublePFBTagCSV_2p2		
HLT_PFHT430_SixPFJet40_PFBTagCSV_1p5		
HLT_PFHT380_SixJet32_DoubleBTagCSV_p075		
HLT_PFHT430_SixJet40_BTagCSV_p080		
2018 Jet Triggers	**2018 Electron & Muon Triggers**	
HLT_PFJet500	HLT_Ele35_WPTight	HLT_IsoMu24
HLT_PFHT1050	HLT_Ele115_CaloIdVT_GsfTrkIdT	HLT_IsoMu27
HLT_PFHT380_SixPFJet32_DoublePFBTagDeepCSV_2p2	HLT_Photon200	HLT_Mu50
HLT_PFHT380_SixPFJet32_DoublePFBTagCSV_2p2		
HLT_PFHT430_SixPFJet40_PFBTagCSV_1p5		
HLT_PFHT430_SixPFJet40_PFBTagDeepCSV_1p5		
HLT_PFHT400_SixPFJet32_DoublePFBTagDeepCSV_2p94		
HLT_PFHT450_SixPFJet36_PFBTagDeepCSV_1p59		

Jets

The jet trigger efficiency is calculated in the SingleMuon data set. The preselections are given in Table 5.8. for the fully-hadronic channel. Requiring at least one N_{muon}^{good} is to register the SingleMuon data.

Table 5.8. The preselections for measuring the jet trigger efficiencies and jet trigger scale factors

Preselections
SingleMuon data
Pass JetID
HEM 15/16 veto applied to 2018 data
$N_{muon}^{good} \geq 1$ ($p_{\mathrm{T}} > 30$ GeV)
$H_{\mathrm{T}} > 500$ GeV ($p_{\mathrm{T}} > 45$ GeV)
$N_{\mathrm{jets}} \geq 6$ ($p_{\mathrm{T}} > 45$ GeV)
$N_{\mathrm{bjets}} \geq 2$ ($p_{\mathrm{T}} > 45$ GeV, medium DeepFlavor)

Then the jet trigger efficiencies are measured as a function of jet H_{T} and p_{T} of the 6$^{\mathrm{th}}$ jet in $t\bar{t}+$jets simulation and in SingleMuon data. The measuremets of the efficiencies for indivual years of full Run2UL are shown in Figure 5.1.. The jet trigger efficiency scale factor as a function of jet H_{T} and p_{T} of the 6$^{\mathrm{th}}$ jet are given in Figure 5.2..

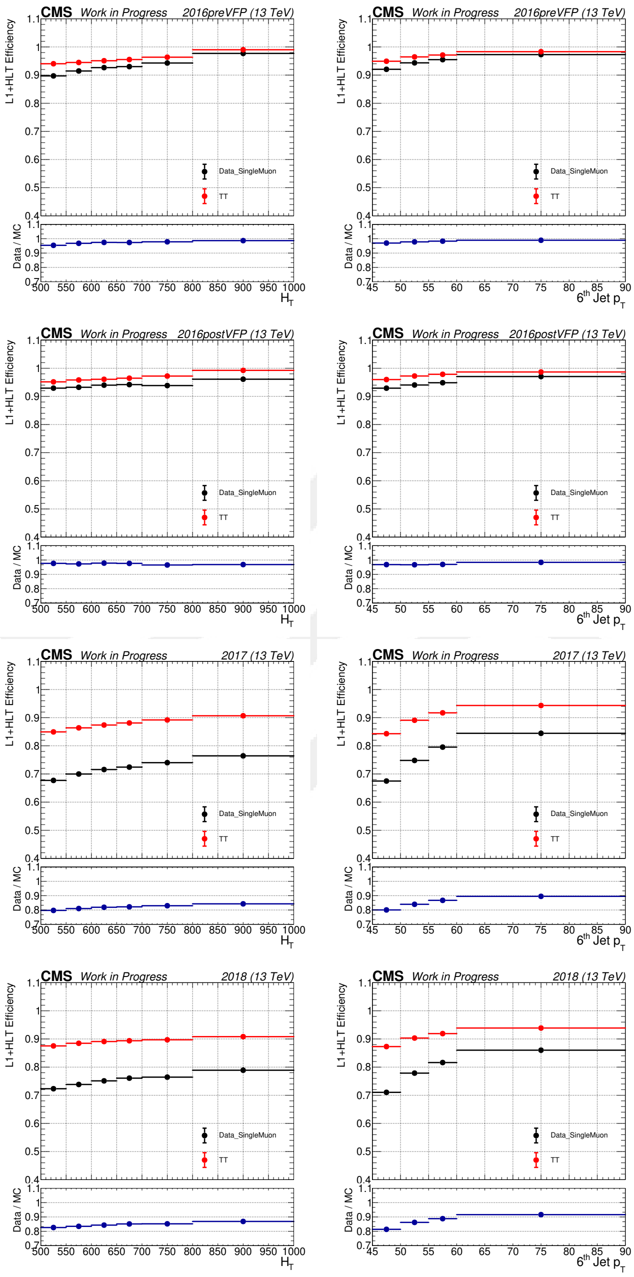


Figure 5.1. For the fully-hadronic channel, the jet trigger efficiency measured as a function of jet H_{T} (left) and p_{T} of the 6^{th} jet (right) in $t\bar{t}+$jets simulation and in SingleMuon data for 2016preVFP (top), 2016postVFP (middle top), 2017 (middle bottom), and 2018 (bottom) of full Run2UL

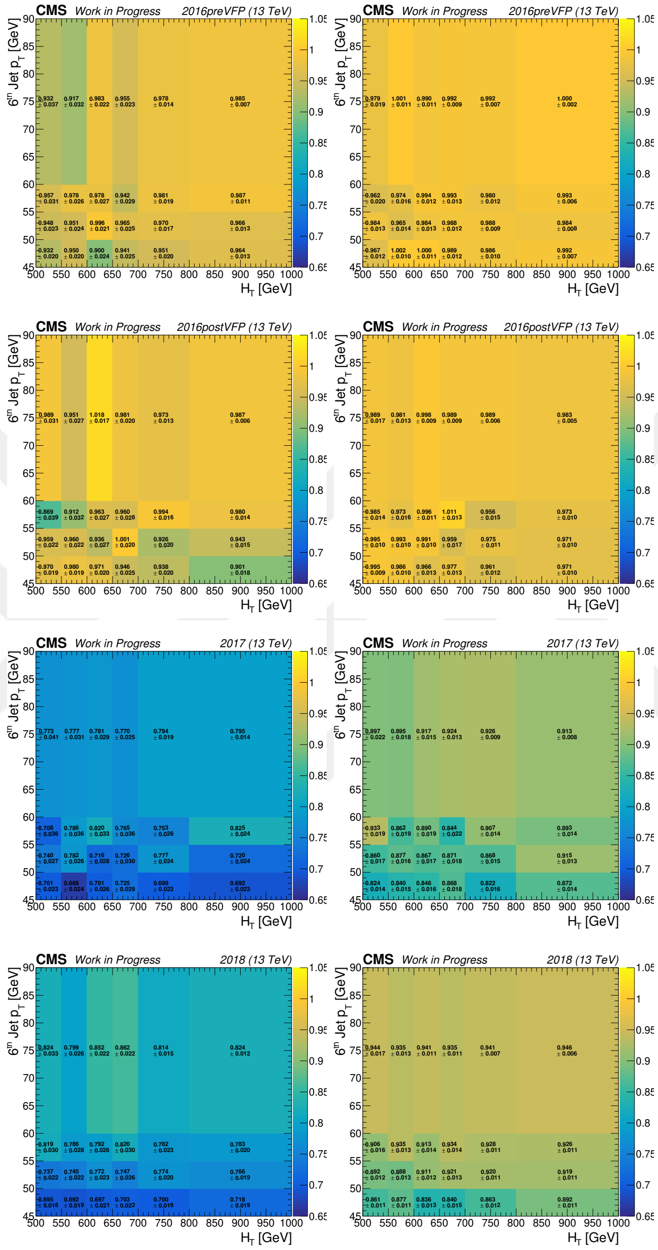


Figure 5.2. For the fully-hadronic channel, the jet trigger efficiency scale factor as a function of jet H_{T} and p_{T} of 6^{th} jet for 2016preVFP, 2016postVFP, 2017, 2018 (top to bottom) of full Run2UL and $N_{\mathrm{bjets}} = 1, \geq 2$ cuts (left to right)

Electrons

The electron trigger efficiency is calculated in the SingleMuon data set. At least one N_{muon}^{good} is required to register the SingleMuon data. Requiring one $N_{electron}^{good}$ is for measurements of the "electron" efficiency and scale factor. The preselcetions are given in Table 5.9. for both semi-leptonic and fully-leptonic channels.

Table 5.9. The preselections for measuring the electron trigger efficiencies and electron trigger scale factors

Preselections
SingleMuon data
Pass JetID
HEM 15/16 veto applied to 2018 data
$N_{muon}^{good} \geq 1$ ($p_{\mathrm{T}} > 40$ GeV)
$N_{electron}^{good} = 1$
$H_{\mathrm{T}} > 300$ GeV ($p_{\mathrm{T}} > 30$ GeV)
$N_{\mathrm{jets}} \geq 5$ ($p_{\mathrm{T}} > 30$)
$N_{\mathrm{bjets}} \geq 1$ ($p_{\mathrm{T}} > 30$ GeV, medium DeepFlavor)

Then the electron trigger efficiencies are measured as a function of electron p_{T} and η in $t\bar{t}+$jets simulation and in SingleMuon data. The measuremets of the efficiencies for indivual years of full Run2UL are shown in Figure 5.3.. The electron trigger efficiency scale factor as a function of electron p_{T} and η are given in Figure 5.4..

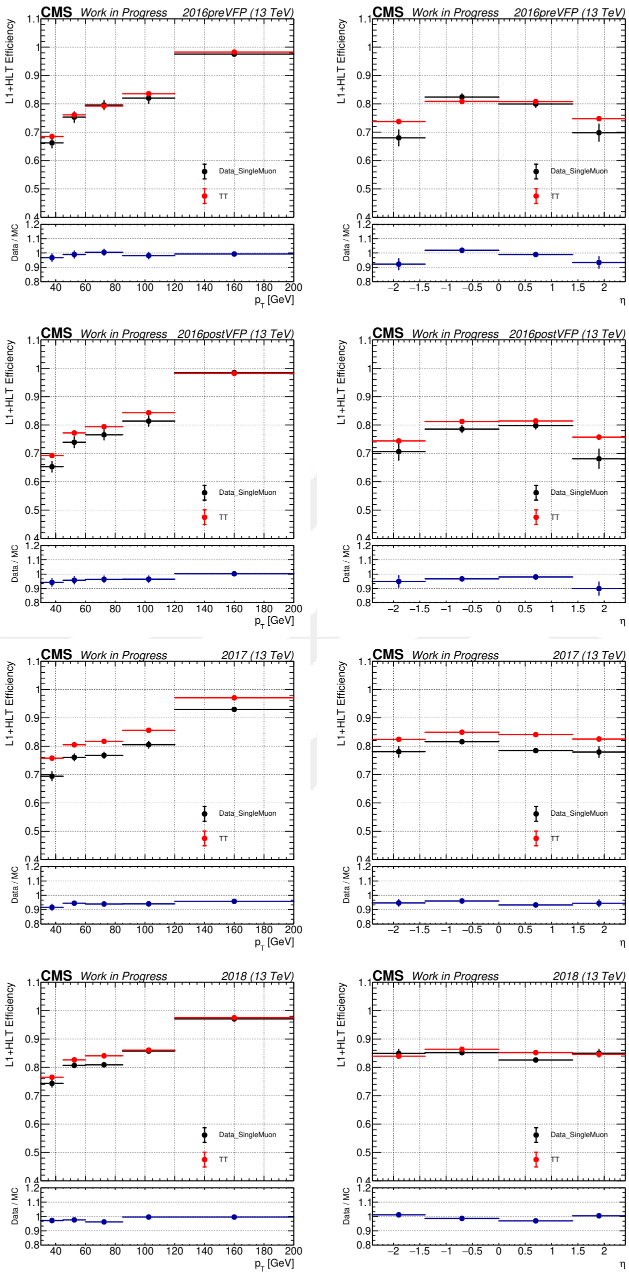


Figure 5.3. For the semi-leptonic and fully-leptonic channels, the electron trigger efficiency as a function of electron p_{T} (left) and η (right) in $t\bar{t}$+jets simulation and in SingleMuon data for 2016preVFP (top), 2016postVFP (middle top), 2017 (middle bottom), and 2018 (bottom) of full Run2UL

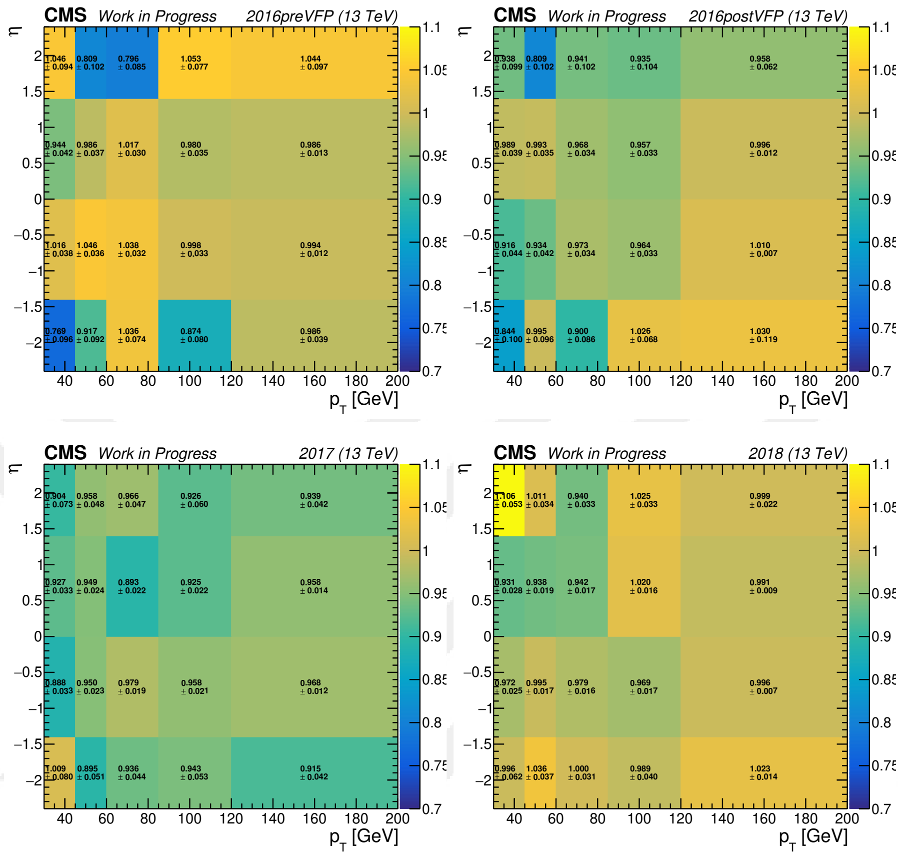


Figure 5.4. For the semi-leptonic and fully-leptonic channels, the electron trigger efficiency scale factor as a function of electron p_T and η for 2016preVFP (top left), 2016postVFP (top right), 2017 (bottom left), and 2018 (bottom right) of full Run2UL

Muons

The muon trigger efficiency is calculated in the SingleElectron data set. At least one $N^{good}_{electron}$ is required to register the SingleElectron data. Requiring one N^{good}_{muon} is for measurements of the "muon" efficiency and scale factor. The preselections are given in Table 5.10. for both semi-leptonic and fully-leptonic channels.

Table 5.10. The preselections for measuring the muon trigger efficiencies and muon trigger scale factors

Preselections
SingleElectron data
Pass JetID
HEM 15/16 veto applied to 2018 data
$N^{good}_{electron} \geq 1$ ($p_\mathrm{T} > 40$ GeV)
$N^{good}_{muon} = 1$
$H_\mathrm{T} > 300$ GeV ($p_\mathrm{T} > 30$ GeV)
$N_\mathrm{jets} \geq 5$ ($p_\mathrm{T} > 30$)
$N_\mathrm{bjets} \geq 1$ ($p_\mathrm{T} > 30$ GeV, medium DeepFlavor)

Then the muon trigger efficiencies are measured as a function of muon p_T and η in $t\bar{t}+$jets simulation and in SingleElectron data. The measuremets of the efficiencies for indivual years of full Run2UL are shown in Figure 5.5.. The muon trigger efficiency scale factor as a function of muon p_T and η are given in Figure 5.6..

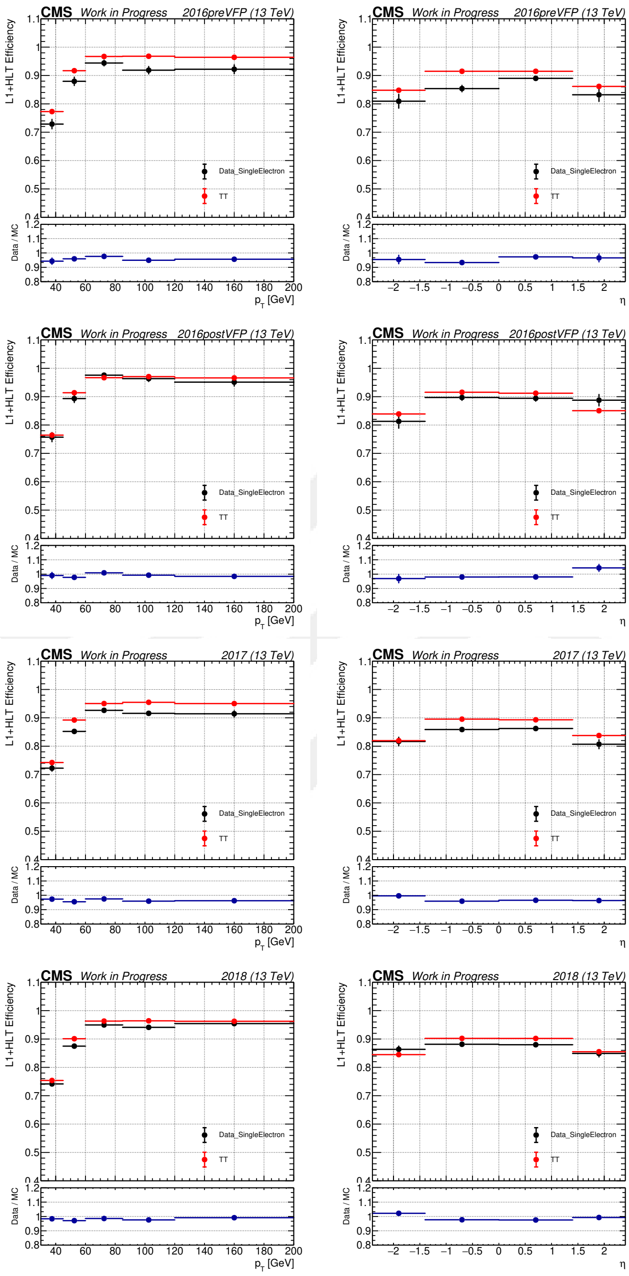


Figure 5.5. For the semi-leptonic and fully-leptonic channels, the muon trigger efficiency as a function of muon p_{T} (left) and η (right) in $t\bar{t}+$jets simulation and in SingleElectron data for 2016preVFP (top), 2016postVFP (middle top), 2017 (middle bottom), and 2018 (bottom) of full Run2UL

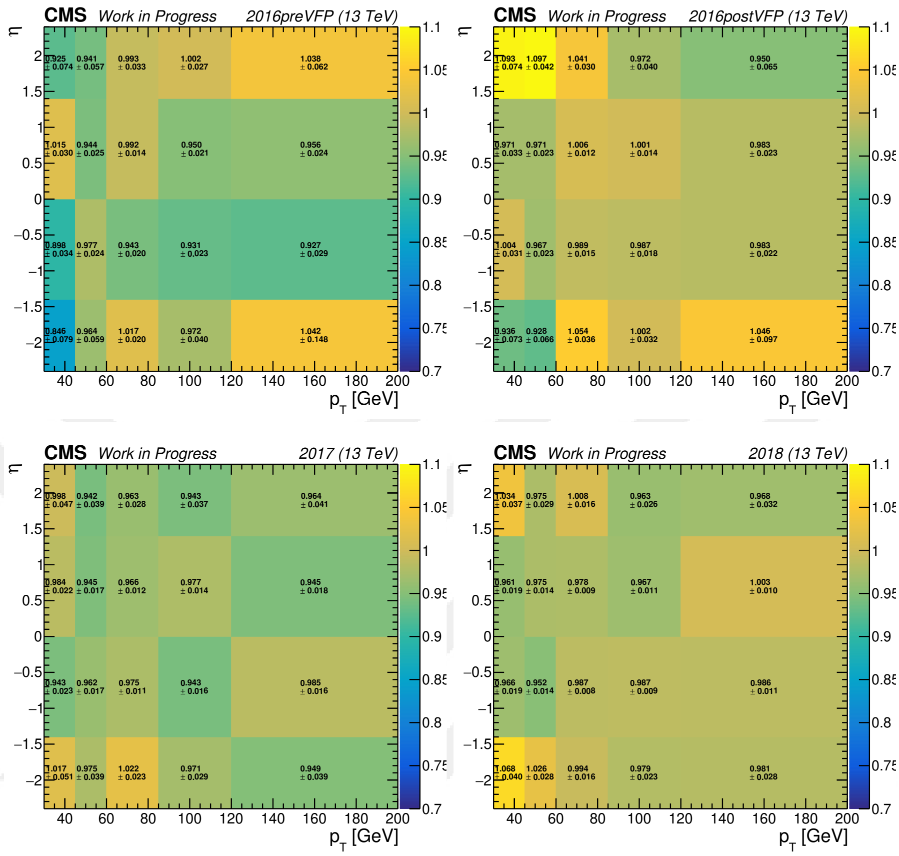


Figure 5.6. For the semi-leptonic and fully-leptonic channels, the muon trigger efficiency scale factor as a function of muon p_{T} and η for 2016preVFP (top left), 2016postVFP (top right), 2017 (bottom left), and 2018 (bottom right) of full Run2UL

5.3.2. top-tagging Scale Factors

Differences between the performance of the top tagger when used in data versus when used in simulation are accommodated by calculating and applying an associated scale factor. These data-simulation differences are parameterized in terms of two performance metrics: the overall tagging efficiency and the mistag rate of the tagger. Here, tagging efficiency refers to how often the tagger can identify a real top quark, whereas mistag rate is related to how often the tagger identifies a top quark when none is actually present.

For measuring the tagging efficiency and mistag rate in simulation and data, a unique event selection is defined for each performance metric. To test the tagging efficiency, an event selection is defined to preferentially select for $t\bar{t}+$jets events where one top quark decays leptonically and the other decays hadronically. These kind of events are selected for based on the properties of the leptonically decaying top where the selection specifically targets a muon as the leptonic decay product and an associated b quark. After requiring a muon and b jet in an event, several kinematic requirements are imposed on a b + muon system to try and identify the leptonically decaying top quark. With these semi-leptonic $t\bar{t}+$jets events identified, the performance of the top tagger can be measured by how well it identifies the hadronically decaying top in the event. The SingleMuon data set is used for this measurement.

Conversely, for measuring the mistag rate in simulation and data, an event selection is defined to preferentially select for QCD multi-jet events. The selections select for events with many jets, no leptons, and some amount of H_{T}. In these events, there are no real top quarks produced, and so any top identification by the tagger is considered a fake. The JetHT data set is used for this measurement. The exact selections used for defining both the semi-leptonic $t\bar{t}+$jets control region and QCD multi-jet control region are shown in Table 5.11..

Table 5.11. The selections for defining the semi-leptonic $t\bar{t}$ + jets control region (left) and QCD multi-jet control region (right). Definitions of objects and triggers follow from Section 5.2. and Section 5.3.1., respectively. Requirements involving a "loose b jet" refer to some other b-tagged jet (the required medium b jet trivially passes the loose WP) in the event that passes the loose Deep Flavour WP

Semi-leptonic $t\bar{t}$ + jets Control Region	**QCD multi-jet Control Region**
Single muon triggers	Jet triggers
$N_\mu = 1$ (with $p_\mathrm{T} > 30$ GeV, $\lvert\eta\rvert < 2.4$)	$N_e = 0$ (with $p_\mathrm{T} > 30$ GeV, $\lvert\eta\rvert < 2.4$)
$N_e = 0$ (with $p_\mathrm{T} > 30$ GeV, $\lvert\eta\rvert < 2.4$)	$N_\mu = 0$ (with $p_\mathrm{T} > 30$ GeV, $\lvert\eta\rvert < 2.4$)
$N_\mathrm{jets} \geq 6$ (with $p_\mathrm{T} > 30$ GeV, $\lvert\eta\rvert < 2.4$)	$N_\mathrm{jets} \geq 6$ (with $p_\mathrm{T} > 30$ GeV, $\lvert\eta\rvert < 2.4$)
$N_\mathrm{bjets} \geq 1$ (medium DeepFlavor)	$H_\mathrm{T} > 500$ GeV
$\Delta R(\mathrm{loose\ b\ jet, muon}) < 1.5$	
$30 \leq m(\mathrm{loose\ b\ jet, muon}) \leq 180$ GeV	
$\Delta\phi(E_T^{miss}, \mathrm{muon}) < 0.8$	
$m_T(\mathrm{muon}, E_T^{miss}) < 100$ GeV	

The measurement of tagging efficiency and mistag rate proceeds by selecting all the simulation events and appropriate data events that pass the respective selections in Table 5.11.. For each event, the top tagger, which runs the merged tagger and then the resolved tagger, is used to tag any possible top candidates. From all possible candidates (which necessarily pass any selections described for DeepResolved and DeepAK8 algorithms in Section 6.1.1.), the "best" top candidate is chosen as the one with a mass closest to 173.2 GeV, and could either be a merged top tag or a resolved top tag. The tagging efficiency for the relevant WPs, is defined as the fraction of events with a "best" top candidate with a discriminator value above the chosen WP relative to all events with a top candidate. Likewise, for the mistag rate, the ratio is number of events with a top candidate (by definition fake) having a discriminator above the given working point to the number of events with a "best" top candidate. Both of these quantities are separately computed for exclusively merged top tags and resolved top tags. Note that WPs are given for tagging each resolved and merged top candidates are given in Table 5.6.. The governing equations for the tagging efficiency and mistag rate are given in Equation 5.3 and Equation 5.4, respectively.

$$\text{Tagger Efficiency} = \frac{\text{Events passing semi-leptonic } t\bar{t} \text{ selection with best top candidate and disc.} > \text{ WP}}{\text{Events passing semi-leptonic } t\bar{t} \text{ selection with best top candidate}} \tag{5.3}$$

$$\text{Mistag Rate} = \frac{\text{Events passing QCD multi-jet selection with best top candidate and disc.} > \text{ WP}}{\text{Events passing QCD multi-jet selection with best top candidate}} \tag{5.4}$$

With these quantities measured for simulation events and data events, the scale factor is simply the tagging efficiency and mistag rate for data divided by the corresponding quantity for simulation as shown in Equation 5.5 and Equation 5.6.

$$\text{Tagging Efficiency SF} = \frac{\text{Tagging Efficiency in Data}}{\text{Tagger Efficiency in Simulation}} \tag{5.5}$$

$$\text{Mistag Rate SF} = \frac{\text{Mistag rate in Data}}{\text{Mistag rate in Simulation}} \tag{5.6}$$

It is decided to parameterize both the tagging efficiency and mistag rate as a function of the top tag candidates transverse momentum. Figure 5.7. and Figure 5.8. show the tagging efficiency and mistag rate, respectively, for simulation and data as well as their ratio as a function of top candidate p_{T} for events with $N_{\mathrm{jets}} \geq 6$. When calculating the tagging efficiency, the conglomerate "TT MC" consists of simulation events from semi-leptonic $t\bar{t}+$jets, single top, $t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$. All other simulation events (DY + jets, W + jets, Diboson, Triboson, QCD multijet, and rare processes) are subtracted from data. Conversely, when calculating the mistag rate, the conglomerate

"Non-TT MC" consists of DY + jets, W + jets, Diboson, Triboson, QCD multijet, and rare processes. The remaining top and $t\bar{t}$ simulated processes are then subtracted from data.

Consistent performance is seen across the different years with the top tagging efficiency scale factor in the range 0.7 to 1.0 for resolved top tags and on the order of unity for merged top tags. The same consistency across the years is also seen for the top mistag rate scale factor, which ranges from 0.85 to 1.25 for resolved top tags and around 1.2 to 1.3 for merged top tags.

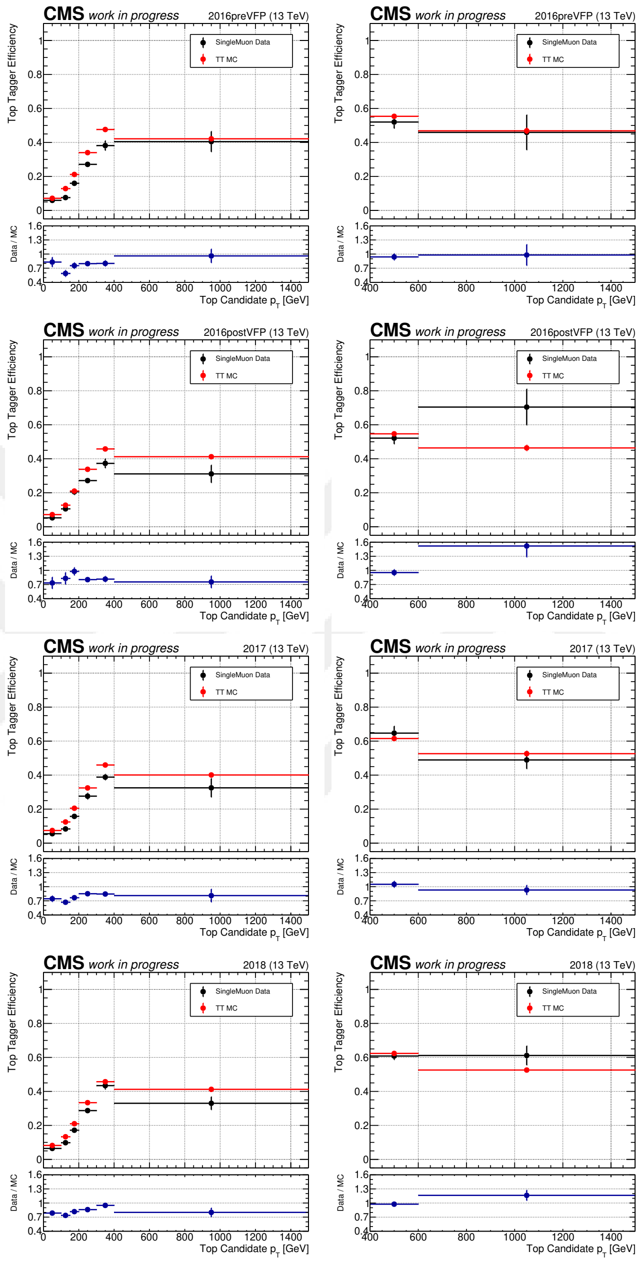


Figure 5.7. The top tagger tagging efficiency in data (black) and simulation (red), and accompanying scale factor (blue) for resolved (left) and merged (right) top candidates, as a function of the candidates transverse momentum. Results are for 2016preVFP (top), 2016postVFP (middle-top), 2017 (middle-bottom), and 2018 (bottom) of full RunUL

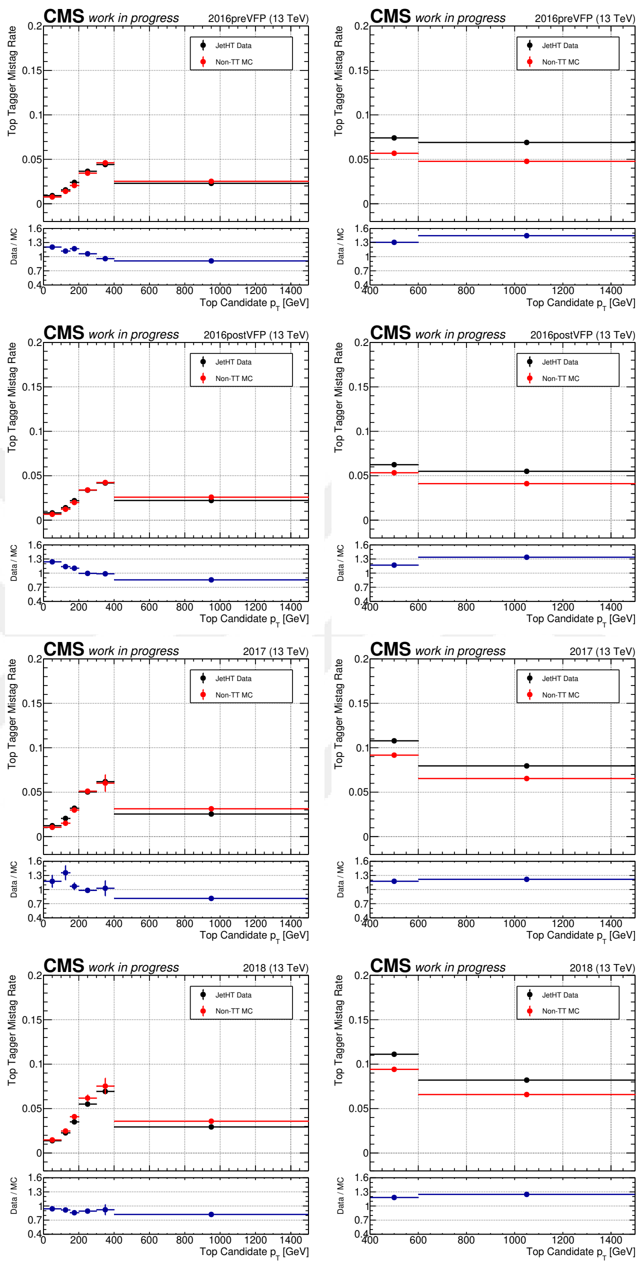


Figure 5.8. The top tagger mistag rate in data (black) and simulation (red), and accompanying scale factor (blue) for resolved (left) and merged (right) top candidates, as a function of the candidates transverse momentum. Results are for 2016preVFP (top), 2016postVFP (middle-top), 2017 (middle-bottom), and 2018 (bottom) of full RunUL

5.4. Event Selections

5.4.1. Signal Region (SR) Selections

Fully-Hadronic Channel: Each event is required to have exactly zero good leptons, as defined in Section 5.2., to make fully-hadronic event selection orthogonal to the semi-leptonic and fully-leptonic selections. Those events which pass the zero lepton requirement are also required to pass the jet triggers denoted in Table 5.7.. Data events must be part of the JetHT data set. For all MC studies and data-MC comparisons, simulated events must also pass the MC emulation of the corresponding triggers. For 2018 data, due to known data quality issues related to a hardware failure in the region of HEM 15/16, a veto is applied to events with an electron in the corresponding detector region.

Furthermore, required exactly zero non-isolated muon to make the signal region (SR) selection also orthogonal to the QCD control region (QCD CR) selection. Additionally, the presence of at least eight jets is required in each event with $p_T > 30$ GeV (at least six of them with $p_T > 45$ GeV based on trigger used) and all with $|\eta| < 2.4$. All of these jets in a candidate event must pass the JetID as defined by the JetMET POG (CMS Collaboration, 2022b). There is also at least two medium DeepFlavor b-tagged jets requirement in each event with $p_T > 30$ GeV (at least one of them with $p_T > 45$ GeV based on trigger used). Then, there is a requirement for at least two top-tagged jets to implement the hadronic top tagger neural network with the WP 0.95 for reducing the QCD multi-jet background. The further details of the hadronic top tagger can be seen in Section 6.. The ΔR between the b-jets is required to be greater than or equal to 1.0 for also reducing the QCD multi-jet background. The ΔR cut also requires at least two medium DeepFlavor b-tagged jets with $p_T > 30$ coming from the decay of the top quarks.

To further eliminate QCD multi-jet events and ensure that the analysis is not impacted by missing QCD multi-jet event simulation, the H_T of an event should be greater than 500 GeV, where H_T is computed using all jets with $p_T > 30$ GeV and $|\eta| < 2.4$. All these selections are listed in Table 5.12..

Table 5.12. The list of the SR selections for the fully-hadronic channel

Fully-hadronic channel
JetHT data used
Jet triggers applied
HEM 15/16 veto applied to 2018 data
Jet trigger and top-tagging SFs locally computed
Pass JetID
$N^{good}_{leptons} = 0$
$N^{non-iso}_{muon} = 0$
$H_{\mathrm{T}} > 500$ ($p_{\mathrm{T}} > 30$ GeV)
$N_{\mathrm{jets}} \geq 8$ ($p_{\mathrm{T}} > 30$ GeV, $\geq$6 of them with $p_{\mathrm{T}} > 45$ Gev)
$N_{\mathrm{bjets}} \geq 2$ ($p_{\mathrm{T}} > 30$ GeV, $\geq$1 of them with $p_{\mathrm{T}} > 45$ GeV)
$N_{tops} \geq 2$ (for hadronic top tagging)
$\Delta R_{b\,jets} \geq 1$ (req. $N_{\mathrm{bjets}} \geq 2$ with $p_{\mathrm{T}} > 30$ GeV)

Figure 5.9. shows some data and simulation comparisons, which are N_{jets}, N_{bjets}, N_{tops}, mass distribution of leading p_{T} top, $p_{\mathrm{T}}/\ H_{\mathrm{T}}$ distribution of leading p_{T} jet, and ΔR distributions between any two b-tagged jets . These comparisons are for the lowest two N_{jets} = 8,9 bins, since the analysis is still blinded to for high N_{jets} bins. Generally, there is a good agreement between data and simulation. Disagreements in modeling of the H_{T} and the jet p_{T} distributions were observed, but the $p_{\mathrm{T}}/H_{\mathrm{T}}$ is well modeled. So, $p_{\mathrm{T}}/\ H_{\mathrm{T}}$ distribution are shown instead H_{T} and there is a good agreement between data and simulation.

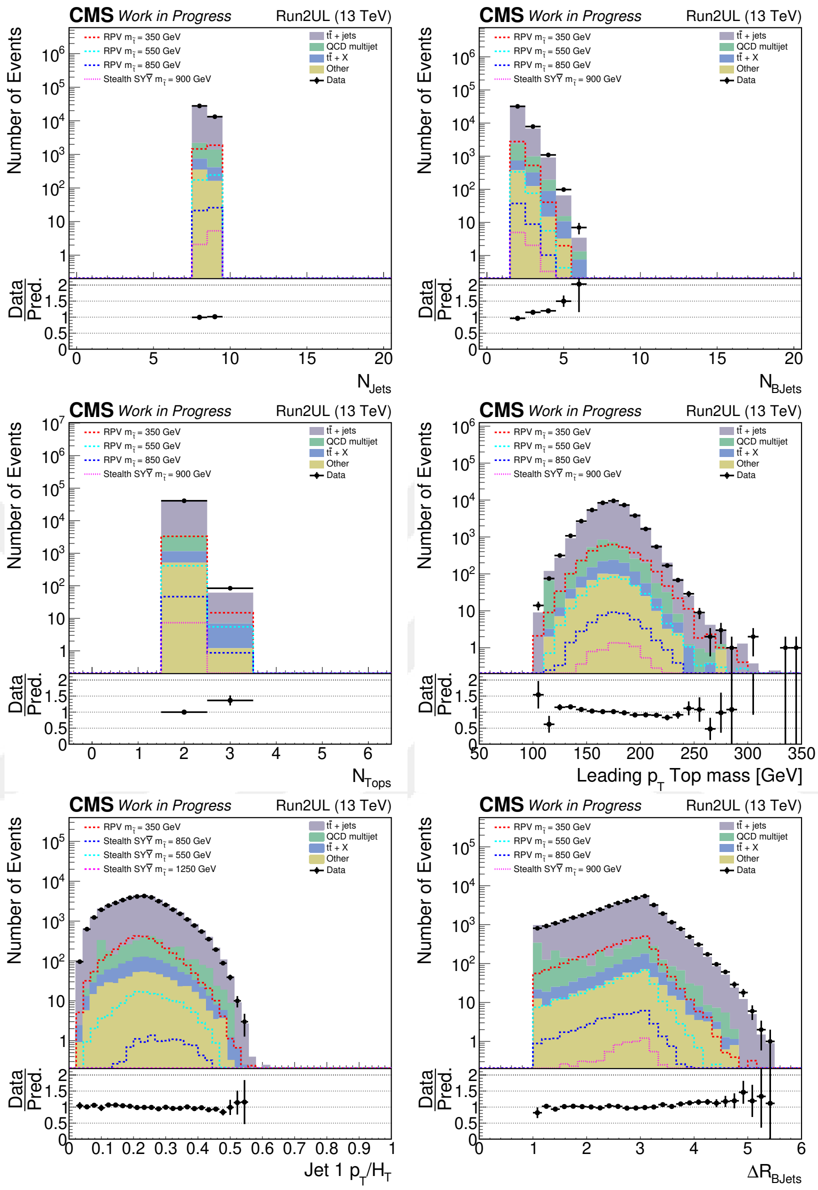


Figure 5.9. For the fully-hadronic channel, N_{jets} (top left), N_{bjets} (top right), N_{tops} (middle left), top mass (middle right), p_T/H_T (bottom left) and ΔR_{bjets} (bottom right) distributions for Run2UL. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield

Semi-Leptonic Channel: Events are required to have the presence of exactly one good electron or good muon. Those events with exactly one good electron and exactly zero good muons are also required to pass the single electron trigger set as described in Table 5.7.. Data events are additionally required to be part of the SingleElectron/EGamma data set. For events with exactly one good muon, analogous requirements are imposed: one good muon, zero good electrons, and the passing of the muon triggers for the appropriate year as listed in Table 5.7.. Data events must be part of the SingleMuon data set. For all MC studies and data-MC comparisons, simulated events must also pass the MC emulation of the corresponding triggers. For 2018 data, due to known data quality issues related to a hardware failure in the region of HEM 15/16, a veto is applied to events with an electron in the corresponding detector region.

Furthermore, a requirement of exactly zero non-isolated muons is made to introduce orthogonality between the SR selection and QCD CR selection. To select for signal, events are required to have at least seven jets, each with p_T greater than 30 GeV and $|\eta| < 2.4$. All of these jets in a candidate event must pass the JetID as defined by the JetMET POG (ibid.). There is also a requirement for events to have at least one medium DeepFlavor b-tagged jet (coming from the decay of the top quark) with p_T greater than 30 GeV. To further select for signal and the dominant background of $t\bar{t}+$jets, a requirement that the b-tagged jet have an invariant mass with the good lepton ($M(b,l)$) between 50 and 250 GeV is imposed. Physically, this acts as a loose semi-leptonic top tag because the b-tagged jet and the good lepton are expected to come from the same top quark. In the case that there is more than one b-tagged jet in the event, the $M(b,l)$ value closest to 105 GeV is chosen.

Finally, events passing all the above criteria are required to have an event H_T greater than 500 GeV to further eliminate QCD multi-jet events and ensure that the analysis is not impacted by missing QCD multi-jet event simulation at low H_T. H_T as defined in this analysis is computed using all jets with $p_T > 30$ GeV and $|\eta| < 2.4$. All these selections are listed in Table 5.13..

Table 5.13. The list of the SR selections for the semi-leptonic channel

Semi-leptonic channel
SingleElectron, SingleMuon data sets used
Electron and muon triggers applied
HEM 15/16 veto applied to 2018 data
Electron, muon trigger SFs locally computed
Pass JetID
$N^{good}_{leptons} = 1(e/\mu)$
$N^{non-iso}_{muon} = 0$
$H_T > 500$ ($p_T > 30$ GeV)
$N_{jets} \geq 7$ ($p_T > 30$ GeV)
$N_{bjets} \geq 1$ ($p_T > 30$ GeV)
$50 < M_{b,l} < 250$ GeV (for loose leptonic top tagging)

Figure 5.10. some data and simulation comparisons, which are N_{jets}, N_{bjets}, p_T/ H_T distribution of leading p_T jet and $M_{b,l}$ distributions. These comparisons are for the lowest two N_{jets} = 7,8 bins, since the analysis is still blinded to for high N_{jets} bins. There is a good agreement between data and simulation.

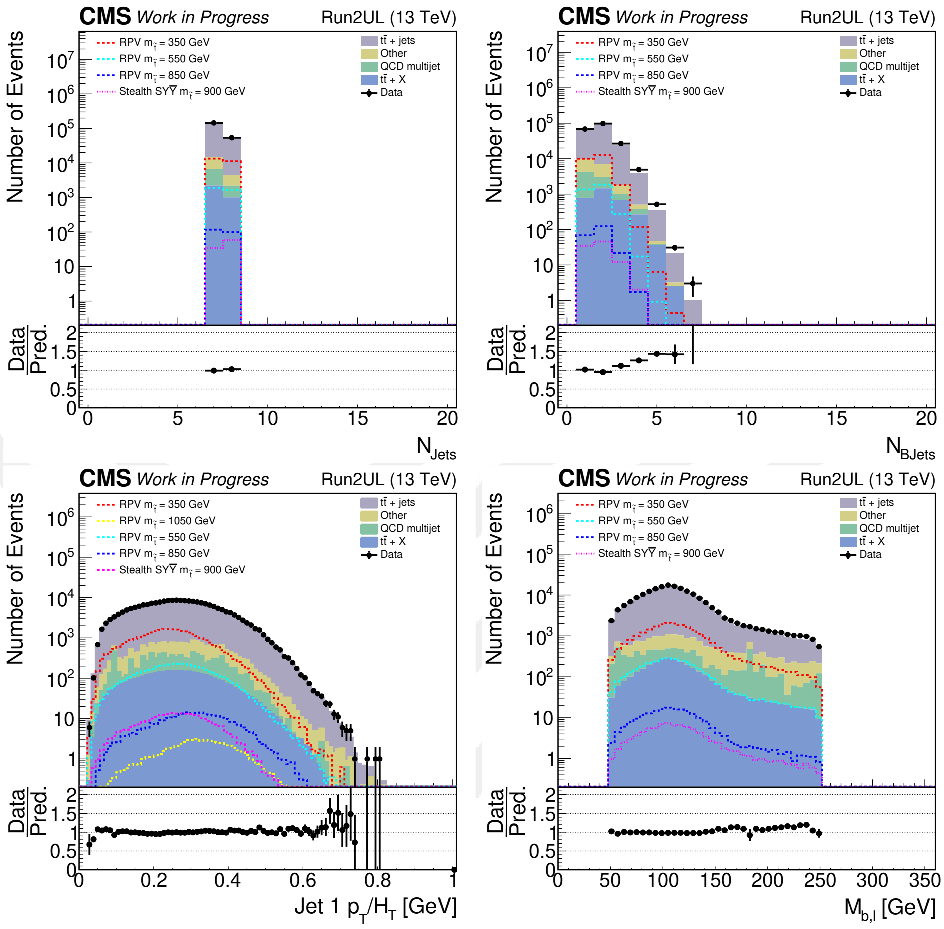


Figure 5.10. For the semi-leptonic channel, N_{jets} (top left), N_{bjets} (top right), $p_{\mathrm{T}}/\ H_{\mathrm{T}}$ (bottom left) and $M_{\mathrm{b,l}}$ (bottom right) distributions for Run2UL. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield

Fully-Leptonic Channel: Each event is required to have exactly two oppositely-charged good leptons (any combination of muon and electron). Those events which pass the two lepton requirement are also required to pass electron or muon triggers denoted in Table 5.7.. Data events are additionally required to be part of the SingleMuon or SingleElectron data set. To avoid overlap between the two data sets, all muon-muon and electron-muon events are only selected from the SingleMuon data set. Likewise, if exactly both leptons are electrons, then the event is only selected from the SingleElectron data set. For all MC studies and data-MC comparisons, simulated events must also pass the MC emulation of the corresponding triggers. For 2018 data, due to known data quality issues related to a hardware failure in the region of HEM 15/16, a veto is applied to events with an electron in the corresponding detector region.

Furthermore, a requirement of exactly zero non-isolated muons is made to introduce orthogonality between the SR selection and QCD CR selection. The presence of at least six jets is required in each event with each jet possessing $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.4$. There is also a requirement for at least one of these jets to be tagged as a medium DeepFlavor b-tagged jet. All of these jets in a candidate event must pass the JetID as defined by the JetMET POG (ibid.). To further eliminate QCD multi-jet events and ensure that the analysis is not impacted by missing QCD multi-jet event simulation, the H_{T} of an event should be greater than 500 GeV, where H_{T} is computed using all jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.4$.

Lastly, to minimize the contribution from the DY + jets background, a requirement is made to reject events where the two leptons form an invariant mass ($M_{l,l}$) near that of the Z with $81 < M_{l,l} < 101$ GeV being the criteria to veto events. This requirement only applies when both leptons are of the same flavor. All these selections are listed in Table 5.14..

Table 5.14. The list of the SR selections for the fully-leptonic channel

Fully-leptonic channel
SingleElectron, SingleMuon data sets used
Electron and muon triggers applied
HEM 15/16 veto applied to 2018 data
Electron, muon trigger SFs locally computed
Pass JetID
$N_{leptons}^{good}$ = 2(2e/2μ/1e1μ), oppositely charged
$N_{muon}^{non-iso} = 0$
$H_{\mathrm{T}} > 500$ ($p_{\mathrm{T}} > 30$ GeV)
$N_{\mathrm{jets}} \geq 6$ ($p_{\mathrm{T}} > 30$ GeV)
$N_{\mathrm{bjets}} \geq 1$ ($p_{\mathrm{T}} > 30$ GeV)
$M_{l,l} < 81 \;\|\|\; M_{l,l} > 101$ GeV (for vetoing Z mass)

Figure 5.11. some data and simulation comparisons, which are N_{jets}, N_{bjets}, p_{T}/ H_{T} distribution of leading p_{T} jet and $M_{l,l}$ distributions. These comparisons are for the lowest two N_{jets} = 6,7 bins, since the analysis is still blinded to for high N_{jets} bins. There is a good agreement between data and simulation excepting an empty bin in $M_{l,l}$ distribution based on the Z mass vetoing.

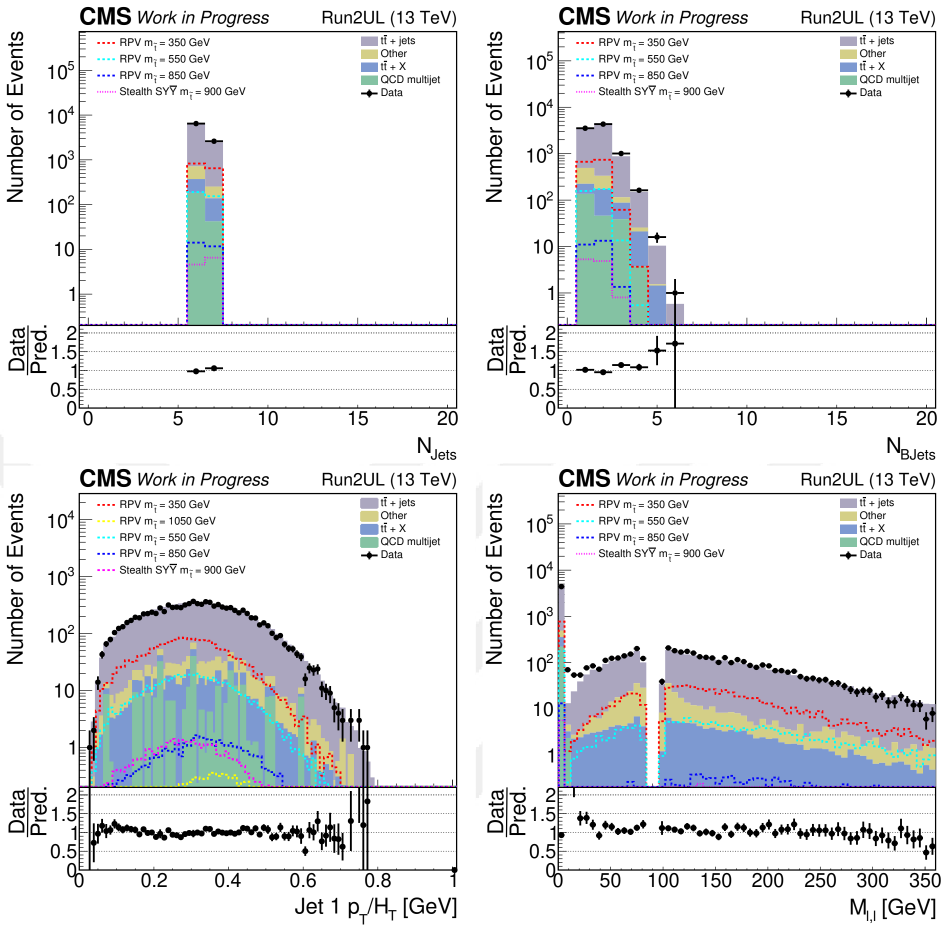


Figure 5.11. For the fully-leptonic channel, N_{jets} (top left), N_{bjets} (top right), H_{T} (bottom left) and $M_{\text{l,l}}$ (bottom right) distributions for Run2UL. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield

5.4.2. QCD Control Region (QCD CR) Selections

Two QCD CR selections are defined orthogonal to the three signal regions specified above, given in Table 5.15.. The two different selections are used to estimate the QCD multi-jet background for the fully-hadronic and semi-leptonic signal regions. The fully-leptonic channel does not have a measurable amount of QCD multi-jet background thus a background estimation is not needed. The overall selections for both of the control regions are almost identical with some changes to the jet selection to match the kinematics of their respective signal regions.

For these QCD multi-jet selections, events are required to have the presence of exactly zero good leptons. At the same time, an event should contain exactly one non-isolated muon, which is defined in Section 5.2.2.. These two types of lepton requirements guarantee orthogonality with the signal region selections. Data events are additionally required to be part of the SingleMuon data set and pass some single muon triggers. For all years, events must pass the HLT_Mu50 trigger path, while for 2016, events may alternatively pass the HLT_TkMu50. For all MC studies and data-MC comparisons, simulated events must pass the MC emulation of this trigger. Events are required to have at least seven or eight jets, each with p_{T} greater than 30 GeV and $|\eta| < 2.4$, where the lepton cleaning procedure only considers non-isolated muons. Additionally, the QCD CR specific to the fully-hadronic channel requires six jets with p_{T} greater than 45 GeV. All of these jets in a candidate event must pass the JetID as defined by the JetMET POG (ibid.). To keep these QCD CRs close to the signal region, the H_{T} of an event should be greater than 500 GeV, where H_{T} is computed using all non-isolated-muon-cleaned-jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.4$ and are cleaned from non-isolated muons.

Data-simulation comparisons for each of QCD CR selections, given in Table 5.15., are shown in Figure 5.12. for the fully-hadronic semi-leptonic channels.

Table 5.15. The two QCD CR selections used for the fully-hadronic and semi-leptonic QCD multi-jet background estimation

Fully-Hadronic	**Semi-Leptonic**
SingleMuon data set used	SingleMuon data set used
Non-isolated muon triggers applied	Non-isolated muon triggers applied
Pass JetID	Pass JetID
$N^{good}_{lepton} = 0$	$N^{good}_{lepton} = 0$
$N^{non-iso}_{muon} = 1$ ($p_{\mathrm{T}} > 55$ GeV)	$N^{non-iso}_{muon} = 1$ ($p_{\mathrm{T}} > 55$ GeV)
$H_{\mathrm{T}} > 500$ GeV ($p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.4$)	$H_{\mathrm{T}} > 500$ GeV ($p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.4$)
$N_{\mathrm{jets}} \geq 8$ ($p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.4$)	$N_{\mathrm{jets}} \geq 7$ ($p_{\mathrm{T}} > 30$ GeV, $\lvert\eta\rvert < 2.4$)
$N_{\mathrm{jets}} \geq 6$ ($p_{\mathrm{T}} > 45$ GeV, $\lvert\eta\rvert < 2.4$)	

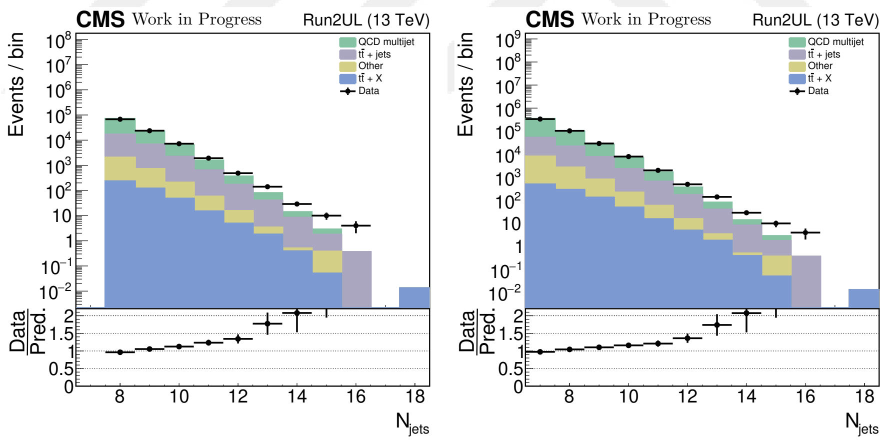


Figure 5.12. For the fully-hadronic (left) and semi-leptonic (right) channels, data-simulation comparison for QCD CR selections. These plots contain data and simulation from all of the four data taking years combined as Run2UL. The total background yield in simulation is normalized to the total data yield

6. HADRONIC TOP TAGGING

The identification (tagging) of the hadronically decaying heavy objects (top quarks and W, Z and Higgs bosons) is an important component in SM searches and in many extensions thereof. The decay products of these heavy objects are hard to reconstruct, because of separating them from the vast amount of background, e.g. coming from QCD multi-jet production. There is a tagger, which was developed for a SUSY analysis (Sirunyan, 2021a), to tag those top objects. For the fully-hadronic channel, based on the signal topology including $t\bar{t}+$jets in the final state, the hadronic top tagger is implemented as a part of the SR selection to reduce specifically the QCD multi-jet background.

The top quark is one of the heaviest fundamental particle. Its phenomenology is driven by its large mass and because of this, the top quark is extremely short-lived. It essentially always decays as $t \rightarrow bW$. The top quark, being heavier than a W boson, is the only quark that decays semi-weakly into a real W boson and a b quark before hadronization occurs with 0.91$\pm$0.04 branching ratio. In addition, it is the only quark whose Yukawa coupling to the Higgs boson is order of unity. For these reasons, the top quark plays a special role in many researches for new physics. An accurate knowledge of its properties (mass, couplings, production cross section, decay branching ratios, etc.) can bring key information on fundamental interactions at the electroweak breaking scale and beyond.

The hadronic decays of top quarks are where $W^{\pm}$ bosons decay through $W^{\pm} \rightarrow q\bar{q}$ and the leptonic decays of the top quarks are where $W^{\pm}$ bosons decay through $W^{\pm} \rightarrow l\nu$. The top quark decays are illustrated in Figure 6.1.. Regarding to the fully-hadronic channel, the leptonic decays of top quarks are ignored in this tagger. Combination of two quark jets from the W and the b quark jet can be easily identified as a fake top candidate. So, the characteristics of a top trijet system, which includes exactly one b jet and an expected resonant mass of known value, are exploited by tagging algorithms which try to identify top candidates in an event. For this reason, the invariant mass of the top quark and b-tagging information are applied as a part of the Hadronic Top Tagger.

The details of the hadronic top tagger is given in Section 6.1.. The following section, Section 6.2., is a dedicated study to optimization study for one of the tagging algorithm.

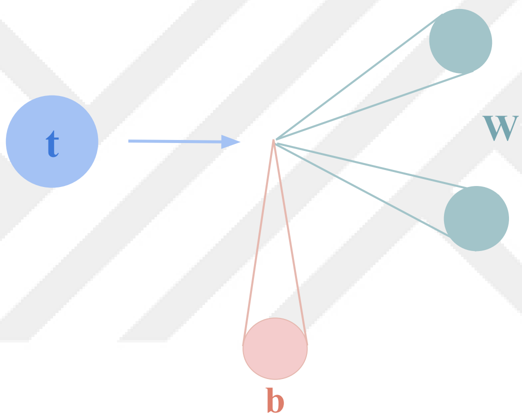


Figure 6.1. Top quark decays, $t \to bW$. The hadronic decays of top quarks where $W^{\pm} \to q\bar{q}$. The leptonic decays of top quarks where $W^{\pm} \to l\nu$

6.1. Hadronic Top Tagger

As the LHC is a high energy collider leading to final states with boosted top quarks, it is desirable to tag top quarks in all p_{T} ranges. The hadronic top tagger is a NN including a combination of two approaches: one where is based on the top trijet system using low boosted regime and another where all the decay products are collimated together using high boosted regime. These two algorithms are run consecutively, with high boosted regime running first, where constituents assigned to viable candidates can only be counted in a top candidate once. Also, there are special considerations in the hadronic top tagger to allow combinations of both the AK4 and AK8 jets within the same event. As an example of the tagging algorithms, Figure 6.2. shows the resolved and merged top decays based on the boosting regimes.

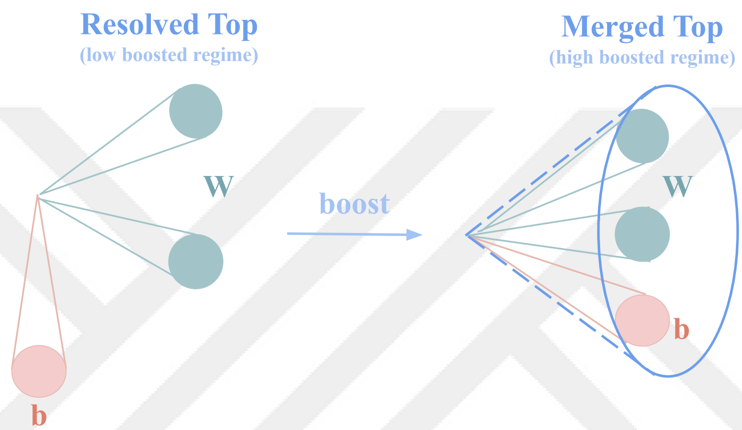


Figure 6.2. Resolved and merged top decays based on the boosting regimes. Resolved top consists of three individual AK4 jets. Merged top consists of three jets which are clustered into one AK8 jet

6.1.1. Tagging Algorithms

DeepResolved Tagging Algorithm

DeepResolved tagging algorithm (CMS HOT Working Group, 2019) is designed to tag low to medium p_T top quarks in the range 0 to 500 GeV. In this p_T regime, the top quark decay products will be three indivudual AK4 jets. These indivudual AK4 jets can be highly separated in the detector and may not be contained inside a single AK8 jet. This approach suffers from a very large background of fakes arising from trijet combinations. In order to reduce the effect of this large background, a NN is used to distinguish trijet combinations which match to a top quark versus those which do not.

The algorithm is broken down into several steps based on using combinations of three AK4 jets. First, a loose preselection on AK4 jets are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.8$. Then, to create the top candidates, three indivudual AK4 jets are selected by requiring (a) $p_{\mathrm{T}} > 40, 30, 20$ GeV on the three jets respectively, (b) the trijet invariant mass must be between 100 and 250 GeV and (c) the three jets must all lie within ΔR < 3.14 of the total summed momentum of the three jets. Finally, the algorithm must remove any overlap arising from top candidates which may share the same jets with DeepAK8 algorithm.

The final list of top quark candidates is then found by placing a requirement on the neural network discriminator. If at least one top quark candidate has a discriminator value above a threshold (0.92 - original one), then this event has a resolved top tag. In addition to the discriminator value, an eta cut ($|\eta| < 2.0$) on the trijet system and no more than one of the three jets be tagged as a b jet by the medium WP of the DeepCSV algorithm (original one) are applied to reduce fakes.

DeepAK8 Tagging Algorithm

DeepAK8 tagging algorithm (CMS JetMET Group, 2019) is designed to tag high p_{T} top quarks in the range above 400 GeV. The top quark decay products become more collimated as the top quark p_{T} increases and they are clustered into one AK8 jet. Like DeepResolved, DeepAK8 is a NN classifier algorithm, and utilizes many features of each of the subjets and total AK8 jet.

This algorithm imposes some basic selections on AK8 jets when classifying. A soft-drop mass (Larkoski et al., 2014) cut of $105 < M_{SD} < 210$ GeV are applied on top quark candidates (CMS JetMET Group, 2020). In addition, an eta cut ($|\eta| < 2.0$) on the trijet system to reduce contributions from high η fakes with little effect on the efficiency.

6.1.2. Working Point (WP) and ROC Curve for Performance

A discriminator value is output of a NN which is a binary number, given as “0” or “1”. As the hadronic top tagger tags the top quarks, the discriminator value can be “1” if the top quark is tagged and the discriminator value can be “0” if the fake top quark is tagged.

To test the performance of the tagger, a selection is made on this discriminator value to define the **WP**, where the highest tagging effiency with fewer fake rate. The Table 6.1. shows the WPs based on each DeepResolved and DeepAK8 algorithms are required for each year. These WPs are based on the original version of the hadronic top tagger. DeepResolved tagger also requires one of the decay products of top quark as b-jet by implementing “medium” DeepCSV WP.

Table 6.1. WPs for each algorithm corresponding to each year. The top WP for DeepResolved tagger is threshold value based on the original version of the tagger. DeepResolved tagger also requires one of the decay products of top quark as b-jet by implementing “medium” DeepCSV WP

	2016	**2017**	**2018**
DeepResolved	WP	WP	WP
top	0.92	0.92	0.92
b-jet	0.63	0.49	0.41
DeepAK8	WP	WP	WP
top	0.937	0.895	0.895

The general performance of the top tagger can be seen by looking at Receiving Operating Characteristic (ROC) curve. The ROC curve is a parametric curve plotted as a function of the WP. Efficiency and fake rate of the top tagger are plotted on independent axes as the values of the discriminator. For a perfect tagger, efficiency would be 1 and fake rate would be 0 for the entire range of discriminator values. The tagger performance is also reviewed in terms of generator level efficiency and fake rate. These quantities are defined as:

$$\text{Efficiency} = \frac{\text{Number of generator-level top quarks matched to a tagged top}}{\text{Number of generator-level top quarks in an event}} \tag{6.1}$$

$$\text{Fake Rate} = \frac{\text{Number of events having at least one tagged top}}{\text{Total number of events}} \tag{6.2}$$

The ”tagged top” represents the trijet combination (top candidate) with the discriminator value above the given threshold. A top candidate is matched with a generator-level top quark within a cone of $\Delta R < 0.6$ in (η, ϕ) space.

The effeiciency and fake rate can also be described as ”true positive rate” (TPR) and ”false positive rate” (FPR) respectively. TPR is defined as the number of true gen matched tops which could be tagged passing the discriminator cuts. FPR is defined as the fraction of top candidates which are not gen matched passing the discriminator cuts. These quantities are defined as:

$$\text{TPR} = \frac{\text{Number of tagged top candidates matched to a generator-level top quark}}{\text{Number of top candidates matched to a generator-level top quark}} \tag{6.3}$$

$$\text{FPR} = \frac{\text{Number of tagged top candidates not matched to a generator-level top quark}}{\text{Number of top candidates not matched to a generator-level top quark}} \tag{6.4}$$

6.2. Working Point Optimization for DeepResolved Tagger

Based on a dedicated study for fully-hadronic channel given as in Appendix A, there are more resolved top quarks in the signal region for low mass points of top squark whereas more merged top quarks for high mass points of top squark. The DeepResolved tagger is not public used in the CMS, since it was developed for a specific SUSY analysis (Sirunyan, 2021a). So, a dedicated WP optimization study is done for DeepResolved tagger to have better efficieny to tag the hadronically decaying top quarks in the signal region. Also, to get better signal-background discrimination in the Double Disco NN trainings—because four-vector of reconstructed top quarks are fed into the Double Disco NN, more details given in Section 7.. The following studies in this section are based on the WP optimization for DeepResolved Tagger using only simulated events.

The discriminator value 0.92 is threshold value based on the original version of the DeepResolved tagger. It is compared with tighter WPs which have corresponding discriminator values of 0.95, 0.96, 0.97, 0.98 in order to determine the optimal one to be used in the further steps of the fully-hadronic analysis. For this study, first a region is defined, which is very similar to the signal region to measure efficiency and fake rate. This region is also used to measure TPR and FPR. The selection criterias for this region are given in Table 6.2..

The efficiency is measured as a function of generator-level top quark p_{T} obtained from simulated $t\bar{t}$ events. A fake rate calculation requires the presence of no hadronically decaying top quarks in the events. Therefore, it is measured as a function of N_{jets} using simulated QCD events. Also, ROC curves are made. To make the ROC curve, the TPR is measured in the $t\bar{t}$ simulated events and the QCD sample is chosen for FPR calculation. TPR and FPR are calculated for 100 discriminator values between 0 to 1 with the constant difference of 0.01.

Table 6.2. The selection criterias to measure efficiency, fake rate, TPR and FPR

Selections
Jet triggers applied
Pass JetID
$N_{leptons}^{good} = 0$
$N_{muon}^{non-iso} = 1$
$H_T > 500$ ($p_T > 30$ GeV)
$N_{jets} \geq 8$ ($p_T > 30$ GeV, $\geq$6 of them with $p_T > 45$)
$N_{bjets} \geq 2$ ($p_T > 30$ GeV, $\geq$1 of them with $p_T > 45$)
$\Delta R_{bjets} \geq 1$ (req. $N_{\mathrm{bjets}} \geq 2$ with $p_T > 30$ GeV)

Tagging effiecies, fake rates and ROC curves are given for different WPs corresponding to discriminator values of 0.95, 0.96, 0.97, 0.98 in Figure 6.3., Figure 6.4., Figure 6.5. and Figure 6.6. for 2016preVFP, 2016postVFP, 2017 and 2018, respectively. Note that each year is as part of the full Run2UL. While comparing the original WP 0.92 with other WPs, it has the most fake rate even if it has the best efficiency. WP 0.96, 0.97 and 0.98 have generally less fake rate while comparing with the WP 0.95. So, the WP 0.95 is chosen as optimized one, since it has better effiency than the tighter ones and lower fake rate than 0.92.

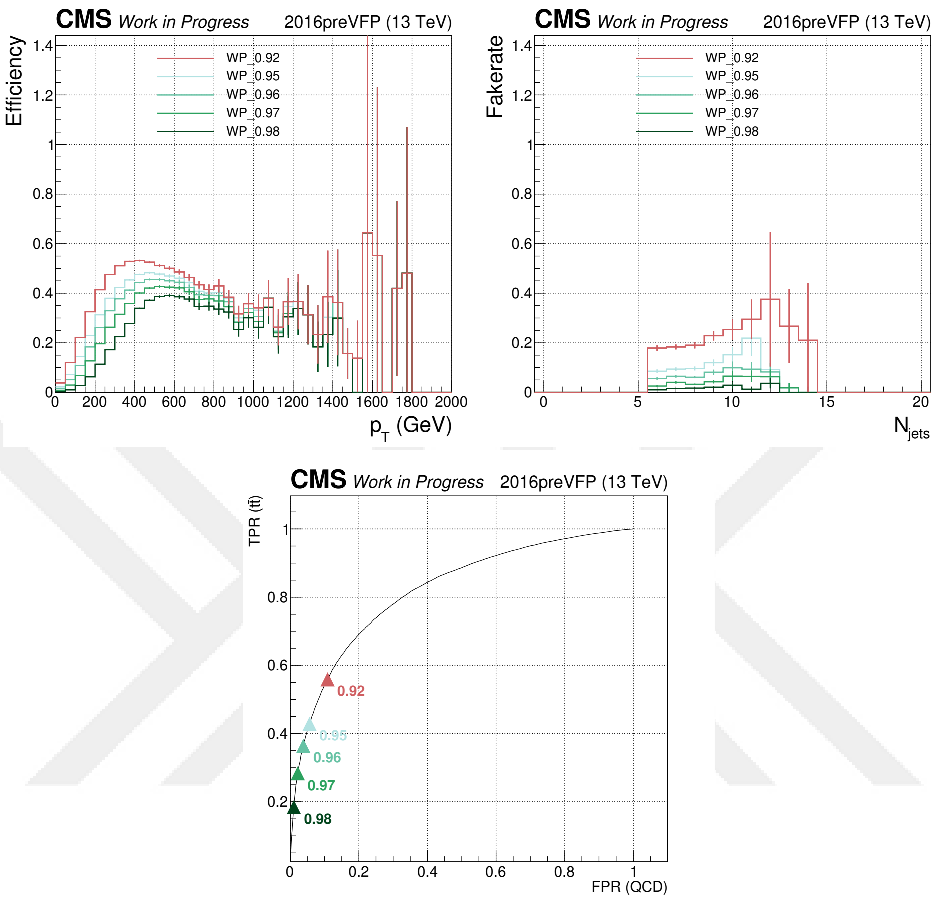


Figure 6.3. For 2016preVFP, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_T from simulated $t\bar{t}$ + jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}$ + jets events and FPR is measured in simulated QCD multi-jet events

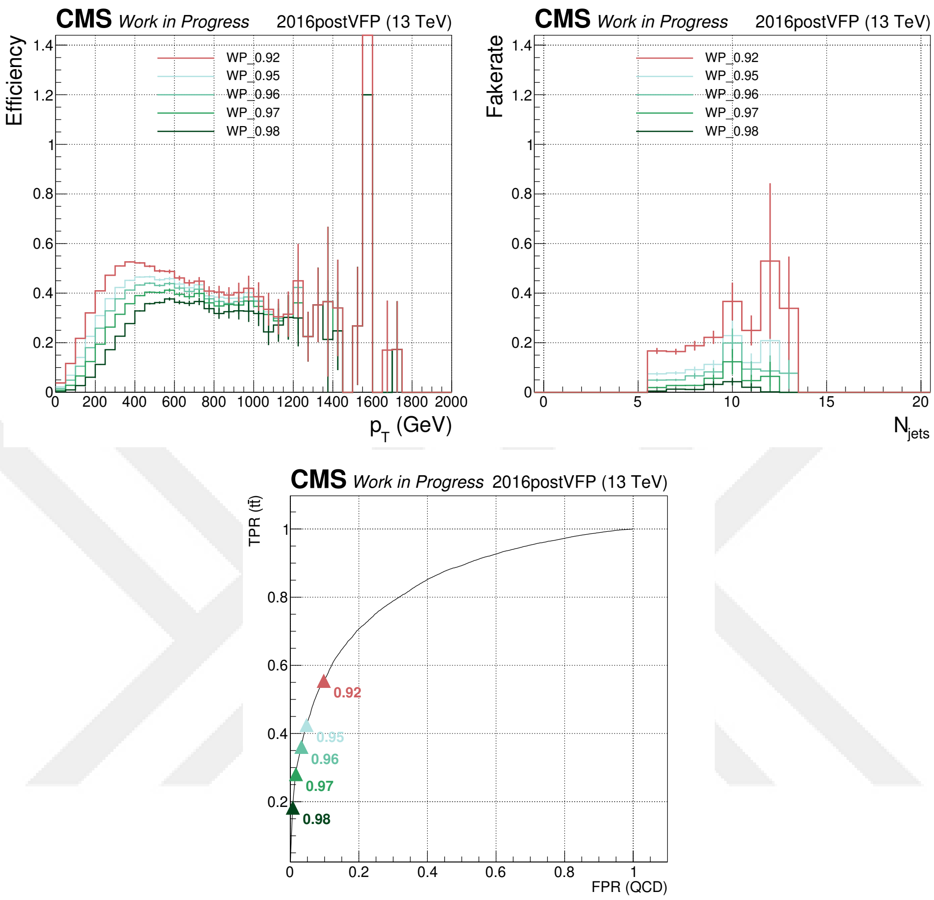


Figure 6.4. For 2016postVFP, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_{T} from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events

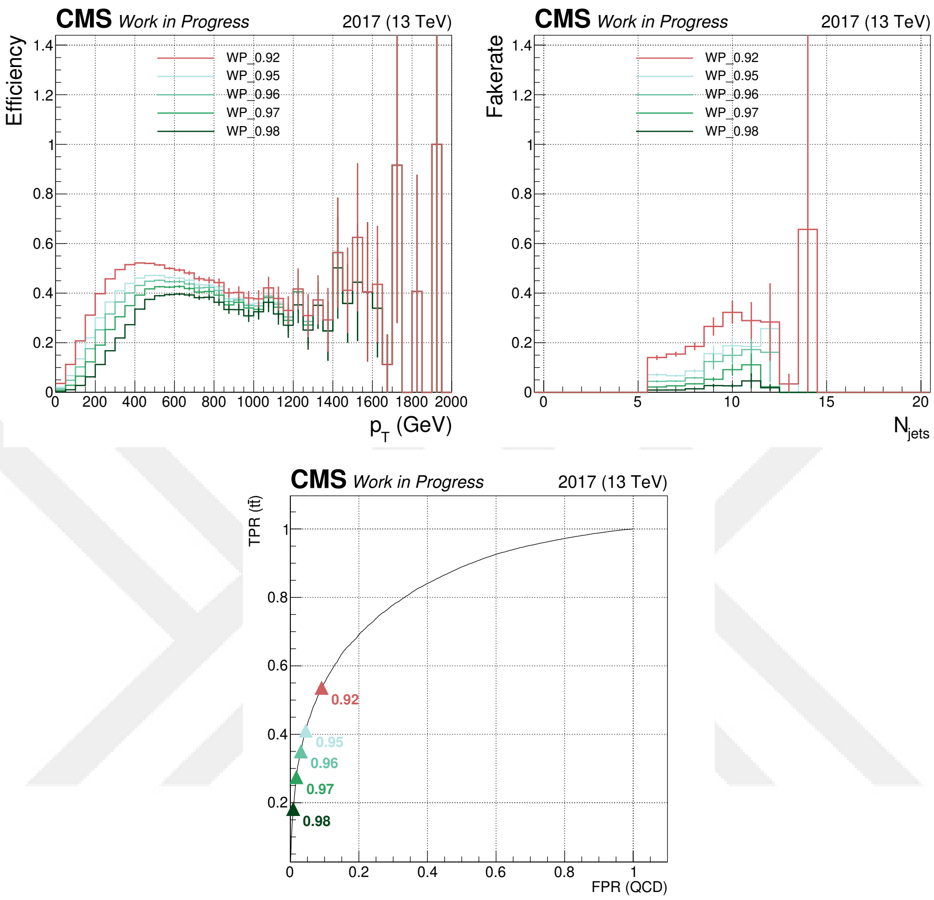


Figure 6.5. For 2017, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_{T} from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events

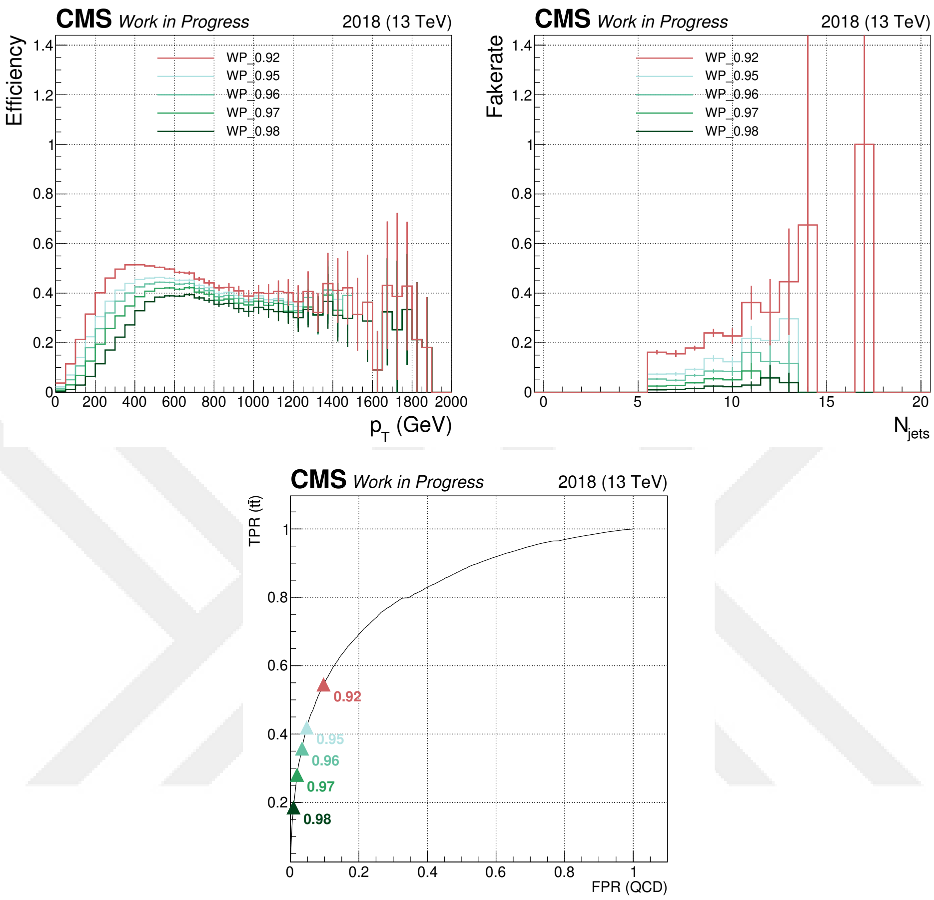


Figure 6.6. For 2018, tagging efficiency (top left) is measured as a fuction of generator-level top quark p_T from simulated $t\bar{t}+$jets events. Fake rate (top right) is measured as a fuction of N_{jets} using simulated QCD events. To make ROC curve (bottom), TPR is measured in simulated $t\bar{t}+$jets events and FPR is measured in simulated QCD multi-jet events

7. TOP SQUARK RECONSTRUCTION

As a part of the $t\bar{t}+$jets background estimation, Section 8., top squark reconstruction is implemented in this analysis. The purpose of this reconstruction is to gain another set of variables, which are directly related to the signal topologies considered in this analysis. This set of variables is feed into the Double DisCo Neural Network to be helpful for the better discrimination between signal and background events.

Motivated by the M_{T2} analysis (S. Chatrchyan et al., 2012), a reconstruction of top squarks in each event is performed by using the **Hemisphere Method** (CMS TDR, 2007). The method is split into the two parts: one is the **seeding method**, another is **association method**. Relying on the hemisphere method, the seeding method determines the two objects which serve as the seed axes of each of the hemispheres. These axes are used for grouping of objects together. Then, the objects are grouped into two hemispheres by the association method.

Two seeding methods are defined for this analysis. Invariant mass seeding method is the one used by M_{T2} analysis and this method chooses two axes as the directions of the objects having the largest invariant mass. The top seeding method is created specifically for this analysis, which chooses two axes as the directions of the top quarks having the highest p_{T} by implementing hadronic top tagger.

Only one association method, which is minimization of Lund Distance, is used in this analysis. Based on this method, each object is associated with either of the seeding axes by minimizing the Lund distance, given in Equation 7.1. The association continues until there is no more object in the event to cluster.

$$(E_i - p_i \cos\theta_{ik})\frac{E_i}{(E_i+E_k)^2} \geq (E_j - p_j \cos\theta_{jk})\frac{E_j}{(E_j+E_k)^2} \tag{7.1}$$

Each hemisphere should represent ”loose reconstruction of the top squark four-vector” in a signal event. For top seeding method, each hemisphere has exactly one top quark by leveraging the reconstruction power of the top tagger. This allows better identification of top squarks. Also, an AK4 jet filter is applied in each seeding method—requirements of the filter are $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$. Because, it is important to get the same objects like general jet selection in the analysis. Also, this filter is similar to preselection of hadronic top tagger algorithms.

Based on a study for choosing the input variables of Double DisCo NN, these four-vector of the top squarks are useful variables to distinguish signal events from the background events. The top seeding method is used to cluster top squarks and four-vector of the these clustered top squarks are fed in the Double DisCo NN for fully-hadronic channel. For the fully-hadronic channel, the four-vector of each clustered top squark are given in Figure 7.1. and Figure 7.2. for full Run2UL simulation where from all of the four data taking years are combined as Run2UL.

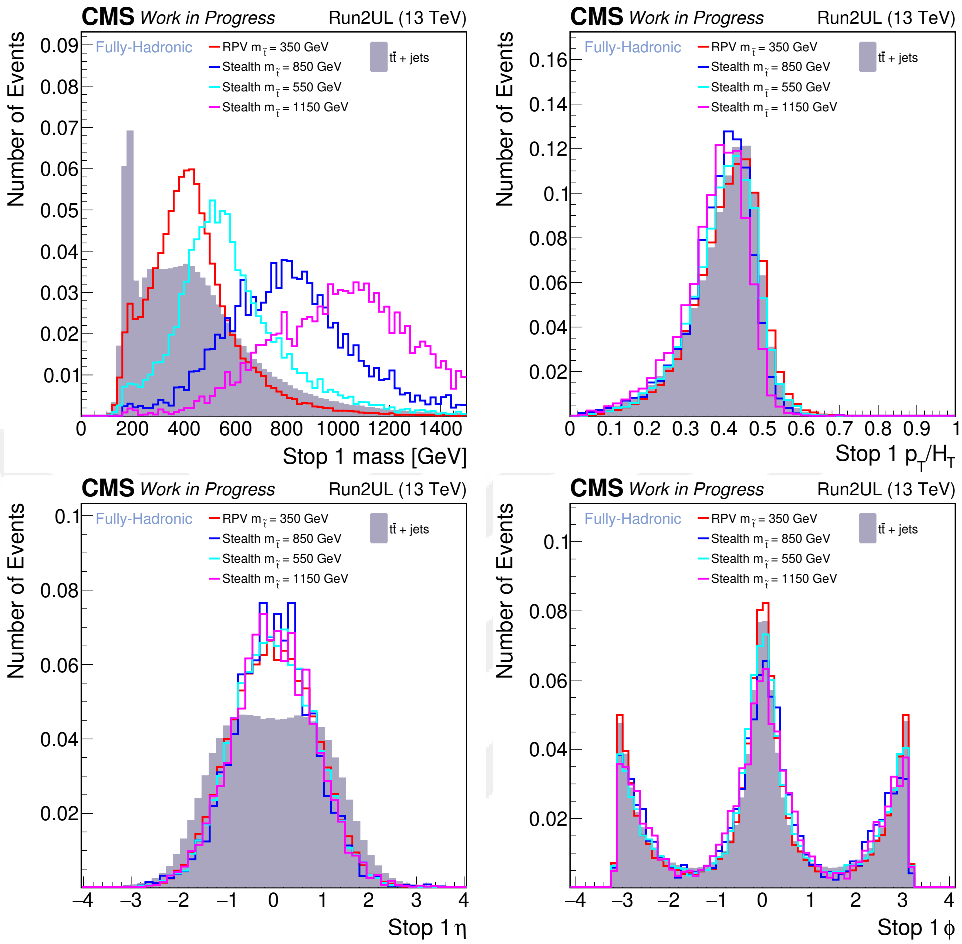


Figure 7.1. For the fully-hadronic channel, four-vector of clustered top squark 1 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The top seeding method is used to cluster top squarks. Top squark 1 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL

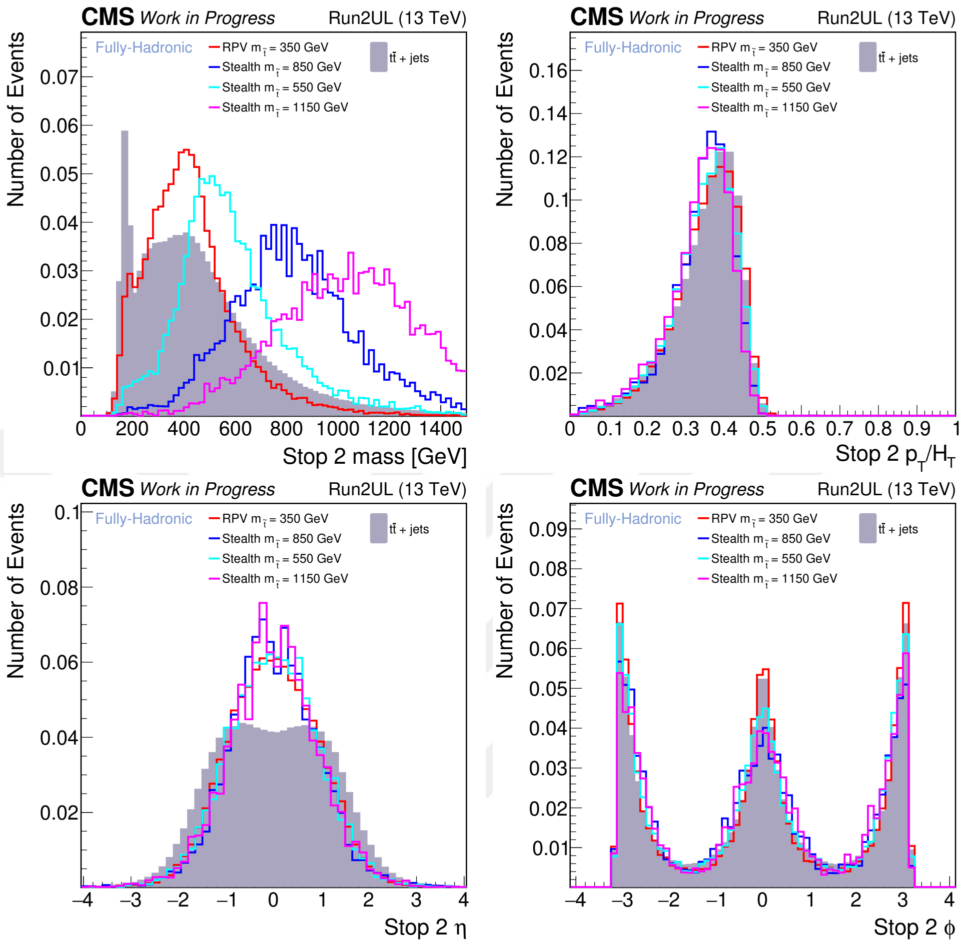


Figure 7.2. For the fully-hadronic channel, four-vector of clustered top squark 2 for $t\bar{t}$ + jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The top seeding method is used to cluster top squarks. Top squark 2 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL

The invariant mass seeding method is used to cluster top squarks and four-vector of the these clustered top squarks are fed in the Double DisCo NN for each of semi-leptonic and fully-leptonic channels. The four-vector of each clustered top squark are given for the semi-leptonic channel in Figure 7.3. and Figure 7.4.; given for the fully-leptonic channel in Figure 7.5. and Figure 7.6. for full Run2UL simulation where from all of the four data taking years are combined as Run2UL.

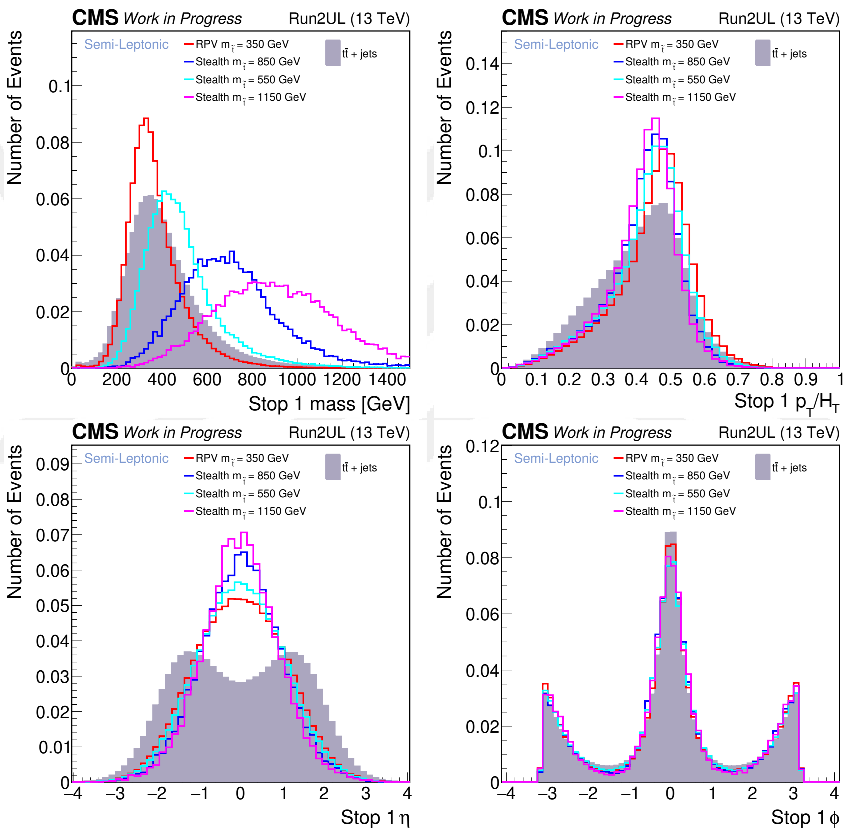


Figure 7.3. For the semi-leptonic channel, four-vector of clustered top squark 1 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 1 mass (top left), p_T/H_T (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL

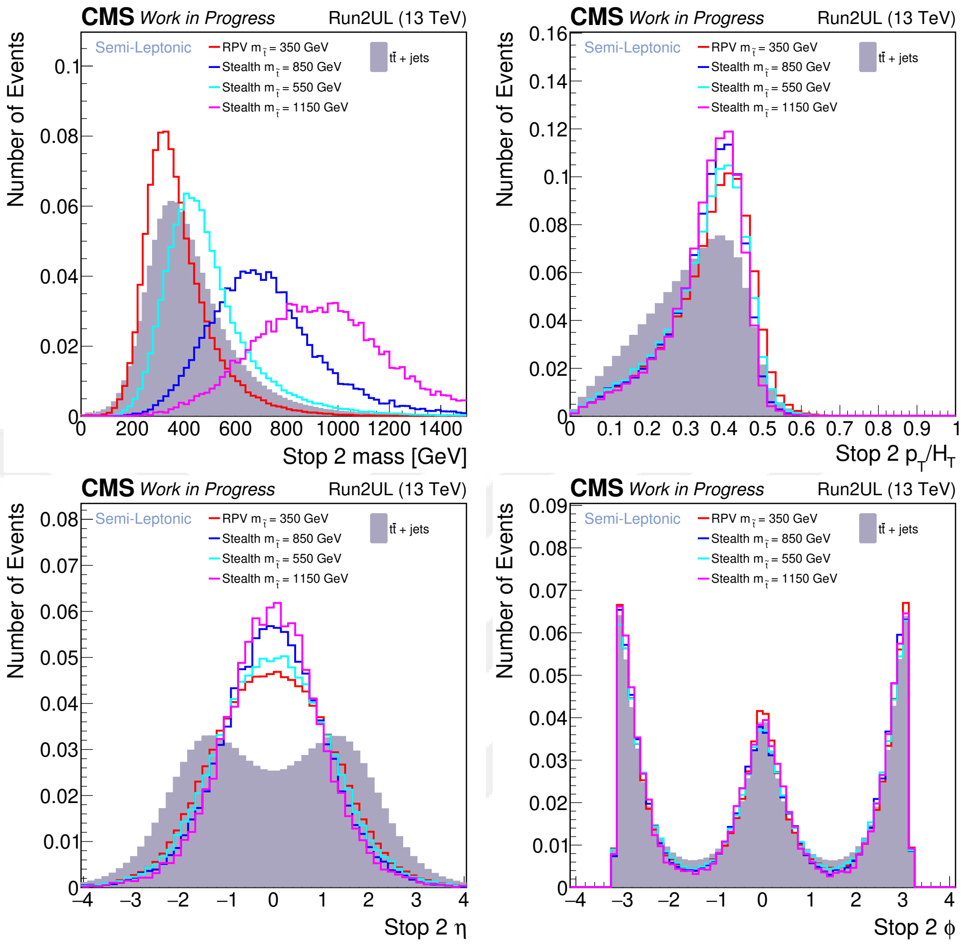


Figure 7.4. For the semi-leptonic channel, four-vector of clustered top squark 2 for $t\bar{t}$ + jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 2 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL

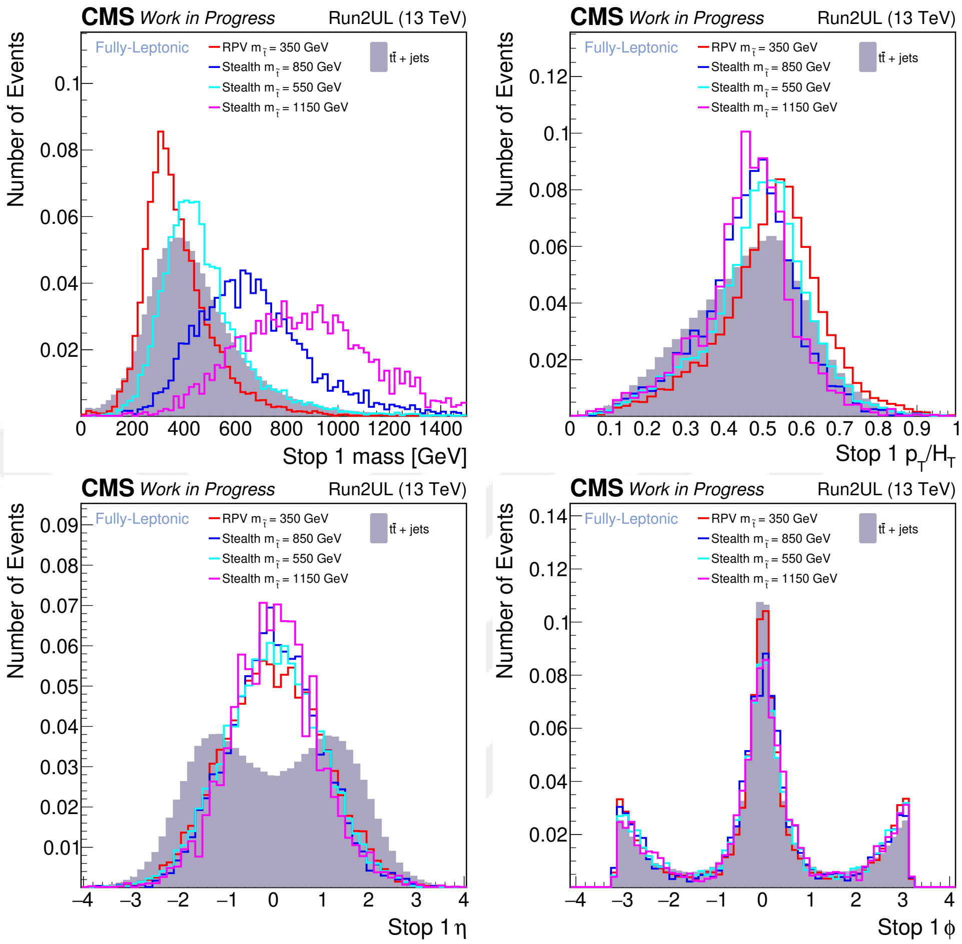


Figure 7.5. For the fully-leptonic channel, four-vector of clustered top squark 1 for $t\bar{t}+$jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 1 mass (top left), $p_{\mathrm{T}}/H_{\mathrm{T}}$ (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL

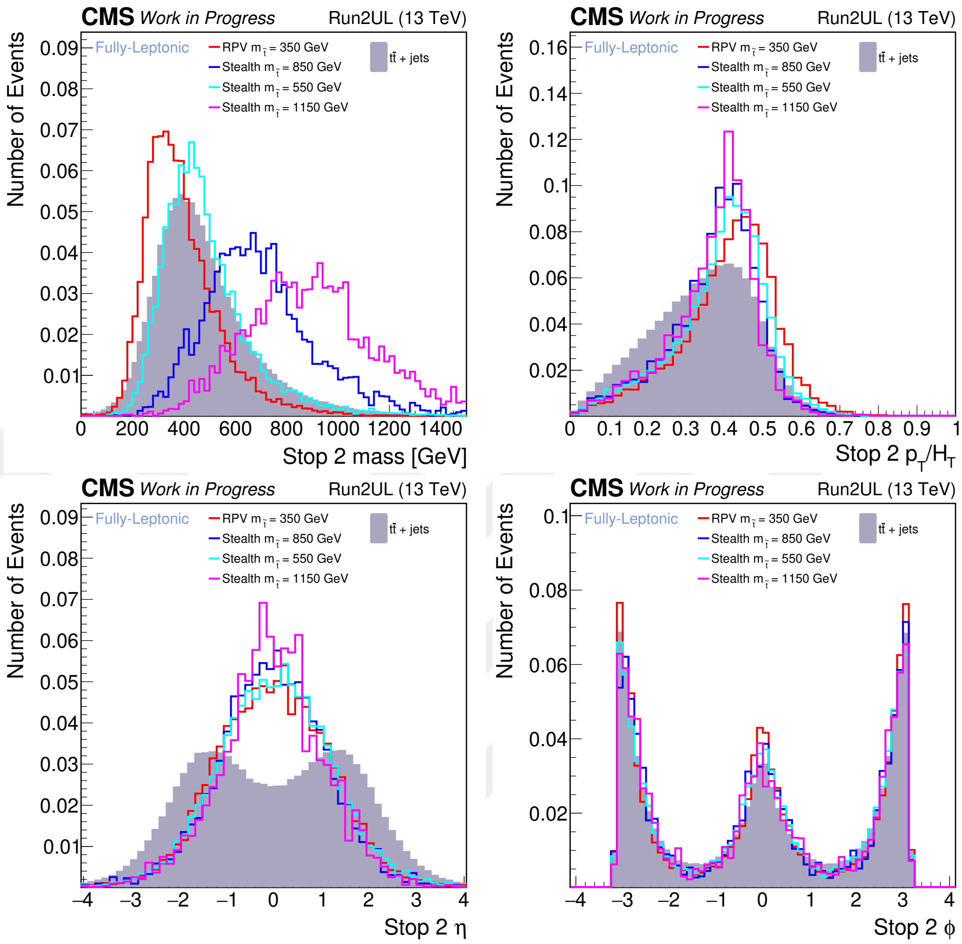


Figure 7.6. For the fully-leptonic channel, four-vector of clustered top squark 2 for $t\bar{t}$ + jets, one RPV and couple of Stealth SY$\bar{Y}$ masses. The invariant mass seeding method is used to cluster top squarks. Top squark 2 mass (top left), p_T/H_T (top right), η (bottom left) and ϕ (bottom right) for full Run2UL simulation where from all of the four data taking years are combined as Run2UL

8. $t\bar{t}+$jets ESTIMATION AND SYSTEMATICS

This entire chapter is dedicated to estimating the $t\bar{t}+$jets background and its associated systematic uncertainties. In this analysis, in contrast to the legacy analysis, using a basic estimation method is prefered for fewer challenges based on N_{jets} distributions. One of the common data-driven estimation methods is used. However, this analysis implements a novel NN to maximize performance.

First the ABCD method is described in Section 8.1.. Then, ”Customized” Double DisCo NN model is presented in Section 8.2. and the details of network trainings are given in Section 8.3.. For understanding the robustness of the ABCD method and to verify the modeling of background in data, validation regions are defined in Section 8.4.. A correction defined is for calculating the systematics, as given in Section 8.5.. Last, the systematic uncertainties are calculated for the $t\bar{t}+$jets background in Section 8.6..

8.1. ABCD Method

In this analysis, the automated ABCD method using Double DisCo NN (Kasieczka et al., 2021) is used to estimate the $t\bar{t}+$jets component of the background in data. The ABCD method is one of the common data-driven background estimation techniques in high energy physics. The data-driven background estimation is to extrapolate or interpolate from some control regions into a signal region of interest.

The idea of the ABCD method is to choose two variables, which are independent for background and also demonstrate separation between signal and background. With the two variables that satisfy the above criteria, four regions can be composed. Three of these regions: **B**, **C**, and **D** are control regions and they mostly are background dominated. The last region **A** is the signal region. If these two variables are independent then the amount of background in the signal region can be predicted by using the control regions via the relationships:

$$N_{\mathrm{pred.},\,A} = (N_B N_C)/N_D \tag{8.1}$$

where N_i is the number of events in region i and $N_{\mathrm{pred.},\,A}$ is the number of predicted (estimated) background events in the signal region. The accuracy of the ABCD method is based on improving its performance in terms of ABCD closure, background rejection, and signal contamination. So, if the method works well, then we expect to see that low signal contamination in **B**, **C**, **D** regions and reasonable signal fraction in **A** region. Also, based on the Double DisCo NN we expect to see good discrimination between signal and background. Then, ABCD closure can be defined as:

$$\text{ABCD Closure} = \frac{N_{\mathrm{pred.},\,A}}{N_{\mathrm{obs.},\,A}} \tag{8.2}$$

where $N_{\mathrm{pred.},\,A}$ is number of predicted background events in the signal region and $N_{\mathrm{obs.},\,A}$ is the number of observed background events in the signal region. The closure is to measure how well the method works. If the closure is about ~ 1, then the ABCD method works well. Figure 8.1. shows a dramatization of the ABCD method for two arbitrary variables.

Typically, these variables are chosen by hand to be simple, physically well-motivated variables such as invariant mass, p_{T}, H_{T}, etc. For this analysis, the primary background $t\bar{t}+$jets is very similar to the final state of the signal process, and thus finding two simple object- or event-level variables to differentiate signal from background – let alone be independent – is very challenging.

To automate the ABCD method and obtain the two necessary variables, a Double Disco NN is used, which uses a well-known statistical measure of nonlinear dependence known as Distance Correlation (DisCo) (Székely and Rizzo, 2014). The network has two binary classifiers, which are simultaneously trained to discriminate

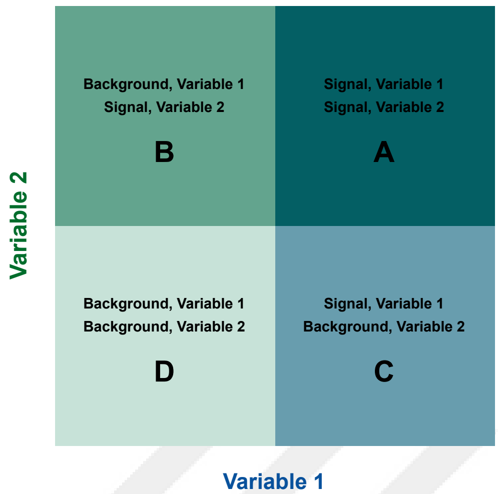


Figure 8.1. A dramatization of the ABCD method for two arbitrary variables. The number of background events in region A, the signal region, can be calculated as $N_{\text{pred.},A} = (N_B N_C)/N_D$

signal from background. Each classifier output value is a number between 0 and 1 represents each of two independent variables. The loss function of the network includes a decorrelation term, forcing the two output discriminants to be independent.

8.2. Double DisCo Neural Network Model

A neural network is a tool including some components to evaluate the input information, which the network learns from, for making predictions or clasifying things as output. Based on its purpose, a network consists of (1) some models which are algorithms to make a certain prediction, (2) loss functions which quantify how well a network performs and (3) hyperparameters which is a list of configuration parameters based on each model and each loss function. For the network functionality, there are input layer for feeding input information to the network, hidden layer for implementing the network components, and output layer for giving the output. Each of these layers has some nodes.

In this thesis, a Double DisCo NN is implemented via Tensorflow (Python, 2023) Python packages including 2 binary classifiers as models and loss functions based on each model. A mass regression model is also added to the network to get extra information for better discrimination between signal and background in the binary classifiers. So, the mass regression model provides a useful input variable to each binary classifier. In the total loss function, there are three terms based on each model ($L_{Reg.}$, $L_{Disc.,1}$, $L_{Disc.,2}$) and a decorrelation term ($L_{DeCorr.}$) forcing the two output discriminants to be independent. The $L_{DeCorr.}$ includes extra term based on the ABCD method which also makes the neural network custom. More details for the total loss function are given in Section 8.2.3.. The hyperparameter list for each model of the network is given in Section 8.2.4..

This "customized Double DisCo NN", illustrated in Figure 8.2. with three models, is built and trained simultaneously as one conglomerate neural network model on a set of $t\bar{t}+$jets and Stealth SY$\bar{Y}$ signal events to produce the final background prediction. There are three individual trainings based on each analysis channel. For each training, the input variables listed in Section 8.3.1. are fed into the network and the training samples are shown in Section 8.3.2..

8.2.1. Mass Regression Model

How the regression output differs between signal and background is the most important aspect. So, the regression is trying to predict masses of either the top squark (signal) or top quark (background) for an event. The output of the regression is an additional input variable to each binary classifier in order to help them for better discrimination. The "mass label" is defined as an event label for the regression. This event label is the per-event truth mass based on the leading-p_{T} reconstructed top squark for a signal event and top quark for a background event. For this label, a generator

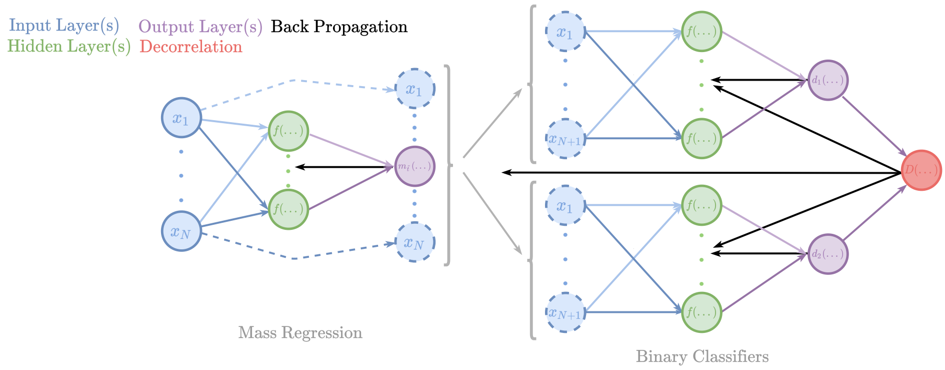


Figure 8.2. The ”customized” Double DisCo neural network is shown with its three models. Mass regression and two binary clasifier models are shown with their input, hidden and output layers. Output of mass regression model is an input for each binary classifier

level matching procedure is used to identify all reconstructed objects coming from the decay of top squark or the top quark.

The mass regression model is composed of the input layer containing the object and event variables, a dropout layer preventing overfitting, a single hidden layer, and the regression output layer, which is connected to the input layers of the binary classifiers. A total of 100/200 nodes (dependent on which analysis channel) are used on the single hidden layer. The loss function, $L_{Reg.}$, is a mean squared error term calculated between the truth and predicted mass as given in Equation 8.3, where λ_{Reg} is a hyperparameter scaling the mass regression loss; $M^{pred.}_{t/\tilde{t}}$ is the predicted top quark (top squark) mass for a background (signal) event; $M^{truth}_{t/\tilde{t}}$ is the pre-computed “truth” mass (mass label); the sum is over the N events in a given batch.

$$L_{Reg.} = \lambda_{Reg.} \sum_{i}^{N} \left(M^{pred.}_{t/\tilde{t}} - M^{truth}_{t/\tilde{t}} \right)^2 \tag{8.3}$$

The set of hyperparameters is given in Section 8.2.4. and the $L_{Reg.}$ as a part of the total loss function is given in Section 8.2.3..

8.2.2. Binary Classifier Models

The binary clasifier models are the main body of the whole network, where background is discriminated from signal. These classifiers are identical including the same number of the hidden layers with the same number of nodes.

Each has an input layer to take in all the input variables and the output of the mass regression providing an additional handle to distinguish background from signal. Both classifiers then have a single hidden layer containing 100/200 nodes (dependent on which analysis channel), a dropout layer and an output layer. A binary cross entropy (BCE) term is used for each of the classifier loss function, $L_{Disc.}$, to quantify how well the classifiers discriminate background from signal. Explicitly, the functional form of this loss contribution is given in Equation 8.4, where $d_{j\in 1,2}$ is one of the two discriminant values for event i; sum over the N events in a batch is performed.

$$L_{\mathrm{Disc.},\,j} = \lambda_{Disc.,j} \sum_{i}^{N} (1 - d_{i,j}) \log d_{i,j} + d_{i,j} \log (1 - d_{i,j}) \tag{8.4}$$

A truth value of 0 is used to label background events and a truth label of 1 is used to label signal events. Note that all signal events, regardless of the assumed mass of the top squark, share the same signal label. More details of the hyperparameters for the classifiers are given in Section 8.2.4., and the $L_{Disc.,1}$ and $L_{Disc.,2}$ loss terms are given as a part of the total loss function in Section 8.2.3..

8.2.3. Total Loss Function

The total loss function for this NN is partially comprised of simple contributions coming from $L_{Reg.}$, $L_{Disc.,1}$ and $L_{Disc.,2}$. The mass regression contributes a term based on how well it is able to predict the mass of the top squark or top quark in signal or background events, respectively. Likewise, each classifier contributes a term quantifying its ability to correctly identify signal and background. In addition to the simple loss terms for each model, there is a more complex term combining the output from both classifiers, which is $\boldsymbol{L_{DeCorr.}}$. As it is desired for the two classifier outputs to be independent, this additional term is added to the total loss function in the form of two components. It is given as in Equation 8.5, where $L_{DisCo.}$ is defined as a distance correlation term and L_{ABCD} is defined as an ABCD non-closure term. Then the total loss function can be written as in Equation 8.6.

$$L_{DeCorr.} = [L_{DisCo.} + L_{ABCD}] \tag{8.5}$$

$$\mathcal{L} = L_{Reg.} + L_{Disc.,1} + L_{Disc.,2} + L_{DeCorr.} \tag{8.6}$$

Distance Correlation Term

Double DisCo NN uses a well-known statistical measure of non-linear dependence known as "distance correlation" (Székely and Rizzo, 2014). Distance correlation is a function of two random variables, given in Equation 8.7 and it is zero if and only if the variables are statistically independent, otherwise these two variables are not independent. This function can be added as a regularization term in a loss term of the total loss function.

$$dCorr^2(X,Y) = \frac{dCov^2(X,Y)}{\sqrt{dCov^2(X,X)dCov^2(Y,Y)}} \tag{8.7}$$

The "Distance Correlation" term of the loss function, $\boldsymbol{L_{DisCo.}}$, directly encodes the correlation between the outputs of the two classifiers to be independent. Specifically, the term calculates the distance correlation between discriminant 1 and discriminant 2 for background events in each batch during training. Then, $\boldsymbol{L_{DisCo.}}$ term of the total loss function is written as in Equation 8.8.

$$L_{DisCo.} = \lambda_{DisCo.} dCorr^2 \tag{8.8}$$

ABCD Non-Closure Term

The "ABCD" term of the loss function, $\boldsymbol{L_{ABCD}}$, is based on the closure of the ABCD method, which is quantified as the difference between the predicted (via the ABCD method) and observed number of $t\bar{t}+$jets events in the **A** region with respect to the number of $t\bar{t}+$jets events observed. The ABCD non-closure as defined in Equation 8.9 and the ABCD loss term of the total loss function is given as in Equation 8.10.

$$\text{ABCD non-closure} = 1 - \frac{N_{\text{pred.},A}}{N_{\text{obs.},A}} \tag{8.9}$$

$$L_{ABCD} = \lambda_{ABCD}\left|1 - \frac{N_{\text{pred.},A}}{N_{\text{obs.},A}}\right| \tag{8.10}$$

During training of the NN, this term is calculated using all the $t\bar{t}+$jets events in the current batch being passed through the network. To perform the calculation, the two bin edges defining the ABCD regions must be chosen. Each time the ABCD non-closure is calculated the bin edge values are randomly and uniformly chosen from the range [0, 1]. This way of choosing the edges prevents the network from simply learning to achieve perfect ABCD closure for a particular location and nowhere else in the 2D plane. The network must try and achieve good ABCD closure across the 2D plane, which by definition, assists in keeping the two output discriminants independent, by more exactly informing the network of how the 2D output discriminants will be used.

With discrete bin edge values randomly chosen per-batch, there may be instances during the training where no events appear in one of the four bins. This will either result in the ABCD closure term going to zero or going to infinity for that batch. To avoid this discontinous varying of the loss term spoiling the network's ability to learn via gradient descent, a "smearing" procedure is used. Two sigmoid functions, one for each discriminant dimension, are used to calculate a "smeared" number of counts in each of the four regions.

$$
\begin{aligned}
N_A &= \sum_i^N S(a_0(x_i - x_0))\, S(a_1(y_i - y_0)) \\
N_B &= \sum_i^N S(a_1(y_i - y_0)) - N_A \\
N_C &= \sum_i^N S(a_0(x_i - x_0)) - N_A \\
N_D &= N_{\text{total}} - N_A - N_B - N_C
\end{aligned}
\tag{8.11}
$$

Equation 8.11 shows how the two sigmoid functions are used to calculate the number of counts in each region, where the randomly chosen bin edges are x_0 and y_0. Scale parameters $a_0 = a_1$ are used to define the sharpness of the sigmoid and in the limit $a \to \infty$, the original counts in each of the ABCD regions are recovered. Both scale parameters are set to 100 and this procedure restores stability to this loss term, allowing the NN to learn from it.

8.2.4. Hyperparameters

The hyperparameters are a set of configuration parameters based on each model and each loss term that allow the network to learn optimally. In terms of the fundemental concept to any network, there are couple of hyperparameters: Learning rate (LR), λ value of the loss function, number of hidden layers and number of hidden layer nodes.

The LR controls how quickly the network is trained. There is one LR for the whole network for each training. It is chosen to try and ensure optimal performance.

The λ hyperparameters control the relative importance of each term in the total loss function of the NN. For this model, it is a matter of balancing the importance of maintaining good separation of signal and background in the 2D discriminant plane, but at the same time, producing a 2D background shape that is conducive to performing the ABCD method. The four main hyperparameters are $\lambda_{Reg.}$, controlling importance of accurate prediction of mass; $\lambda_{Disc.}$, controlling importance of signal vs background discrimination for both classifiers; λ_{DisCo}, controlling importance of keeping the two classifier outputs independent via distance correlation; and λ_{ABCD}, controlling the importance of how well ABCD method works.

Lastly, number of hidden layers and number of hidden layer nodes are the parameters based on the functionality of the network, which tell how complex the network is. The list of the hyperparameters for each of the fully-hadronic, semi-leptonic and fully-leptonic channel is given in Table 8.1..

Table 8.1. The list of hyper parameters for each of the fully-hadronic, semi-leptonic and fully-leptonic channel. Note that the given LR for the whole network for a given channel

Fully-Hadronic Channel		
Mass Regression	**Binary Classifiers**	**DeCorr**
Regression Nodes = 200	**Disc. Nodes** = 200	**DeCorr Nodes** = 200
Regression Layers = 1	**Disc. Layers** = 1	**DeCorr Layers** = 1
$\lambda_{\text{Reg.}}$ = 1.0	$\lambda_{Disc.}$ = 1.0	λ_{DisCo} = 10.0 | λ_{ABCD} = 10.0
LR = 0.0001 for the whole network		
Semi-Leptonic Channel		
Mass Regression	**Binary Classifiers**	**DeCorr**
Regression Nodes = 200	**Disc. Nodes** = 200	**DeCorr Nodes** = 200
Regression Layers = 1	**Disc. Layers** = 1	**DeCorr Layers** = 1
$\lambda_{Reg.}$ = 1.0	$\lambda_{Disc.}$ = 1.0	λ_{DisCo} = 10.0 | λ_{ABCD} = 10.0
LR = 0.0001 for the whole network		
Fully-Leptonic Channel		
Mass Regression	**Binary Classifiers**	**DeCorr**
Regression Nodes = 200	**Disc. Nodes** = 200	**DeCorr Nodes** = 200
Regression Layers = 1	**Disc. Layers** = 1	**DeCorr Layers** = 1
$\lambda_{Reg.}$ = 1.0	$\lambda_{Disc.}$ = 2.0	λ_{DisCo} = 10.0 | λ_{ABCD} = 10.0
LR = 0.0001 for the whole network		

8.3. Double DisCo Trainings

8.3.1. Input Variables

With a final state containing many jets, the properties of them are important and useful to the network. Properties for the first eight (for the fully-hadronic), the first seven (for the semi-leptonic), and the first six (for the fully-leptonic) leading p_{T} jets (channel dependent), such as four-vector information and flavor information are used as input variables. Also, four-vectors of leptons are used as inputs where applicable for the semi-leptonic and fully-leptonic channels. In addition to this lower level information, there are some higher level variables including Fox-Wolfram Moments (Fox and Wolfram, 1978) and eigenvalues of the jet energy-momentum tensor (Bjorken and Brodsky, 1970). These high level variables are calculable from the lower level variables; however, with a finite training sample and a finite-complexity network, pre-computing the variables for the network is advantageous to its learning. The moments and eigenvalues help give additional information about energy flow in an event and the relative angular distribution and separation of all the jets. Additionally, the four-vector of top squark, details are given in Section 7., also gives information on how the event is clustered into two halves, which can be distinct for signal and background. Note that one set of the four-vector of top squark variables are reconstructed with "Old Seeding Method" and it is used as input for trainings of the semi-leptonic and fully-leptonic channels. Another set of the four-vector of top squark variables are reconstructed with "Top Seeding Method" and it is used as input for training of the fully-hadronic channel. The input variables of the network for each channel of training are listed in Table 8.2.. Distributions of some other input variables derived from simulation events are shown in Appendix C1..

8.3.2. Training Samples

The $t\bar{t}+$jets background events used for training are those that pass the fully-hadronic, semi-leptonic, and fully-leptonic baseline selections as defined in Section 5.4.1.. Signal events are all SY$\bar{Y}$ model events that pass the baseline selections, including all mass points available between 300 GeV and 1400 GeV. In order to increase the total size of the training sample (80% of all events eligible to be used in training), several variations are also used as background events. For $t\bar{t}+$jets, events with variations to the ME-PS matching scale are used as well as events generated with

Table 8.2. The input variables of the network for each of the fully-hadronic, semi-leptonic and fully-leptonic channel. The four-vector of top squark with "Old Seeding Method" is used for trainings of the semi-leptonic and fully-leptonic channels. The four-vector of top squark with "Top Seeding Method" is used for training of the fully-hadronic channel

Fully-Hadronic channel
Jet p_T/H_T (1-7)
Jet η and ϕ (1-7)
Jet DeepFlavor Discriminators (1-7)
Combined 8th+ Jet p_T/H_T
Combined 8th+ Jet η and ϕ
Fox-Wolfram Moments
Jet Momentum Tensor Eigenvalues
Four-Vector of top squark with Top Seeding
Semi-Leptonic Channel
Jet p_T/H_T (1-6)
Jet η and ϕ (1-6)
Jet DeepFlavor Discriminators (1-6)
Combined 7th+ Jet p_T/H_T
Combined 7th+ Jet η and ϕ
Fox-Wolfram Moments
Jet Momentum Tensor Eigenvalues
Four-Vector of top squark with Old Seeding
Four-vector of lepton (1)
Fully-Leptonic Channel
Jet p_T/H_T (1-5)
Jet η and ϕ (1-5)
Jet DeepFlavor Discriminators (1-5)
Combined 6th+ Jet p_T/H_T
Combined 6th+ Jet η and ϕ
Fox-Wolfram Moments
Jet Momentum Tensor Eigenvalues
Four-Vector of top squark with Old Seeding
Four-vector of lepton (1-2)

tune variations, and events where color reconnection is allowed in the parton shower. For all of $t\bar{t}+$jets events and for all signal events, jet energy correction (JEC) and jet energy resolution (JER) variations of the events are also added to the training set. Besides just increasing the total number of available events to train on, the intent with using all of these variations is to "desensitize" the network to the variations and make it more robust.

The batches of events are carefully prepared so that the network sees equal populations of events constantly while training. A batch size of 2048 is used where half of the events are background and half are signal. For each half, five categories are equally represented, corresponding to the number of jets in an event, i.e. ~ 200 events each for all N_{jets} categories (for the fully-hadronic channel: 8-12+ jets, for the semi-leptonic channel: 7-11+ jets, for the fully-leptonic channel: 6-10+ jets).

8.3.3. Training Performance

There is an individual NN training for each of three analysis channels. For each of these NN trainings, the Double DisCo NN is trained on simulated events which include both $t\bar{t}+$jets background and all Stealth $SY\bar{Y}$ signal events as input root files. Note that these simulated events for each of $t\bar{t}+$jets background and Stealth $SY\bar{Y}$ signal are combination from all of the four data taking years as Run2UL.

An important aspect of the NN training is verifying that it is not overtrained and that the outputs behave as expected. Loss as a function of epoch curves are produced using the training data set and the validation data set. If the validation curve increases above the train curve after some number of epochs, it is suggestive that the network is starting to overtrain. These comparisons are made in Figure 8.3., Figure 8.4. and Figure 8.5. for the fully-hadronic, semi-leptonic and fully-leptonic channels, respectively. These are separated by the different components of the total loss function:

$L_{Reg.}$, $L_{Disc.}$, L_{DisCo} and L_{ABCD}. No signs of overtraining are observed for any of given loss terms.

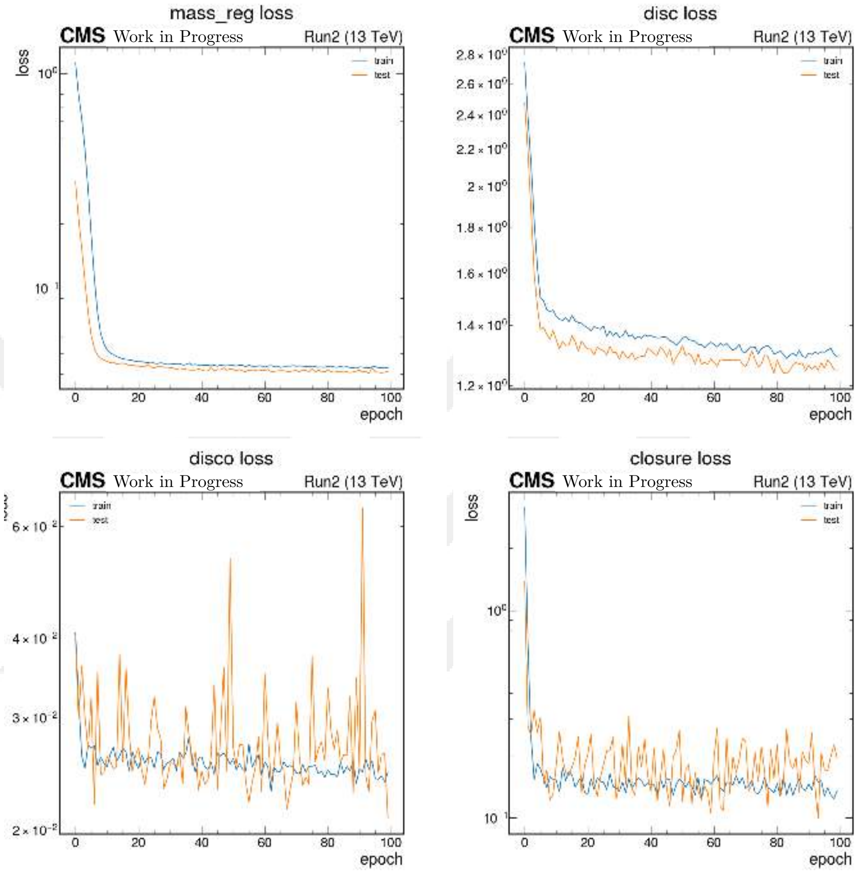


Figure 8.3. For the fully-hadronic channel, comparisons of loss as a function of epoch for the training (blue) and validation (yellow) data sets. The comparisons are shown for the mass regression component (top left), the classifier component (top right), the distance correlation component (bottom left) and ABCD closure component (bottom right)

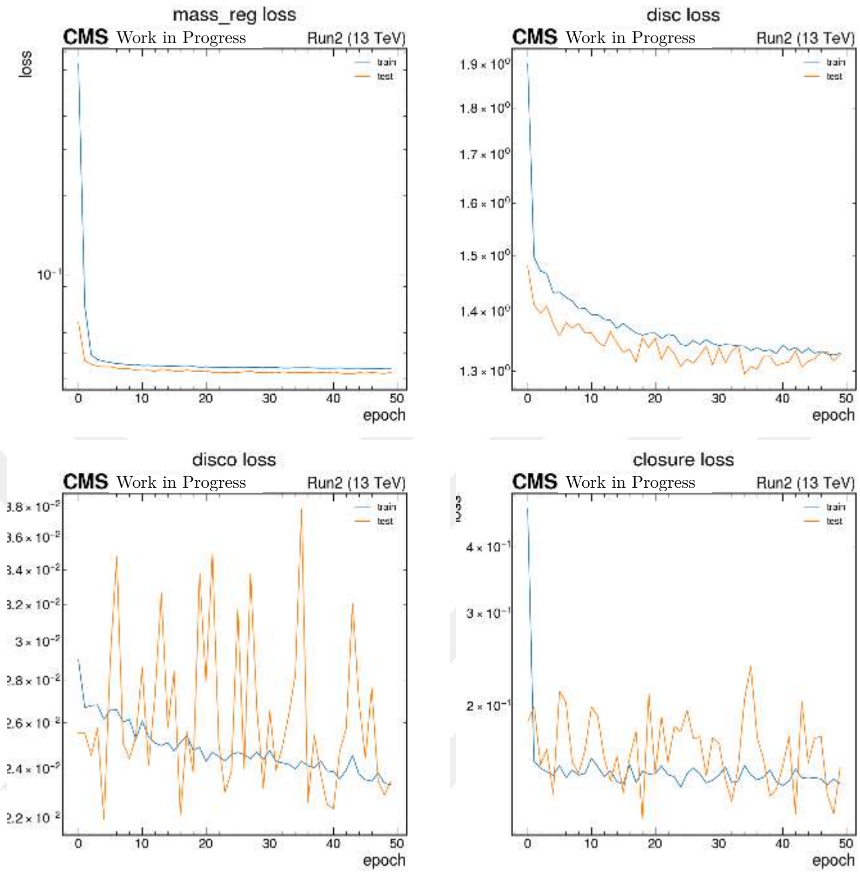


Figure 8.4. For the semi-leptonic channel, comparisons of loss as a function of epoch for the training (blue) and validation (yellow) data sets. The comparisons are shown for the mass regression component (top left), the classifier component (top right), the distance correlation component (bottom left) and ABCD closure component (bottom right)

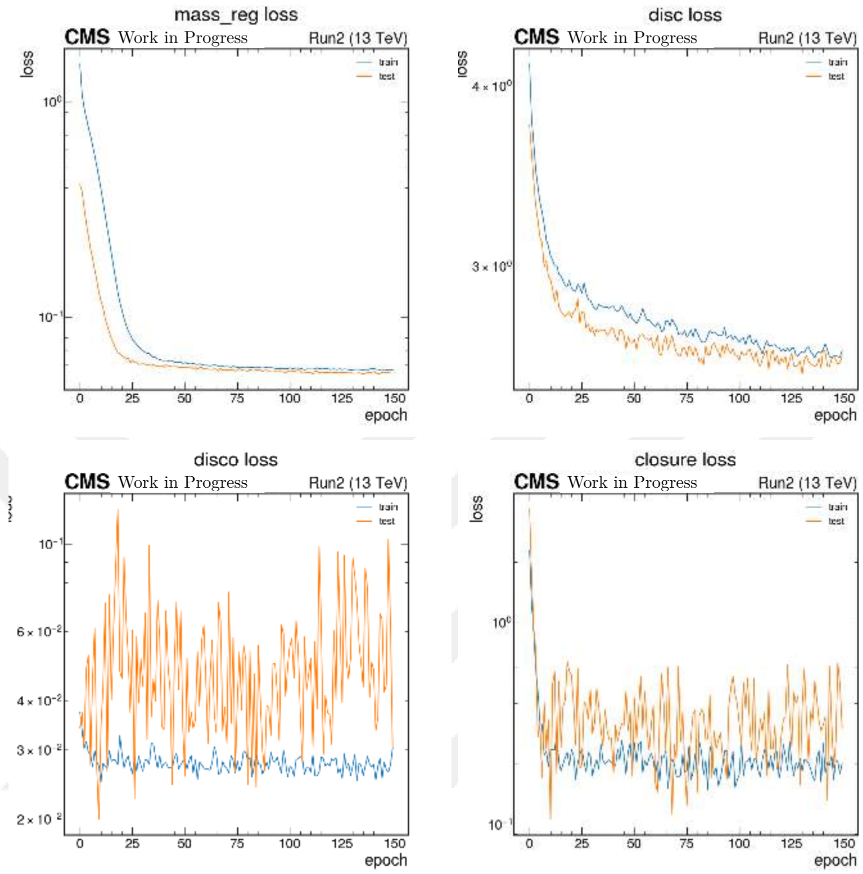


Figure 8.5. For the fully-leptonic channel, comparisons of loss as a function of epoch for the training (blue) and validation (yellow) data sets. The comparisons are shown for the mass regression component (top left), the classifier component (top right), the distance correlation component (bottom left) and ABCD closure component (bottom right)

Additionally, another comparisons between 2D discriminant outputs and mass regression outputs to check if there is any overtraining. Figure 8.6. for the fully-hadronic channel, Figure 8.7. for the semi-leptonic channel and Figure 8.8. for the fully-leptonic channel show these comparisons and the network is not overtrained, since the outputs used for training have the same behavior as the outputs used for validating of the network.

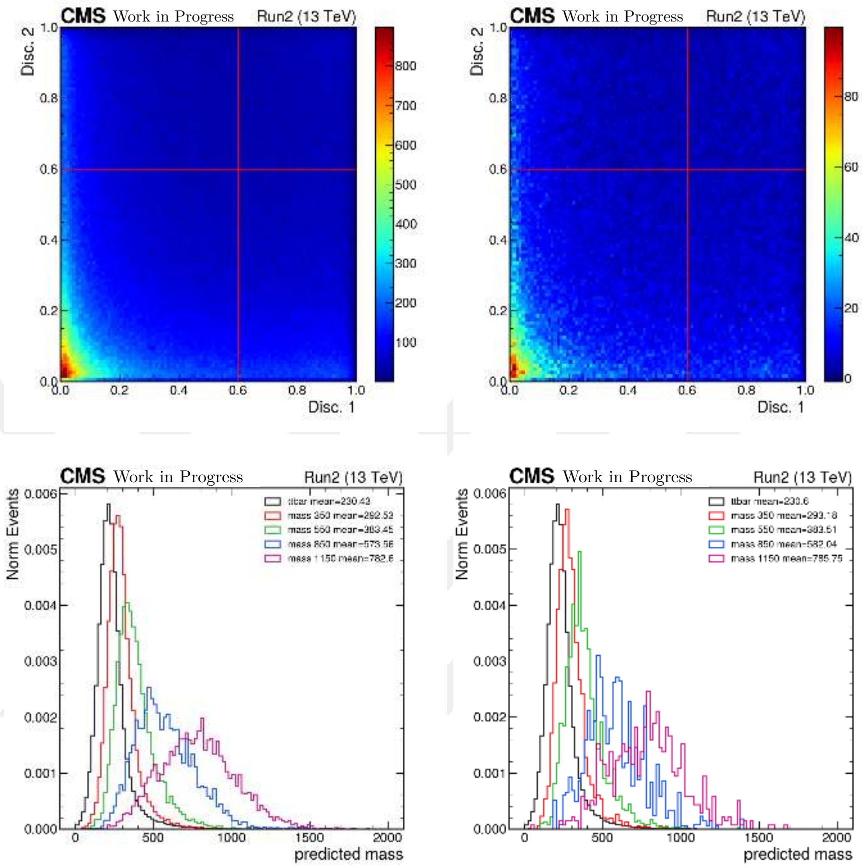


Figure 8.6. For the fully-hadronic channel, the 2D background discriminant output used for training of the network (top left), the 2D background discriminant output used for validating of the network (top right) to check for over training. The mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for training of the network (bottom left) and the mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for validating of the network (bottom right). The truth label in the mass regression was chosen as the GEN-matched reconstructed top mass for background and the GEN-matched reconstructed top squark mass for signal

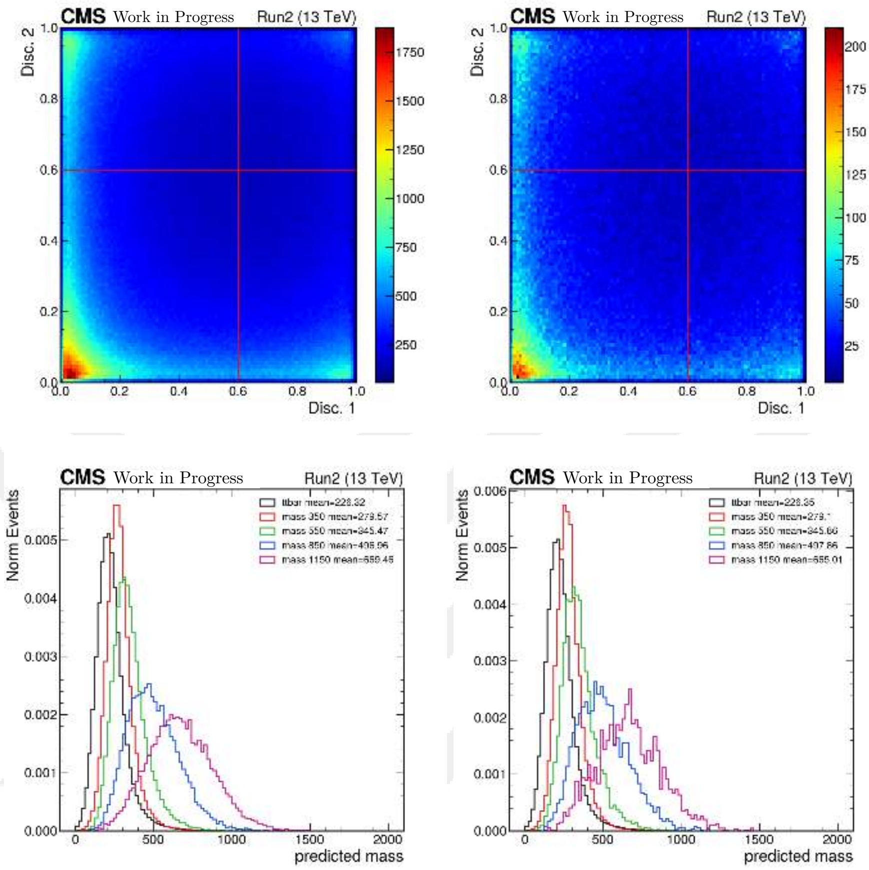


Figure 8.7. For the semi-leptonic channel, the 2D background discriminant output used for training of the network (top left), the 2D background discriminant output used for validating of the network (top right) to check for over training. The mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for training of the network (bottom left) and the mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for validating of the network (bottom right). The truth label in the mass regression was chosen as the GEN-matched reconstructed top mass for background and the GEN-matched reconstructed top squark mass for signal

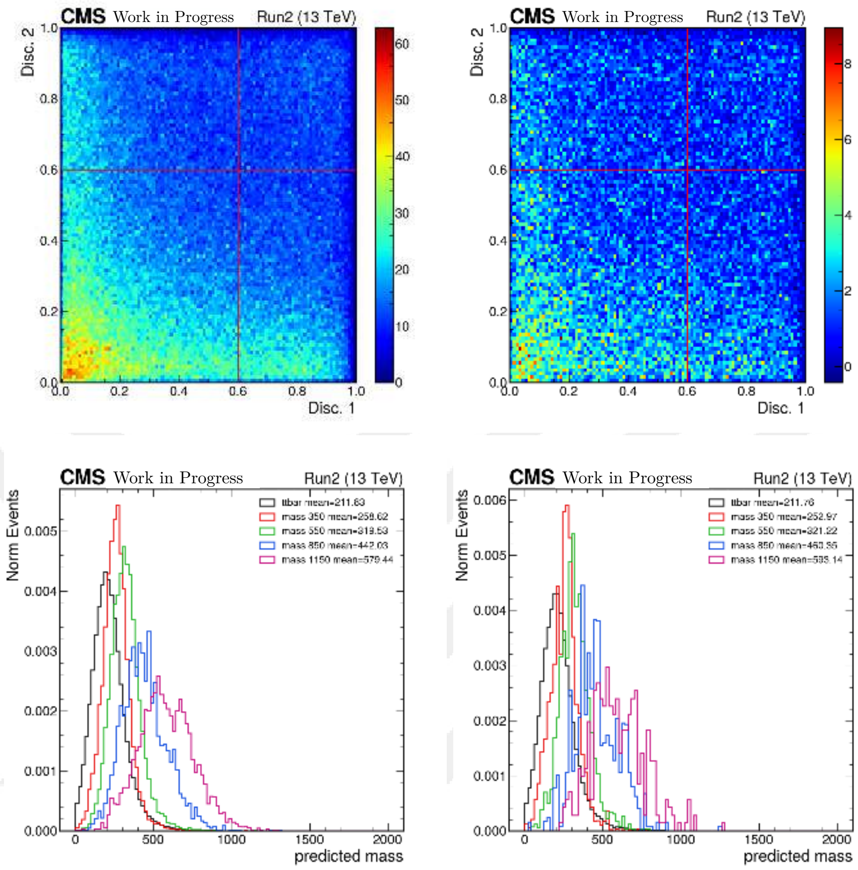


Figure 8.8. For the fully-leptonic channel, the 2D background discriminant output used for training of the network (top left), the 2D background discriminant output used for validating of the network (top right) to check for over training. The mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for training of the network (bottom left) and the mass regression output for background and several Stealth SY$\bar{Y}$ signal mass points used for validating of the network (bottom right). The truth label in the mass regression was chosen as the GEN-matched reconstructed top mass for background and the GEN-matched reconstructed top squark mass for signal

With the aforementioned configuration and setup, the NN is trained as specified and the resulting performance plots are obtained. Shown in Figure 8.9. for the fully-hadronic channel, Figure 8.10. for the semi-leptonic channel and Figure 8.11. for the fully-leptonic channel are the output of the mass regression for $t\bar{t}+$jets as well as for four signal mass points spanning the range of mass values for the network trained on the Stealth SY$\bar{Y}$ signal model. The regression is able to discern the difference between $t\bar{t}+$jets and each background and predicts a mass close to the “truth” value. This behavior is accomplished equally well for events in any N_{jets} bin. In general, the predicted mass is systematically lower than the “truth” mass, regardless of signal mass point or background or N_{jets}. As the accuracy of the actual predicted value is not more important than the difference in the value for different mass points, it is not troubling.

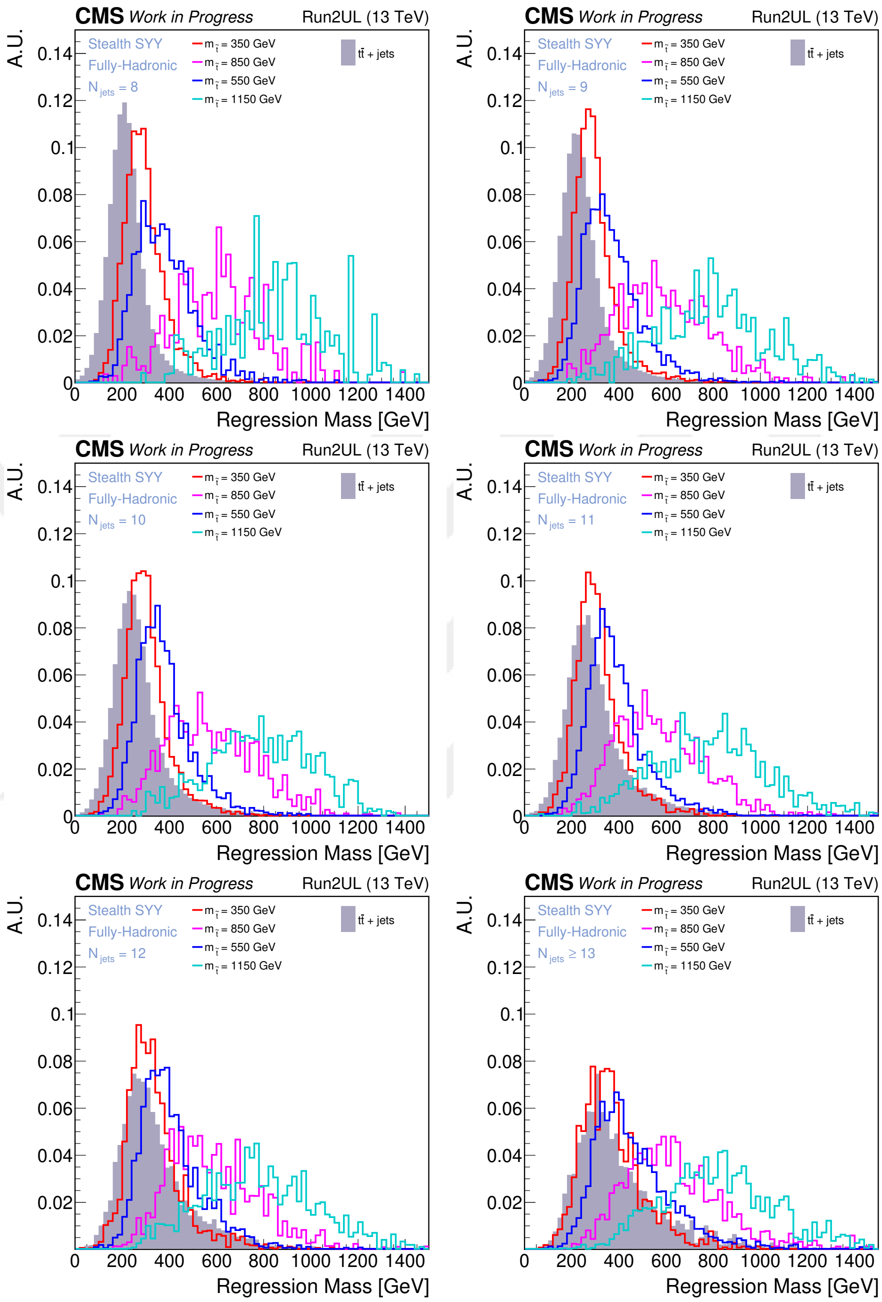


Figure 8.9. For the fully-hadronic channel, mass regression output (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 8,9,10,11,12$\geq$ 13 (left to right, top to bottom)

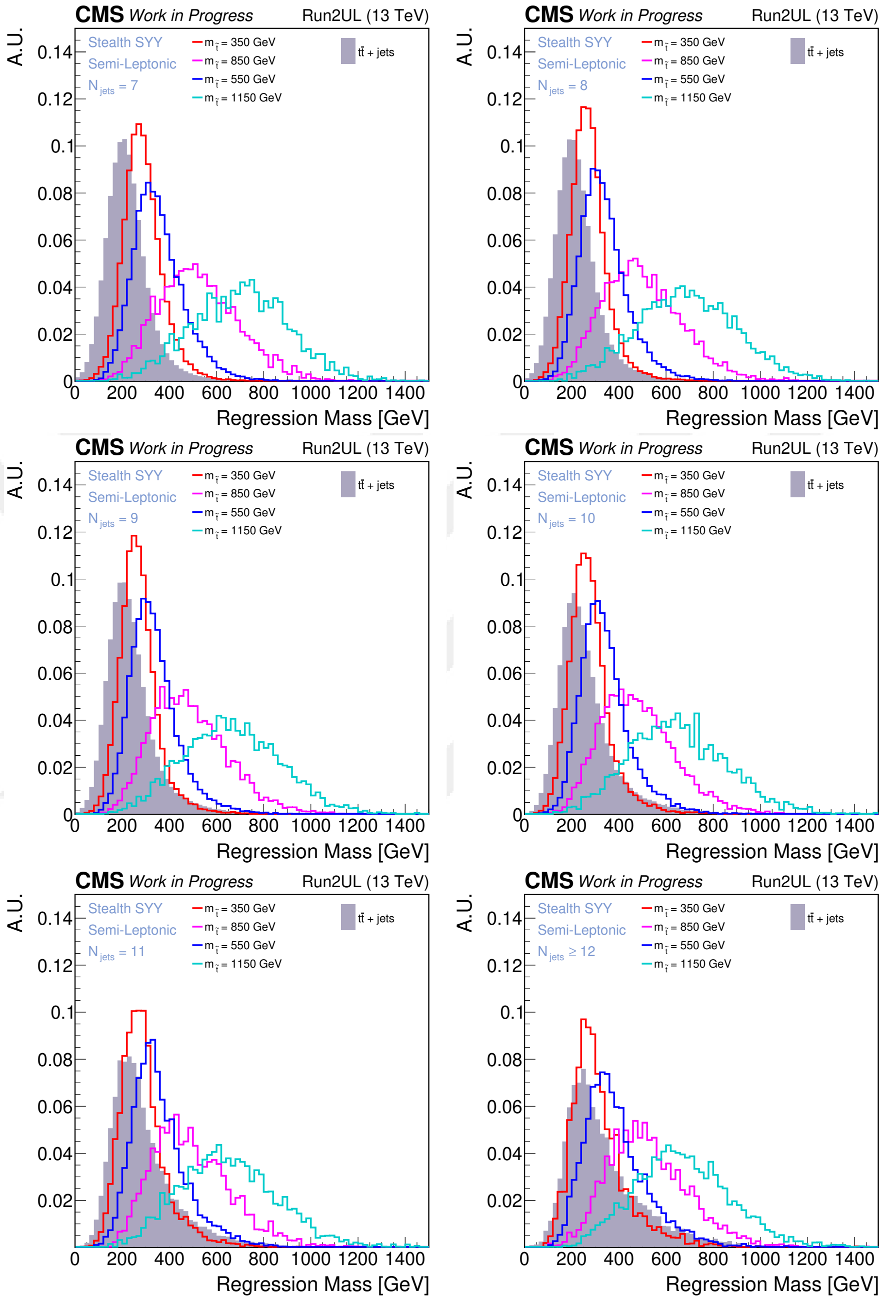


Figure 8.10. For the semi-leptonic channel, mass regression output (NN trained on Stealth SY$\bar{Y}$ signal) for $N_{\text{jets}} = 7,8,9,10,11,\geq 12$ (left to right, top to bottom)

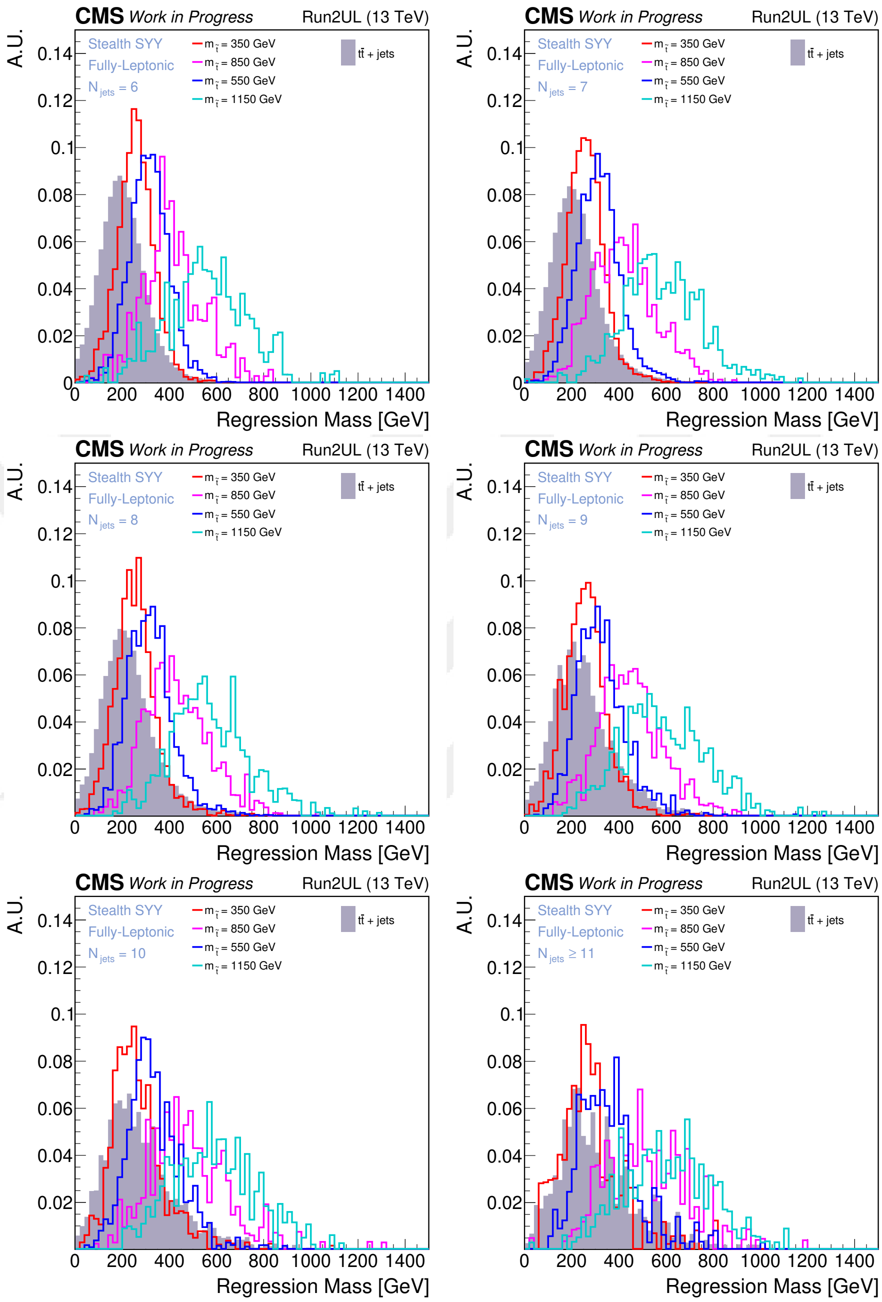


Figure 8.11. For the fully-leptonic channel, mass regression output (NN trained on Stealth SY$\bar{Y}$ signal) for $N_{\text{jets}} = 6{,}78{,}9{,}10{,}\geq 11$ (left to right, top to bottom)

The outputs of two classifiers are compared for $t\bar{t}+$jets and a few stop masses for the Stealth SY$\bar{Y}$ signal, shown in Figure 8.12. for the fully-hadronic channel, Figure 8.13. for the semi-leptonic channel, and Figure 8.14. for the fully-leptonic channel. Again, these results are for the networks trained on the Stealth SY$\bar{Y}$ signal model. It is demsontrated that the classifiers can discern signal from background in a similar manner for events with any number of N_{jets}. Compared to one another, both of the classifiers perform equally well, denoted by the similar shape of the $t\bar{t}+$jets distributions.

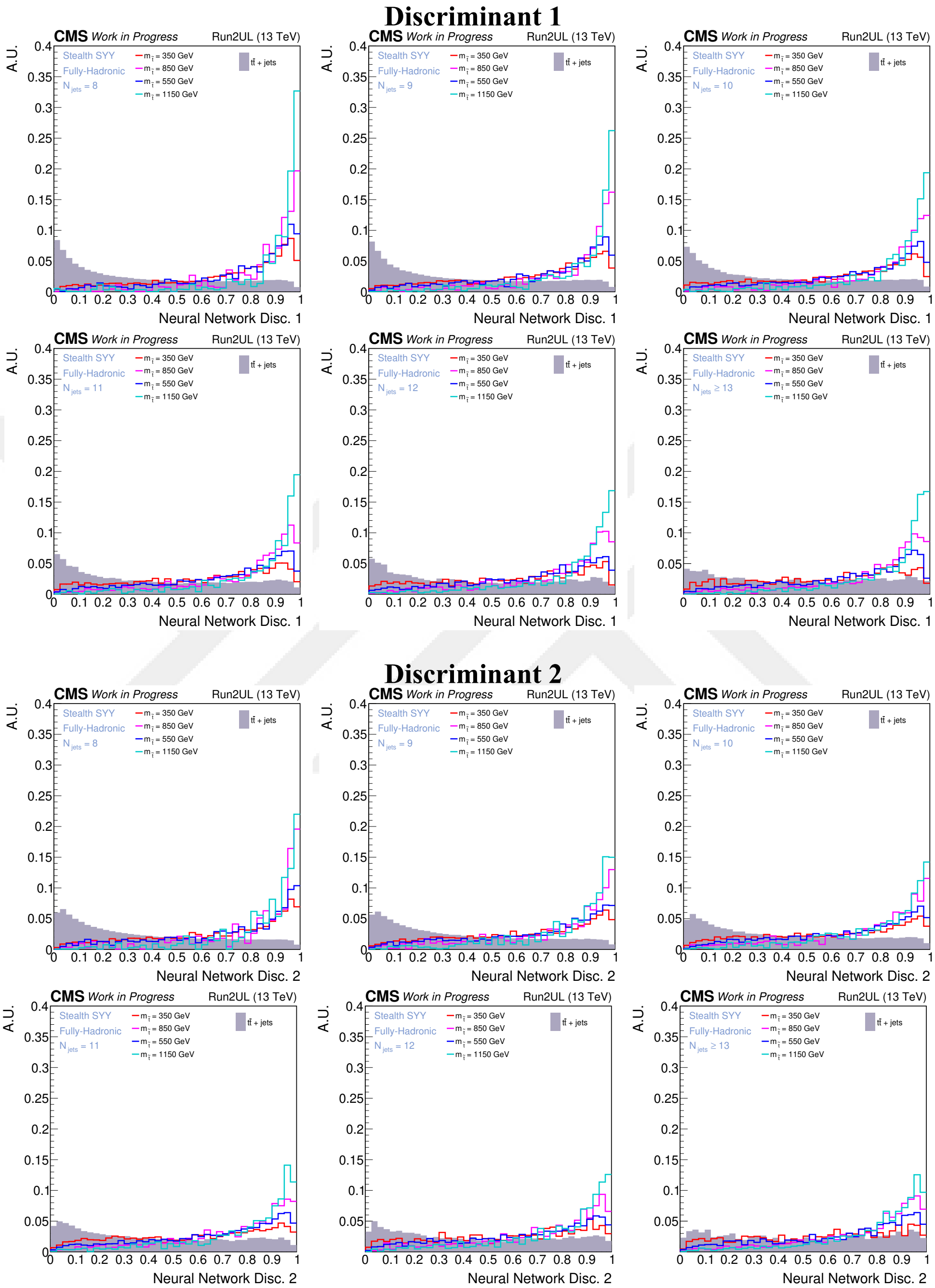


Figure 8.12. For the fully-hadronic channel, Discriminant 1 and 2 outputs (NN trained on Stealth SY$\bar{Y}$ signal) for $N_{\text{jets}} = 8,9,10,11,12,\geq 13$ (left to right, top to bottom)

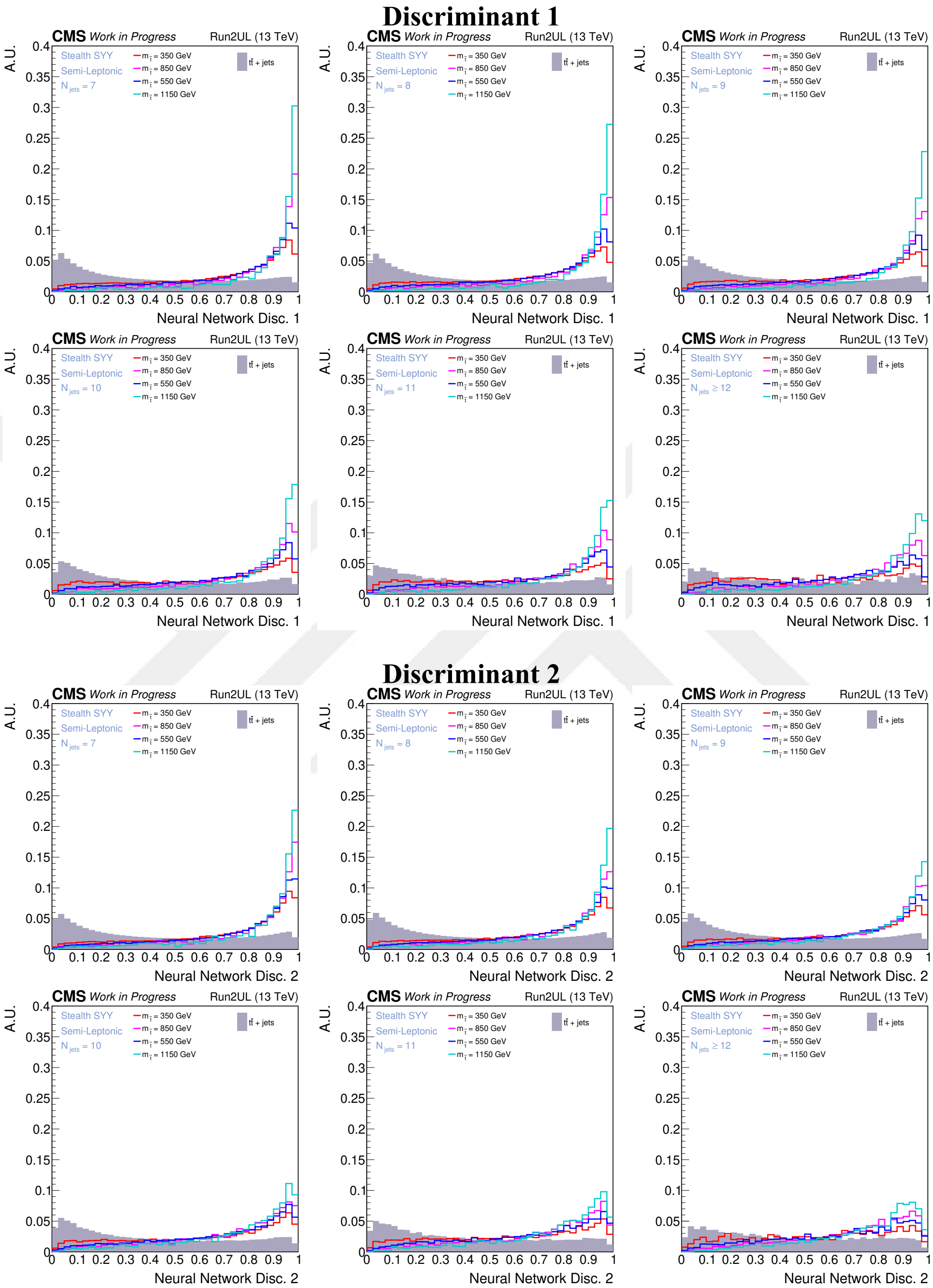


Figure 8.13. For the semi-leptonic channel, Discriminant 1 and 2 outputs (NN trained on Stealth SY$\bar{Y}$ signal) for $N_{\text{jets}} = 7,8,9,10,11,\geq 12$ (left to right, top to bottom)

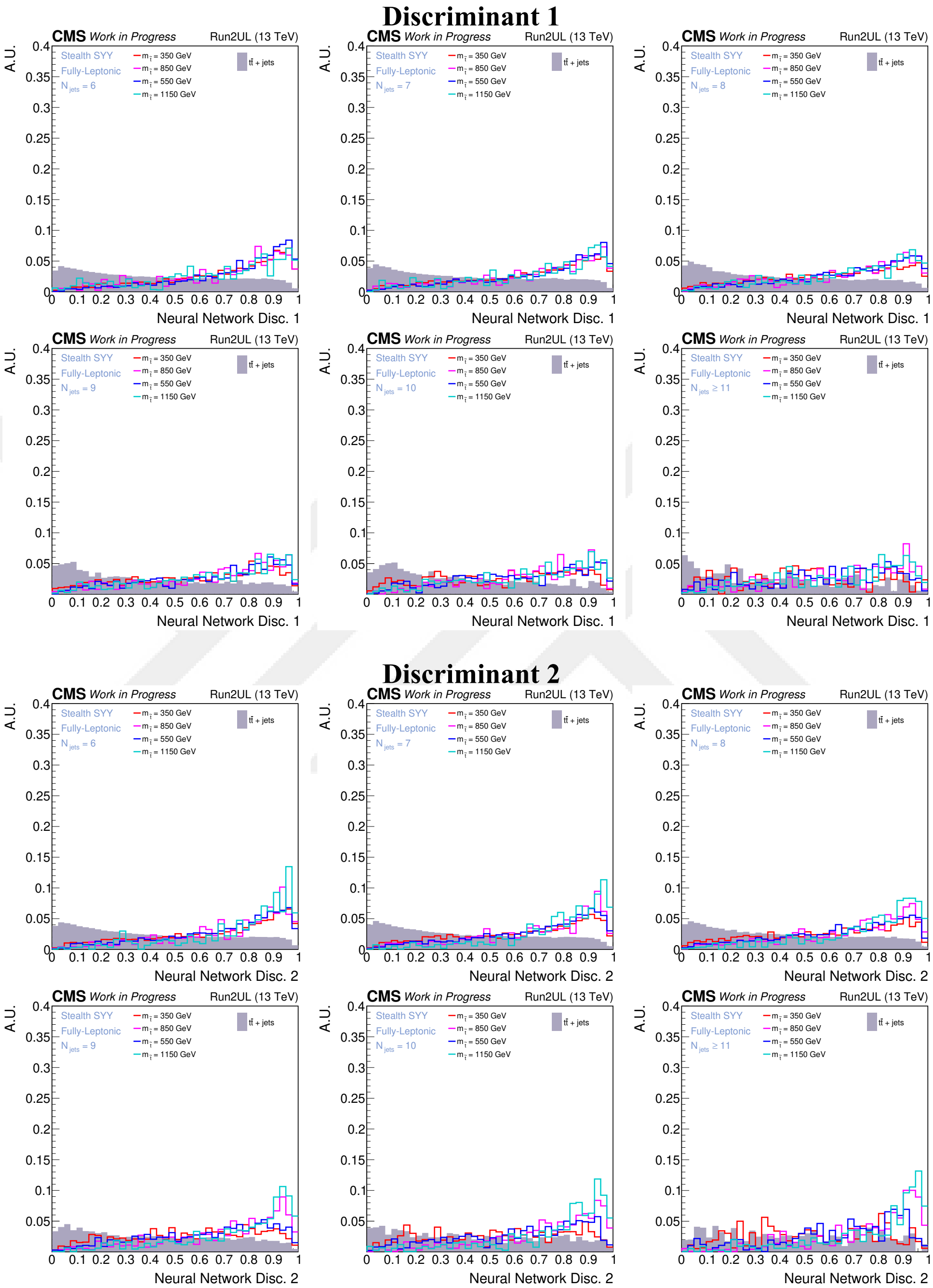


Figure 8.14. For the fully-leptonic channel, Discriminant 1 and 2 outputs (NN trained on Stealth SY$\bar{Y}$ signal) for N_{jets} = 6,7,8,9,10,$\geq$ 11 (left to right, top to bottom)

The ability for the network to distinguish signal from background is shown in Figure 8.15. as Receiver-Operator Curves (ROCs). Figure 8.15. shows that the network, for specific signal mass $m_{\tilde{t}} = 550$ GeV, achieves roughly equal performance for events with any $N_{\rm jets}$. These patterns of behavior can be attributed to the fact that the network is shown equal numbers of events from each $N_{\rm jets}$ category, which prevents the network from optimizing for a category with many events at the expense of a category with fewer events that it sees less often.

The another type ROC curves are shown in Figure 8.16., Figure 8.17. and Figure 8.18. for the fully-hadronic, semi-leptonic and fully-leptonic channels, respectively. The different mass points of the Stealth SY$\bar{Y}$ model for a given $N_{\rm jets}$ are compared. These ROC plots demonstrate that the network performs best for higher $m_{\tilde{t}}$ signals, regardless of $N_{\rm jets}$. Likewise, the network performs equally well for either of the two classifiers.

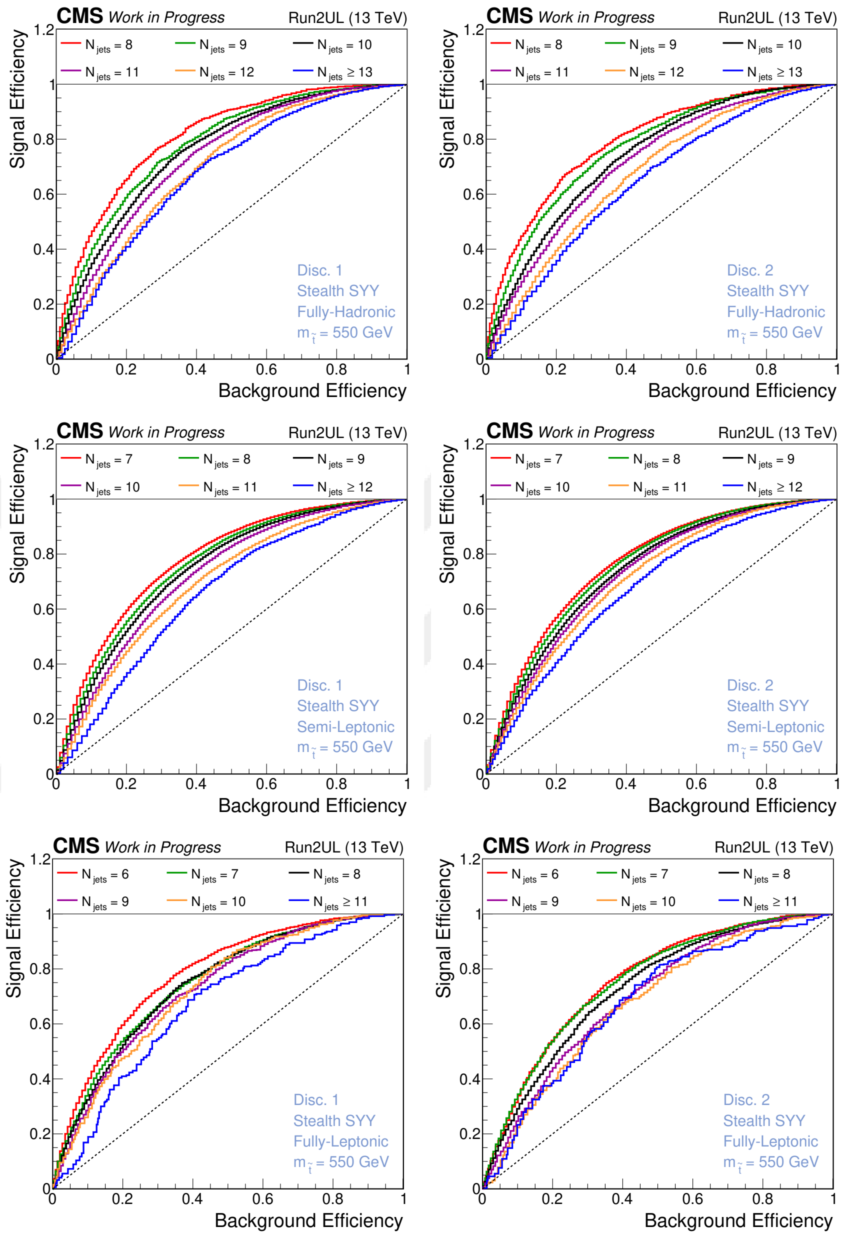


Figure 8.15. For the fully-hadronic (top), semi-leptonic (middle), and fully-leptonic (bottom) channels, the ROC plot for discriminant 1 (left) and discriminant 2 (right), for each N_{jets} separately assuming Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ GeV as the signal. Performance is similar for the different N_{jets} categories and between the two discriminants

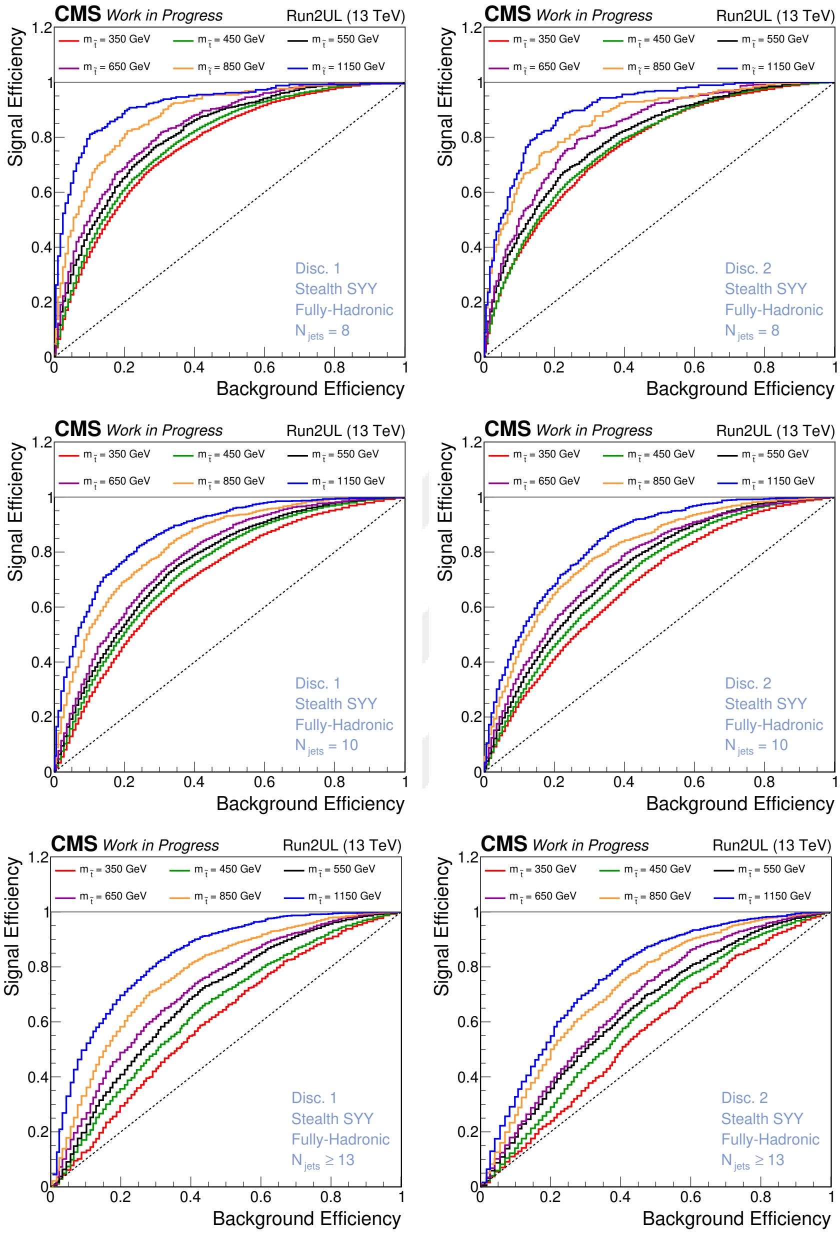


Figure 8.16. For the fully-hadronic channel, the ROC for discriminant 1 (left) and discriminant 2 (right) for $N_{\text{jets}} = 8, 10, \geq 13$ (top to bottom), showing results for several Stealth SY$\bar{Y}$ signal mass points

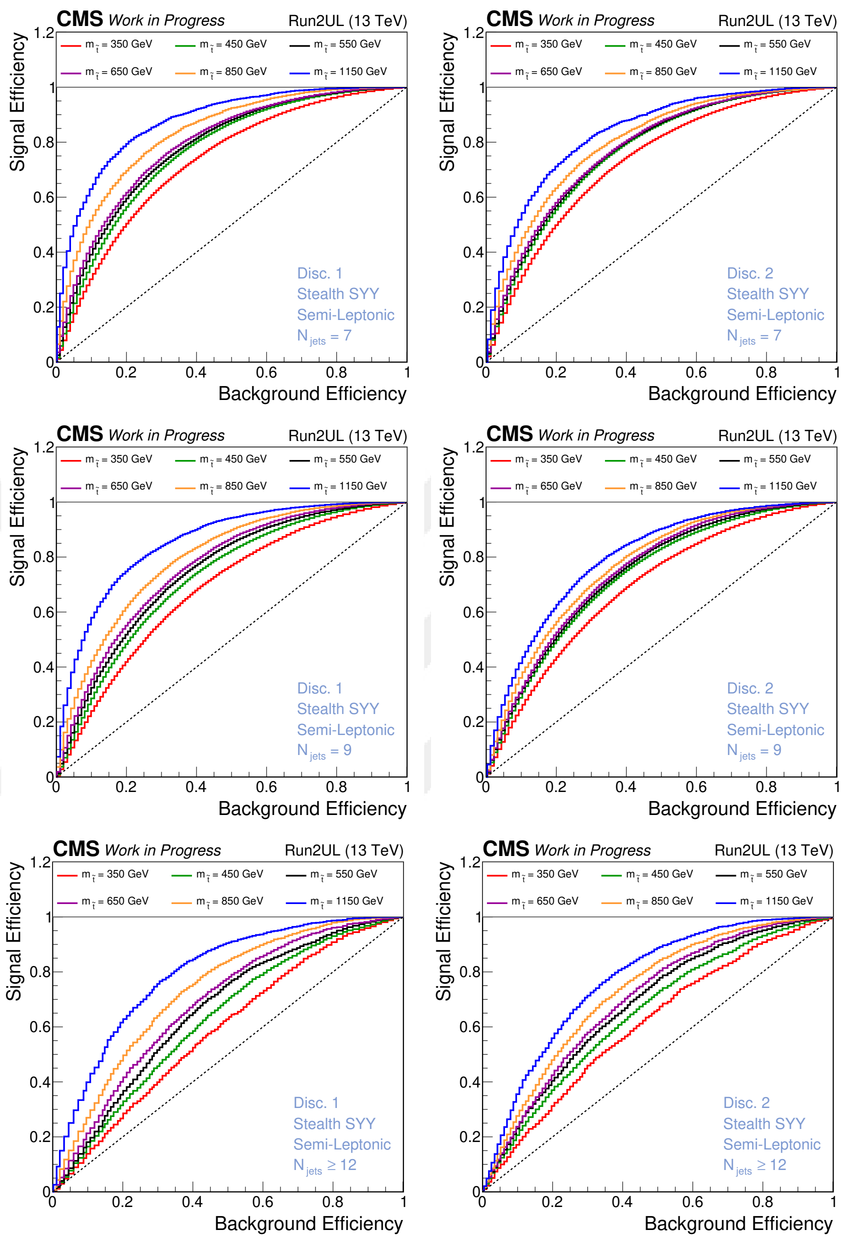


Figure 8.17. For the semi-leptonic channel, the ROC for discriminant 1 (left) and discriminant 2 (right) for $N_{\text{jets}} = 7,9,\geq 12$ (top to bottom), showing results for several Stealth SY$\bar{Y}$ signal mass points

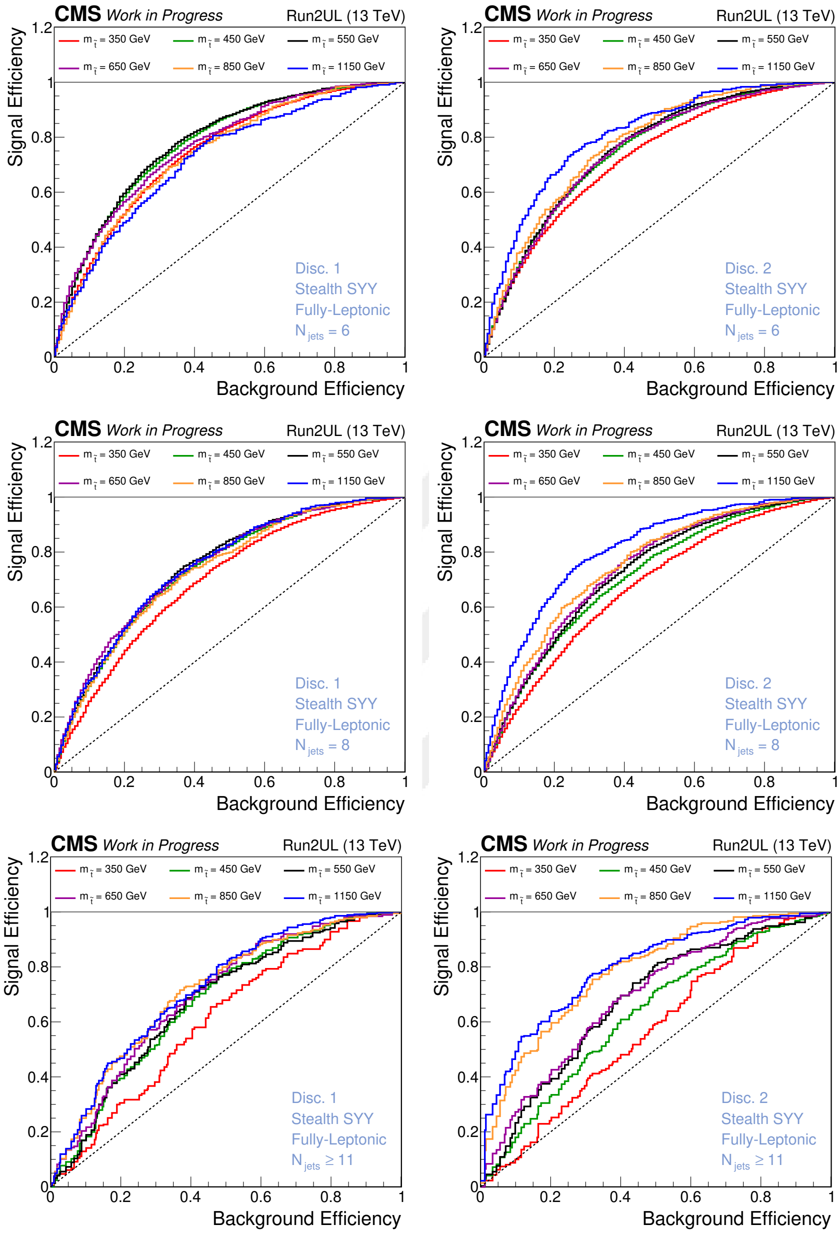


Figure 8.18. For the fully-leptonic channel, the ROC for discriminant 1 (left) and discriminant 2 (right) for $N_{\text{jets}} = 6, 8, \geq 11$ (top to bottom), showing results for several Stealth SY$\bar{Y}$ signal mass points

A better qualitative understanding of the 2D discriminant plane is revealed in Figure 8.19., Figure 8.20., and Figure 8.21. for both $t\bar{t}+$jets and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ GeV. The background has a desirable shape, peaking towards (0.0, 0.0) and then both discriminants decaying equally towards (1.0, 1.0). This behavior is conducive to performing the ABCD method. In addition, the signal peaks towards the (1.0, 1.0) corner, showing some separation from background. Both of these behaviors for background and signal also hold across the N_{jets} bins, again, a result of showing same number of N_{jets} events during training. This 2D plane is used to validating the ABCD method, given in next part.

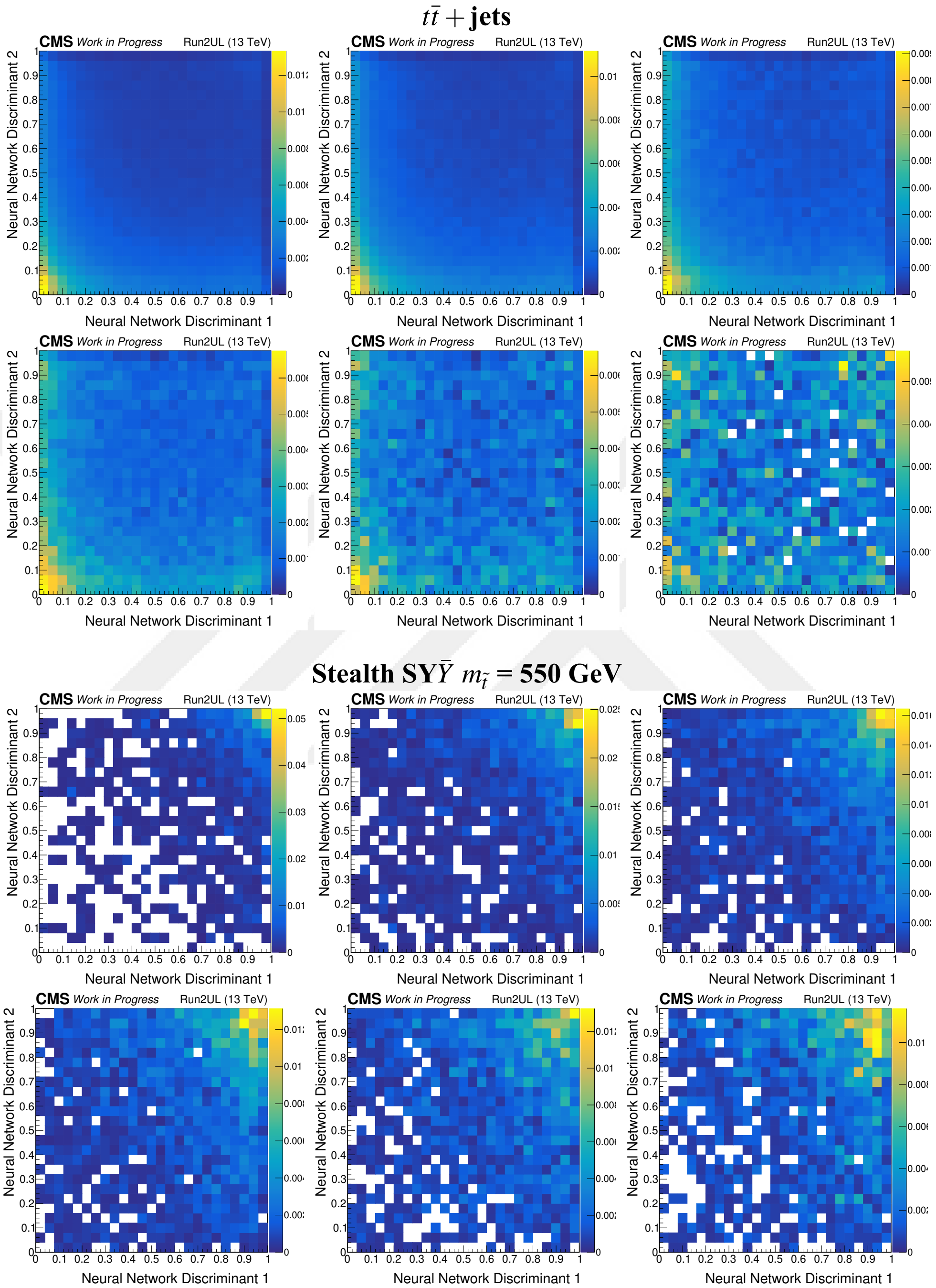


Figure 8.19. For the fully-hadronic channel, the 2D discriminant plane for $t\bar{t}+$jets (top two rows) and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ (bottom two rows) separated in $N_{\text{jets}} = 8,9,10,11,12,\geq 13$ (left to right, top to bottom)

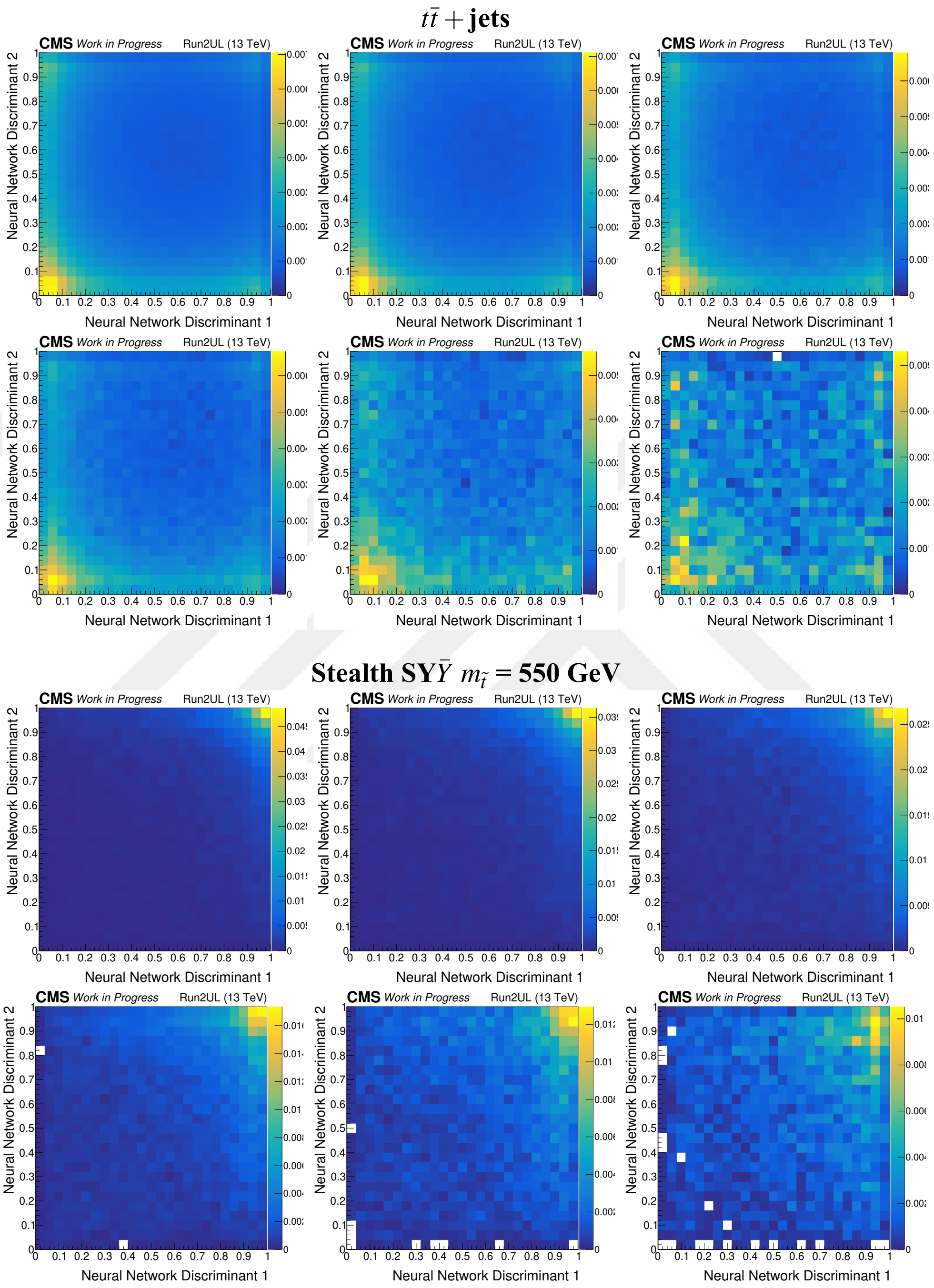
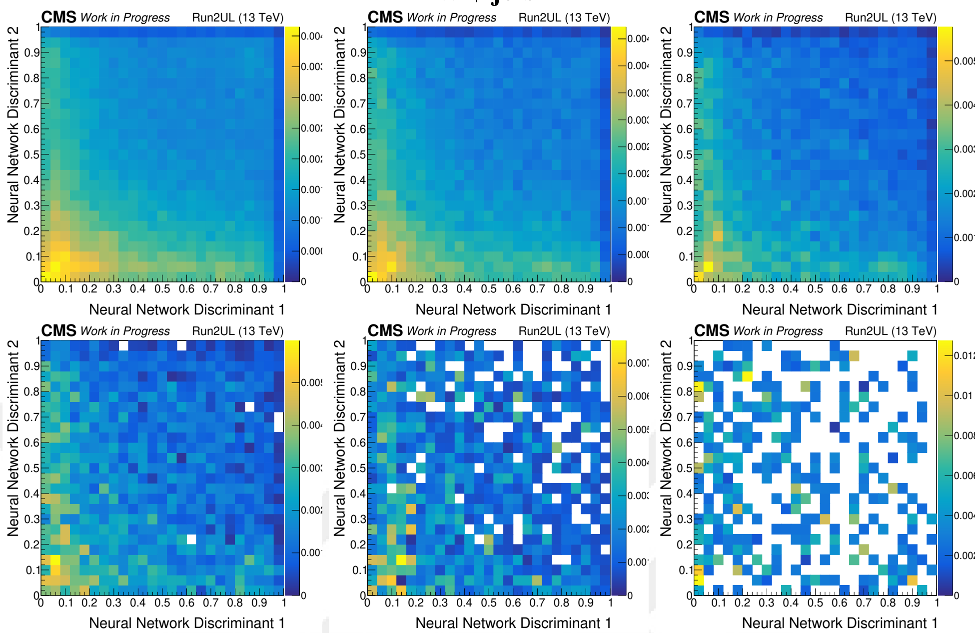


Figure 8.20. For the semi-leptonic channel, the 2D discriminant plane for $t\bar{t}+$jets (top two rows) and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ (bottom two rows) separated in $N_{\text{jets}} = 7,8,9,10,11,\geq 12$ (left to right, top to bottom)

$t\bar{t}+$jets



Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV

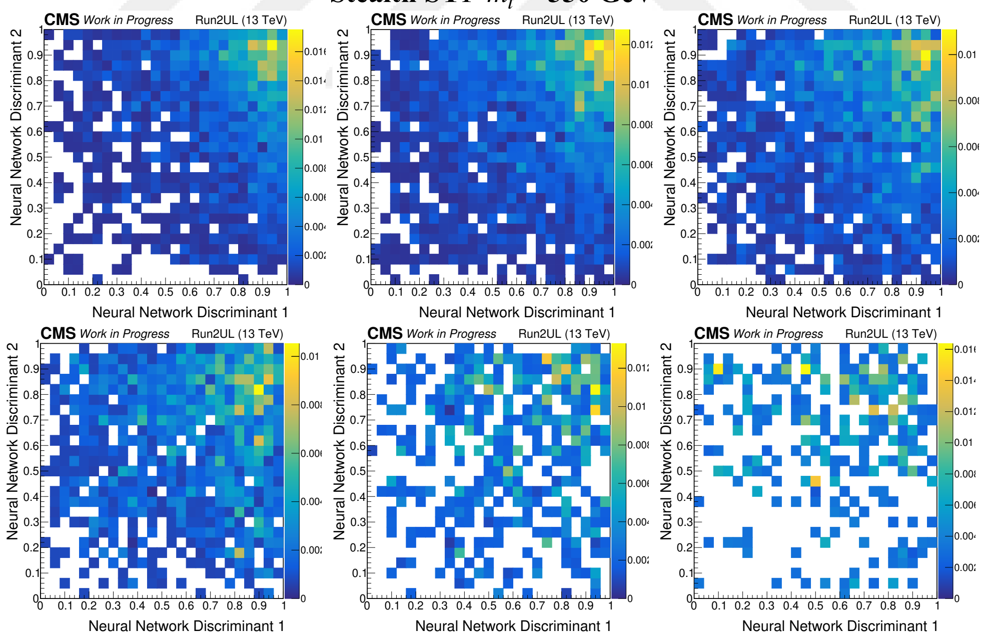


Figure 8.21. For the fully-leptonic channel, the 2D discriminant plane for $t\bar{t}+$jets (top two rows) and Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ (bottom two rows) separated in $N_{\text{jets}} = 6,7,8,9,10,\geq 11$ (left to right, top to bottom)

8.4. Validating the ABCD Method

The importance of validating the ABCD method is to prove that the ABCD method works well for estimating the $t\bar{t}+$jets events in data. For validating the method, a couple of steps are needed. One of the steps is that investigating how robust the ABCD method is in simulation and in the same time clarifying that the Double DisCo NN behaviour is good enough. For this step, the ABCD bin edges are chosen to divide the 2D plane formed from network's output as **A**, **B**, **C**, **D** regions. If the ABCD method is robust, then validation regions can be defined to check the method outside of the main signal region **A** and via these validation regions $t\bar{t}+$jets events in data is estimated.

Note that all the results for validating the ABCD method are for Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV model.

8.4.1. Defining the ABCD Regions

Before defining the final ABCD bin edges, an initial study was done to test how the ABCD method works in each N_{jets} bin. With this initial study we had some idea how we optimally choose the final ABCD edges and have a general view of the robustness of the ABCD method. Then, the ABCD bin edges were optimized based on some criterias. The details of this initial study and the optimization are given in the following subsections.

Initial Study for Understanding How to Optimize the ABCD Bin Edges

For testing the robustness of the ABCD method, some quantities can be mapped as a function of choice of bin edges in the Disc. 1 vs Disc. 2 2D plane. From lower to higher N_{jets} bins, referring to Equation 8.9, non-closure should be close to zero in the sense of the method working perfectly when non-closure is exactly zero.

Pull is one of the important quantity to understand how significant the ABCD closure is. It can be calculated in terms of number of predicted and observed events in region **A** with their (statistical only) uncertainties, as shown in Equation 8.12. The pull is the difference between the number of predicted and number of observed events, which is scaled by the uncertainties on both. The uncertainties are counted as a part of the pull, since they give a sense of how significant a certain value of closure is.

$$\text{pull } = (N_{\text{pred.},A} - N_{\text{obs.},A}) / \sqrt{\delta^2_{\text{pred.}} + \delta^2_{\text{obs.}}} \tag{8.12}$$

where $N_{\text{pred.},A}$ is number of predicted events in region **A**; $N_{\text{obs.},A}$ is number of observed events in region **A**; $\delta_{\text{pred.}}$ is uncertainty on the predicted background events in region A and $\delta_{\text{obs.}}$ is uncertainty on the observed background events in region **A**.

To understand how much signal sensitivity is achievable with the ABCD method, significance is also calculated as a function of choice of ABCD bin edges, where N_{sig} is number of signal events; *sys.* is 30% uncertainty in the total number of events; N_{bkg} is number of background events.

$$\text{significance} = N_{sig} \ / \ \sqrt{N_{bkg} + (sys. * N_{bkg})^2 + (non - closure * N_{bkg})^2} \tag{8.13}$$

Figure 8.22. shows non-closure and pull in the 2D plane of bin edge choices for the fully-hadronic, semi-leptonic, fully-leptonic channels. The non-closure and pull are given for $N_{\text{jets}} = 8$ (for fully-hadronic), $N_{\text{jets}} = 7$ (for semi-leptonic), $N_{\text{jets}} = 6$ (for fully-leptonic) as examples. Overall, the non-closure and pull are $\sim 10\%$ or less across a majority of the 2D plane. The non-closure and pull for the full N_{jets} set are given in Appendix D1.1.. For the higher N_{jets} bins, the quantities are affected by low number of event counts.

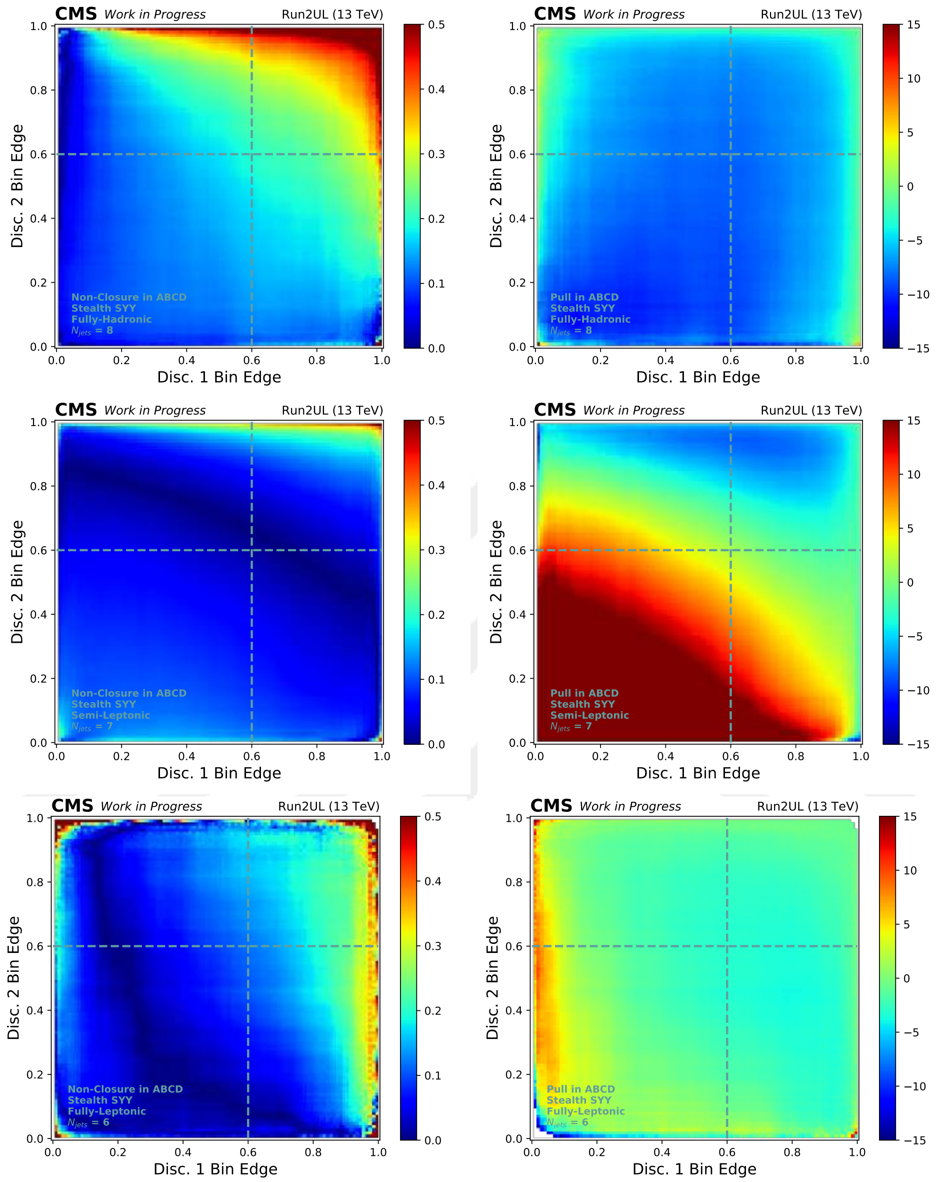


Figure 8.22. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the non-closure (left) and pull (right) as a function of "non-optimized" ABCD bin edges for $N_{\text{jets}} = 8$ (for fully-hadronic channel), $N_{\text{jets}} = 7$ (for semi-leptonic channel) and $N_{\text{jets}} = 6$ (for fully-leptonic channel). Overall non-closure and pull are about ~ 0 across a majority of the 2D plane of ABCD edge choices

Figure 8.23., Figure 8.24., Figure 8.25. show significance in 2D plane for the fully-hadronic, semi-leptonic, fully-leptonic channels, respectively. There is fractionally more signal events at high N_{jets}, so it is expected to see higher significance for these bins. For lower N_{jets} bins, on the top right corner in 2D plane, it is still reasonable to see the significance higher than 0. Because, as the NN places signal events in the upper righthand corner, this region is the signal region.

Based on the observed non-closure, pull and significance, the “non-optimized” ABCD bin edges are chosen as (0.600, 0.600) for all fully-hadronic, semi-leptonic and fully-leptonic channels. Then **A**, **B**, **C**, and **D** regions are defined with the choice of the “non-optimized” ABCD bin edges. Another quantity to understand signal sensitivity is signal fraction, while also giving information about how well network can separate signal from background. The signal fraction is calculated in each **A**, **B**, **C**, and **D** region as:

$$\text{sigFrac in i} = \left[\frac{N_{sig,i}}{N_{tot,i}}\right], \; N_{tot,i} = [N_{sig,i} + N_{bkg,i}] \tag{8.14}$$

where i is each **A**, **B**, **C**, and **D** region; $N_{sig,i}$ is number of signal events in region i; $N_{bkg,i}$ is number of background events in region i and $N_{tot,i}$ is total number of events in region i. Lower signal fractions are desired in each **B**, **C**, and **D** region for lower N_{jets} bins since they are control regions, where there are much more background events than signal events. The $t\bar{t}+$jets fraction is calculated as the ratio between number of total $t\bar{t}+$jets events and number total background events in a given N_{jets} bin.

With all these checks for each of the quantities above, the choice of the non-optimized ABCD bin edges was very useful to figure out how optimizes the ABCD bin edges. In the same time, parallel to the optimizing the edges, it was used to explore on calculating the systematics.

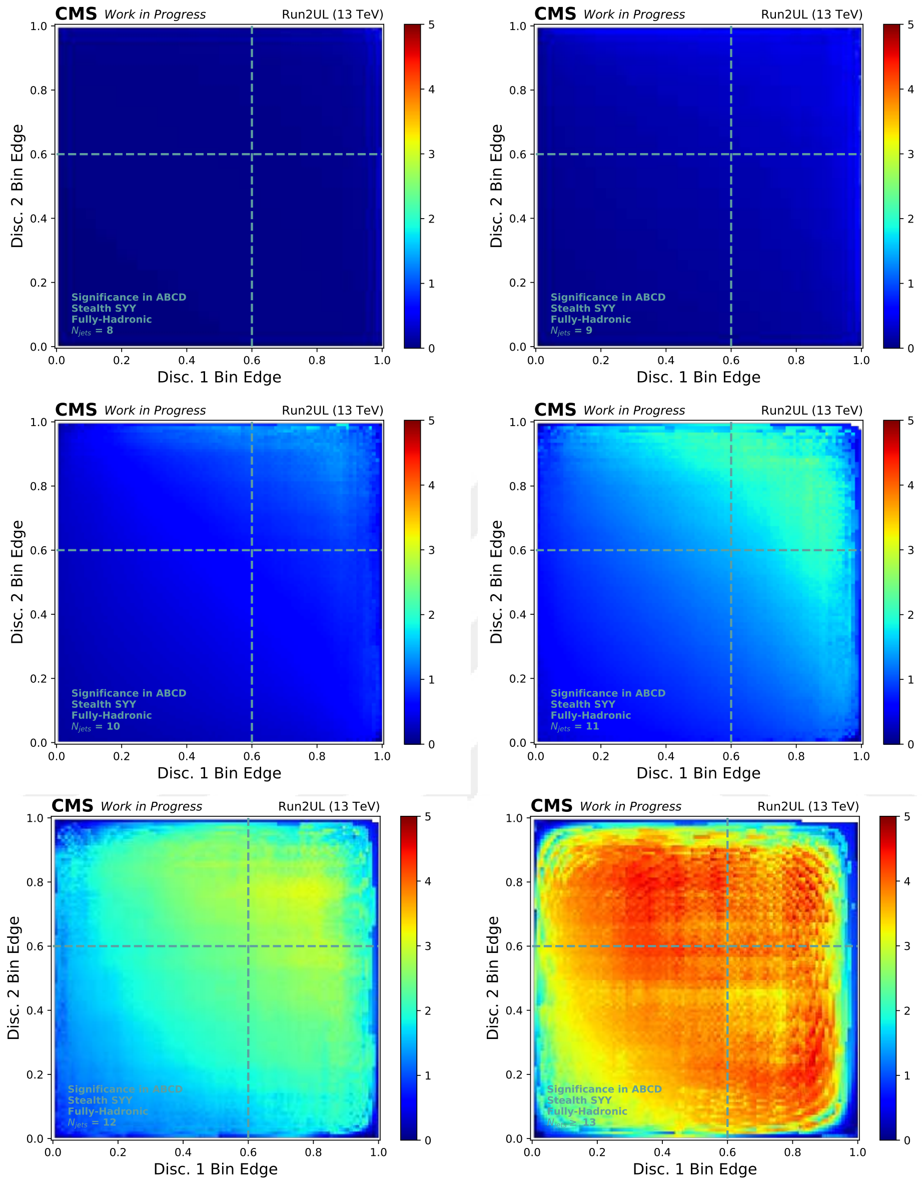


Figure 8.23. For the fully-hadronic channel, significance as a function of "non-optimized" ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 8,9,10,11,12, $\geq$ 13 bins (left to right, top to bottom)

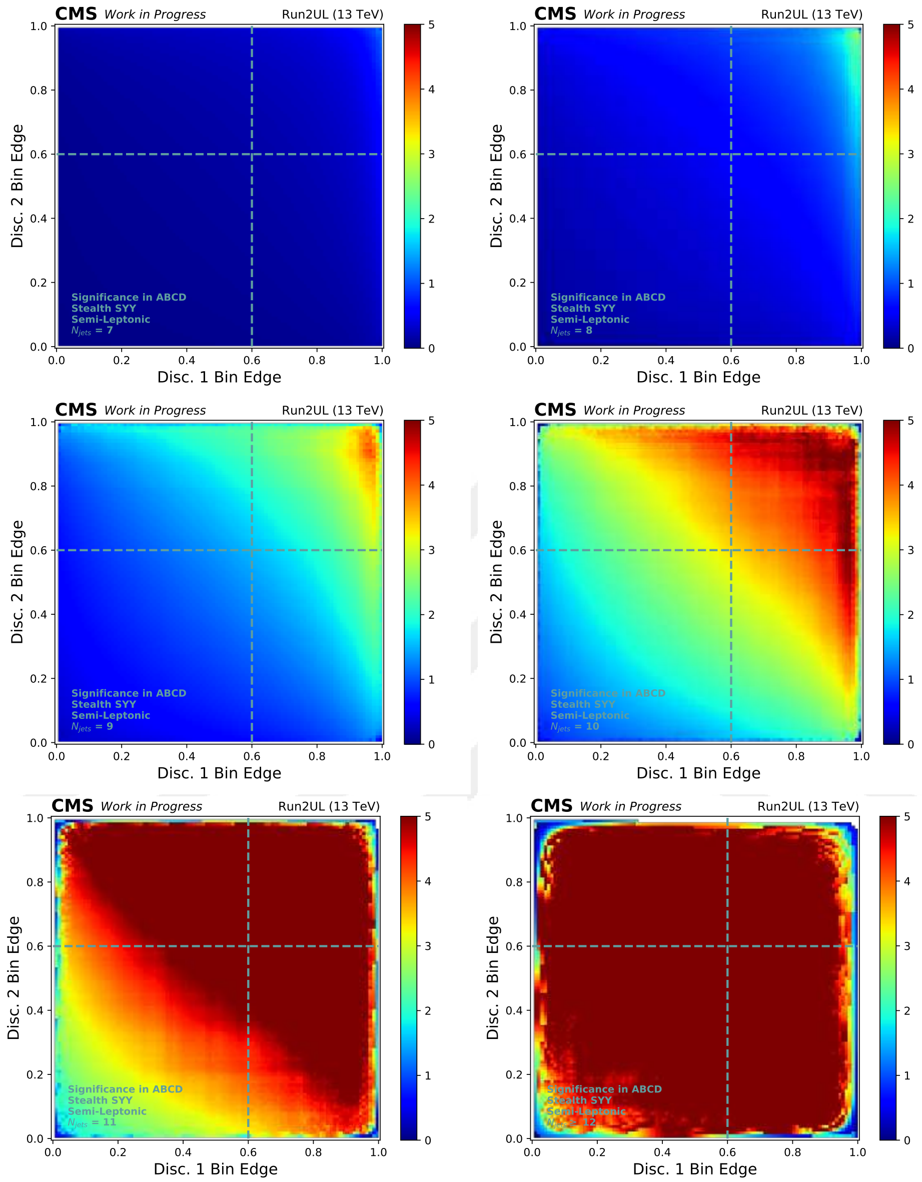


Figure 8.24. For the semi-leptonic channel, significance as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. $N_{\text{jets}} = 7,8,9,10,11,\geq 12$ bins (left to right, top to bottom)

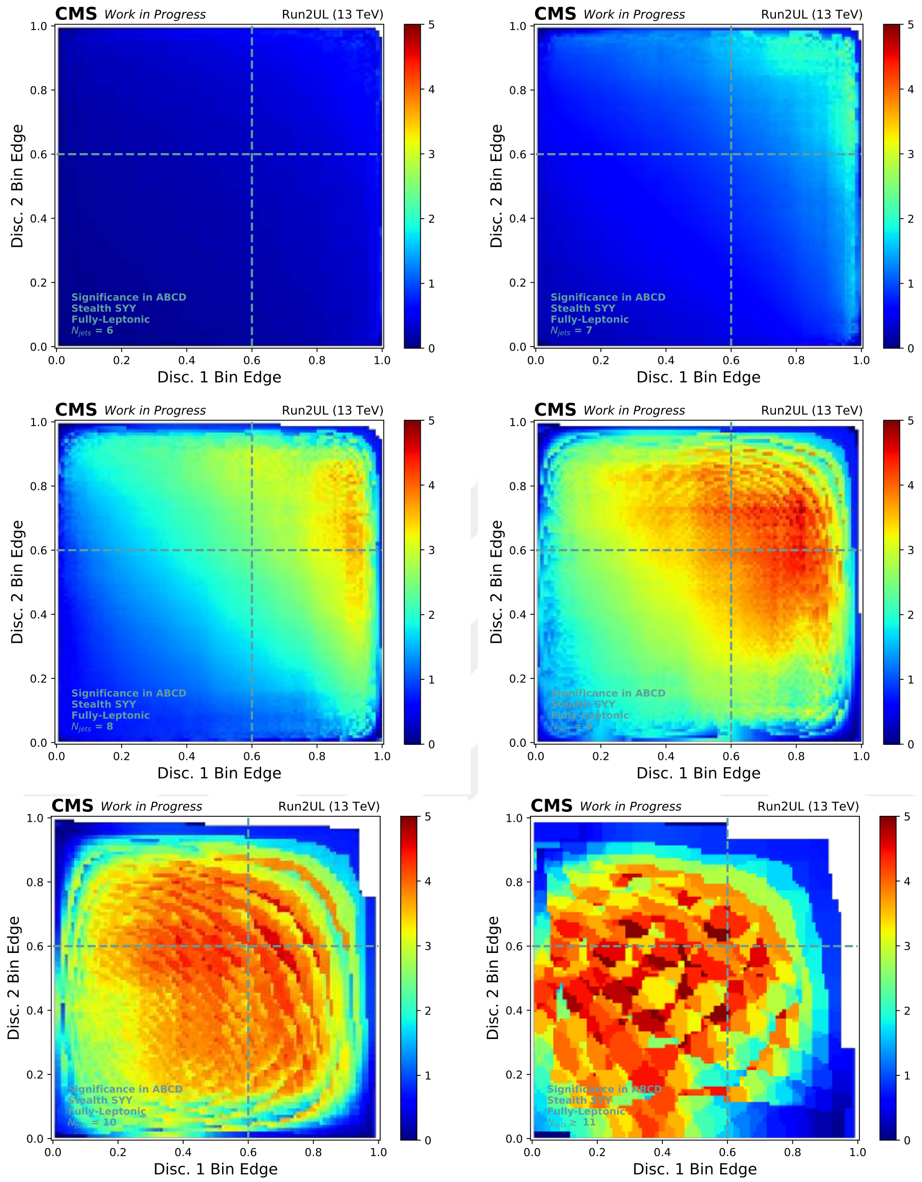


Figure 8.25. For the fully-leptonic channel, significance as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. $N_{\text{jets}} = 6,7,8,9,10,\geq 11$ bins (left to right, top to bottom)

Opimizing the ABCD Bin Edges

Based on the initial study in the previous part, reasonable non-closure value is less than 30% and pull value is less than 2.0 for all fully-hadronic, semi-leptonic and fully-leptonic channels. With the choice of "non-optimized" ABCD edges (0.600, 0.600), the maximum signal fractions among the **B**, **C**, and **D** regions are about $\sim 50\%$ for all of the channels—just for Stealth SY$\bar{Y}$ model. Considering all of these observed values for the quantities of interest, an optimization metric is constructed and applied for determining the optiminal choice of ABCD bin edges.

The first step of applying the optimization metric is to go through all bin edges options in steps of 0.01 between 0.5 and 0.95 along Disc. 1 and Disc. 2 edge values. Then, non-closure $< 30\%$ and pull < 2.0 requirements are OR'd together for all N_{jets} bins simultaneously to balance the contributions from lower and higher N_{jets} bins. Since, the OR'ing of these conditions is more robust against statistical fluctuations that could drive non-closure far from 0, most likely in high N_{jets} bins. Due to the inherent differences between the all three channels, the fraction of the $t\bar{t}+$jets is expected to be different for each channel and this ends up with different signal contaminations. Considering the differences based on channels and signal models, the signal fraction (as defined in Equation 8.14) in each **B**, **C**, and **D** region is required to be less than 100% simultaneously for all N_{jets} bins for a given channel. Among all the bin edge options that satisfy the above selections, the best ABCD bin edges choice is the one with the maximum significance, adding per-N_{jets} significances in quadrature. While calculating the maximum significance for Stealth SY$\bar{Y}$ model, $m_{\tilde{t}}$ = 550 GeV is applied. After applying this metric, the top four choices are ranked by significance in Table 8.3., Table 8.4., Table 8.5. for the fully-hadronic, semi-leptonic and fully-leptonic channels, respectively. Also, the choice of the non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance and added to the each table.

Table 8.3. For the fully-hadronic channel, top four choices for the ABCD bin edges, ranked by significance. Also, the choice of non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance. For each ABCD bin edges choice, the values of the metric elements are also listed for each N_{jets} bin. Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV is applied while optimizing the edges

ABCD edges	Significance	N_{jets}	Non-closure	Pull	Sig. Fracs B	Sig. Fracs C	Sig. Fracs D
('0.600', '0.600')	4.923	8	0.195	-8.046	0.002	0.002	0.001
		9	0.054	-1.478	0.009	0.011	0.003
		10	0.021	-0.366	0.034	0.043	0.011
		11	0.002	0.020	0.093	0.106	0.032
		12	0.009	-0.058	0.190	0.222	0.095
		≥ 13	0.227	-1.014	0.370	0.429	0.196
('0.850', '0.740')	5.903	8	0.297	-6.275	0.004	0.005	0.001
		9	0.170	-2.343	0.020	0.020	0.004
		10	0.134	-1.214	0.062	0.069	0.019
		11	0.077	-0.438	0.151	0.152	0.055
		12	0.118	-0.408	0.283	0.292	0.136
		≥ 13	0.211	-0.517	0.472	0.500	0.290
('0.810', '0.720')	5.889	8	0.272	-6.664	0.004	0.004	0.001
		9	0.144	-2.329	0.018	0.018	0.004
		10	0.106	-1.121	0.057	0.062	0.018
		11	0.073	-0.481	0.141	0.147	0.051
		12	0.075	-0.298	0.263	0.271	0.131
		≥ 13	0.176	-0.498	0.453	0.483	0.275
('0.800', '0.690')	5.879	8	0.256	-6.758	0.003	0.003	0.001
		9	0.122	-2.099	0.017	0.016	0.004
		10	0.113	-1.289	0.054	0.058	0.017
		11	0.052	-0.367	0.134	0.139	0.048
		12	0.010	-0.042	0.259	0.262	0.126
		≥ 13	0.180	-0.540	0.436	0.464	0.267
('0.800', '0.740')	5.865	8	0.278	-6.755	0.004	0.004	0.001
		9	0.141	-2.232	0.018	0.018	0.004
		10	0.109	-1.133	0.056	0.062	0.018
		11	0.066	-0.427	0.140	0.149	0.052
		12	0.004	-0.017	0.261	0.266	0.134
		≥ 13	0.235	-0.666	0.458	0.484	0.277

Table 8.4. For the semi-leptonic channel, top five choices for the ABCD bin edges, ranked by significance. Also, the choice of non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance. For each ABCD bin edges choice, the values of the metric elements are also listed for each N_{jets} bin. Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV is applied while optimizing the edges

ABCD edges	Significance	N_{jets}	Non-closure	Pull	Sig. Fracs B	Sig. Fracs C	Sig. Fracs D
('0.600', '0.600')	9.811	7	0.014	1.705	0.007	0.008	0.002
		8	0.023	1.607	0.026	0.027	0.008
		9	0.012	0.449	0.076	0.081	0.025
		10	0.027	-0.528	0.181	0.178	0.068
		11	0.085	-0.815	0.357	0.342	0.152
		≥ 12	0.247	-1.398	0.536	0.518	0.308
('0.700', '0.850')	12.186	7	0.093	-6.709	0.011	0.014	0.003
		8	0.071	-3.017	0.040	0.048	0.012
		9	0.070	-1.574	0.110	0.130	0.039
		10	0.113	-1.264	0.249	0.272	0.099
		11	0.172	-0.928	0.434	0.449	0.215
		≥ 12	0.252	-0.757	0.630	0.601	0.395
('0.810', '0.840')	12.168	7	0.107	-6.524	0.014	0.016	0.004
		8	0.083	-2.986	0.050	0.053	0.014
		9	0.089	-1.715	0.133	0.142	0.045
		10	0.156	-1.519	0.288	0.288	0.112
		11	0.176	-0.805	0.475	0.470	0.235
		≥ 12	0.282	-0.747	0.651	0.606	0.423
('0.690', '0.840')	12.144	7	0.083	-6.218	0.011	0.014	0.003
		8	0.064	-2.838	0.039	0.046	0.012
		9	0.064	-1.504	0.107	0.127	0.038
		10	0.117	-1.379	0.245	0.268	0.096
		11	0.183	-1.037	0.433	0.442	0.212
		≥ 12	0.218	-0.691	0.604	0.594	0.392
('0.730', '0.840')	12.141	7	0.091	-6.454	0.012	0.015	0.003
		8	0.068	-2.826	0.042	0.048	0.013
		9	0.068	-1.513	0.115	0.131	0.040
		10	0.134	-1.496	0.259	0.272	0.102
		11	0.200	-1.078	0.444	0.451	0.219
		≥ 12	0.214	-0.642	0.623	0.597	0.402

Table 8.5. For the fully-leptonic channel, top five choices for the ABCD bin edges, ranked by significance. Also, the choice of non-optimized ABCD bin edges (0.6, 0.6) is ranked by significance. For each ABCD bin edges choice, the values of the metric elements are also listed for each N_{jets} bin. Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV is applied while optimizing the edges

ABCD edges	**Significance**	N_{jets}	**Non-closure**	**Pull**	**Sig. Fracs B**	**Sig. Fracs C**	**Sig. Fracs D**
('0.600', '0.600')	7.420	6	0.120	-3.044	0.015	0.018	0.004
		7	0.068	-1.019	0.056	0.055	0.016
		8	0.021	0.165	0.139	0.151	0.046
		9	0.066	-0.274	0.268	0.308	0.114
		10	0.000	0.000	0.389	0.429	0.214
		≥ 11	0.000	0.000	0.556	0.500	0.400
('0.690', '0.570')	8.449	6	0.136	-3.136	0.017	0.019	0.004
		7	0.074	-1.013	0.062	0.058	0.018
		8	0.005	0.036	0.152	0.161	0.050
		9	0.060	0.215	0.286	0.325	0.117
		10	0.130	0.226	0.409	0.455	0.233
		≥ 11	0.333	0.333	0.636	0.600	0.400
('0.660', '0.540')	8.414	6	0.122	-3.072	0.015	0.017	0.004
		7	0.071	-1.067	0.058	0.052	0.017
		8	0.014	0.107	0.142	0.145	0.046
		9	0.036	-0.143	0.268	0.325	0.117
		10	0.114	-0.227	0.409	0.400	0.214
		≥ 11	0.333	0.333	0.600	0.600	0.333
('0.690', '0.610')	8.293	6	0.144	-3.175	0.017	0.021	0.004
		7	0.084	-1.088	0.066	0.060	0.019
		8	0.002	-0.011	0.158	0.174	0.053
		9	0.003	-0.011	0.302	0.349	0.127
		10	0.040	-0.069	0.400	0.455	0.242
		≥ 11	0.333	0.333	0.600	0.600	0.400
('0.700', '0.590')	8.276	6	0.146	-3.242	0.017	0.020	0.004
		7	0.084	-1.103	0.065	0.061	0.018
		8	0.029	-0.196	0.158	0.169	0.051
		9	0.008	0.029	0.288	0.350	0.122
		10	0.000	0.000	0.429	0.455	0.226
		≥ 11	0.333	0.333	0.600	0.600	0.400

Based on the tables, the "optimized" ABCD bin edges are chosen as (0.85, 0.74), (0.70, 0.85) and (0.69, 0.57) for the fully-hadronic, semi-leptonic and fully-leptonic channels, respectively. The significance as a function of the "optimized" ABCD bin edges is given for each channel in Figure 8.26.. Acording to the signal fractions, the best significance is observed in the highest N_{jets} bin for each channel where the Disc. 1 and Disc. 2 edges are crossed. The significance for all N_{jets} bins (for each channel) are given in Appendix D1.2..

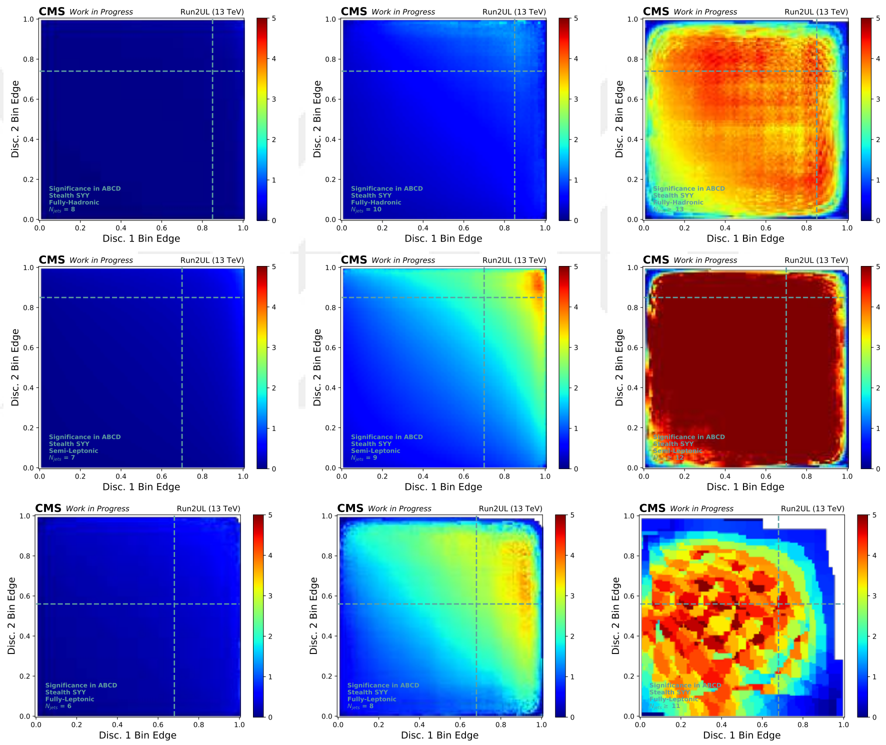


Figure 8.26. For the fully-hadronic (top), semi-leptonic (middle), fully-leptonic (bottom) channels, significance as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,$\geq$13 bins (left to right). N_{jets} = 7,9,$\geq$12 bins (left to right). N_{jets} = 6,8,$\geq$11 bins (left to right). According to signal fraction, the best significance is observed in the highest N_{jets} bin for each channel where the Disc. 1 and Disc. 2 edges are crossed

Figure 8.27. shows the non-closure as a function of "optimized" ABCD bin edges for each channel. As expected for lower N_{jets} bins, non-closure is less then $\sim 5\%$, where the Disc. 1 and Disc. 2 edges are crossed for a given channel For all N_{jets} bins, the non-closure are given in Appendix D1.2..

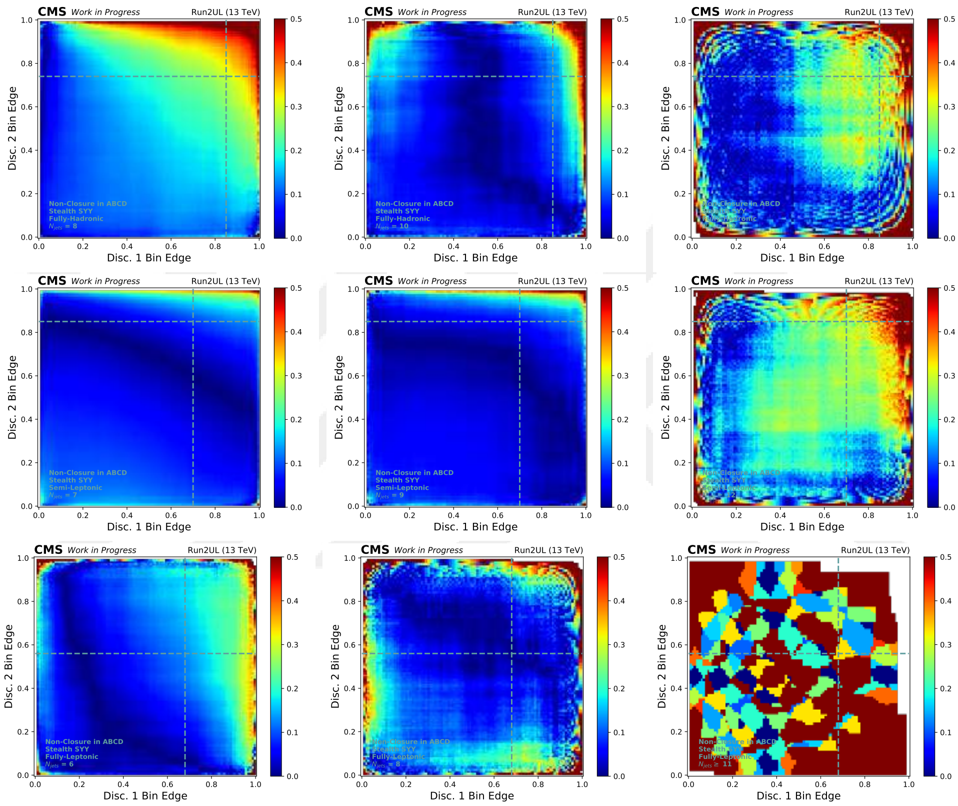


Figure 8.27. For the fully-hadronic (top), semi-leptonic (middle), fully-leptonic (bottom) channels, non-closure as a function of "optimized" ABCD bin edges choice. $N_{\text{jets}} = 8,10,\geq 13$ bins (left to right). $N_{\text{jets}} = 7,9,\geq 12$ bins (left to right). $N_{\text{jets}} = 6,8,\geq 11$ bins (left to right). As expected for lower N_{jets} bins, non-closure is less then $\sim 5\%$, where the Disc. 1 and Disc. 2 edges are crossed

In general, the pull is about zero as expected where the Disc. 1 and Disc. 2 edges are crossed for a given channel as shown in Figure 8.28.. In Appendix D1.2., more plots are shown for each channel.

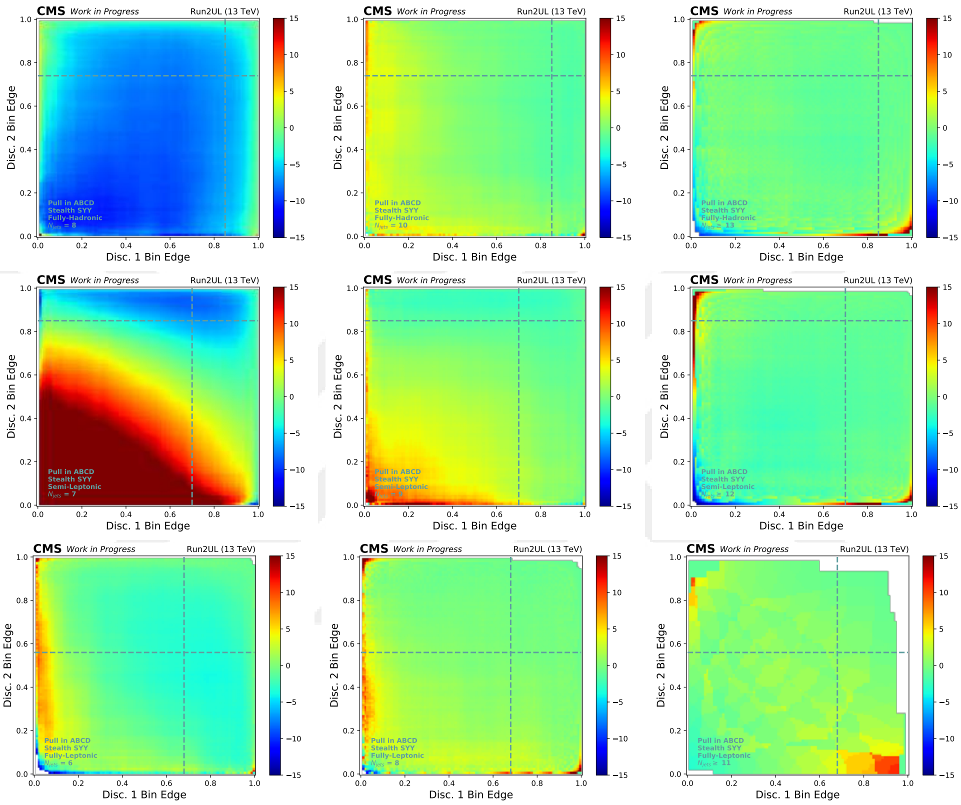


Figure 8.28. For the fully-hadronic (top), semi-leptonic (middle), fully-leptonic (bottom) channels, pull as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,≥13 bins (left to right). N_{jets} = 7,9,≥12 bins (left to right). N_{jets} = 6,8,≥11 bins (left to right). In general, the pull is about zero as expected where the Disc. 1 and Disc. 2 edges are crossed for a given channel

Also, signal fractions in **B**, **C**, and **D** regions are given in Figure 8.29. for the fully-hadronic channel, Figure 8.30. for the semi-leptonic channel and Figure 8.31. for the fully-leptonic channel. For a given channel, the signal fractions for **B**, **C**, and **D** regions are about zero for lower N_{jets} bins whereas they are greater than zero for higher N_{jets} bins. The signal fractions in **B**, **C**, and **D** regions for all N_{jets} bins (for each channel) are given in Appendix D1.2..

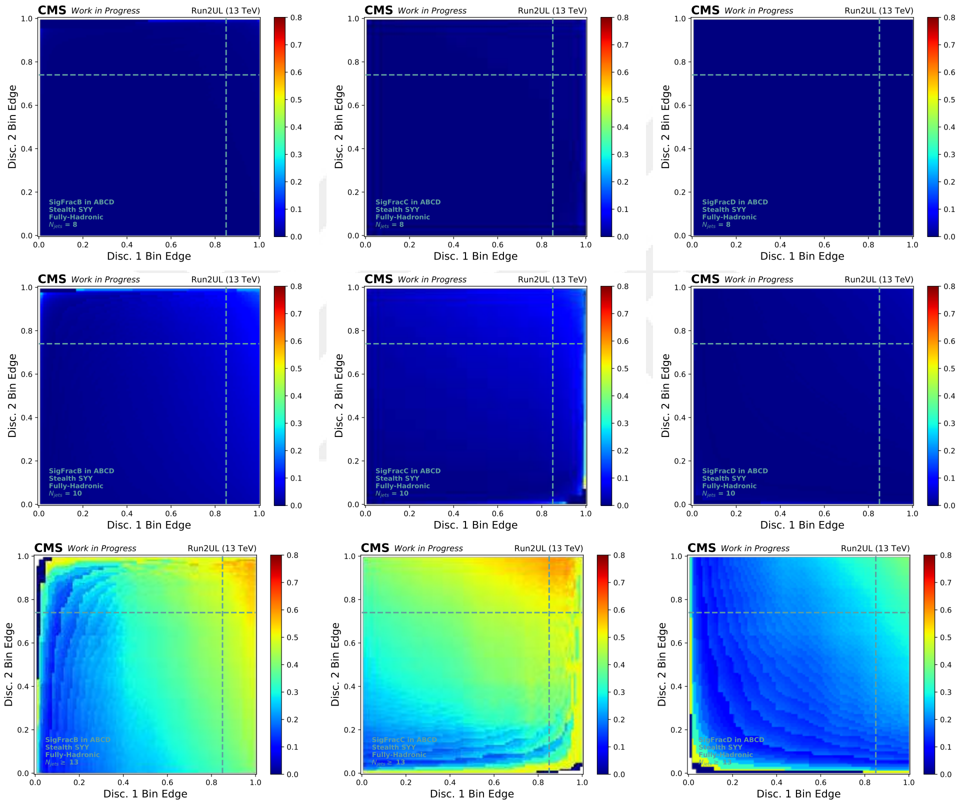


Figure 8.29. For the fully-hadronic channel, signal fractions in each B, C, D region (left to right) as a function of "optimized" ABCD bin edges choice. N_{jets} = 8,10,≥13 bins (top to bottom). The signal fractions for **B**, **C**, and **D** regions are about zero for lower N_{jets} bins for a given channel

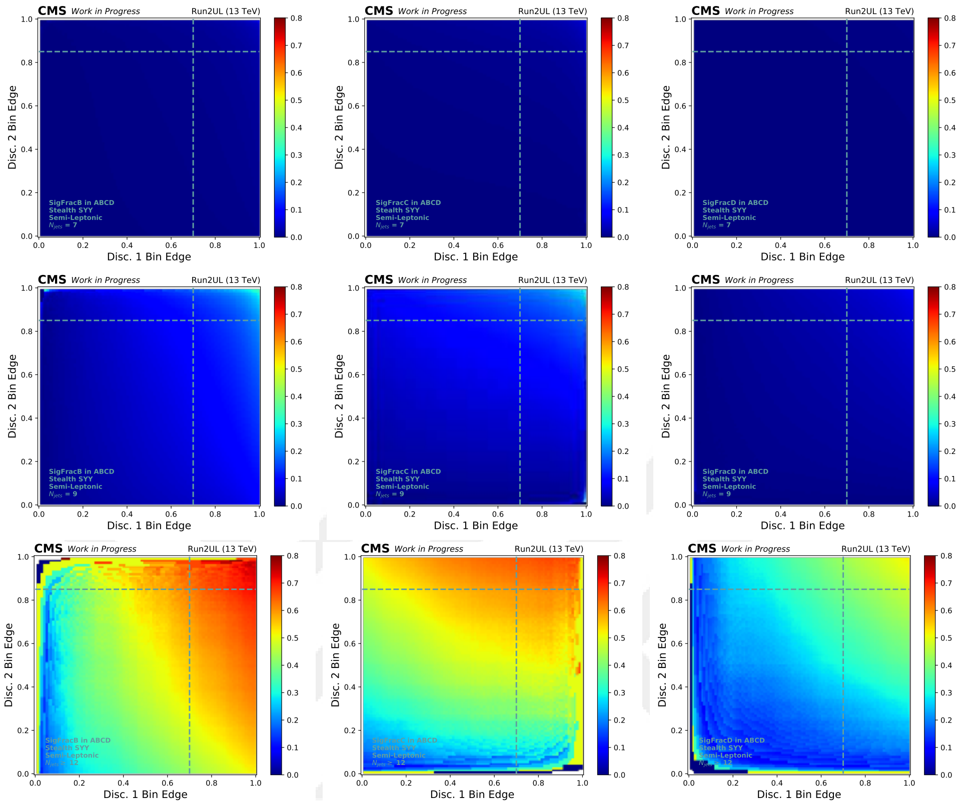


Figure 8.30. For the semi-leptonic channel, signal fractions in each B, C, D region (left to right) as a function of "optimized" ABCD bin edges choice. N_{jets} = 7,9,$\geq$12 bins (top to bottom)

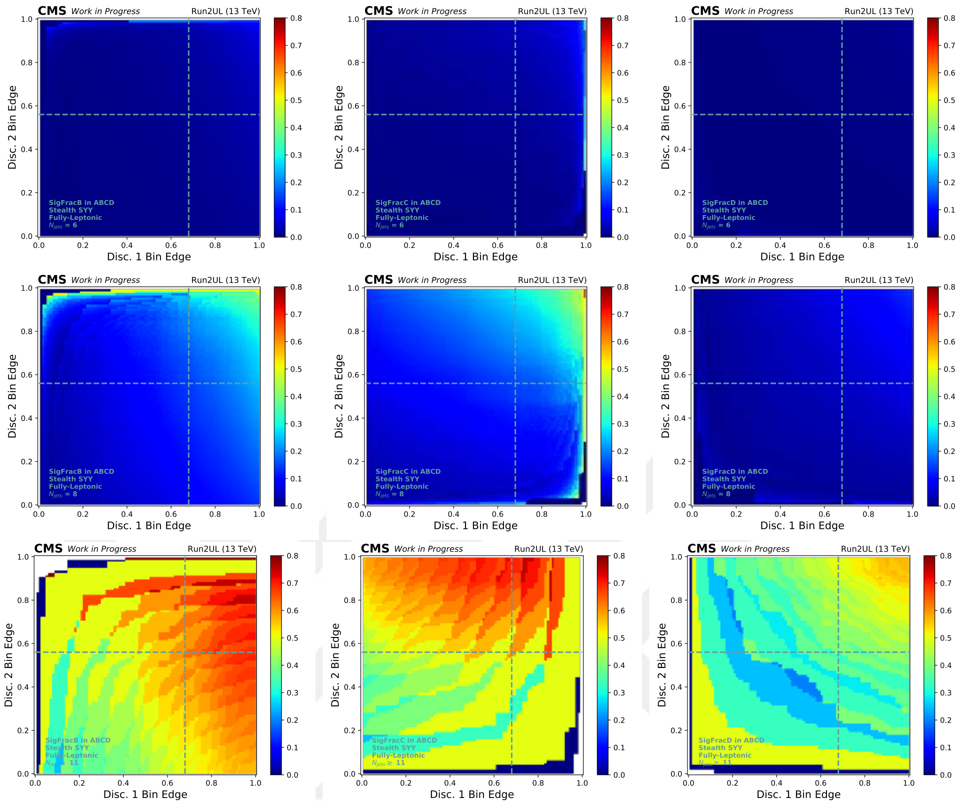


Figure 8.31. For the fully-leptonic channel, signal fractions in each B, C, D region (left to right) as a function of "optimized" ABCD bin edges choice. N_{jets} = 6,8,≥11 bins (top to bottom)

8.4.2. Defining the Validation Regions

The importance of the validation regions is to check the ABCD method outside of the main signal region **A**. This tests the robustness of the ABCD method, i.e. that is able to close in arbitrary subregions of the 2D plane. Also, these regions can be checked in data to verify the modeling of $t\bar{t}+$jets in data. Because, the signal contamination is low enough.

To get the validation regions, the bin edges for the ABCD regions are first determined. Three validation regions are then created. The validation regions are defined based on the sub-division of each ”BD” regions, ”CD” regions and ”D” region as shown in Figure 8.32..

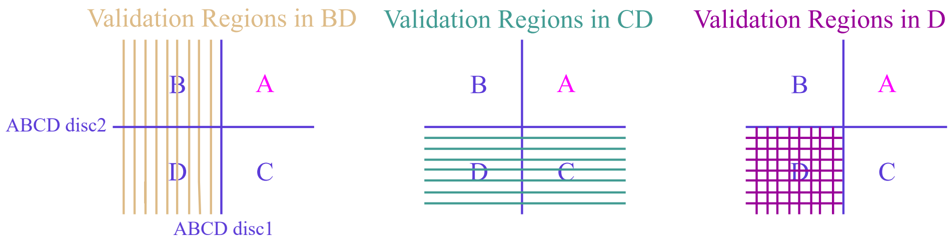


Figure 8.32. Validation regions are defined based on the sub-division of each ”BD” regions, ”CD” regions and ”D” region

Validation Region I : Sub-Division of BD

The idea is to keep the **BD** region based on ABCD edges and getting the best choice of the four sub regions to validate the ABCD method. This region helps to understand how well the ABCD method works while taking into account the full Disc. 2 range. So, the control regions **B** and **D** are subdivided into four sub regions as validation region I. To get these sub regions, values are scanned from 0 to ABCD Disc. 1 edge to get the Disc. 1 edge of the validation region I and then the ABCD Disc. 2 edge is chosen as Disc. 2 edge of the validation region I.

Validation Region II : Sub-Division of CD

The idea is to keep the **CD** region based on ABCD edges and getting the best choice of the four sub regions to validate the ABCD method. This region helps to understand how well the ABCD method works while taking into account the full Disc. 1 range. The control regions **C** and **D** are subdivided into four sub regions as validation region II. To get these sub regions, the ABCD Disc. 1 edge is chosen as the Disc. 1 edge of the validation region II and values are scanned from 0 to ABCD Disc. 2 edge to get the Disc. 2 edge of the validation region II.

Validation Region III : Sub-Division of D

In the control region **D**, on bottom left corner the least signal contamination is expected based on the output of the Double DisCo. The region explores both discriminants equally as well. The control region **D** is subdivided into four sub regions as validation region III. To get these sub regions, the edges are chosen as [(ABCD Disc. 1) / 2] and [(ABCD Disc. 2) / 2].

The exact definitions of the validation regions (as detailed in the subsequent paragraphs) was motivated by initial studies of signal contamination. For the validation regions to be useful at all, they need to be sufficiently devoid of signal in order for the regions to be investigated in data. Thus, when choosing the values of the edges for each validation region, a signal contamination value of 5% is imposed as a criteria, as computed in the equivalent "A" region of each validation region. With respect to the ABCD bin edges, two types of boundaries are defined for the validation regions. The "right-most" and "top-most" boundaries for each respective validation region are shown in Figure 8.33.. The final choices for the boundaries result in a "gap" between the respective validation region boundaries and the ABCD bin edges, since accommodating the events in the gap, would lead to too high signal contamination.

For Validation Region I, the right-most boundary is at 0.4 and this region is framed with 0.4 and ABCD Disc. 2 edge. Then, the edges are chosen as $\text{Disc. }1, \text{Disc. }2 = [0.2, \text{ABCD} \quad \text{Disc. }2]$ to get the sub-division of the validation region I. For Validation Region II, the top-most boundary is at 0.4 and this region is framed with 0.4 and ABCD Disc. 1 edge. Then, the edges are chosen as $\text{Disc. }1, \text{Disc. }2 = [\text{ABCD} \quad \text{Disc. }1, 0.2]$ to get the sub-division of the validation region II. For Validation Region III, the right-most boundary is at 0.6 and top-most boundary is at 0.6. This region is framed with 0.6 and 0.6. Then, the edges are chosen as Disc. 1, Disc. 2 $= [0.6/2, 0.6/2]$ to get the sub-division of validation region III. Subdivisions of each validation region I, II, III are shown in Figure 8.33..

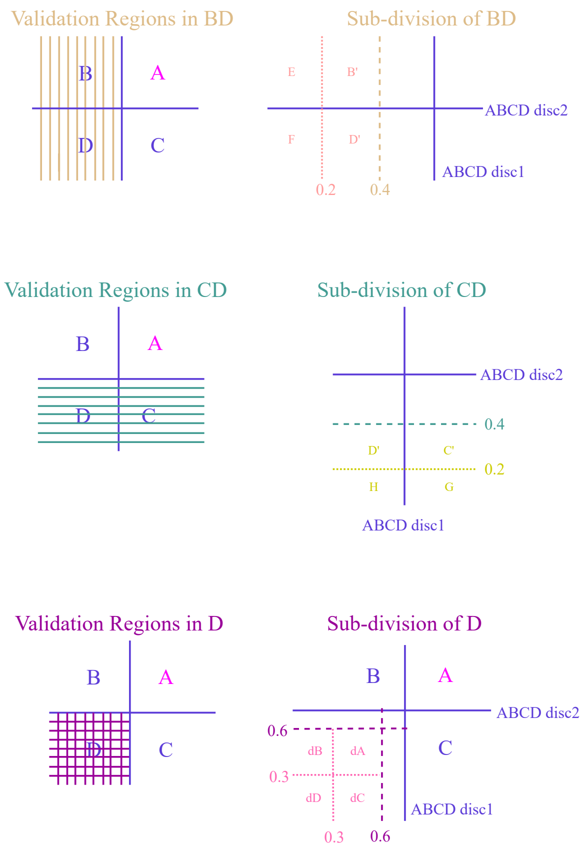


Figure 8.33. Validation region I with ABCD bin edges and right-most boundary as sub-division of BD regions (top). Validation region II with ABCD bin edges and top-most boundary as sub-division of CD regions (middle). Validation region III with ABCD bin edges and right-most and top-most boundaries as sub-division of D region (bottom)

8.4.3. Verifying the Validation Regions

The non-closure for validation regions are given in Figure 8.34. for the fully-hadronic channel, in Figure 8.35. for the semi-leptonic channel and in Figure 8.36. for the fully-leptonic channel. In any of the three validation regions, the non-closure as a function of ABCD bin edges demonstrates that the method continues working well outside of the main ABCD region. This helps to build confidence about the robustness of the method. Good closure of the ABCD method is seen for all channels. For the higher N_{jets} bins, it is expected to see the higher value of the non-closure. The non-closure for other N_{jets} bins for each channel are also given in Appendix D2., Figure D19., Figure D20., Figure D21..

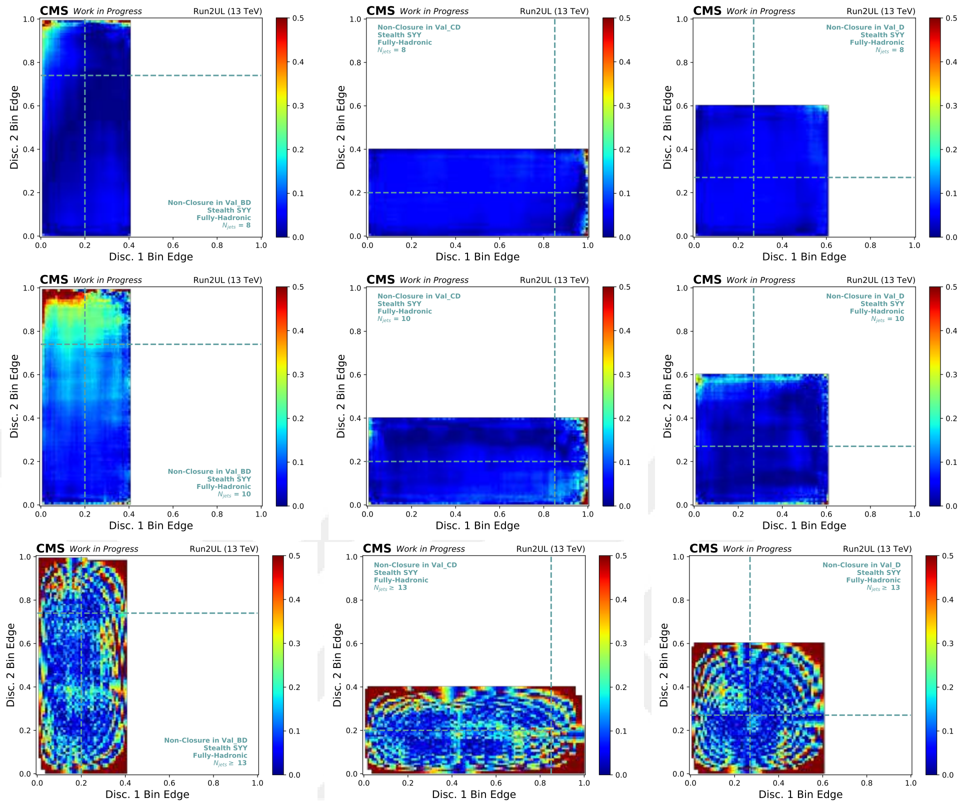


Figure 8.34. For the fully-hadronic channel, non-closure as a function of validation I edges, validation II edges, and validation III edges (left to right). Top to bottom N_{jets} = 8,10,$\geq$ 13 bins. The large values of the non-closure are observed for the higher N_{jets} bins due to the statistical uncertainty on the closure

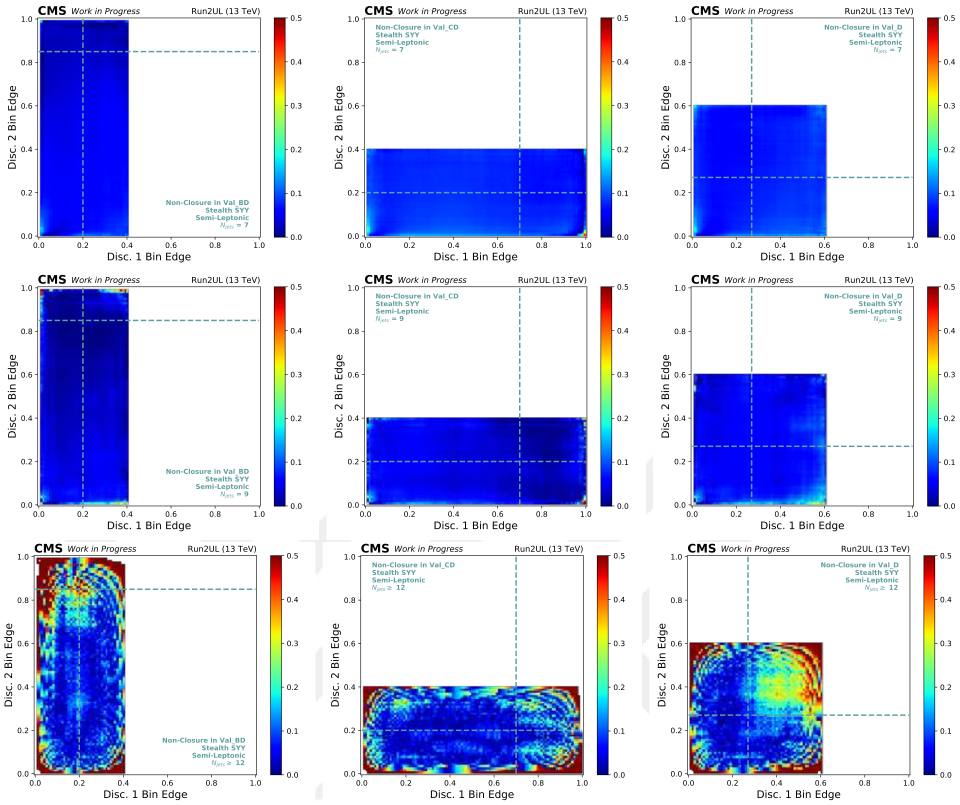


Figure 8.35. For the semi-leptonic channel, non-closures as a function of validation I edges, validation II edges, and validation III edges (left to right). Top to bottom N_{jets} = 7,9,≥ 12 bins. The large values of the non-closure are observed for the higher N_{jets} bins due to the statistical uncertainty on the closure

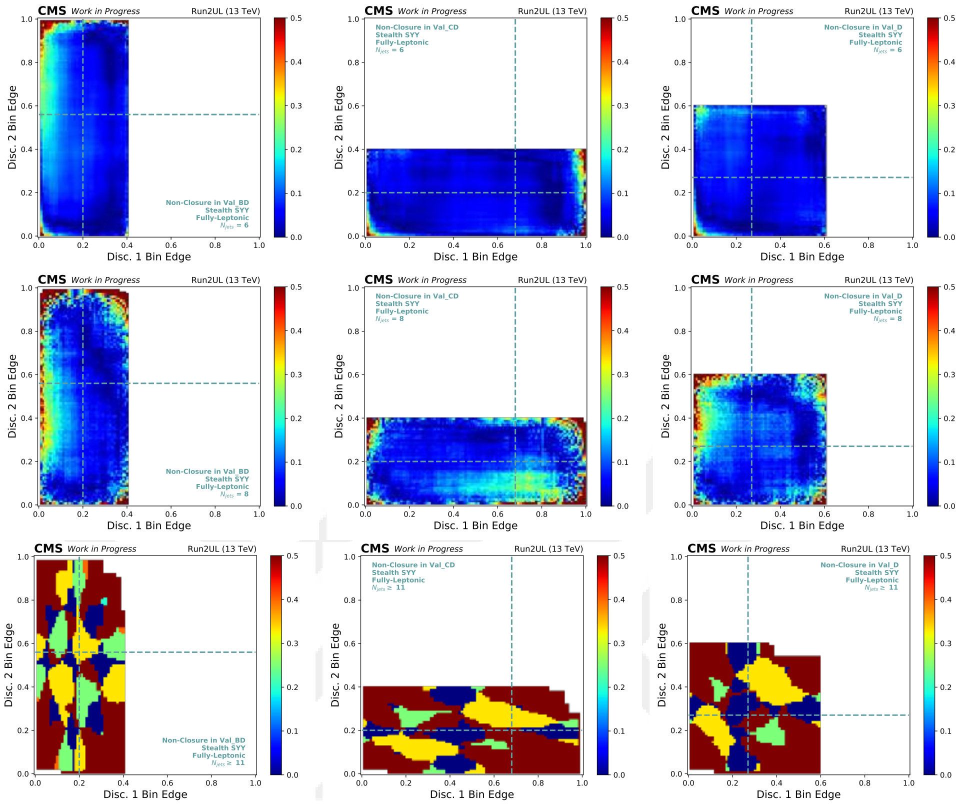


Figure 8.36. For the fully-leptonic channel, non-closures as a function of validation I edges, validation II edges, and validation III edges (left to right). Top to bottom $N_{\text{jets}} = 6, 8, \geq 11$ bins. The large values of the non-closure are observed for the higher N_{jets} bins due to the statistical uncertainty on the closure

The non-closure is calculated per N_{jets} by using $t\bar{t}$ + jets simulation to investigate how robust the ABCD method is in simulation. At the same time, to clarify that the Double DisCo NN behaviour is good. The results are given in Figure 8.37. (for the fully-hadronic channel), Figure 8.38. (for the semi-leptonic channel), and Figure 8.39. (for the fully-leptonic) and show that the ABCD method is robust and the validation regions improve that the ABCD method also works well outside of the signal region. This gives us confidence to look at data.

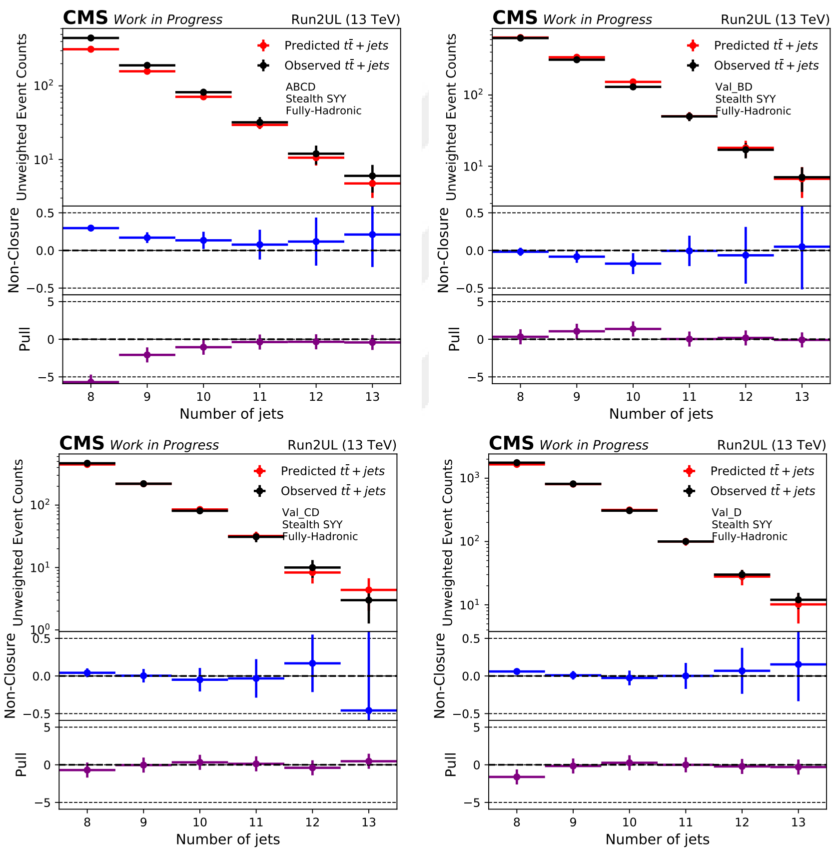


Figure 8.37. For the fully-hadronic channel, the closure per-N_{jets} for $t\bar{t}$ + jets in the ABCD region (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right)

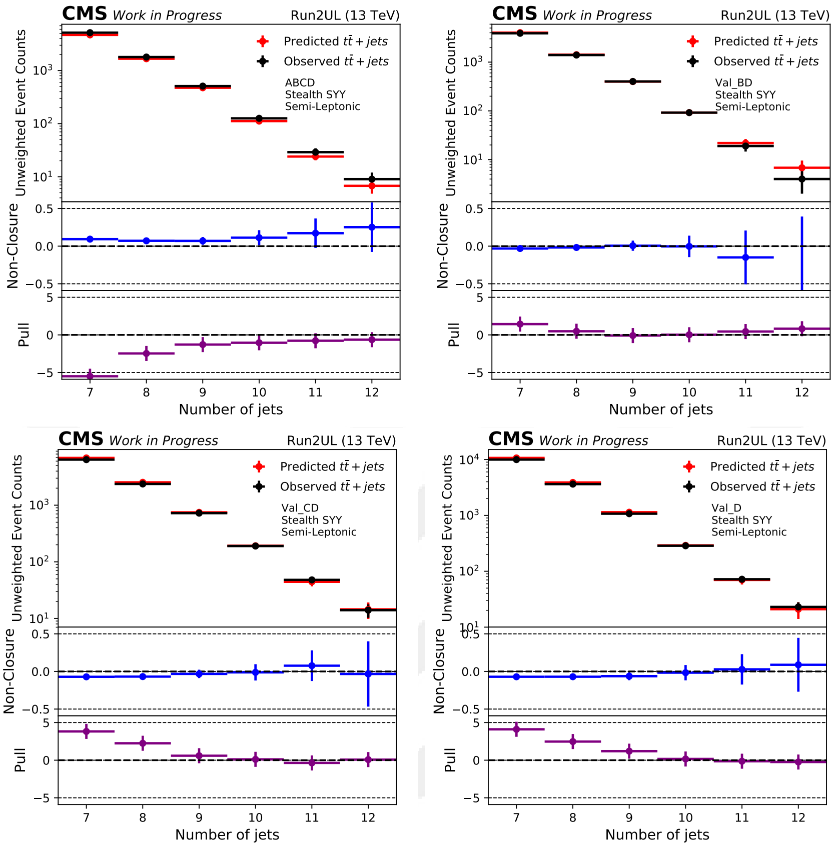


Figure 8.38. For the semi-leptonic channel, the closure per-N_{jets} for $t\bar{t}$ + jets in the ABCD region (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right) In ABCD, Validation CD and Validation D regions, the non-closure values are still reasonable even if pull values are far from 0

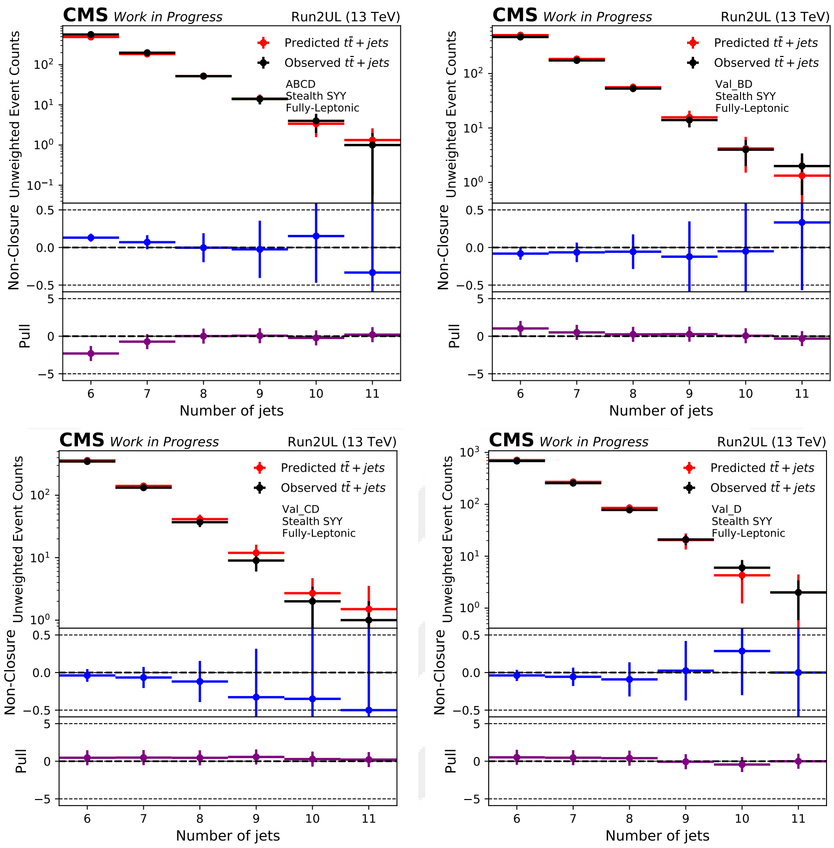


Figure 8.39. For the fully-leptonic channel, the closure per-N_{jets} for $t\bar{t}+$jets in the ABCD region (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right)

8.5. $t\bar{t}+$jets Closure Correction

The $t\bar{t}+$jets background prediction is performed by using the ABCD method. There are two fundamental assumptions that must be upheld to ensure that this method accurately estimates the $t\bar{t}+$jets background. One assumption is that double discriminant distributions for $t\bar{t}+$jets allow for a well-closing ABCD method. Another assumption is that the closure of the ABCD method with $t\bar{t}+$jets simulation is representative of that with data. If these two assumptions are satisfied, then any non-closure of the method in data can be corrected via a correction factor derived in simulation.

According to the ideas above, a simulation-based **closure corrrection** is calculated from $t\bar{t}+$jets simulation, as given in Equation 8.15. The closure corrections as function of several boundary values of the validation regions are given for fully-hadronic, semi-leptonic and fully-leptonic channels in Figure 8.40., Figure 8.41. and Figure 8.42., respectively.

$$\text{Closure Correction} \equiv \frac{1}{\text{ABCD Closure}} = \left.\frac{N_{\text{obs.},A}}{N_{\text{pred.},A}}\right|_{\text{MC}} \tag{8.15}$$

Figure 8.40., Figure 8.41. and Figure 8.42. show the nominal closure correction for several boundary values, which start from right-most(top-most) boundary and go through boundary value “1.0” according to the each validation region. The left most point in each plot corresponds to the right-most(top-most) boundary for validation I(II). For the validation region III, the left most point in each plot corresponds to where right-most and left-most boundaries are crossed. At that right-most point, the three regions are identically the signal region, and so the closure corrections are identical. At the end, the closure correction will be applied to the final $t\bar{t}+$jets background prediction in the signal region.

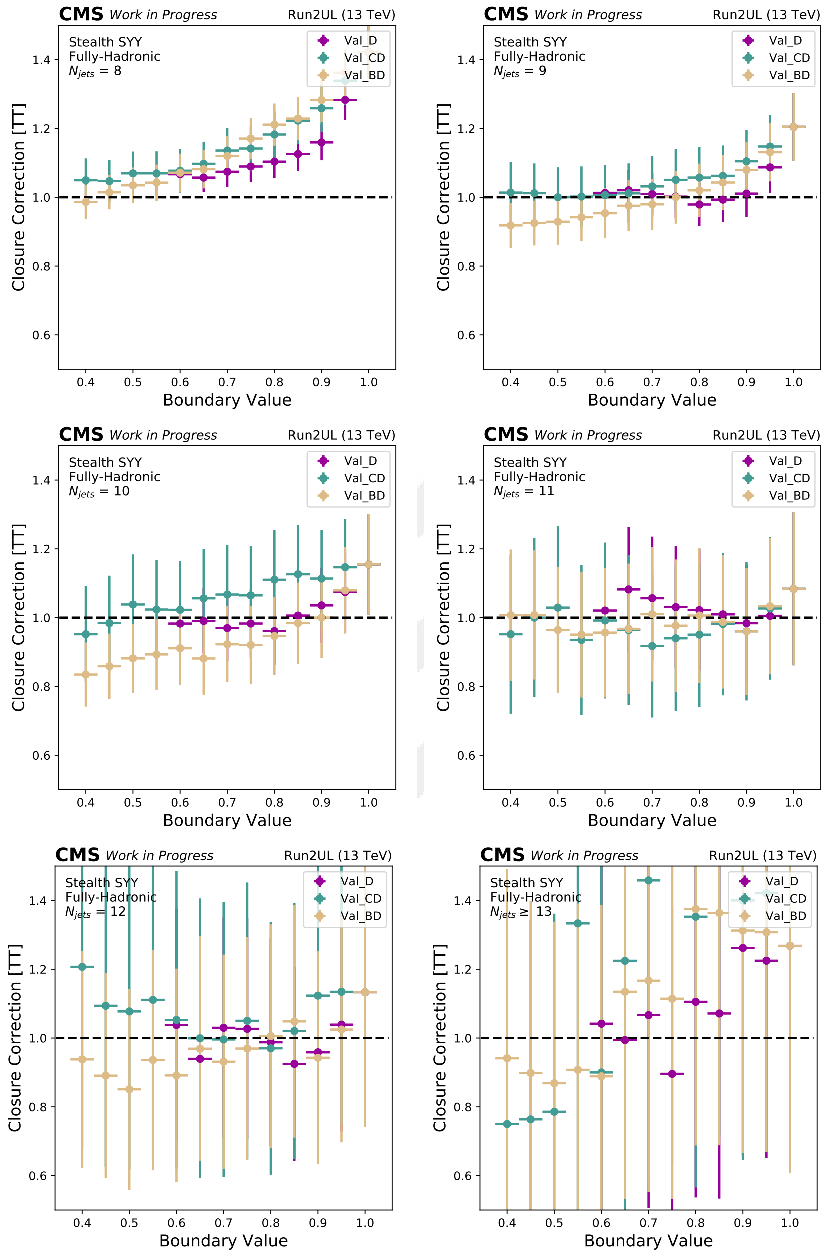


Figure 8.40. For the fully-hadronic channel, the closure corrections as a function of boundary value of the validation regions. The closure corrections for $N_{\text{jets}} = 8, 9, 10, 11, 12, \geq 13$ (left to right, top to bottom) and calculated from $t\bar{t}$ + jets simulation

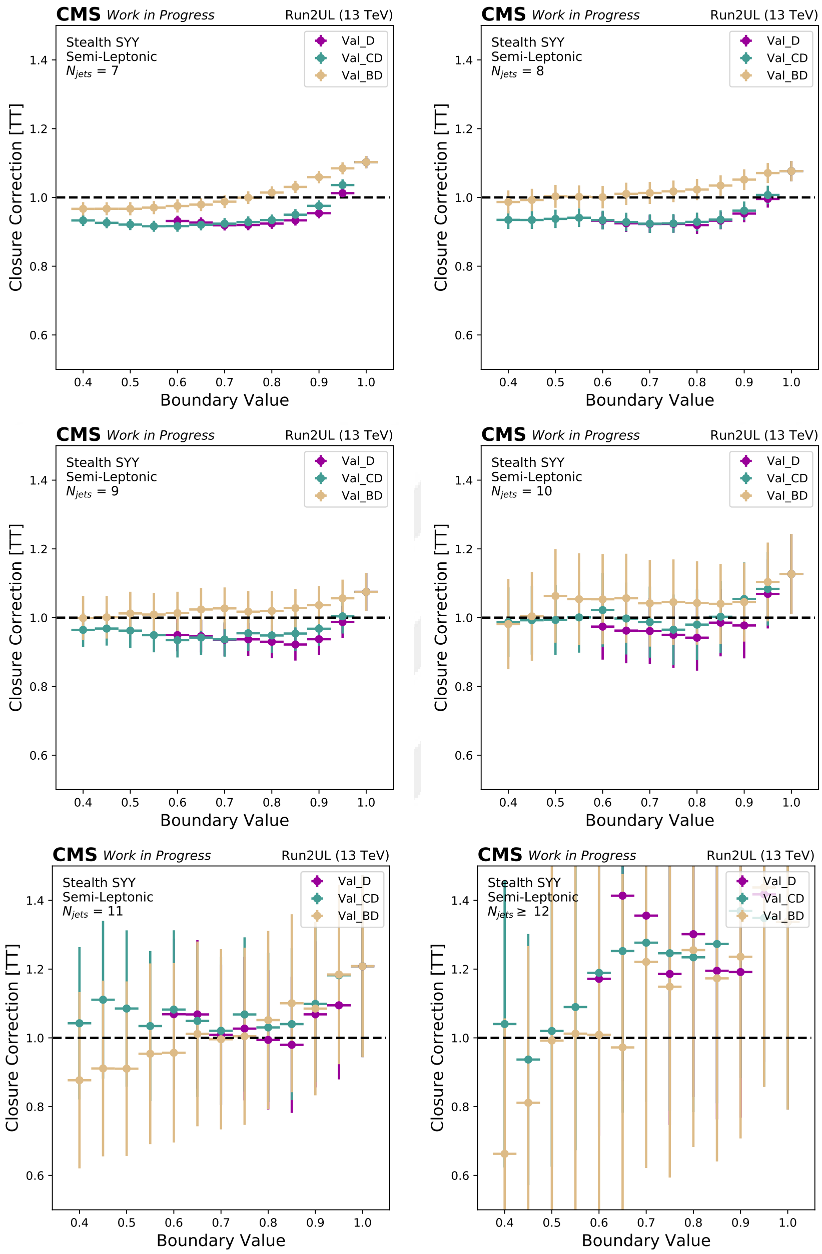


Figure 8.41. For the semi-leptonic channel, the closure corrections as a function of boundary value of the validation regions. The closure corrections for N_{jets} = 7, 8, 9, 10, 11, $\geq$ 12 (left to right, top to bottom) and calculated from $t\bar{t}+$jets simulation

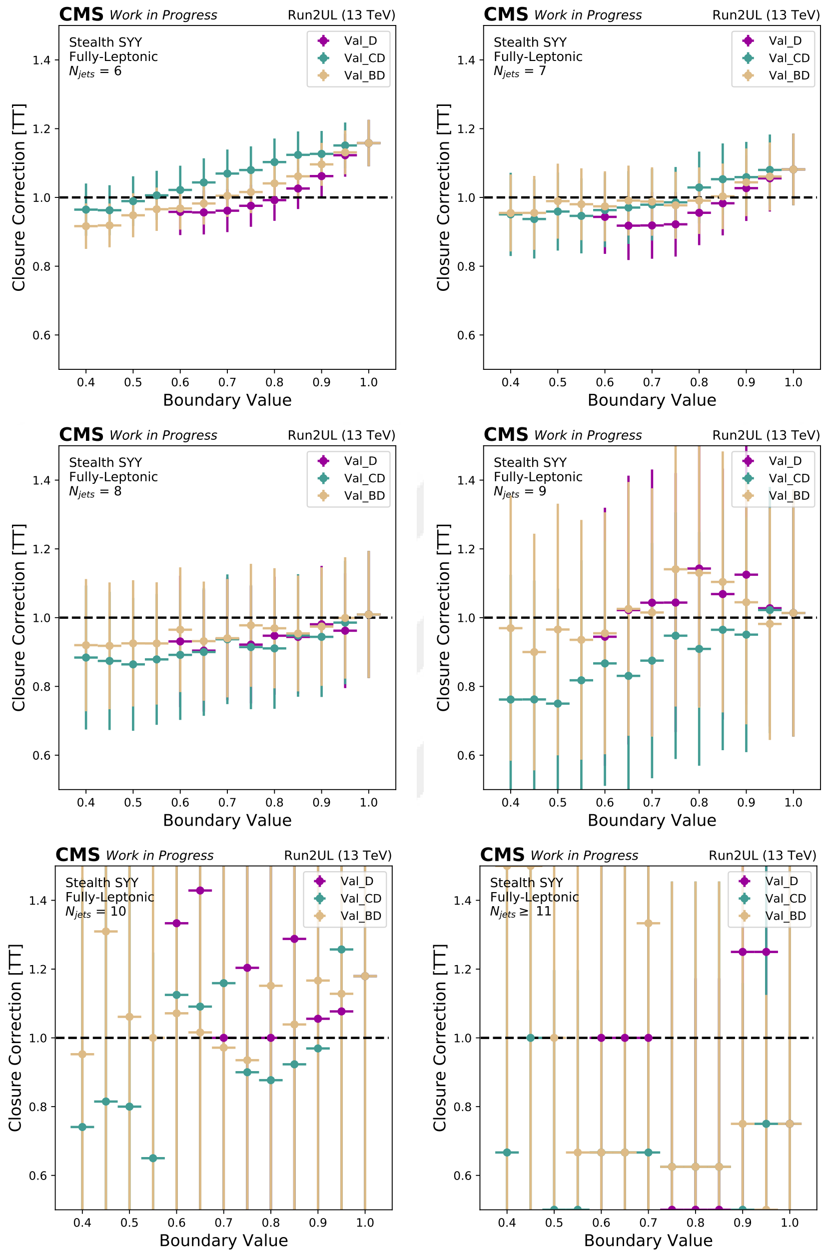


Figure 8.42. For the fully-leptonic channel, the closure corrections as a function of boundary value of the validation regions. The closure corrections for N_{jets} = 6, 7, 8, 9, 10, ≥ 11 (left to right, top to bottom) and calculated from $t\bar{t}$ + jets simulation

8.6. $t\bar{t}+$jets Systematics

Based on the estimation method for the $t\bar{t}+$jets background and the associated closure correction, there are two types of systematics: simulation-based, which is derived using simulation and data-based derived when analyzing data in the validation regions. The simulation-based systematic is related to the closure correction, as detailed in Section 8.6.1.. The data-based systematic is related to the ABCD method itself, as detailed in Section 8.6.2..

8.6.1. Simulation-Based Systematics

As the closure correction is derived in simulation, there is inherent "uncertainty" in the calculated value related to uncertainty in the simulation and how the network treats differently simulated events. This uncertainty is directly linked to the uncertainty in the physics modeling (variations in POWHEG and PYTHIA parameters) and detector modeling (jet energy correction (JEC) and jet energy resolution (JER) variations) of the $t\bar{t}+$jets simulation.

A unique version of the closure correction, **closure correction ratio**, is derived using each variation of the $t\bar{t}+$jets background in the signal region, where the right-most(top-most) boundary values are 1.0 which is the full ABCD region. The closure correction ratio is given in Equation 8.16. The systematic uncertainty for each variation is computed by taking the ratio between the closure correction using the nominal $t\bar{t}+$jets and the closure correction computed with a variation of the $t\bar{t}+$jets events.

$$\text{Closure Correction Ratio} = \left.\frac{\text{variation of } t\bar{t}+\text{jets}}{t\bar{t}+\text{jets}}\right|_{\text{Closure Corr.}} \tag{8.16}$$

Likewise, there is a natural statistical uncertainty of the closure correction based on the finite number of events in each of the **A**, **B**, **C**, and **D** regions as can be inferred from Equation 8.15. As a comparison, these simulation-based systematics for each $t\bar{t}+$jets variation are compared to the propagated uncertainty on the nominal closure correction itself.

In Figure 8.43., the left plots (for given channels) show the closure correction ratios (simulation-based systematics) derived when using different $t\bar{t}+$jets events based on modeling or detector variations. The gray band is the pure statistical uncertainty on the nominal closure correction divided by the nominal closure correction itself to allow direct comparison with the $t\bar{t}+$jets variation systematics. The right plots (for given channels) show a single example of closure correction for the modeling variation that affects “color reconnection” and the top panel compares it with the nominal closure correction derived per N_{jets}, while the bottom panel then shows the closure correction ratio between them. This ratio then the literal systematic uncertainty on the closure correction for that particular variation.

Given how the individual closure correction ratio for each variation are encompassed within the gray band, it is understood that the pure statistical uncertainty on the nominal closure correction itself is the dominant one. Thus, when incorporated into the final fit setup, the statistical uncertainty on the nominal closure correction is used as the sole simulation-based systematic.

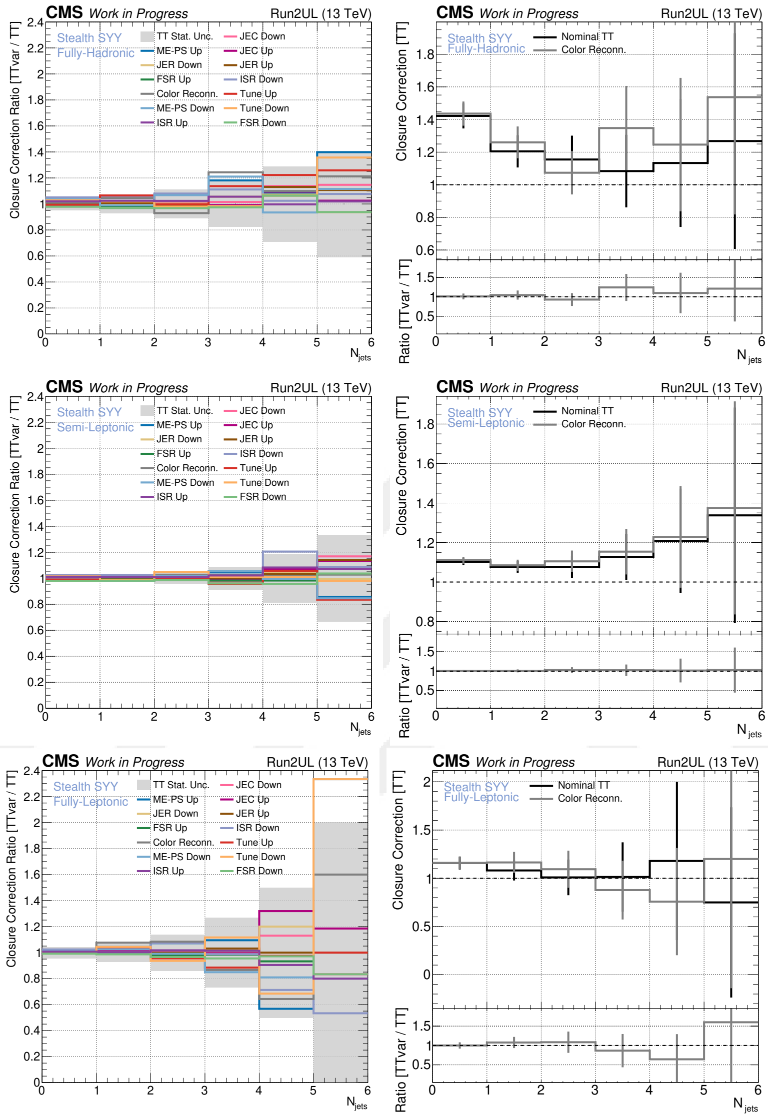


Figure 8.43. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the closure correction ratios for all variations (left) and the gray band given on this plot is the pure statistical uncertainty on the nominal closure correction. A single example of closure correction for the modeling variation that affects "color reconnection" (right) and the top panel compares it with the nominal closure correction derived per N_{jets}, while the bottom panel then shows the closure correction ratio between them. Note that there is an offset to be applied for numbering the N_{jets} bins on the x-axis of all plots

8.6.2. Data-Based Systematics

The data-based systematic is related to the ABCD method itself and they are derived by using data. Even after applying the closure correction (derived in simulation) to data, the ABCD method still does not close perfectly when used in data. This residual non-closure or difference from a perfect closure of 1.0 is taken as a separate systematic, which is data-based.

To get the data-based systematic, data closure is first calculated by applying the closure correction to the predicted data events in **A** region. Then, the corrected data closure is given in Equation 8.18. Data closure and corrected data closure values are only calculated for validation boundary values where the estimated signal contamination in each of the validation region's effective **A** region is less than 5%.

$$\text{Data Closure} = \left.\frac{N_{\text{pred.},A}}{N_{\text{obs.},A}}\right|_{\text{Data}} \tag{8.17}$$

$$\text{Corrected Data Closure} = \text{Closure Correction} * \left.\frac{N_{\text{pred.},A}}{N_{\text{obs.},A}}\right|_{\text{Data}} \tag{8.18}$$

Figure 8.44. shows how well the ABCD method closes in the data[1] after applying the closure correction derived in simulation. If the closure correction works well, then the corrected data closure has smooth distribution around this ideal value of 1.0. Even if some non-closure remains after the correction, it is still seen that the closure correction ameliorates non-closure seen in data while looking in the different validation regions.

[1] When calculating closure of the ABCD method in data, the predicted $t\bar{t}+$jets in data is calculated by subtracting non-$t\bar{t}+$jets simulation from data. Then, the predicted data in region **A** is defined as [(predicted $t\bar{t}+$jets in the data) + (observed non-$t\bar{t}+$jets)].

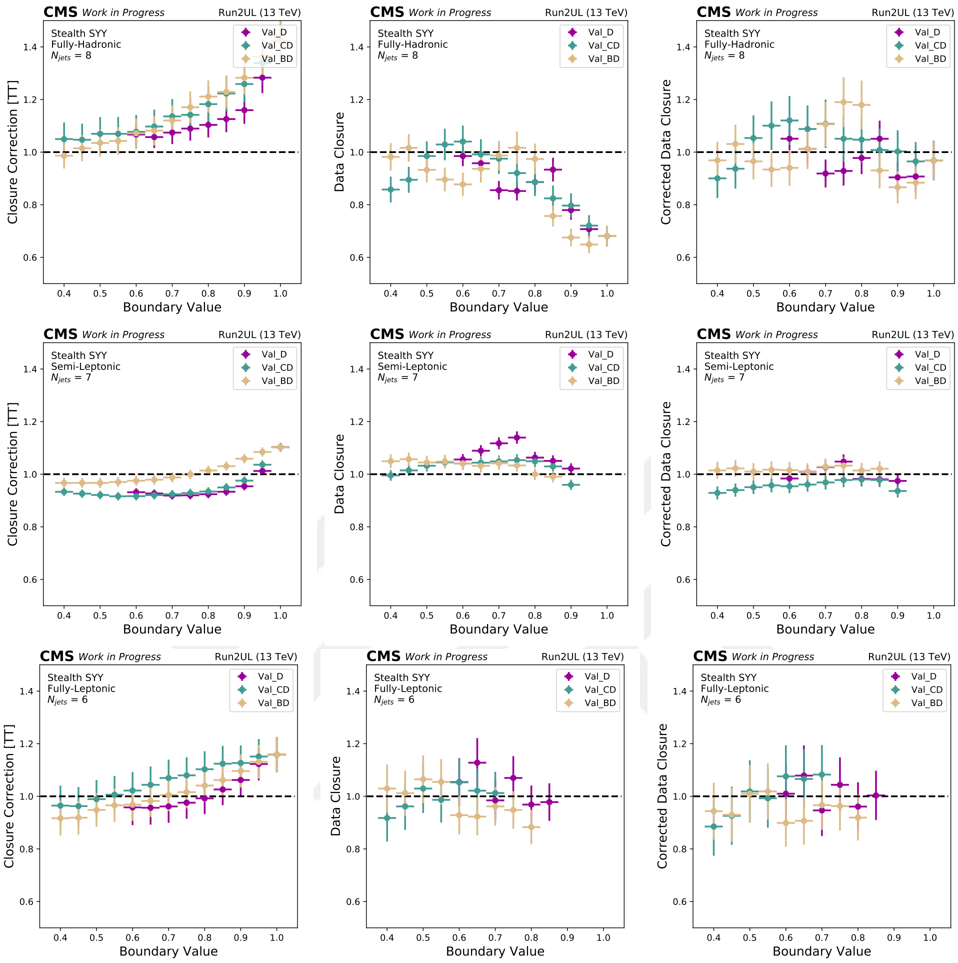


Figure 8.44. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels; the closures corrections (left), data closures (middle) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 8$ (for fully-hadronic channel), $N_{\text{jets}} = 7$ (for semi-leptonic channel), $N_{\text{jets}} = 6$ (for fully-leptonic channel)

The closure corrections per-N_{jets} for ABCD and all three validation regions are shown Figure 8.45., Figure 8.46., Figure 8.47. for the fully-hadronic, semi-leptonic and fully-leptonic channels, respectively. For the two lowest N_{jets} bins for each channel (this is because signal is sufficiently small there), data closure and corrected data closure are also shown. If the closure correction behaves well, then the corrected data closure in the ABCD region may be $\sim$1.0. It is seen that the closure correction works and lessens the non-closure in data—the remaining non-closure deviation from 1.0 is not more than 20% amongst all regions for a given channel.

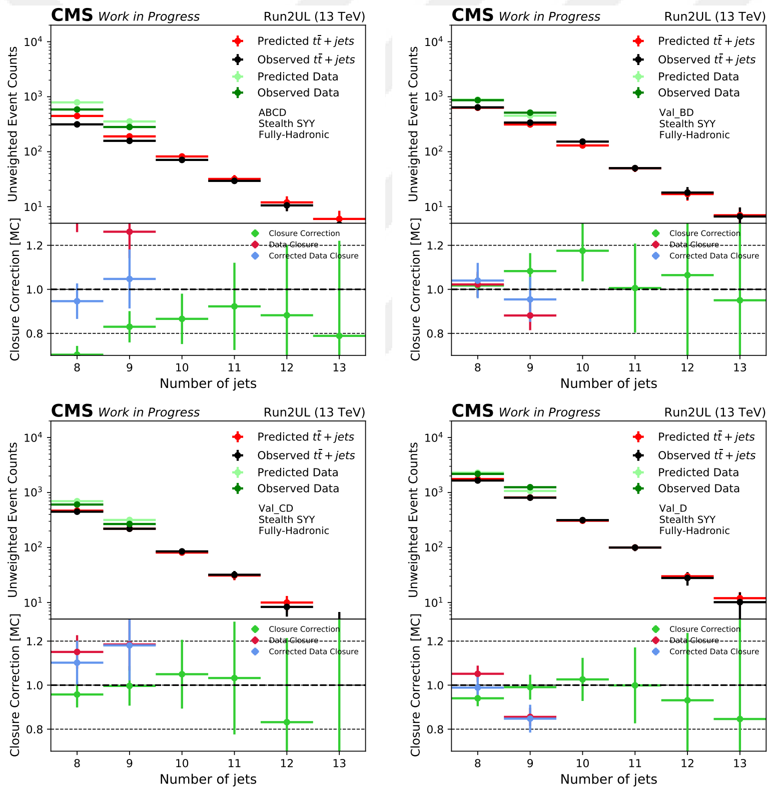


Figure 8.45. For the fully-hadronic channel, the closure correction per-N_{jets} for ABCD (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right)

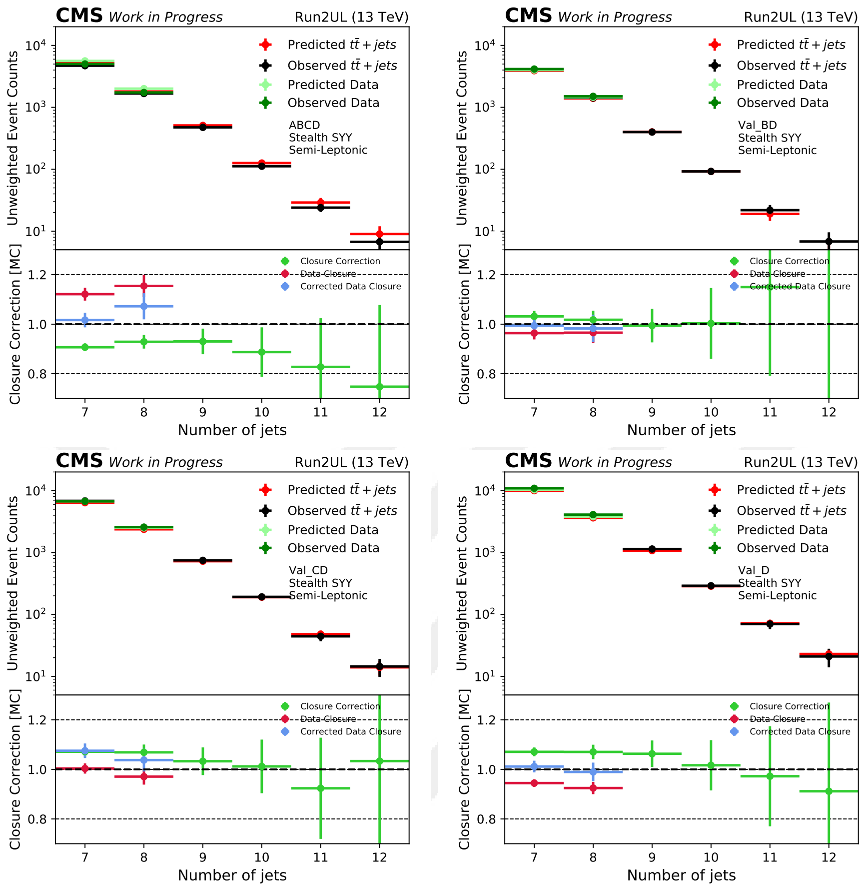


Figure 8.46. For the semi-leptonic channel, the closure correction per-N_{jets} for ABCD (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right)

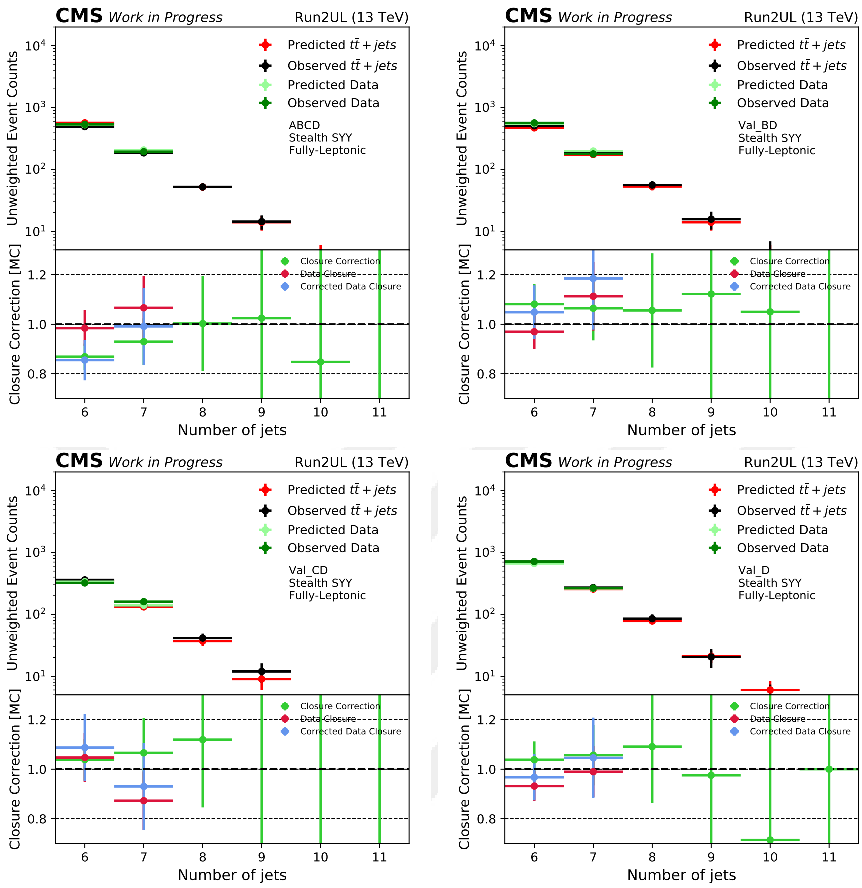


Figure 8.47. For the fully-leptonic channel, the closure correction per-N_{jets} for ABCD (top left), validation region I (top right), validation region II (bottom left) and validation region III (bottom right)

Then, the “maximum” value of the corrected data closure for a given N_{jets} bin, which is the furtest from the ideal value of 1.0 across all validation regions and all boundary values, is identified. This identification is performed for only the two lowest N_{jets} bins (for each channel) where signal contamination is below 5% for a majority of the validation region boundary choices. The data-based systematic for each of these two lowest N_{jets} bins is calculated as “[1 / max corrected data closure]”.

The maximum value of the corrected data closures and data-based systematics are given in Table 8.6. for the fully-hadronic channel, Table 8.7. for the semi-leptonic channel and Table 8.8. for the fully-leptonic channel. Note that the data-based systematics for the lowest two N_{jets} bins are whatever they are calculated. But, based on the satisfying signal contemination (5%) in each validation region, the data-based systematic is not calculated for the higher N_{jets} bins. So, the data-based systematic, which has the larger value of the "maximum" corrected data closure in between the lowest two N_{jets} bins, is assigned to all higher N_{jets} bins.

Table 8.6. For the fully-hadronic channel, the values of the maximum corrected data closures and data-based systematics for all N_{jets}

N_{jets}	**max Corrected Data Closure**	$t\bar{t}$ **syst.**
8	1.190	0.840
9	1.298	0.770
10	-	0.770
11	-	0.770
12	-	0.770
≥ 13	-	0.770

Table 8.7. For the semi-leptonic channel, the values of the maximum corrected data closures and data-based systematics for all N_{jets}

N_{jets}	**max Corrected Data Closure**	$t\bar{t}$ **syst.**
7	0.929	1.077
8	1.039	0.962
9	-	1.077
10	-	1.077
11	-	1.077
≥ 12	-	1.077

Table 8.8. For the fully-leptonic channel, the values of the maximum corrected data closures and data-based systematics for all N_{jets}

N_{jets}	**max Corrected Data Closure**	$t\bar{t}$ **syst.**
6	0.885	1.130
7	0.913	1.095
8	-	1.130
9	-	1.130
10	-	1.130
≥ 11	-	1.130

The corrected data closures as a function of boundary value for $t\bar{t}+$jets and all variations are shown in Figure 8.48., Figure 8.49. and Figure 8.50. for the fully-hadronic, semi-leptonic and fully-leptonic channels, respectively. The corrected data closure plots show that applying the closure correction is worked well—since the variations are covered after correction. Also, the data closure with and without correction can be comapred in these figures. Note that the corrected data closure are also given for second lowest N_{jets} bins (N_{jets} = 9 for the fully-hadronic channel, N_{jets} = 8 for the semi-leptonic channel, N_{jets} = 7 for the fully-leptonic channel) in Appendix D3..

For the semi-leptonic channel, comparing Table 8.7. and Figure 8.49., the maximum corrected data value for N_{jets} = 7 is taken from Val_CD (validation region II) where the boundary value is 0.4. This is as an example how the maximum value of the corrected data closure is chosen across all validation regions and all boundary values.

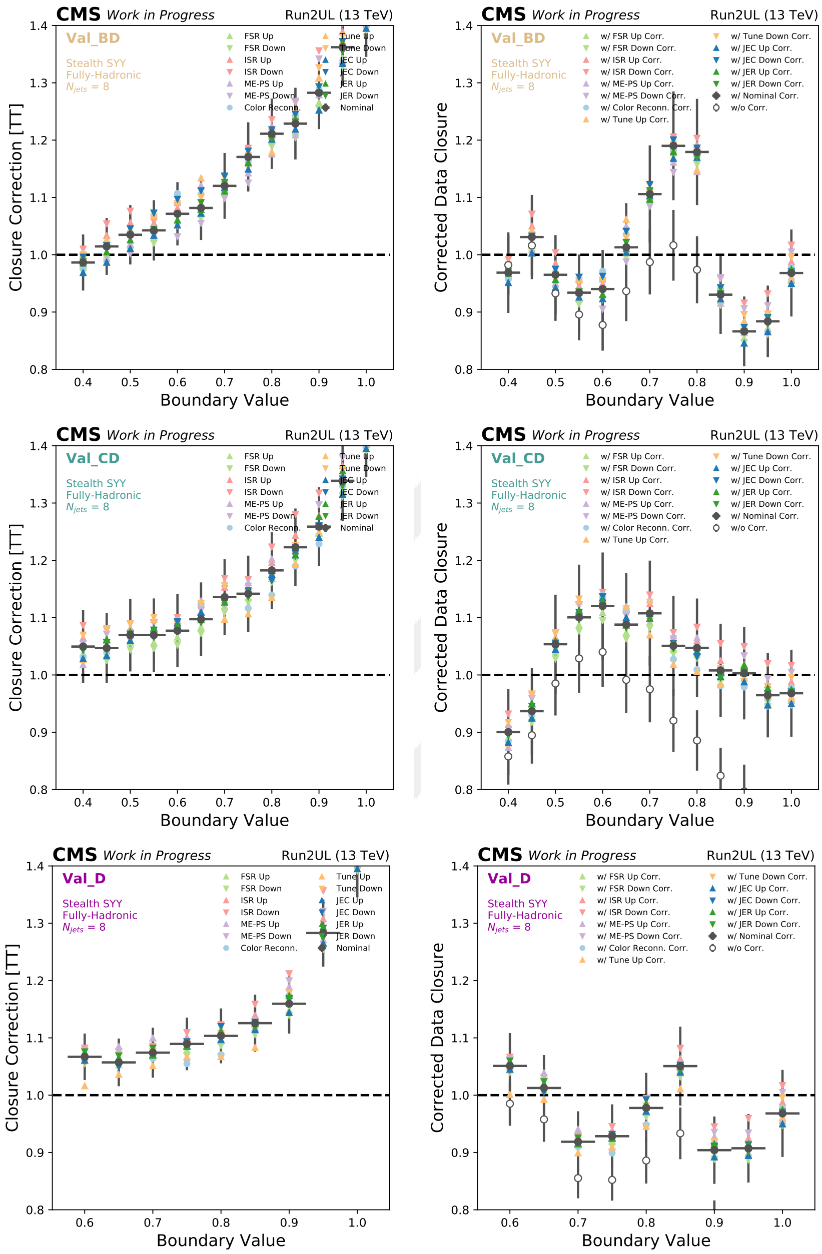


Figure 8.48. For the fully-hadronic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 8$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom)

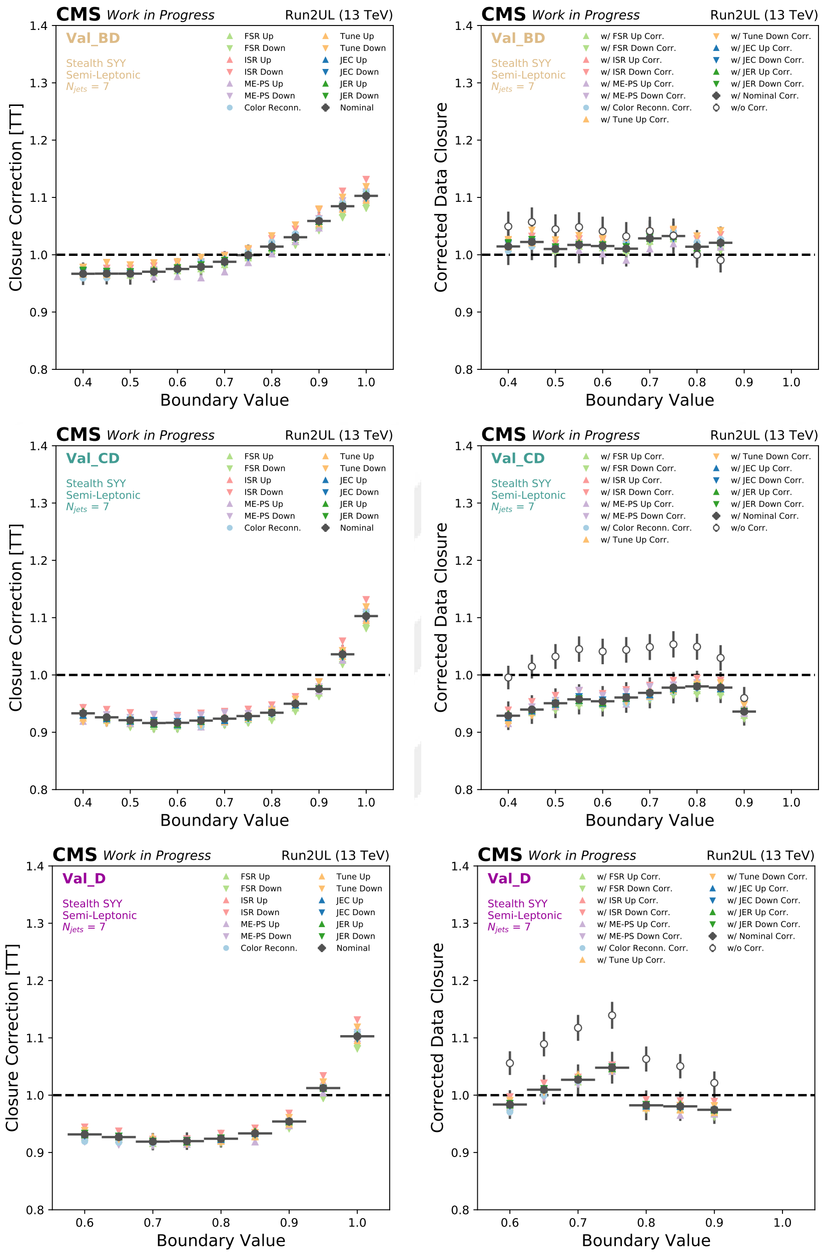


Figure 8.49. For the semi-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 7$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom)

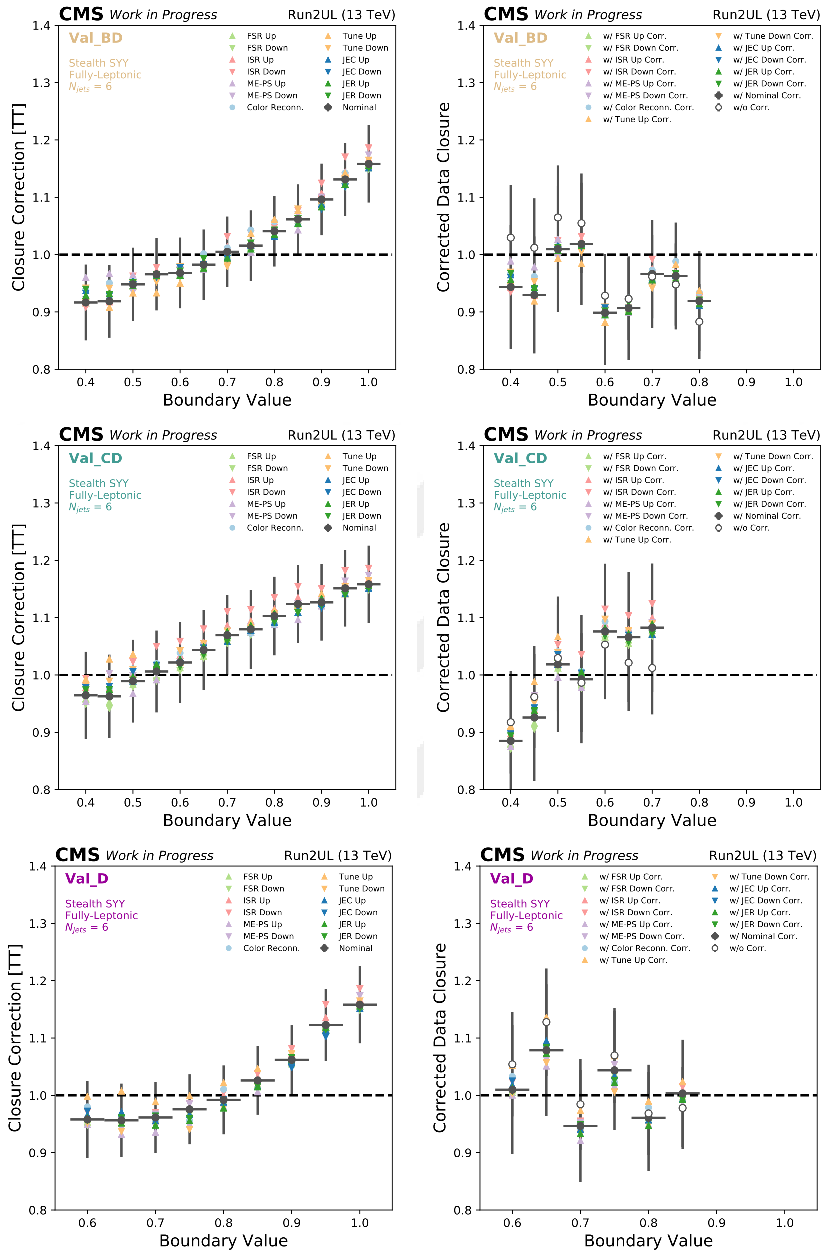


Figure 8.50. For the fully-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 6$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom)

9. OTHER ESTIMATIONS AND SYSTEMATICS

Most of the simulated events in the signal regions are expected to be $t\bar{t}+$jets, with contributions of less than 7% from QCD multi-jet and a few percent from the remaining minor backgrounds including "$t\bar{t}+$X " and "Other" backgrounds (given in Table 5.2.). Note that "$t\bar{t}+$X " is to where X is one or more W/Z/H bosons and "Other" is to which includes DY$+$jets, diboson, triboson, SingleTop$+$jets and $W+$jets. These processes are grouped into "$t\bar{t}+$X " and "Other" by the analysis. In this section, background estimations and related systematics of the QCD multi-jet in Section 9.1. and minor backgrounds in Section 9.2. are presented.

9.1. QCD multi-jet Estimation and Systematics

The QCD multi-jet background is the second largest background, after $t\bar{t}+$jets, for the fully-hadronic and semi-leptonic channels. However, the number of simulation events for this subdominant background that passes the signal region (SR) selections is low and some of these simulated events have a relatively large weight, which causes the statistical uncertainties associated with these events to be large. In fact, these large statistical uncertainties can hide a signal, because they give the fit too much freedom, which decreases the overall sensitivity of the analysis. Therefore, the QCD multi-jet background is estimated using data in a QCD-dominated control region (QCD CR) (defined in Section 5.4.2.) to avoid factoring in these large statistical fluctuations. For the fully-leptonic channel, total number of the QCD multi-jet events is negligible based on the simulation, and thus it is not considered in the final fit. Accordingly, the procedure defined in this section only applies to the fully-hadronic and semi-leptonic channels.

The **A**, **B**, **C**, **D** NN bins for this analysis are defined based on selections on the output of the Double DisCo NN used to estimate the $t\bar{t}+$jets background described in Section 8.2.. Therefore, the QCD multi-jet background must be estimated in these **A**, **B**, **C**, **D** NN bins, where there are different network trainings used for the fully-hadronic and semi-leptonic channels, for a given N_{jets} bin. For the semi-leptonic channel, the network discriminants are calculated for the QCD CR by using the non-isolated muon in place of the isolated lepton as an input to the NN. However, for the fully-hadronic channel, the non-isolated muon is not cleaned from the QCD CR jet collection and the jet collection is passed directly to the NN in the same manner as the SR. It is important to stress that the NN used in these QCD CR are the exact same NN used to evaluate the SR events. The data-simulation comparisons for the analysis search bins using the QCD CR selections is shown in Figure 9.1..

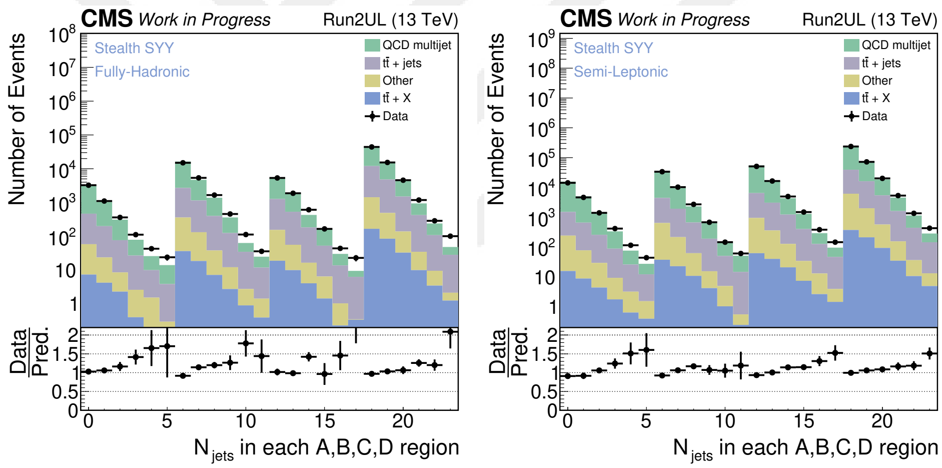


Figure 9.1. For the fully-hadronic (left) and semi-leptonic (right) channels, data *vs* simulation comparisons of the A, B, C, D NN bins in the two QCD CRs for the Stealth SY$\bar{Y}$ trainings

A comparison of the N_{jets} distributions of the QCD multi-jet simulation between the QCD CR and the SR demonstrates that the QCD CR is a good proxy for the SR. If the QCD CR is well-defined, the QCD CR must be a good proxy and the N_{jets} distributions of the QCD multi-jet simulation in QCD CR is matched with the N_{jets} distributions of the QCD multi-jet simulation in SR. The comparisons of the N_{jets} distributions for the fully-hadronic and semi-leptonic channels are shown in Figure 9.2.. It is considered that there is a good agreement of the N_{jets} distributions for QCD CR and SR.

An additional comparison of the Double DisCo NN Disc.1 and Disc. 2 outputs are also shown in Figure 9.2.. This additional comparison gives further confidence in the agreemnet that QCD CR selection is matched to SR selection. Note that these are normalized QCD multi-jet events for all N_{jets}, Disc. 1 and Disc. 2 distributions in each QCD CR and SR.

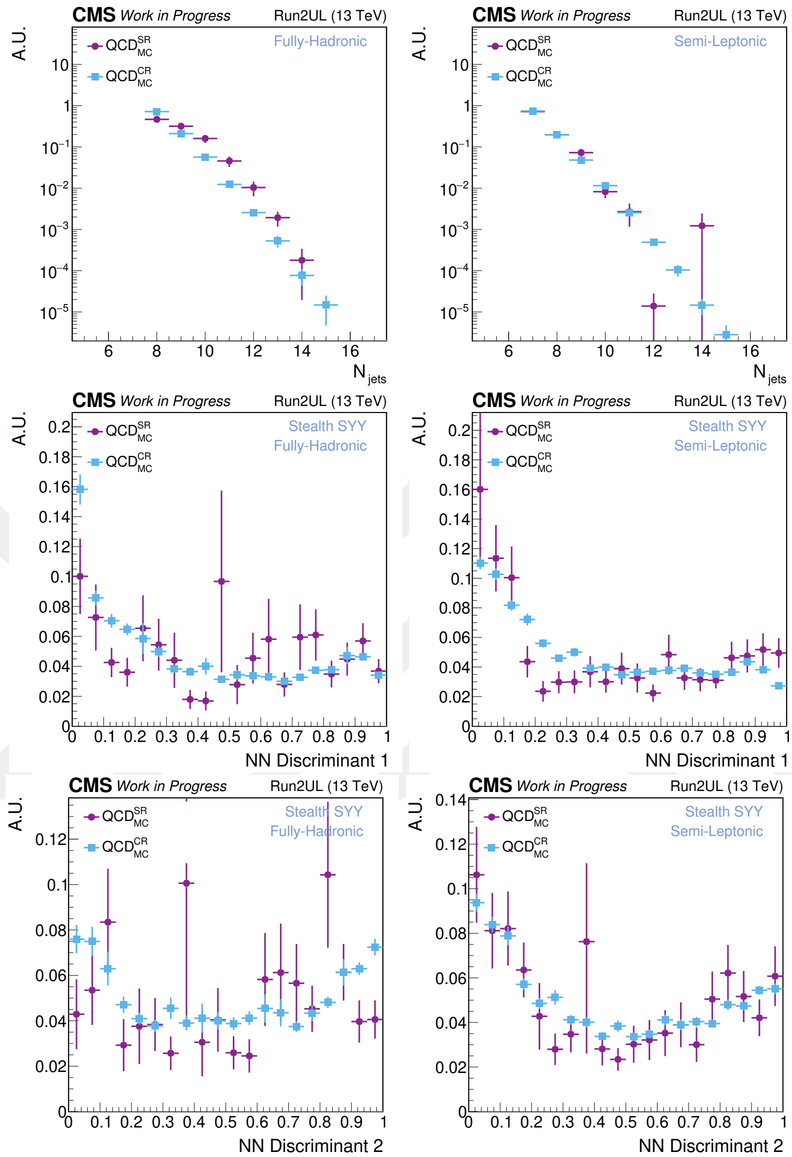


Figure 9.2. For the fully-hadronic (left) and semi-leptonic (right) channels, comparisons of the N_{jets} (top), Double DisCo NN Disc.1 (middle) and Disc.2 (bottom) distributions of the QCD multi-jet simulation between QCD CR and SR. Good agreement of the N_{jets} distributions for QCD CR and SR. Double DisCo NN trained on the Stealth SY$\bar{Y}$ model. Comparisons of Disc. 1 and Disc. 2 give further confidence that QCD CR selection is matched to SR selection

9.1.1. Estimation Method

To estimate the QCD multi-jet background, all the events that pass the QCD CR selection are divided into **A**, **B**, **C**, **D** NN bins utilizing the bin edges derived in the SR. Next, the non-QCD multi-jet events, taken from simulation, are subtracted from the data in the QCD CR. The QCD multi-jet simulation is then used to define a transfer factor (TF) for each of the A, B, C, D NN bins, which is the ratio of the QCD yield in the SR and QCD CR in simulation, given in Equation 9.1. Where i runs over the A, B, C, D bins, N_{MC}^{SR} is the number of weighted QCD multi-jet simulation events in the SR and the N_{MC}^{CR} is the number of weighted QCD multi-jet simulation events in the QCD CR.

$$\mathrm{TF}_i = \frac{N_{\mathrm{MC}}^{\mathrm{SR}}}{N_{\mathrm{MC}}^{\mathrm{CR}}} \tag{9.1}$$

The TF is used as an normalization factor and it is given for each of the **A**, **B**, **C**, **D** NN and for all N_{jets} bins for a given channel (all N_{jets} bins means that ≥ 8 for the fully-hadronic channel, ≥ 7 for the semi-leptonic channel, ≥ 6 for the fully-leptonic channel). The TF is given for each of the **A**, **B**, **C**, **D** NN bins in Table 9.1.. The exact equation demonstrating the transfer from the QCD CR to the SR is given in Equation 9.2. Here, the TF is multiplied by the number of QCD multi-jet CR events for data ($N_{\mathrm{Data},\,ij}^{\mathrm{CR}}$) for a given N_{jets} bin. This then gives a prediction for QCD multi-jet background for a particular bin in the SR (N_{ij}^{SR}), where i runs over the **A**, **B**, **C**, **D** NN bins and j runs over the N_{jets} bins.

$$N_{ij}^{\mathrm{SR}} = \mathrm{TF}_{\mathrm{MC},\,i}\, N_{\mathrm{Data},\,ij}^{\mathrm{CR}} \tag{9.2}$$

Table 9.1. For the fully-hadronic and semi-leptonic channels, TF values used for each of **A**, **B**, **C**, **D** NN bins for the QCD multi-jet background estimation. Note that **A**, **B**, **C**, **D** bins are separated by the NN trainings, targeting the Stealth SY$\bar{Y}$ signal

Year	channel	TF A (%)	TF B (%)	TF C (%)	TF D (%)
Run2	fully-hadronic	7.1 ± 1.2	4.3 ± 0.7	3.7 ± 0.8	4.8 ± 0.6
Run2	semi-leptonic	2.1 ± 0.4	2.1 ± 0.3	2.7 ± 0.3	2.0 ± 0.2

The QCD multi-jet background prediction in data (QCD^{CR}_{Data}) is compared to the QCD multi-jet background simulation in the SR (QCD^{SR}_{MC}) is shown in Figure 9.3. for the A, B, C, D bins separated by the Stealth SY$\bar{Y}$ NN trainings for each channel. The statistical uncertainty on the QCD multi-jet background prediction has improved and generally agrees with the prediction from the simulation.

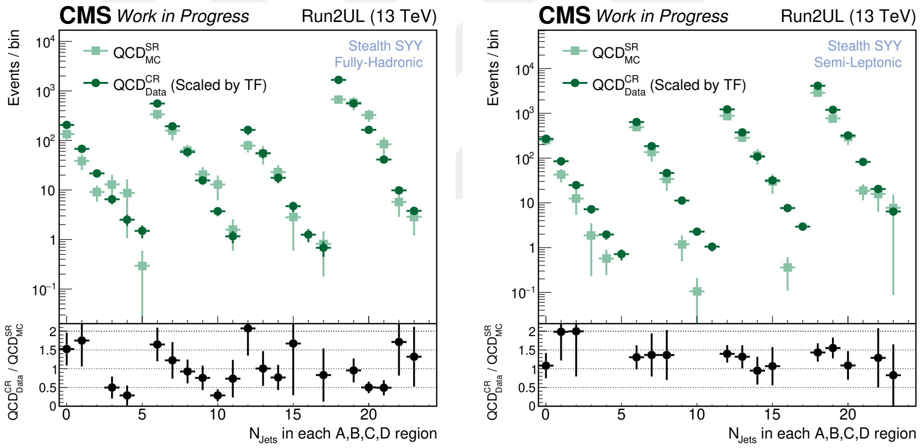


Figure 9.3. For the fully-hadronic (left) and semi-leptonic (right) channels, comparison between the QCD multi-jet background prediction and the QCD multi-jet background simulation in the SR

9.1.2. Systematics

The systematic uncertainty on the TF is taken to be the propagated statistical uncertainty on the two components of the ratio in Equation 9.1. This propagated statistical uncertainty in each of the **A**, **B**, **C**, **D** is given in Table 9.1., where it is roughly on the order of 10 to 30%. More details are given in Table 10.1..

9.2. Minor Background Estimations and Systematics

The minor backgrounds "$t\bar{t}+$X" and "Other", given in Table 5.2., are estimated directly by using simulation for each of **A**, **B**, **C**, **D** NN bin for a given in N_{jets} bin. Each of those is then put into the fit as a separate component. A range of systematic uncertainties is assigned to cover for potential mismodeling of simulation for these non-$t\bar{t}+$jets and non-QCD multi-jet background events, and they are included as nuisance parameters in the fit.

Sources of systematic uncertainty in the predictions for these minor backgrounds are given in four groups: **(1)** the luminosity uncertainty and the variation in pileup reweighting, **(2)** related to modeling of the detector simulation, which are prefiring efficiency, trigger efficiency, b tagging efficiency, top tagging efficiency, JEC and JER, **(3)** related to physics modeling, which are initial state radiation (ISR) modeling, final state radiation (FSR) modeling, parton distribution function (PDF), renormalization (μ_R) and refactorization (μ_F) scales, **(4)** the last one is "MC statistics" systematic. Note that more details for these systematic uncertainties are given with the other systematics based on signal, $t\bar{t}+$jets and QCD multi-jet backgrounds in Table 10.1. (in Section 10.1.).

10. FIT PROCEDURE AND VALIDATION

The main background in this analysis is $t\bar{t}+$jets and in the end we want to estimate it via the ABCD method in the "real data". So, a fit is designed to get the final prediction of $t\bar{t}+$jets background events in the real data. But, the fit still needs to be able to accommodate the non-$t\bar{t}+$jets backgrounds, because the real data includes more than just $t\bar{t}+$jets background events. Likewise, the fit can be allowed to also include a signal component, and in such a scenario, the most likely number of signal events in each analysis bin is also determined by the fit. Then, the fit is performed as a simultaneous binned fit to estimate the number of events for all processes. More details are given in Section 10.1.. The analysis is still ongoing (still blinded) and thus cannot yet look at real data. So, in this thesis, the fit is performed with simulated events and the preliminary fit results are given in Section 10.2. with some details.

After establishing the estimation method for each background given in the previous sections, the specific systematic uncertainties based on the each estimation method are derived for the $t\bar{t}+$jets and QCD multi-jet backgrounds, as given in Section 8.6. and Section 9.1.2., respectively. For each of the minor backgrounds and signal processes, the systematic uncertainties are just based on the mismodeling of simulation. For performing the fit, the fit components given in Table 10.1. are implemented in the Higgs Combine tool (Higgs Combine Tool, 2021) framework, which is based on the fitting software package RooStat (RooStat, 2022).

10.1. Fit Procedure

The fit is performed as a simultaneous binned fit to estimate the total number of events for each process in each N_{jets} category. There is a total of 6 N_{jets} bins, ranging from $N_{\text{jets}} = 8$ to $N_{\text{jets}} \geq 13$ for the fully-hadronic channel, from $N_{\text{jets}} = 7$ to $N_{\text{jets}} \geq 12$ for the semi-leptonic channal and from $N_{\text{jets}} = 6$ to $N_{\text{jets}} \geq 11$ for the fully-leptonic channel in one jet incriments. Events in each N_{jets} bin then are subdivided into the **A**, **B**, **C**, and **D** subregions as defined by the two neural network discriminant values, where the division lines of **A**, **B**, **C**, and **D** are the same for each N_{jets} bin, and determined through an optimziation procedure, as described in Section 8.4.1.. These subregions are defined by splitting the double discriminant two-dimensional plane along both of the discriminant ranges.

Fit Function

The maximum likelihood method (Tanabashi, 2018) is used as the fit method. In this method a probability distribution function is used to determine probability of observed data given a set of parameters. The likelihood in particle physiscs is donated as in Equation 10.1, where r is the signal strength, θ is the nuisance parameters used to define the probability distribution function, s and b are the number of final predicted signal and background events, respectively.

$$\mathcal{L}(data|r \cdot s(\theta)\ +\ b(\theta)) \tag{10.1}$$

For performing the fit, the fit components are fed into the Higgs combine tool. A systematic uncertainty is one of the fit components and each process has own systematic uncertainty. The systematic uncertainty is a way of quantifying how much expected data is different from actual data and any difference can be due to imperfect modeling of the data. Then, it is included in the fit as given in Equation 10.2.

$$N_{\text{final}} = N_{\text{initial}} R^{\theta} \tag{10.2}$$

where N_{final} is the number of final predicted events for a given process, $N_{initial}$ is the number of initial predicted events for a given process, R is an actual value of the systematic uncertainty for a given process, and θ is a nuisance parameter, which is used to adjust the size of R and it gives additional freedom to the fit. The Higgs combine automatically figures out the value of the θ in between (-1,1). If it is 1, the systematic is fully applied in the fit. If it is 0, this means that systematic is not used in the fit.

Typically, there is more than one systematic uncertainy for a given process. For example, the $t\bar{t}+$jets background has a simulation-based and a data-based systematic uncertainties, that would multiply out in the form given in Equation 10.3, where the θ_i are then nuisance parameters introduced into the fit.

$$N^{t\bar{t}\text{+jets}}_{\mathbf{A}\text{, final pred.}} \to N^{t\bar{t}\text{+jets}}_{\mathbf{A}\text{, initial pred.}} R_1^{\theta_1} R_2^{\theta_2} \ldots \tag{10.3}$$

Systematic Uncertainties

In this analysis, the six group of systematic uncertainties are applied for each analysis channel, excepting that QCD multi-jet systematic is not applied for the fully-leptonic channel. All of those groups of systematics are outlined with their associated nuisance parameters in Table 10.1.. All these systematics are explained in the text.

Note that some of the systematic uncertainties are applied for a specific channel. Also, some of them are not applied for signal. The first and forth groups are applied for signal and all background processes. The second group is applied for signal and all minor background processes whereas the third group is applied for all minor background processes.

Note that the “value” is the initial value used in the fit for the rateParam cases, but that it is free to float. This means that fit can change these values.

Table 10.1. The systematic uncertainties in the final fit for all backgrounds processes and Stealth SY$\bar{Y}$ signal model. The first group for all background and signal processes excepting that lumi is not for $t\bar{t}+$jets and QCD multi-jet. The second groups for minor backgrounds and signal processes, the third group for only minor background processes, the forth group for signal and all background processes. The last the two groups for $t\bar{t}+$jets and QCD multi-jet background processes, respectively. Abbreviated names for each source of uncertainty are explained in the text, denoting the related N_{jets} and ABCD bins

Systematic Uncertainty	**Type**	**Value**
Lumi Uncertainty	lnN	1.6%
Pile Up	lnN	Up/Down variations
Prefiring	lnN	Up/Down variations
Trigger	lnN	Up/Down variations
b-Tag	lnN	Up/Down variations
top-Tag	lnN	Up/Down variations
JEC	lnN	Up/Down variations
JER	lnN	Up/Down variations
ISR	lnN	Up/Down variations
FSR	lnN	Up/Down variations
PDF	lnN	Up/Down variations
Scale (μ_R, μ_F)	lnN	Up/Down variations
MC Stat.	gmN	Statistical uncertainty on raw event number and avg. event weight
$t\bar{t}+$jets Background Prediction in B,C,D	rateParam	0 to 10 times MC prediction
$t\bar{t}+$jets Background Prediction in A	rateParam	$N_{\text{pred.},A} = (N_{\text{pred.},B}\, N_{\text{pred.},C}) \,/\, N_{\text{pred.},D}$
$t\bar{t}+$jets Closure Correction	fixed	$N_{\text{obs.},A} \,/\, N_{\text{pred.},A}$
$t\bar{t}+$jets Closure Correction	lnN	Statistical uncurtainty on Closure Correction
$t\bar{t}+$jets Corrected Data Closure	lnN	Closure Correction $* \left.\frac{N_{\text{pred.},A}}{N_{\text{obs.},A}}\right\rvert_{\text{Data}}$
QCD multi-jet TF (Transfer Factor)	fixed	$\text{TF}_i = \frac{N_i^{SR}}{N_i^{\text{CR}}}$
QCD multi-jet TF	gmN	Statistical uncertainty on TF

The first group of systematics: The same luminosity uncertainty R_{lumi} value is applied globally to all N_{jets}, ABCD bins for signal and all considered minor background processes and acts as a global modification of the total number of predicted events. In a similar way, the reweighting of events based on the simulated amount of pileup is varied "up" and "down" to yield $R_{\text{PU up/down}}$ unique to each N_{jets}, ABCD bin for minor background and signal processes.

The second group of systematics: They are all related to modeling of the detector simulation. The prefiring efficiency is to varied "up" and "down" from its nominal value in order to produce $R_{\text{prefire up/down}}$. The trigger efficiency (for a given channel) is varied "up" and "down" from its nominal value in order to produce $R_{\text{prefire up/down}}$. The b tagging efficiency is varied up and down from its nominal value in order to produce $R_{\text{btagging up/down}}$. The top tagging efficiency (for only fully-hadronic channel) is varied up and down from its nominal value in order to produce $R_{\text{ttagging up/down}}$. All these above are scale factors for fixing disagreement between simulation and data, and they are applied for each N_{jets}, ABCD bin. Lastly, there are variations of the JEC and JER corrections. These variations are slightly different than the previously mentioned ones as they modify the 4-vector information of jets in event. Thus, the number of jets that are within acceptance are allowed to change and/or the neural network can score the event differently in the 2D discriminant plane. In either case, the number of events in each N_{jets}, ABCD bin will change allowing the calculation of an $R_{\text{JEC/JER up/down}}$. All these systematics are applied for minor background and signal processes.

The third group of systematics: They are related to physics modeling. These purely event weight-based variations correspond to changes in initial state radiation (ISR) and final state radiation (FSR) modeling, parton distribution function (PDF), and different renormalization (μ_R) and factorization (μ_F) scale choices. The event weight is modified allowing the calculation of $R_{\text{ISR, FSR, etc... up/down}}$. They are applied for each N_{jets}, ABCD bin for all minor background processes.

The forth group of systematics: A singular "MC statistics" systematic is also present. This corresponds to uncertainty on raw event number and average event weight in each N_{jets}, ABCD bin due to finite counting statistics. A unique, independent R value for each N_{jets}, ABCD bin is present in the fit. This systematic is applied for signal and all background processes.

The fifth group of systematics: Specifically for $t\bar{t}+$jets, there is the statistical uncertainty on the closure correction as well as the data-based systematic, discussed in Section 8.6.1. and Section 8.6.2., respectively.

The sixth group of systematics: Specifically for QCD multi-jet, the statistical uncertainty in the TFs are used in the fit as well.

Final Predictions for All Processes

The final prediction from the fit is the total number of expected events in region **A** for each N_{jets} bin. The fit itself is comprised of several different components as given in Equation 10.4. In total, Equation 10.4 is the prediction of number of events in region **A**, with an equivalent equation for each N_{jets} bin.

$$N_{\mathbf{A}}^{\text{Data}} \sim N_{\mathbf{A}}^{\text{Pred.}} = \frac{N_{\mathbf{B}}^{t\bar{t}+\text{jets}} N_{\mathbf{C}}^{t\bar{t}+\text{jets}}}{N_{\mathbf{D}}^{t\bar{t}+\text{jets}}} \text{CC}_{\mathbf{A}}^{t\bar{t}+\text{jets}} + N_{\mathbf{A}}^{\text{QCD}} + N_{\mathbf{A}}^{\text{minor}} \tag{10.4}$$

The main $t\bar{t}+\text{jets}$ background component is estimated via the ABCD relationship given as the first term in Equation 10.4, where the number of $t\bar{t}+\text{jets}$ events in data in the **B**, **C**, and **D** regions is constrained by the simultaneous prediction of all the other backgrounds in those three regions, as specified in Equation 10.5. This leaves the $N_{\mathbf{B},\mathbf{C},\mathbf{D}}^{t\bar{t}+\text{jets}}$ parameters for each N_{jets} bin as floating parameters in the fit. These parameters can float between 0 and 10 times their value predicted directly by simulation, and their final value is then simply determined in the best fit. Note that the closure correction ($\text{CC}_{\mathbf{A}}^{t\bar{t}+\text{jets}}$) derived in Section 8.5. is also multiplied on the ABCD expression.

$$\begin{aligned} N_{\mathbf{B}}^{\text{Data}} &= N_{\mathbf{B}}^{t\bar{t}+\text{jets}} + N_{\mathbf{B}}^{\text{QCD}} + N_{\mathbf{B}}^{\text{minor}} \\ N_{\mathbf{C}}^{\text{Data}} &= N_{\mathbf{C}}^{t\bar{t}+\text{jets}} + N_{\mathbf{C}}^{\text{QCD}} + N_{\mathbf{C}}^{\text{minor}} \\ N_{\mathbf{D}}^{\text{Data}} &= N_{\mathbf{D}}^{t\bar{t}+\text{jets}} + N_{\mathbf{D}}^{\text{QCD}} + N_{\mathbf{D}}^{\text{minor}} \end{aligned} \tag{10.5}$$

The QCD multi-jet component of the fit (for only the fully-hadronic and semi-leptonic channels) is the number of QCD data events in the QCD CR, scaled by the TF taken as the ratio of the amount of simulation-predicted QCD multi-jet events in the SR to that in the QCD CR for each **A**, **B**, **C**, **D** region. Additionally, Equation 10.6 gives that the QCD multi-jet predictions comes through a relation between the SR and

QCD CR where the TF is from Equation 9.1.

$$
\begin{aligned}
N^{\mathrm{QCD}}_{\mathbf{A},\,\mathrm{SR}} &= \mathrm{TF}_{\mathbf{A}} N^{\mathrm{QCD\ Data}}_{\mathbf{A},\,\mathrm{CR}} \\
N^{\mathrm{QCD}}_{\mathbf{B},\,\mathrm{SR}} &= \mathrm{TF}_{\mathbf{B}} N^{\mathrm{QCD\ Data}}_{\mathbf{B},\,\mathrm{CR}} \\
N^{\mathrm{QCD}}_{\mathbf{C},\,\mathrm{SR}} &= \mathrm{TF}_{\mathbf{C}} N^{\mathrm{QCD\ Data}}_{\mathbf{C},\,\mathrm{CR}} \\
N^{\mathrm{QCD}}_{\mathbf{D},\,\mathrm{SR}} &= \mathrm{TF}_{\mathbf{D}} N^{\mathrm{QCD\ Data}}_{\mathbf{D},\,\mathrm{CR}}
\end{aligned} \tag{10.6}
$$

Last, the number of events coming from each of the minor backgrounds and signal processes are estimated directly from simulation for all N_{jets} bins in each **A**, **B**, **C**, and **D** region.

10.2. Fit Validation

Several fitting scenarios are performed to test the fit setup and validate its behavior. Also, these are the preliminary results, because the analysis is still ongoing (still blinded) and thus cannot yet look at real data. So, constructed data, which are based on simulation, are used in place real data to fit. There are two types of fit. One is that fit includes only background component. Another is that fit includes both background and signal components. For the fits shown below, "pseudoData" and "pseudoDataS" are constructed date and used in place of real data. To contruct "pseudoData", the number simulated events for all background processes are added together for each N_{jets}, ABCD bins. Additionally, a "pseudoDataS" is constructed, which the "pseudoData" and the number simulated events for signal are added together for each N_{jets}, ABCD bins. There are three scenarios are presented in this thesis: (1) "Background Only" fit to pseudoData, (2) "Background Only" fit to pseudoDataS, (3) "Background + Signal" fit to pseudoDataS.

“Background Only” fit to pseudoData

This is the most basic test for the fit setup and it checks how stable the fit is and the number of final predicted background events are reasonable in each N_{jets}, ABCD bins.

“Background Only” fit to pseudoData are given for the fully-hadronic, semi-leptonic and fully-leptonic channels in Figure 10.1.. In each plot, the purple line shows the fit. The black point shows the number of total observed background events in each **A**, **B**, **C**, **D** region for a given N_{jets} bin and it refers pseudoData. The green line is the number of signal events given as reference. Also, each bottom panel shows the goodness-of-fit which is a quatity to tell how much good the fit is. If the fit is good, then number of events predicted by fit should be matched with the number of observed background events in each **A**, **B**, **C**, **D** region for a given N_{jets} bin.

Background Only Fit to pseudoData for each channel shows that the fit is stable and the number of total background events can be predicted accurately in each **A**, **B**, **C**, **D** region for a given N_{jets} bin. Also, the goodness-of-fit is about 0, which means that the fit set up is good.

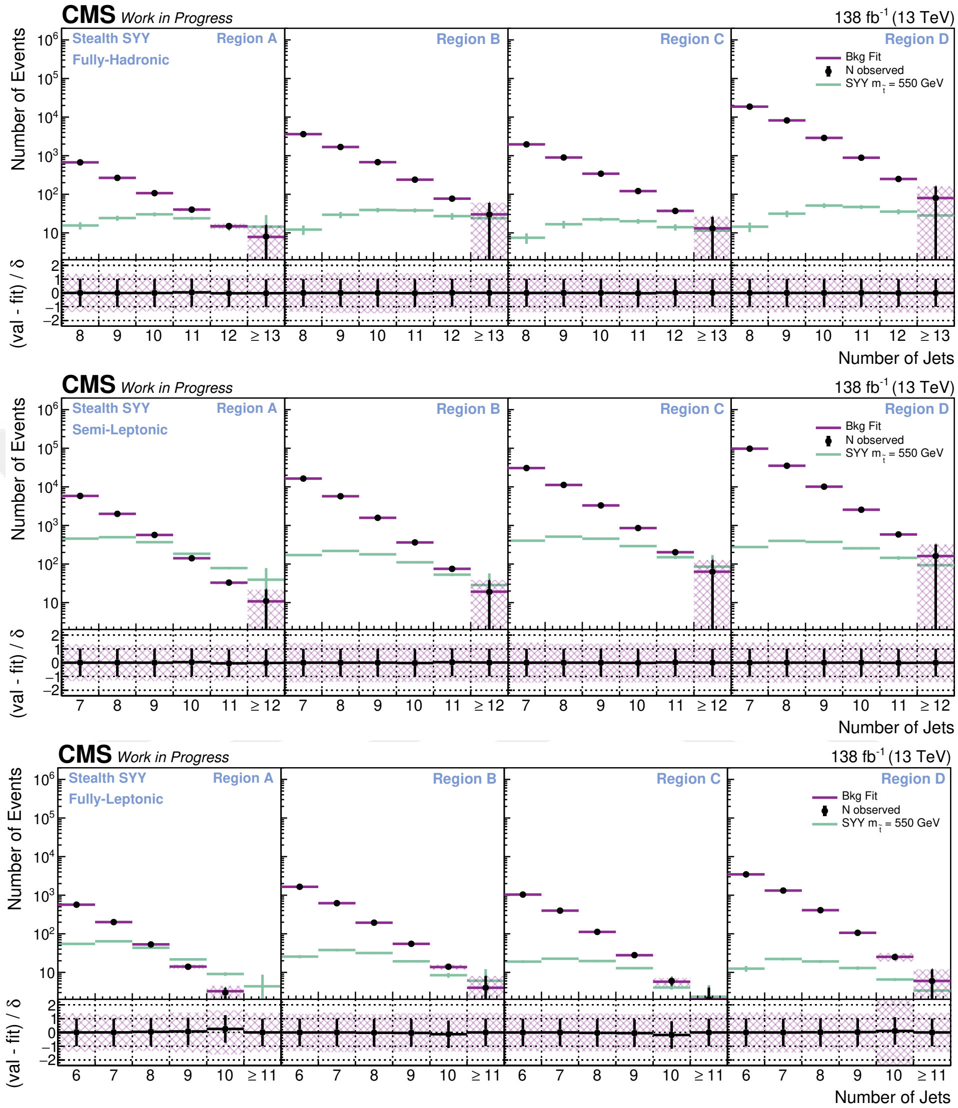


Figure 10.1. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the "Background Only" fit to pseuodaData. In each plot, the bottom panel shows the goodness-of-fit. The signal is overlaid for illustration purposes here, it is not included in the fit

“Background Only” fit to pseudoDataS

This can show if the fit sensitive to the presence of the signal. If the “Background Only” fit no longer works well when including signal in the data, this demonstrates sensitivity to the signal which means that the fit can detect the signal in the data.

“Background Only” fit to pseudoData are given for the fully-hadronic, semi-leptonic and fully-leptonic channels in Figure 10.2.. In each plot, the purple line shows the fit. The black point shows the number of total observed background and signal events in each each **A**, **B**, **C**, **D** region for a given N_{jets} bin, since it refers pseudoDataS. The green line is the number of signal events given as reference.

Background Only Fit to pseudoDataS shows whether the fit is sensitive to signal or not. It means whether goodness-of-fit will be changed or not when the signal events will be included. If the fit sensitive to the signal, then the less fluctuation could be on the goodness-of-fit in respect of that its value is zero. Considering the goodness-of-fit for each channel, the semi-leptonic channel is the most sensitive to the signal, whereas the fully-hadronic channel is the least sensitive. Based on these results, we can say that the semi-leptonic channel leads the analysis.

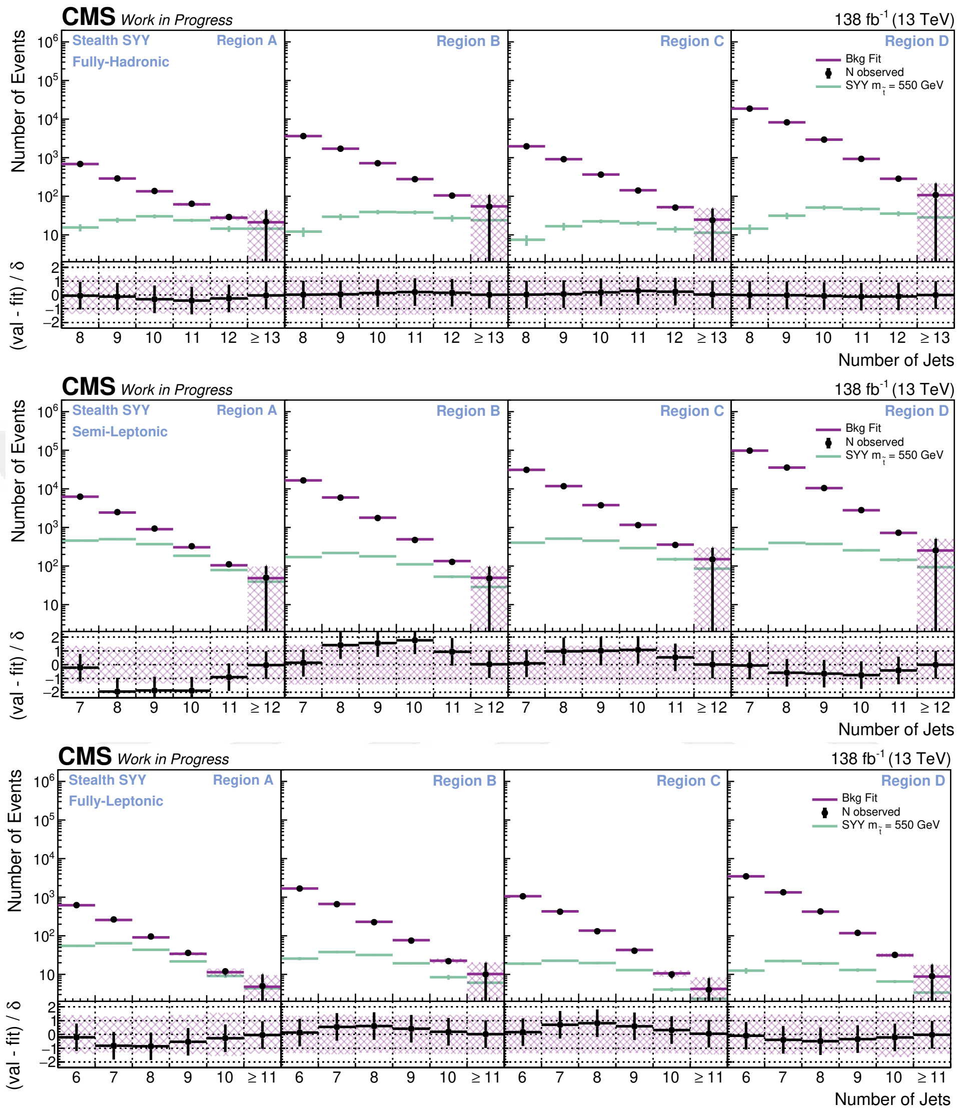


Figure 10.2. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the "Background Only" fit to pseudaDataS. In each plot, the bottom panel shows the goodness of fit. Note that the pseudaDataS is with Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV

“Background + Signal” fit to pseudoDataS

This type of fit completes the result for “Background Only fit to pseudoDataS”. By including a signal component in the fit, the prediction for number of events is more accurate, because the fit is allowed to add signal events.

“Background + Signal” fit to pseudoDataS are given for the fully-hadronic, semi-leptonic and fully-leptonic channels in Figure 10.3.. In each plot, the purple line shows the fit. The black point shows the number of total observed events in each each **A**, **B**, **C**, **D** region for a given N_{jets} bin, since it refers pseudoDataS. Based on the “Background + Signal” fit, the green line is the number of observed signal events whereas the blue line is the number of observed background events.

In this type of fit, the signal component is added to the fit. When this type of fit is performed on pseudoDataS, the fit introduces signal to obtaion a good quality overall fit. The good quality “Background + Signal” fit compared to “Background Only” fit, both using pseudoDataS, further demonstrates the sensitivity of the fit to signal. In terms of comaparing each channel, they all behaive similarly.

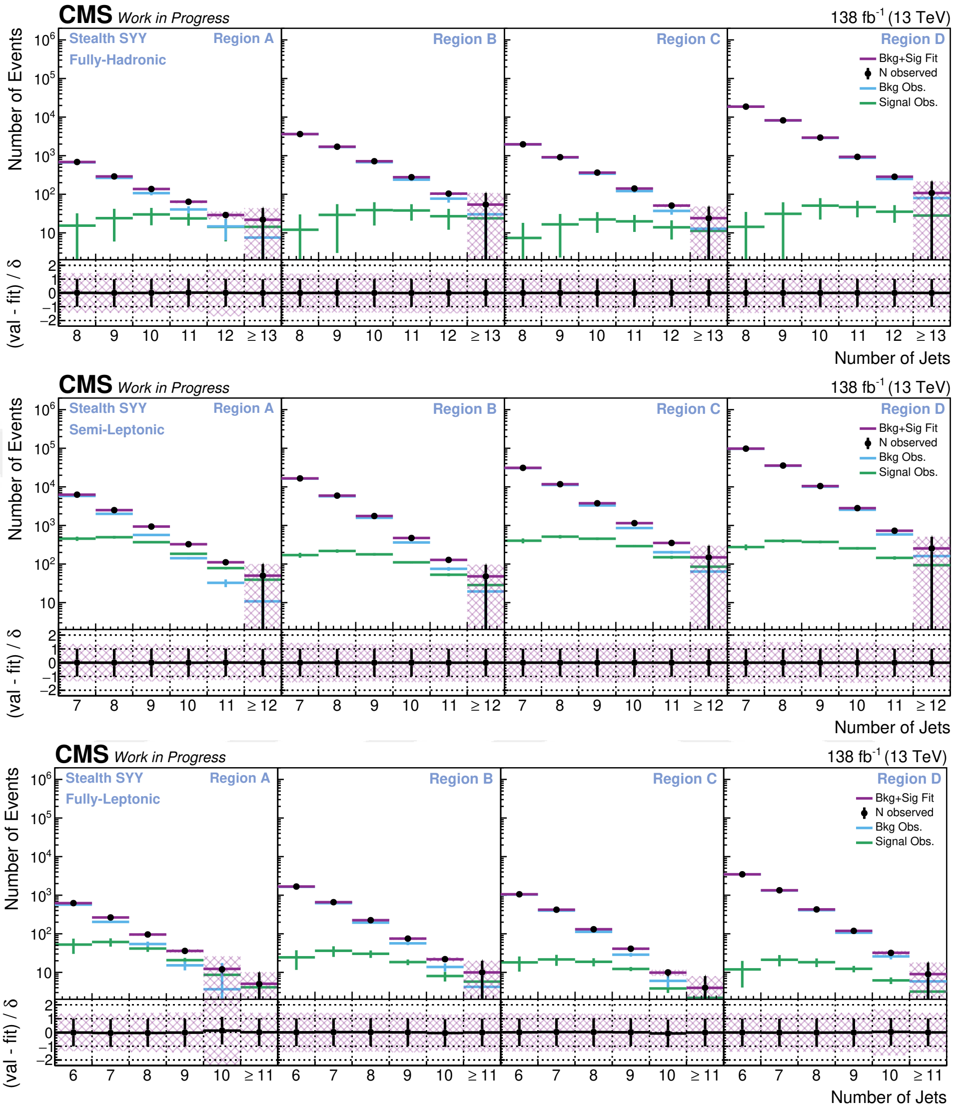


Figure 10.3. For the fully-hadronic (top), semi-leptonic (middle) and fully-leptonic (bottom) channels, the “Background + Signal” fit to pseudaDataS. In each plot, the bottom panel shows the goodness of fit. Note that the pseuodaDataS is with Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV

Fit Parameter Impacts

Impact plots show the relative importance of the fit parameters. These plots are generated for the "Background + Signal" fit with the "Asimov style" data sets (Cowan et al., 2011). The "Asimov style" data sets are generated by setting all fit parameters to their default values and by evaluating the fit. All nuisance parameters are set to zero and one at a time the nuisance parameters are set up by fit. For each nuisance parameter that is set up by the fit, the signal strength is determined. The importance is quantified based on how much the signal strength ($\hat{r}$) changes when a given fit parameter is changed from its nominal value. Here a signal strength of 0 implies that the fit does not include any signal events. In contrast, a signal strength of 1 means that the fit includes exact number of signal events as predicted by simulation.

For the Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV, the "Asimov style" impact plots for the semi-leptonic channel are given in Figure 10.4.. The top 30 parameters are shown in each plot. Some of the most important parameters are based on $t\bar{t}+$jets background estimation. Following them, most of the others are related to the $t\bar{t}+$jets background systematics. Consediring both pseuodaData and pseuodaDataS, the expected signal strength is obtained, where $\hat{r}$ = 0 with $\pm$11% uncerstainty for pseuodaData and $\hat{r}$ = 1 with $\pm$14% uncerstainty for pseuodaDataS.

Note that this is "Asimov style" impact plot and generally it is relevant to check the impact of the fit parameters (red and blue bars) on the signal strength. Also, these plots are shown for the fully-hadronic and fully-leptonic channels in Appendix E.

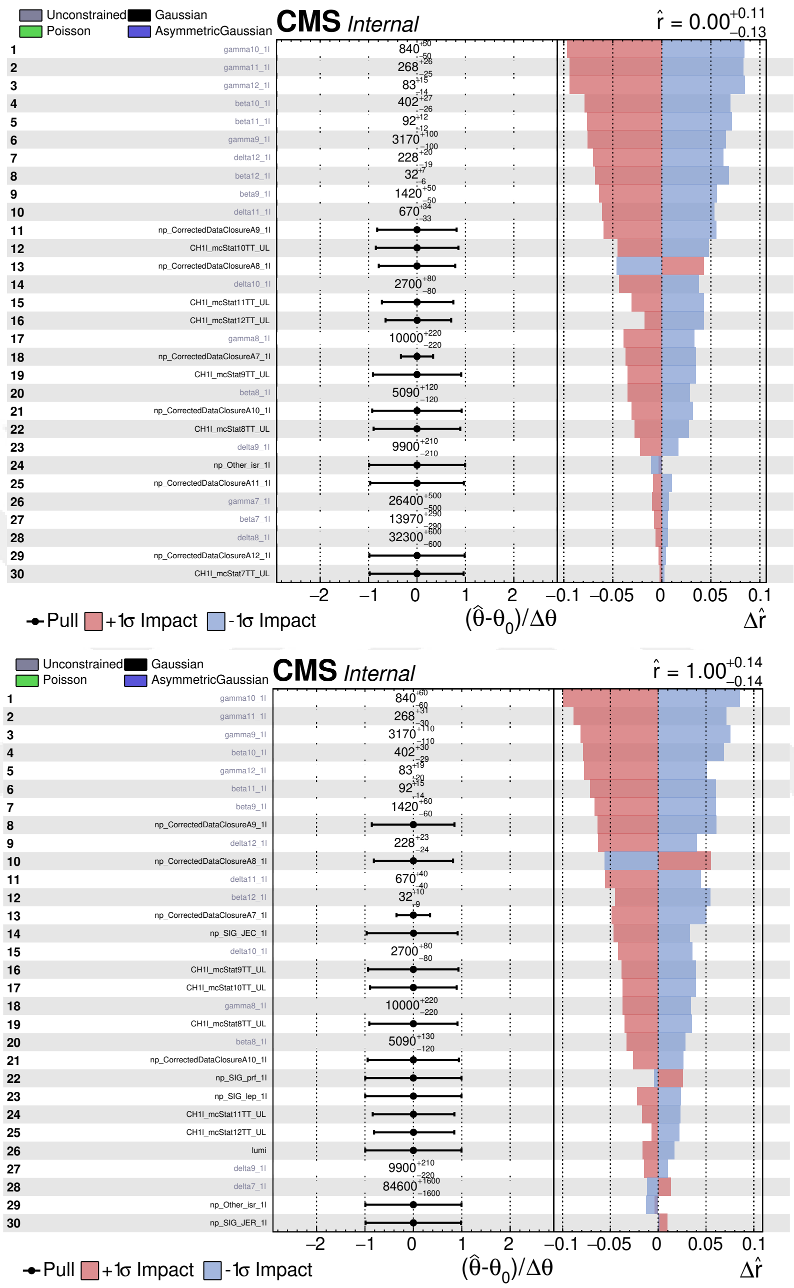


Figure 10.4. For the semi-leptonic channel, “Asimov style” impact plots for Stealth $SY\bar{Y}$ $m_{\tilde{t}} = 550$ GeV with pseudoData (top) and pseudoDataS (bottom). $\hat{r} = 0$ with $\pm 11\%$ uncerstainty for pseudoData and $\hat{r} = 1$ with $\pm 14\%$ uncerstainty for pseudoDataS

11. PRELIMINARY RESULTS AND INTERPRETATION

The expected exclusion limits for the Stealth SY$\bar{Y}$ model are shown in Figure 11.1.. Results are separated for each channel, as well as their combination. The limits are plotted as a function of top squark mass, where certain assumptions are made about how the top squarks decay. The “red-orange line” ($\sigma_{\tilde{t}\,\bar{\tilde{t}}}$) denotes the cross section for the top squark pair production. The “black dots” are the observed limit for each mass point and the dark blue dashed line is the expected limit. The “green and yellow bands” show different levels of uncertainty on the expected limit.

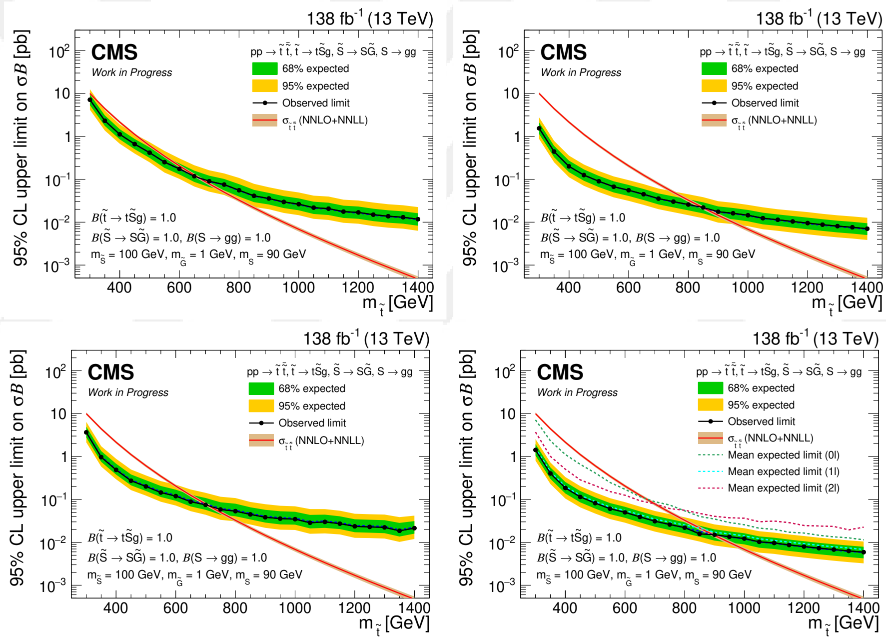


Figure 11.1. For the fully-hadronic (top left), semi-leptonic (top right), fully-leptonic (bottom left) and combination of all channels (bottom right), the observed limits for Stealth SY$\bar{Y}$ model. The analysis is sensitive to the Stealth SY$\bar{Y}$ model up to $m_{\tilde{t}}$ = 900 GeV. Note that limits are with “Asimov style”

Limits are the smallest pair production cross section for a given mass of the top squark, where the analysis could still tell if the signal exists. If the cross section were smaller than the limit, the analysis will not to be sensitive to a presence of the signal. Thus, the analysis is sensitive to the Stealth SY$\bar{Y}$ model from a $m_{\tilde{t}}$ = 300 GeV to a $m_{\tilde{t}}$ = 900 GeV considering the statistically combination of the all three channel results. Likewise, the analysis is unable to conclude anything about the Stealth SY$\bar{Y}$ model for above $m_{\tilde{t}}$ = 900 GeV.

The local p-values as a function of top squark mass for the Stealth SY$\bar{Y}$ model with Run2 simulated events are shown in Figure 11.2.. The local p-values are shown for each channel separately, as well as their combination. Results are shown for separate fits ("Background + Signal" fit to pseudoDataS which is the observed data that since the analysis is still blinded and thus cannot yet look at real data) of the fully-hadronic (green dashed line), semi-leptonic (blue dashed line), fully-leptonic (red dashed line) channels, and combination of all tree channels (black line). Equivalent significance values are denoted on the far right y-axis of the plot. The bottom panel of the plots shows the measured signal strengths ($\hat{r}$) for combination of the all channels and the green band is the uncertainty on the measured signal strength.

These local p-values directly quantify the analysis sensitivity to the Stealth SY$\bar{Y}$ model for different masses of the top squark. Assuming the signal exists with the specified cross section for the top squark pair production, a small p-value means that the analysis can significantly tell that the signal is present. In this way, the analysis is most sensitive to low mass of the top squark in particular $m_{\tilde{t}}$ = 400 GeV. Given that p-value results are made with using pseudoDataS, the expected measured signal strength is $\hat{r} = 1$. This expected value is exactly what is found, confirming that the fit setup works well.

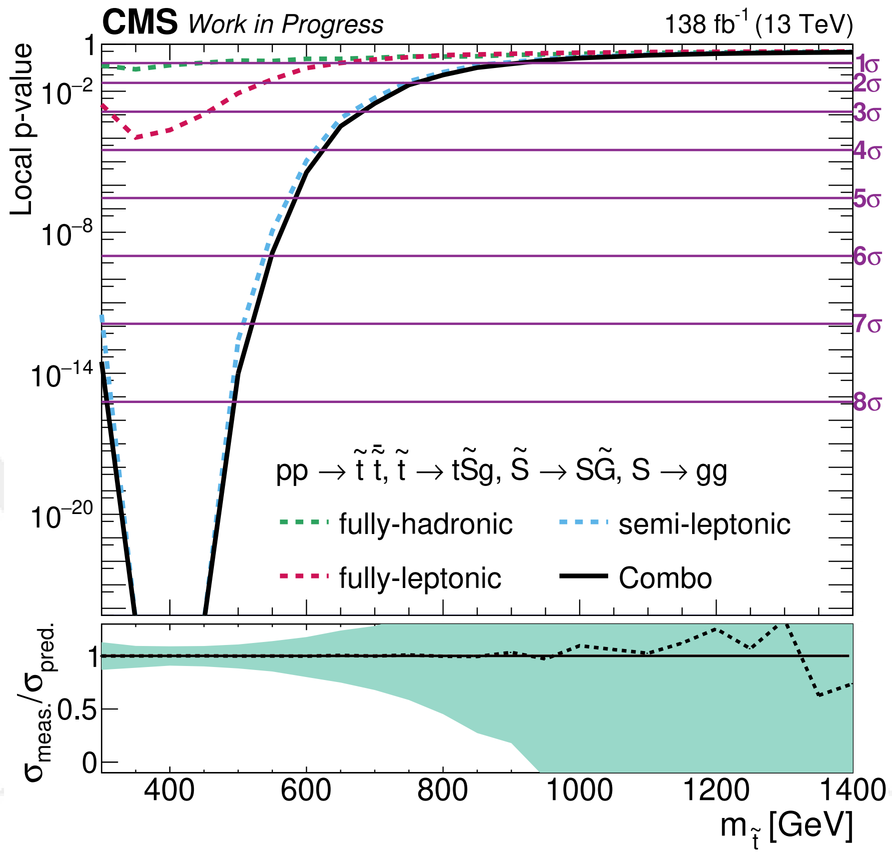


Figure 11.2. The local p-values as a function of top squark mass for the Stealth SY$\bar{Y}$ model with Run2 simulated events. Results are shown for separate fits of the fully-hadronic (green dashed line), semi-leptonic (blue dashed line), fully-leptonic (red dashed line) channels, and combination of all three channels (black line). The bottom panel shows the signal strength for combination of the all channels with uncertainty denoted by the green band. Note that limits are with "Asimov style"

12. CONCLUSION

This thesis presented a novel search for any sign of new physics beyond the SM, including a version of SUSY characterized by Stealth SY$\bar{Y}$ SUSY. This was the preliminary results of a pair production of top squark featuring a final state with two top quarks, many extra light-flavor jets and no extra missing transverse momentum (p_{T}^{miss}). This search was mostly unexplored top squark decays, since high-p_{T}^{miss} top squark decays generally are focused at the LHC.

The analysis was conducted, using Run2UL data collected with the CMS detector at the LHC from 2016 to 2018 and correspond to a total integrated luminosity of 138 fb^{-1}. With a non-negligible motivation from legacy analysis, the fully-hadronic and fully-leptonic channels were also included. Combining the final results for the fully-hadronic, semi-leptonic and fully-leptonic channels could make a stronger statement about the "past results". On the other hand, adding the fully-hadronic channel brings some challenges, since most of the backgrounds come from this channel based on the decay branching fraction of the top quarks in collision. There are also extra challenges for this channel due to QCD multijet events, which contain gluon splitting to pairs of b quarks. So, the hadronic top tagger, which is a NN including a combination of two approaches to tag the hadronically decaying top quarks in all p_{T} ranges, was implemented as part of the SR selection of the fully-hadronic channel.

Considering the legacy analysis, a more standard background estimation technique called the ABCD method was used for the main irreducible background ($t\bar{t}+$jets) in order to try and have fewer systematic uncertainties regarding to the "past results". The ABCD method was used to choose two variables, which are independent for background and also demonstrate separation between signal and background. As the $t\bar{t}+$jets background and signal have very similar final-state topologies, a double discriminant NN was used to generate the two signal-vs-background-discriminating. This analysis was novel because of its implementation of the Double DisCo NN. With the Double DisCo NN trained on simulated events for all years of Run2 for each of the three analysis channels individually, the ABCD method can be performed. To attain the best signal sensitivity, events were categorized by the number of jets (N_{jets}), to make use of the signal's high jet multiplicity. Then, the ABCD method was performed simultaneously for each N_{jets} bin where sensitivity is maximized in the high N_{jets} bins.

The analysis was matured and analysis strategy was well-defined, which included all necessary analysis-level decisions and studies which were necessary before unblinding (soon analysis will be unblinded). Thus far, the analysis was performed with simulated events including several fit scenarios to test the fit setup and validate its behavior. The fit was a simultaneous binned fit to predict the number of $t\bar{t}+$jets background events in each **A**, **B**, **C**, **D** region for a given N_{jets} bin. The analysis has included the systematics related to the $t\bar{t}+$jets, QCD multi-jet, and minor background processes. The systematics for the signal processes have not been included yet. Also, QCD multi-jet systematics (not expected to be dominant) are still to be finalized. However, the simulation-based fit results and expected limits show that the analysis setup works well.

For the further steps, the fit will be performed again with real data and all systematics included for the latest results of the Stealth SY$\bar{Y}$ model. Also, analysis will be performed for the RPV model.

REFERENCES

Ade, P. et al. (Sept. 2016). "Planck 2015 results. XIII. Cosmological parameters". In: 594, A13. doi: 10.1051/0004-6361/201525830. url: https://doi.org/10.1051%2F0004-6361%2F201525830.

Aitchison, I. and Hey, A. (2013). *Gauge Theories in Particle Physics: A Practical Introduction, Volume 1 : From Relativistic Quantum Mechanics to QED, Fourth Edition*. Taylor and Francis. 454p. isbn: 978-1-4665-1302-0, 978-0-429-18538-0, 978-1-4665-1299-3. doi: {10.1201/b13717}.

Aitchison, I. J. R. and Hey, A. J. G. (2013). *Non-Abelian Gauge Theories: QCD and The Electroweak Theory*. Taylor and Fancis. 518p. isbn: 9780429185397. doi: {10.1201/9781466513105}. url: https://library.oapen.org/handle/20.500.12657/50884.

Altarelli, Guido (2005). "The Standard Model of Particle Physics". doi: 10.48550/ARXIV.HEP-PH/0510281. url: https://arxiv.org/abs/hep-ph/0510281.

Barbier, R. et al. (Nov. 2005). "R-Parity-violating supersymmetry". In: *Physics Reports* 420.1-6, pp. 1–195. doi: 10.1016/j.physrep.2005.08.006. url: https://doi.org/10.1016%2Fj.physrep.2005.08.006.

Biino, Cristina (Feb. 2015). "The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections". In: *Journal of Physics: Conference Series* 587, p. 012001. doi: 10.1088/1742-6596/587/1/012001. url: https://doi.org/10.1088/1742-6596/587/1/012001.

Biswas, Ambalika, Anirban Kundu, and Poulami Mondal (Oct. 2020). "Hierarchy problem and dimension-six effective operators". In: *Physical Review D* 102.7. doi: 10.1103/physrevd.102.075022. url: https://doi.org/10.1103%2Fphysrevd.102.075022.

Bjorken, James D. and Stanley J. Brodsky (Mar. 1970). "Statistical Model for Electron-Positron Annihilation into Hadrons". In: *Phys. Rev. D* 1 (5), pp. 1416–1420. doi: 10.1103/PhysRevD.1.1416. url: https://link.aps.org/doi/10.1103/PhysRevD.1.1416.

Bols, E. et al. (Dec. 2020). "Jet flavour classification using DeepJet". In: *Journal of Instrumentation* 15.12, P12012–P12012. doi: 10.1088/1748-0221/15/12/p12012. url: https://doi.org/10.1088%2F1748-0221%2F15%2F12%2Fp12012.

Cabrera, Luis Andres (Feb. 2017). "Status and performance of the CMS muon system in Run2". In: p. 733. doi: 10.22323/1.282.0733.

Cacciari, Matteo, Gavin P Salam, and Gregory Soyez (Apr. 2008). "The anti-k_t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04, pp. 063–063. doi: 10.1088/1126-6708/2008/04/063. url: https://doi.org/10.1088/1126-6708/2008/04/063.

Chatrchyan, S. et al. (Mar. 2012). "Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV". In: *Physics Letters B* 710.1, pp. 26–48. doi: 10.1016/j.physletb.2012.02.064. url: https://doi.org/10.1016%2Fj.physletb.2012.02.064.

Chatrchyan, S. et al. (2008). "The CMS Experiment at the CERN LHC". In: *JINST* 3, S08004. doi: 10.1088/1748-0221/3/08/S08004.

Chatrchyan, S. et al. (Oct. 2012). "Search for supersymmetry in hadronic final states using MT2 in pp collisions at $\sqrt{s} = 7$ TeV". In: *Journal of High Energy Physics* 2012.10. issn: 1029-8479. doi: 10.1007/jhep10(2012)018. url: http://dx.doi.org/10.1007/JHEP10(2012)018.

CMS Collaboration (June 2020a). “Cut Based Electron ID for Run 2”. url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2.

— (2020b). “Electron scale factors”. url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSLeptonSF#Electrons.

— (2020c). “Electron scale factors”. url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSLeptonSF#Muons.

— (May 2021). “LHC SUSY Cross Section”. url: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections#Cross_section_plots.

— (2022a). “Baseline muon selections for Run-II”. url: https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2.

— (Oct. 2022b). “Jet Identification for the 13 TeV UL data”. url: https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVUL.

— (Apr. 2022c). “RunII Data Sets”. url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/RA2b13TeVProduction.

— (Apr. 2022d). “RunII Ultra Legacy Data Sets”. url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/RA2b13TeVProduction#Run2ProductionV20_UltraLegacy.

— (Mar. 2023). “Run-2 UltraLegacy Datasets for Analysis”. url: https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmVRun2LegacyAnalysis.

CMS HOT Working Group (2019). “Resolved top quark tagging using neural networks”. url: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2018/273.

CMS JetMET Group (2019). “Machine learning-based identification of highly Lorentz-boosted hadronically decaying particles at the CMS experiment”. In: url: `https://cds.cern.ch/record/2683870`.

— (Apr. 2020). “DeepAK8 (V2) Working Points and Scale Factors for 13 TeV data”. url: `https://twiki.cern.ch/twiki/bin/view/CMS/DeepAK8Tagging2018WPsSFs`.

CMS TDR (2007). “CMS Technical Design Report Volume 2: Physics Performance”. In: *J. Phys. G* 34.CERN-LHCC-2006-021. CMS-TDR-8-2, p. 995. url: `http://cds.cern.ch/record/942733`.

Collaboration, CMS (Mar. 2010). “Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays”. In: *Journal of Instrumentation* 5.03, T03021–T03021. doi: `10.1088/1748-0221/5/03/t03021`. url: `https://doi.org/10.1088/1748-0221/5/03/t03021`.

Collaboration, The CMS (Oct. 2014). “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *Journal of Instrumentation* 9.10, P10009–P10009. doi: `10.1088/1748-0221/9/10/p10009`. url: `https://doi.org/10.1088/1748-0221/9/10/p10009`.

Cowan, Glen et al. (Feb. 2011). “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2. doi: `10.1140/epjc/s10052-011-1554-0`. url: `https://doi.org/10.1140%2Fepjc%2Fs10052-011-1554-0`.

Djouadi, A. (Apr. 2008). “The anatomy of electroweak symmetry breaking. II: The Higgs bosons in the Minimal Supersymmetric Model”. In: *Physics Reports* 459.1-6, pp. 1–241. doi: `10.1016/j.physrep.2007.10.005`. url: `https://doi.org/10.1016%2Fj.physrep.2007.10.005`.

Dordevic, Milos (Jan. 2018). “The CMS Particle Flow Algorithm”. In: *EPJ Web of Conferences* 191, p. 02016. doi: 10.1051/epjconf/201819102016.

Englert, F. and R. Brout (Aug. 1964). “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9), pp. 321–323. doi: 10.1103/PhysRevLett.13.321. url: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

Evans, Jared A. and Yevgeny Kats (2013). “LHC Coverage of RPV MSSM with Light Stops”. In: *JHEP* 04, p. 028. doi: 10.1007/JHEP04(2013)028. arXiv: 1209.0764 [hep-ph].

Fan, JiJi, Rebecca Krall, et al. (2016). “Stealth Supersymmetry Simplified”. In: *JHEP* 07, p. 016. doi: 10.1007/JHEP07(2016)016. arXiv: 1512.05781 [hep-ph].

Fan, JiJi, Matthew Reece, and Joshua T. Ruderman (2012). “A Stealth Supersymmetry Sampler”. In: *JHEP* 07, p. 196. doi: 10.1007/JHEP07(2012)196. arXiv: 1201.4875 [hep-ph].

Fox, Geoffrey C. and Stephen Wolfram (Dec. 1978). “Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes”. In: *Phys. Rev. Lett.* 41 (23), pp. 1581–1585. doi: 10.1103/PhysRevLett.41.1581. url: https://link.aps.org/doi/10.1103/PhysRevLett.41.1581.

Ganin, Yaroslav and Victor Lempitsky (2014). “Unsupervised Domain Adaptation by Backpropagation”. doi: 10.48550/ARXIV.1409.7495. url: https://arxiv.org/abs/1409.7495.

Gerwick, Erik et al. (Oct. 2012). “Scaling patterns for QCD jets”. In: *Journal of High Energy Physics* 2012.10. issn: 1029-8479. doi: 10.1007/jhep10(2012)162. url: http://dx.doi.org/10.1007/JHEP10(2012)162.

Glashow, Sheldon L. (1961). “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4, pp. 579–588. issn: 0029-5582. doi: https://doi.org/10.

1016/0029-5582(61)90469-2. url: https://www.sciencedirect.com/science/article/pii/0029558261904692.

Goldstone, Jeffrey, Abdus Salam, and Steven Weinberg (Aug. 1962). “Broken Symmetries”. In: *Phys. Rev.* 127 (3), pp. 965–970. doi: 10.1103/PhysRev.127.965. url: https://link.aps.org/doi/10.1103/PhysRev.127.965.

Grojean, C. (2015). “Beyond the standard Higgs after the 125 GeV Higgs discovery”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373.2032, p. 20140042. doi: 10.1098/rsta.2014.0042. url: https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2014.0042.

Gross, David J. and Frank Wilczek (Nov. 1973). “Asymptotically Free Gauge Theories. I”. In: *Phys. Rev. D* 8 (10), pp. 3633–3652. doi: 10.1103/PhysRevD.8.3633. url: https://link.aps.org/doi/10.1103/PhysRevD.8.3633.

Guthoff, Moritz (2016). “Instrumentation for beam radiation and luminosity measurement in the CMS experiment using novel detector technologies”. In: url: https://indico.cern.ch/event/391665/contributions/1827269/attachments/1229512/1801668/Guthoff_VCI16_Poster.pdf.

— (2017). “Instrumentation for beam radiation and luminosity measurement in the CMS experiment using novel detector technologies”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 845. Proceedings of the Vienna Conference on Instrumentation 2016, pp. 565–569. issn: 0168-9002. doi: https://doi.org/10.1016/j.nima.2016.06.028. url: https://www.sciencedirect.com/science/article/pii/S0168900216305708.

Higgs Combine Tool (2021). “Asymptotic Significance”. url: https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit.

Higgs, Peter W. (Oct. 1964). “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16), pp. 508–509. doi: 10.1103/PhysRevLett.13.508. url: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

Hoch, Michael (Dec. 2007). “Muon Chambers”. CMS Collection. url: http://cds.cern.ch/record/1274467.

Jeitler, Manfred (2013). “Upgrade of the trigger system of CMS”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 718. Proceedings of the 12th Pisa Meeting on Advanced Detectors, pp. 11–15. issn: 0168-9002. doi: https://doi.org/10.1016/j.nima.2012.08.091. url: https://www.sciencedirect.com/science/article/pii/S0168900212009941.

Jet algorithms performance in 13 TeV data (2017). Tech. rep. Geneva: CERN. url: http://cds.cern.ch/record/2256875.

Karimäki, V et al. (1997). *The CMS tracker system project: Technical Design Report*. Technical design report. CMS. Geneva: CERN. url: http://cds.cern.ch/record/368412.

Kasieczka, Gregor et al. (Feb. 2021). “Automating the ABCD method with machine learning”. In: *Physical Review D* 103.3. issn: 2470-0029. doi: 10.1103/physrevd.103.035021. url: http://dx.doi.org/10.1103/PhysRevD.103.035021.

Klinkhamer, F. (Feb. 2013). “Neutrino mass and the Standard Model”. In: *Modern Physics Letters A* 28.05, p. 1350010. doi: 10.1142/s0217732313500107. url: https://doi.org/10.1142%2Fs0217732313500107.

Larkoski, Andrew J. et al. (May 2014). “Soft drop”. In: *Journal of High Energy Physics* 2014.5. doi: 10.1007/jhep05(2014)146. url: https://doi.org/10.1007%2Fjhep05%282014%29146.

Lopienska, Ewa (Feb. 2022). “The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022”. In: General Photo. url: https://cds.cern.ch/record/2800984.

Mans, J et al. (Sept. 2012). *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*. Tech. rep. url: https://cds.cern.ch/record/1481837.

Martin, Stephen P. (July 1998). “A Supersymmetry Primer”. In: *Perspectives on Supersymmetry*. World Scientific, pp. 1–98. doi: 10.1142/9789812839657_0001. url: https://doi.org/10.1142%2F9789812839657_0001.

Nastase, Horatiu (2011). “Introduction to supergravity”. doi: 10.48550/ARXIV.1112.3502. url: https://arxiv.org/abs/1112.3502.

Neutelings, I. (June 2017). In: url: https://tikz.net/axis3d_cms/.

OPERA Collaboration (July 2014). “Observation of tau neutrino appearance in the CNGS beam with the OPERA experiment”. In: *Progress of Theoretical and Experimental Physics* 2014. doi: 10.1093/ptep/ptu132.

“Public CMS Luminosity Information” (Mar. 2022). In: url: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Multi_year_plots.

Python (2023). “Create production-grade machine learning models with TensorFlow”. url: https://www.tensorflow.org/.

RooStat (2022). “RooFit”. url: https://twiki.cern.ch/twiki/bin/view/RooStats.

Sakharov, A D (Jan. 1967). “Violatin of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe.” In: *JETP Lett. (USSR)(Engl. Transl.), 5: 24-7(Jan. 1, 1967)*. url: https://www.osti.gov/biblio/4449128.

Sakuma, Tai (May 2016). "3D SketchUp images of the CMS detector (120918)". In: url: https://cds.cern.ch/record/2628527.

Salam, A. and J.C. Ward (1964). "Electromagnetic and weak interactions". In: *Physics Letters* 13.2, pp. 168–171. issn: 0031-9163. doi: https://doi.org/10.1016/0031-9163(64)90711-5. url: https://www.sciencedirect.com/science/article/pii/0031916364907115.

Seo, H. (2013). "RPC hit contribution to CMS muon reconstruction at LHC". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 718. Proceedings of the 12th Pisa Meeting on Advanced Detectors, pp. 437–439. issn: 0168-9002. doi: https://doi.org/10.1016/j.nima.2012.08.056. url: https://www.sciencedirect.com/science/article/pii/S0168900212009412.

Siddireddy, Prasanna Kumar (June 2018). *The CMS ECAL Trigger and DAQ system: electronics auto-recovery and monitoring*. Tech. rep. 7 pages, 6 figues, 1 table, conference record. arXiv: 1806.09136. url: http://cds.cern.ch/record/2630618.

Sirunyan, A. M. et al. (Oct. 2017a). "Particle-flow reconstruction and global event description with the CMS detector". In: *Journal of Instrumentation* 12.10, P10003–P10003. doi: 10.1088/1748-0221/12/10/p10003. url: https://doi.org/10.1088%2F1748-0221%2F12%2F10%2Fp10003.

— (Jan. 2017b). "The CMS trigger system". In: *Journal of Instrumentation* 12.01, P01020–P01020. doi: 10.1088/1748-0221/12/01/p01020. url: https://doi.org/10.1088%2F1748-0221%2F12%2F01%2Fp01020.

— (June 2018). "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$=13 TeV". In: *Journal of Instrumentation*

13.06. doi: 10.1088/1748-0221/13/06/p06015. url: https://doi.org/10.1088%2F1748-0221%2F13%2F06%2Fp06015.

— (May 2020). "Calibration of the CMS hadron calorimeters using proton-proton collision data at √s=13 TeV". In: *Journal of Instrumentation* 15.05, P05002–P05002. doi: 10.1088/1748-0221/15/05/p05002. url: https://doi.org/10.1088/1748-0221/15/05/p05002.

— (Sept. 2021a). "Search for top squark production in fully hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 104 (5), p. 052001. doi: 10.1103/PhysRevD.104.052001. url: https://link.aps.org/doi/10.1103/PhysRevD.104.052001.

— (Aug. 2021b). "Search for top squarks in final states with two top quarks and several light-flavor jets in proton-proton collisions at √s=13 TeV". In: *Physical Review D* 104.3. issn: 2470-0029. doi: 10.1103/physrevd.104.032006. url: http://dx.doi.org/10.1103/PhysRevD.104.032006.

Székely, Gábor J. and Maria L. Rizzo (2014). "Partial distance correlation with methods for dissimilarities". In: *The Annals of Statistics* 42.6, pp. 2382–2412. doi: 10.1214/14-AOS1255. url: https://doi.org/10.1214/14-AOS1255.

T2K Collaboration (2014). "Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-axis Beam". In: doi: 10.48550/ARXIV.1403.1532. url: https://arxiv.org/abs/1403.1532.

Tanabashi, M. et al. (Aug. 2018). "Review of Particle Physics". In: *Phys. Rev. D* 98 (3), p. 030001. doi: 10.1103/PhysRevD.98.030001. url: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

The KamLAND Collaboration (June 2008). "Precision Measurement of Neutrino Oscillation Parameters with KamLAND". In: *Phys. Rev. Lett.* 100 (22), p. 221803.

doi: 10.1103/PhysRevLett.100.221803. url: https://link.aps.org/doi/10.1103/PhysRevLett.100.221803.

The Tracker Group Of The CMS Collaboration (2020). “The CMS Phase-1 Pixel Detector Upgrade”. doi: 10.48550/ARXIV.2012.14304. url: https://arxiv.org/abs/2012.14304.

Thomson, M. (2013). *Modern Particle Physics*. Cambridge University Press. 570p. isbn: 978-1107034266. doi: {10.1017/CBO9781139525367}.

Trocino, Daniele (June 2014). “The CMS High Level Trigger”. In: *Journal of Physics: Conference Series* 513.1, p. 012036. doi: 10.1088/1742-6596/513/1/012036. url: https://doi.org/10.1088/1742-6596/513/1/012036.

Veszpremi, V (Mar. 2014). “Operation and performance of the CMS tracker”. In: *Journal of Instrumentation* 9.03, pp. C03005–C03005. doi: 10.1088/1748-0221/9/03/c03005. url: https://doi.org/10.1088%2F1748-0221%2F9%2F03%2Fc03005.

Virdee, T. S. (2016). “Beyond the standard model of particle physics”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2075, p. 20150259. doi: 10.1098/rsta.2015.0259. url: https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2015.0259.

Wadud, Mohammad Abrar (2019). “The CMS ECAL Calibration and Monitoring”. url: https://inspirehep.net/files/2c22c843e51818174012fef627877dca.

Weinberg, Steven (Nov. 1967). “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21), pp. 1264–1266. doi: 10.1103/PhysRevLett.19.1264. url: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

Workman, R. L. et al. (Aug. 2022). “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2022.8. 083C01. issn: 2050-3911. doi: 10.1093/ptep/ptac097. url: https://doi.org/10.1093/ptep/ptac097.

CURRICULUM VITAE

Semra TÜRKÇAPAR graduated from Dicle University with a BSc in Physics in 2008. In fall 2009, she participated in the English Language Program at Çukurova University Foreign Languages Teaching Research and Application Center—got an achievement certificate in 2010. She got her MSc degree in High Energy Physics at Çukurova University in early 2014. During her MSc term, she started working on the CMS HCAL Phase-1 upgrades at CERN. In 2012 and 2013, she tested the the Silicon Photo Multipliers (SiPMs) for the HCAL Outer (HO) detector and also studied stability analysis to monitor temperature, current, led and pedestal values of SiPMs under the supervision of Dr. Benjamin Lutz. In 2013, LS1 at CERN, she was also involved with assembling and testing the 4-anode Photo Multiplier Tubes (PMTs) for the HCAL Forward (HF) detector. Additionally, she studied pedestal analysis for HF to monitor pedestal values of PMTs. In fall 2014, she has started her Ph.D. in Particle Physics at Çukurova University. During her Ph.D. term, she had direct involvement in various aspects of detector development for CMS HCAL Phase-1 upgrades including leading the testing and integrating of the calibration units for HCAL Barrel (HB) detector, installating and commissionining of the HF front-end electronics. Additionally, she took several shifts including the HCAL Detector On-Call (DOC) and CMS BRIL shifts, which are central ones during the CMS data taking. She was funded a couple of times by the U.S. CMS group to work on different parts of the Phase-1 upgrades. In addition to this, she was invited to Fermilab as a visitor student of the CMS LHC Physiscs Center (LPC) Guest and Visitor program in 2019. Since then, she also has been working on supersymmetry (SUSY) analyses. Her main dissertation research is based on one of these SUSY analyses, which is ”Search for Stealth $SY\bar{Y}$ Top Squark Decays with an Automated ABCD Method using a Double DisCo Neural Network” under the supervision of Scientist Dr. Jim Hirschauer.

APPENDIX

A Hadronic Top Tagging

A1. Number of Tops Populations

It is studied how many resolved and merged tops appear in events that pass the fully-hadronic event selection. The DeepAK8 and DeepResolved algorithms are run in tandum, and an event passes the selection if there are some combination of at least two tops tagged by either of the algorithms. To understand where merged top decays (tagged with the DeepAK8 algorithm) are more important than resolved top decays (tagged with the DeepResolved algorithm)—and vice versa—events are categorized by all possible combinations of merged and resolved top tags i.e. 0 merged, 2 resolved; 1 merged, 1 resolved; 2 merged, 0 resolved, etc.

Figure A1. contains 2D histograms for the main $t\bar{t}$ + jets background, a background with no tops—QCD multi-jet—and the Stealth SY$\bar{Y}$ signal for select stop mass points, where each bin corresponds to a different combination of merged and resolved top tags. The histograms are normalized so the number reported in each colored bin is the fraction of all events in the histogram. It is seen that for low stop masses 300-500 GeV, a majority of the events contain at least two resolved tops. As the top squark mass increases, the top decay products become more boosted and a larger fraction of the tops are tagged by the DeepAK8 algorithm. At the maximum top squark mass of 1400 GeV, a vast majority of events contain at least one merged top and a simple majority contain at least two. This observation demonstrates that it is important to utilize both the DeepAK8 and DeepResolved algorithms to tag tops for the range of considered top squark masses. A loss of signal efficiency would be incurred at the low (high) top squark mass range if either of the taggers were excluded.

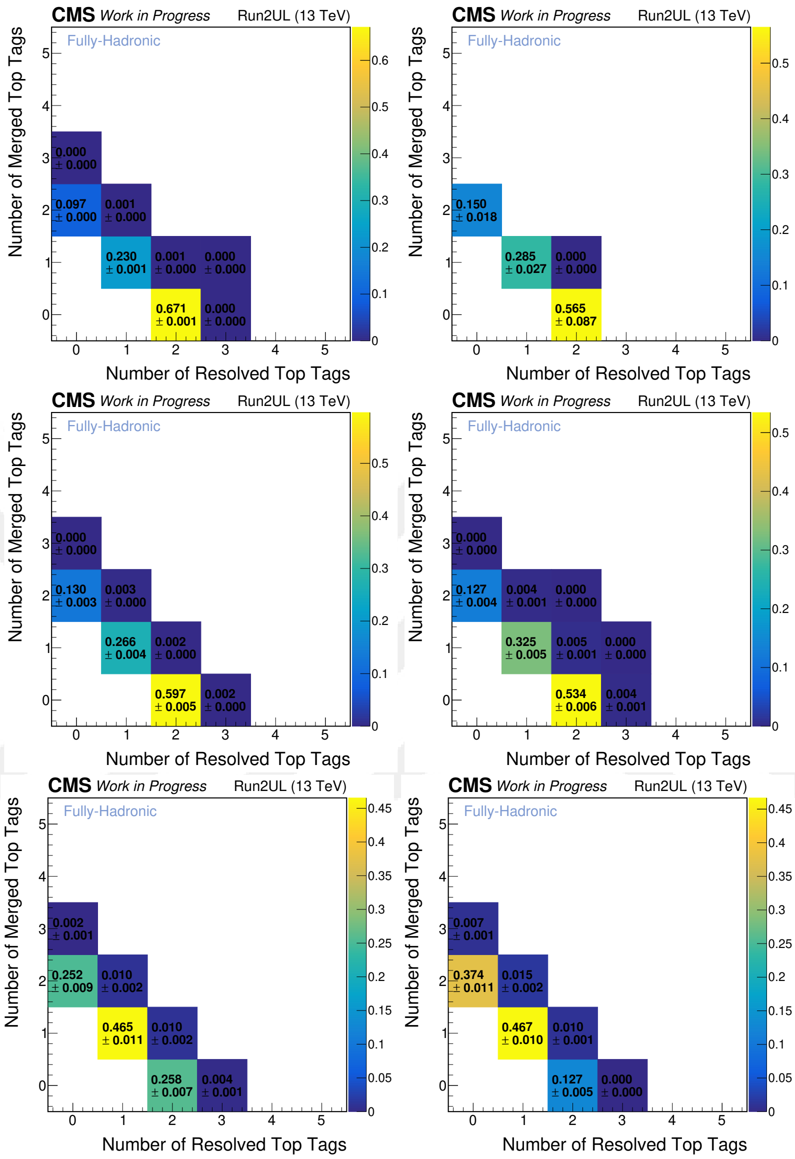


Figure A1. Plots of the number of resolved top tags versus number of merged top tags for $t\bar{t}+$jets (top-left), QCD multi-jet (top-right), and Stealth SY$\bar{Y}$ signal events that pass the fully-hadronic event selection. The selected top squark mass points are 350 (middle left), 550 (middle right), 850 (bottom left), and 1150 (bottom right) GeV. Each colored square represents a possible combination of resolved and merged top tags, while the number reports the fraction of events for the specific top squark mass

B Event Reconstruction and Selection

B1. Simulated Background and Signal Samples

Table B1. Simulated signal samples and related cross section for each sample are listed for this analysis. Cross sections are the same for all 2016preVFP, 2016postVFP, 2017 and 2018

Process	Full data set name	Cross section (pb)
RPV	RPV_2t6j_mStop-300_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	10.0
	RPV_2t6j_mStop-350_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	4.43
	RPV_2t6j_mStop-400_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	2.15
	RPV_2t6j_mStop-450_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	1.11
	RPV_2t6j_mStop-500_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.609
	RPV_2t6j_mStop-550_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.347
	RPV_2t6j_mStop-600_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.205
	RPV_2t6j_mStop-650_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.125
	RPV_2t6j_mStop-700_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0783
	RPV_2t6j_mStop-750_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.05
	RPV_2t6j_mStop-800_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0326
	RPV_2t6j_mStop-850_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0216
	RPV_2t6j_mStop-900_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0145
	RPV_2t6j_mStop-950_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00991
	RPV_2t6j_mStop-1000_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00683
	RPV_2t6j_mStop-1050_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00476
	RPV_2t6j_mStop-1100_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00335
	RPV_2t6j_mStop-1150_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00238
	RPV_2t6j_mStop-1200_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0017
	RPV_2t6j_mStop-1250_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00122
	RPV_2t6j_mStop-1300_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.000887
	RPV_2t6j_mStop-1350_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.000646
	RPV_2t6j_mStop-1400_mN1-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.000473
Stealth $SY\bar{Y}$	StealthSYY_2t6j_mStop-300_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	10.0
	StealthSYY_2t6j_mStop-350_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	4.43
	StealthSYY_2t6j_mStop-400_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	2.15
	StealthSYY_2t6j_mStop-450_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	1.11
	StealthSYY_2t6j_mStop-500_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.609
	StealthSYY_2t6j_mStop-550_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.347
	StealthSYY_2t6j_mStop-600_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.205
	StealthSYY_2t6j_mStop-650_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.125
	StealthSYY_2t6j_mStop-700_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0783
	StealthSYY_2t6j_mStop-750_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.05
	StealthSYY_2t6j_mStop-800_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0326
	StealthSYY_2t6j_mStop-850_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0216
	StealthSYY_2t6j_mStop-900_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0145
	StealthSYY_2t6j_mStop-950_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00991
	StealthSYY_2t6j_mStop-1000_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00683
	StealthSYY_2t6j_mStop-1050_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00476
	StealthSYY_2t6j_mStop-1100_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00335
	StealthSYY_2t6j_mStop-1150_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00238
	StealthSYY_2t6j_mStop-1200_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.0017
	StealthSYY_2t6j_mStop-1250_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.00122
	StealthSYY_2t6j_mStop-1300_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.000887
	StealthSYY_2t6j_mStop-1350_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.000646
	StealthSYY_2t6j_mStop-1400_mSo-100_TuneCP5_13TeV-madgraphMLM-pythia8	0.000473

Table B2. Simulated background samples and the production cross section for each sample are listed for this analysis. Due to identical physics simulation settings for each year, the cross section for a given sample is the same for 2016preVFP, 2016postVFP, 2017 and 2018

Process	Full data set name	Cross section (pb)
$t\bar{t}$ + jets	TTToHadronic_TuneCP5_13TeV-powheg-pythia8	380.1326
	TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	364.3284
	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	87.339
QCD	QCD_HT50to100_TuneCP5_PSWeights_13TeV-madgraph-pythia8	186100000.0
	QCD_HT100to200_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	23630000.0
	QCD_HT200to300_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	1554000.0
	QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	325000.0
	QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraph-pythia8	30350.0
	QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraph-pythia8	6403.0
	QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	1117.0
	QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	108.4
	QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	21.93
$t\bar{t}$ + X	TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.19580072
	TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.40499928
	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2439
	TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.5113
	ttHJetTobb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	0.293693
	ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	0.215307
	TTHH_TuneCP5_13TeV-madgraph-pythia8	0.0006683
	TTTT_TuneCP5_13TeV-amcatnlo-pythia8	0.01197
	TTTW_TuneCP5_13TeV-madgraph-pythia8	0.0007299
	TTWH_TuneCP5_13TeV-madgraph-pythia8	0.001143
	TTWW_TuneCP5_13TeV-madgraph-pythia8	0.007003
	TTWZ_TuneCP5_13TeV-madgraph-pythia8	0.002453
	TTZH_TuneCP5_13TeV-madgraph-pythia8	0.001134
	TTZZ_TuneCP5_13TeV-madgraph-pythia8	0.001389
	TTTJ_TuneCP5_13TeV-madgraph-pythia8	0.000401
Other	WW_TuneCP5_13TeV-pythia8	118.7
	WZ_TuneCP5_13TeV-pythia8	47.13
	ZZ_TuneCP5_13TeV-pythia8	16.523
	WWG_TuneCP5_13TeV-amcatnlo-pythia8	0.3369
	WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.2158
	WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.1707
	WZG_TuneCP5_13TeV-amcatnlo-pythia8	0.07876
	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05709
	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01476
	ST_s-channel_4f_InclusiveDecays_TuneCP5_13TeV-amcatnlo-pythia8	10.32
	ST_t-channel_antitop_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	71.74
	ST_t-channel_top_5f_InclusiveDecays_TuneCP5_13TeV-powheg-pythia8	119.7
	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	32.51
	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	32.45
	tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8	0.07561
	DYJetsToLL_M-50_HT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	139.9
	DYJetsToLL_M-50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	140.3
	DYJetsToLL_M-50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	38.37
	DYJetsToLL_M-50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	5.212
	DYJetsToLL_M-50_HT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.267
	DYJetsToLL_M-50_HT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.5678
	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.1332
	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.002988
	DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8	6077.22
	WJetsToLNu_HT-70To100_TuneCP5_13TeV-madgraphMLM-pythia8	1264.0
	WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1256.0
	WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	335.5
	WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	45.25
	WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	10.97
	WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	4.933
	WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.16
	WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.02627
	WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8	61526.7
	WJetsToQQ_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	2549.0
	WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	276.5
	WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	59.25
	WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8	28.75

C Double DisCo Neural Network

C1. Choosing the Input Variables

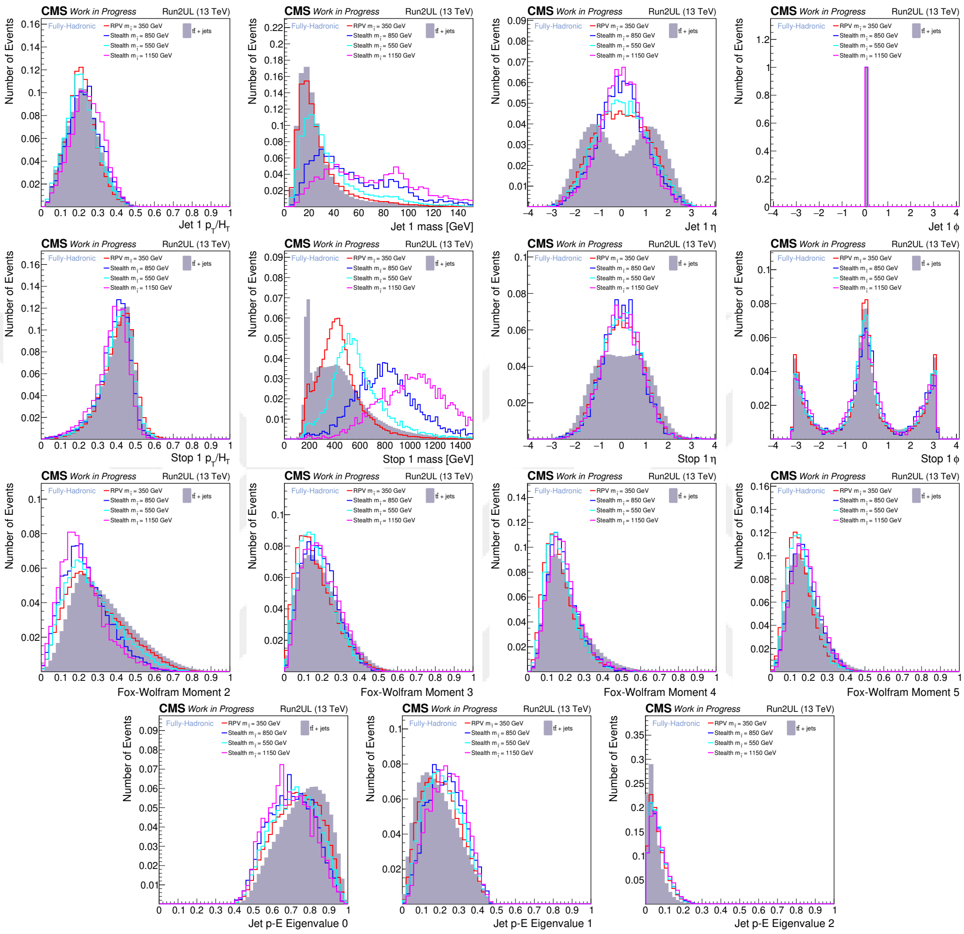


Figure C1. For the fully-hadronic channel, a sample of input variables used in the final training of the neural network. The top row is the 4-vector of the leading p_{T} jet. N.B. This jet is rotated to $\phi = 0$ and all other objects in an event rotated accordingly. To avoid passing a delta function to the network, the jet 1 ϕ is not passed as an input. The top middle row is the 4-vector of the leading p_{T} top squark object. The bottom middle row shows Fox-Wolfram moments 2 through 5. The bottom row shows the energy-momentum tensor eigenvalues

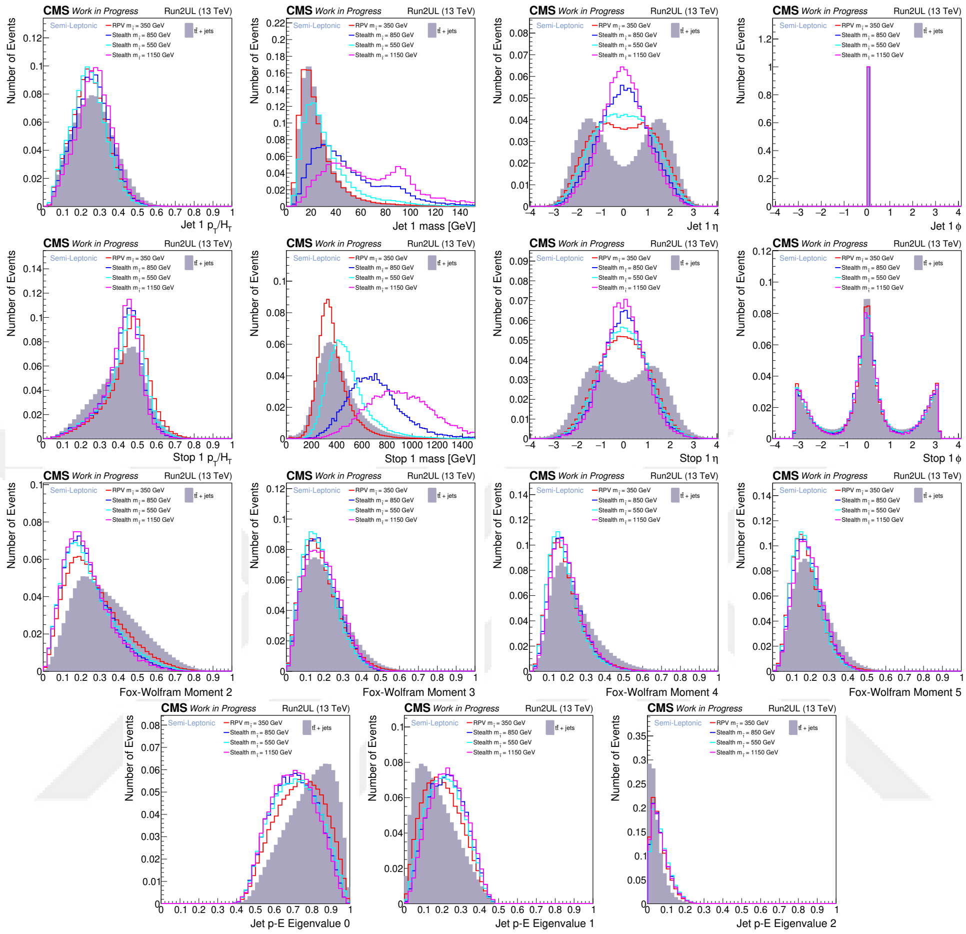


Figure C2. For the semi-leptonic channel, a sample of input variables used in the final training of the neural network. The top row is the 4-vector of the leading p_T jet. N.B. This jet is rotated to $\phi = 0$ and all other objects in an event rotated accordingly. To avoid passing a delta function to the network, the jet 1 ϕ is not passed as an input. The top middle row is the 4-vector of the leading p_T top squark object. The bottom middle row shows Fox-Wolfram moments 2 through 5. The bottom row shows the energy-momentum tensor eigenvalues

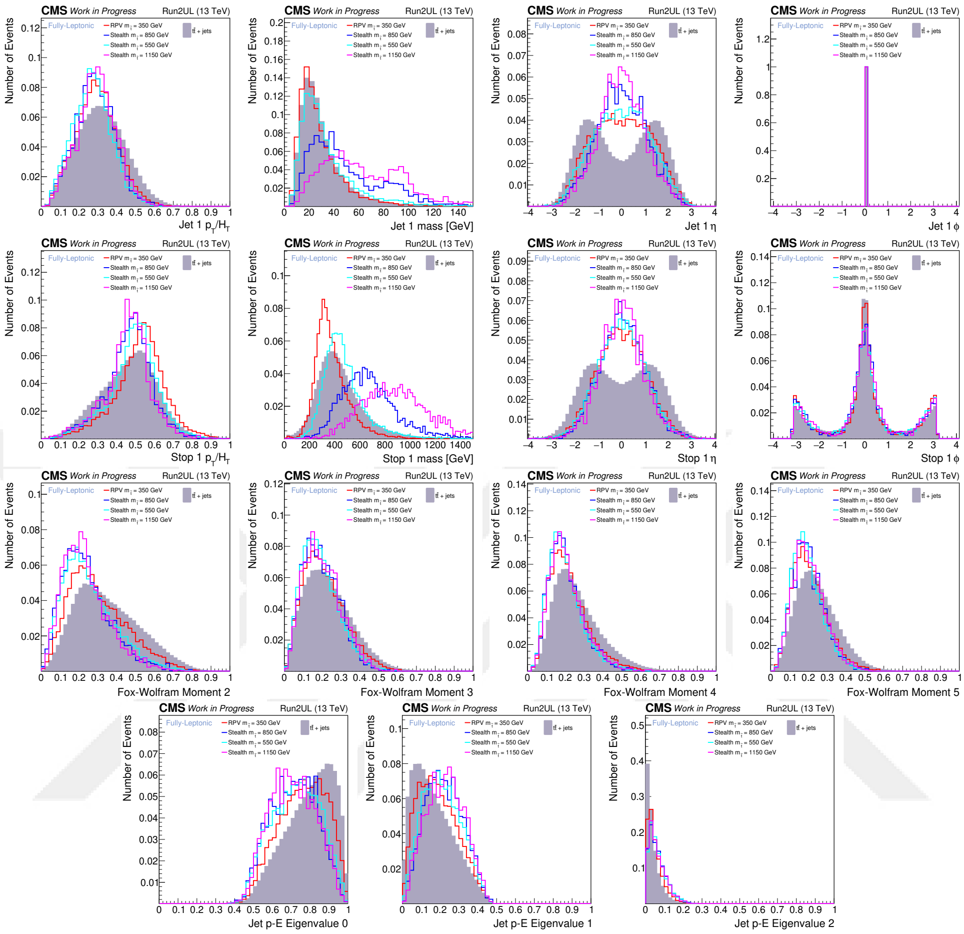


Figure C3. For the fully-leptonic channel, a sample of input variables used in the final training of the neural network. The top row is the 4-vector of the leading p_{T} jet. N.B. This jet is rotated to $\phi = 0$ and all other objects in an event rotated accordingly. To avoid passing a delta function to the network, the jet 1 ϕ is not passed as an input. The top middle row is the 4-vector of the leading p_{T} top squark object. The bottom middle row shows Fox-Wolfram moments 2 through 5. The bottom row shows the energy-momentum tensor eigenvalues

D Validating the ABCD Method

D1. Defining the ABCD Regions

D1.1. Initial Study for Optimizing the ABCD Bin Edges

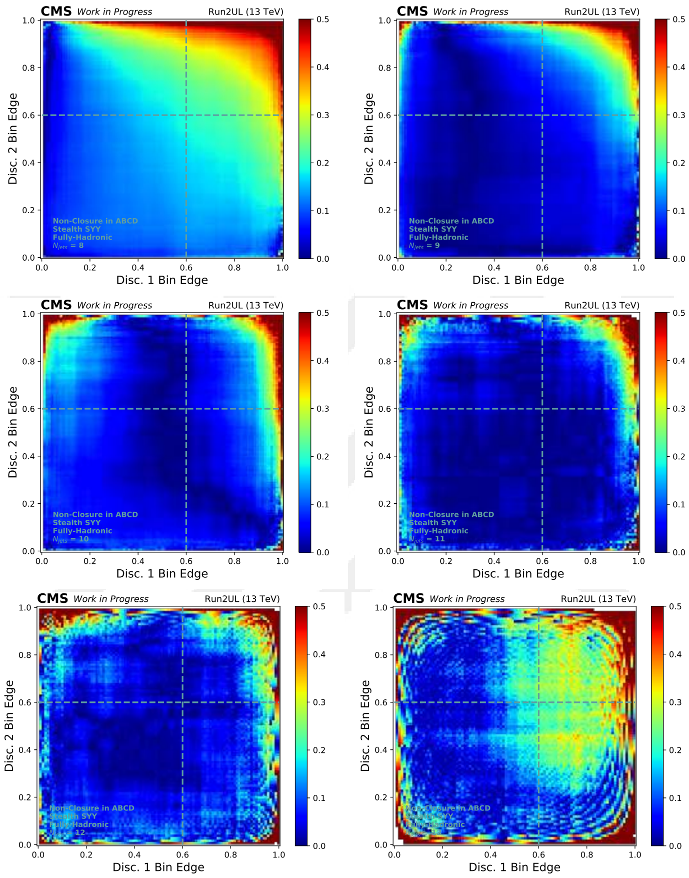


Figure D1. For the fully-hadronic channel, the non-closure as a function of "non-optimized" ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 8,9,10,11,12,≥13 bins (left to right, top to bottom). Overall non-closure is about ~ 0 as expected by the sense of perfect ABCD closure in the lower N_{jets} bins

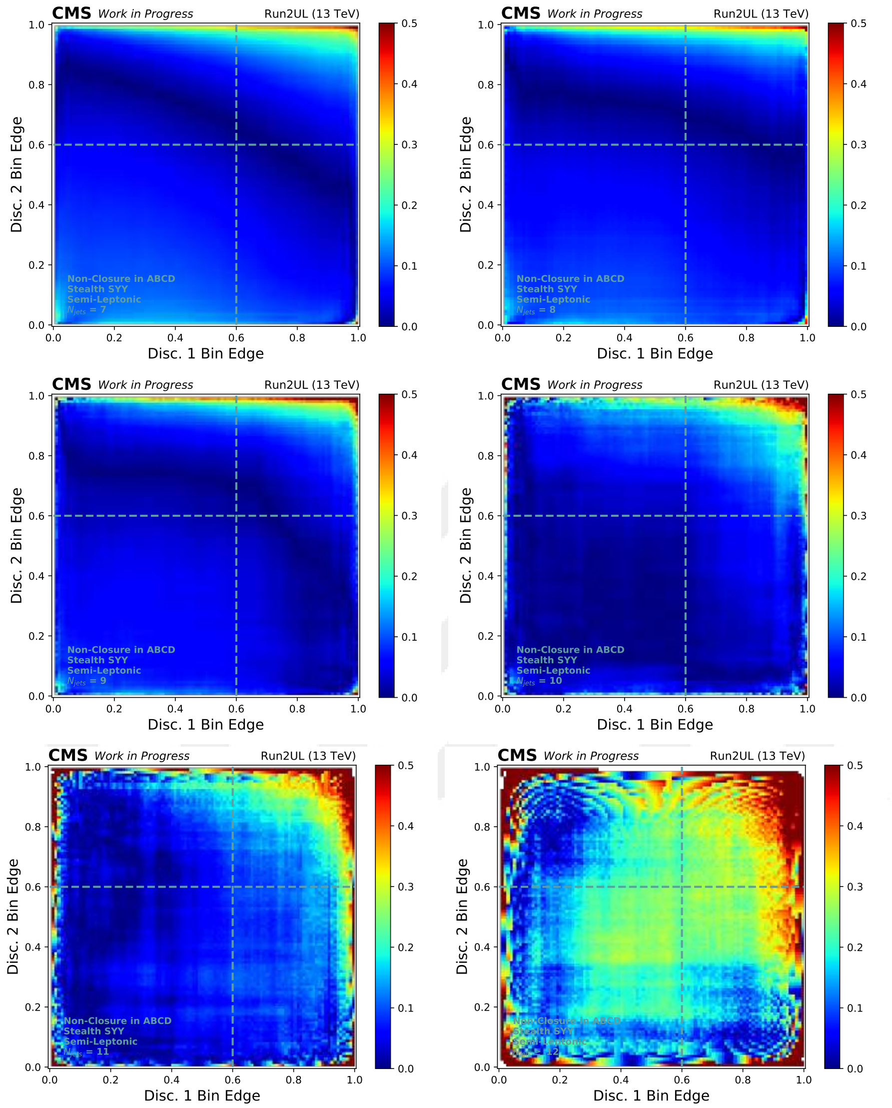


Figure D2. For the semi-leptonic channel, the non-closure as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+$jets. N_{jets} = 7,8,9,10,11,$\geq$12 bins (left to right, top to bottom). Overall non-closure is about ~ 0 as expected by the sense of perfect ABCD closure in the lower N_{jets} bins

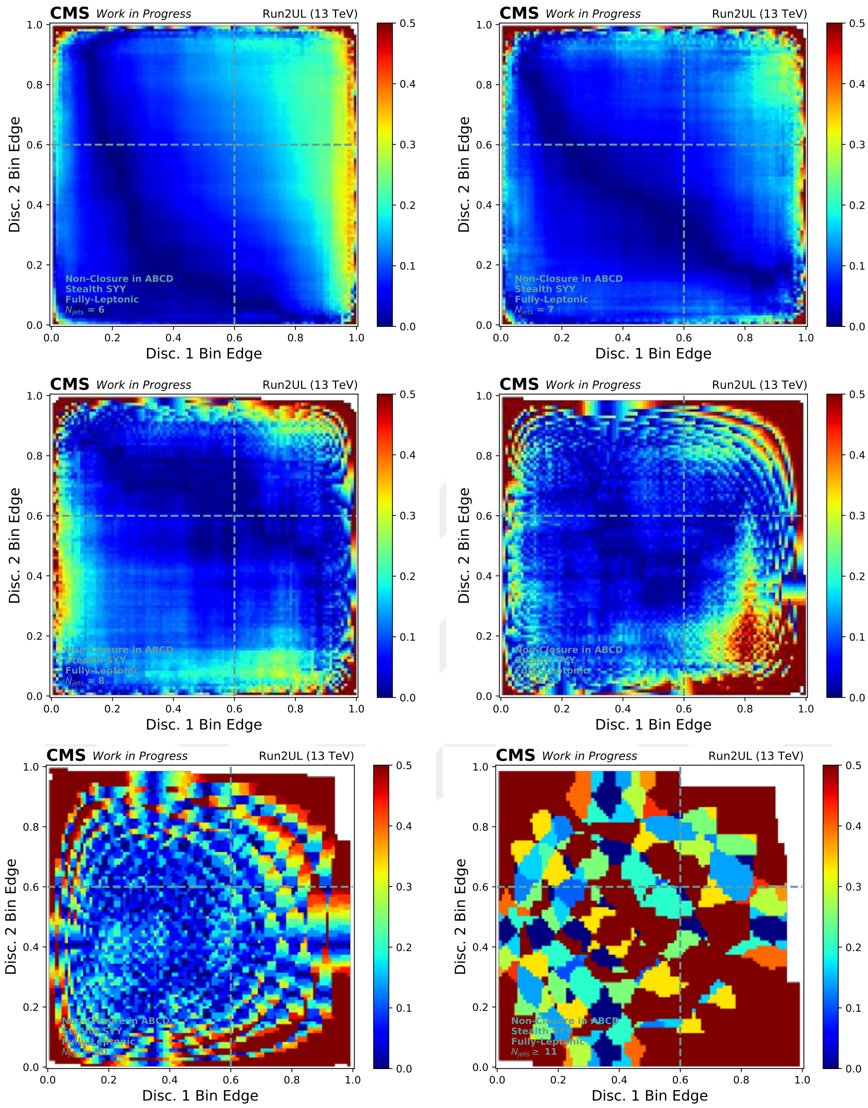


Figure D3. For the fully-leptonic channel, the non-closure as a function of "non-optimized" ABCD bin edges based on $t\bar{t}+\text{jets}$. N_{jets} = 6,7,8,9,10,≥11 bins (left to right, top to bottom). Overall non-closure is about ~ 0 as expected by the sense of perfect ABCD closure in the lower N_{jets} bins

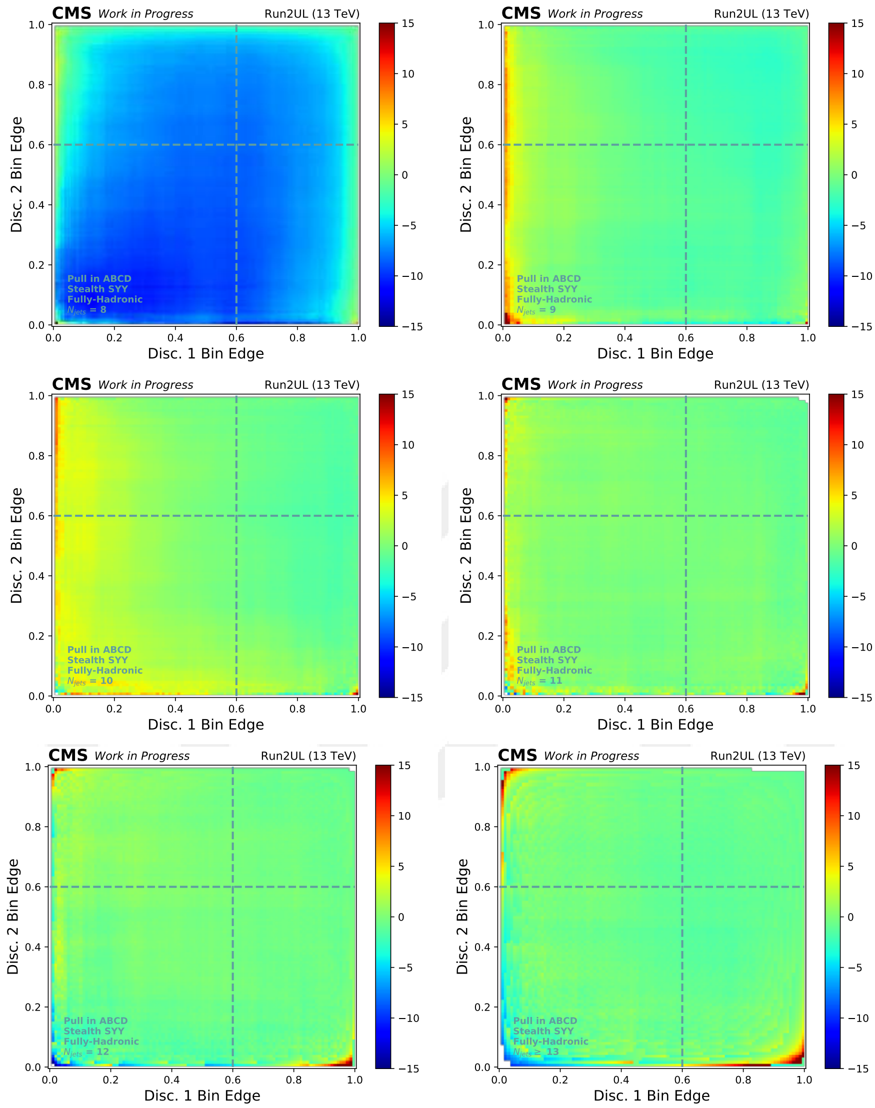


Figure D4. For the fully-hadronic channel, pull as a function of “non-optimized” ABCD bin edges based on $t\bar{t}+\text{jets}$. N_{jets} = 8,9,10,11,12,≥13 bins (left to righti, top to bottom)

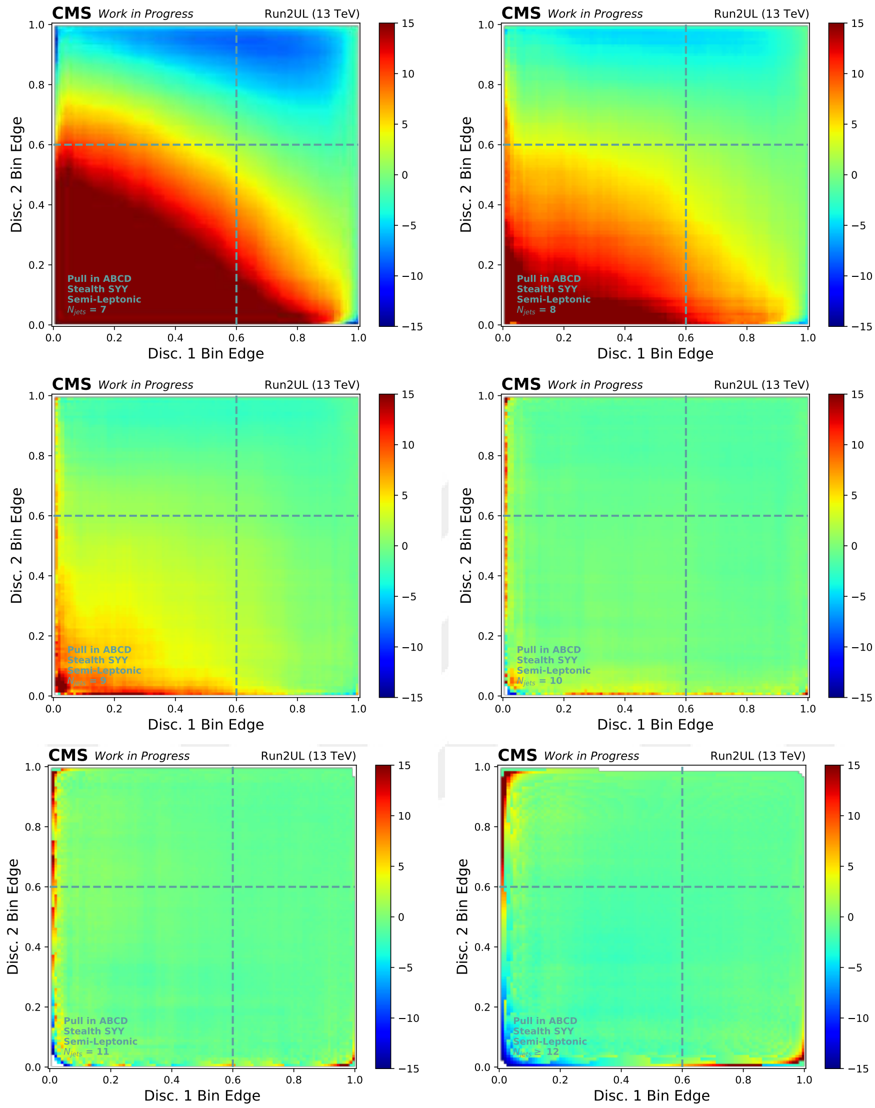


Figure D5. For the semi-leptonic channel, pull as a function of "non-optimized" ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 7,8,9,10,11,≥12 bins (left to right, top to bottom)

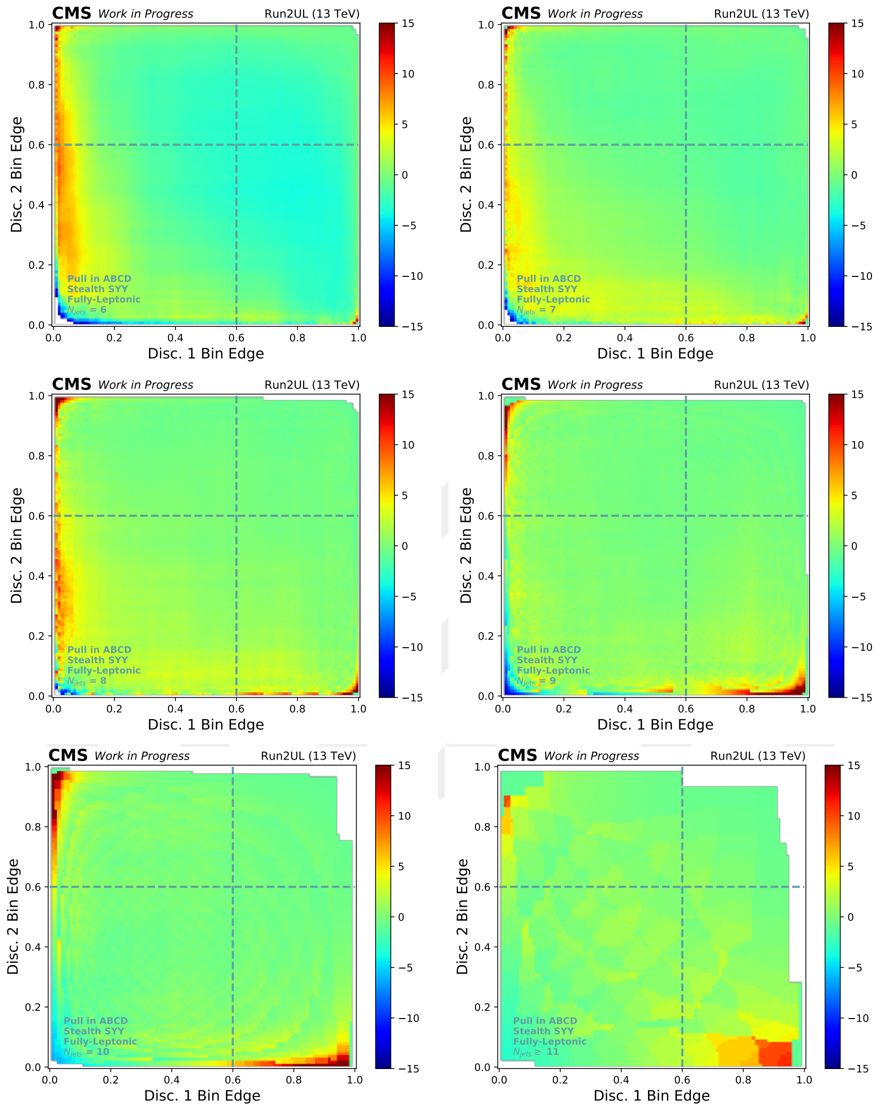


Figure D6. For the fully-leptonic channel, pull as a function of “non-optimized” ABCD bin edges based on $t\bar{t}$ + jets. N_{jets} = 6,7,8,9,10,≥11 bins (left to right, top to bottom)

D1.2. Optimizing the ABCD Bin Edges

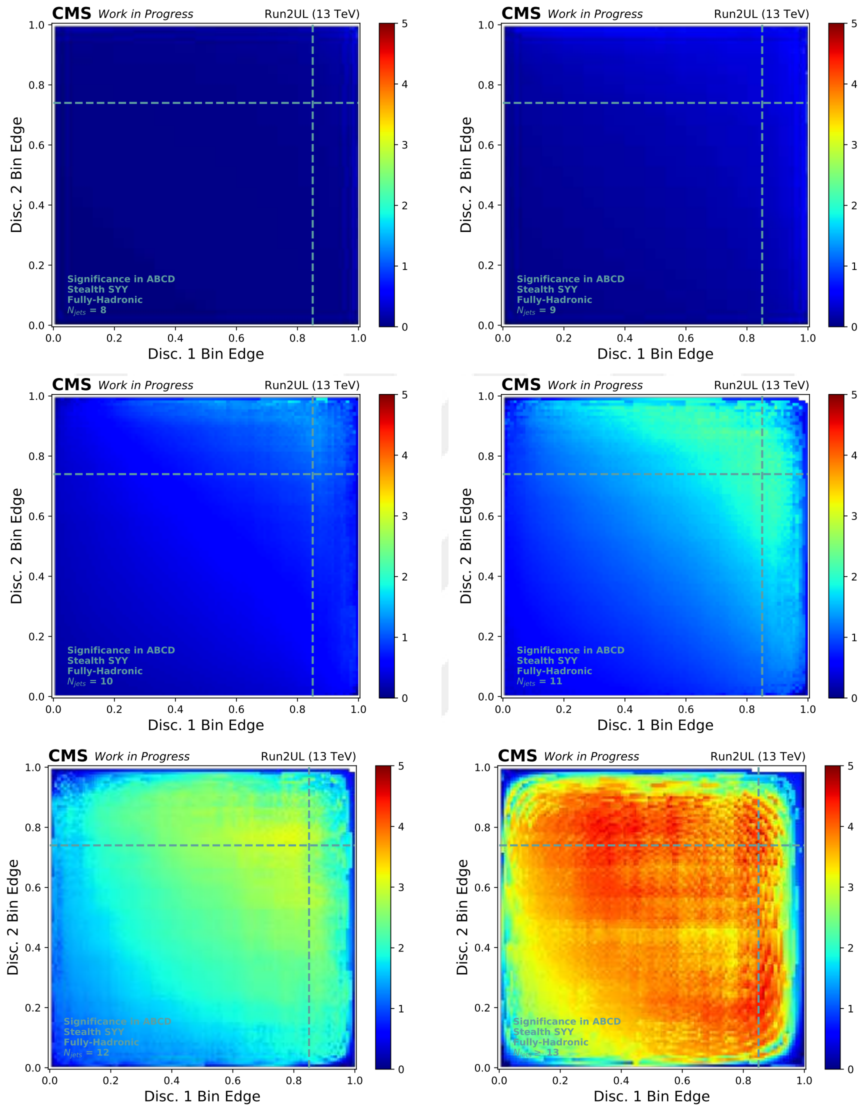


Figure D7. For the fully-hadronic channel, the significance as a function of "optimized" ABCD bin edges. $N_{\text{jets}} = 8,9,10,11,12 \geq 13$ bins (left to right, top to bottom)

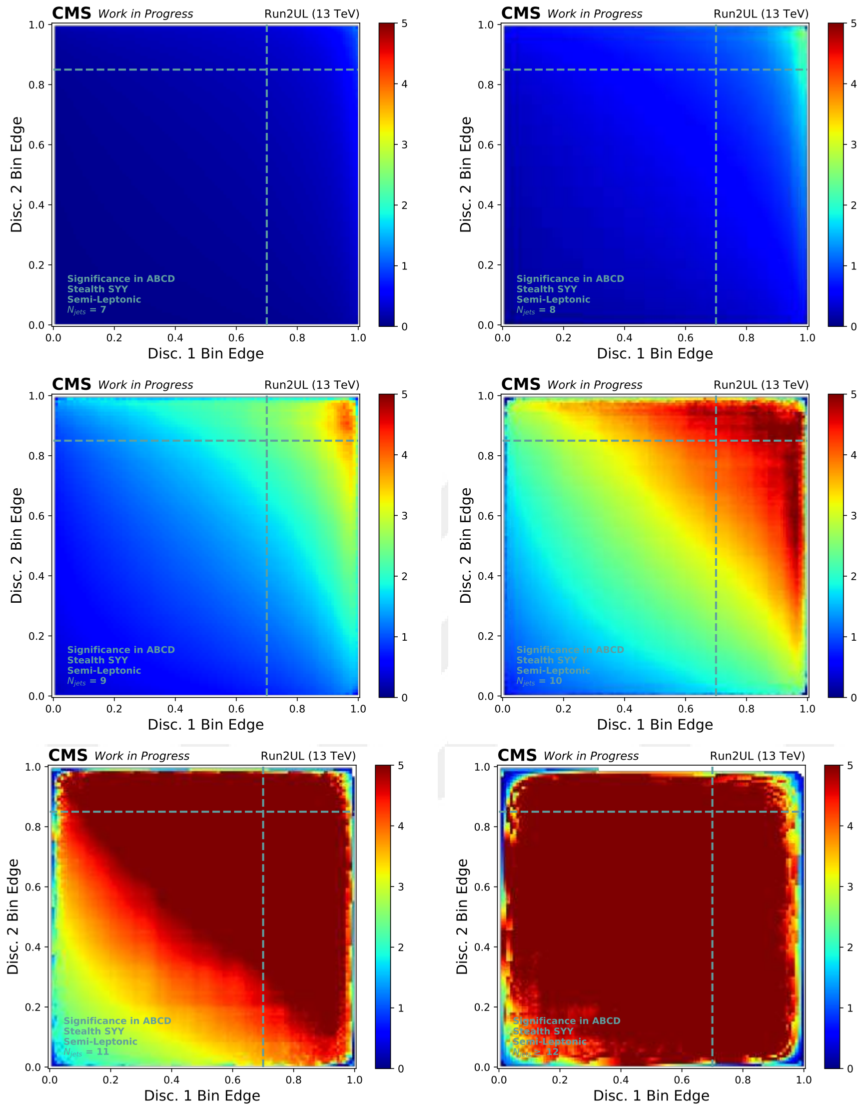


Figure D8. For the semi-leptonic channel, the significance as a function of "optimized" ABCD bin edges. N_{jets} = 7,8,9,10,11,$\geq$12 bins (left to right, top to bottom)

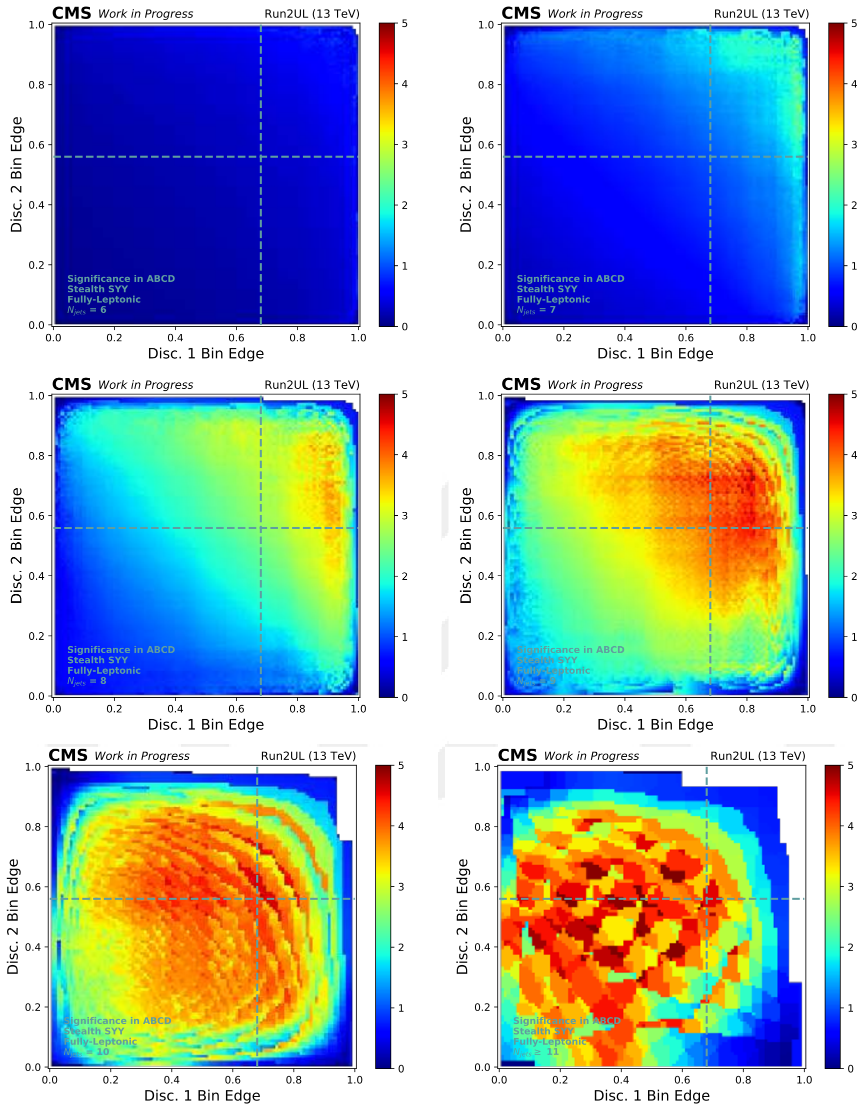


Figure D9. For the fully-leptonic channel, the significance as a function of “optimized” ABCD bin edges. $N_{\text{jets}} = 6,7,8,9,10,\geq 11$ bins (left to right, top to bottom)

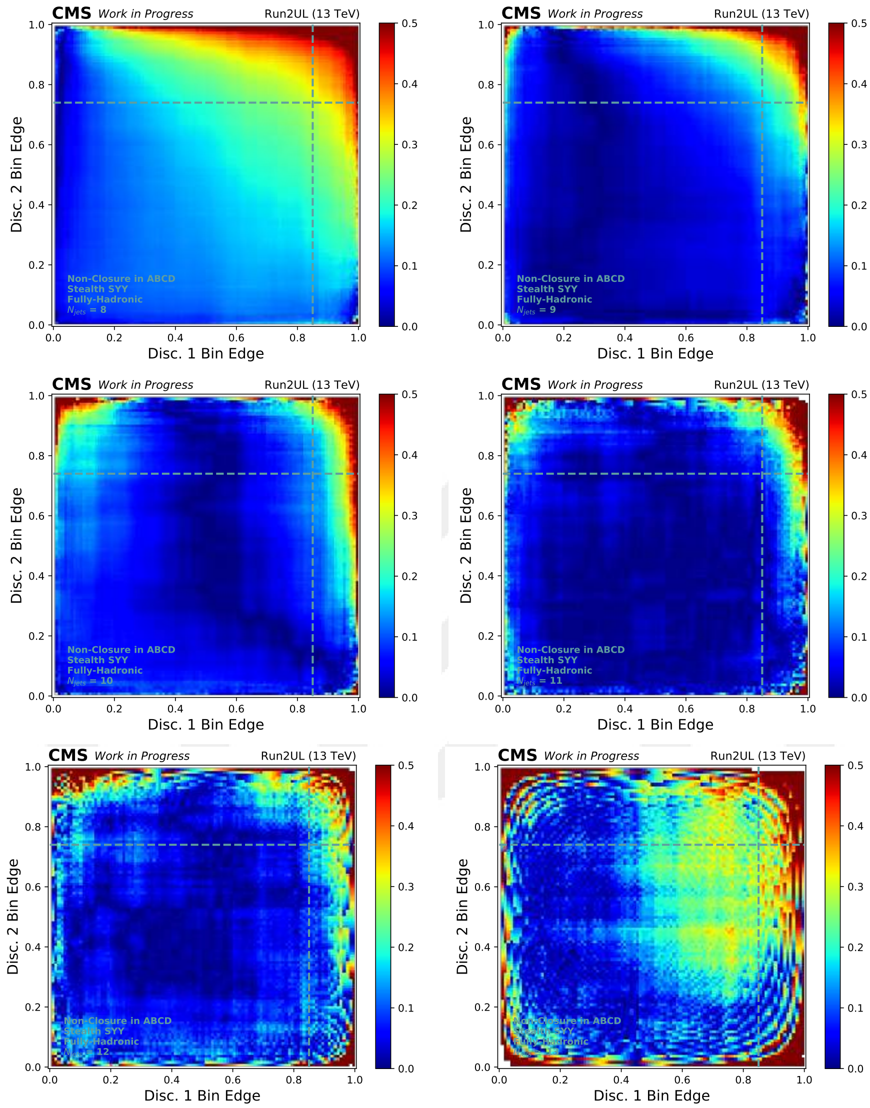


Figure D10. For the fully-hadronic channel, the non-closure as a function of “optimized” ABCD bin edges. $N_{\text{jets}} = 8,9,10,11,12\geq 13$ bins (left to right, top to bottom)

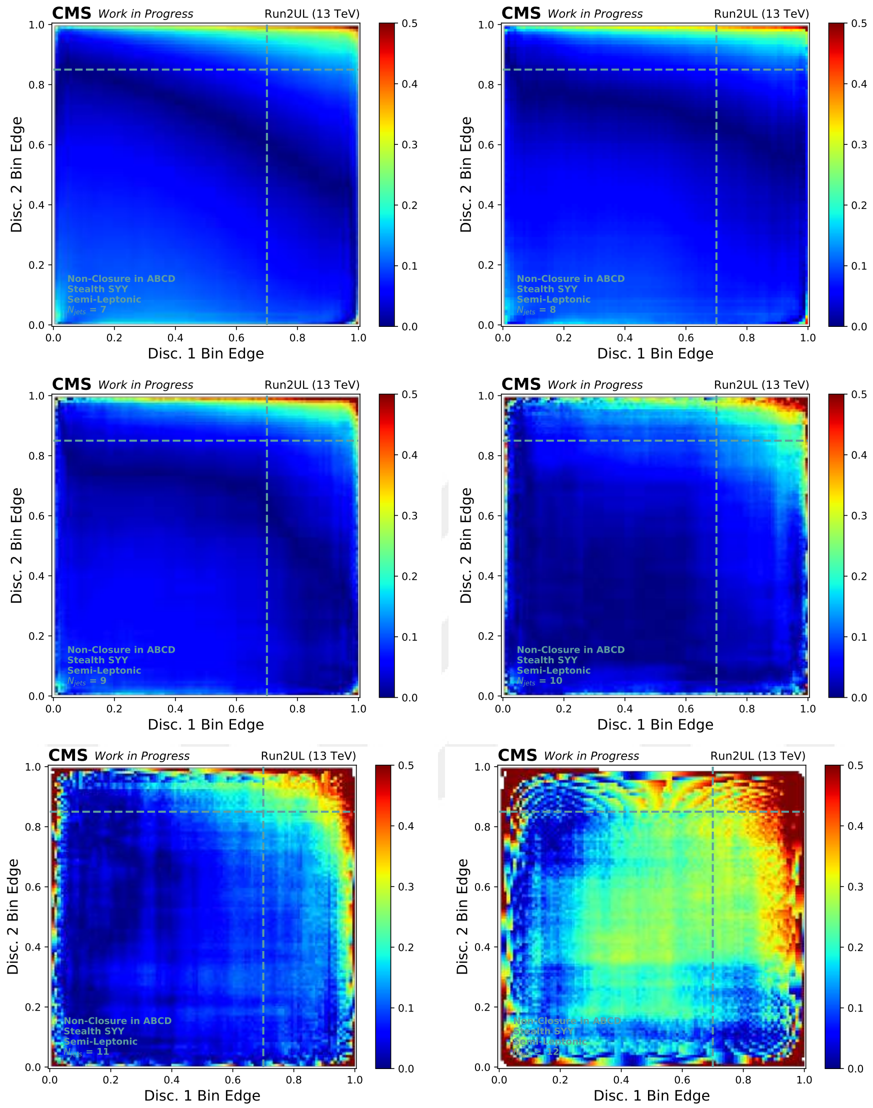


Figure D11. For the semi-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. $N_{\text{jets}} = 7,8,9,10,11,\geq 12$ bins (left to right, top to bottom)

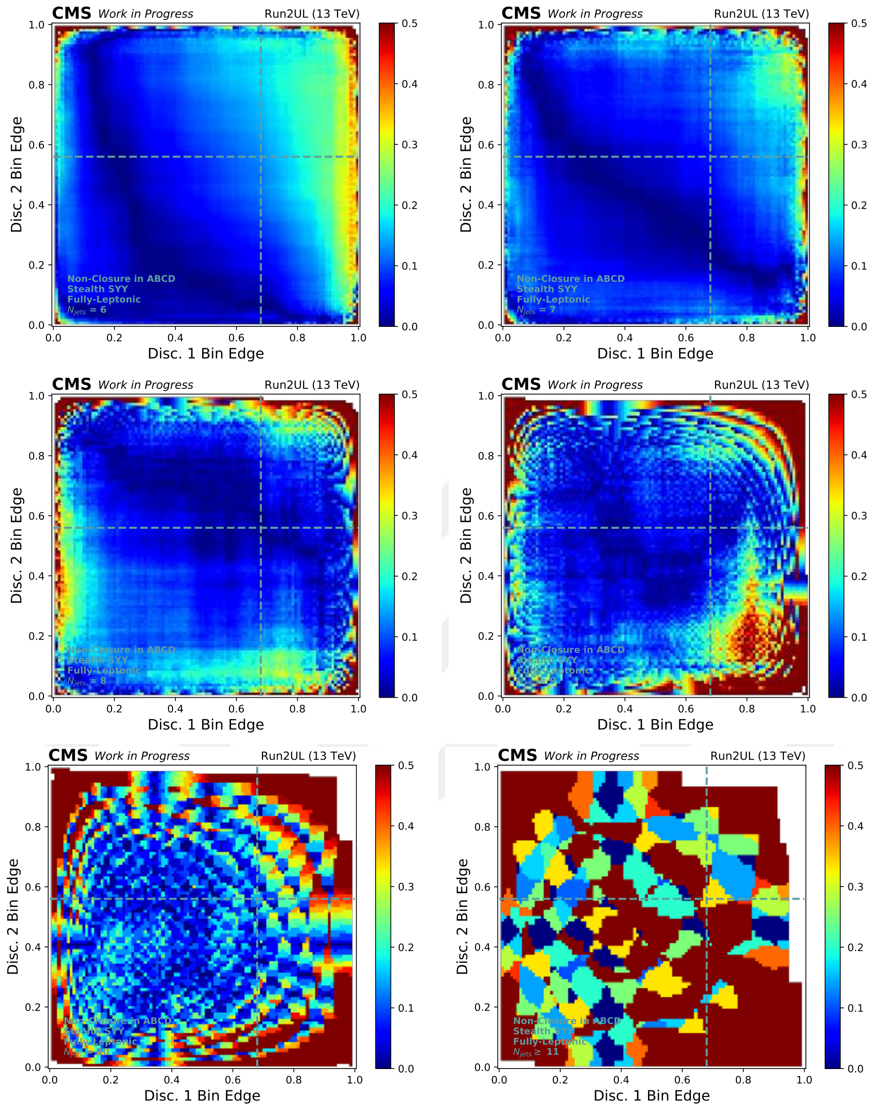


Figure D12. For the fully-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 6,7,8,9,10,≥11 bins (lest to right, top to bottom)

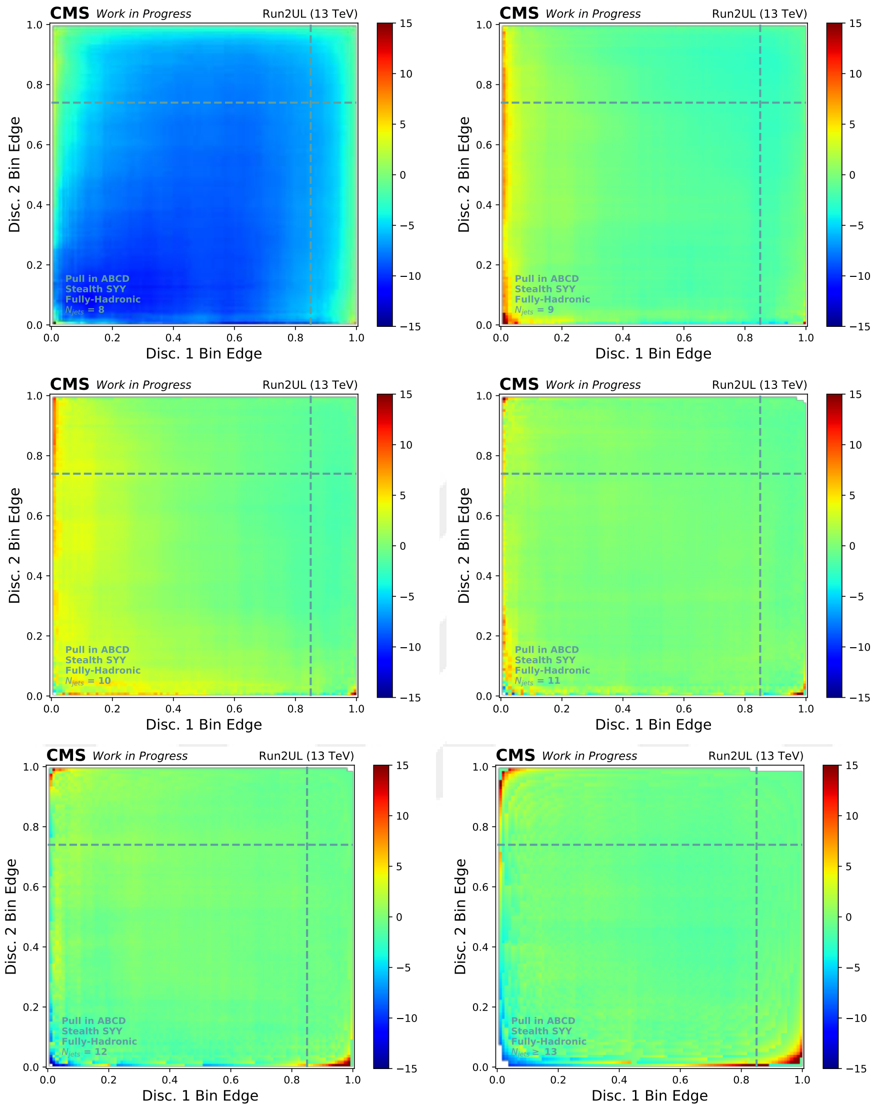


Figure D13. For the fully-hadronic channel, the non-closure as a function of "optimized" ABCD bin edges. $N_{\text{jets}} = 8,9,10,11,12\geq 13$ bins (left to right, top to bottom)

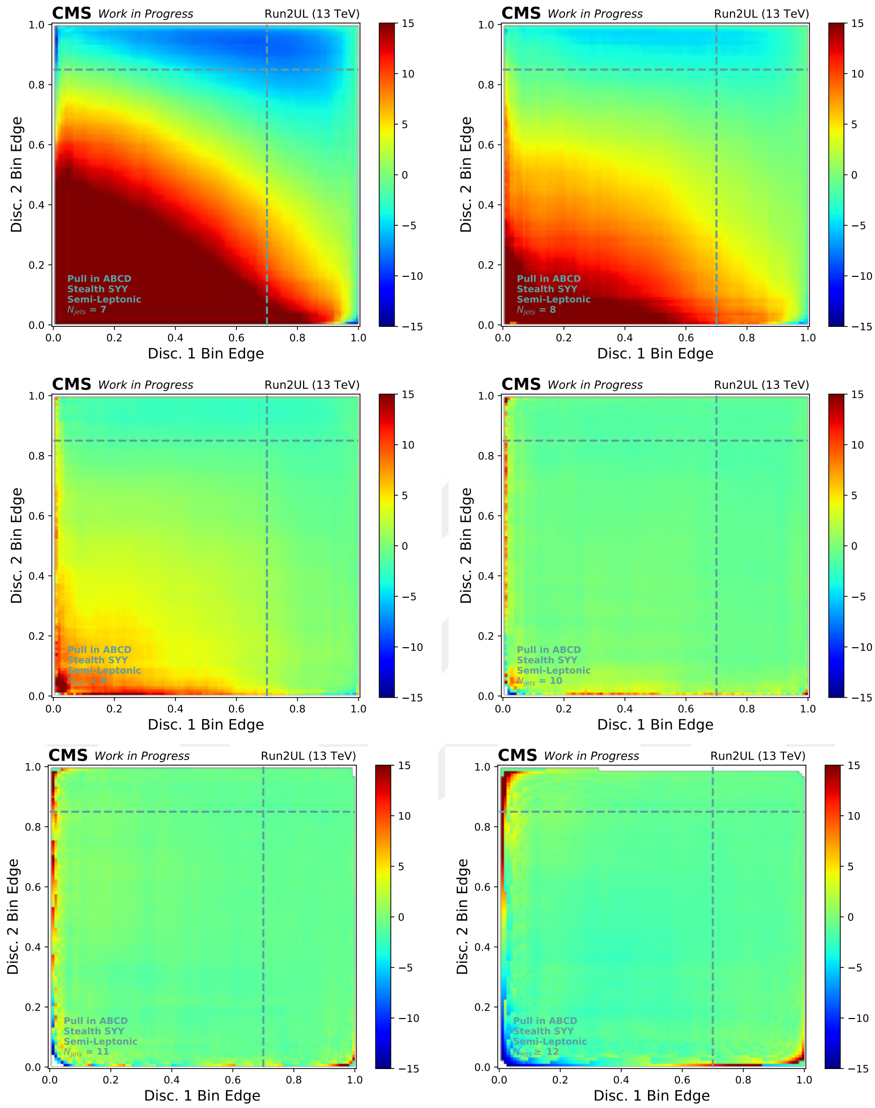


Figure D14. For the semi-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. $N_{\text{jets}} = 7,8,9,10,11,\geq 12$ bins (left to right, top to bottom)

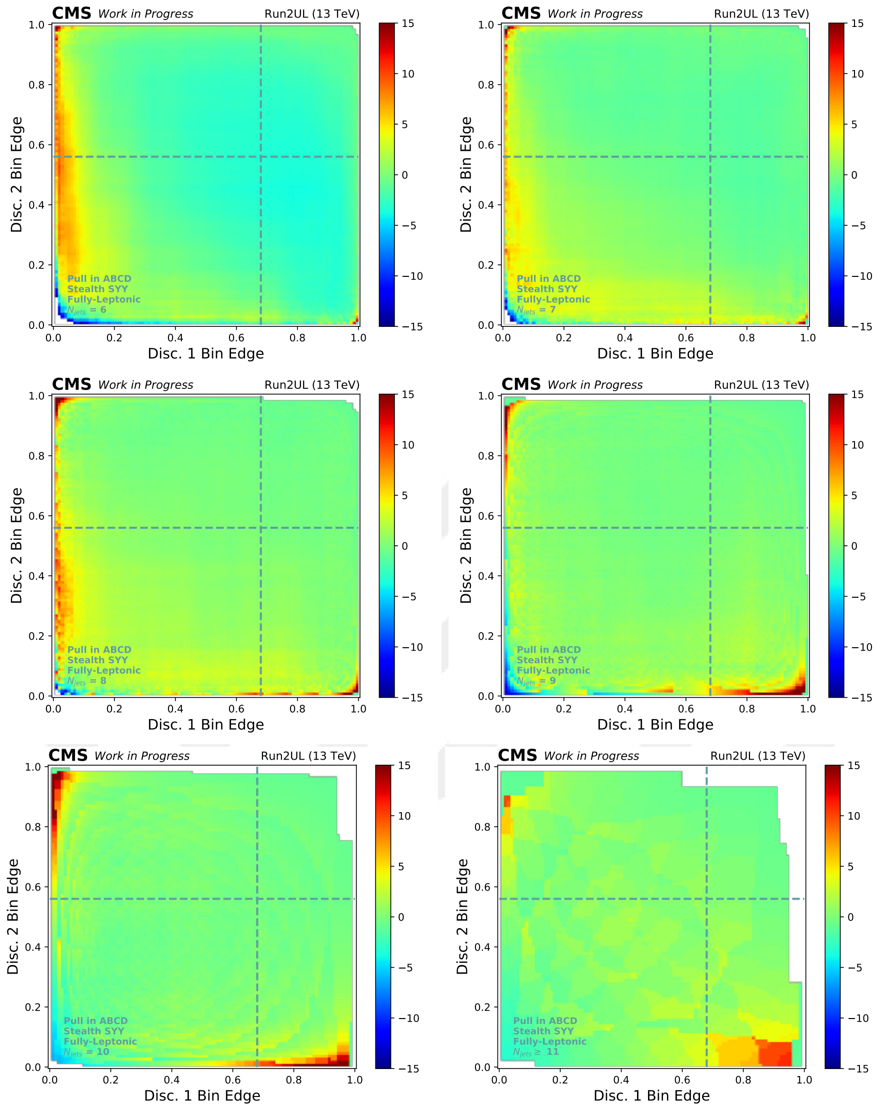


Figure D15. For the fully-leptonic channel, the non-closure as a function of "optimized" ABCD bin edges. N_{jets} = 6,7,8,9,10,≥11 bins (lest to right, top to bottom)

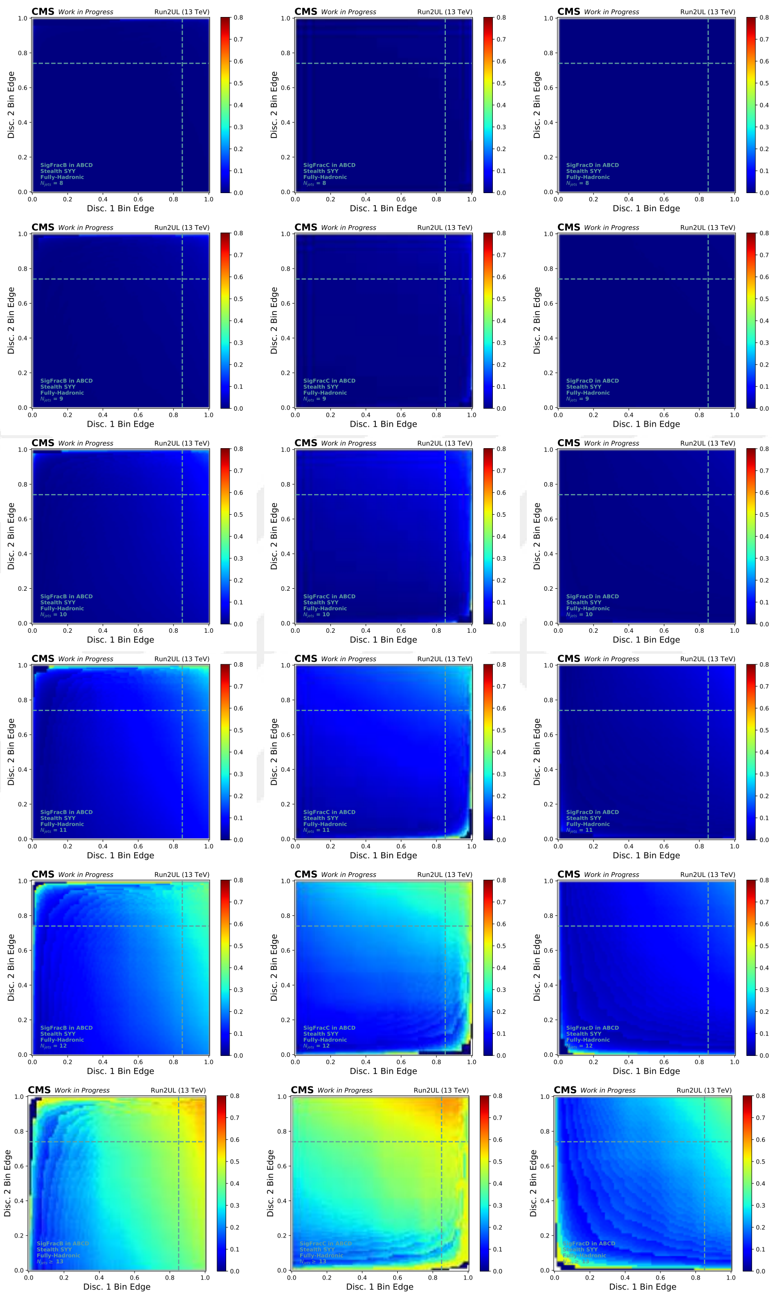


Figure D16. For the fully-hadronic channel, the signal fractions in each **B**, **C**, **D** region (left to right) as a function of "optimized" ABCD bin edges. N_{jets} = 8,9,10,11,12≥13 bins (top to bottom)

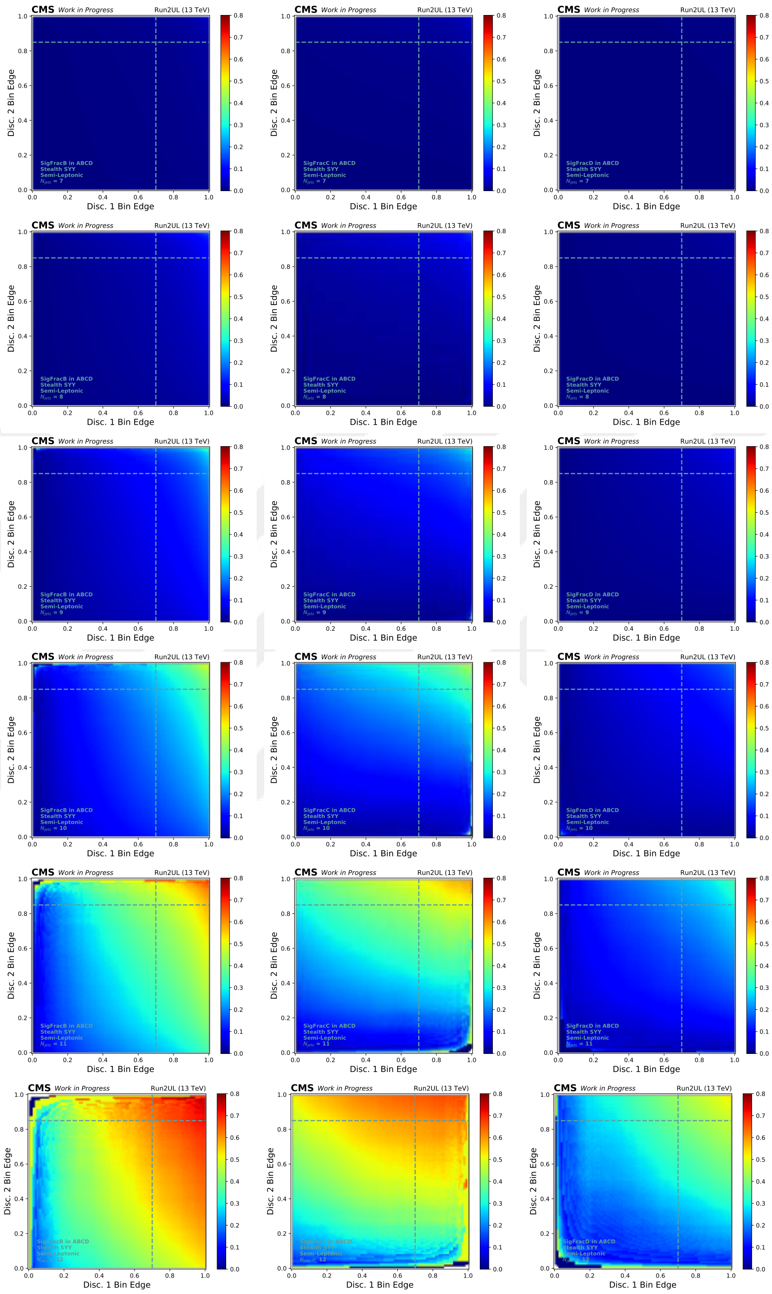


Figure D17. For the semi-leptonic channel, the signal fractions in each **B**, **C**, **D** region (left to right) as a function of "optimized" ABCD bin edges. N_{jets} = 7,8,9,10,11,$\geq$12 bins (top to bottom)

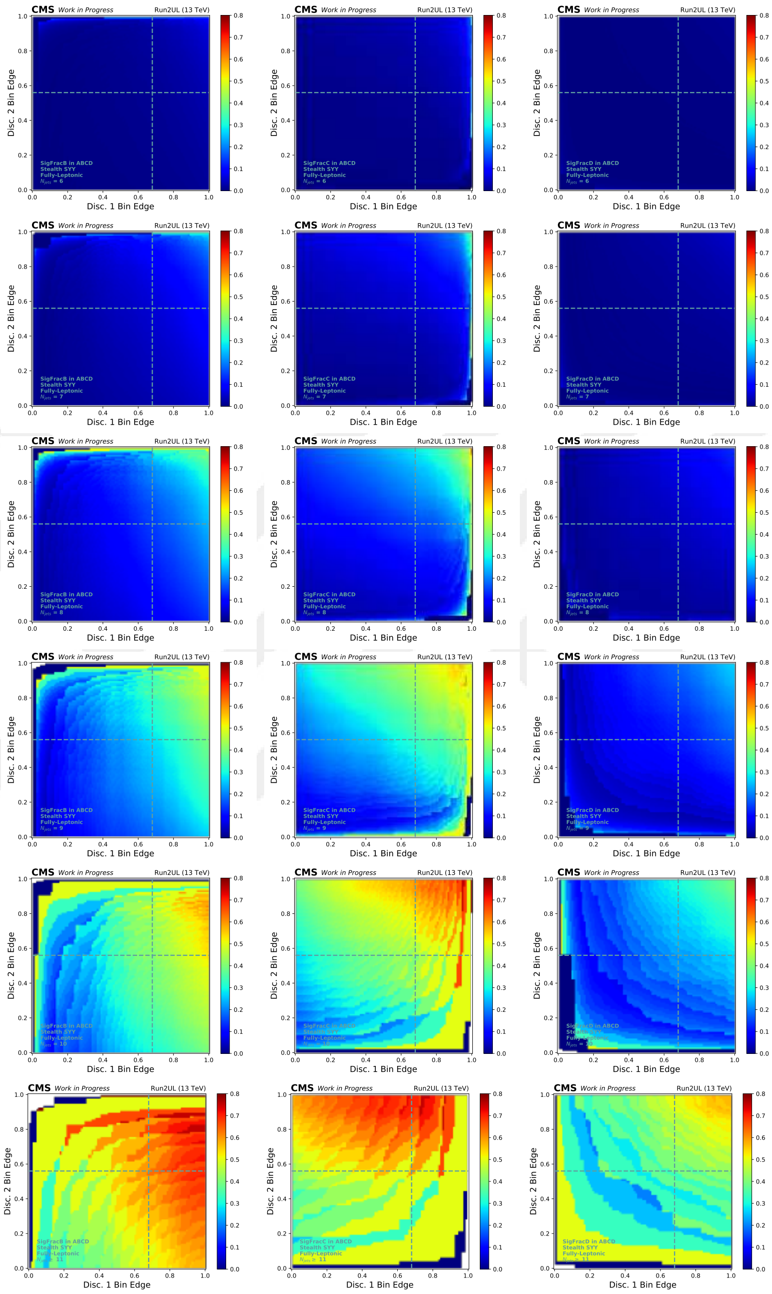


Figure D18. For the fully-leptonic channel, the signal fractions in each **B**, **C**, **D** region (left to right) as a function of “optimized” ABCD bin edges. N_{jets} = 6,7,8,9,10,≥11 bins (top to bottom)

D2. Verifying the Validation Regions

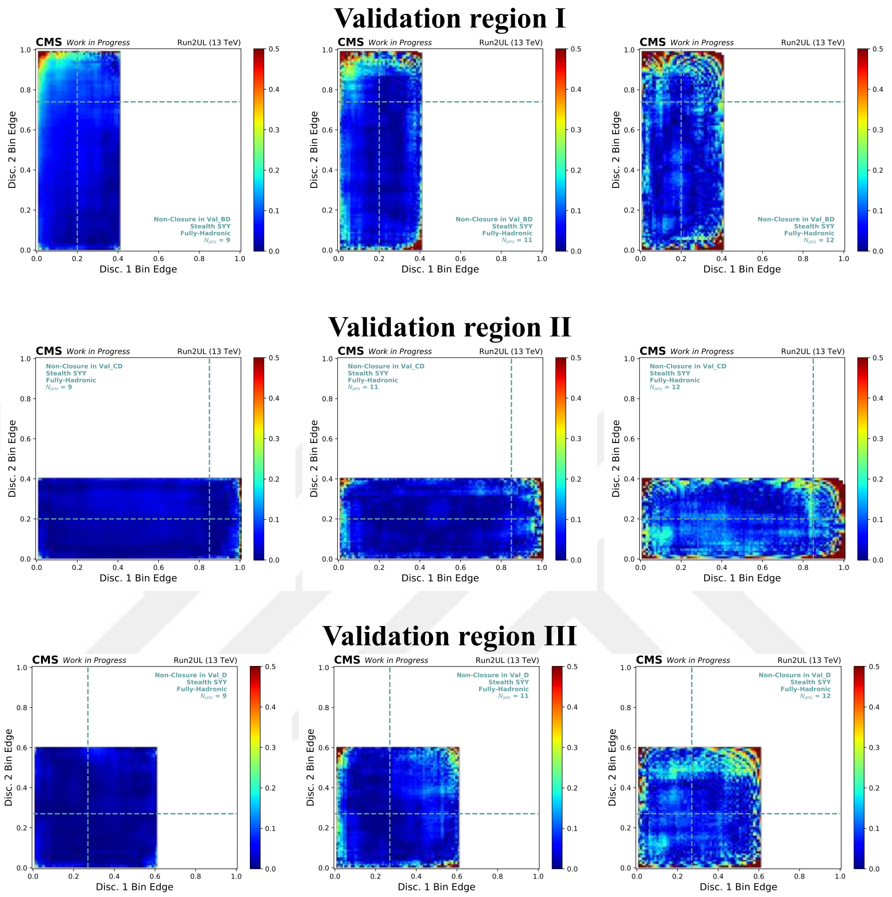


Figure D19. For the fully-hadronic channel, the non-closure as a function of validation I (top) edges, validation II (middle) edges, and validation III (bottom) edges. Left to right $N_{\text{jets}} = 9,11,12$ bins. Note that these are for the Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ GeV model

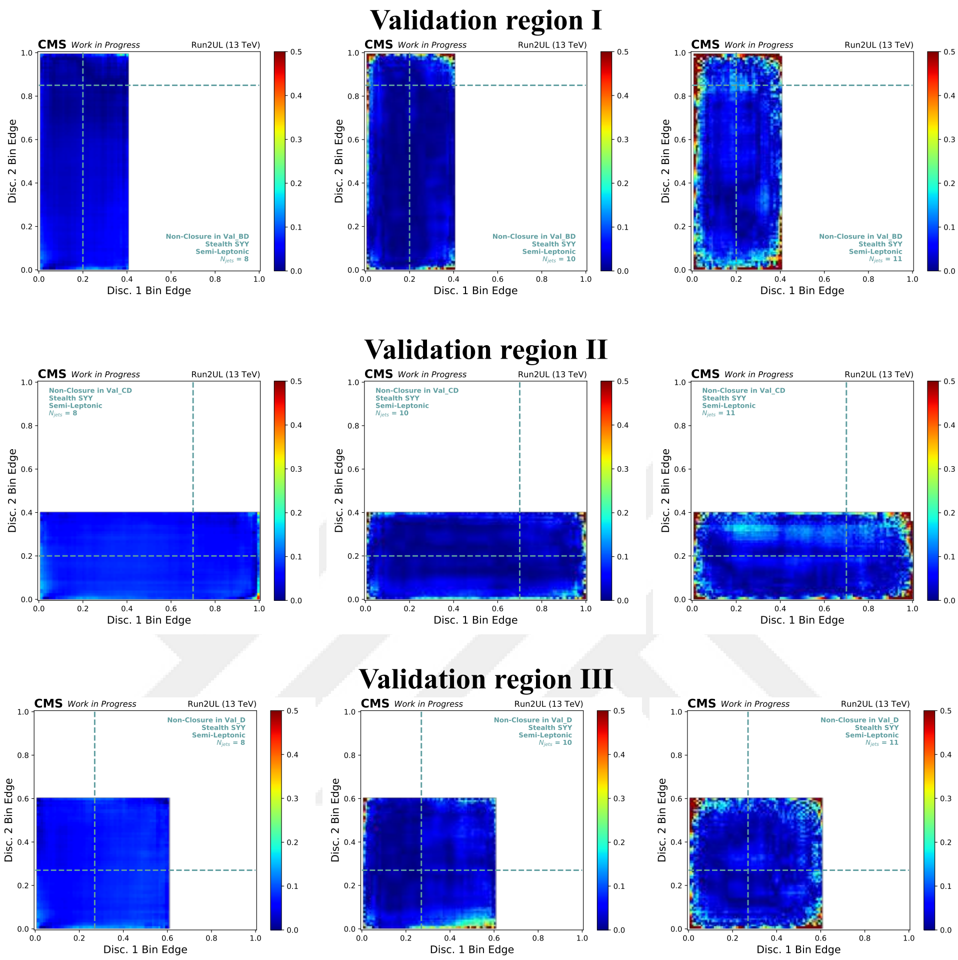


Figure D20. For the semi-leptonic channel, the non-closure as a function of validation I (top) edges, validation II (middle) edges, and validation III (bottom) edges. Left to right N_{jets} = 8,10,11 bins. Note that these are for the Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV model

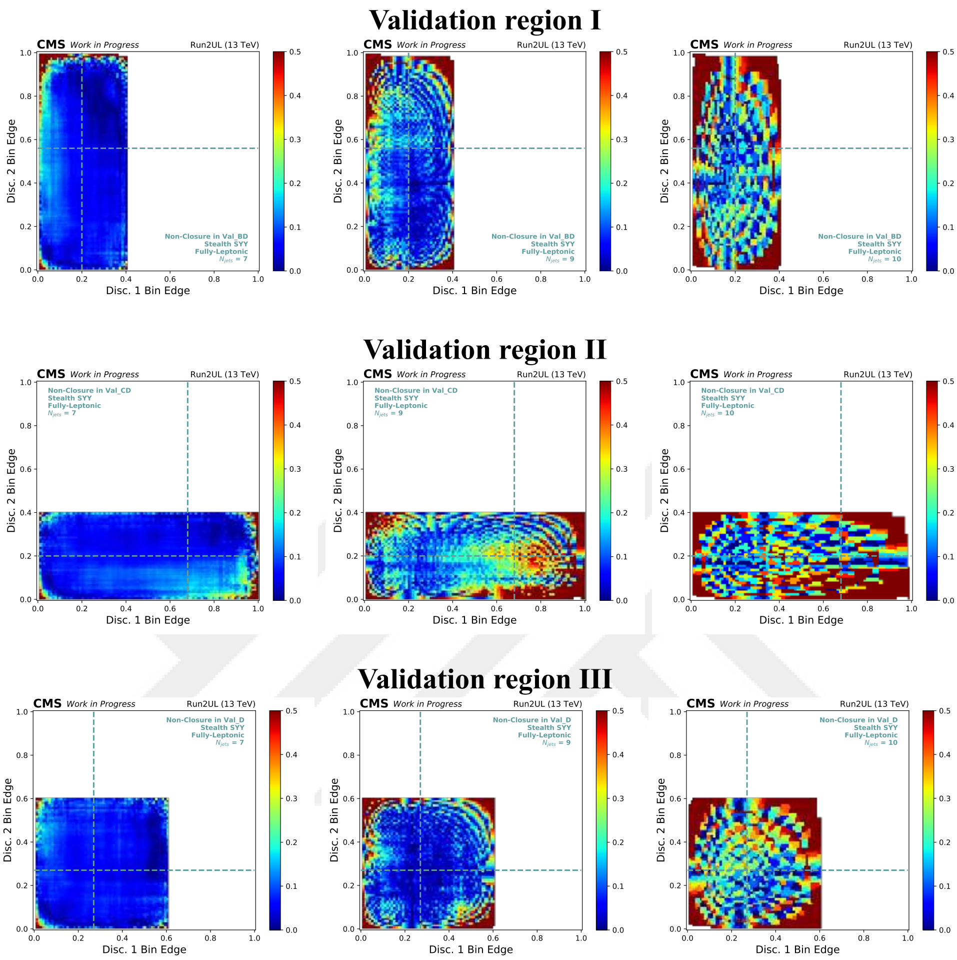


Figure D21. For the fully-leptonic channel, the non-closure as a function of validation I (top) edges, validation II (middle) edges, and validation III (bottom) edges. Left to right $N_{\text{jets}} = 7,9,10$ bins. Note that these are for the Stealth SY$\bar{Y}$ $m_{\tilde{t}} = 550$ GeV model

D3. $t\bar{t}$ + jets Data-Based Systematics

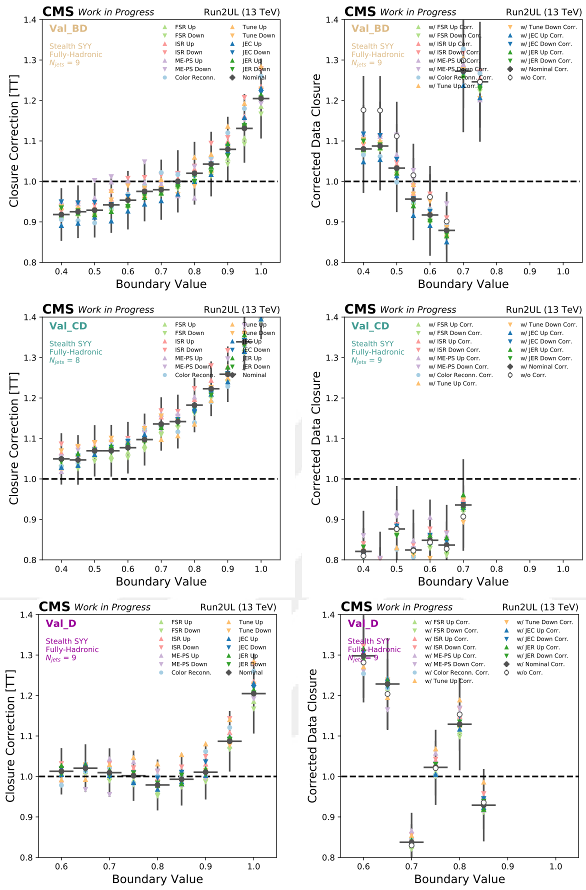


Figure D22. For the fully-hadronic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 9$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom)

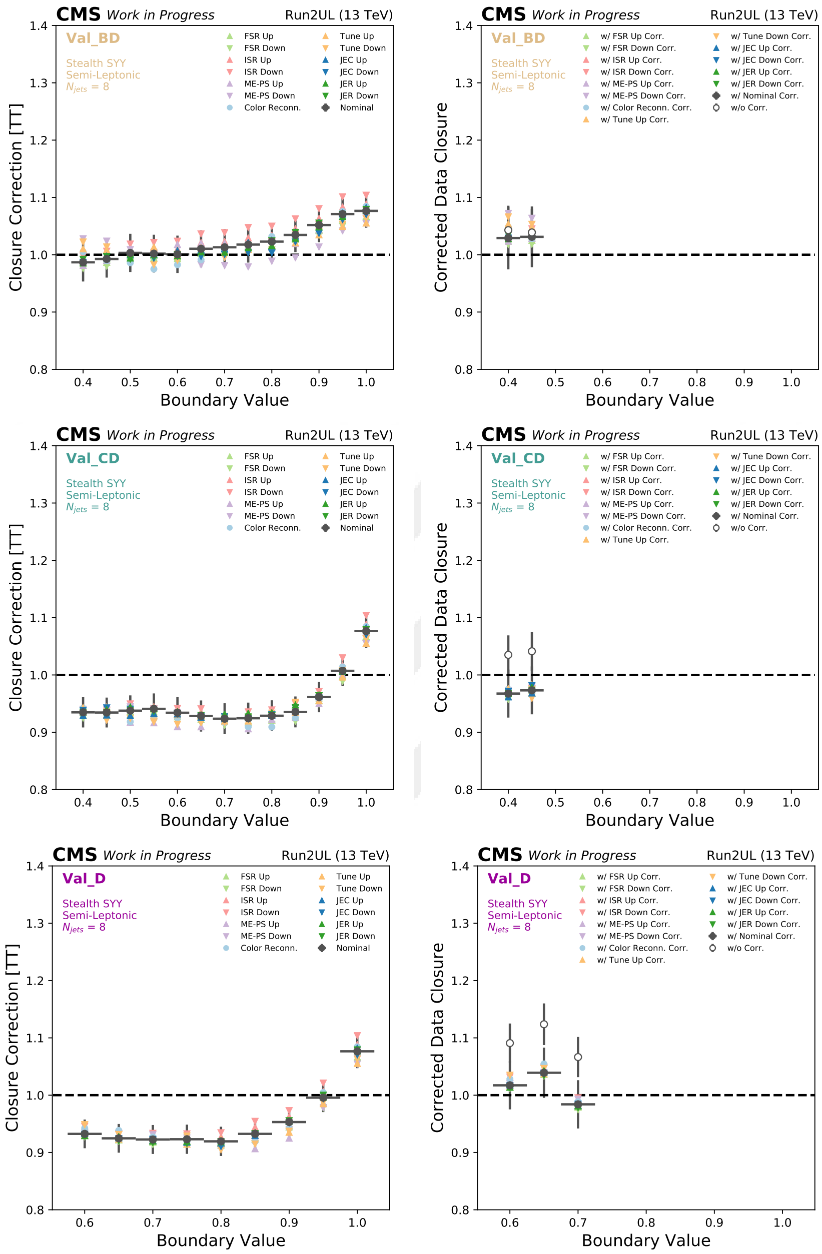


Figure D23. For the semi-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 8$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom)

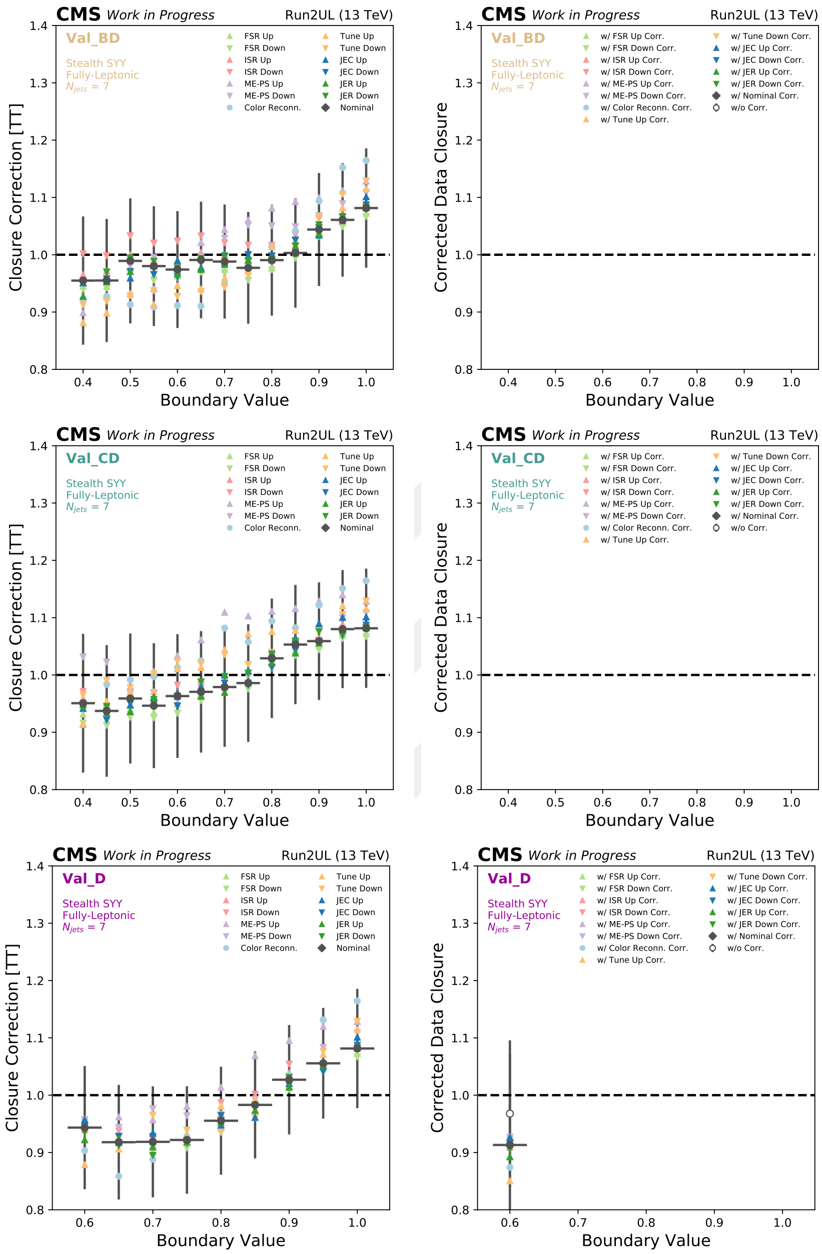


Figure D24. For the fully-leptonic channel, the closure corrections (left) and corrected data closures (right) as a function of boundary value for $N_{\text{jets}} = 7$. Top to bottom validation region I (top), validation region II (middle) and validation region III (bottom)

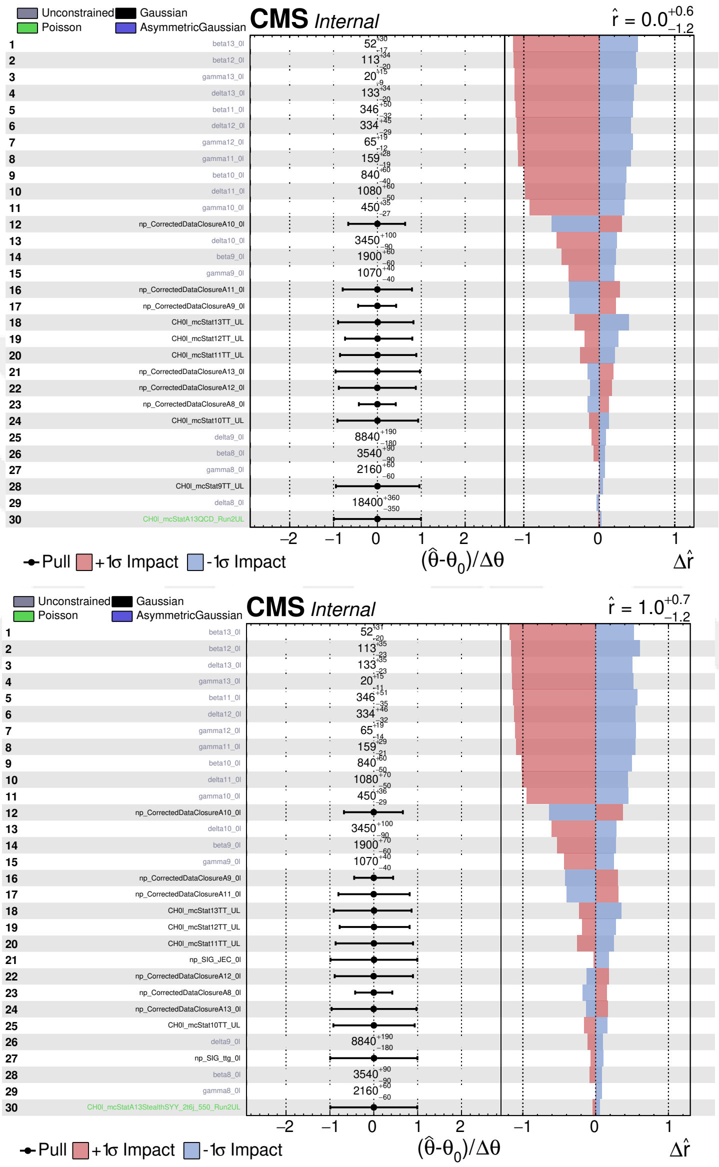


Figure E1. For the fully-hadronic channel, asimov style impact plots for Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV with pseuodaData (top) and pseuodaDataS (bottom)

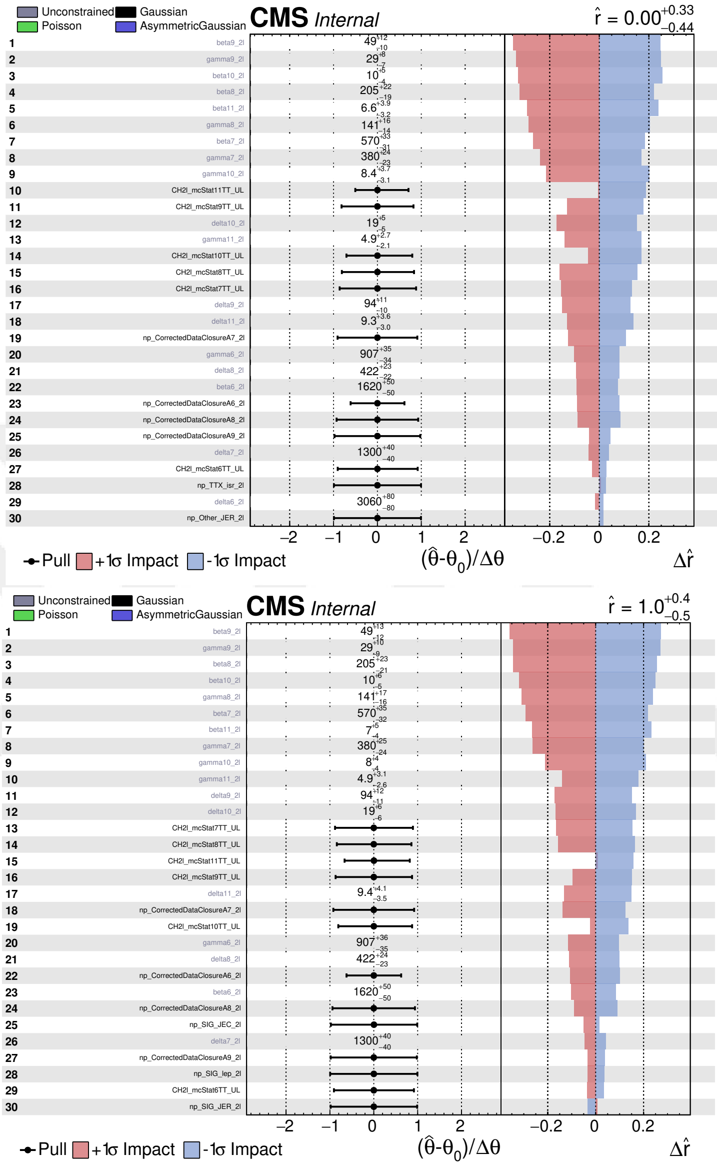


Figure E2. For the fully-leptonic channel, asimov style impact plots for Stealth SY$\bar{Y}$ $m_{\tilde{t}}$ = 550 GeV with pseuodaData (top) and pseuodaDataS (bottom)