



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Defying 4D Gravity: Novel Searches for Signatures of String Theory Via Non-minimal Dark Sectors, Scalar Cascades, Noncommutative Black Holes, and Deep Learning

Elena Villhauer



Doctor of Philosophy
The University of Edinburgh
January 2024

Abstract

The existence of extra dimensions could lower the fundamental Planck scale to the low TeV scale and very excitingly allow string theory to be probed at the LHC. This thesis presents novel searches for signatures of string theory via three models that have never been searched for by any experiment: scalar cascades, noncommutative black holes, and non-minimal dark sectors. This thesis includes the first-time exploration of a novel final state. Deep learning is employed through the use of particle flow networks to learn jet substructure and design analysis regions. The background estimation is automated through the implementation of distance correlation in the neural network loss function. A novel jet reclustering algorithm is used to overcome the challenges associated with a soft final state.

Lay Summary

The Standard Model (SM) of particle physics seeks to explain the most fundamental building blocks of the universe and the forces that govern them. It describes all of known particles, as well as three of the four forces: the strong force, weak force, and electromagnetic force. The strong force forms atoms from subatomic particles, the weak force causes the radioactive decay of atoms, and the electromagnetic force controls visible light and interactions between atoms. The Standard Model, however, does not explain gravity. String theory is an alternative model of the universe that does try to explain gravity. According to string theory, the fundamental building blocks of nature are not the point-like particles that the Standard Model predicts. Instead, they are vibrating strings. String theory includes fascinating ideas such as extra dimensions (beyond the 3 spatial and 1 time dimension that we know about) and the existence of parallel universes.

Extra dimensions could help explain why gravity is so much weaker than the three other forces. Gravity might be spread over each of the extra dimensions, thereby diluting its strength. This would help solve this pressing problem of modern physics: how to explain, in a single theory, gravity and the three other forces. Excitingly, the ATLAS experiment at the Large Hadron Collider in Geneva, Switzerland can be used to search for evidence of string theory. String theory could potentially appear in many different ways. This thesis focuses on three ideas of how extra dimensions may behave. When particles collide at the LHC, the production of a very large number of light particles may be possible evidence of string theory. Sometimes these particles will be scattered widely in all directions and sometimes they will be collected into a narrow spray (referred to as a jet). The patterns formed by these particles can be studied.

The analyses in this thesis use a machine learning system to attempt to identify the patterns of particles that are most interesting. The machine learning system is trained with computer simulations of evidence for string theory. It is then used to

try to identify the same things in real data from the ATLAS experiment. Though the searches in the thesis did not provide any evidence for string theory, they have developed novel techniques that will help future searches. Additionally, these searches help future searches by having eliminated some of potential places that evidence for string theory may be hiding.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Elena Villhauer, January 2024)

Acknowledgements

I would like to thank my PhD supervisor, Christos Leonidopoulos, for allowing me the intellectual freedom to propose and pursue my own ideas throughout the entirety of my degree and to bring something entirely new to the LHC. This has meant absolutely everything to me. Thank you for all of your support and for always making time for me.

Thank you to Mark Kruse for the incredible entrance into the world of high energy physics with Duke and for allowing me to propose my first ever idea in high energy physics while just less than a week old in the field. It made me think that, if it was okay for me to do that less than a week old in the field, it should be okay when I am older and know more. Thank you.

Thank you to Ariel Schwartzman for joining me in the quest to defy 4D gravity. Your genuine excitement for this work has made every meeting something to look forward to. Thank you for viewing a challenging final state as an opportunity for innovation and excitement. In Ariel, I found a tireless and endlessly selfless mentor. From spending over three hours helping me perfect my seminar at CERN during Upgrade Week as Upgrade Physics convener to consistently carving out time for me each week in a calendar that seemed to have eight overlapping meetings at a given time, I have never felt anything other than infinitely cared for and as though I was one of his own students. The environment that Ariel creates is the most conducive to research. Thank you for always being there for me with consistent, kind, meticulous, and impactful feedback, and for helping me come back stronger from setbacks. It has been the highest privilege to work with you, and I hope to continue working with you for years to come.

Thank you to theorists Matthew Low and Raffaele Tito D'Agnolo, for being so kind and open-minded to my wildest ideas and helping me make them a reality. Collaborating with theorists has always been the highlight of my time in ATLAS, and I am grateful as well to the non-minimal dark sector theorists, especially Doojin Kim, and Huayang Song, for fielding all of my questions.

Thank you to Christian Sander for his kindness and his invaluable feedback and advice as JDM convener.

Thank you to Ben Wynne, to whom I am very grateful for his kindness and help.

Thank you also to Sara Alderweireldt for her kindness and help during these last two months.

Thanks to my dad for the thermite reactions, the infamous dry ice b*mb, the homemade glowsticks via chemiluminescence reactions, all of the trips to the Liberty Science Center as a kid, for helping me use Woofie as the test chihuahua for my 5th grade science project, and for the vildagliptin stories. Thanks to my mom, Eric, Cindy, and my friends for the support. Thanks especially to Amanda Murphy and Ms. Judith Spinola for the decades of friendship and to Jaime Coello de Portugal - Martínez Acacio Vázquez for his friendship since I first ever arrived at CERN.

Contents

Abstract	i
Lay Summary	ii
Declaration	iv
Acknowledgements	v
Contents	vii
List of Figures	xiii
List of Tables	xxxi
1 Introduction	1
2 Theoretical Background	7
2.1 The Standard Model.....	7
2.2 Novel searches for String Theory	11
2.3 ADD-type extra dimensions.....	11
2.4 RS-type extra dimensions and Universal Extra Dimensions.....	12
2.5 Black Hole Formation at the LHC.....	12
2.6 Semiclassical Black Holes	13
2.7 Noncommutative Black Holes.....	15

2.8	Scalar cascades	16
2.8.1	3 Body Stable Chain Decay	19
2.8.2	3 Body Stable Cascade Decay	20
2.8.3	2 Body Stable Cascade Decay	21
2.8.4	3 Body Unstable Chain Decay	22
2.8.5	3 Body Unstable Cascade Decay	23
2.8.6	2 Body Unstable Cascade Decay	24
2.9	Non-minimal Dark Sectors	25
2.10	Summary	29
3	The Large Hadron Collider and the ATLAS Detector	30
3.1	The Large Hadron Collider	30
3.2	Coordinate System and Geometry	35
3.3	Inner Detector	37
3.4	Calorimeters	39
3.5	Electromagnetic Calorimeter	39
3.6	Hadronic Calorimeter	42
3.7	Forward Calorimeter	43
3.8	Muon Spectrometer	43
3.9	Forward Detectors	44
3.10	Magnet System	45
3.11	Trigger and Data Acquisition System	45
3.11.1	L1 trigger	46
3.11.2	HLT	47

3.11.3	Summary	47
4	Object Reconstruction	49
4.1	Object Reconstruction	50
4.1.1	Tracks, Vertexing, Topo-clusters.....	50
4.1.2	Electrons and Photons	51
4.1.3	Muons.....	51
4.1.4	Missing Energy	52
4.1.5	Jets	52
4.1.6	Small R jets	57
4.1.7	Large $\mathbf{R}$ jets and Jet Substructure.....	58
4.1.8	Summary	60
5	Trigger Upgrade and jFEX Jet Calibrations	61
5.1	Upgrade.....	61
5.2	jFEX Jet Calibration	65
5.3	Summary	73
6	Unexcluded Phase Space, Kinematics, and Search Strategy for Non-minimal Dark Sectors, Scalar Cascades, and Noncommuta- tive Black Holes	75
6.1	Noncommutative Black Hole Parameters and MC Samples	76
6.2	Non-minimal Dark Sector Parameters and MC Samples	91
6.3	Scalar Cascade Parameters and MC Samples.....	105
6.4	Summary	114

7	A jet-substructure-based analysis strategy and a novel jet-reconstruction algorithm	115
7.1	Event Displays	115
7.2	Large R jet Distributions	137
7.3	Novel Large R jet Reconstruction Algorithm (Hybrid Jet Reclustering) vs. Standard Large R Jet Reclustering.....	142
7.3.1	Technical details of Hybrid Jet Reclustering.....	144
7.3.2	Determining a radius for Hybrid Jet Reclustering.....	145
7.3.3	Validating the Hybrid Jet Reclustering Algorithm	146
7.4	Summary	148
8	Datasets and Event Selection	149
8.1	Datasets	149
8.1.1	Background Processes.....	150
8.1.2	Object Definition and Selection.....	154
8.1.3	Cuts Specific to Background and Cuts Specific to Data	156
8.1.4	Scalar Cascade Cuts.....	158
8.1.5	Noncommutative Black Hole and Non-minimal Dark Sector Cuts	163
8.2	Summary	166
9	Machine Learning for Signal Region Definition and Background Estimation	167
9.1	Principles of Machine Learning and Neural Networks	168
9.2	Deep Sets Based Architectures	174
9.3	Distance Correlation (DisCo).....	176
9.4	Particle Flow Network - DisCo Architecture	178

9.5	Introduction to event shape variables and other event level characterizations.....	181
9.6	Scalar Cascades Neural Network Results	184
9.7	Noncommutative Black Holes and Non-minimal Dark Sector Neural Networks.....	197
9.7.1	Noncommutative Black Holes and Non-minimal Dark Sector Neural Networks - 200 GeV E_T^{miss} cut	198
9.7.2	Noncommutative Black Holes Networks - 500 GeV E_T^{miss} cut....	208
9.7.3	Non-minimal Dark Sector Networks - 500 GeV E_T^{miss} cut with 1400 GeV H_T cut	218
9.7.4	Noncommutative Black Holes Neural Network Results.....	228
9.7.5	Non-minimal Dark Sector Neural Network Results	237
9.8	Summary	245
10	Background Estimation and Statistical Analysis	247
10.1	Background Estimation	248
10.2	ABCD Analysis for Scalar Cascades	253
10.3	NCBH ABCD Analysis.....	272
10.4	NMDS ABCD Plane	287
10.5	Summary	299
11	Statistical Analysis	301
11.1	Signal Contamination.....	301
11.2	Likelihood Fit	304
11.3	Frequentist Approach.....	306
11.4	NCBH Benchmark NCBH1 Exclusion Limit	310
11.5	NMDS Exclusion Limit	312

11.6 Summary	313
12 Conclusion	314
13 Appendix	316
13.1 Effects of tight cleaning and bad tile event veto on data.....	316
13.2 Systematic Uncertainties for Signal Monte Carlo Models	318
13.3 Full names of background MC samples, signal MC samples, and data samples used.....	326
13.4 Cross sections, filter efficiencies, and k Factors for MC Background samples used.....	330
13.5 Conservative Nonclosure Uncertainties	333
13.6 Trigger Efficiencies for Signal Monte Carlo.....	335
Bibliography	352

List of Figures

2.1	The 17 particles of the Standard Model of Particle Physics. Figure taken from [10].	8
2.2	Scalar cascade generation via (a) pair production $\sigma \sim (\lambda^{\text{SM}})^2$ (where h^* is the Higgs) and (b) single production $\sigma \sim (a^{\text{SM}})^2$. Figure modified from [3].	17
2.3	Scalar cascade collider phenomenology. Scalar Cascade decays via (a) hidden sector quartic $\Gamma \sim \lambda^2$, (b) Higgs portal $\Gamma \sim (\lambda^{\text{SM}})^2$, (c) hidden sector trilinear $\Gamma \sim a^2$, and (d) SM trilinear $\Gamma \sim (a^{\text{SM}})^2$. Figure modified from [3].	18
2.4	Scalar Cascades 3 Body Stable Chain Decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. Dashed lines indicate that there can be more steps in the decay. The lightest ϕ is stable.	19
2.5	Scalar Cascades 3 Body Stable Cascade Decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. Dashed lines indicate that there can be more steps in the decay. The lightest ϕ is stable.	20
2.6	Scalar Cascade 2 Body Stable Cascade Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is stable.	22
2.7	Scalar Cascade 3 Body Unstable Chain Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is unstable.	23
2.8	Scalar Cascade 3 Body Unstable Cascade Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is unstable.	24

2.9	Scalar Cascade 2 Body Unstable Cascade Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is unstable.	25
2.10	$\chi\phi$ Production Channel of Non-minimal Dark Sectors	26
2.11	Decay chain in which each χ (ensemble constituent) decays into a χ and two quarks. The process ends with the production of two quarks and a stable, lightest χ	26
2.12	$\phi\phi$ Production Channel of Non-minimal Dark Sectors	27
2.13	$\chi\chi$ Production Channel of Non-minimal Dark Sectors	27
3.1	Aerial view of Geneva, Switzerland and the surrounding area with a drawing of the Large Hadron Collider (LHC) and associated experiments overlaid. Figure taken from [59].	31
3.2	CERN Accelerator Complex. Figure taken from Reference [63].	32
3.3	LHC timeline. Figure taken from Reference [64].	33
3.4	ATLAS Integrated Luminosity delivered in Run 2. Figure taken from Reference [68].	34
3.5	Mean Number of Interactions in the ATLAS detector during Run 2. Figure taken from Reference [68].	35
3.6	ATLAS Detector depicted in figure on the left, taken from Reference [69]. Coordinate system used by ATLAS on the right, taken from Reference [70].	36
3.7	ATLAS Inner Detector. Figure taken from Reference [69].	38
3.8	View of the accordion structure of the Electromagnetic Calorimeter and the composition of the ECAL. Figure taken from Reference [71].	40
3.9	Wedge-shaped module in the HEC. Figure taken from Reference [72].	42
3.10	Muon Spectrometer	44
3.11	ATLAS Muon Spectrometer. Figure taken from Reference [69].	44
3.12	ATLAS Magnet System. Figure taken from Reference [69].	45
3.13	ATLAS Trigger System. Figure taken from Reference [74].	46
3.14	Level 1 trigger tower. Figure taken from Reference [74].	47

4.1	Illustration of jet clustering with three different jet reclustering algorithms applied to the same event. Each colored region corresponds to a different reconstructed jet. Figure taken from [81]	54
4.2	Distribution of Jet Vertex Tagger scores for jets originating from pileup (PU) and the hard scatter (HS) and with transverse momentum ranges from 20 GeV to 30 GeV. Figure taken from [83]	57
4.3	Energy spread of a boosted QCD jet with $R = 0.6$. Figure taken from [85]	58
4.4	Energy spread of a boosted W jet with $R = 0.6$. Figure taken from [85]	59
4.5	Energy spread of a boosted top jet with $R = 0.8$. Figure taken from [85]	59
5.1	L1 Calorimeter from Run 2 and Run 3. Figure taken from [88] . . .	62
5.2	jFEX prototype. Figure taken from Reference [88]	62
5.3	Small R jFEX jet for Run 3. Figure modified from [89]	63
5.4	Large R jFEX jet for Run 3. Figure taken from [89]	64
5.5	High variance of jet energy response over η . Figure taken from [74] .	65
5.6	Response map for hardware. The p_T response is plotted as a function of η and jFEX reconstructed jet p_T	67
5.7	Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing p_T response as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-4.9 < \eta < -1.6$, $-0.8 < \eta < 0$, and $0.8 < \eta < 1.6$. The p_T of the jFEX truth jet is labeled as p_T True.	68
5.8	Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing p_T response as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-1.6 < \eta < -0.8$, $0 < \eta < 0.8$, and $1.6 < \eta < 4.9$. The p_T of the jFEX truth jet is labeled as p_T True.	69
5.9	Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing resolution as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-4.9 < \eta < -1.6$, $-0.8 < \eta < 0$, and $0.8 < \eta < 1.6$. The p_T of the jFEX truth jet is labeled as p_T True.	70

5.10	Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing resolution as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-1.6 < \eta < -0.8$, $0 < \eta < 0.8$, and $1.6 < \eta < 4.9$. The p_T of the jFEX truth jet is labeled as p_T True.	71
5.11	Comparison of the performance of the jFEX calibration using fine η binning to the performance of the jFEX calibration using coarser η binning by showing resolution as a function of p_T of jFEX truth jets. Results shown for coarse η bins $0 < \eta < 0.8$ and $0.8 < \eta < 1.6$ and the following fine η bins: $0.5 < \eta < 1.0$ and $-0.5 < \eta < 0$. The p_T of the truth jet is labeled as p_T True.	72
5.12	Comparison of the performance of the jFEX calibration using fine η binning to the performance of the jFEX calibration using coarser η binning by showing resolution of as a function of p_T of jFEX truth jets. Results shown for coarse η bins $-4.9 < \eta < -1.6$ and $-0.8 < \eta < 0$ and the following fine η bins: $-4.5 < \eta < -4.0$ and $-0.5 < \eta < 0$. The p_T of the truth jet is labeled as p_T True.	73
6.1	Decay chain in which each χ (ensemble constituent) decays into a χ and two quarks. The process ends with the production of two quarks and a stable, lightest χ	92
7.1	Scalar cascade generation via pair production. Figure taken from [3]	117
7.2	Legend for all Scalar Cascade Event Displays. The final state particles are tagged as coming from ϕ_i or ϕ_j of the scalar cascade model or as not coming from the scalar cascade process. The legend includes a breakdown of seven main categories: E_T^{miss} , leptons, hadrons, anti- k_t truth jets of $R = 0.2$, $R = 0.4$, and $R = 1.0$, and p_T indicators (the smallest colored circles in the display) that indicate whether the truth jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS. The anti- k_t truth jets are additionally color coded to indicate the jet with the highest p_T (leading jet), the jet with the second highest p_T (subleading jet), etc.	118
7.3	Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC1 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.	120

7.4	Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC1 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.	122
7.5	Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC2 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.	123
7.6	Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC2 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.	124
7.7	Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC3 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.	125
7.8	Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC4 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.	126
7.9	$\phi\phi$ Production Channel of Non-minimal Dark Sectors	127
7.10	$\chi\phi$ Production Channel of Non-minimal Dark Sectors	127

7.11	$\chi\chi$ Production Channel of Non-minimal Dark Sectors	127
7.12	Legend for all NMDS Event Displays. The final state particles are tagged as coming from branch 1 or branch 2 of the NDMS model or as not coming from the NMDS process. The legend includes a breakdown of seven main categories: E_T^{miss} , leptons, hadrons, anti- k_t truth jets of $R = 0.2$, $R = 0.4$, and $R = 1.0$, and p_T indicators (the smallest colored circles in the display) that indicate whether the truth jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS. The anti- k_t truth jets are additionally color coded to indicate the jet with the highest p_T (leading jet), the jet with the second highest p_T (subleading jet), etc.	128
7.13	Distribution of final state particles in the η - ϕ plane for NMDS Benchmark NMDS1 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from branch 1 or branch 2 of the NMDS model. The corresponding legend for this event display is found in Figure 7.12.	130
7.14	Distribution of final state particles in the η - ϕ plane for NMDS Benchmark NMDS2 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from branch 1 or branch 2 of the NMDS model. The corresponding legend for this event display is found in Figure 7.12.	131
7.15	Distribution of final state particles in the η - ϕ plane for NMDS Benchmark NMDS3 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from branch 1 or branch 2 of the NMDS model. The corresponding legend for this event display is found in Figure 7.12.	132

7.16	Legend for all NCBH Event Displays. The final state particles are tagged as either coming from the black hole or as not coming from the black hole. The legend includes a breakdown of seven main categories: E_T^{miss} , leptons, hadrons, anti- k_t truth jets of $R = 0.2$, $R = 0.4$, and $R = 1.0$, and p_T indicators (the smallest colored circles in the display) that indicate whether the truth jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS. The anti- k_t truth jets are additionally color coded to indicate the jet with the highest p_T (leading jet), the jet with the second highest p_T (subleading jet), etc.	133
7.17	Distribution of final state particles in the η - ϕ plane for NCBH Benchmark NCBH1 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from the black hole. The corresponding legend for this event display is found in Figure 7.16. . .	134
7.18	Distribution of final state particles in the η - ϕ plane for NCBH Benchmark NCBH4 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from the black hole. The corresponding legend for this event display is found in Figure 7.16. . .	136
7.19	Noncommutative Black Hole, Non-minimal Dark Sector, and Scalar Cascades Large R jet Sum of masses with a 200 GeV E_T^{miss} cut applied.	139
7.20	Scalar Cascades Large R jet p_T distributions with a 200 GeV E_T^{miss} cut applied	140
7.21	Noncommutative Black Hole Large R jet p_T distributions with a 200 GeV E_T^{miss} cut applied	141
7.22	Non-minimal Dark Sectors Large R jet p_T distributions with a 200 GeV E_T^{miss} cut applied	142
7.23	Left plot: A visual of large R jet reclustering showing (a) Left: the large R jets and all of its associated truth particles (b) Right: all of the small R jets associated with the large R jet. Figure taken from [95]	143
7.24	ΔR between final state particles and ϕ_i of the scalar cascade process as a function of (top left) E_T^{miss} without any cuts applied, (top right) leading small R jet p_T without any cuts applied, (bottom) leading small R jet p_T with a 200 GeV E_T^{miss} cut applied. This is for the Scalar Cascades Benchmark SC1.	146

7.25	Hybrid Jet Reclustering vs. Standard Large R jet Reclustering performance comparison for Scalar Cascade Benchmark SC1: Fractions of exotic particles captured by (Top left) the reclustered jet that is most efficient at capturing exotic particles (Top right) the reclustered jet that is second most efficient at capturing exotic particles (Bottom left) the leading p_T reclustered jet (Bottom right) second leading p_T reclustered jet. A 200 GeV E_T^{miss} cut applied.	147
8.1	W production in association with a jet	152
8.2	Diboson production, where V can be a W boson or Z boson	152
8.3	Top quark pair production	153
8.4	Single Top Production	154
8.5	Log of E_T^{miss} ϕ distribution after bad tile module event veto and after applying tight cleaning. Full Run 2 dataset is used.	157
8.6	Discriminating Variables of Scalar Cascade Benchmark SC1. Depicted here are $\Delta\phi_{\text{min}}(j, E_T^{\text{miss}})$, $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$, and the number of tracks in the second leading reclustered jet.	161
8.7	Lead small R jet p_T with basic 200 GeV cuts applied.	165
9.1	Neuron structure Reference [106]	169
9.2	Architecture of a neural network with one or multiple hidden layers. This figure is a personal remake of a figure from Reference [107] . . .	170
9.3	Activation functions in ML. Taken from Reference [108]	171
9.4	Illustration of Φ providing a per particle mapping into latent space and F acting on the sum of the latent space representations to map out of latent space and provide a single output. Figure taken from Reference [112]	175
9.5	Particle Flow Network - DisCo NN Architecture	178
9.6	Variables inputted into event level tagger of NN for Scalar Cascades. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied to these variables. Data and Monte Carlo distributions are overlaid for the distributions of event shape variable C and the mass of the second leading reclustered jet.	186

9.7	Top plot: Comparison of Monte Carlo and data distributions of number of tracks in the second leading reclustered jet for Scalar Cascades. The comparisons are made for track multiplicities smaller than 80 tracks in order to avoid the signal region. This distribution is normalized. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied. Bottom plot: The Monte Carlo to Data ratio of the number of tracks in the second leading reclustered jet. This distribution was created from the normalized distributions in the top plot.	187
9.8	Top plot: Data and Monte Carlo comparison for the η distributions of the leading reclustered jets for Scalar Cascades. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the η of the leading reclustered jets for Scalar Cascades. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied to these variables.	188
9.9	Top plot: Data and Monte Carlo comparison for the η distributions of the second leading reclustered jets for Scalar Cascades. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the η of the second leading reclustered jets for Scalar Cascades. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied to these variables.	189
9.10	Kinematics of all tracks within the first and second leading reclustered jets for Scalar Cascades and Background with Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 applied	190
9.11	Kinematics of all small R jets in the event for Scalar Cascades and Background with Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 applied	191
9.12	ROC Curve of Background Rejection vs. Signal Efficiency for the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet.	192
9.13	Evolution of NN validation accuracy and training accuracy as a function of epochs for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet.	194
9.14	Evolution of NN validation loss and training loss as a function of epochs for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet.	195

9.15	NN prediction on signal events for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet. (A green circle highlights the bin containing the majority of the signal.)	196
9.16	NN prediction on background events for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet. (A green circle highlights the bin containing the majority of the background.)	197
9.17	Discriminating Variables for Noncommutative Black Hole and Non-minimal Dark Sector with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. Monte Carlo and Data distributions are overlaid for all variables except transverse sphericity, which serves as a decorrelated variable for a non-minimal dark sector benchmark.	200
9.18	Discriminating Variables for Noncommutative Black Hole and Non-minimal Dark Sector with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. Monte Carlo and Data distributions are overlaid for all variables except $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, which serves as a decorrelated variable for noncommutative black holes.	201
9.19	Discriminating Variables for Noncommutative Black Hole and Non-minimal Dark Sector with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied	203
9.20	Top plot: Data and Monte Carlo comparisons for the track multiplicity in the second leading reclustered jet with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the number of tracks in the second leading reclustered jet.	204
9.21	Top plot: Data and Monte Carlo comparisons for $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$	205
9.22	Top plot: Data and Monte Carlo comparisons for the transverse sphericity with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the transverse sphericity.	206

9.23	Kinematics of all tracks within the first and second leading reclustered jets for Noncommutative Black Hole and Nonminimal Dark Sector with Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts Applied	207
9.24	Kinematics of all tracks within the first and second leading reclustered jets for Noncommutative Black Hole and Nonminimal Dark Sector with Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts Applied	208
9.25	E_T^{miss} distribution of noncommutative black holes and non-minimal dark sectors with basic 200 GeV cuts applied.	209
9.26	Discriminating Variables for Noncommutative Black Hole with 500 GeV E_T^{miss} cut from Table 9.1 Applied.	212
9.27	Discriminating Variables for Noncommutative Black Hole with 500 GeV E_T^{miss} cut from Table 9.1 Applied.	213
9.28	Discriminating Variables for Noncommutative Black Hole with 500 GeV E_T^{miss} cut from Table 9.1 Applied.	215
9.29	Top plot: Comparison of Monte Carlo and data distributions for event shape variable C . This distribution is normalized. The 500 GeV Noncommutative Black Hole E_T^{miss} cuts from Table 9.1 have been applied. Bottom plot: The Monte Carlo to Data ratio of event shape variable C . This distribution was created from the normalized distributions in the top plot.	216
9.30	Top plot: Comparison of Monte Carlo and data distributions for the η distribution of the leading reclustered jet. This distribution is normalized. The 500 GeV Noncommutative Black Hole E_T^{miss} cuts from Table 9.1 have been applied. Bottom plot: The Monte Carlo to Data ratio of the η distribution of the leading reclustered jet. This distribution was created from the normalized distributions in the top plot.	217
9.31	Top plot: Comparison of Monte Carlo and data distributions for the distribution of $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$. This distribution is normalized. The 500 GeV Noncommutative Black Hole E_T^{miss} cuts from Table 9.1 have been applied. Bottom plot: The Monte Carlo to Data ratio of $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$. This distribution was created from the normalized distributions in the top plot.	218
9.32	H_T distribution of NMDS benchmark NMDS3	219
9.33	Discriminating Variables for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied.	222

9.34	Discriminating Variables for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied.	223
9.35	Discriminating Variables for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied.	225
9.36	Top plot: Comparison of Monte Carlo and data distributions for event shape variable C for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied. This distribution is normalized. Bottom plot: The Monte Carlo to Data ratio of event shape variable C . This distribution was created from the normalized distributions in the top plot.	226
9.37	Top plot: Comparison of Monte Carlo and data distributions for the η distribution of the leading reclustered jet for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied. This distribution is normalized. Bottom plot: The Monte Carlo to Data ratio of the η distribution of the leading reclustered jet. This distribution was created from the normalized distributions in the top plot.	227
9.38	Top plot: Comparison of Monte Carlo and data distributions for the η distribution of transverse sphericity for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied. This distribution is normalized. Bottom plot: The Monte Carlo to Data ratio of the transverse sphericity. This distribution was created from the normalized distributions in the top plot.	228
9.39	ROC Curve of Background Rejection vs. Signal Efficiency for NCBH benchmark NCBH4 with DisCo hyperparameter $\lambda = 2$ and 200 E_T^{miss} cut and basic cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss}	229
9.40	ROC Curve of Background Rejection vs. Signal Efficiency for NCBH benchmark NCBH1 with DisCo hyperparameter λ value of 2 and 200 GeV E_T^{miss} cut and basic cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss}	230
9.41	NN prediction for background events using NN for NCBH benchmark NCBH1 with a DisCo hyperparameter λ value of 2 and applying the 200 GeV E_T^{miss} cut and basic cuts from Table 8.6. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss}	231

9.42	NN prediction for signal events using NN for NCBH benchmark NCBH1 with a DisCo hyperparameter λ value of 2 and applying the 200 GeV E_T^{miss} cut and basic cuts from Table 8.6. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} .	232
9.43	ROC Curve of Background Rejection vs. Signal Efficiency using NN for NCBH benchmark NCBH1 with 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.	233
9.44	Evolution of accuracy of NN validation accuracy and training accuracy as a function of epochs for NN using NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.	234
9.45	Evolution of NN validation loss and training loss as a function of epochs for NN using NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.	235
9.46	NN prediction for background events using NN for NCBH benchmark NCBH1 with 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.	236
9.47	NN prediction for signal events using NN for NCBH benchmark NCBH1 with 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.	237
9.48	ROC Curve of Background Rejection vs. Signal Efficiency using NN for NMDS benchmark NMDS2 with 200 GeV E_T^{miss} cut and basic cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} . The DisCo hyperparameter λ value is 2.	238
9.49	NN prediction for background events using NN for NMDS benchmark NMDS2 with a 200 GeV E_T^{miss} cut and cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} . The DisCo hyperparameter λ value is 2.	239
9.50	NN prediction for signal events using NN for NMDS benchmark NMDS2 with a 200 GeV E_T^{miss} cut and cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} . The DisCo hyperparameter λ value is 2.	240

9.51	ROC Curve of Background Rejection vs. Signal Efficiency using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.	241
9.52	Evolution of NN validation accuracy and NN training accuracy as a function of epochs using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.	242
9.53	Evolution of accuracy of NN validation loss and NN training loss as a function of epochs using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.	243
9.54	NN prediction for background events using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.	244
9.55	NN prediction for signal events using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.	245
10.1	Example of an ABCD plane. Figure modified from Reference [115] . .	248
10.2	Example of splitting an ABCD plane into additional regions for nonclosure and validation studies. Figure modified from Reference [115]	251
10.3	MC Background ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis.	255
10.4	MC Signal ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis.	255
10.5	MC Background ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis. ABCD region definitions are drawn on plane.	256
10.6	MC Signal ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis. ABCD region definitions are drawn on plane. . . .	257

10.7	Top plot: Number of tracks in the second leading reclustered jet in Control Region B of the ABCD plane for the Scalar Cascade Analysis. Distribution is normalized. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in Control Region B of the ABCD plane. This ratio is created from the normalized distribution shown in the top plot.	259
10.8	Top plot: Number of tracks in the second leading reclustered jet in Control Region C of the ABCD plane for the Scalar Cascade Analysis. Distribution is normalized. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in Control Region C of the ABCD plane. This ratio is created from the normalized distribution shown in the top plot.	260
10.9	Top plot: Number of tracks in the second leading reclustered jet in Control Region D of the ABCD plane for the Scalar Cascade Analysis. Distribution is normalized. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in Control Region D of the ABCD plane. This ratio is created from the normalized distribution shown in the top plot.	261
10.10	Number of tracks in the second leading reclustered jet in Signal Region (Region A) of the ABCD plane for the Scalar Cascade Analysis. Comparison of distributions for MC background and MC signal. Data is absent. Distributions are normalized.	261
10.11	Top plot: number of tracks in the second leading reclustered jet for all regions of the ABCD plane (Region A, B, C, and D) combined for Monte Carlo. This is compared to the number of tracks in the second leading reclustered jet for the control regions (Regions B, C, and D) for data. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in the four regions of the ABCD plane for Monte Carlo vs. the three control regions of the ABCD plane for data. This ratio is created from the normalized distribution shown in the top plot. This is for the Scalar Cascade Analysis.	262
10.12	Top plot: NN output for all regions of the ABCD plane (Region A, B, C, and D) combined for Monte Carlo. This is compared to all control regions (Regions B, C, and D) for data. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in the four regions of the ABCD plane for Monte Carlo vs. the three control regions of the ABCD plane for data. This ratio is created from the normalized distribution shown in the top plot. This is for the Scalar Cascade Analysis.	263
10.13	Data ABCD plane with the signal Region blinded. This is for the Scalar Cascade Analysis. Full Run 2 dataset of 139 fb^{-1} used.	264

10.14	Nonclosure and Validation Regions of ABCD plane for MC background for Scalar Cascade analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.	265
10.15	Nonclosure and Validation Regions of ABCD plane for data for Scalar Cascade analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.	265
10.16	Plane showing NN output vs. $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for MC background for full range of NN output and full range of ϕ for NCBH Analysis of NCBH benchmark NCBH1	273
10.17	Plane showing NN output vs. $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for full range of NN output and range of ϕ . This is for the NCBH Analysis of NCBH benchmark NCBH1	274
10.18	1D Distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for MC background. This is for the NCBH Analysis of NCBH benchmark NCBH1	275
10.19	1D Distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for MC signal. This is for the NCBH Analysis of NCBH benchmark NCBH1	276
10.20	MC background ABCD plane as a function of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ and NN output. This is for the NCBH Analysis of NCBH benchmark NCBH1.	277
10.21	Signal ABCD plane as a function of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ and NN output. This is for the NCBH Analysis of NCBH benchmark NCBH1.	278
10.22	Top plot: Normalized distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ in Control Region B with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This ratio is created from the normalized distribution shown in the top plot. This is for the NCBH Analysis of NCBH benchmark NCBH1.	279
10.23	Top plot: Normalized distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ in Control Region C with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This ratio is created from the normalized distribution shown in the top plot. This is for the NCBH Analysis of NCBH benchmark NCBH1.	280
10.24	Top plot: Normalized distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ in Control Region D with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This ratio is created from the normalized distribution shown in the top plot. This is for the NCBH Analysis of NCBH benchmark NCBH1.	281
10.25	Nonclosure and Validation Regions of ABCD plane for MC background for NCBH analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2	282

10.26	Nonclosure and Validation Regions of ABCD plane for data for NCBH analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2	283
10.27	NN output vs. transverse sphericity for MC background for full NN output range and full transverse sphericity range for NMDS analysis of benchmark NMDS3.	288
10.28	NN output vs. transverse sphericity for MC signal for full NN output range and full transverse sphericity range for NMDS analysis of benchmark NMDS3.	289
10.29	MC background ABCD plane as a function of transverse sphericity and NN output for NMDS analysis of NMDS benchmark NMDS3. . .	290
10.30	MC signal ABCD plane as a function of transverse sphericity and NN output for NMDS analysis of NMDS benchmark NMDS3.	291
10.31	Top plot: Normalized distributions of transverse sphericity in Control Region B of NMDS ABCD plane with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of transverse sphericity in control Region B. This ratio is created from the normalized distribution shown in the top plot. This is for the NMDS Analysis of NMDS benchmark NMDS3	292
10.32	Normalized distributions of transverse sphericity in Control Region C of NMDS ABCD plane with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of transverse sphericity in control Region C. This ratio is created from the normalized distribution shown in the top plot. This is for the NMDS Analysis of NMDS benchmark NMDS3	293
10.33	Normalized distributions of transverse sphericity in Control Region D of NMDS ABCD plane with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of transverse sphericity in control Region D. This ratio is created from the normalized distribution shown in the top plot. This is for the NMDS Analysis of NMDS benchmark NMDS3	294
10.34	Nonclosure and Validation Regions of ABCD plane for MC background for NMDS analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.	295
10.35	Nonclosure and Validation Regions of ABCD plane for data for NMDS analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.	296

11.1	Distribution displaying the Gaussian nature the expected (6.6) data events in the signal region of the Scalar Cascades ABCD plane. This serves as sufficient evidence for the number of events in the signal region to be approximated as a Gaussian instead of a Poisson distribution. This plot was created by drawing random values for the number of events in regions B, C, and D assuming Gaussian distributions, calculating the resulting value of the number of expected events in region A, and producing the probability density distribution of the number of expected events in region A	306
11.2	Scalar Cascade Exclusion Limit. Result obtained via use of the Frequentist method.	309
11.3	NCBH Benchmark NCBH1 exclusion limit. Result obtained via use of the Frequentist method.	311
11.4	NMDS Benchmark NMDS3 exclusion limit. Result obtained via use of the Frequentist method.	312
13.1	Data 2017. Distribution of the ϕ_i value of E_T^{miss} in 4 cases: (top left) without the bad tile module event veto and without tight cleaning, (top right) without the bad tile module event veto and with tight cleaning, (bottom left) with the tile module event veto and without tight cleaning, and (bottom right) with the bad tile module event veto and with tight cleaning	317
13.2	Cross-check rerunning exclusion limits for NCBH Benchmark NCBH1 with nonclosure uncertainty of 9.35%	333
13.3	Cross-check rerunning exclusion limits for NCBH Benchmark NCBH1 with nonclosure uncertainty of 16.19%	334
13.4	Cross-check rerunning exclusion limits for NMDS Benchmark NMDS3 with nonclosure uncertainty of 52%	335

List of Tables

3.1	Inner Detector Parameters. Table taken from Reference [69].	38
3.2	Parameters of the Electromagnetic Calorimeter. Table taken from Reference [69].	41
3.3	Parameters of various components of the ATLAS detector. Table taken from Reference [69].	44
6.1	Range of $\sqrt{\theta}M_D$ in units of M_D that can be probed at the LHC. The $(r_S)_{\min}$ is the minimum Schwarzschild radius needed for black hole formation. Table modified from Reference [91]	76
6.2	Limits on M_D from ATLAS search for ADD-type extra dimensions via monojet final state with full Run 2 data at 139 fb^{-1} . Figure taken from Reference [7]	78
6.3	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 2$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Blue: benchmarks that yield under 5 events at 139 fb^{-1} at 13 TeV. Purple: out of reach (OOR) benchmarks.	80
6.4	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 3$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Blue: benchmarks that yield under 5 events at 139 fb^{-1} at 13 TeV. Purple: out of reach (OOR) benchmarks.	81
6.5	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 4$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Blue: benchmarks that yield under 5 events at 139 fb^{-1} at 13 TeV. Purple: out of reach (OOR) benchmarks.	82

6.6	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 5$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOR) benchmarks.	83
6.7	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 5$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOR) benchmarks.	84
6.8	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 6$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOR) benchmarks.	86
6.9	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 7$ Extra Dimensions. The color-coding scheme is defined as follows. Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOR) benchmarks.	88
6.10	Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 7$ Extra Dimensions. The color-coding scheme is defined as follows. Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOR) benchmarks.	89
6.11	Parameters and Cross Sections (σ) for Noncommutative Black Hole Samples	90
6.12	Parameters and Cross Sections (σ) for Non-minimal Dark Sector Signal Samples	92
6.13	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS4	95
6.14	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS4	96
6.15	Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS4	96

6.16	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS2	98
6.17	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS2	98
6.18	Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS2	98
6.19	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS3	99
6.20	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS3	99
6.21	Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS3	99
6.22	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS5	101
6.23	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS5	102
6.24	Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS5	102
6.25	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS1	103
6.26	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS1	103
6.27	Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS1	104
6.28	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS6	104

6.29	Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS6	105
6.30	Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS6	105
6.31	Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of Standard Model particles (SMP), number of stable particles (StaP), and cross sections (σ) for 2 Body Stable Cascade Benchmarks. The benchmarks chosen for the truth-level analysis are highlighted.	108
6.32	Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of Standard Model particles (SMP), and cross sections (σ) for 2 Body Unstable Cascade Benchmarks. The benchmark chosen for the truth-level analysis is highlighted.	109
6.33	Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of Standard Model particles (SMP), number of stable particles (StaP), and cross sections (σ) for 3 Body Stable Cascade Benchmarks. The benchmarks chosen for the truth-level analysis are highlighted.	110
6.34	Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of stable particles (StaP), and cross sections (σ) for 3 Body Stable Chain Benchmarks up to 15 Scalars. Each benchmark yields 2 stable scalars.	112
6.35	Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), decay particle multiplicity, and cross sections (σ) for 3 Body Stable Chain Benchmarks up to 25 Scalars. Each benchmark yields 2 stable scalars. The benchmark chosen for the truth-level analysis is highlighted.	113
6.36	Scalar Cascade Signal Samples Names	113
8.1	Good Run Lists for Run 2 data-taking	149
8.2	$E_{\text{T}}^{\text{miss}}$ triggers by year of Run 2 data-taking	150
8.3	Standard Model MC Background Processes used in these searches and their corresponding DSIDs	151
8.4	Scalar Cascades cuts prior to Neural Network Cuts	162

8.5	MC Event Yields for Scalar Cascades cuts prior to the SC Neural Network Cuts defined in Table 8.4. Background composition, signal significance, and Tile Efficiency Correction resulting from these cuts are also included.	163
8.6	Most basic set of cuts for NCBH and NMDS prior to Neural Network Cuts. Some NCBH and NMDS benchmarks have additional cuts prior to the NN, which will be discussed in the next chapter	164
8.7	MC Event Yields for Noncommutative Black Holes, Nonminimal Dark Sectors, and Scalar Cascades and Background processes after applying the basic NCBH and NMDS cuts defined in Table 8.6. Background composition, signal significance, and tile efficiency correction resulting from these cuts are also included.	166
9.1	Selection cuts for analysis of Noncommutative Black Hole Benchmark NCBH1 with a 500 GeV E_T^{miss} cut	210
9.2	Background yield, signal yield, background composition, significance increases, and tile efficiency correction at 500 GeV E_T^{miss} cut for Noncommutative Black Holes	211
9.3	Selection cuts for analysis of Non-minimal Dark Sector benchmark NMDS3 with a 500 GeV E_T^{miss} and 1400 GeV HT cut.	220
9.4	MC background yields, MC background composition, Signal Yields, Significances, Tile Efficiency Correction with 500 GeV E_T^{miss} and 1400 GeV HT cut.	221
10.1	Definitions of regions in the ABCD plane for the Scalar Cascade Analysis	256
10.2	Monte Carlo Yields, Data Yields, and Signal Yields for the A, B, C, and D Regions of the ABCD plane for the Scalar Cascade Analysis. Column “% of Total Signal” gives the percent of total signal events in the ABCD plane that are contained in a given region. The column “ $100 \times \frac{\text{Signal}}{\text{Background}}$ ” turns into a percent the fraction of signal events in a given region with respect to the number of background events in the region. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	257

10.3	Monte Carlo Yields and Data Yields in subvalidation regions A_{nc} , B_{nc} , C' , D' used to test for nonclosure. The definitions of A_{nc} , B_{nc} , C' , D' are found in Figure 10.2. Monte Carlo and Data Nonclosure Uncertainties (MC NC Uncertainty, Data NC Uncertainty), Pearson Correlation Coefficients for MC (MC Corr. Coeff) and Data (Data Corr. Coeff), expected MC and Data event yields (Exp MC events and Exp Data yields) are also included. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN. This is for the NCBH Analysis.	266
10.4	Monte Carlo Yields and Data Yields in subvalidation regions A_V , B_V , $A_{nc} + D'$, and $B_{nc} + C'$ used to validate nonclosure uncertainty. The definitions of A_V , B_V , $A_{nc} + D'$, and $B_{nc} + C'$ are found in Figure 10.2. Monte Carlo and Data Nonclosure Uncertainties (MC NC Uncertainty, Data NC Uncertainty), Pearson Correlation Coefficients for MC (MC Corr. Coeff) and Data (Data Corr. Coeff), expected MC and Data event yields (Exp MC events and Exp Data events) are also included. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	267
10.5	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D in the ABCD plane used for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	268
10.6	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation.	269
10.7	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	269
10.8	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	270

10.9	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	270
10.10	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	270
10.11	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	271
10.12	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	271
10.13	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	272
10.14	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.	272
10.15	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D of the ABCD plane used for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is abbreviated as ‘ ϕ ’ and the NN output is denoted as NN.	277

10.16	Monte Carlo Yields and Signal Yields for the A, B, C, and D Regions of the ABCD plane for the NCBH Analysis of NCBH benchmark NCBH1. Column “% of Total Signal” gives the percent of total signal events in the ABCD plane that are contained in a given region. The column “ $100 \times \frac{\text{Signal}}{\text{Background}}$ ” turns into a percent the fraction of signal events in a given region with respect to the number of background events in the region. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is abbreviated as ‘ ϕ ’ and the NN output is denoted as NN.	278
10.17	Monte Carlo Yields, Data Yields, Monte Carlo and Data Nonclosure Uncertainties (MC NC Uncertainty, Data NC Uncertainty), Pearson Correlation Coefficients for MC (MC Corr. Coeff) and Data (Data Corr. Coeff), expected MC and Data event yields (Exp MC events and Exp Data yields) for regions of an ABCD plane used to calculate the nonclosure uncertainty for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.	284
10.18	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions of an ABCD plane used to validate the nonclosure uncertainty for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.	284
10.19	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.	285
10.20	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.	286
10.21	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.	286
10.22	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.	287

10.23	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of the ABCD plane for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN.	289
10.24	Monte Carlo Yields and Signal Yields for the A, B, C, and D Regions of the ABCD plane for the NMDS Analysis of NMDS benchmark NMDS3. Column “% of Total Signal” gives the percent of total signal events in the ABCD plane that are contained in a given region. The column “ $100 \times \frac{\text{Signal}}{\text{Background}}$ ” turns into a percent the fraction of signal events in a given region with respect to the number of background events in the region. The transverse sphericity is abbreviated as ‘S’ and the NN output is denoted as NN.	291
10.25	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions in the ABCD plane used calculate the nonclosure uncertainty. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN. . . .	297
10.26	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions in the ABCD plane used to validate the nonclosure uncertainty. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN . . .	297
10.27	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN . . .	298
10.28	Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN . . .	298
10.29	MC and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN . . .	299

10.30	MC and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘ S ’ and the NN output is denoted as NN . . .	299
13.1	Systematic values (in percentages) in order from greatest to lowest (in percentages) for non-minimal dark sector benchmark NMDS3 - Part 1 of 3 Tables	318
13.2	Systematic values (in percentages) in order from greatest to lowest (in percentages) for non-minimal dark sector benchmark NMDS3. Part 2 of 3 Tables. Systematics are continued from Table 13.1	319
13.3	Systematic values (in percentages) in order from greatest to lowest (in percentages) for non-minimal dark sector benchmark NMDS3. Part 3 of 3 Tables. Systematics are contained from Table 13.2	320
13.4	Systematic values (in percentages) in order from greatest to lowest (in percentages) for noncommutative black hole benchmark NCBH1. Part 1 of 2 Tables	321
13.5	Systematic values (in percentages) in order from greatest to lowest (in percentages) for noncommutative black hole benchmark NCBH1. Part 2 of 2 Tables, continued from Table 13.4	322
13.6	Systematic values (in percentages) in order from greatest to lowest (in percentages) for Scalar Cascade Benchmark SC1. Part 1 of 3 Tables .	323
13.7	Systematic values (in percentages) in order from greatest to lowest (in percentages) for Scalar Cascade Benchmark SC1. Part 2 of 3 Tables, continued from Table 13.6	324
13.8	Systematic values (in percentages) in order from greatest to lowest (in percentages) for Scalar Cascade Benchmark SC1. Part 3 of 3 Tables, continued from Table 13.7	325
13.9	Cross sections, filter efficiencies, and k Factors for MC background samples corresponding to Run 2 MC Campaign MC16a	330
13.10	Cross sections, filter efficiencies, and k Factors for MC background samples corresponding to Run 2 MC Campaign MC16d	331
13.11	Cross sections, filter efficiencies, and k Factors for MC background samples corresponding to Run 2 MC Campaign MC16e	332
13.12	Trigger efficiencies for Scalar Cascade Benchmark SC4 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	336

13.13	Trigger efficiencies for Scalar Cascade Benchmark SC2 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	337
13.14	Trigger efficiencies for Scalar Cascade Benchmark SC6 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	338
13.15	Trigger efficiencies for Scalar Cascade Benchmark SC5 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	339
13.16	Trigger efficiencies for Scalar Cascade Benchmark SC1 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	340
13.17	Trigger efficiencies for Scalar Cascade Benchmark SC3 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	341
13.18	Trigger efficiencies for noncommutative black hole benchmark NCBH1 for Run 2 data-taking years: 2015, 2016	342
13.19	Trigger efficiencies for noncommutative black hole benchmark NCBH1 for Run 2 data-taking years: 2017, 2018	343
13.20	Trigger efficiencies for noncommutative black hole benchmark NCBH4 for all Run 2 data-taking years: 2015, 2016	344
13.21	Trigger efficiencies for noncommutative black hole benchmark NCBH4 for all Run 2 data-taking years: 2017, 2018	345
13.22	Trigger efficiencies for NMDS benchmark NMDS1 for all Run 2 data- taking years: 2015, 2016, 2017, 2018	346
13.23	Trigger efficiencies for NMDS benchmark NMDS2 for all Run 2 data- taking years: 2015, 2016, 2017, 2018	347
13.24	Trigger efficiencies for non-minimal dark sector benchmark NMDS5 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	348
13.25	Trigger efficiencies for non-minimal dark sector benchmark NMDS4 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	349
13.26	Trigger efficiencies for non-minimal dark sector benchmark NMDS6 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	350
13.27	Trigger efficiencies for non-minimal dark sector benchmark NMDS3 for all Run 2 data-taking years: 2015, 2016, 2017, 2018	351

Chapter 1

Introduction

The Standard Model (SM) of particle physics is a quantum field theory that unifies classical field theory, quantum mechanics, and special relativity. It serves as the most accurate description of particle physics to date, describing all known particles and their interactions with three of the four forces: the electromagnetic force, the weak force, and the strong force. However, despite its success, the Standard Model has some notable shortcomings. For example, it does not account for gravity. Additionally, it does not describe, on a particle level, dark matter or dark energy, both of which are known to exist due to astronomical observations. The inability to explain the discrepancy in strength between the weak force and the gravitational force, wherein gravity is a full 24 orders of magnitude weaker than the weak force, is deemed the hierarchy problem and remains a flagrant shortcoming of the Standard Model.

The Large Hadron Collider (LHC) in Geneva, Switzerland collides two beams of protons to create proton-proton collisions. These collisions are recorded by detectors. One such detector is the ATLAS experiment. These collisions, of either quarks or gluons from the protons, result in the generation of additional particles. These particles and the signatures that they leave in the detector can be analyzed to search for new physics. This thesis presents novel searches for three possible models of new physics that have never been searched for by any experiment: scalar cascades, non-minimal dark sectors, and noncommutative black holes. These models could address some of the unresolved problems of the Standard Model: the incorporation of gravity in Quantum Field Theory, the nature of dark matter, and the hierarchy problem.

The discovery of the Higgs Boson in 2012 [1] [2], a mere two years after the LHC

began, gave hope to particle physicists that more discoveries were just around the corner. However, over a decade later, hundreds of searches have failed to discover new physics. There are several possible explanations for this. New physics could be beyond the energy reach of the LHC (or any future collider). Alternatively, the LHC may require more data to uncover new physics. More excitingly, it could be that new physics has already been produced at the LHC, but has been hiding in signatures different than those we have used in searches. The last possibility served as the inspiration for the development of the scalar cascade model that this thesis targets. With no evidence of new physics, the goal was to target unexplored phase space.

The scalar cascade models used in this thesis were designed specifically for this analysis in collaboration with the theorists from [3]. The couplings in the models in this thesis were engineered such that the internal structure of the models from [3] was broken so as to force and allow maximum multiplicity of final state particles. This arose from the author of this thesis desiring to design a novel final state and, thus, a novel search targeting challenging phase space. The goal was to design a final state with at least 30 soft (low transverse momentum) jets (the collider object definition of the collimated decay products of quarks and gluons) and missing transverse energy (non-interacting energy in the detector), so as to have the signature fail all existing data filtering mechanisms (triggers), except the trigger that utilizes missing energy (non-interacting particles), and have its analysis and potential discovery be missed by the collaboration.

The incentive behind targeting a minimum of 30 soft jets was that soft jets can blend into Standard Model background processes. If the desired signature looks like background, perhaps new analysis strategies have to be developed in order to identify the signal. The reason behind trying to have the scalar cascades fail all triggers except the missing energy trigger is so the scalar cascades would avoid all constraints posed by existing multijet searches (searches requiring many jets). Similarly, the goal was to avoid any energetic leptons appearing in the model, so as to evade limits from processes with an energetic lepton. Of course, a trigger is needed in order to save the event and be able to have data to analyze. Therefore, the goal was to have the scalar cascades produce enough missing energy in order to survive the filtering of the corresponding trigger.

The detailed path to discovering which set of parameters and which class of scalar cascades yielded a novel and striking signature, and how the most striking signature arose in the most counterintuitive region of phase space, will be described

systematically in a later chapter. A brief summary is provided here. The scalar cascade model is a model of new scalars (spin-0 particles). Each scalar decays into additional scalars, which then decay into more scalars. The last step of the cascade results in the generation of a lightest scalar, which is either stable or unstable. Due to the nature of the model, in which there are a number of scalars simultaneously decaying into lightest stable scalars, there can be many lightest stable scalars. The lightest stable particle would pass through the detector undetected and would be registered as missing energy. The intuitive assumption is that the heavier the mass of the lightest scalar, the higher the missing energy. By having many lightest scalars, however, the total amount of missing energy is not equal to the sum of the masses of the lightest stable scalars, as there would be some cancellation of the missing energy due to a number of stable particles moving into opposite directions. Setting the mass of the lightest stable scalar to be large, while requiring high particle multiplicity, however, may cause the mass of the heaviest scalar to be too large to yield cross sections suitable for discovery. The most counter-intuitive region of phase space for finding high missing energy is lightest stable scalars with very low masses (roughly 0.5 GeV). As the missing energy is expected to come from the lightest scalar, if the scalar is very light, one would not expect substantial missing energy. The benchmark used in the analysis yields 54 stable particles. Naively, this would be expected to yield 27 GeV in missing energy. However, enough missing energy (for example, at least 200 GeV) is produced so as to allow for acceptance by the missing energy trigger. A possible explanation could be the generation of tracks that are too soft to be reconstructed (meaning below 500 MeV), which would then appear as missing energy. The final state yielded is one of very high track multiplicity and missing energy.

Six different classes of scalar cascades were studied in order to select a final benchmark for the analysis. The benchmark with the highest particle multiplicity, highest missing energy, and thereby highest acceptance by the missing energy trigger, was chosen. Due to the number of scalars in the model, all scalar cascade benchmarks charted in this thesis have evaded exclusion limits due to their small cross sections and low trigger efficiencies. If, however, the scalar cascade model was forced to have very few scalars, and therefore very small particle multiplicity, constraints may be imposed by dijet searches [4] (searches for two jets). This thesis, however, focuses on high multiplicity states.

The scalar cascade model could be a model of baryogenesis. Baryogenesis refers to the creation of an excess of baryons over antibaryons. Additionally, the scalar cascade

model can be a model of extra dimensions. Scalar cascades share the same final states (high particle multiplicity) as fermionic cascades. (Fermions are spin half particles.) Light fermions arise in string theory. String theory is the theory that states that nature's fundamental building blocks are strings, rather than particles. These strings vibrate. Different particles are associated with different vibrational patterns. For example, the vibrational pattern associated with an electron is different from the vibrational pattern associated with the down quark. One vibrational pattern in string theory is associated to the graviton, which is a particle that carries the gravitational force. String theory, therefore, provides a way to explain gravity, a feat that the Standard Model does not accomplish. String theory requires extra dimensions and would allow for the unification of quantum mechanics and general relativity. By virtue of having the same final states as fermionic cascades, scalar cascades would be a signature of string theory. Additionally, scalar cascades are a model of hidden sectors. Hidden sectors (or dark sectors) [5] [6] are quantum fields and their corresponding particles that do not interact strongly with the Standard Model and remain hidden. Hidden sectors are dark matter models.

The two other models that the analyses in this thesis targets are noncommutative black holes and non-minimal dark sectors. These are models of extra dimensions, string theory, and dark matter. They are characterized by final states with high particle multiplicity, though not as high as those seen in scalar cascades. These particles form jets. Noncommutative black holes are of interest because they have a softer transverse momentum spectrum than the highly energetic black holes that ATLAS has previously searched for. The constraints on noncommutative black holes come primarily from the model-independent limits on the Planck scale in higher dimensions. The Planck scale is the scale at which the quantum effects of gravity become important. The limits on the Planck scale in 2, 3, 4, 5, 6, and 7 extra dimensions are 11.2, 8.5, 7.1, 6.4, 5.9 TeV, and 0.8 TeV respectively [7] [8]. A future chapter uses these limits to derive the viable phase space for noncommutative black holes. The non-minimal dark sector model has 6 free parameters. A single benchmark for non-minimal dark sectors can require the generation of hundreds or even thousands of subprocesses. As it is not possible to map out the entire unexcluded phase space for non-minimal dark sectors, 6 benchmarks were designed to ensure evasion of model-independent limits imposed by the mono-jet analysis [7] (a search requiring an energetic jet) and multijet analyses [9].

Chapter 2 introduces the theory that governs particle physics, the Standard Model, describes its limitations, and proposes potential solutions. Three models that propose

the existence of more than four dimensions are introduced: noncommutative black holes, scalar cascades, and non-minimal dark sectors. The rest of the thesis details dedicated and novel searches for these new physics models. These searches will target final states with many particles, particularly targeting jets. The author proposed these searches in ATLAS and served as the sole analysis contact. This work is not yet published.

Chapter 3 introduces the accelerator at which theoretical models proposing solutions to the unresolved problems of the Standard Model can be tested: the Large Hadron Collider (LHC). The chapter details how the LHC works and introduces the experiments it provides collisions for. Special emphasis is given to the ATLAS experiment as this thesis uses ATLAS data to perform each search. Each subdetector of the ATLAS experiment is described and a description of the trigger system, the mechanism by which ATLAS records and filters its data, is provided.

Chapter 4 introduces event reconstruction within ATLAS, the process of transforming raw and uncalibrated hits across a number of detector subsystems into meaningful reconstructed physics objects. It describes the advanced algorithms that are used to identify all stable Standard Model particles. Special emphasis is placed on the algorithms used to reconstruct jets. Two main types of jets are described: small radius jets (small R jets) and large radius jets (large R jets). The concept of jet substructure, the internal structure of large R jets, is introduced as a means of distinguishing signal from background. The analysis strategy for all three models centers on large R jets and jet substructure.

Chapter 5 describes calibrations derived by the author for jets in the upgraded ATLAS trigger system for Run 3 (the higher energy LHC run). These jets are called jFEX jets. This chapter describes why calibrations are needed, discusses how the optimal calibration was chosen, and shows the improvements afforded by the calibration. In order to qualify for ATLAS authorship, each member of ATLAS must complete a service task, or a qualification task. The jFEX studies served as the author's qualification task. These studies are for use for Run 3 and, therefore, are not used in the Run 2 searches presented in this thesis.

In Chapter 6, a set of benchmarks is selected for each model. Though no experiment has ever searched for noncommutative black holes, scalar cascades, or non-minimal dark sectors, model-independent limits from other ATLAS searches are taken into account in order to design searches that target unexplored phase space. This chapter determines the unexcluded and potentially reachable phase space for each model.

Chapter 7 presents 2D representations of the collisions, called event displays, to depict how final state particles are spread throughout the detector and give insight into the number of jets each exotic model makes. The event displays show that each exotic model exhibits jet substructure. This chapter presents studies that demonstrate that the thresholds imposed by standard ATLAS reconstruction on large R jets are often insufficient to capture the exotics process, and details the design and optimization of a novel large R jet reconstruction algorithm as a remedy.

Chapter 8 details the event selection for each search. It also describes the simulated Monte Carlo samples used to model Standard Model background processes and the exotic processes and lists the datasets used to analyze Run 2 data. It quantifies the background rejection afforded by the initial cuts.

Chapter 9 describes how machine learning is used for signal region definition and background estimation in each search. It introduces the machine learning architecture, which relies on permutation invariant neural networks called particle flow networks. Track information from the most energetic and second most energetic large R jets is provided to the neural network so that the jet substructure can be learned and used to distinguish signal from background. Two decorrelated variables, the NN output and another variable that discriminates signal from background, are used for the background estimation of the Standard Model background processes.

Chapter 10 describes the derivation of the data-driven background estimation via the ABCD background estimation method. A plane is created using the 2 decorrelated variables from the previous chapter and four quadrants are defined. The background contribution in the signal region is determined from the data counts in the other 3 quadrants, called control regions. The background and data distributions are compared in the control regions to ensure that the neural network is generalizing well and behaving as expected. The chapter then describes the expected event yield uncertainties on the ABCD plane, providing an expected background estimate with uncertainties.

Chapter 11 introduces the statistical analysis performed to estimate the search sensitivity on the exotic models. The systematic uncertainties incorporated in the analysis are discussed.

Chapter 12 presents the conclusion of the three searches.

Chapter 2

Theoretical Background

This chapter begins with an introduction to the theoretical framework that governs particle physics: the Standard Model of particle physics. The fundamental particles and forces of the Standard Model are introduced and the limitations of the model are described. The chapter then describes theories that could provide solutions to the limitations of the Standard Model. Some of these theories contain extra dimensions, namely additional dimensions to the four dimensions (3 spatial and 1 time) that the Standard Model predicts. The chapter describes the three models that the searches in this thesis target: noncommutative black holes, scalar cascades, and non-minimal dark sectors. No experiment has ever searched for these models. As noted in the introduction chapter, the scalar cascade models were designed specifically for this analysis in collaboration with theorists. The commonality between all three models is that they yield a similar experimental signature: final states with many particles. This similarity will be exploited throughout the thesis to design similar, though distinct, search strategies for each model.

2.1 The Standard Model

The Standard Model (SM) of particle physics is a quantum field theory that unifies classical field theory, quantum mechanics, and special relativity. It serves as the most accurate description of particle physics to date, describing all known particles and their interactions with three of the four forces: the strong force, electromagnetic force, and weak force.

The Standard Model describes, as seen in Figure 2.1, 6 quarks and 6 leptons, totaling 12 spin half particles called fermions. The four integer spin gauge bosons predicted by the Standard Model are gluon, photon, Z boson, and W boson. A scalar is a spin 0 particle. The Higgs Boson is a scalar particle. Fermions serve as the building blocks of everyday matter, while bosons serve as force carriers, or a particle that mediates an interaction - a force - between other particles. As seen in Figure 2.1, all particles have a characteristic mass, electric charge, and spin. For each particle, the Standard Model also describes an antiparticle. Though an antiparticle has the same mass as its particle counterpart, it has opposite electric charge. The gluon, photon, and Z boson are their own antiparticles. These particles are electrically neutral. In addition to being electrically neutral, the photon is massless.

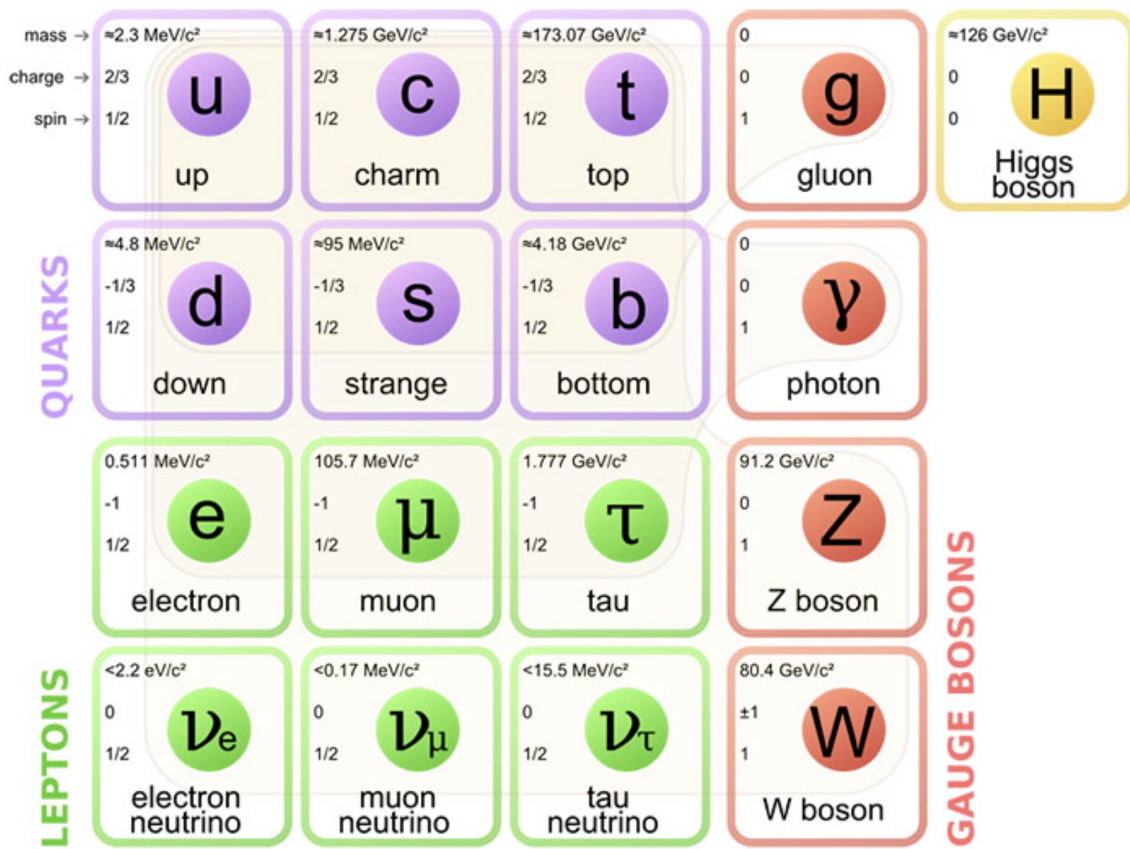


Figure 2.1 The 17 particles of the Standard Model of Particle Physics. Figure taken from [10].

There are three generations of fermions. Each generation contains two quarks and two leptons. The generations are ordered by increasing mass such that the first generation is the lightest, while the third generation is the heaviest. The first generation of fermions consists of the down quark, up quark, electron, and electron

neutrino. The second generation of fermions consists of the strange quark, charm quark, muon, and muon neutrino. The third generation of fermions consists of the bottom quark, the top quark, the tau, and the tau neutrino. Second and third generation particles decay to first generation particles due to the mass hierarchy, which explains why the first generation particles are the particles that constitute everyday matter. Each fermion is characterized by a flavor. The electron, muon, and tau have electric charges of $-1e$, where e is the magnitude of the electric charge of the electron. Each charged lepton has a neutrino with no charge. The up quark, charm quark, and top quark all have an electric charge of $+2/3e$, while the down quark, strange quark, and bottom quark all have an electric charge of $-1/3e$.

The electromagnetic force is a force that occurs between electrically charged particles and is mediated by the photon. The range of the electromagnetic force is infinite and the relative strength of the electromagnetic force is 10^{-2} times smaller than that of the strong force, which will be discussed shortly.

The weak force via the charged current changes the flavor of quarks, such that a down quark becomes an up quark or an up quark becomes a down quark. Consequently, the weak force allows a neutron to turn into a proton, an electron and an anti-neutrino. (Protons are comprised of two up quarks and one down quark, while a neutron is comprised of one up quark and two down quarks). The range of the weak force is 10^{-18} m, which is on the subatomic scale. The strength of the weak force is 10^{-6} times that of the strong force. The W^- , W^+ , and Z bosons serve as mediators of the weak force. The weak force and electromagnetic force are predicted to be the same force at an energy scale of 246 GeV according to the unified theory of electroweak interactions.

The strong force is the force that both confines quarks into hadrons (ie: protons and neutrons) and binds protons and neutrons to form atomic nuclei. As indicated earlier, it is 100 times greater than the electromagnetic force and 10^6 times stronger than the weak force. The range of the strong force is 10^{-15} m. The strong force is mediated by the gluon. The field theory that describes the strong interaction is called Quantum Chromodynamics (QCD). The strong force does not act on leptons, but instead only acts on quarks. This is because quarks and gluons alone have color and the strong force only acts between particles that have color. There are three types of color: blue, green, and red. Color carrying particles, as noted earlier, are confined into hadrons. This is because the total color charge must be zero. In the event that a colored particle is separated from a hadron, other color carrying objects are produced in order to ensure that the total color charge is zero. This showering

of particles, or collimated spray of particles, is called a jet. Jets will be described in detail in the subsequent chapters.

The Standard Model, in its original formalism, predicted all particles to be massless. This was in stark contrast to observation. In 1964, Peter Higgs, François Englert, and Robert Brout independently proposed a solution. Particles could be given mass via the coupling of massive Standard Model particles to the Higgs field, where the Higgs is a scalar field. The Higgs boson was discovered in 2012 at the Large Hadron Collider by the ATLAS and CMS collaborations [1] [2].

In addition to the most recent discovery of the Higgs Boson, the Standard Model has a rich history of achieving validation through experimental observation. The up quark, strange quark, down quark, electron, electron neutrino, muon, and muon neutrino were discovered in the 1960s while the charm quark was discovered in 1974 [11] [12]. The tau was discovered in 1975 [13], the bottom quark was discovered in 1975 [14], the W and Z bosons were discovered in 1983 [15] [16], the top quark was discovered in 1995, and the tau neutrino was discovered in 2000 [17] [18]. As noted previously, the Higgs boson was discovered in 2012 [1] [2].

However, despite the success of the Standard Model, it has some notable shortcomings. While it successfully describes the electromagnetic force, the weak force, and the strong force, the Standard Model does not account for gravity. The Standard Model predicts neutrinos to be massless in contrast of observations indicating a non-zero mass [19]. Additionally, the Standard Model does not describe, on a particle level, dark matter or dark energy, both of which are known to exist due to astronomical observations. Astrophysical experiments have indicated that roughly only 5% of the universe consists of observable matter, while roughly 26% consists of dark matter and roughly 69% consists of dark energy. Dark matter is a non-interacting form of matter that is believed to be responsible for the structure of the universe, gravitational lensing, and the movement of galaxies [20]. Dark energy neither emits or interacts with light and is believed to stretch the fabric of spacetime. Another shortcoming of the Standard Model is the inability to explain the abundance of matter over antimatter. The inability to explain the discrepancy in strength between the weak force and the gravitational force, wherein gravity is a full 24 orders of magnitude weaker than the weak force, is deemed the hierarchy problem and remains a flagrant shortcoming of the Standard Model. The searches in this thesis seek new fermions and new scalars beyond that which the Standard Model predicts. These new fermions and scalars seek to address some of the unresolved problems of the Standard Model: the incorporation of gravity in Quantum Field

Theory, the nature of dark matter, and the hierarchy problem.

2.2 Novel searches for String Theory

The sections below will provide the theoretical background for the three models targeted in the novel searches in this thesis: scalar cascades, non-minimal dark sectors, and noncommutative black holes. These models can be searched for at the Large Hadron Collider (LHC) with the ATLAS experiment. More detail on both the LHC and the ATLAS experiment will be given in the next chapter. However, it is useful to note here that the LHC collides two beams of protons to create proton-proton collisions. These collisions, of either quarks or gluons from the protons, result in the generation of additional particles. These particles and the signatures that they leave in the detector can be analyzed to search for new physics. The models considered in this thesis predict final states with many particles. Specifically, these models predict the formation of jets, which were introduced earlier as collimated sprays of particles originating from the decay of quarks and gluons. Scalar cascades, non-minimal dark sectors, and noncommutative black holes are all models of extra dimensions and serve as signatures of string theory. Additionally, they all are dark matter models. String theory, introduced earlier in the Introduction chapter, is the theory that states that nature's fundamental building blocks are strings, rather than particles.

2.3 ADD-type extra dimensions

The ADD model of extra dimensions was developed by Dimopoulos, G. Dvali, and N. Arkani-Hamed. It proposes extra dimensions that are large and flat. The ADD model is an example of a bulk-brane model. In a bulk-brane model, the brane (which is a term that originated from the word membrane) is a 4D (3 spatial dimensions and 1 time dimension) hyperspace that lives in a higher dimensional spacetime called the bulk. In the ADD model of extra dimensions, all Standard Model fields are confined to the 4D brane. However, gravity alone is believed to propagate in more than four dimensions (into the higher dimensional bulk). An important term to introduce is the Planck scale. The Planck scale is the scale at which the quantum effects of gravity become important. In four dimensions, the Planck scale is believed to be at 10^{19} GeV. If more than four dimensions exist, the Planck scale can be

lowered, perhaps even to the TeV scale. In the ADD model, extra dimensions are compactified on an n -dimensional torus, where n is the number of extra dimensions. The radius of the torus upon which the extra dimensions are compactified is called the compactification radius, which is denoted as R_c .

In ADD-type extra dimensions, the fundamental Planck scale in higher dimensions, M_D , is related to the 4D Planck scale, M_{4D} , by the volume of the compactified extra dimensions [21], V_n , such that

$$M_{4D}^2 = V_n M_D. \quad (2.1)$$

The definition of the volume of extra dimensions, V_n , is

$$V_n = (2\pi R_C)^n. \quad (2.2)$$

Consequently, as can be seen in Equation 2.2, the Planck scale in higher dimensions, M_D , is smaller than the Planck scale that is predicted in 4D, M_{4D} . The convention $\hbar = c = 1$ has been used here.

2.4 RS-type extra dimensions and Universal Extra Dimensions

The RS model [22] of extra dimensions, developed by Raman Sundrum and Lisa Randall, predicts one tiny, warped extra dimension that is compactified. As in ADD-type extra dimensional models, Standard Model fields in RS-type extra dimensions are confined to the brane. Universal Extra Dimensions (UED), however, allow all fields to propagate in the bulk.

2.5 Black Hole Formation at the LHC

A black hole is an object that has a singularity at its center and that is surrounded by an event horizon. A singularity is a break down of spacetime. An event horizon is the point, on the outer edge of the black hole, beyond which nothing can escape (due to the strength of the gravitational pull). The term “black hole” originated

from the fact that not even light can escape beyond the event horizon. The classical Hoop Conjecture states that a black hole is formed when an object is compressed in all directions into a space with a circumference less than $2\pi r_s$, where the object's Schwarzschild radius is denoted as r_s [23]. The Schwarzschild radius is the distance between the center of the black hole and the event horizon. The criteria for black hole formation at the LHC is that the hoop conjecture must be satisfied by two partons from a proton-proton collision [24], [23]. Black hole generators at the LHC utilize the classical hoop conjecture. A quantum version of the hoop conjecture, however, was proposed by Casadio, Micu, and Scardigli [24] and serves as a more precise description of the hoop conjecture for the production of quantum black holes.

Semiclassical black holes [25], string balls, and quantum black holes are the main types of black hole and signatures of TeV-scale predicted for formation at the LHC and searched for so far. (Unlike stellar or supermassive black holes, black holes at the LHC are microscopic and decay instantly.) Semiclassical black holes, which will be discussed in more detail shortly, are characterized by energetic final states with many particles. String balls are the excited string states that are produced from the decay of semiclassical black holes in scenarios of weakly coupled string theory. Like semiclassical black holes, string balls [26] are characterized by energetic final states with many particles. Quantum black holes [27] are black holes with masses that are predicted to be at the Planck scale. Quantum black holes are characterized by a primarily two particle final state. Both ATLAS [28] and CMS [29] have performed searches for semiclassical black holes yielding many jets characterized by high energy. Additionally, ATLAS searched for semiclassical black holes decaying to jets and at least one energetic lepton [30]. ATLAS have searched for quantum black holes decaying to two energetic jets [4] as well as quantum black holes decaying to a lepton and jet [31], quantum black holes decaying to two leptons ($e\mu$, $e\tau$, and $\mu\tau$) [32], and quantum black holes decaying to a photon and a jet [33]. CMS has searched for quantum black holes decaying to two leptons ($e\mu$, $e\tau$, and $\mu\tau$) [34], quantum black holes decaying to a photon and jet [35], and quantum black holes decaying to two energetic jets [36]. None of these searches have found any evidence of signature of TeV-scale at the LHC.

2.6 Semiclassical Black Holes

If the black hole mass is much greater than the Planck scale, semiclassical black holes are predicted to form. However, because the maximum energy design of the LHC is

currently 14 TeV, semiclassical black holes are likely beyond the reach of the LHC. Semiclassical black holes are arbitrarily decaying black holes, decaying thermally into many particles via Hawking radiation.

The experiments at the LHC model semiclassical black hole production primarily with the BlackMax [37] Monte Carlo generator and the Charybdis2 [38] Monte Carlo generator, both of which use ADD-type extra dimensions. This thesis uses the Charybdis2 black hole generator. The black hole cross section, σ , is modeled by the Charybdis2 generator as approximately equal to πr_S^2 , where the black hole's Schwarzschild radius is denoted as r_S . The black hole radius is modeled by Charybdis2 as

$$r_S = \frac{1}{\sqrt{\pi}M_D} \left(\frac{m_{\min}}{M_D} \left(\frac{8\Gamma\left(\frac{n+3}{2}\right)}{n+2} \right) \right)^{\frac{1}{n+1}}. \quad (2.3)$$

where n is the number of extra dimensions, $m_{\min}$ is the minimum mass required for the black hole to be produced, M_D is the fundamental Planck scale in higher dimensions, and Γ is the Gamma function defined below as

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt. \quad (2.4)$$

The Charybdis2 generator forces a black hole to be formed in every collision. For every 50,000 events (collisions) generated, 50,000 black holes will be formed. The four phases of black hole decay are modeled by the Charybdis2 generator. First is the balding (formation) phase, second is the spin-down phase, third is the Schwarzschild phase, and fourth is the Planck phase. Due to gravitational radiation, the black hole loses its asymmetry and multipole moments in the balding phase [38]. Gravitational radiation is the emission of ripples in space-time called gravitational waves. Hawking radiation is the radiation emitted outside a black hole's event horizon. Due to Hawking radiation, the black hole loses its angular momentum and mass in the spin-down phase. Also due to Hawking radiation, the black hole loses mass in the Schwarzschild phase. In the final stage of black hole decay, the Planck phase, the black hole mass and perhaps the black hole temperature reach the Planck scale. More details on the Charybdis2 black hole generator can be found in [38].

2.7 Noncommutative Black Holes

Noncommutative black holes are different from all other black holes previously searched for by ATLAS in that they are not arbitrarily decaying. Noncommutative black holes decay up until a certain point and then stop decaying. What remains of the black hole is called a black hole remnant. The black hole remnant is stable and is a dark matter candidate.

This remnant appears as missing energy, E_T^{miss} . Noncommutative black holes result from noncommutativity being incorporated into extra dimensions. The noncommutativity is defined in [39] as

$$\left[x_A, x_B \right] = i\theta^{AB} = i \frac{\epsilon^{AB}}{\Lambda_{\text{NC}}^2}, \quad (2.5)$$

where Λ_{NC} is the noncommutative scale, ϵ^{AB} is the dimensional, frame-independent antisymmetric matrix with real and constant entries, $\sqrt{\theta}M_D$ is the noncommutative parameter and is approximately equal to $\frac{1}{\Lambda_{\text{NC}}}$.

The effect of noncommutativity is that mass distributions in the energy-momentum tensor [39] become smeared. Instead of being characterized by a δ function (δM), a point mass, m , becomes a spherically symmetric, static, Gaussian smeared matter source ρ [39]

$$\rho = \frac{m}{(4\pi\theta)^{\left(\frac{n+3}{2}\right)}} e^{\frac{-r_S^2}{(4\theta)}}, \quad (2.6)$$

where $\sqrt{\theta}$ is the noncommutative parameter, Λ_{NC} is the noncommutative mass scale, n is the extra dimensions (beyond the four known dimensions), and r_S is the Schwarzschild radius.

Noncommutative black holes have three free parameters: number of extra dimensions, minimum black hole mass (m_{min} in units of TeV), and noncommutative scale. An additional parameter is the fundamental Planck scale M_D (in units of TeV). The noncommutativity causes the black hole to be cold and have a softer transverse momentum spectrum than the highly energetic black holes that ATLAS

has previously searched for. Noncommutative black holes emit many relatively soft particles. This stark contrast served as the motivation to search for noncommutative black holes, allowing for the opportunity to chase a challenging final state at colliders. Noncommutativity is motivated by string theory. Therefore, noncommutative black holes are a signature of string theory.

2.8 Scalar cascades

Another model this thesis targets is the scalar cascade model. The scalar cascade model [3] is a model of N real scalars, ϕ , with the following Lagrangian, $\mathcal{L}_\phi$:

$$\mathcal{L}_\phi = \partial_\mu \phi_i \partial^\mu \phi_i - \frac{m_i^2}{2} - a_{ijk} \phi_i \phi_j \phi_k - \lambda_{ijkl} \phi_i \phi_j \phi_k \phi_l. \quad (2.7)$$

Here, m_i is the mass of the scalar, a_{ijk} is the strength of the trilinear interaction, and λ_{ijkl} is the strength of the quartic interaction.

These are hidden sector models. Hidden sectors (or dark sectors) [5] [6] are quantum fields and their corresponding particles that do not interact strongly with the Standard Model and remain hidden. The interaction of hidden sectors with Standard Model particles is believed to be mediated by gravity or via a mediator particle called a portal. Both the ATLAS experiment and the CMS experiment have rich search programs for hidden sectors [40] [41] [42] [43] [44] [45] [46] [47] [48] [49] [50] [51] [52] [53] [54] [55] [56].

The portal in the scalar cascade model is the Higgs. New scalars ϕ are connected to the Standard Model coupling through the Higgs portal couplings as seen in the following Lagrangian term:

$$\mathcal{L}_{\phi H} = -a^{\text{SM}}_i \phi_i |H|^2 - \lambda^{\text{SM}}_{ij} \phi_i \phi_j |H|^2. \quad (2.8)$$

These scalars share the same final states (high particle multiplicity) as the light fermions that arise in string theory and, therefore, by virtue, would be signatures of string theory.

The new scalars are produced via either a scalar-mediated pair production process or scalar-mediated single production, as seen in Figure 2.2.

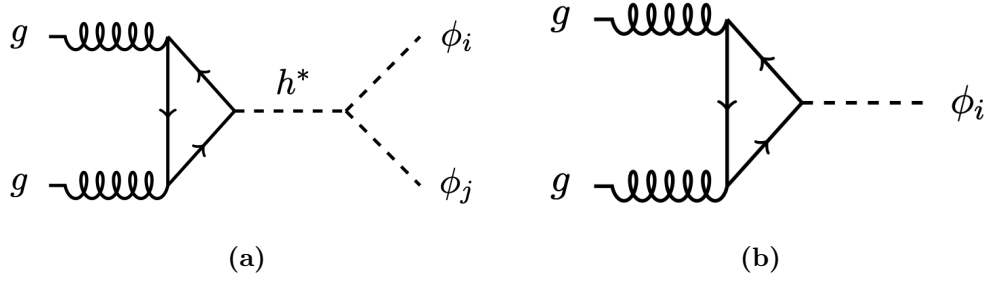


Figure 2.2 *Scalar cascade generation via (a) pair production $\sigma \sim (\lambda^{\text{SM}})^2$ (where h^* is the Higgs) and (b) single production $\sigma \sim (a^{\text{SM}})^2$. Figure modified from [3].*

The new scalars can decay via either a hidden sector quartic coupling, a Higgs portal coupling, a hidden sector trilinear coupling, or a Standard Model trilinear coupling, as seen in Figure 2.3.

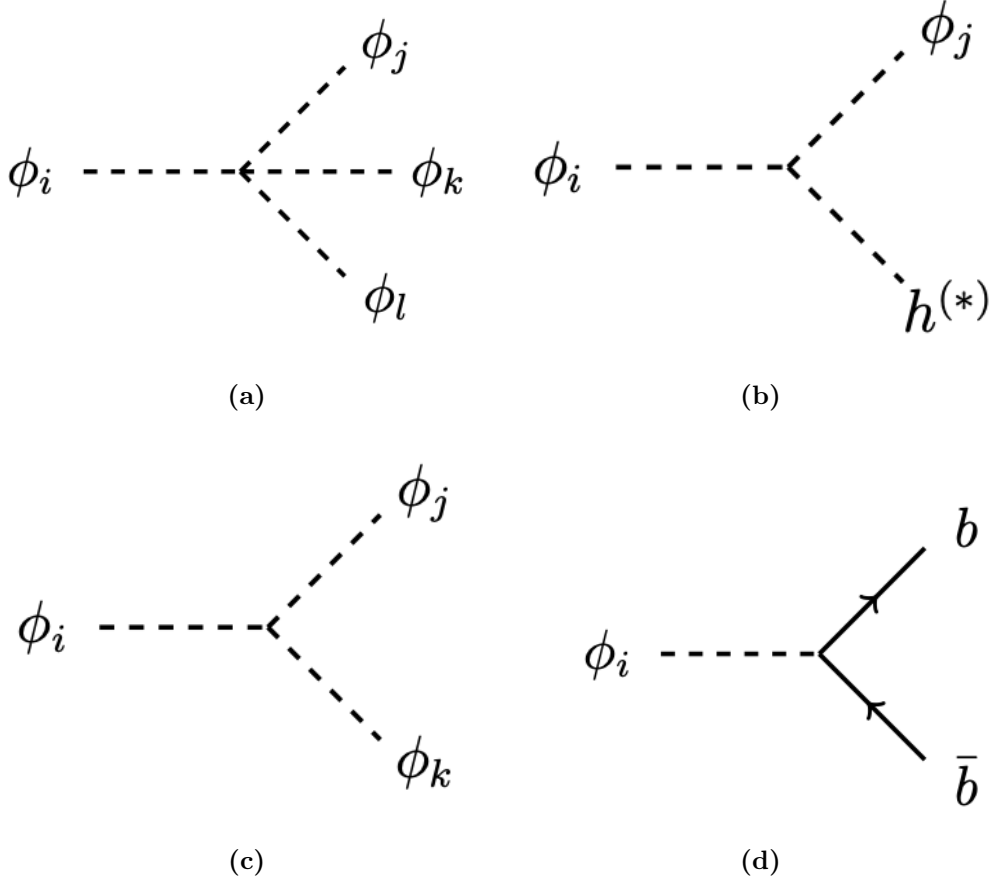


Figure 2.3 *Scalar cascade collider phenomenology. Scalar Cascade decays via (a) hidden sector quartic $\Gamma \sim \lambda^2$, (b) Higgs portal $\Gamma \sim (\lambda^{\text{SM}})^2$, (c) hidden sector trilinear $\Gamma \sim a^2$, and (d) SM trilinear $\Gamma \sim (a^{\text{SM}})^2$. Figure modified from [3].*

This analysis studies the pair production of scalar cascades. The two new scalars, ϕ_i and ϕ_j , created via pair production, as seen in Figure 2.2, will be referred to as ϕ_i and ϕ_j . As they are the first scalars produced in the cascade, they will also be referred to in future diagrams and tables as ϕ_1 . Future chapters in this thesis will identify final state particles as having originated from ϕ_i or ϕ_j .

There are 6 classes of scalar cascades studied in this thesis: a 3 Body Stable Chain Decay where the lightest scalar is stable, a 3 Body Unstable Chain Decay where the lightest scalar is unstable, a 3 Body Stable Cascade Decay where the lightest scalar is stable, a 3 Body Unstable Cascade Decay where the lightest scalar is unstable, a 2 Body Stable Cascade Decay where the lightest scalar is stable, and a 2 Body Unstable Cascade Decay where the lightest scalar is unstable. As noted in the introduction, the scalar cascade models used in this thesis were designed specifically for this analysis

in collaboration with the theorists from [3] to allow for final states with maximum particle multiplicity. The final state of each class depends on the number of scalars included in the model. The phenomenology implications of the number of scalars in each model will be detailed in a future chapter. The search conducted in this thesis focuses on a scalar cascade model with a final state characterized by high particle multiplicity and containing substantial missing energy.

2.8.1 3 Body Stable Chain Decay

In the case of the 3-body chain decay where the lightest scalar is stable, two ϕ 's are produced via pair production and each ϕ decays to another ϕ and two Standard Model particles at every step of the cascade. The exception is at the last step of the decay, where the ϕ decays to a stable ϕ and two Standard Model particles. Figure 2.4 shows the decay of one of the pair produced ϕ 's.

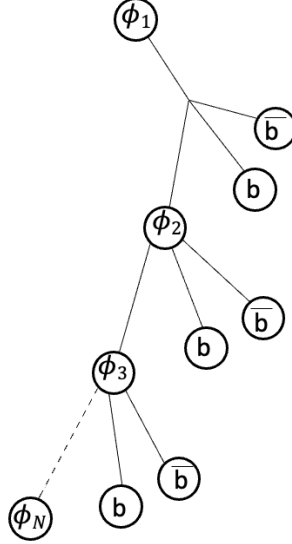


Figure 2.4 *Scalar Cascades 3 Body Stable Chain Decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. Dashed lines indicate that there can be more steps in the decay. The lightest ϕ is stable.*

The mass, m_i , of each subsequent ϕ must abide by the relation $m_i \gtrsim m_{i+1} + 2m_{\text{SMP}}$, where $2m_{\text{SMP}}$ denotes the mass of two Standard Model particles, in order to energetically accommodate the generation of two Standard Model particles. For example, in the decay to two b quarks, this requirement would be $m_i \gtrsim m_{i+1} + 10$ GeV because the mass of the b quark is 4.18 GeV.

The Standard Model coupling constant, λ^{SM} , for the 3 Body Stable Chain Decay class, must satisfy $\lambda^{\text{SM}} \lesssim \frac{4\pi}{N}$. The 3-Body Stable Chain class does not couple to the hidden sector, so $\lambda^{\text{HS}} = 0$.

2.8.2 3 Body Stable Cascade Decay

In the case of the 3 Body Stable Cascade decay where the lightest scalar is stable, two ϕ 's are produced via pair production and each ϕ decays to three ϕ 's. At each step of the cascade, each ϕ decays to three ϕ 's. The exception is at the last step of the decay where the ϕ decays to a stable ϕ and two Standard Model particles [57]. Figure 2.5 shows the decay of one of the pair produced ϕ 's. The Standard Model particles are illustrated in the diagram as b quarks, but can be other particles. For example, the Standard Model particles can be other particles such as charm quarks, taus, electrons, and muons. The type of Standard Model particle allowed is directly related to the mass spacing between the lightest scalar and the second lightest scalar. For example, a 10 GeV mass spacing would allow for the generation of two b quarks. A 3 GeV mass spacing would allow for the generation of two charm quarks, but would not allow for the generation of two b quarks. In the studies performed for this thesis, in which the focus was on models with a large number of scalars, the preferred coupling of the scalars was to charm quarks, b quarks, and taus. No filter was applied at the simulation level for the decays of these models. All final state particles were permitted to be generated.

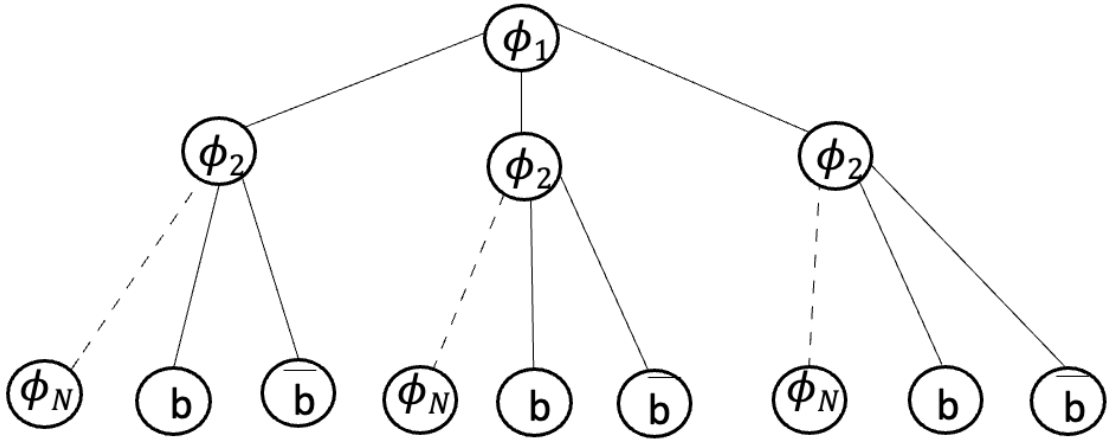


Figure 2.5 *Scalar Cascades 3 Body Stable Cascade Decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. Dashed lines indicate that there can be more steps in the decay. The lightest ϕ is stable.*

The mass of each scalar in the decay must abide by the relation $m_i \gtrsim 3m_{i+1}$ in order to accommodate the generation of three new scalars. The last scalar, however, need only be generated from a scalar that is heavier by at least the mass of two Standard Model particles in order to allow for the production of two Standard Model particles at the last step of the decay. As will be seen in a future chapter, the mass scale of the scalars depends on how many scalars are in the model, as well as the particular class the model belongs to. For models in the 3 Body Stable Cascade Decay class that contain fewer scalars, the heaviest scalars can be below, for example, even 50 GeV. However, for models with more scalars, the heaviest scalars will necessarily have higher masses such as 200 GeV, 300 GeV, 400 GeV, or higher. This will be charted out in a future chapter.

The Standard Model coupling constant, λ^{SM} , for the 3 Body Stable Cascade Decay class, must be satisfy $\lambda^{\text{SM}} \lesssim \frac{4\pi}{N}$. The limit of the hidden sector coupling constant, λ^{HS} , for the 3 Body Stable Cascade Decay class is such that $\lambda^{\text{HS}} \lesssim \frac{16\pi^2}{N^2}$ [57].

2.8.3 2 Body Stable Cascade Decay

In the case of the 2-body cascade decay where the lightest scalar is stable, two ϕ 's are produced via pair production and each ϕ decays to two ϕ 's. At each step of the cascade, each ϕ decays to two ϕ 's. The exception is at the last step of the decay where the ϕ decays to a stable ϕ and two Standard Model particles. Figure 2.6 shows the decay of one of the pair produced ϕ 's.

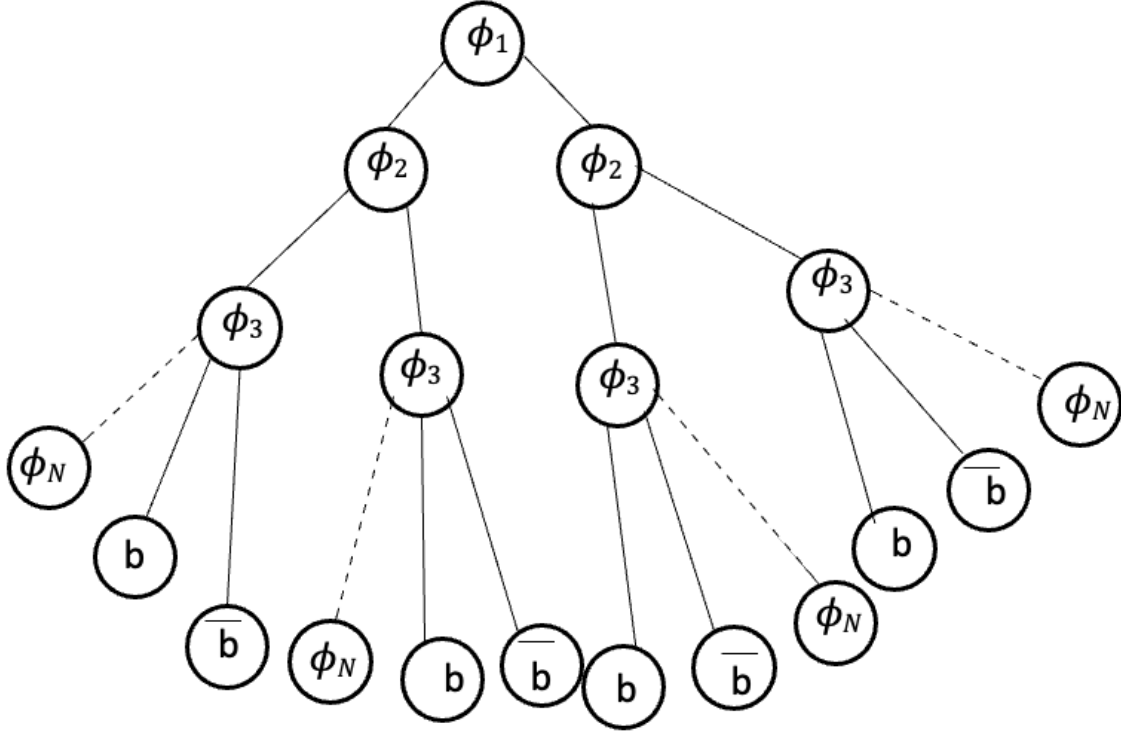


Figure 2.6 *Scalar Cascade 2 Body Stable Cascade Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is stable.*

The mass of each scalar in the decay must abide by the relation $m_i \gtrsim 2m_{i+1}$ in order to accommodate the generation of two new scalars. The last scalar, however, need only be generated from a scalar that is heavier by at least the mass of two Standard Model particles in order to allow for the production of two Standard Model particles at the last step of the decay.

The Standard Model coupling constant, λ^{SM} , for the 2 Body Stable Cascade class must be satisfy $\lambda^{\text{SM}} \lesssim \frac{4\pi}{N}$. The hidden sector coupling constant of the 2 Body Stable Cascade class, λ^{HS} , must fulfill $\lambda^{\text{HS}} \lesssim \frac{4\pi m_\phi}{N^{3/2}}$ [57].

2.8.4 3 Body Unstable Chain Decay

In the case of the 3-body chain decay where the lightest scalar is unstable, two ϕ 's are produced via pair production and each ϕ decays to a ϕ and two Standard Model particles at every step of the cascade. The exception is at the last step of the decay,

where the ϕ decays to a stable ϕ and two Standard Model particles. Then, the unstable ϕ decays to two Standard Model particles. Figure 2.7 shows the decay of one of the pair produced ϕ 's.

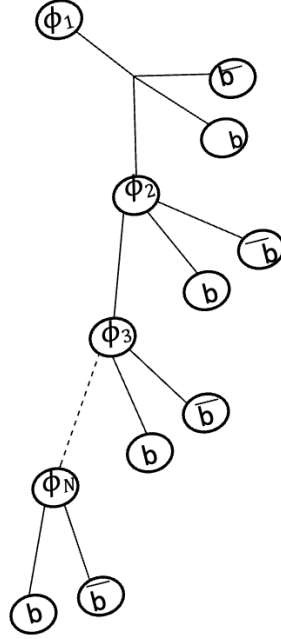


Figure 2.7 *Scalar Cascade 3 Body Unstable Chain Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is unstable.*

The limit of the Standard Model coupling constant, λ^{SM} , for the 3 Body Unstable Chain Decay class is such that $\lambda^{\text{SM}} \lesssim \frac{4\pi}{N}$. The 3 Body Unstable Chain Decay class does not couple to the hidden sector, so $\lambda^{\text{HS}} = 0$ [57].

2.8.5 3 Body Unstable Cascade Decay

In the case of the 3-body cascade decay where the lightest scalar is unstable, two ϕ 's are produced via pair production and each ϕ decays to three ϕ 's. At each step of the cascade, each ϕ decays to three ϕ 's. The exception is at the second to last step of the decay, in which the ϕ decays to an unstable ϕ and two Standard Model particles. Then, the unstable ϕ decays to two Standard Model particles. Figure 2.8 shows the decay of one of the pair produced ϕ 's.

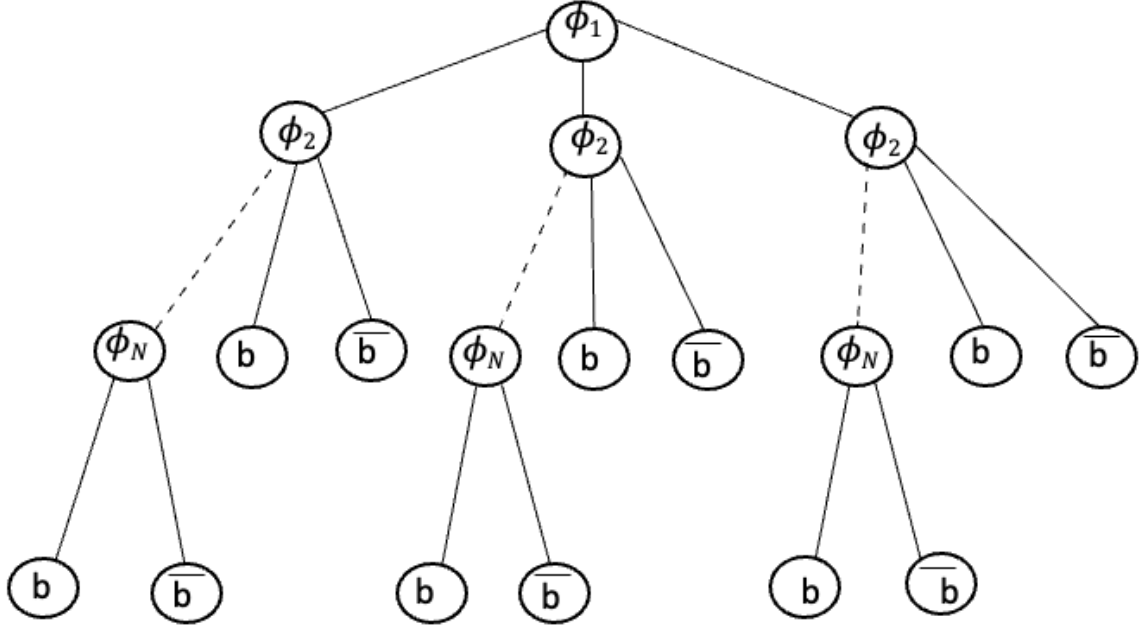


Figure 2.8 *Scalar Cascade 3 Body Unstable Cascade Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is unstable.*

The limit of the Standard Model coupling constant, λ^{SM} , for the 3 Body Unstable Cascade Decay class is such that $\lambda^{\text{SM}} \lesssim \frac{4\pi}{N}$. The hidden sector coupling constant, λ^{HS} , for the 3 Body Unstable Cascade Decay class must satisfy $\lambda^{\text{HS}} \lesssim \frac{16\pi^2}{N^2}$ [57].

2.8.6 2 Body Unstable Cascade Decay

In the case of the 2-body cascade decay where the lightest scalar is unstable, two ϕ 's are produced via pair production and each ϕ decays to two ϕ 's. At each step of the cascade, each ϕ decays to two ϕ 's. The exception is at the second to last step of the decay, in which the ϕ decays to an unstable ϕ and two Standard Model particles. Then, the unstable ϕ decays to two Standard Model particles. Figure 2.9 shows the decay of one of the pair produced ϕ 's.

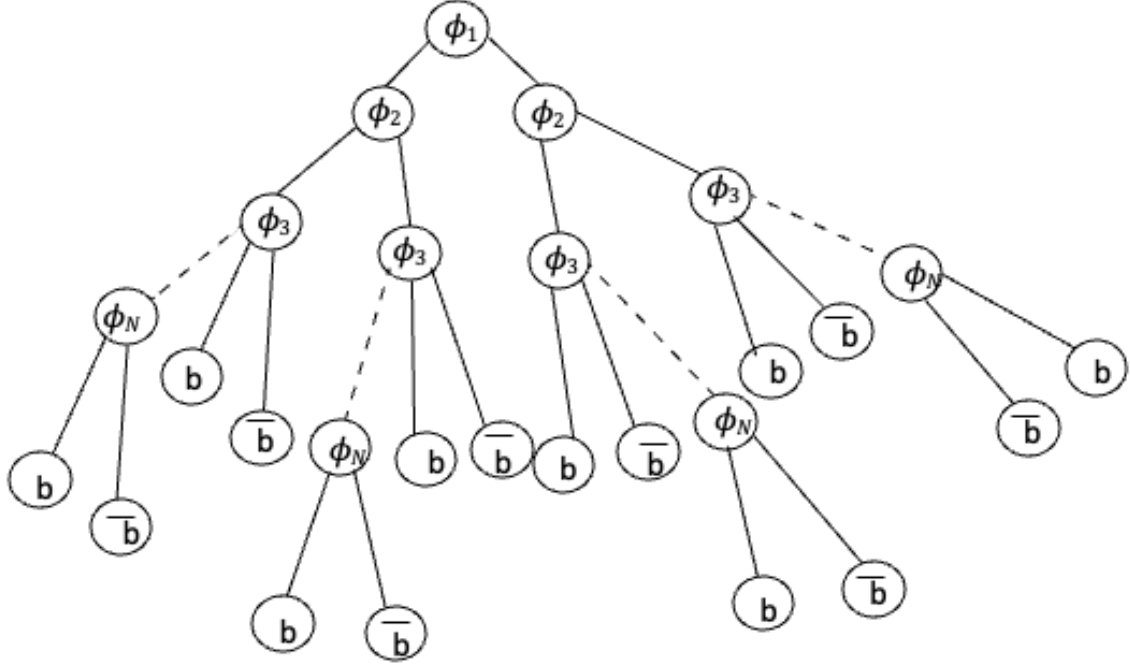


Figure 2.9 *Scalar Cascade 2 Body Unstable Cascade Decay. Dashed lines indicate that there can be more steps in the decay. New scalars are denoted as ϕ 's, while b 's (for b quarks) are illustrative and represent Standard Model particles. The lightest ϕ is unstable.*

The limit of the Standard Model coupling constant, λ^{SM} , for the 2 Body Unstable Cascade Decay class is such that $\lambda^{\text{SM}} \lesssim \frac{4\pi}{N}$. The hidden sector coupling constant, λ^{HS} , for the 2 Body Unstable Cascade Decay class must satisfy $\lambda^{\text{HS}} \lesssim \frac{4\pi m_\phi}{N^{3/2}}$ [57].

2.9 Non-minimal Dark Sectors

Non-minimal dark sectors [58] are a model of large ensembles. They are extra dimensional models that are also models of dark matter. They yield long decay chains of many particles.

The non-minimal dark sectors have three production mechanisms. These mechanisms rely on new scalars, referred to here as ϕ , and new fermions, referred to as χ . They are particles beyond that which the Standard Model predicts. To be clear, while also spin 0 particles, the new scalars in the non-minimal dark sector model are not the same new scalars that appear in the scalar cascades model.

Figure 2.10 shows the production of non-minimal dark sectors via one χ and one

scalar mediator, ϕ . This process is referred to as the $\chi\phi$ process. The ϕ produces a quark and a χ . This leaves two χ 's and a quark.

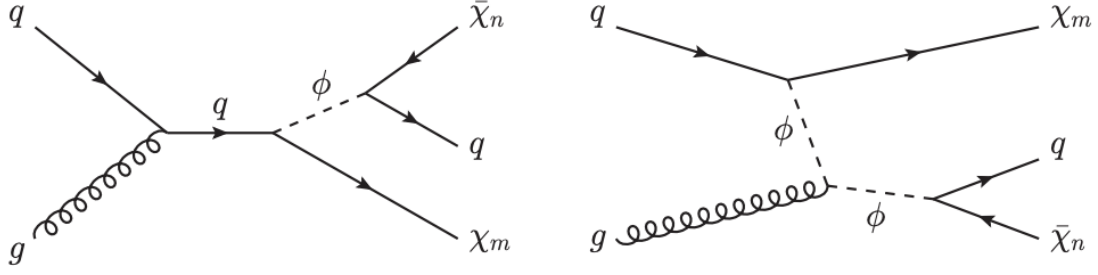


Figure 2.10 $\chi\phi$ Production Channel of Non-minimal Dark Sectors

Throughout this document, the decay products of the ϕ and the decay products of the χ will be referred to as “branch 1” and “branch 2”. It does not matter which one is referred to as “branch 1” or “branch 2”. All that matters is that the decay products from the χ are known to originate separately from the ϕ .

The χ 's, referred to as χ_m and χ_n , then both decay via the χ cascade chain depicted in Figure 6.1. The non-minimal dark sectors in this analysis can be described as an ensemble of N χ_n , where n ranges from 0 to $N - 1$. (N is the number of dark fermions.) The subscript n or m for a given χ_m or χ_n increases by one every time a new χ is produced. The decay chain stops upon the production of a lightest, stable χ . Each step of the decay yields two quarks. For simplicity, the theorists in [58] have demanded that the quark be an up quark.

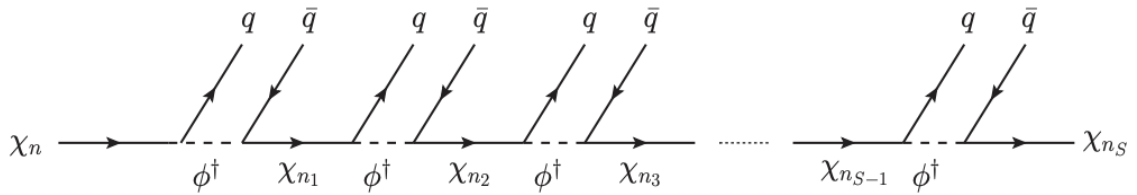


Figure 2.11 Decay chain in which each χ (ensemble constituent) decays into a χ and two quarks. The process ends with the production of two quarks and a stable, lightest χ .

The second production occurs via two ϕ mediators. This process is referred to as the $\phi\phi$ process. It is depicted in Figure 2.12. Each ϕ mediator yields a quark and a χ . Again, the χ 's are referred to as either χ_m or χ_n .

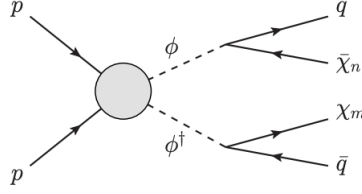


Figure 2.12 $\phi\phi$ Production Channel of Non-minimal Dark Sectors

Throughout this document, the decay products of one of the two ϕ 's and the decay products of the other ϕ will be referred to as "branch 1" and "branch 2". These χ 's both decay via the non-minimal dark sector decay chain depicted in Figure 6.1 and in the same manner as described above.

The third production mechanism for non-minimal dark sectors, depicted in Figure 2.13, occurs via the ϕ mediated production of two χ 's, χ_m or χ_n . This is called the $\chi\chi$ production channel.

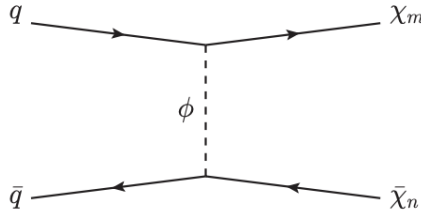


Figure 2.13 $\chi\chi$ Production Channel of Non-minimal Dark Sectors

Throughout this document, the decay products of one of the two χ 's and the decay products of the other χ will be referred to as "branch 1" and "branch 2". These χ 's both decay via the chain depicted in Figure 6.1.

A single simulated signal sample must contain all 3 processes properly weighted by production cross sections.

Non-minimal dark sectors have 6 free parameters: Δm , δ , γ , c_0 , m_0 , and, m_ϕ . The equations presented here that describe the physics of non-minimal dark sectors all come from [58]. The mass of the scalar mediator is denoted as m_ϕ . The mass scaling of the ensemble is described by m_n , where

$$m_n = m_0 + n^\delta \Delta m \quad (2.9)$$

and m_0 represents the mass of the lowest ensemble constituent, a stable particle

which appears as missing energy, δ is the mass scaling parameter, n ranges from 0 to $N - 1$, and Δm represents the mass gap between ensemble constituents. The mass m_2 , for example, is greater than m_1 .

The coupling relation to describe the coupling of the dark fermion, χ , to quarks at different steps in the ensemble depicted in Figure 6.1 is c_{nq} , where

$$c_{nq} = c_{0q} \left(\frac{m_n}{m_0} \right)^\gamma. \quad (2.10)$$

Here, γ is a coupling scaling exponent. The coupling is different for different flavors of quarks. However, as noted, the theorists restricted the quark to be an up quark for simplification purposes only.

Three benchmark values of δ that are of particular theoretical interest are $\delta = \frac{1}{2}$, $\delta = 1$, and $\delta = 2$. A δ value of $\frac{1}{2}$ describes scenarios dictated by string theory [58]. A δ value of 1 and a δ value of 2 describe scenarios of a single tiny, warped compactified extra dimension from RS-type extra dimensions but are not indicative of string theory [58].

The set of parameter values chosen to model the ensemble must satisfy certain consistency checks. One consistency check involves the number of ensemble constituents, N . N should satisfy the following condition:

$$N \leq 1 + \left(\frac{m_\phi - m_0 - m_q}{\Delta m} \right)^2, \quad (2.11)$$

where m_q is the mass of the up quark in the ensemble depicted in Figure 6.1. This constraint is imposed by the theorists in [58] for simplification purposes only.

For γ greater or equal to 0, the coupling constant of the lowest state must satisfy

$$c_0 = 4\pi \left[1 + \frac{\Delta m}{m_0} (N - 1)^\delta \right]^{-\gamma}. \quad (2.12)$$

due to the perturbative assumptions made. For the same reason, γ less than or equal to 0, c_0 must satisfy $c_0 = 4\pi$.

In the case where χ_m and χ_n are produced via the direct production of the scalar

mediator ϕ , seen in Figure 2.12, an extra condition is applied to c_0 in order to kinematically allow for the direct production of the scalar mediator. Specifically,

$$c_0 \ll 4\sqrt{\pi} \left[\sum_{n=0}^{N-1} \left(\frac{m_n}{m_0} \right)^\gamma \left(1 - \frac{m_n^2}{m_\phi^2} \right)^2 \right]^{-\frac{1}{2}}. \quad (2.13)$$

A more detailed discussion of these parameters will occur later in this thesis.

2.10 Summary

This chapter introduced the theory that governs particle physics, the Standard Model, described its limitations, and proposed potential solutions. Three models that propose the existence of more than four dimensions were introduced: noncommutative black holes, scalar cascades, and non-minimal dark sectors. The rest of the thesis will detail dedicated searches for these new physics models. These searches will target final states with many particles, particularly targeting jets, the collimated sprays of particles originating from the decay of quarks and gluons.

Chapter 3

The Large Hadron Collider and the ATLAS Detector

The previous chapter detailed the unresolved problems posed by the Standard Model and introduced new physics models that could serve as solutions. This chapter introduces the machine at which these new models could be discovered: the Large Hadron Collider (LHC). The chapter details how the LHC works and introduces the experiments it provides collisions for. Special emphasis is given to the ATLAS experiment as this thesis uses ATLAS data to perform each search. Out of all the detectors ever made for particle colliders, the ATLAS detector is the largest. Starting from the innermost layer of the detector, this chapter goes through each subdetector of the ATLAS experiment, working its way to outermost layer of the detector and ending with a description of the trigger system, the mechanism by which ATLAS records and filters its data.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at the European Organization of Nuclear Research, CERN, is the world's biggest and highest energy particle accelerator. It is installed in a 27 km ring 100 m underground and is located near Geneva, Switzerland. CERN seeks to investigate the Standard Model of particle physics, its problems, and possible solutions.

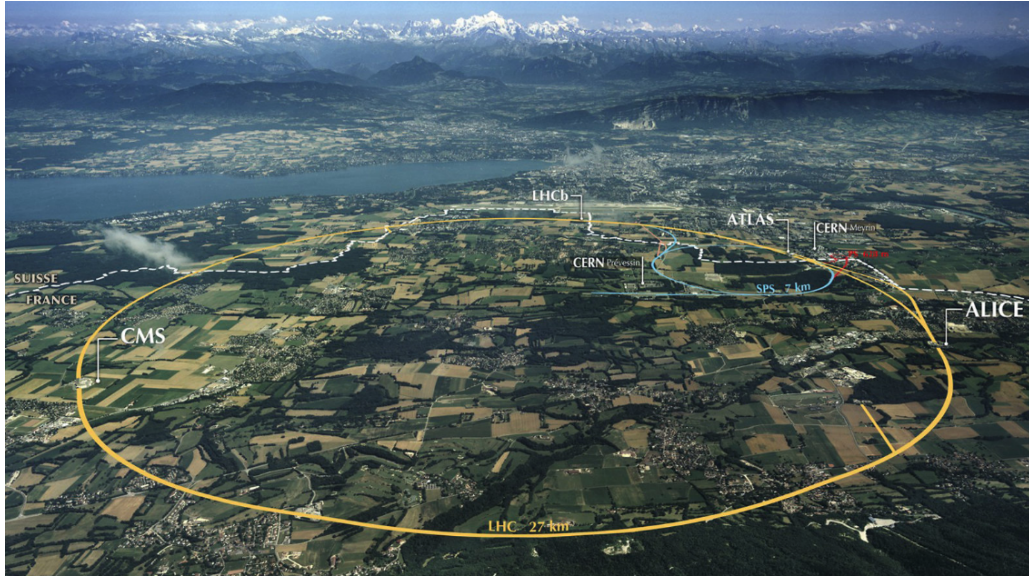


Figure 3.1 *Aerial view of Geneva, Switzerland and the surrounding area with a drawing of the Large Hadron Collider (LHC) and associated experiments overlaid. Figure taken from [59].*

Protons are accelerated around the 27 km ring to nearly the speed of light. A bottle of hydrogen gas serves as the source of these protons. An electric field strips the hydrogen atoms of their electrons, leaving only protons. Before reaching the LHC, these hydrogen atoms pass through Linac 2, the first linear accelerator in a series of accelerators to accelerate the protons [60] [61] [62].

The CERN accelerator complex *Complexe des accélérateurs du CERN*

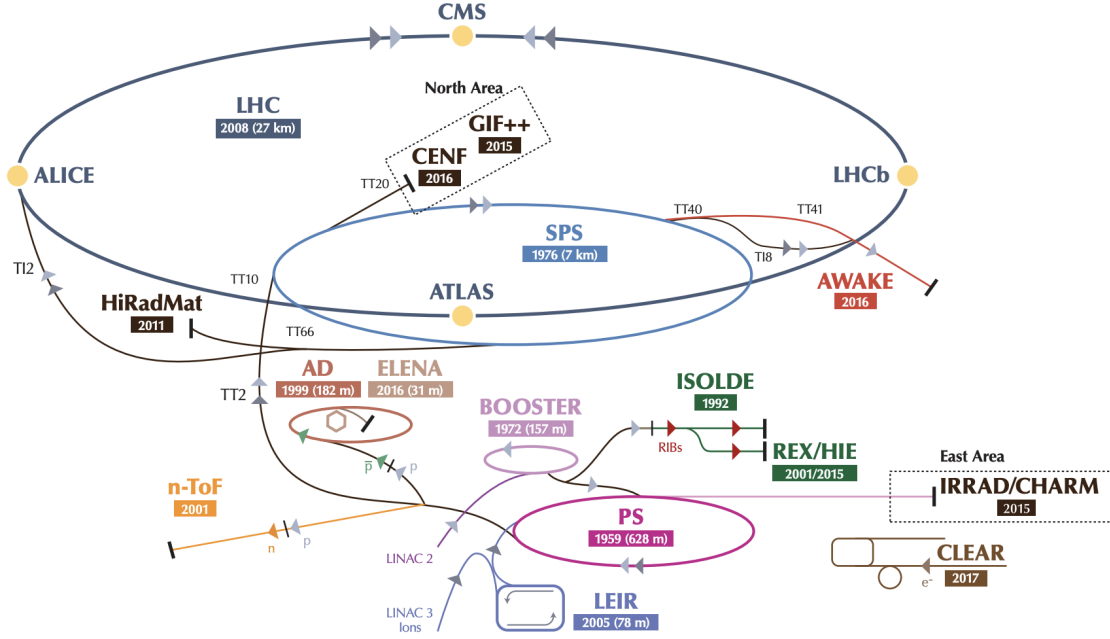


Figure 3.2 *CERN Accelerator Complex. Figure taken from Reference [63].*

The protons are accelerated up to 50 MeV by Linac 2. The protons are further accelerated, this time to an energy of 1.4 GeV, by the Proton Synchrotron Booster (PSB), which is comprised of four superimposed synchrotron rings. The protons are next accelerated by the Proton Synchrotron (PS) to an energy of 25 GeV. The protons are accelerated then to an energy of 450 GeV by the Super Proton Synchrotron. The protons are then circulated in two beams around the LHC ring. Under nominal operating conditions, each proton beam contains 2808 bunches, where there are 10^{11} protons per bunch. The proton beams are controlled by a combination of dipole and quadrupole magnets. Dipole magnets steer the beams along the trajectory of the ring. Quadrupole magnets focus and squeeze the proton beams so as to allow for more particles to collide. The protons collide at four interaction points. A detector lies at each of these four interaction: the ATLAS (A Toroidal LHC Apparatus) detector, ALICE (A Large Ion Collider Experiment) detector, LHCb (Large Hadron Collider beauty) detector, and CMS (Compact Muon Solenoid) detector [60] [61] [62]. Along with the CMS detector, the ATLAS detector is one of the two general purpose detectors at the LHC designed to seek evidence of physics beyond the Standard Model and to make precision measurements of Standard Model processes.

The instantaneous luminosity of the LHC, $\mathcal{L}$, is given by

$$\mathcal{L} = \frac{n_b f_{rev} n_1 n_2}{2\pi \sigma_x \sigma_y}. \quad (3.1)$$

where n_b is the number of bunches per beam, f_{rev} is the revolution frequency, n_1 and n_2 are the number of particles per bunch for each of the two beams, and σ_x and σ_y are the x and y components, respectively, of the beam width.

At the LHC, the number of events that occur for a given process is given by

$$N = \sigma \int \mathcal{L}(t) dt. \quad (3.2)$$

where σ is the cross section of the process and the integrated luminosity is $\int \mathcal{L}(t) dt$. The integrated luminosity is quoted in inverse femtobarns (fb^{-1}).

As seen in Figure 3.3, the LHC physics schedule from 2010 to 2041 can be split into 6 stages: Run 1, Run 2, Run 3, Run 4, Run 5, and Run 6.

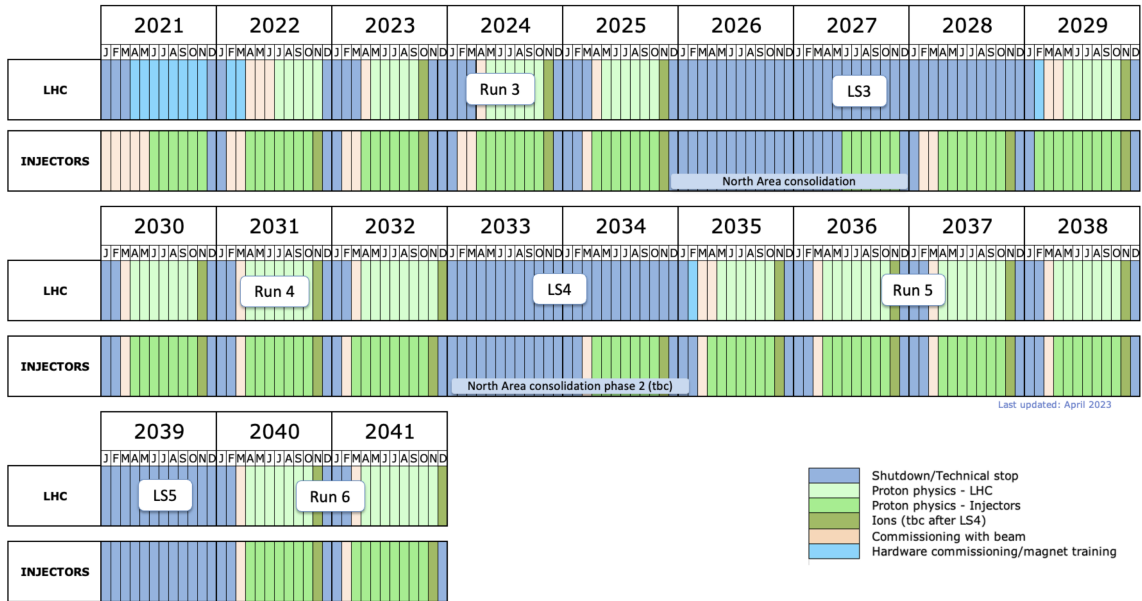


Figure 3.3 *LHC timeline. Figure taken from Reference [64].*

In 2010, Run 1 began and the LHC produced collisions with a total center of mass energy, $\sqrt{s}$, of 7 TeV by accelerating protons to an energy of 3.5 TeV per beam. Run 1 continued through 2010 and 2011 with collisions of $\sqrt{s} = 7$ TeV. In 2012, the LHC continued the Run 1 campaign, but with collisions of a heightened total center of mass

energy of 8 TeV. The LHC delivered 25 fb^{-1} of data to the ATLAS detector when running in 2011-2012. In 2015, after a long shutdown consisting of many detector upgrades, Run 2 began and the LHC provided proton collisions of a total center of mass energy of 13 TeV [65] [66] [67]. Run 2 finished at the end of 2018, thereby marking the beginning of Long Shutdown 2 (LS2), which lasted until the spring of 2022. As seen in Figure 3.5, the LHC delivered 156 fb^{-1} of integrated luminosity during Run 2, with 139 fb^{-1} of data determined to be good for physics studies. The nominal instantaneous luminosity that the LHC was designed to operate with was $1 \times 10^{34} \text{ cm}^{-2}$. During Run 2, the LHC operated with an instantaneous luminosity of $2 \times 10^{34} \text{ cm}^{-2}$, which is twice the instantaneous luminosity that the LHC was designed to operate with. The mean number of interactions per bunch crossing, or the pileup, is shown in Figure 3.4. Pileup increases with increased luminosity. Pileup presents challenges to experiments insofar as experiments have to distinguish events of interest from pileup events that look like potential signal events.

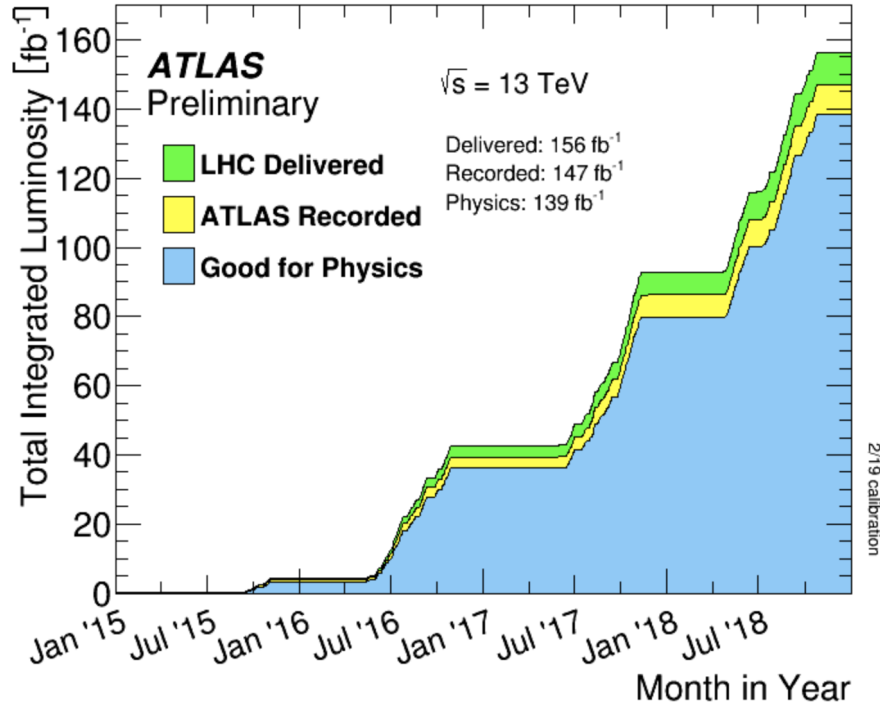


Figure 3.4 *ATLAS Integrated Luminosity delivered in Run 2. Figure taken from Reference [68].*

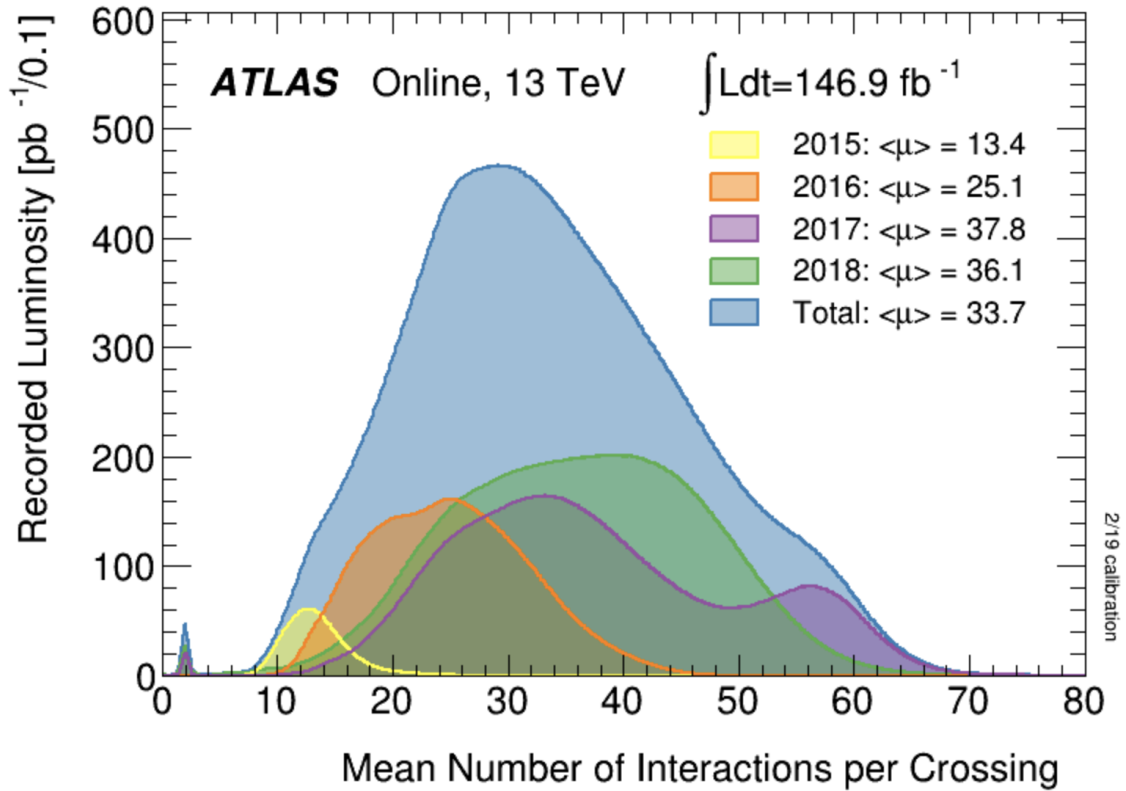


Figure 3.5 *Mean Number of Interactions in the ATLAS detector during Run 2. Figure taken from Reference [68].*

Run 3 started in July 2022 and protons were collided at a record high energy of 13.6 TeV. Run 3 will continue until 2026, at which point Long Shutdown 3 will begin and upgrades for the High Luminosity - Large Hadron Collider (HL-LHC) will commence. The HL-LHC will begin in 2031 and continue until 2041 and is expected to operate with an instantaneous luminosity 5 to 7.5 times greater than the nominal LHC instantaneous luminosity.

3.2 Coordinate System and Geometry

The ATLAS detector is 44 m long, weighs 7000 t, and has a diameter of 25 m. It consists of a series of subdetectors and subsystems: the Trigger system, the Calorimeters, the Inner Detector, and the Muon Systems. The Inner Detector consists of three sub-detectors: the pixel, transition radiation tracker, and semiconductor tracker. The Inner Detector measures the momentum, charge, and direction of charged particles. The Calorimeter contains the Tile Hadronic Calorimeter and Liquid Argon Calorimeter and measures energies carried by neutral

and charged particles. The Muon Spectrometer contains Resistive Plate Chambers, Thin Gap Chambers, Monitored Drift Tubes, and Cathode Strip Chambers and identifies and measures muons momenta. The Trigger and Data Acquisition System carefully selects interesting physics events.

The ATLAS detector uses a right-handed coordinate system, as seen in Figure 3.6

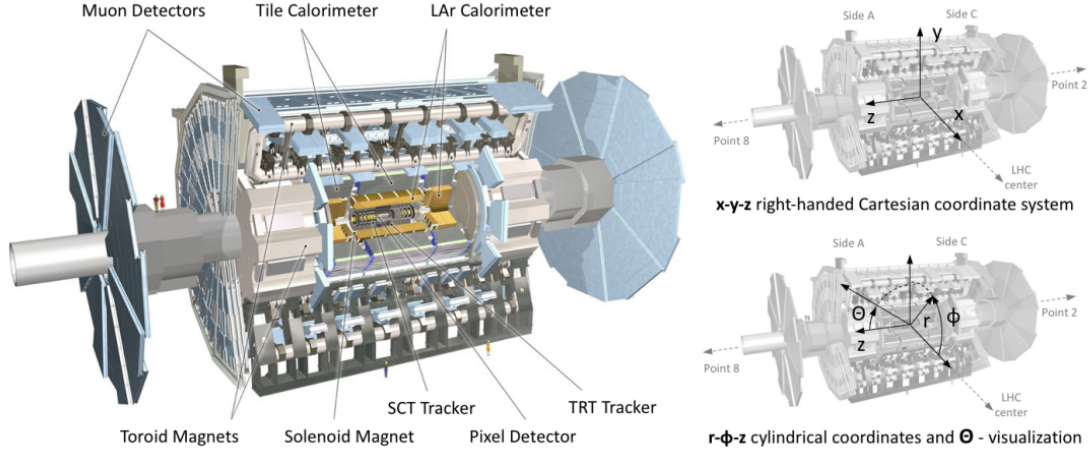


Figure 3.6 *ATLAS Detector depicted in figure on the left, taken from Reference [69]. Coordinate system used by ATLAS on the right, taken from Reference [70].*

The coordinate system's origin is the nominal interaction point, while the z axis is the direction of the beam, the x axis points to the center of the LHC from the interaction point, the y axis points upwards from the interaction point, and the xy -plane is perpendicular to the direction of the beam. The azimuthal angle, ϕ , starts from the positive x axis and extends into the transverse plane. The polar angle, θ , starts from the z axis and extends into the transverse plane. The quantity η is defined in terms of θ as

$$\eta = -\ln\left(\tan \frac{\theta}{2}\right). \quad (3.3)$$

The η is more frequently used than ϕ because it is invariant under Lorentz boost. The distance between particles in the $\eta \phi$ plane can be expressed as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (3.4)$$

In proton-proton collisions, the net momentum in the transverse plane (p_T , defined

as $\sqrt{p_x^2 + p_y^2}$, must be zero. Any additional momentum in the transverse plane at the end of the collision is termed missing energy (E_T^{miss}). The existence of E_T^{miss} could indicate Standard Model particles that pass through the detector undetected or BSM particles that similarly pass undetected through the detector.

3.3 Inner Detector

The Inner Detector is the detector subsystem closest to the interaction point. Its immersion in a 2 T solenoid magnet causes charged particles to bend. The three concentric subdetectors - the pixel, the semiconductor tracker, and the transition radiation tracker - that comprise the Inner Detector enable vertex reconstruction, track finding, and momentum measurements of charged particles. The structure of the Inner Detector can be seen in Figure 3.7 while Table 3.1 shows some of its important parameters.

The pixel detector is the innermost part of the Inner Detector, spanning from radii of 33 mm to 122.5 mm with respect to the interaction point. Its original design consisted of three concentric cylindrical layers surrounding the beamline and 3 layers at each end of the detector, perpendicular to the beamline. It contains 1744 pixel-sensor modules, where each module contains 46080 pixels. The typical pixel size is $50 \mu\text{m}$ ($R - \phi$) $\times$ $400 \mu\text{m}$ (z) each. The accuracy of the pixels in the barrel is $10 \mu\text{m}$ in $R - \phi$ and $115 \mu\text{m}$ in z , while the accuracy in the endcaps is $115 \mu\text{m}$ in R . In 2014, a fourth layer, the insertable B-layer (IBL) was added to the barrel. The IBL contains 12 million silicon pixels. The typical size of the pixel is $50 \mu\text{m}$ in $R - \phi$ and $250 \mu\text{m}$ in z . The IBL aids in the identification of secondary vertices for B hadrons.

The SCT spans from radii of 299 mm to 554 mm with respect to the interaction point and contains roughly 6.3 million readout channels. The accuracy of modules in the barrel is $17 \mu\text{m}$ ($R - \phi$) and $580 \mu\text{m}$ in Z , while the accuracy of the modules in the endcap is $17 \mu\text{m}$ ($R - \phi$) and $580 \mu\text{m}$ (R).

Surrounding the SCT is the TRT, which spans from radii of $R = 554 \text{ mm}$ to $R = 1082 \text{ mm}$. The TRT consists of 351,000 straw tube readout channels filled with gas consisting of 3% O_2 , 27% CO_2 , and 70% Xe . Each straw tube has a diameter of 4 mm and contains a $31 \mu\text{m}$ gold-plated tungsten anode wire that collects the ionized particles that results from the gas interacting with charged particles. Straws within the barrel have a length of 144 cm along the z axis while straws are positioned

radially in the endcaps and have a length of 37 cm. Valuable tracking information is obtained due to the TRT allowing for 36 hits per track.

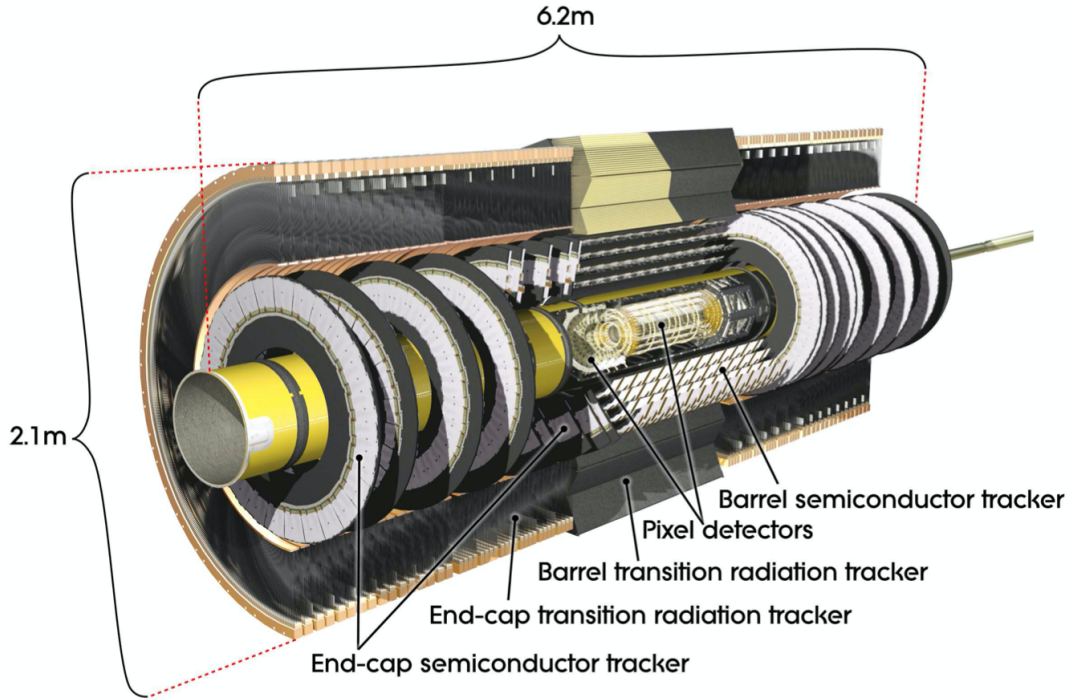


Figure 3.7 *ATLAS Inner Detector. Figure taken from Reference [69].*

Item		Radial extension (mm)	Length (mm)
Overall ID envelope		$0 < R < 1150$	$0 < z < 3512$
Beam-pipe		$29 < R < 36$	
Pixel 3 cylindrical layers 2×3 disks	Overall envelope	$45.5 < R < 242$	$0 < z < 3092$
	Sensitive barrel	$50.5 < R < 122.5$	$0 < z < 400.5$
	Sensitive end-cap	$88.8 < R < 149.6$	$495 < z < 650$
SCT 4 cylindrical layers 2×9 disks	Overall envelope	$255 < R < 549$ (barrel)	$0 < z < 805$
		$251 < R < 610$ (end-cap)	$810 < z < 2797$
	Sensitive barrel	$299 < R < 514$	$0 < z < 749$
	Sensitive end-cap	$275 < R < 560$	$839 < z < 2735$
TRT 73 straw planes 160 straw planes	Overall envelope	$554 < R < 1082$ (barrel)	$0 < z < 780$
		$617 < R < 1106$ (end-cap)	$827 < z < 2744$
	Sensitive barrel	$563 < R < 1066$	$0 < z < 712$
	Sensitive end-cap	$644 < R < 1004$	$848 < z < 2710$

Table 3.1 *Inner Detector Parameters. Table taken from Reference [69].*

3.4 Calorimeters

The ATLAS calorimeter system is compromised of two separate calorimeters: the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). The ECAL measures the energy of electrons and photons while the HCAL measures the energy of hadrons. These measurements are done by absorbing the particles. The ATLAS calorimeter system extends to $\eta < 4.9$, having finer granularity in the region of $\eta < 2.5$, where the measurements of electrons and photons are undertaken, and coarser granularity in $\eta > 2.5$ for measurements of jets and missing energy.

3.5 Electromagnetic Calorimeter

The Electromagnetic Calorimeter consists of a component in the barrel from $-1.475 < \eta < 1.475$, a component in the endcap from $-1.375 < \eta < -2.5$ and a component in the endcap from $1.375 < \eta < 2.5$. Electrons interacting with the ECAL lose energy via Bremsstrahlung radiation, yielding a subsequent energy, E , equal to

$$E = E_0 e^{\frac{-x}{X_0}}. \quad (3.5)$$

Here, E_0 is the electron's initial energy, X_0 is the radiation length, and x is the distance traveled. Photons interacting with the ECAL lose energy via e^+e^- pair production, such that the resulting intensity of the photon, I , is equal to

$$I = I_0 e^{\frac{-7x}{9X_0}}, \quad (3.6)$$

where I_0 is the photon's initial intensity and, as before, X_0 is the radiation length, and x is the distance traveled. The radiation length is a characteristic of the material used in the calorimeter. The radiation length is defined as

$$X_0 = \frac{716.4A}{Z(Z+1)\ln(\frac{287}{\sqrt{Z}})}, \quad (3.7)$$

where Z is the atomic number and A is the atomic weight. The Electromagnetic Calorimeter uses lead as the stopping material, as the large Z value yields a small radiation length. The barrel component of the Electromagnetic Calorimeter has a thickness of 22 radiation lengths, while the end-caps have a thickness of 24 radiation lengths. The Electromagnetic Calorimeter has an accordion-like structure, with layers of liquid argon and lead interleaved throughout, as seen in Figure 3.8.

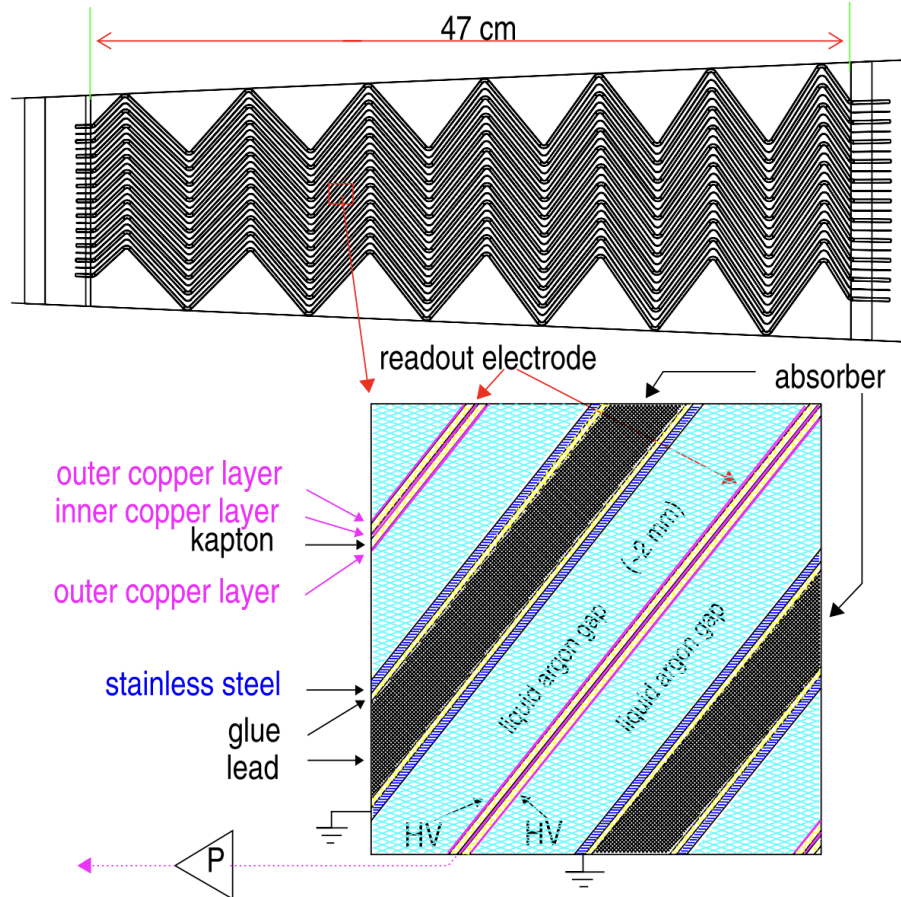


Figure 3.8 View of the accordion structure of the Electromagnetic Calorimeter and the composition of the ECAL. Figure taken from Reference [71].

The liquid argon acts as the active medium, measuring the energy of the particles. The calorimeter is kept at -184 degrees Celcius in order to ensure that the argon remains a liquid. Details of the parameters of the Electron Calorimeter can be found in Table 3.2.

		Barrel	End-cap	
EM calorimeter				
Number of layers and $ \eta $ coverage				
Presampler	1	$ \eta < 1.52$	1	$1.5 < \eta < 1.8$
Calorimeter	3	$ \eta < 1.35$	2	$1.375 < \eta < 1.5$
	2	$1.35 < \eta < 1.475$	3	$1.5 < \eta < 2.5$
			2	$2.5 < \eta < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $ \eta $				
Presampler	0.025×0.1	$ \eta < 1.52$	0.025×0.1	$1.5 < \eta < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$ \eta < 1.40$	0.050×0.1	$1.375 < \eta < 1.425$
	0.025×0.025	$1.40 < \eta < 1.475$	0.025×0.1	$1.425 < \eta < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \eta < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \eta < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \eta < 2.4$
			0.025×0.1	$2.4 < \eta < 2.5$
			0.1×0.1	$2.5 < \eta < 3.2$
Calorimeter 2nd layer	0.025×0.025	$ \eta < 1.40$	0.050×0.025	$1.375 < \eta < 1.425$
	0.075×0.025	$1.40 < \eta < 1.475$	0.025×0.025	$1.425 < \eta < 2.5$
			0.1×0.1	$2.5 < \eta < 3.2$
Calorimeter 3rd layer	0.050×0.025	$ \eta < 1.35$	0.050×0.025	$1.5 < \eta < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$ \eta $ coverage			$1.5 < \eta < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \eta < 2.5$
			0.2×0.2	$2.5 < \eta < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$ \eta $ coverage			$3.1 < \eta < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \eta < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \eta < 3.15,$
				$4.30 < \eta < 4.83$
			FCal2: 3.3×4.2	$3.24 < \eta < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \eta < 3.24,$
				$4.50 < \eta < 4.81$
			FCal3: 5.4×4.7	$3.32 < \eta < 4.60$
		FCal3: $\sim$ four times finer	$3.29 < \eta < 3.32,$	
			$4.60 < \eta < 4.75$	
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$ \eta $ coverage	$ \eta < 1.0$		$0.8 < \eta < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

Table 3.2 *Parameters of the Electromagnetic Calorimeter. Table taken from Reference [69].*

3.6 Hadronic Calorimeter

The Hadronic Calorimeter contains the Tile Calorimeter (TileCal) and the Hadronic End-cap Calorimeter (HEC). The TileCal spans $-1.7 < \eta < 1.7$, while the HEC spans $-3.2 < \eta < -1.5$ and $1.5 < \eta < 3.2$. Hadrons interact via the strong force with the nuclei of the calorimeter material, resulting in hadronic showers. The nuclear interaction length, λ , that dictates the hadronic shows is defined as $\frac{A}{N_A \sigma \rho}$, where ρ is the density of the material, N_A is Avogadro's number, and σ is the inelastic nuclear cross section. The TileCal uses 420,000 plastic scintillating tiles to measure the energy of particles, while steel is used to stop the particles. The TileCal spans from a radius of 2.28 m to 4.25 m, having a total thickness of 9λ .

The HEC contains two wheels for each end-cap, where each wheel consists of 32 wedge-shaped modules shown in Figure 3.9. These wheels are made of copper plates, where the copper acts as the stopping material. Liquid argon is interleaved between the copper plates and serves to measure the energy of the particles.

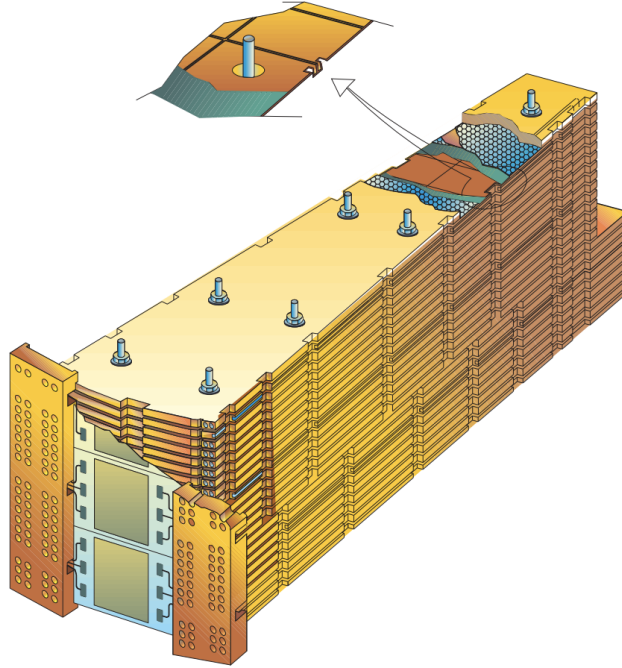


Figure 3.9 *Wedge-shaped module in the HEC. Figure taken from Reference [72].*

3.7 Forward Calorimeter

The Forward Calorimeter (FCAL) ranges from $-4.9 < \eta < -3.1$ and $3.1 < \eta < 4.9$. Liquid argon is similarly used to measure the energy of the particles, while tungsten and copper are used to stop the particles. The FCAL has a thickness of 10 interaction lengths.

3.8 Muon Spectrometer

The Muon Spectrometer consists of four subsystems, as seen in Figure 3.11: the Monitored Drift Tubes (MDTs), the Cathode Strip Chambers (CSCs), the Resistive Plate Chambers (RPC), and the Thin Gap Chambers (TGC). As muons are minimally ionizing particles, muons interact only minimally with the calorimeters. The Muon Spectrometer, consequently, play a vital role in identifying muons, measuring the momentum of muons, and providing valuable information to the trigger. The Monitored Drift Tubes extend from $-2.7 < \eta < 2.7$ and providing tracking information for muons. The MDTs consist of 33 m diameter aluminum tubes with lengths varying from 70 cm to 630 cm. The tubes contain gas with the following composition: 91% Argon, 5% CH₄, and 4% N₂. A tungsten rhenium wire lies within each tube. The ionization of muons in the MDT results in the release of electrons, which are then measured by the wire. The MDTs have a resolution in the z direction of 35 μ m. The CSCs are comprised of two rings of eight chambers, where each chamber contains four wire planes. The CSCs span from $-2.7 < \eta < -2.0$ and $2.0 < \eta < 2.7$, providing a resolution of 40 μ m in the z direction and a resolution of 5 mm radially. The MDTs and CSCs both provide precise measurements of the muon. The RPCs and TGCs are used for triggering. The RPCs span $-1.5 < \eta < 1.5$ and consist of two resistive plates that are parallel and separated by a gas gap. The TGCs are similarly used for triggering and span $-2.4 < \eta < -1.0$ and $1.0 < \eta < 2.4$. TGCs are multiwire chambers containing a 2.8 mm gas gap and provides a resolution of roughly 2 to 7 mm in both the radial and z directions.

Now that the Inner Detector, Electromagnetic Calorimeter, Hadronic Calorimeter, and Muon Spectrometer have all been introduced, a summary of the performance values for different ATLAS subdetectors is provided in Table 3.3. While coverage extends to $\eta \leq 4.9$, the resolution is optimal in the central $\eta < 4.9$ region.

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets) barrel and end-cap forward	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \eta < 4.9$	$3.1 < \eta < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 3.3 *Parameters of various components of the ATLAS detector. Table taken from Reference [69].*

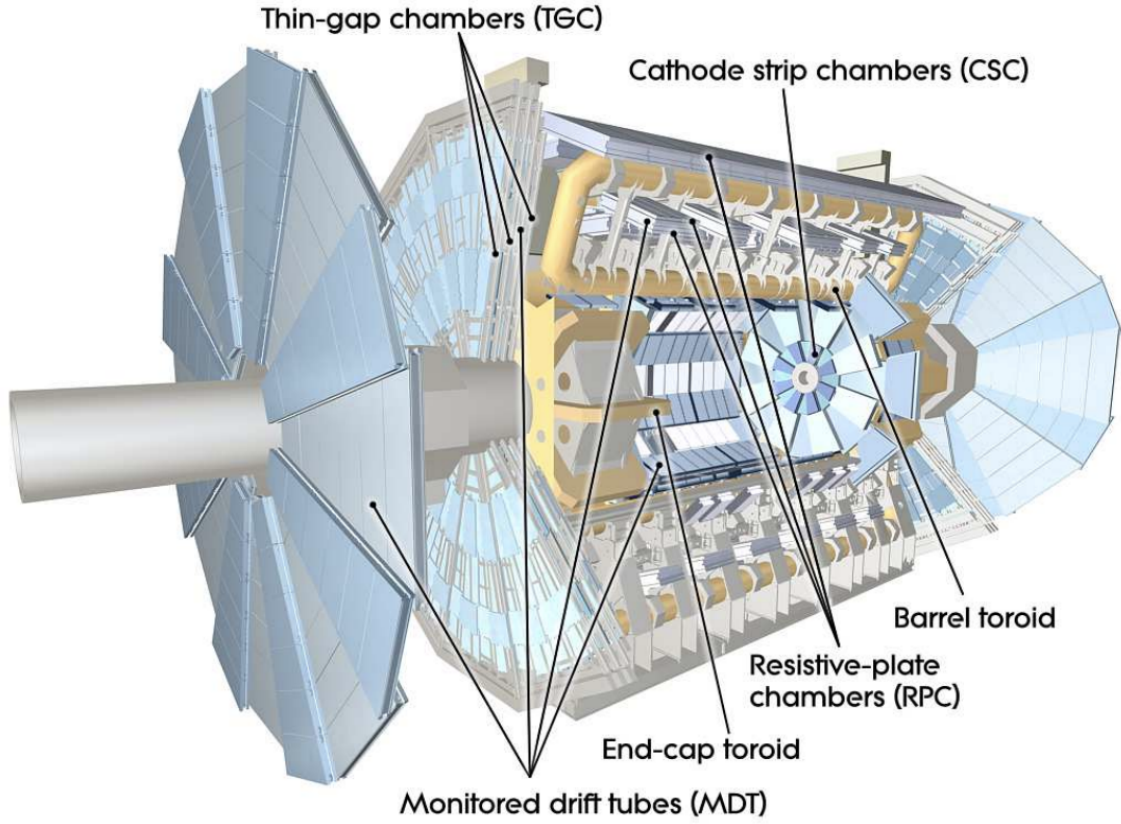


Figure 3.10 *Muon Spectrometer*

Figure 3.11 *ATLAS Muon Spectrometer. Figure taken from Reference [69].*

3.9 Forward Detectors

The Forward Detectors are the ALFA (Absolute Luminosity For ATLAS) detector, the ZDC (Zero Degree Calorimeter) detector, and the LUCID (LUMinosity mea-

surement using Cerenkov Integrating Detector) detector. As the names suggest, these detectors are used to precisely measure the luminosity delivered from the beams. LUCID detector measures, in the forward region, the proton-proton inelastic scattering and, with only a 5% error, gives ATLAS's primary online luminosity measurement. ALFA gives measurements of the absolute luminosity from elastic proton-proton collisions at very small angles. The ZDC detector provides luminosity measures for heavy ion collisions.

3.10 Magnet System

The magnet system within ATLAS consists of both a solenoid and toroid in the barrel and two sets of toroids in the end-caps, as depicted in Figure 3.12. The barrel solenoid extends 1.2 m radially from the beamline and is 5.3 m in length along the beamline. It creates a 2T magnetic field that forces the trajectory of charged particles to bend. The barrel toroid, which produces a 0.5 T magnetic field, and the two end-cap toroids, which produce a 1 T magnetic field, bend the trajectories of muons passing through the Muon Spectrometer.

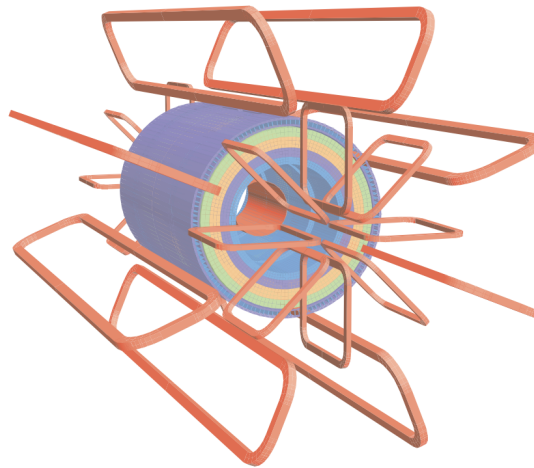


Figure 3.12 *ATLAS Magnet System. Figure taken from Reference [69].*

3.11 Trigger and Data Acquisition System

With an LHC collision rate of 40 MHz, there are far too many events than can be stored. The Trigger and Data Acquisition System, TDAQ, is the means by which ATLAS filters out uninteresting data and stores interesting physics events. The

ATLAS trigger system, seen in Figure 3.13, is comprised of two levels: a hardware level 1 (L1) trigger and a software higher level trigger (HLT) [73]. The HLT is comprised of two levels: the Level 2 trigger and an event filter.

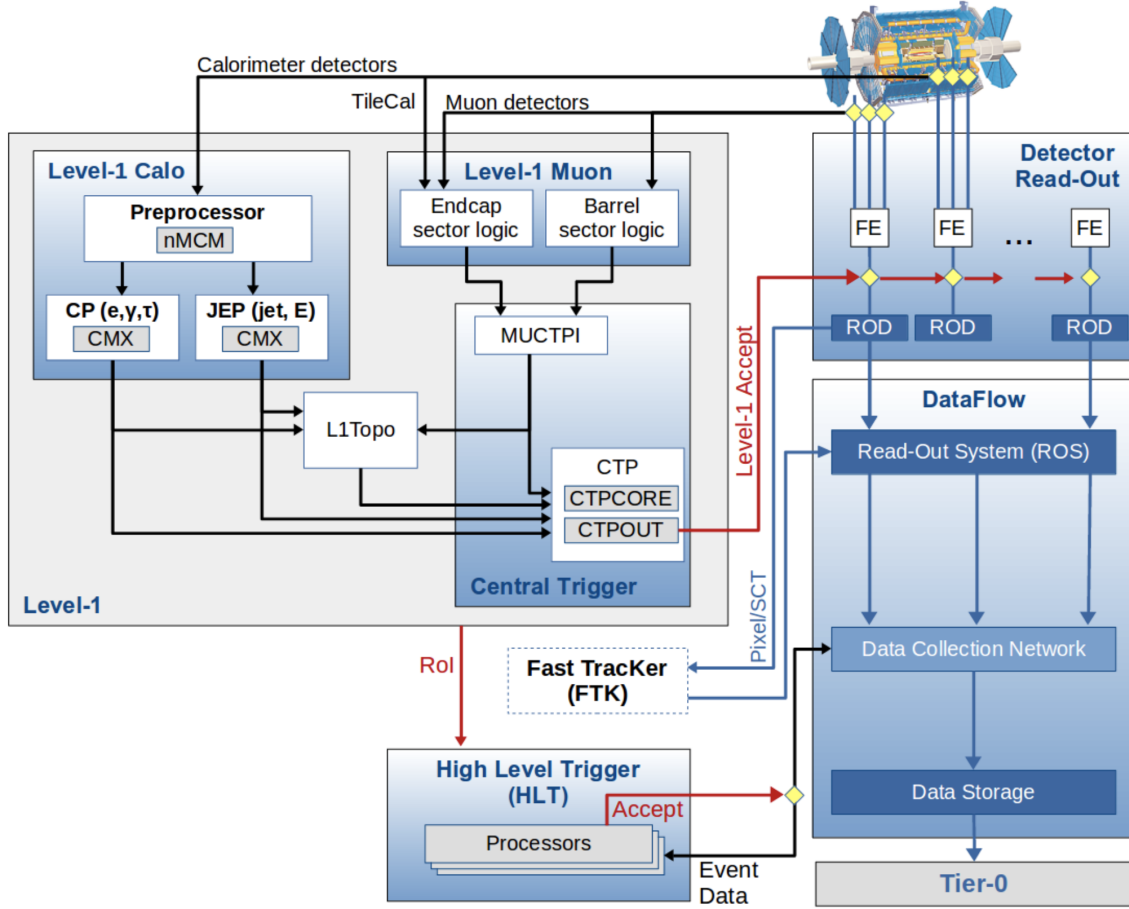


Figure 3.13 ATLAS Trigger System. Figure taken from Reference [74].

3.11.1 L1 trigger

The L1 trigger runs on Field Programmable Gate Arrays (FPGAs) and reduces the event rate to 100 kHz from the original rate of 40 MHz within a span of $2.5 \mu\text{s}$. The L1 trigger takes as input information from the calorimeter and muon spectrometer systems. The L1 trigger reads information from the calorimeter with a granularity of $\Delta\eta = 0.1 \times 0.1$. These regions are called trigger towers and can be seen in Figure 3.14.

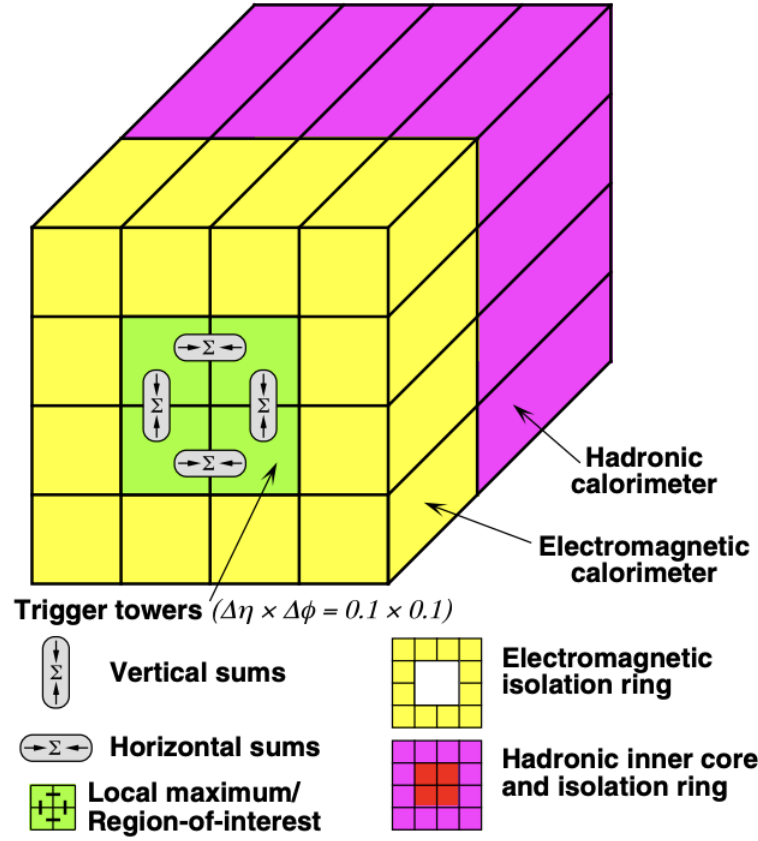


Figure 3.14 *Level 1 trigger tower. Figure taken from Reference [74].*

The RPC and TGC subsystems in the muon spectrometer provide triggering for energetic muons. The central trigger processor (CTP) makes the final decision for the L1 trigger, selecting events which are then sent to the HLT.

3.11.2 HLT

The Level 2 trigger performs further reduction, passing events to the event filter. The event filter reconstructs the event and provides additional selections, yielding a final event rate from the HLT of roughly 1.5 kHz. These events are then passed to permanent storage.

3.11.3 Summary

The LHC is the highest energy particle collider in the world. This chapter explored the different components of the ATLAS detector, all to set the stage for the next

chapter, which explores how particles interact with different parts of the detector leaving different experimental signatures. The next chapter details how the ATLAS experiment uses different algorithms to reconstruct particles so that data can be analyzed.

Chapter 4

Object Reconstruction

The previous chapter introduced the ATLAS detector and its various subdetectors and components. This chapter introduces event reconstruction, the process of transforming raw and uncalibrated hits across a number of detector subsystems into meaningful reconstructed physics objects. Advanced algorithms are used to identify all stable Standard Model particles. As the following chapters will show, the most important physics objects for the analyses in this thesis are tracks (the trajectory calculated from reconstructed hits of charged particles), missing energy (particles that escape the detector without a trace), and jets (the collimated decay products of gluons and quarks). This chapter provides an introduction to all of these objects and places special emphasis on the algorithms used to reconstruct jets. Two main types of jets are described: small radius jets (small R jets) and large radius jets (large R jets). Small R jets take up less area in the η - ϕ plane than large R jets. Both types of jets are key to the analysis. The concept of jet substructure, the internal structure of large R jets, is introduced as a means of distinguishing signal from background. As will be seen, the analysis strategy for all three models will center on large R jets and jet substructure. This chapter provides the building blocks upon which the rest of the thesis will be based.

4.1 Object Reconstruction

4.1.1 Tracks, Vertexing, Topo-clusters

Event reconstruction is the process of transforming raw and uncalibrated hits across a number of detector subsystems into meaningful physics objects. Tracks, vertices, and “topo-clusters” serve as the fundamental building blocks of event reconstruction. The Inner Detector measures charged particles near the interaction point. Charged particles leave hits in the Inner Detector that are then used to form tracks [75]. Tracks are the reconstructed trajectories of charged particles. A particle’s momentum and electromagnetic charge can be indicated by the curvature of the track. The points on the beamline where large quantities of tracks intersect are used to determine collision vertices. The primary vertex is identified as the vertex with the highest scalar sum of transverse momenta of all tracks associated to it [76]. Topo-clusters are formed by the three dimensional topological grouping of calorimeter cells. Each calorimeter cell is assigned a significance $\zeta_{\text{cell}}^{\text{EM}} = \frac{|E_{\text{cell}}^{\text{EM}}|}{\sigma_{\text{cell}}^{\text{EM}}}$, where $\sigma_{\text{cell}}^{\text{EM}}$ includes pileup and noise uncertainties and $|E_{\text{cell}}^{\text{EM}}|$ is the energy of the cell [77]. A topo-cluster seed is a cell in which the significance is greater than 4σ . Cells that are both adjacent (in 3 dimensions) to the seed cluster and that have a significance of 2σ are subsequently added to the topo-cluster seed. The iterative process is continued until all adjacent cells have a significance of less than 2σ . In order to correct for the chance that multiple independent signatures are added to one topo-cluster, a test for a local maximum is performed. A local maximum occurs when a cell has greater than 500 GeV and there are no neighboring cells with greater energy. In the event that a single topo-cluster has more than one local maximum, the topo-cluster is divided along the energy peaks. When a cell neighbors more than one local maximum, the two clusters with the highest energy are assigned the cell. The local cell weighting (LCW) scale and the electromagnetic (EM) scale are two distinct approaches to measuring the energy of a topo-cluster. The sum of energy of all cells in a topo-cluster serves as the definition of total energy for the EM scale. The LCW approach, however, determines whether a cluster is hadronic or electromagnetic. In the case of a hadronic cluster, the LCW applies Monte Carlo based corrections to account for the energy dependent response of the calorimeter to hadronic showers.

4.1.2 Electrons and Photons

Both photons and electrons produce electromagnetic showers and deposit energy in the Liquid Argon Calorimeter. The key distinction, however, comes from the electrically neutral nature of the photons and the electrically charged nature of electrons. Photons, due to their neutrality, do not leave a charged track in the Inner Detector. Identification of electrons is done by matching a cluster of energy in the Liquid Argon detector to a charged track in the Inner Detector. A multivariate algorithm is used to identify interesting electrons (prompt electrons) from uninteresting electrons resulting from background processes such as heavy flavor decays or converted photons [78]. Electron identification is categorized by either a Loose Working Point, Medium Working Point, or Tight Working Point. The shape of the electromagnetic shower and the matching of tracks to clusters are key components upon which the multivariate algorithm makes its distinction. The Loose Working Point offers the weakest background rejection while the Tight Working Point offers the strongest background rejection. Care must be taken when choosing which Working Point to use, as a stronger background rejection can also lead to a higher rejection of signal electrons. Because prompt electrons have a higher chance of being isolated than background events, Isolation Working Points are assigned such that electrons can be chosen or rejected based on the distance between the electron and other objects in the event.

Photons are distinguished from background processes, such as non-prompt photons that arise from, for example, hadron decays, by an identification process that relies on the shape of the electromagnetic shower and isolation requirements.

4.1.3 Muons

Measurements from the Muon Spectrometer, Inner Detector, and calorimeter are used as input by the muon reconstruction algorithm. Sensitivity to muons with transverse momentum ranging the low GeV scale to the low TeV scale is afforded by the use of measurements from multiple subsystems. Tracks from the Inner Detector are reconstructed separately from tracks from the Muon Detector. There are four ways by which a muon candidate is reconstructed. Combined muons are formed from reconstructed tracks in the Inner Detector and reconstructed tracks in the Muon Spectrometer that satisfy a global fit procedure [79]. Extrapolated muons use tracks from the Muon Spectrometer which have a high probability of having

originated from the primary vertex. Extrapolated muons improve acceptance in the central pseudorapidity region of $2.5 < \eta < 2.7$. Segmented muons have Inner Detector tracks that are compatible with at least one segment from the CSC or MDT. Calorimeter-tagged muons use Inner Detector tracks that are matched to a calorimeter energy deposit that is consistent with a minimum-ionizing particle. This is the muon type with the lowest purity. Quality cuts are used to reject background from pion and kaon decays.

4.1.4 Missing Energy

Protons collide in ATLAS perpendicular to the transverse plane. Since the initial transverse momentum is zero, conservation of momentum dictates that the final transverse momentum must be zero. An imbalance in transverse momentum is called missing energy (E_T^{miss}). Missing energy is calculated by taking the magnitude of the negative vector sum of all final state objects in the collision. The final state objects used to calculate the missing energy in this analysis are electrons, muons, and jets, which constitute what ATLAS refers to as the “hard term”, and then objects in what ATLAS refers to as the “soft term”. The “soft term” refers to any detector signal that is unassociated with objects in the “hard term”. These unassociated objects can either be related to unassociated soft calorimeter clusters or unassociated tracks. The definition of the soft term used in this analysis is the track-based soft term, which uses unassociated objects based on tracks and that is more robust against pileup than soft term based on soft calorimeter clusters [80].

The source of E_T^{miss} can be either real or fake. Real E_T^{miss} comes from particles that are invisible to the detector, such as potential Beyond the Standard Model particles that are stable or neutrinos. Fake E_T^{miss} arises from instrumental errors such as noisy detector regions, errors due to calibration, and dead channels.

4.1.5 Jets

Energetic quarks and gluons resulting from the collisions of protons are called partons. These partons can radiate gluons, which can in turn produce quarks via pair production. This process is called fragmentation. This process continues until the particles become less energetic. So as to conserve color and provide color neutral states for these particles, more gluons and quarks are created, but this time from

vacuum. This is called hadronization. The resulting collimated spray of particles from hadronization is called a jet. The parton shower is the phrase that describes this process from start to finish.

Jets leave energy deposits in the calorimeter. These energy deposits are matched with the tracks the jets leave in the Inner Detector. Jet reclustering algorithms use both tracking information and calorimeter information as input. The most common jet reclustering algorithm used in ATLAS is the anti- k_t jet reclustering algorithm [81], an reclustering algorithm that relies on sequential recombination algorithms. Sequential recombination algorithms seek to identify pairs of final state particles that are most likely to have originated from the same initial process [82]. To do so, two distances, d_{ij} and d_{iB} are utilized. The distance d_{ij} is defined in Equation 4.1, while the distance d_{iB} is defined in Equation 4.2.

The distance d_{ij} is defined as

$$d_{ij} = \min(p_{T,i}^{2p} - p_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}. \quad (4.1)$$

Here, i and j represent final state particles or pseudojets. Pseudojets are the term used to refer to particles and their four vector combinations. The distance d_{ij} is the distance between the i^{th} and j^{th} pseudojet. The transverse momentum of the i^{th} pseudojet is denoted as $p_{T,i}$, while the transverse momentum of the j^{th} pseudojet is denoted as $p_{T,j}$. The distance between the i^{th} and j^{th} pseudojet in the η - ϕ plane is denoted as ΔR_{ij} and defined as $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$. The parameter p can take values of -1, 0, and 1 depending on the type of reclustering algorithm, as will be discussed further later. The radius parameter is R .

The distance between the i^{th} pseudojet and the beamline is denoted as d_{iB} and is defined as

$$d_{iB} = p_{T,i}^{2p}. \quad (4.2)$$

The sequential recombination algorithm begins by taking all final state particles and picking a particle (pseudojet) i and a particle (pseudojet) j . If the distance d_{ij} between these two pseudojets is less than the distance d_{iB} between the i^{th} pseudojet and the beamline, then pseudojet i and pseudojet j are combined - ie: the four momenta of pseudojet i and pseudojet j are combined. If the distance d_{iB} is smaller

than the distance d_{ij} , then the pseudojet i is removed from the list and final state particles and declared a jet. This procedure is continued until there are no final state particles left.

As noted earlier, the value of the parameter p depends on the type of reclustering algorithm. For the k_t reclustering algorithm, the value of p is 1. For the Cambridge/Aachen algorithm, the value of p is 0. For the anti- k_t algorithm, the value of p is -1. Given the value of p , the k_t reclustering algorithm necessarily combines softer final state particles first in the sequential recombination process. Due to the value of p , the anti- k_t algorithm preferentially combines final state particles with higher transverse momentum, resulting in largely circular and compact jets. The k_t algorithm, however, preferentially combines final state particles with softer (lower) transverse momentum first. The jets resulting from the anti- k_t algorithm, the k_t algorithm, and the Cambridge/Aachen algorithm can be seen in the figure below.

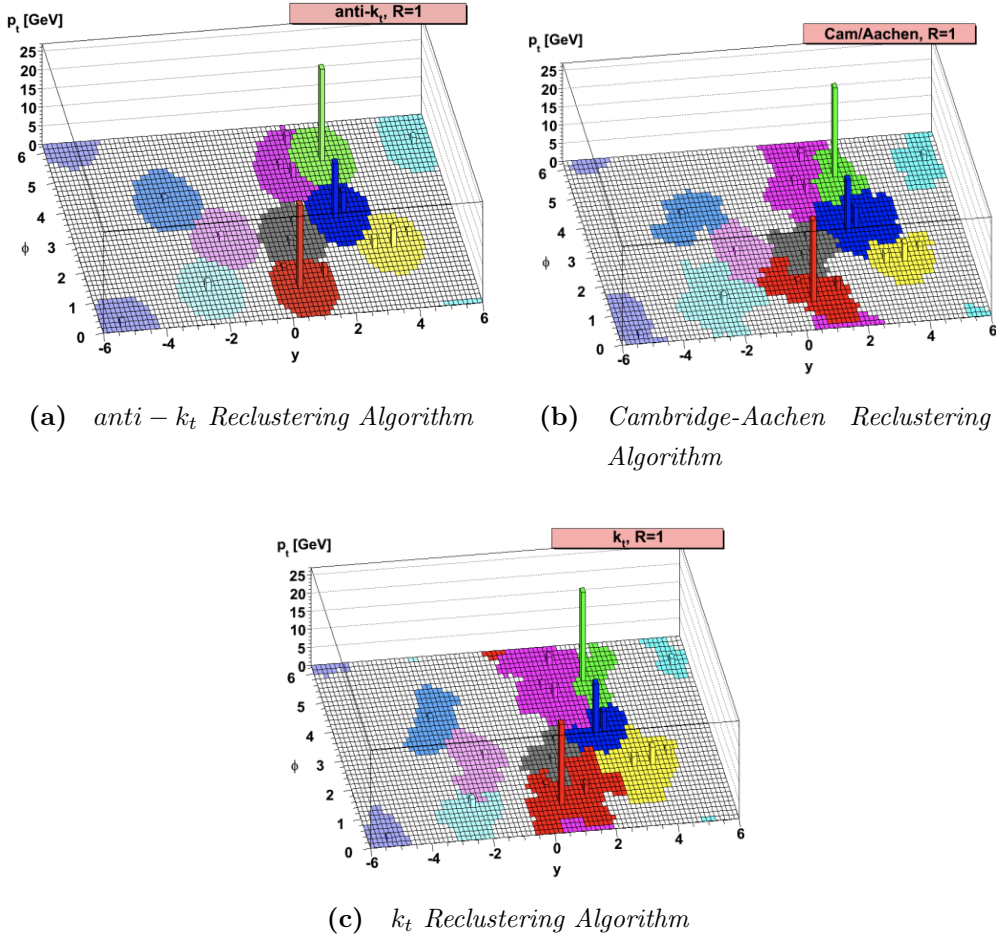


Figure 4.1 *Illustration of jet clustering with three different jet reclustering algorithms applied to the same event. Each colored region corresponds to a different reconstructed jet. Figure taken from [81]*

To be robust, a jet reclustering algorithm must be both infrared safe and collinear safe. To be infrared safe, the presence of additional particles with low transverse momentum must not alter the set of jets output by the reclustering algorithm. To be collinear safe, the set of jets output by the reclustering algorithm must be unaffected by the collinear splitting of a jet. For example, in the event that the hardest particle is split into a pair of particles, the algorithm must recognize that pair as the jet with the highest transverse momentum and not consider a particle that has lower transverse momentum than the original hardest particle but higher transverse momenta than the particles resulting from the collinear splitting to be the hardest jet. Since the anti- k_t algorithm clusters particles with the highest transverse momenta first, the anti- k_t algorithm is both infrared safe and collinear safe.

The Jet Vertex Tagger (JVT) [83] is a variable that ATLAS uses to distinguish and reject jets arising from pileup events. The primary vertex was introduced earlier in this chapter as the vertex with the highest scalar sum of transverse momenta of all tracks associated to it. The hard scatter is the interaction of interest. Pileup is additional soft (low transverse momentum) radiation in an event, arising from proton-proton interactions other than the collision of interest, and is typically much softer than the hard scatter. In order to determine the primary vertex, multiple vertex candidates are reconstructed. The highest ranking candidate (which will be referred to in the equations below as PV_0) is taken to be the hard scatter. The other candidates for the primary vertex are attributed to pileup and are referred to as PV_n in the subsequent discussion.

The JVT is a multivariate combination of two independent track-based variables: the corrected Jet Vertex Fraction (corrJVT), depicted in Equation 4.3, and the ratio of track p_T to jet p_T (R_{p_T}), depicted in Equation 4.4.

The corrJVT is defined as

$$\text{corrJVT} = - \frac{\sum_k p_T^{\text{trk}_k}(PV_0)}{\sum_l p_T^{\text{trk}_l}(PV_0) \frac{\sum_{n \geq 1} \sum_l p_T^{\text{trk}_l}(PV_n)}{k \cdot n_{\text{trk}}^{\text{PU}}}}. \quad (4.3)$$

The scalar transverse momentum sum of tracks that originate from the hard scatter and that are associated with the jet is denoted by $\sum_k p_T^{\text{trk}_k}(PV_0)$. The scalar transverse momentum sum of tracks that stemming from any of the pileup (PU) interactions is written as $\sum_{n \geq 1} \sum_l p_T^{\text{trk}_l}(PV_n)$. This quantity, $\sum_{n \geq 1} \sum_l p_T^{\text{trk}_l}(PV_n)$, is divided by $k \cdot n_{\text{trk}}^{\text{PU}}$, where $n_{\text{trk}}^{\text{PU}}$ is the total number of pileup tracks per event and k is

0.01. Here, the k value of 0.01 is approximately the slope of the linear relationship of $\sum_{n \geq 1} \sum_l p_T^{\text{trk}_l}(\text{PV}_n)$ with total number of pileup tracks per event. This is a standard ATLAS calculation. This division occurs in order to correct for the linear increase of $\sum_{n \geq 1} \sum_l p_T^{\text{trk}_l}(\text{PV}_n)$ as the total number of pileup tracks per event increases.

The track-based variable R_{p_T} is the scalar transverse momentum sum of tracks originating from the hard scatter and vertex and associated with the jet, divided by the transverse momentum of the calibrated jet. It is defined as

$$R_{p_T} = -\frac{\sum_k p_T^{\text{trk}_k}(\text{PV}_0)}{p_T^{\text{jet}}}. \quad (4.4)$$

The ability to distinguish between jets arising from pileup and jets arising from the hard scatter can be seen in Figure 4.2, where jets originating from pileup are peaked at JVT values of 0 and jets originated from the hard scatter are peaked at JVT values of 1. Both jets originating from the hard scatter and from pileup do, however, have tails that extend to lower and higher JVT scores, respectively, indicating that the JVT can only reduce the contribution of jets arising from pileup and not fully eliminate them.

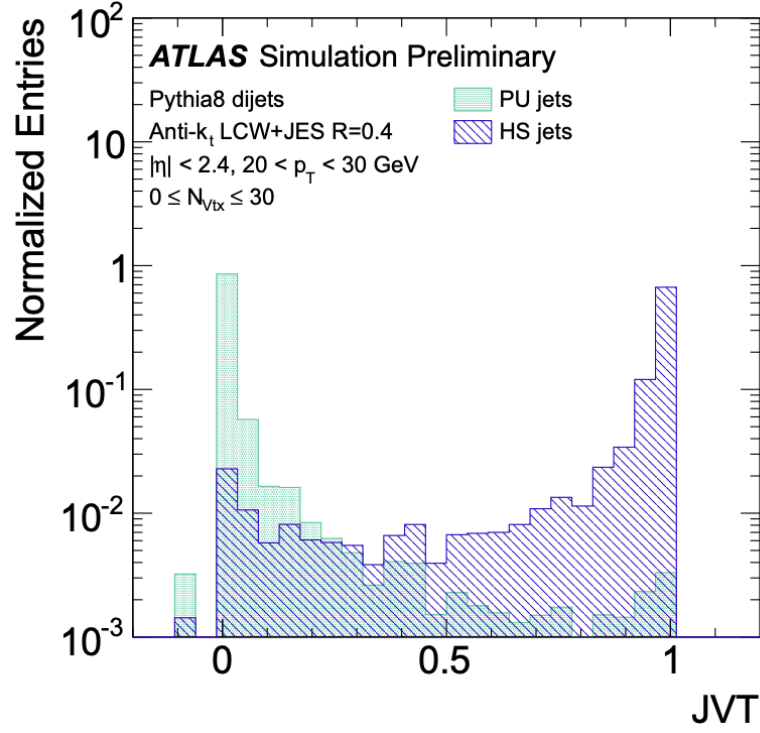


Figure 4.2 *Distribution of Jet Vertex Tagger scores for jets originating from pileup (PU) and the hard scatter (HS) and with transverse momentum ranges from 20 GeV to 30 GeV. Figure taken from [83]*

Due to the JVT being a track-based quantity, its use is restricted to the η range of the tracker: $-2.4 < \eta < 2.4$. The forward Jet Vertex Tagger (fJVT), however, identifies and distinguishes forward pileup jets from jets originating from the hard scatter.

4.1.6 Small R jets

One of the most common jets used in ATLAS is the small radius, or “small R ”, jet. This jet is created by using a jet reclustering algorithm with a radius, R , of 0.4 in the η - ϕ plane. There exist different types of small R jets in ATLAS. The searches in this thesis uses small R Particle Flow jets (PFlow jets), or jets created using the Particle Flow algorithm. The Particle Flow algorithm is an algorithm that runs the anti- k_t jet reclustering algorithm with both topo-clusters and tracks provided as input [84].

4.1.7 Large R jets and Jet Substructure

Another prevalent jet used in ATLAS is the large radius, or “large R ”, jet. The generic, standard large R jet used in ATLAS is created via the anti- k_t jet reclustering algorithm and has a radius, R , of 1.0. (It is not practical for ATLAS to provide standard large R jets with a variety of different radii as it takes dedicated teams of dozens of people to provide even one fully validated and calibrated large R jet to the collaboration. The choice to use $R = 1.0$ is simply because that radius fits the physics needs of most analyses and groups in ATLAS reliant on large R jets.)

Standard large R jets with a transverse momentum of at least 200 GeV are fully supported within ATLAS insofar as they are calibrated. Large R jets are of use when the physics process of interest has a collimated spray of particles spanning wider than the radius, R , of 0.4 in the η - ϕ plane. Additionally, large R jets are of particular interest and value when the physics process of interest produces distinct clusters, or subjets, within a given radius, where the radius is greater than $R = 0.4$. The existence of subjets within large R jets is referred to as jet substructure. Jet substructure provides a means of differentiating non-interesting Standard Model background processes from interesting physics processes. A QCD jet (or a jet arising from Quantum Chromodynamics processes) exhibits a one prong structure, in which there is no distinctive substructure. This can be seen in Figure 4.3 [85].

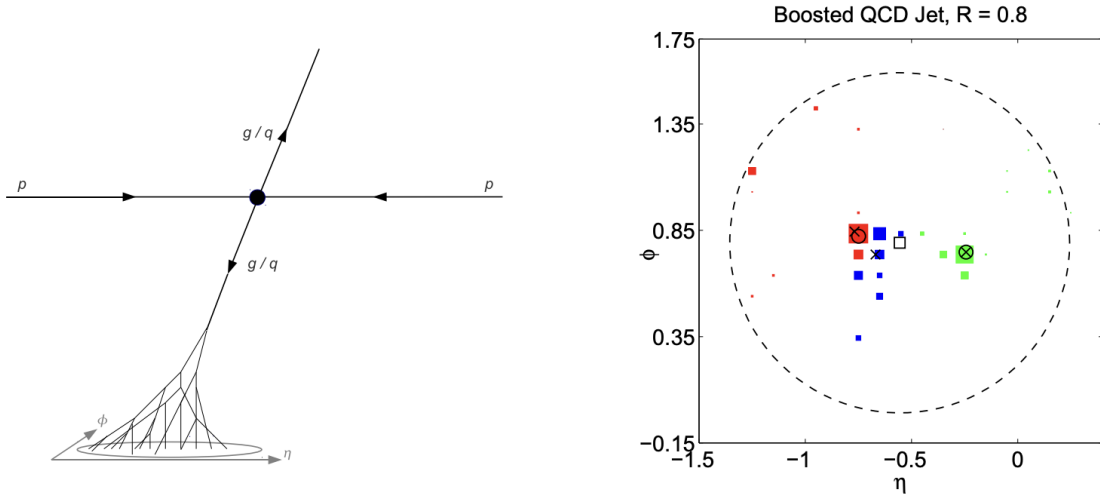


Figure 4.3 *Energy spread of a boosted QCD jet with $R = 0.6$. Figure taken from [85]*

W and Z processes, when characterized by high transverse momentum, are considered “boosted” and can exhibit a two prong structure, indicating jet substructure.

This can be seen in Figure 4.4.

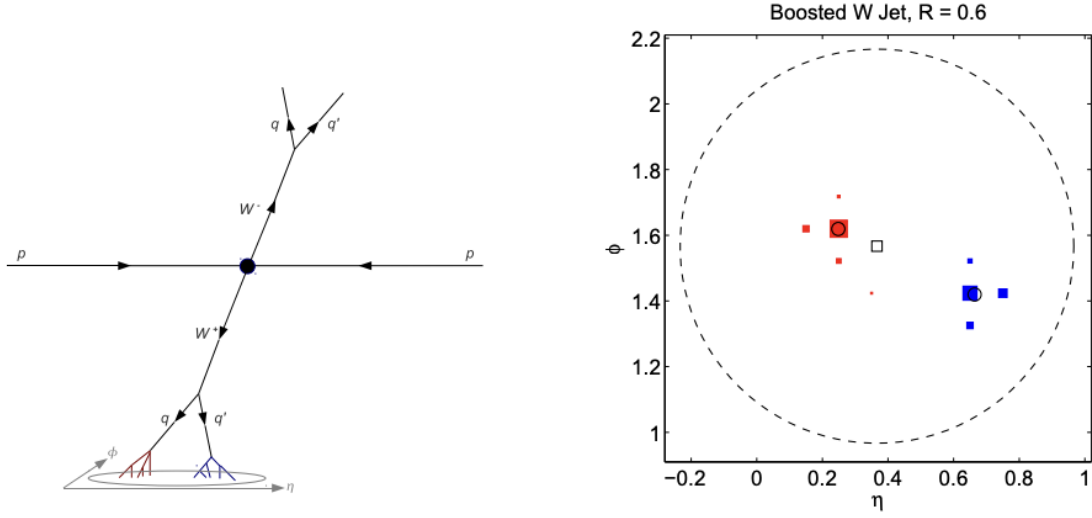


Figure 4.4 Energy spread of a boosted W jet with $R = 0.6$. Figure taken from [85]

Top processes, when boosted, can exhibit a 3 prong structure, indicating jet substructure. This can be seen in Figure 4.5. Exotic processes with two or three prong structures can readily be distinguished from QCD multijets. Exotic processes with greater than 3 prong structure thereby gain distinction from boosted W , Z , and top processes, as well as QCD multijets.

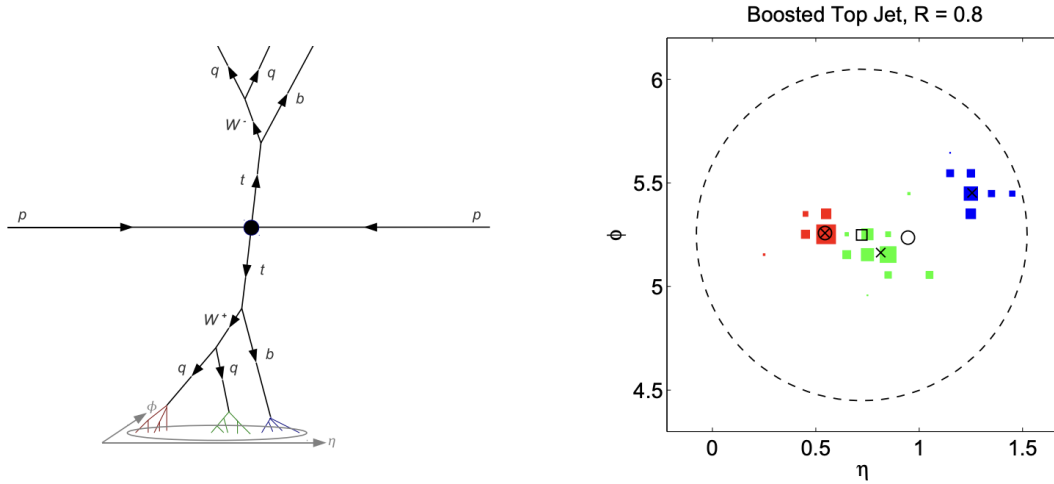


Figure 4.5 Energy spread of a boosted top jet with $R = 0.8$. Figure taken from [85]

Exotic processes with a two prong structure can be distinguished from a W or Z with a two prong structure and exotics processes with a three prong structure can

be distinguished from a top with a three prong structure in a number of ways. For example, the mass or transverse momentum of the large R jet can be used to distinguish the exotics process for the Standard Model processes.

The strategy for the analyses in this thesis relies on the use of large R jets. However, as will be shown in future chapters, the mass or transverse momentum of the large R jet are not powerful discriminants for the analyses in this thesis to distinguish signal from background. Instead, this analysis uses the position and transverse momentum of the tracks in the large R jet to differentiate exotic signal from background. A number of standard jet substructure observables exist for studying the energy distributions within large R jets. Instead of using such observables, this analysis uses machine learning to learn the jet substructure of the large R jets from the track information and distinguish signal from background.

An important point to highlight is the 200 GeV threshold requirement for the transverse momentum for standard ATLAS large R jets. As will be seen in future chapters, this 200 GeV requirement on transverse momentum is often too high for large R jets containing the decay products of the exotic processes considered in this thesis. This, and the reliance of the analyses on tracks within the large R jets, will motivate the creation of a new reclustering algorithm for large R jets.

4.1.8 Summary

This chapter discussed the algorithms used in ATLAS to reconstruct particles so that data can be analyzed. The key physics objects for the analyses in this thesis, namely tracks, jets, and missing energy, were introduced. Jet substructure, upon which the analysis strategy will center, was also introduced. High energy collider experiments are constantly developing and improving algorithms in order to adapt to the changing detector conditions and in order to best identify interesting events from uninteresting background events. The next chapter discusses algorithms developed to accommodate the increase in luminosity from Run 2 to Run 3.

Chapter 5

Trigger Upgrade and jFEX Jet Calibrations

The previous chapter described the various reconstruction algorithms used to identify physics objects so that ATLAS data can be analyzed. This chapter describes reconstruction algorithms designed to improve jet identification in the ATLAS trigger system for Run 3. The chapter starts with an introduction to the upgrades for the ATLAS trigger system for Run 3, paying special emphasis on the inclusion of the jet Feature EXtractor (jFEX), a hardware-based jet finder for the upgraded level 1 (L1) trigger system. The jFEX identifies both small R jets at L1 (called small R jFEX jets) and large R jets at L1 (called large R jFEX jets). Due to the detector being nonuniform in η , jFEX jets need to be calibrated. This chapter describes calibrations derived by the author for small R jFEX jets, discusses how the optimal calibration was chosen, and shows the improvements afforded by the calibration.

5.1 Upgrade

During the Long Shutdown 2 that took place between late 2018 and early 2022, the ATLAS trigger system underwent several key upgrades [86]. Two major sets of upgrades to the ATLAS level 1 trigger system were the LAr supercell [87] and the Feature EXtractors (FEXs): Jet FEX (jFEX), global FEX (gFEX), and electromagnetic FEX (eFEX) [88]. These upgrades can be seen in Figure 5.1, which shows the old L1 Calo system in dashed red lines, with the new upgrades outside

the lines.

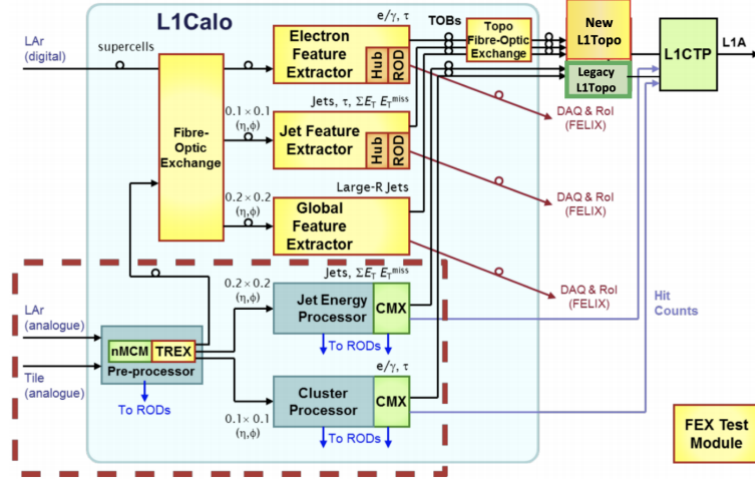


Figure 5.1 *L1 Calorimeter from Run 2 and Run 3. Figure taken from [88]*

There are six jFEX modules. Each jFEX module, seen in Figure 5.2, is an ATCA (Advanced Telecommunication Computing Architecture) board with 4 processor Xilinx UltraScale+ FPGAs (field-programmable gate array) and can process a total input data rate of 3.6 Tb/s [88].

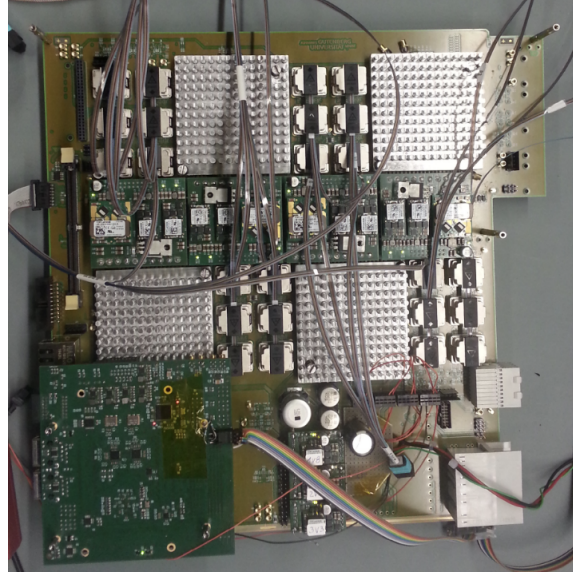


Figure 5.2 *jFEX prototype. Figure taken from Reference [88]*

The jFEX identifies jets, large-area taus (leptons with conical shape footprint in the detector), and calculates missing transverse momentum and other energy sums. The new jFEX jets have much finer resolution in η - ϕ than Run 2 L1 jets, and have more

calibration options available. A jFEX jet is a 0.7×0.7 rounded-off square with a 0.1×0.1 tower granularity and 0.3×0.3 seed, as shown below. The tower depicted in figure below is a jTower. A L1 jet region of interest (ROI) is a 0.8×0.8 square with a 0.2×0.2 tower granularity and 0.4×0.4 seed [88].

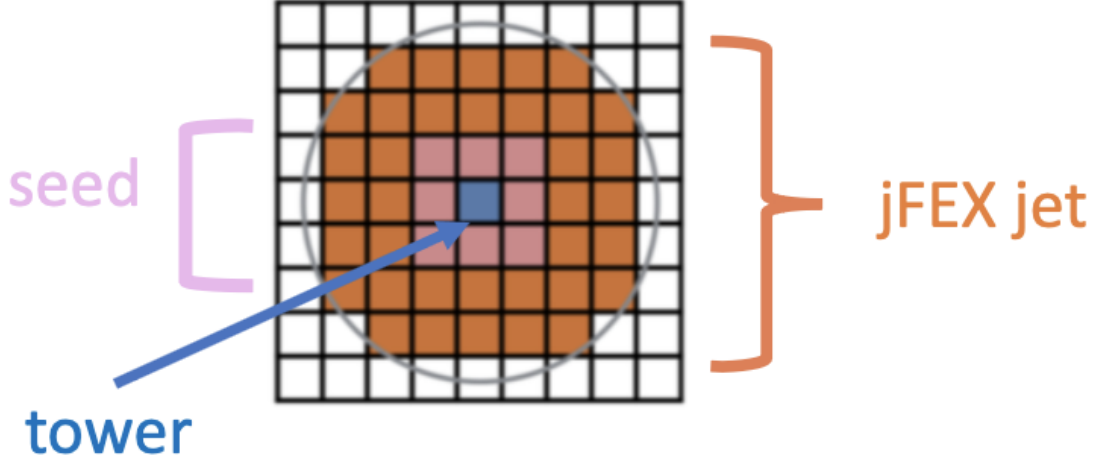


Figure 5.3 *Small R jFEX jet for Run 3. Figure modified from [89]*

A small R jFEX jet is reconstructed using the sliding window algorithm [89]. A noise threshold of 2 GeV is required for any jTower so that noise from pileup and electronics can be suppressed. If a jTower does not meet the noise threshold, the E_T of the jTower is set to 0. Each 0.1×0.1 jTower is surrounded by a 0.3×0.3 seed that consists of 3×3 jTowers. The local maximum, or the seed with the highest E_T , is determined by scanning all seeds that live within a range of 5×5 jTowers and finding the seed with the highest E_T [89]. A small R jFEX jet is taken to be a cone of $\Delta R = 0.4$, where R is $\sqrt{\eta^2 + \phi^2}$, with the $\Delta R = 0$ point originating from the center of the local maximum. The E_T of each jTower in the cone is summed to yield the jet E_T [89]. L1 jet ROIs are massless, so p_T equals E_T for L1 jets. (At L1, ATLAS works with energy sums and does not consider any mass components, doing massless jet reconstruction by design. Then, later, at the offline reconstructed level, more information is available, such as the fully reconstructed four-vector jet, as well as additional information from other detector parts, allowing for a fully optimized object. ATLAS, therefore, transitions from energy calibrations at L1 to calibrations that incorporate mass at the offline reconstructed level. At LHC energies, L1 jet E_T and offline jet p_T would correlate reasonably well.)

Large R jFEX jets are formed using the sliding window algorithm, but with a window of 1.7 to reflect that a cone of $\Delta R = 0.8$ is used for Large R jFEX jets, shown in

Figure 5.4.

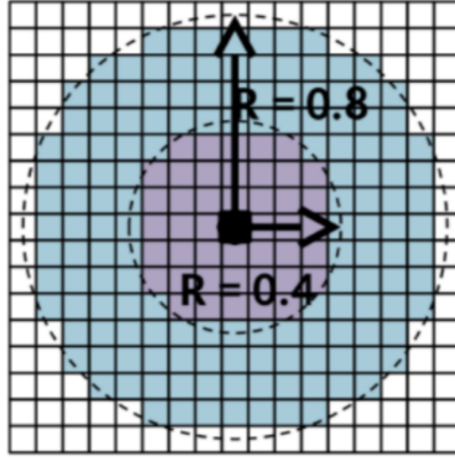


Figure 5.4 Large R jFEX jet for Run 3. Figure taken from [89]

This differs from the Run 2 L1 large R jet approach, which implemented a simple cone algorithm that drew a circle of radius 1.0 around each small R jet and calculated the large R jet E_T by summing the E_T for each small R jet within the circle, vetoing an E_T from the sum if the E_T was less than 15 GeV [89]. The incorporation of jFEX and gFEX in Run 3 has caused L1 jet triggers to change significantly [89]. To maintain the efficiency of the L1 and HLT jet chains and to maximize use of L1 rate, it is crucial that L1 jet triggers perform as well as possible. To that end, it is vital to study the performance of the new L1 jet triggers, derive potential calibrations, and determine the resulting L1 jets that should seed each HLT chain. Improved granularity and more-accurate shape of jFEX jets allows for improvements in jFEX trigger turn-ons and gives better prospects for L1 jet calibration in Run 3 than in Run 2. Jet triggers for low-threshold multijet triggers and for Trigger-level analyses benefit from polished jFEX jets and L1 calibration. Calibration of the jFEX jets is needed because the ATLAS detector is not uniform in η , so the jet response is not uniform in η . The goal for the trigger is to select all jets that had true p_T above X . For a response, R , that is constant everywhere, the L1 selection would be at $X \cdot R$. If R is a function of η , a cut at a single value $X \cdot R$ will select truth jets at different p_T across η . Figure 5.5 shows the high variance of energy response over η without calibrations and thereby demonstrates the need to calibrate. (A description of truth particles: output of MC event generator with full information of the physics all of the particles involved. This is an unphysical level of detail because we know exactly what happened in our simulation. Then reconstructed quantities are the non-idealized versions of this data that would be reproduced by a real experiment.)

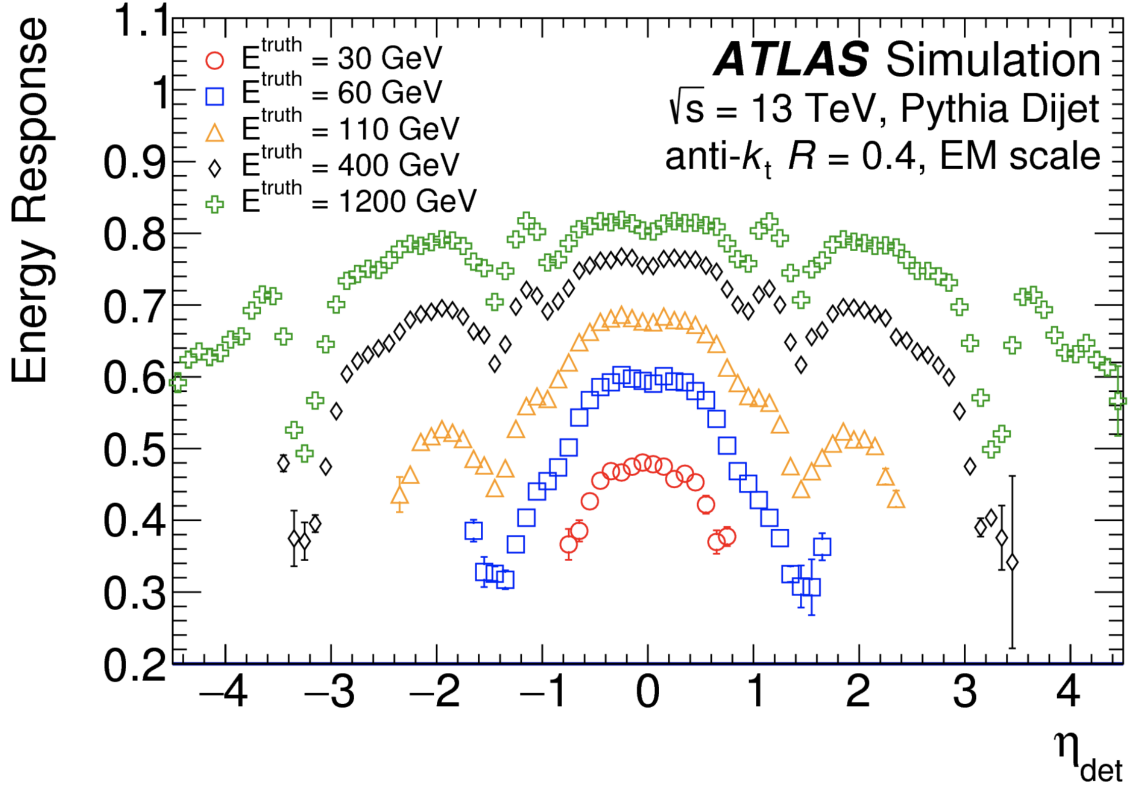


Figure 5.5 *High variance of jet energy response over η . Figure taken from [74]*

5.2 jFEX Jet Calibration

A calibration can be derived to correct for the nonuniformity of jet response across η . The ideal goal of calibrations is to allow for a response to be 1. Isolated, reconstructed jFEX jets are matched to truth jFEX jets, serving as an initial jet energy scale calibration. A central ATLAS framework called IsolatedJetTree [90] is used to perform the isolation of the reconstructed jFEX jets and match these jets to truth jFEX jets. These outputted isolated and truth-matched reconstructed jFEX jets are then used as input to code developed by the author to derive the calibration via steps that will be explained now. All subsequent plots in this chapter were created by the author.

The distribution of $(\text{reconstructed energy})/(\text{truth energy})$ for jFEX jets is binned in η and truth energy and is fitted with a Gaussian. The mean of Gaussian fit is the average energy response. The average energy response is parameterized as a function of reconstructed energy and η . The inverse of the average energy response in each bin of reconstructed energy and η becomes the jet calibration factor for each

bin of reconstructed energy and η . Two approaches for the fit were used for the jFEX jet calibration: the first being the Gaussian fit already described, and the second being more naive, simply taking the mean value of the bin of (reconstructed energy)/(truth energy), deemed the response, vs. η . The second approach, simply utilizing the mean, provided a better calibration, so this is the approach used in the final calibration.

An entirely separate calibration that can also be derived using p_T instead of energy. The procedure for the p_T calibration is exactly the same as the energy calibration with the only difference being that p_T is used instead of energy. As the p_T calibration proved to be more effective than the energy calibration, only the results of the p_T calibration follow. Calibration of jFEX jets allows for better jFEX jets and ultimately the ability to probe lower p_T .

The p_T calibration was performed with the following p_T binning: minimum p_T to 20 GeV, 20 GeV to 30 GeV, 30 GeV to 40 GeV, 40 GeV to 50 GeV, 50 GeV to 65 GeV, 65 GeV to 80 GeV, 80 GeV to 110 GeV, 110 GeV to 150 GeV, 150 GeV to maximum p_T . The p_T calibration used the following η binning: -4.9 to -1.6, -1.6 to -0.8, -0.8 to 0, 0 to 0.8, 0.8 to 1.6, and 1.6 to 4.9. The response map provided to hardware experts is depicted in Figure 5.6. The calibration factor is the inverse of the value in each bin of p_T and η .

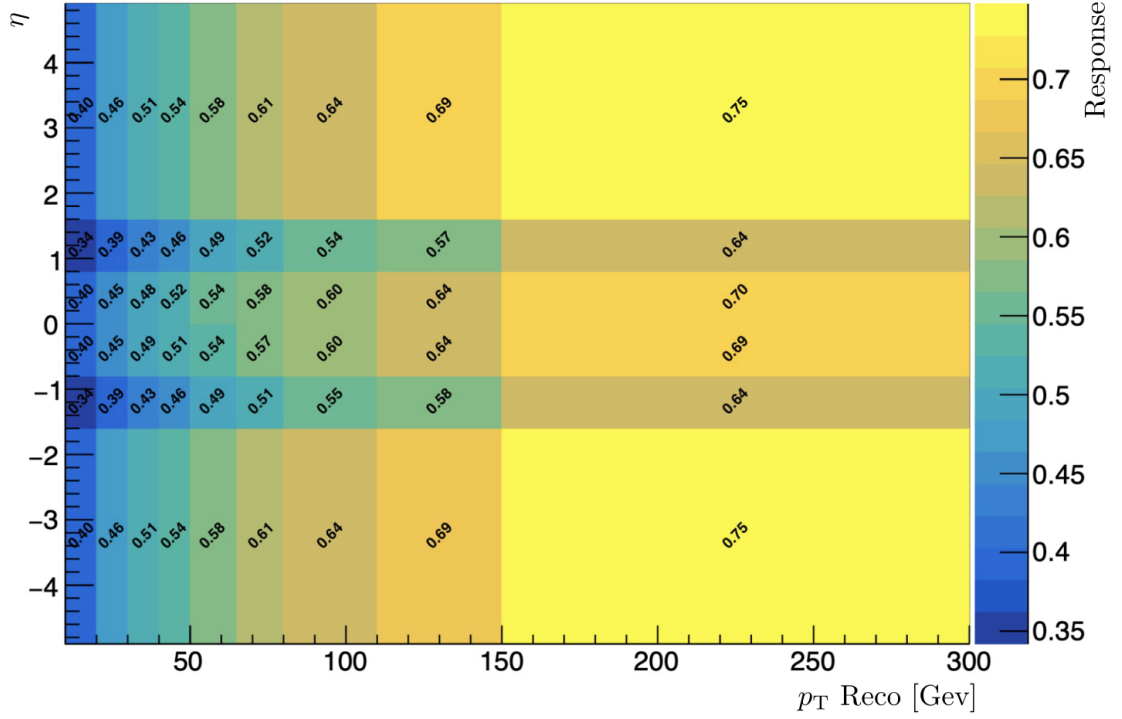


Figure 5.6 *Response map for hardware. The p_T response is plotted as a function of η and jFEX reconstructed jet p_T .*

The power of the calibration can be seen in Figure 5.7 and Figure 5.8. The p_T response, or the reconstructed p_T over truth p_T , of the uncalibrated jFEX jets can be seen to be highly variant and well below 1 for all p_T - η bins. The calibrated jFEX jets have a p_T response of roughly 1 across all bins of p_T and η , except for the first bin, which extends from 0.85 towards 1. As the goal of the calibration was to make the p_T response of the jFEX jets roughly 1, the calibration is shown to be successful.

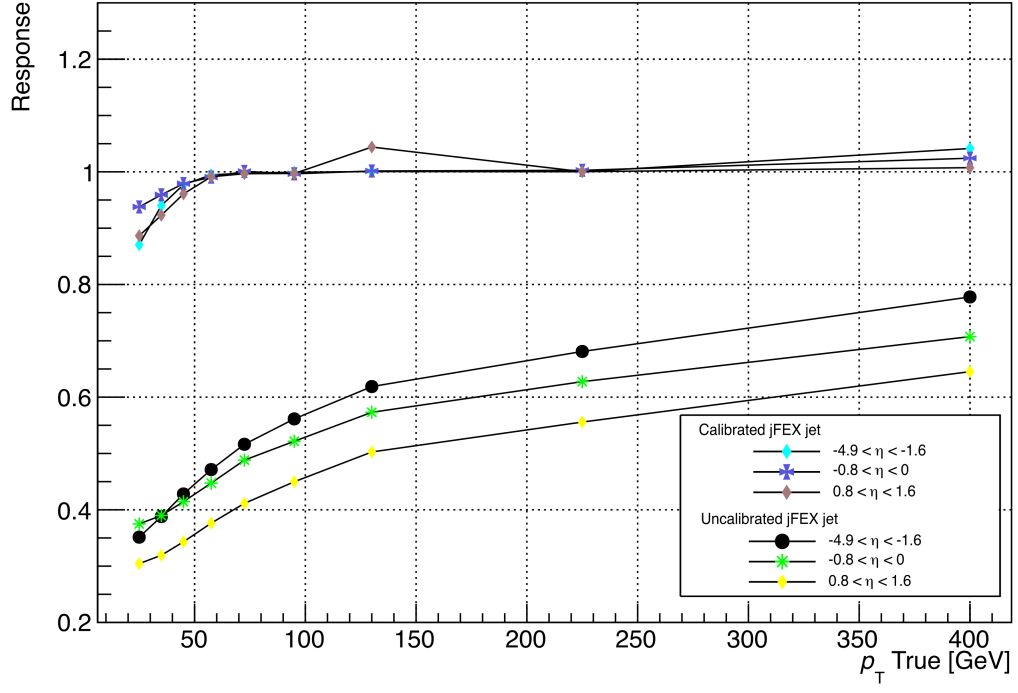


Figure 5.7 Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing p_T response as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-4.9 < \eta < -1.6$, $-0.8 < \eta < 0$, and $0.8 < \eta < 1.6$. The p_T of the jFEX truth jet is labeled as $p_T \text{ True}$.

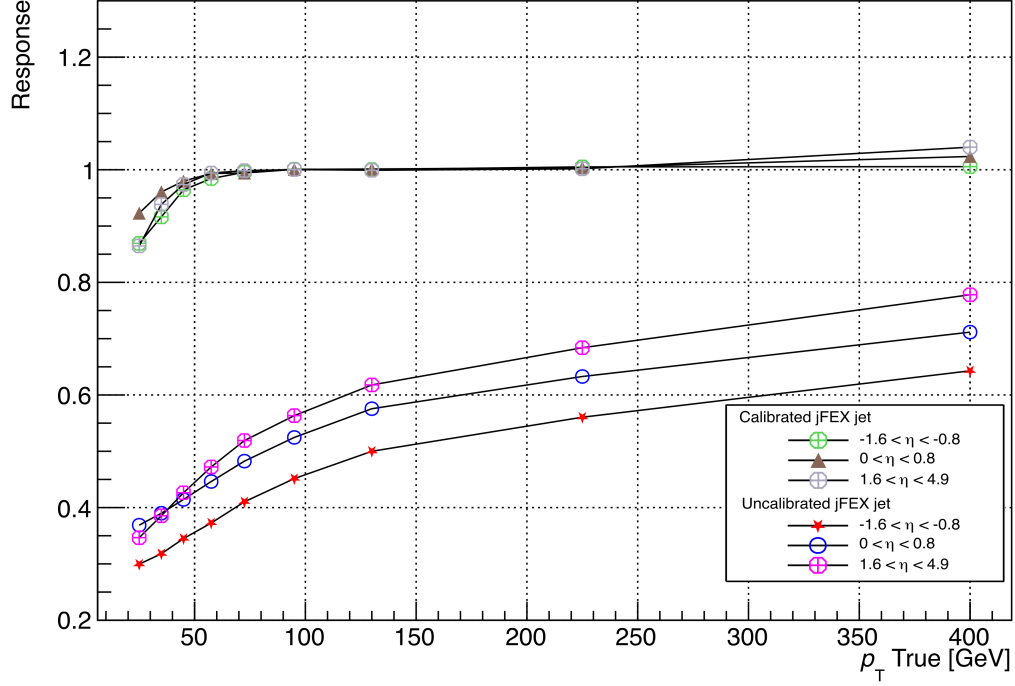


Figure 5.8 Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing p_T response as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-1.6 < \eta < -0.8$, $0 < \eta < 0.8$, and $1.6 < \eta < 4.9$. The p_T of the jFEX truth jet is labeled as $p_T \text{ True}$.

The resolution in each p_T bin is calculated by dividing the standard deviation of the bin by the mean of the bin. (In the Gaussian fit approach, the resolution was calculated by dividing the standard deviation of the fit in the bin by the mean of the Gaussian fit.) A successful calibration would allow for an improved resolution - or a smaller resolution - in each bin, which is exactly what is observed in Figure 5.9 and Figure 5.10.

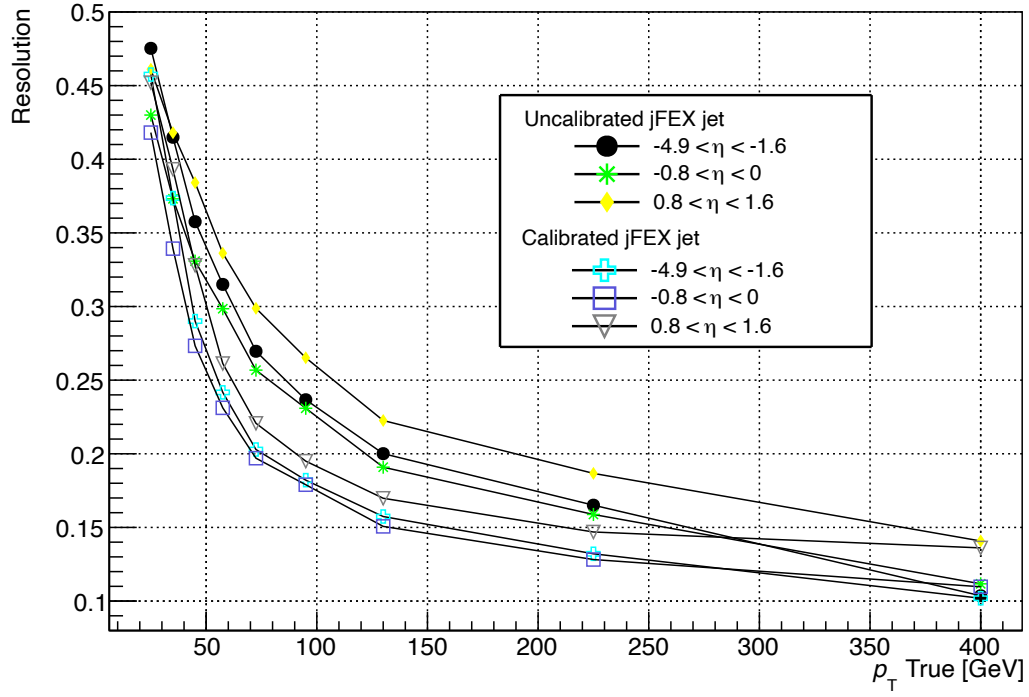


Figure 5.9 Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing resolution as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-4.9 < \eta < -1.6$, $-0.8 < \eta < 0$, and $0.8 < \eta < 1.6$. The p_T of the jFEX truth jet is labeled as p_T True.

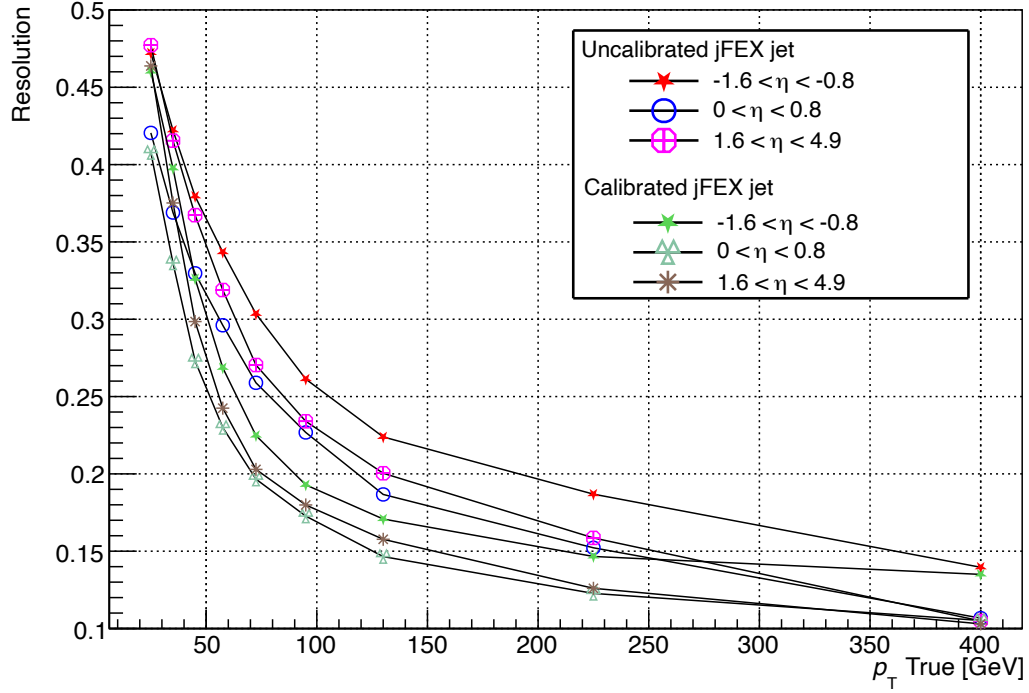


Figure 5.10 Comparison of the performance of the calibrated jFEX jets to the performance of the uncalibrated jFEX jets by showing resolution as a function of p_T of jFEX truth jets. Results shown for the following η bins: $-1.6 < \eta < -0.8$, $0 < \eta < 0.8$, and $1.6 < \eta < 4.9$. The p_T of the jFEX truth jet is labeled as p_T True.

For each $p_T - \eta$ bin, the calibrated jFEX jets have a smaller resolution than the uncalibrated jFEX jets. This, coupled with the improved responses, indicates a successful calibration.

The binning of p_T calibration was chosen to make the firmware implementing easier. A calibration that was more finely binned in η was also performed. The η binning for the fine calibration was the following: -4.5 to -4.0 , -4.0 to -3.5 , -3.5 to -3.0 , -3.0 to -2.5 , -2.5 to -2.0 , -2.0 to -1.5 , -1.5 to -1.0 , -1.0 to -0.5 , and -0.5 to 0 . The same binnings hold for the positive values of η . The performance of the more coarsely binned p_T calibration and the more finely binned p_T calibration are shown in Figure 5.11 and Figure 5.12. Both calibrations have responses close to 1 and improved resolution. There is no significant degradation resulting from using coarser η bins, so the p_T calibration with the coarser η binning was used to make things easier for the hardware experts.

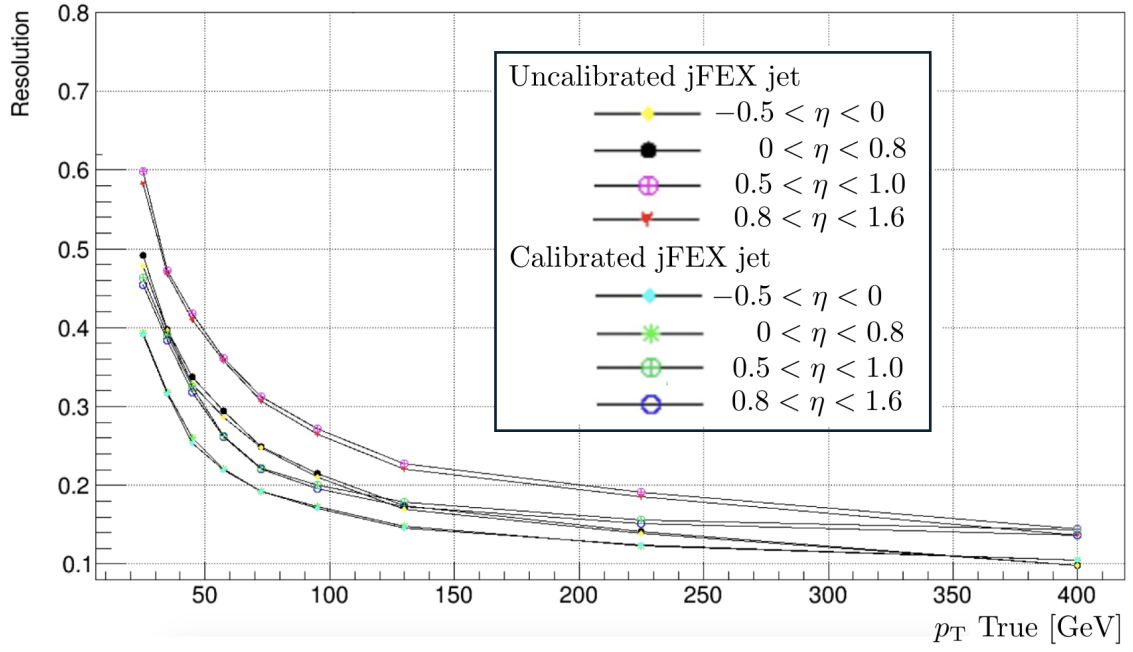


Figure 5.11 Comparison of the performance of the jFEX calibration using fine η binning to the performance of the jFEX calibration using coarser η binning by showing resolution as a function of p_T of jFEX truth jets. Results shown for coarse η bins $0 < \eta < 0.8$ and $0.8 < \eta < 1.6$ and the following fine η bins: $0.5 < \eta < 1.0$ and $-0.5 < \eta < 0$. The p_T of the truth jet is labeled as p_T True.

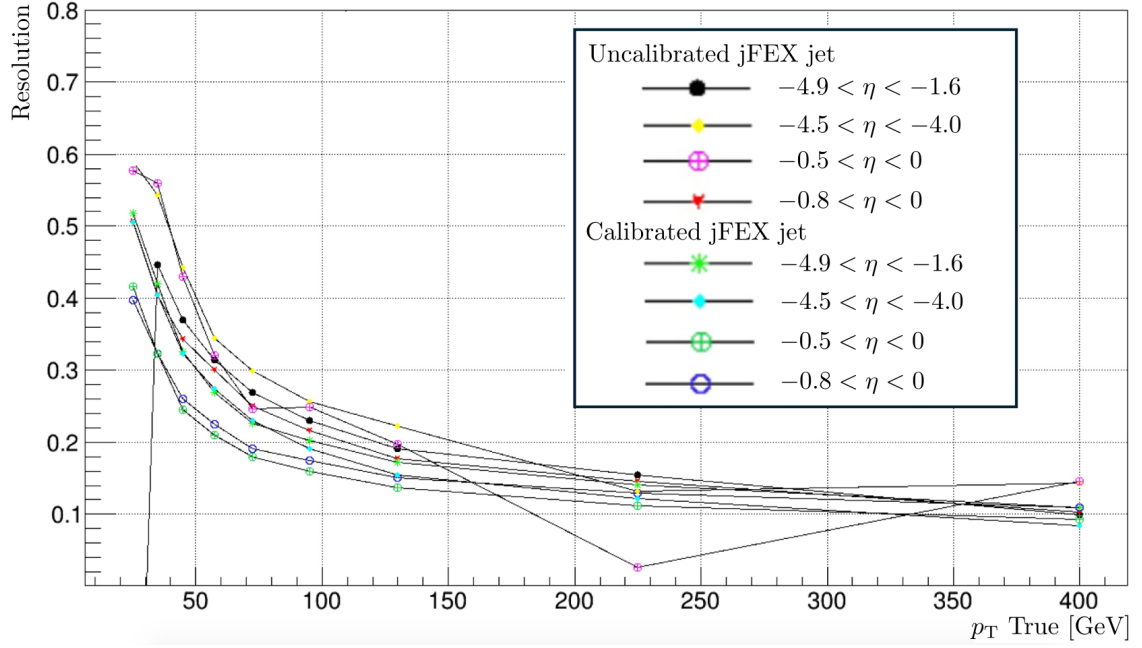


Figure 5.12 Comparison of the performance of the jFEX calibration using fine η binning to the performance of the jFEX calibration using coarser η binning by showing resolution of as a function of p_T of jFEX truth jets. Results shown for coarse η bins $-4.9 < \eta < -1.6$ and $-0.8 < \eta < 0$ and the following fine η bins: $-4.5 < \eta < -4.0$ and $-0.5 < \eta < 0$. The p_T of the truth jet is labeled as p_T True.

5.3 Summary

A series of calibrations were derived to make calibrated jFEX jets have a flat response in η . Out of all of the calibrations derived, the one provided to hardware experts was a calibration of the jFEX jet p_T response in bins of jFEX jet p_T and η , where the η binning was defined as $-4.9 < \eta < -1.6$, $-1.6 < \eta < -0.8$, $-0.8 < \eta < 0$, $0 < \eta < 0.8$, $0.8 < \eta < 1.6$, and $1.6 < \eta < 4.9$. The calibration corrected the jFEX jet p_T response to the desired value of 1 in almost every $p_T - \eta$ bin with the exception of the bins with the lowest p_T values, which deviated from 1 by no more than 15%, and the highest p_T values, which deviated from 1 by no more than roughly 5%. The resolution saw significant improvement, decreasing across each p_T bin for each η bin. Depending on the $p_T - \eta$ bin, the p_T response for calibrated jFEX jets is roughly 1.5 to 2.5 times that of the p_T response for uncalibrated jFEX jets. The resolution of calibrated jFEX jets is as small as 0.71 times the resolution of uncalibrated jFEX jets in higher p_T bins or can be up to 0.95 times the resolution of uncalibrated jFEX jets in lower p_T bins. The objective was achieved and the calibration improved both

the response of jFEX jets across η and the resolution of the jFEX jets.

Chapter 6

Unexcluded Phase Space, Kinematics, and Search Strategy for Non-minimal Dark Sectors, Scalar Cascades, and Noncommutative Black Holes

No experiment has ever searched for noncommutative black holes, scalar cascades, or non-minimal dark sectors. However, model-independent constraints from other ATLAS searches are taken into account in order to design searches that target unexplored phase space. This chapter determines the unexcluded and potentially reachable phase space for each model. The first section categorizes hundreds of potential noncommutative black hole benchmarks into benchmarks that are excluded by the model-independent limits on the Planck scale in ADD-type extra dimensions, benchmarks that are unexcluded and yield at least 5 events at 139 fb^{-1} at 13 TeV, benchmarks that yield under 5 events and are, therefore, unviable at 139 fb^{-1} at 13 TeV, and benchmarks that require greater than the LHC collisional energy of 13 TeV to be produced. The second section outlines 6 benchmarks of interest for non-minimal dark sectors. The non-minimal dark sector model has 6 free parameters. This means that the possible combinations of parameters is vast. Additionally, a single non-minimal dark sector benchmark can require the generation of hundreds or even thousands of subprocesses. It is, therefore, not possible on the timescale

of a PhD to determine the full unexcluded phase space for this model. Therefore, 6 benchmarks yielding potentially reachable phenomenology and that have evaded exclusion limits are presented and discussed. The third part of this chapter describes the various classes of scalar cascades and presents collectively over a hundred different benchmarks that have not been excluded due to small cross sections and small trigger efficiencies. For each model, a set of benchmarks is selected for each analysis.

6.1 Noncommutative Black Hole Parameters and MC Samples

This section details the determination of the unexcluded phase space of noncommutative black holes. In order to determine the unexcluded phase space, the author designed hundreds of potential benchmarks. This entailed a series of calculations by the author. As a reminder, the key parameters in the noncommutative black hole model are the Planck scale in higher dimensions, M_D , the noncommutative parameter $\sqrt{\theta}$, and the minimum black hole mass required for the black hole to be produced, $m_{\min}$.

In this chapter, the noncommutative parameter $\sqrt{\theta}$ will be written in units of M_D as $\sqrt{\theta}M_D$. The value $\sqrt{\theta}M_D$ will be referred to as the noncommutative scale. The first step in designing a potential noncommutative black hole benchmark is to consider the range of $\sqrt{\theta}M_D$ that can be probed at the LHC. Table 6.1 shows the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC for a given n , where n is the number of extra dimensions.

n	$(r_S)_{\min}/\sqrt{\theta}$	$(m_{\min}/M_D)(\sqrt{\theta}M_D)^{-(n+1)}$	$\sqrt{\theta}_{\min}M_D$	$\sqrt{\theta}_{\max}M_D$
2	2.51	65.2	0.248	0.377
3	2.41	58.8	0.361	0.667
4	2.34	48.6	0.460	0.789
5	2.29	37.9	0.546	0.869
6	2.26	28.2	0.621	0.929
7	2.23	20.3	0.686	0.982

Table 6.1 *Range of $\sqrt{\theta}M_D$ in units of M_D that can be probed at the LHC. The $(r_S)_{\min}$ is the minimum Schwarzschild radius needed for black hole formation. Table modified from Reference [91]*

The maximum allowed value of $\sqrt{\theta}M_D$ is denoted as $\sqrt{\theta}_{max}M_D$. The minimum allowed value of $\sqrt{\theta}M_D$ is denoted as $\sqrt{\theta}_{min}M_D$. The next step required is to pick a value for n . For example, for 2 extra dimensions, n is 2. Then the value of $\sqrt{\theta}M_D$ is chosen between $\sqrt{\theta}_{min}M_D$ and $\sqrt{\theta}_{max}M_D$ for the chosen n . For 2 extra dimensions, for example, $\sqrt{\theta}M_D$ can be any value from 0.248 to 0.377. The third column in Table 6.1 shows the value of the relation $(\frac{m_{min}}{M_D})(\sqrt{\theta}M_D)^{-(n+1)}$ for a given n . For 2 extra dimensions, this value is 65.2. By setting $(\frac{m_{min}}{M_D})(\sqrt{\theta}M_D)^{-(n+1)}$ equal to the value in the third column, the ratio m_{min} to M_D can be calculated by plugging in the chosen number of extra dimensions, n , and the chosen value of $\sqrt{\theta}M_D$. The only remaining step is to choose a value of M_D and use the value calculated for $\frac{m_{min}}{M_D}$ to determine the value of m_{min} .

After performing these calculations, the author inputted the m_{min} , M_D , and n into the Charybdis2 generator to simulate the production and decay of noncommutative black holes and produce an output file called a Les Houches Event (LHE) file. These files contain event information which includes the total predicted cross section of the black hole. The procedure detailed above was systematically used to create LHE files and obtain cross sections for noncommutative black holes for various combinations of $\sqrt{\theta}M_D$, m_{min} , and n . All of the remaining tables in this section were generated by the author using this procedure.

In order to determine unexcluded and reachable phase space, a number of additional considerations were taken into account. The black hole minimum mass threshold m_{min} cannot exceed the total energy of the LHC. Any benchmark that has a m_{min} that corresponds to a M_D of greater than 13 TeV for a given number of extra dimensions and a given noncommutative scale, $\sqrt{\theta}M_D$, cannot be formed at 13 TeV. As the value of the noncommutative scale increases, $\frac{m_{min}}{M_D}$ increases and the values of M_D at which the black hole cannot form decreases, restricting the values of M_D . The exclusion limits on the Planck scale in higher dimensions are considered for all numbers of dimensions for which the LHC has set explicit limits [7]. The limit from the full Run 2 (at 139 fb^{-1}) ATLAS mono-jet analysis is the following for 2, 3, 4, 5, and 6 extra dimensions, respectively: 11.2, 8.5, 7.1, 6.4, and 5.9 TeV [7]. This can be seen in Table 6.2. The ATLAS mono-jet analysis was performed by requiring no more than 4 small R jets above 30 GeV, requiring at least 200 GeV in E_T^{miss} , and requiring the closest small R jet to the E_T^{miss} to be greater than a value of 0.4, among other selections.

ADD Model Limits on M_D (95% CL)		
	Expected [TeV]	Observed [TeV]
$n = 2$	$11.6^{+1.0}_{-1.0}$	$11.2^{+1.0}_{-1.3}$
$n = 3$	$8.6^{+0.6}_{-0.6}$	$8.5^{+0.6}_{-0.8}$
$n = 4$	$7.2^{+0.4}_{-0.4}$	$7.1^{+0.4}_{-0.6}$
$n = 5$	$6.4^{+0.3}_{-0.3}$	$6.4^{+0.3}_{-0.5}$
$n = 6$	$5.9^{+0.2}_{-0.2}$	$5.9^{+0.3}_{-0.4}$

Table 6.2 *Limits on M_D from ATLAS search for ADD-type extra dimensions via monojet final state with full Run 2 data at 139 fb^{-1} . Figure taken from Reference [7]*

Additionally, the model independent limits on cross sections set by the latest (full Run 2 at 139 fb^{-1}) multijet SUSY (supersymmetry) analysis [9] and the monojet search [7] are considered. In [9], it can be seen that, for the selection of 8 jets above 50 GeV, no b-jets, and a requirement of a sum of mass greater than 500 GeV (for all large R jets of $R=1.0$ and within $-1.5 < \eta < 1.5$, the model-independent exclusion limit on the cross section $\times$ efficiency is 1.16 fb. Similarly from [9], for the selection of 9 jets above 50 GeV, no b-jets (jets that arise from b quarks), and a requirement of a sum of mass greater than 340 GeV, the model-independent exclusion limit on the cross section $\times$ efficiency is 0.95 fb. Similarly from [9], for a selection of 10 jets above 50 GeV, 1 b-jet, and a requirement of a sum of mass greater than 500 GeV, the model-independent exclusion limit on the cross section $\times$ efficiency is 0.12 fb. These three selections are sufficient to use to determine unexcluded phase space for scalar cascades, noncommutative black holes, and non-minimal dark sectors.

Using all of the information presented so far and using the procedure described at the beginning of the chapter, the unexcluded and viable phase space for noncommutative black holes is systematically charted out below for 2 to 7 extra dimensions. As seen in Table 6.2, the observed exclusion limit from the mono-jet analysis on M_D in 2 extra dimensions is 11.2 TeV. Table 6.3 shows the $m_{\min}$ and cross sections for given values of M_D for (from left to right) $\sqrt{\theta}M_D$ values of 0.25 and 0.3. As a reminder, as seen in Table 6.1, the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC in 2 extra dimensions is $0.248 < \sqrt{\theta}M_D < 0.377$. The following studies will demonstrate that cross sections become smaller for a given value of M_D as $\sqrt{\theta}M_D$ values increase. As benchmarks with the lowest possible $\sqrt{\theta}M_D$ values and, therefore, highest cross sections, provide the best opportunity for sensitivity, it is advantageous to assess the viable phase space for a given number of extra dimensions

starting with the lowest possible $\sqrt{\theta}M_D$ values. Therefore, the assessment of the viability of phase space in 2 extra dimensions begins with a $\sqrt{\theta}M_D$ value of 0.25. The $\sqrt{\theta}M_D$ value is then increased to 0.3. However, as will be seen, it is not always necessary to study $\sqrt{\theta}M_D$ ranging to the maximum allowed $\sqrt{\theta}M_D$ value (which, in the instance of 2 extra dimensions would be 0.377) because there may be no viable phase space left at lower values of $\sqrt{\theta}M_D$. In that case, there would be no viable phase space left at higher $\sqrt{\theta}M_D$ values because higher $\sqrt{\theta}M_D$ values have lower cross sections. The red color-coded benchmarks in Table 6.3 indicate benchmarks that have been excluded by limits on the Planck scale, M_D . The blue color-coded benchmarks indicate benchmarks that have cross sections that yield less than 5 events at 139 fb^{-1} at 13 TeV and that are, therefore, unviable to search for in Run 2. The cross section that corresponds to this threshold is $3.6 \times 10^{-5} \text{ pb}$. The purple color-coded benchmarks indicate benchmarks with a black hole minimum mass m_{min} greater than 13 TeV, which means that the energy required to produce the black hole is out of reach (OOR) of the 13 TeV collider for Run 2. These benchmarks are labeled “OOR”. For $\sqrt{\theta}M_D = 0.25$, benchmarks with M_D values below 11.2 TeV are excluded, while benchmarks listed with M_D values from 11.5 TeV to 12.5 TeV have cross sections that yield too few events, and the benchmark with $M_D = 13 \text{ TeV}$ is out of reach at the 13 TeV collider for Run 2. For $\sqrt{\theta}M_D = 0.3$, black holes are already out of reach by 7.5 TeV. Since cross sections become smaller as the value of $\sqrt{\theta}M_D$ increases, there is no need to look at $\sqrt{\theta}M_D$ values beyond 0.3 to know that no unexcluded and viable phase space remains for noncommutative black hole production in 2 dimensions at 13 TeV.

$\sqrt{\theta}M_D$ 0.25				$\sqrt{\theta}M_D$ 0.3		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
2.0	2.038	2.80E+02		1.0	1.760	3.20E+03
2.5	2.547	6.80E+01		1.5	2.641	2.60E+02
3.0	3.056	1.90E+01		2.0	3.521	3.20E+01
3.5	3.566	5.80E+00		2.5	4.401	4.60E+00
4.0	4.075	1.90E+00		3.0	5.281	7.00E-01
4.5	4.584	6.10E-01		3.5	6.161	1.00E-01
5.0	5.094	2.00E-01		4.0	7.042	1.30E-02
5.5	5.603	6.70E-02		4.5	7.922	1.40E-03
6.0	6.113	2.20E-02		5.0	8.802	1.20E-04
6.5	6.724	5.90E-03		6.0	10.562	1.60E-07
7.0	7.131	2.00E-03		6.5	11.443	1.80E-09
7.5	7.641	5.60E-04		7.0	12.323	7.20E-12
8.0	8.150	1.40E-04		7.5	13.203	OOR
8.5	8.659	3.30E-05		8.0	14.083	OOR
9.0	9.169	6.60E-06				
9.5	9.678	1.10E-06				
10.0	10.188	1.50E-07				
10.5	10.697	1.60E-08				
11.0	11.206	1.20E-09				
11.5	11.716	7.50E-11				
12.0	12.225	3.00E-12				
12.5	12.734	5.50E-14				
13.0	13.244	OOR				

Table 6.3 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 2$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Blue: benchmarks that yield under 5 events at 139 fb^{-1} at 13 TeV. Purple: out of reach (OOR) benchmarks.*

The observed exclusion limit from the mono-jet analysis on M_D in 3 extra dimensions is 8.5 TeV. As seen in Table 6.1, the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC in 3 extra dimensions is $0.361 < \sqrt{\theta}M_D < 0.667$. Table 6.4 shows the $m_{\min}$ and cross sections for given values of M_D for $\sqrt{\theta}M_D$ values of 0.37, 0.4, 0.5, and 0.6 in 3 extra dimensions. All benchmarks for $\sqrt{\theta}M_D = 0.37$ are either excluded, yield too few events, or are out of reach of the LHC for 13 TeV. All benchmarks for $\sqrt{\theta}M_D = 0.4$ are either excluded or out of reach of the LHC for 13 TeV. All benchmarks for $\sqrt{\theta}M_D$ values of 0.5 and 0.6 are excluded. As the noncommutative scale increases, the cross sections get smaller. No viable phase space remains to search for noncommutative black holes in 3 extra dimensions at the LHC with 13 TeV.

$\sqrt{\theta}M_D$ 0.37			$\sqrt{\theta}M_D$ 0.4		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)	M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
1.0	1.102	2.00E+04	1.0	1.505	9.10E+03
1.5	1.653	2.30E+03	1.5	2.258	8.60E+02
2.0	2.204	4.30E+03	2.0	3.011	1.20E+02
2.5	2.755	9.90E+01	2.5	3.763	2.20E+02
3.0	3.306	2.60E+01	3.0	4.516	4.30E+00
3.5	3.857	7.60E+00	3.5	5.268	8.70E-01
4.0	4.408	2.30E+00	4.0	6.021	1.70E-01
4.5	4.959	7.00E-01	4.5	6.773	3.00E-02
5.0	5.510	2.10E-01	5.0	7.526	4.90E-03
5.5	6.061	6.40E-02	5.5	8.279	6.70E-04
6.0	6.612	1.80E-02	6.0	9.032	7.10E-05
6.5	7.163	4.90E-03	6.5	9.784	5.20E-06
7.0	7.714	1.20E-03	7.0	10.537	2.30E-07
7.5	8.265	2.80E-04	7.5	11.290	5.4E-09
8.0	8.816	5.60E-05	8.0	12.042	5.9E-10
8.5	9.367	9.30E-06	8.5	12.795	1.1E-13
9.0	9.918	1.20E-06	9.0	13.547	OOOR
9.5	10.469	1.20E-07	9.5	14.300	OOOR
10.0	11.020	8.90E-09			
10.5	11.571	4.60E-10			
11.0	12.122	1.50E-11			
11.5	12.673	3.00E-13			
12.0	13.224	OOOR			
12.5	13.775	OOOR			
13.0	14.324	OOOR			
$\sqrt{\theta}M_D$ 0.5			$\sqrt{\theta}M_D$ 0.6		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)	M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
1.0	3.625	3.30E+02	1.0	7.582	4.50E-02
1.5	5.513	6.50E+00	1.5	11.372	2.90E-07
2.0	7.350	1.00E-01	2.0	15.163	OOOR
2.5	9.188	6.10E-04	2.5	18.954	OOOR
3.0	11.025	3.10E-07	3.0	22.751	OOOR
3.5	12.863	1.80E-13			
4.0	14.700	OOOR			
4.5	16.538	OOOR			

Table 6.4 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 3$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Blue: benchmarks that yield under 5 events at 139 fb^{-1} at 13 TeV. Purple: out of reach (OOOR) benchmarks.*

The observed exclusion limit from the mono-jet analysis on M_D in 4 extra dimensions is 7.1 TeV. As seen in Table 6.1, the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC in 4 extra dimensions is $0.460 < \sqrt{\theta}M_D < 0.789$. Table 6.5 shows the $m_{\min}$ and cross sections for given values of M_D for $\sqrt{\theta}M_D$ values of 0.37, 0.4, and 0.5 in 4 extra dimensions. Noticeably, the $\sqrt{\theta}M_D$ values of 0.37 and 0.4 are below the allowed values of $\sqrt{\theta}M_D$. However, they exist in Table 6.5 simply as comparison to $\sqrt{\theta}M_D = 0.5$, as $\sqrt{\theta}M_D = 0.5$ is immediately seen to be excluded. No viable phase space remains to search for noncommutative black holes in 4 extra dimensions at the LHC with 13 TeV.

$\sqrt{\theta}M_D$ 0.37			$\sqrt{\theta}M_D$ 0.4			$\sqrt{\theta}M_D$ 0.5		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)	M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)	M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
1.0	1.519	1.30E+04	1.0	3.779	3.80E-03	1.0	8.168	0.17
1.5	2.278	1.20E+03	1.5	5.669	7.00E+00	1.5	12.252	1.6E-09
2.0	3.038	1.80E+02	2.0	7.558	9.30E-02	2.0	16.336	0.00
2.5	3.797	3.20E+01	2.5	9.448	3.90E-04			
3.0	4.556	6.20E+00	3.0	11.337	8.50E-08			
3.5	5.316	1.20E+00	3.5	13.227	0.00			
4.0	6.075	2.30E-01	4.0	15.117	0.00			
4.5	6.834	4.20E-02	4.5	17.006	0.00			
5.0	7.594	6.60E-03	5.0	18.896	0.00			
5.5	8.353	5.30E-03						
6.0	9.113	8.70E-05						
6.5	9.872	6.00E-06						
7.0	10.631	2.40E-07						
7.5	11.391	4.90E-09						
8.0	12.150	4.30E-11						
8.5	12.909	9.50E-15						
9.0	13.669	0.00						

Table 6.5 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 4$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Blue: benchmarks that yield under 5 events at 139 fb^{-1} at 13 TeV. Purple: out of reach (OOR) benchmarks.*

The case of 5 extra dimensions is the first instance in which noncommutative black holes can be viably searched for at the LHC with 13 TeV. The observed exclusion limit from the mono-jet analysis on M_D in 5 extra dimensions is 6.4 TeV. A new color is added to the color-coding scheme. Green indicates benchmarks that are unexcluded and yield at least 5 events at 13 TeV at a luminosity of 139 fb^{-1} . As seen in Table 6.1, the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC in 5 extra dimensions is $0.546 < \sqrt{\theta}M_D < 0.869$. Table 6.6 shows the $m_{\min}$ and cross sections for given values of M_D for $\sqrt{\theta}M_D$ values of 0.546, 0.55, and 0.56 in 5 extra dimensions. The unexcluded phase space of M_D that allows for at least 5 events to be generated at 13 TeV with a luminosity of 139 fb^{-1} in 5 extra dimensions is the following: $6.4 \text{ TeV} < M_D < 9.0 \text{ TeV}$ for $\sqrt{\theta}M_D = 0.546$, $6.4 \text{ TeV} < M_D < 8.5 \text{ TeV}$ for $\sqrt{\theta}M_D = 0.55$, and $6.4 \text{ TeV} < M_D < 8.5 \text{ TeV}$ for $\sqrt{\theta}M_D = 0.56$.

$\sqrt{\theta}M_D$ 0.546				$\sqrt{\theta}M_D$ 0.55				$\sqrt{\theta}M_D$ 0.56		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
4.0	4.017	9.20E+00		5.0	5.255	7.30E-01		4.5	5.260	9.40E-01
4.5	4.519	0.31E+01		5.5	5.770	2.40E-01		5.0	5.844	0.27E+00
5.0	5.021	0.11E+01		6.0	6.293	7.40E-02		5.5	6.429	7.30E-02
5.5	5.523	3.60E-01		6.5	6.819	2.20E-02		6.0	7.013	1.90E-02
6.0	6.025	0.12E+00		7.0	7.344	6.30E-03		6.5	7.598	4.40E-03
6.5	6.527	3.90E-02		7.5	7.868	1.60E-03		7.0	8.182	9.50E-04
7.0	7.029	1.20E-02		8.0	8.392	3.90E-04		7.5	8.767	1.80E-04
7.5	7.531	3.50E-03		8.5	8.917	8.2E-05		8.0	9.351	2.70E-05
8.0	8.033	9.40E-04		9.0	9.442	1.50E-05		8.5	9.935	0.32E-05
8.5	8.535	2.30E-04		9.5	9.966	2.1E-06		9.0	10.520	0.28E-06
9.0	9.037	0.50E-04		10.0	10.49	2.30E-07		9.5	11.104	0.16E-07
9.5	9.539	9.20E-06		10.5	11.016	1.90E-08		10.0	11.689	0.62E-09
10.0	10.041	1.40E-06		11.0	11.54	1.10E-09		10.5	12.273	0.14E-10
10.5	10.543	1.60E-07		11.5	12.065	4.30E-11		11.0	12.858	2.90E-14
11.0	11.046	1.40E-08		12.0	12.589	1.3E-12		11.5	13.442	OOO
11.5	11.548	9.50E-10		12.5	13.114	4.50E-13		12.0	14.026	OOO
12.0	12.050	4.30E-11		13.0	13.638	OOO		12.5	14.611	OOO
12.5	12.552	1.60E-12						13.0	15.195	OOO
13.0	13.054	OOO								

Table 6.6 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 5$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOO) benchmarks.*

Additional values of $\sqrt{\theta}M_D$, seen in Table 6.7, can be considered in order to assess the allowed phase space for noncommutative black holes in 5 dimensions at 13 TeV. However, little remain viable. In 5 extra dimensions, as indicated by the green color-coding scheme, the range of M_D that allows for at least 5 events to be generated at 13 TeV with a luminosity of 139 fb^{-1} is $6.4 < M_D < 7.5$ for $\sqrt{\theta}M_D = 0.57$. For $\sqrt{\theta}M_D = 0.58$, only $M_D = 6.5$ TeV is viable. All values of M_D for $\sqrt{\theta}M_D = 0.59$ and $\sqrt{\theta}M_D = 0.6$ are either excluded, yield too few events, or are out of reach of the LHC for 13 TeV.

$\sqrt{\theta}M_D$ 0.57				$\sqrt{\theta}M_D$ 0.58		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
5.5	7.149	1.80E-02		6.0	8.657	4.40E-04
6.0	7.799	3.50E-03		6.5	9.378	0.44E-04
6.5	8.449	6.00E-04		7.0	10.100	3.00E-06
7.0	9.099	0.83E-04		7.5	10.821	0.12E-06
7.5	9.749	0.88E-05		8.0	11.542	2.60E-09
8.0	10.398	6.40E-07		8.5	12.264	0.24E-10
8.5	11.048	2.90E-08		9.0	12.985	4.50E-18
9.0	11.698	7.80E-10		9.5	13.707	OOOR
9.5	12.348	1.10E-11		10.0	14.428	OOOR
10.0	12.998	1.20E-11		10.5	15.149	OOOR
10.5	13.648	OOOR		11.0	15.871	OOOR
11.0	14.298	OOOR		11.5	16.592	OOOR
11.5	14.948	OOOR		12.0	17.314	OOOR
12.0	15.598	OOOR		12.5	18.035	OOOR
12.5	16.248	OOOR		13.0	18.756	OOOR
13.0	16.897	OOOR				
$\sqrt{\theta}M_D$ 0.59				$\sqrt{\theta}M_D$ 0.6		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
5.5	8.792	3.80E-04		5.0	8.840	0.44E-03
6.0	9.592	2.80E-05		5.5	9.724	2.30E-05
6.5	10.391	1.20E-06		6.0	10.608	0.59E-06
7.0	11.190	2.40E-08		6.5	11.492	5.90E-09
7.5	11.990	2.10E-10		7.0	12.376	1.90E-11
8.0	12.789	3.00E-13		7.5	13.260	OOOR
8.5	13.588	OOOR		8.0	14.144	OOOR
9.0	14.387	OOOR		8.5	15.028	OOOR
9.5	15.187	OOOR		9.0	15.912	OOOR
10.0	15.986	OOOR		9.5	16.796	OOOR
10.5	16.785	OOOR		10.0	17.680	OOOR
11.0	17.585	OOOR		10.5	18.564	OOOR
11.5	18.384	OOOR		11.0	19.448	OOOR
12.0	19.183	OOOR		11.5	20.332	OOOR
12.5	19.983	OOOR		12.0	21.216	OOOR
13.0	20.782	OOOR		12.5	22.100	OOOR
				13.0	22.984	OOOR

Table 6.7 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 5$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOOR) benchmarks.*

Viable phase space for noncommutative black holes in 6 extra dimensions exist for some values of $\sqrt{\theta}M_D$. As seen in Table 6.1, the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC in 6 extra dimensions is $0.621 < \sqrt{\theta}M_D < 0.929$. Table 6.8 shows the $m_{\min}$ and cross sections for given values of M_D for $\sqrt{\theta}M_D$ values of 0.621, 0.622, 0.63, 0.65, 0.67, 0.70, 0.72, 0.74, and 0.78 in 6 extra dimensions. However, $\sqrt{\theta}M_D$ values of 0.67 and higher can immediately be ruled out, as seen in Table 6.8, due to either already being excluded, yielding too few events, or the minimum black hole mass required for black hole creation being out of reach of the LHC. The observed exclusion limit from the mono-jet analysis on M_D in 6 extra dimensions is 5.9 TeV. The viable M_D range for noncommutative black holes in 6 extra dimensions

and with a $\sqrt{\theta}M_D$ value of 0.621 is $5.9 \text{ TeV} < M_D < 9 \text{ TeV}$, while the viable range for $\sqrt{\theta}M_D = 0.622$ is $5.9 \text{ TeV} < M_D < 9.0 \text{ TeV}$. For $\sqrt{\theta}M_D = 0.63$, the range of viable M_D is $5.9 \text{ TeV} < M_D < 8.5 \text{ TeV}$. For $\sqrt{\theta}M_D = 0.65$, the range of viable M_D becomes restricted to $5.9 \text{ TeV} < M_D < 6.5 \text{ TeV}$.

$\sqrt{\theta}M_D$ 0.621				$\sqrt{\theta}M_D$ 0.622				$\sqrt{\theta}M_D$ 0.63		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
4.5	4.52	0.39E+01		4.5	4.57	0.36E+01		4.5	4.99	1.80E+00
5	5.02	1.30E+00		5	5.08	1.20E+00		5	5.54	5.60E-01
5.5	5.52	0.46E+00		5.5	5.59	4.10E-01		5.5	6.09	1.60E-01
6	6.03	1.50E-01		6	6.09	1.40E-01		6	6.65	0.47E-01
6.5	6.53	4.90E-02		6.5	6.60	0.43E-01		6.5	7.20	1.30E-02
7	7.03	1.50E-02		7	7.11	1.30E-02		7	7.76	3.10E-03
7.5	7.53	4.40E-03		7.5	7.62	3.70E-03		7.5	8.31	6.90E-04
8	8.03	1.20E-03		8	8.13	0.96E-03		8	8.86	1.30E-04
8.5	8.54	0.29E-03		8.5	8.63	2.30E-04		8.5	9.42	2.20E-05
9	9.04	6.30E-05		9	9.14	0.47E-04		9	9.97	2.80E-06
9.5	9.54	1.20E-05		9.5	9.65	0.82E-05		9.5	10.53	2.60E-07
10	10.04	0.18E-05		10	10.16	0.11E-05		10	11.08	0.17E-07
10.5	10.55	0.20E-06		10.5	10.67	1.20E-07		10.5	11.63	7.80E-10
11	11.05	0.18E-07		11	11.17	9.50E-09		11	12.19	2.10E-11
11.5	11.55	0.12E-08		11.5	11.68	5.60E-10		11.5	12.74	1.80E-13
12	12.05	5.10E-11		12	12.19	2.10E-11		12	13.29	OOOR
12.5	12.55	0.18E-11		12.5	12.70	4.80E-13		12.5	13.85	OOOR
13	13.06	OOOR		13	13.20	OOOR		13	14.40	OOOR
$\sqrt{\theta}M_D$ 0.65				$\sqrt{\theta}M_D$ 0.67				$\sqrt{\theta}M_D$ 0.70		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
4	5.53	1.00E+00		4	6.836	9.80E-02		5.5	12.77	OOOR
4.5	6.22	2.30E-01		4.5	7.691	1.20E-02		6	13.93	OOOR
5	6.91	4.70E-02		5	8.5455	1.20E-03		6.5	15.10	OOOR
5.5	7.60	0.87E-02		5.5	9.400	8.10E-05		7	16.26	OOOR
6	8.29	1.40E-03		6	10.255	3.20E-06		7.5	17.42	OOOR
6.5	8.99	0.18E-03		6.5	11.109	0.54E-07		8	18.58	OOOR
7	9.68	1.70E-05		7	11.964	3.50E-10		8.5	19.74	OOOR
7.5	10.37	0.11E-05		7.5	12.818	0.20E-12		9	20.90	OOOR
8	11.06	4.00E-08		8	13.673	OOOR		9.5	22.06	OOOR
8.5	11.75	0.80E-09		8.5	14.527	OOOR		10	23.22	OOOR
9	12.44	8.00E-12		9	15.382	OOOR		10.5	24.39	OOOR
9.5	13.13	OOOR		9.5	16.236	OOOR		11	25.55	OOOR
10	13.82	OOOR		10	17.091	OOOR		11.5	26.71	OOOR
10.5	14.52	OOOR		10.5	17.946	OOOR		12	27.87	OOOR
11	15.21	OOOR		11	18.800	OOOR		12.5	29.03	OOOR
11.5	15.90	OOOR		11.5	19.655	OOOR		13	30.19	OOOR
$\sqrt{\theta}M_D$ 0.72				$\sqrt{\theta}M_D$ 0.74				$\sqrt{\theta}M_D$ 0.78		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
5	14.14	OOOR		4	13.71	OOOR		3	14.86	OOOR
5.5	15.56	OOOR		4.5	15.42	OOOR		3.5	17.34	OOOR
6	16.97	OOOR		5	17.13	OOOR		4	19.81	OOOR
6.5	18.39	OOOR		5.5	18.85	OOOR		4.5	22.29	OOOR
7	19.80	OOOR		6	20.56	OOOR		5	24.77	OOOR
7.5	21.21	OOOR		6.5	22.27	OOOR		5.5	27.24	OOOR
8	22.63	OOOR		7	23.99	OOOR		6	29.72	OOOR
8.5	24.04	OOOR		7.5	25.70	OOOR		6.5	32.20	OOOR
9	25.46	OOOR		8	27.41	OOOR		7	34.67	OOOR
9.5	26.87	OOOR		8.5	29.13	OOOR		7.5	37.15	OOOR
10	28.29	OOOR		9	30.84	OOOR		8	39.63	OOOR
10.5	29.70	OOOR		9.5	32.55	OOOR		8.5	42.10	OOOR
11	31.11	OOOR		10	34.27	OOOR		9	44.58	OOOR
11.5	32.53	OOOR		10.5	35.98	OOOR		9.5	47.06	OOOR

Table 6.8 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 6$ Extra Dimensions. The color-coding scheme is defined as follows. Red: benchmarks excluded by the limits on M_D . Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOOR) benchmarks.*

The largest available phase space is found in 7 extra dimensions. Unlike 2 to 6 extra dimensions, no explicit model-independent limits have been set by ATLAS on M_D

in 7 extra dimensions. However, in 2008, the D0 detector at the Fermilab Tevatron Collider reported an exclusion limit of 0.8 TeV on M_D in ADD-type extra dimensions for 7 extra dimensions [8]. Though ATLAS has not explicitly set a limit on M_D in 7 extra dimensions, the limits set on the mass of the string scale, M_s , by a 2012 ATLAS search for dileptons can be translated to limits on the M_D [92].

As seen in Table 6.1, the allowed range of $\sqrt{\theta}M_D$ that can be probed at the LHC in 7 extra dimensions is $0.686 < \sqrt{\theta}M_D < 0.982$. Table 6.10 shows the $m_{\min}$ and cross sections for given values of M_D for $\sqrt{\theta}M_D$ values of 0.686, 0.687, 0.7, 0.71, 0.72, and 0.73 in 7 extra dimensions. The lowest values of M_D are not included in the table because, though no explicit model-independent limit on M_D exists above 0.8 TeV in 7 extra dimensions, the cross sections for those M_D values are large enough that they would have likely already been seen in LHC data. The viable ranges of M_D for $\sqrt{\theta}M_D = 0.686$ and $\sqrt{\theta}M_D = 0.687$ are $3.5 < M_D < 9.0$ TeV, while the viable ranges of M_D for $\sqrt{\theta}M_D = 0.7$ and $\sqrt{\theta}M_D = 0.71$ are $3.5 < M_D < 8.0$ TeV and $3.0 < M_D < 7.0$ TeV, respectively. The viable ranges of M_D for $\sqrt{\theta}M_D = 0.72$ and $\sqrt{\theta}M_D = 0.73$ are $3.0 < M_D < 6.5$ TeV and $3.0 < M_D < 6.0$ TeV. The viable ranges of M_D for $\sqrt{\theta}M_D = 0.74$ and $\sqrt{\theta}M_D = 0.75$ are $2.5 < M_D < 5.5$ TeV and $2.5 < M_D < 4.5$ TeV, respectively.

$\sqrt{\theta}M_L$ 0.686				$\sqrt{\theta}M$ 0.687				$\sqrt{\theta}M_D$ 0.7		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
3.5	3.485	4.40E+01		3.5	3.525	4.20E+01		3.5	4.096	1.70E+01
4.0	3.982	1.50E+01		4.0	4.029	1.40E+01		4.0	4.681	4.90E+00
4.5	4.480	5.00E+00		4.5	4.533	4.60E+00		4.5	5.266	1.40E+00
5.0	4.978	0.17E+01		5.0	5.036	0.16E+01		5.0	5.851	0.40E+00
5.5	5.476	0.60E+00		5.5	5.540	5.40E-01		5.5	6.436	1.10E-01
6.0	5.974	2.00E-01		6.0	6.044	0.18E+00		6.0	7.022	2.80E-02
6.5	6.471	0.66E-01		6.5	6.547	5.70E-02		6.5	7.607	6.60E-03
7.0	6.969	2.10E-02		7.0	7.051	1.80E-02		7.0	8.192	1.40E-03
7.5	7.467	6.10E-03		7.5	7.555	5.10E-03		7.5	8.777	2.60E-04
8.0	7.965	1.70E-03		8.0	8.058	1.40E-03		8.0	9.362	0.39E-04
8.5	8.463	0.43E-03		8.5	8.562	0.33E-03		8.5	9.947	0.46E-05
9.0	8.960	9.50E-05		9.0	9.066	7.00E-05		9.0	10.532	3.80E-07
9.5	9.458	0.18E-04		9.5	9.569	1.30E-05		9.5	11.117	0.21E-07
10.0	9.956	0.29E-05		10.0	10.073	1.90E-06		10.0	11.703	0.79E-09
10.5	10.454	3.50E-07		10.5	10.576	2.10E-07		10.5	12.288	1.60E-11
11.0	10.952	3.30E-08		11.0	11.080	1.70E-08		11.0	12.873	1.30E-14
11.5	11.449	2.30E-09		11.5	11.584	1.10E-09		11.5	13.458	OOO
12.0	11.947	1.10E-10		12.0	12.087	4.50E-11		12.0	14.043	OOO
12.5	12.445	4.10E-12		12.5	12.591	1.40E-12		12.5	14.628	OOO
13.0	12.943	1.10E-16		13.0	13.095	OOO		13.0	15.213	OOO
$\sqrt{\theta}M_L$ 0.71				$\sqrt{\theta}M$ 0.72				$\sqrt{\theta}M_D$ 0.73		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
3.0	3.933	3.30E+01		3.0	4.398	0.16E+02		3.0	4.911	7.10E+00
3.5	4.588	0.80E+01		3.5	5.131	0.33E+01		3.5	5.730	0.12E+01
4.0	5.244	0.20E+01		4.0	5.864	6.90E-01		4.0	6.548	2.00E-01
4.5	5.899	4.80E-01		4.5	6.597	1.30E-01		4.5	7.367	0.29E-01
5.0	6.554	0.11E+00		5.0	7.330	0.24E-01		5.0	8.185	0.34E-02
5.5	7.210	0.24E-01		5.5	8.063	0.36E-02		5.5	9.004	3.10E-04
6.0	7.865	4.50E-03		6.0	8.796	0.44E-03		6.0	9.822	0.18E-04
6.5	8.521	7.40E-04		6.5	9.529	0.40E-04		6.5	10.641	0.57E-06
7.0	9.176	0.99E-04		7.0	10.262	0.23E-05		7.0	11.459	7.80E-09
7.5	9.832	9.70E-06		7.5	10.996	0.74E-07		7.5	12.278	3.80E-11
8.0	10.487	0.64E-06		8.0	11.729	1.20E-09		8.0	13.096	OOO
8.5	11.142	2.50E-08		8.5	12.462	8.60E-12		8.5	13.915	OOO
9.0	11.798	0.58E-09		9.0	13.195	OOO		9.0	14.733	OOO
9.5	12.453	0.71E-11		9.5	13.928	OOO		9.5	15.552	OOO
10.0	13.109	OOO		10.0	14.661	OOO		10.0	16.370	OOO
10.5	13.764	OOO		10.5	15.394	OOO		10.5	17.189	OOO
11.0	14.420	OOO		11.0	16.127	OOO		11.0	18.007	OOO
11.5	15.075	OOO		11.5	16.860	OOO		11.5	18.826	OOO
12.0	15.731	OOO		12.0	17.593	OOO		12.0	19.644	OOO
12.5	16.386	OOO		12.5	18.326	OOO		12.5	20.463	OOO
13.0	17.041	OOO		13.0	19.059	OOO		13.0	21.281	OOO

Table 6.9 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 7$ Extra Dimensions. The color-coding scheme is defined as follows. Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOO) benchmarks.*

$\sqrt{\theta}M_D$ 0.74				$\sqrt{\theta}M_D$ 0.75		
M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)		M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
2.5	4.563	1.90E+01		2.5	5.081	8.60E+00
3.0	5.476	2.80E+00		3.0	6.097	0.96E+00
3.5	6.389	0.38E+00		3.5	7.113	9.40E-02
4.0	7.301	4.50E-02		4.0	8.129	7.10E-03
4.5	8.214	4.30E-03		4.5	9.145	3.50E-04
5.0	9.127	0.28E-03		5.0	10.161	0.86E-05
5.5	10.040	0.11E-04		5.5	11.178	0.67E-07
6.0	10.952	0.17E-06		6.0	12.194	0.12E-09
6.5	11.865	8.70E-10		6.5	13.210	OOOR
7.0	12.778	0.48E-12		7.0	14.226	OOOR
7.5	13.690	OOOR		7.5	15.242	OOOR
8.0	14.603	OOOR		8.0	16.258	OOOR
8.5	15.516	OOOR		8.5	17.274	OOOR
9.0	16.428	OOOR		9.0	18.291	OOOR
9.5	17.341	OOOR		9.5	19.307	OOOR
10.0	18.254	OOOR		10.0	20.323	OOOR
10.5	19.166	OOOR		10.5	21.339	OOOR
11.0	20.079	OOOR		11.0	22.355	OOOR
11.5	20.992	OOOR		11.5	23.371	OOOR
12.0	21.904	OOOR		12.0	24.388	OOOR
12.5	22.817	OOOR		12.5	25.404	OOOR
13.0	23.730	OOOR		13.0	26.420	OOOR

Table 6.10 *Noncommutative Black Hole Parameters M_D , $m_{\min}$, and Cross Section (σ) for a given $\sqrt{\theta}M_D$ for $n = 7$ Extra Dimensions. The color-coding scheme is defined as follows. Green: benchmarks that are unexcluded and yield at 5 events at 139 fb^{-1} at 13 TeV. Blue: benchmarks that yield under 5 events. Purple: out of reach (OOOR) benchmarks.*

The studies detailed above were used to determine benchmarks for the analysis. The noncommutative black holes benchmarks analyzed in this thesis are shown in Table 6.11. All 6 of these benchmarks are used in the subsequent truth-level studies. Truth-level kinematic studies were performed on all benchmarks that were determined to be potentially reachable (meaning, all benchmarks that appear in green in the previous tables). However, for conciseness, the kinematics of only 6 are presented in this thesis. As will be seen in the next chapter (in Figure 7.21), the value of the noncommutative scale can affect the transverse momentum spectrum of the model. To demonstrate the effect of noncommutative scale on transverse momentum, two different benchmarks, NCBH1 and NCBH4 were chosen. Benchmark NCBH1, which has a noncommutative scale of 0.55, has a higher transverse momentum spectrum (demonstrated for large R jets in Figure 7.21) than benchmark NCBH4, which has a noncommutative scale of 0.63. When requesting Monte Carlo samples in the collaboration, a prominent question from convenership was how varying the minimum mass of the black hole ($m_{\min}$), while fixing the number of extra dimensions and noncommutative scale, affects the kinematic distribution of the model. Two additional benchmarks (NCBH2 and NCBH3) were created to demonstrate how the kinematics change when taking the same number of extra dimensions and the same noncommutative scale as NCBH1, but increasing the minimum black hole mass from

6.8 TeV to 7.9 TeV and 8.9 TeV, respectively. Due to the relation between $m_{\min}$ and $\sqrt{\theta}M_D$ seen in Table 6.1 ($(\frac{m_{\min}}{M_D})(\sqrt{\theta}M_D)^{-(n+1)}$), changing the minimum black hole mass ($m_{\min}$) also changes M_D . The M_D values for benchmarks NCBH2 and NCBH3 are 7.5 TeV and 8.5 TeV (an increase from the M_D value of 6.5 TeV for benchmark NCBH1.) Similarly, benchmarks NCBH5 and NCBH6 were created to demonstrate how the kinematics change when taking the same number of extra dimensions and the same noncommutative scale as benchmark NCBH4, but increasing the minimum black hole mass from 6.7 TeV to 7.8 TeV and 8.9 TeV, respectively. The M_D values for benchmarks NCBH2 and NCBH3 are 7 TeV and 8 TeV (an increase from the M_D value of 6 TeV for benchmark NCBH1). As will be seen in the next chapter in Figure 7.21, the minimum black hole mass does not have a significant impact on the kinematics, unlike the noncommutative scale. Since this analysis uses machine learning, 500,000 events were needed for each reconstructed black hole sample. Due to this relatively large requirement for the number of events needed for generation of the sample, convenership limited the number of reconstructed noncommutative black hole samples to 2 samples, noting that the statistical result of the analysis can be scaled, such that statements could be made on the cross sections of other black hole benchmarks. (The number of reconstructed samples for non-minimal dark sectors, which also required 500,000 events was similarly limited to 3. Scalar cascades, which required 2 million events due to low trigger efficiencies, were permitted a single reconstructed sample.) Since this is the first analysis for noncommutative black holes, and since it was not certain how sensitive the analysis would be to the model, benchmarks NCBH1 and NCBH4 were chosen because, out of the samples with a noncommutative scale of 0.55, benchmark NCBH1 had the highest cross section and, out of the benchmarks with a noncommutative scale of 0.63, benchmark NCBH4 had the highest cross section. Lower minimum black hole masses allow for higher cross sections.

Signal Sample	Number of Extra Dimensions	Noncommutative Scale	M_D (TeV)	$m_{\min}$ (TeV)	σ (pb)
NCBH1	5	0.55	6.5	6.819	0.022
NCBH2	5	0.55	7.5	7.868	0.0016
NCBH3	5	0.55	8.5	8.917	0.000082
NCBH4	6	0.63	6	6.668	0.047
NCBH5	6	0.63	7	7.779	0.0031
NCBH6	6	0.63	8	8.886	0.00013

Table 6.11 *Parameters and Cross Sections (σ) for Noncommutative Black Hole Samples*

6.2 Non-minimal Dark Sector Parameters and MC Samples

The non-minimal dark sector benchmarks analyzed in this thesis are shown in Table 6.12. As noted at the beginning of the chapter, the full unexcluded phase space for non-minimal dark sectors is not mapped out in this thesis due to how elaborate and intricate the Monte Carlo generation process is for non-minimal dark sectors, as well as how expansive the phase space is due to the non-minimal dark sector model having 6 free parameters. Instead, 6 benchmarks that have all evaded the exclusion limits from the SUSY multijet and monojet searches mentioned earlier are included. All 6 of these benchmarks are used in the subsequent truth level studies. Two of these benchmarks, NMDS2 and NMDS4, are taken from recommendations from the non-minimal dark sector theorists in [58], where it was noted that they evaded the exclusion limits from the SUSY multijet and monojet searches. The remaining four benchmarks were designed by the author by choosing values for each of the 6 free parameters of the model for each benchmark. This chapter will detail how the various parameters of the non-minimal dark sector model affect the phenomenology of the model. In particular, it will be described how the parameters of the model are adjusted to affect the transverse momentum spectrum of the model. As will be explained in detail later, the non-minimal dark sector benchmarks NMDS2 and NMDS4 provide insight into a regime characterized by moderate p_T . Benchmarks NMDS5, NMDS1, and NMDS6 were designed specifically to provide insight into a regime characterized by low p_T . These benchmarks evade the model-independent limits from the mono-jet analysis due to passing the missing energy triggers at low rates (as seen in Tables 13.26, 13.24, and 13.22 in the appendix) and the model-independent limits from the multijet analysis due to passing the multijet (many jets) triggers at low rates (again, as seen in Tables 13.26, 13.24, and 13.22). Benchmark NMDS3 was designed to probe the high p_T , or severely boosted, region of phase space. The benchmark NMDS3 evades exclusion limits due to its very small cross section, as seen in 6.12. Benchmarks NMDS1, NMDS2, and NMDS3 are used in the reconstructed analysis as they each represent one of the three p_T regimes: low p_T , moderate p_T , and high p_T .

Signal Sample	m_ϕ	Δm	m_0	γ	c_0	δ	σ (pb)
NMDS1	650	50	500	1	0.1	0.5	0.11
NMDS2	800	50	500	1	0.1	1	0.0271
NMDS3	1841	600	500	1	0.1	0.5	0.000133
NMDS4	1000	50	500	1	0.1	1	0.00778
NMDS5	600	10	500	3	0.1	1	0.164
NMDS6	561	25	500	2	0.1	0.5	0.215

Table 6.12 *Parameters and Cross Sections (σ) for Non-minimal Dark Sector Signal Samples*

Tables 6.13, 6.14, and 6.15 show the parameter values and cross sections for all subprocesses of $\chi\chi$ production, all subprocesses of $\phi\chi$ production, and the $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS4. All tables in this section were produced via calculations performed by the author. The next set of passages describe the calculations performed in order to create a benchmark. Though the procedure is outlined for the Non-minimal Dark Sector Benchmark NMDS4, the general procedure is the same for all benchmarks.

As a reminder, dark fermions in the non-minimal dark sector model decay via the following decay chain.

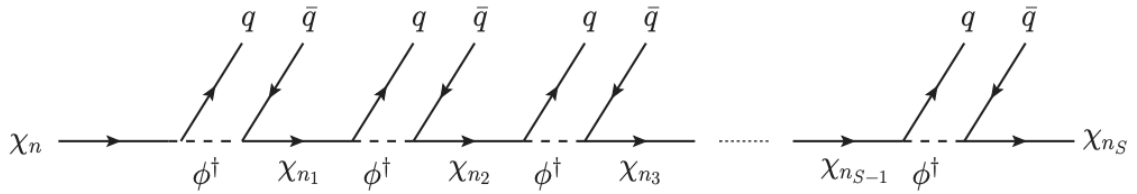


Figure 6.1 *Decay chain in which each χ (ensemble constituent) decays into a χ and two quarks. The process ends with the production of two quarks and a stable, lightest χ .*

As noted previously, the mass relation between each ensemble constituent in the decay is governed by

$$m_n = m_0 + n^\delta \Delta m. \quad (6.1)$$

The value n represents the n^{th} step of the ensemble. The mass of the lowest ensemble constituent, or the mass of the ensemble constituent at the 0^{th} step of the ensemble, is denoted as m_0 . The mass of the ensemble constituent at the n^{th} step of the ensemble is denoted as m_n .

The coupling relation to describe the coupling at different steps in the ensemble is c_{nq} , where

$$c_{nq} = c_{0q} \left(\frac{m_n}{m_0} \right)^\gamma \quad (6.2)$$

and γ is the coupling scaling exponent.

The parameters of benchmark NMDS4 are defined in Table 6.12. The mass of the scalar mediator, m_ϕ , is 1000 GeV. The mass of the lowest ensemble constituent, m_0 , is 500 GeV. The Δm is 50 GeV. As δ is defined as 1, the mass relation becomes

$$m_n = 500 + n(50). \quad (6.3)$$

The number of ensemble constituents theoretically allowed when decaying from a scalar mediator mass of 1000 GeV to the lowest ensemble constituent mass of 500 GeV in steps of 50 GeV is 10. Explicitly, the masses of the ensemble constituents from $n=0$ to $n=10$ are 500 GeV, 550 GeV, 600 GeV, 650 GeV, 700 GeV, 750 GeV, 800 GeV, 850 GeV, 900 GeV, and 950 GeV. In the $\chi\chi$ production process of non-minimal dark sectors, there are two χ 's. Therefore, all possible ensemble constituents from both χ decays need to be considered. Therefore, the number of theoretically possible unique combinations of ensemble constituents that the $\chi\chi$ process can transition to have to be considered. Table 6.13 accounts for each of these possible combinations of ensemble constituents in the $\chi\chi$ production process of non-minimal dark sectors. For example, the scalar mediator of 1000 GeV can decay directly to a χ_m with a mass of 850 GeV and a χ_n with a mass of 650 GeV. Similarly, it could have decayed directly to a χ_m with a mass of 550 GeV and a χ_n with a mass of 500 GeV. All of these possibilities have to be accounted for.

In Table 6.13, the column “ m of χ_m ” indicates that the ensemble constituent is at the m^{th} step of the ensemble. The column “ n of χ_n ” indicates that the ensemble constituent is at the n^{th} step of the ensemble. (There is both a χ_m and a χ_n in the $\chi\chi$ production process because there are two χ 's. In the $\phi\chi$ process, there is only one χ , so the tables for the $\phi\chi$ production process for each benchmark only have a column for “ m of χ_m ”.) The column m_m refers to the mass of χ_m , while the m_n refers to the mass of χ_n . A value of 0 for “ m of χ_m ” and a value of 0 for “ n of χ_n ” corresponds to a scenario in which the scalar mediator (with a mass of

1000 GeV) decays directly to a χ_m with a mass of 500 GeV and a χ_n with a mass of 500 GeV. A value of 4 for “ m of χ_m ” and a value of 8 for “ n of χ_n ” corresponds to a scenario in which the scalar mediator decays directly to a χ_m with a mass of 700 GeV and a χ_n with a mass of 900 GeV. These masses were calculated by the author using Equation 6.1. The c_n and c_m refer to the coupling of that particular χ_n and χ_m . These couplings were calculated by the author using Equation 6.2. The σ (pb) refers to the specific cross section of that particular subprocess. (Each possible decay state of the scalar mediator, meaning each possible combination of χ_m and χ_n , is referred to as a subprocess.) In order to obtain the cross section of a subprocess, the author generated the subprocess with version 2.3.3 of the Monte Carlo generator MadGraph [93]. (MadGraph simulates the proton-proton collision of a given process.) In MadGraph, the author used Universal FeynRules Output (UFO) [94] models provided directly by the theorists from [58]. UFO is a model format that contains the details of the model for processing by MadGraph. (No citation exists for the specific UFO models provided for non-minimal dark sectors.)

The cross section was obtained from the resulting LHE file. The “Relative σ ” is the cross section of a particular subprocess with respect to all processes and subprocesses throughout the entire MC sample. This will be described in more detail later, but it refers to calculating the cross sections of all subprocesses in each of the production processes for non-minimal dark sectors ($\chi\chi$ production, $\phi\chi$ production, and $\phi\phi$ production), summing these cross sections to produce a total cross section for the benchmark, and then calculating the relative cross section of a specific subprocess with respect to the total cross section. This again was calculated by the author. The column “Total number of Events” refers to the number of events in the Monte Carlo sample that were produced for a given subprocess.

$\chi\chi$: $\delta = 1$, $m_\phi = 1000$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$									
m of χ_m	n of χ_n	m_m	m_n	c_m	c_n	σ (pb)	Relative σ (pb)	Subprocess	Total number of Events
0	0	500	500	0.10	0.10	1.50E-06	1.93E-04	S1	21.21
1	0	550	500	0.11	0.10	1.62E-06	2.08E-04	S2	22.90
2	0	600	500	0.12	0.10	1.58E-06	2.04E-04	S3	22.39
3	0	650	500	0.13	0.10	1.70E-06	2.18E-04	S4	24.01
4	0	700	500	0.14	0.10	1.87E-06	2.40E-04	S5	26.38
5	0	750	500	0.15	0.10	1.80E-06	2.31E-04	S6	25.39
6	0	800	500	0.16	0.10	1.76E-06	2.27E-04	S7	24.95
7	0	850	500	0.17	0.10	1.86E-06	2.38E-04	S8	26.23
8	0	900	500	0.18	0.10	1.84E-06	2.37E-04	S9	26.04
9	0	950	500	0.19	0.10	1.71E-06	2.20E-04	S10	24.21
1	1	550	550	0.11	0.11	1.78E-06	2.29E-04	S11	25.15
2	1	600	550	0.12	0.11	1.76E-06	2.26E-04	S12	24.90
3	1	650	550	0.13	0.11	1.87E-06	2.40E-04	S13	26.43
4	1	700	550	0.14	0.11	1.89E-06	2.43E-04	S14	26.71
5	1	750	550	0.15	0.11	2.05E-06	2.63E-04	S15	28.95
6	1	800	550	0.16	0.11	1.93E-06	2.48E-04	S16	27.23
7	1	850	550	0.17	0.11	1.98E-06	2.55E-04	S17	28.00
8	1	900	550	0.18	0.11	2.01E-06	2.59E-04	S18	28.48
9	1	950	550	0.19	0.11	1.91E-06	2.45E-04	S19	26.94
2	2	600	600	0.12	0.12	1.81E-06	2.33E-04	S20	25.61
3	2	650	600	0.13	0.12	1.93E-06	2.48E-04	S21	27.31
4	2	700	600	0.14	0.12	1.97E-06	2.54E-04	S22	27.92
5	2	750	600	0.15	0.12	2.06E-06	2.64E-04	S23	29.06
6	2	800	600	0.16	0.12	1.18E-05	1.52E-03	S24	166.94
7	2	850	600	0.17	0.12	2.01E-06	2.58E-04	S25	28.36
8	2	900	600	0.18	0.12	2.07E-06	2.66E-04	S26	29.25
9	2	950	600	0.19	0.12	1.99E-06	2.56E-04	S27	28.14
3	3	650	650	0.13	0.13	2.08E-06	2.67E-04	S28	29.38
4	3	700	650	0.14	0.13	2.11E-06	2.71E-04	S29	29.79
5	3	750	650	0.15	0.13	2.13E-06	2.74E-04	S30	30.18
6	3	800	650	0.16	0.13	2.30E-06	2.95E-04	S31	32.49
7	3	850	650	0.17	0.13	2.31E-06	2.97E-04	S32	32.67
8	3	900	650	0.18	0.13	2.10E-06	2.70E-04	S33	29.75
9	3	950	650	0.19	0.13	2.09E-06	2.69E-04	S34	29.57
4	4	700	700	0.14	0.14	2.19E-06	2.81E-04	S35	30.95
5	4	750	700	0.15	0.14	2.27E-06	2.92E-04	S36	32.15
6	4	800	700	0.16	0.14	2.36E-06	3.04E-04	S37	33.39
7	4	850	700	0.17	0.14	2.29E-06	2.95E-04	S38	32.42
8	4	900	700	0.18	0.14	2.22E-06	2.85E-04	S39	31.35
9	4	950	700	0.19	0.14	2.31E-06	2.97E-04	S40	32.70
5	5	750	750	0.15	0.15	2.41E-06	3.10E-04	S41	34.09
6	5	800	750	0.16	0.15	2.37E-06	3.05E-04	S42	33.54
7	5	850	750	0.17	0.15	2.41E-06	3.09E-04	S43	34.04
8	5	900	750	0.18	0.15	2.48E-06	3.19E-04	S44	35.10
9	5	950	750	0.19	0.15	2.23E-06	2.87E-04	S45	31.54
6	6	800	800	0.16	0.16	2.41E-06	3.09E-04	S46	34.03
7	6	850	800	0.17	0.16	2.33E-06	2.99E-04	S47	32.92
8	6	900	800	0.18	0.16	2.41E-06	3.10E-04	S48	34.14
9	6	950	800	0.19	0.16	2.36E-06	3.03E-04	S49	33.31
7	7	850	850	0.17	0.17	2.50E-06	3.21E-04	S50	35.30
8	7	900	850	0.18	0.17	2.36E-06	3.04E-04	S51	33.42
9	7	950	850	0.19	0.17	2.52E-06	3.24E-04	S52	35.59
8	8	900	900	0.18	0.18	2.46E-06	3.16E-04	S53	34.78
9	8	950	900	0.19	0.18	2.42E-06	3.11E-04	S54	34.23
9	9	950	950	0.19	0.19	2.58E-06	3.32E-04	S55	36.54

Table 6.13 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS4*

The description of the calculations performed in order to create a benchmark (in this case the non-minimal dark sector benchmark NMDS4) is not yet done, and instead continues with a description of the calculations performed for the $\phi\chi$ production process of benchmark NMDS4. In Table 6.14, the possible decay states of a $\phi\chi$

process are shown. Here, there is only one χ . Therefore, for this benchmark in which the lowest ensemble constituent is 500 GeV and there are only 10 possible ensemble constituents to which the scalar mediation of mass 1000 GeV can decay to, only 10 possible decay states are listed. The author generated each subprocess with version 2.3.3 of MadGraph. The cross section for each subprocess was obtained from the resulting LHE file. The total cross section of the $\phi\chi$ process was calculated by summing the cross sections of the individual subprocesses.

$\phi\chi$: $\delta = 1$, $m_\phi = 1000$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$					
m of χ_m	c_m	m_m	σ (pb)	Relative σ (pb)	Total Events
0	0.10	500	2.59E-04	3.33E-02	3666.52
1	0.11	550	2.97E-04	3.82E-02	4205.91
2	0.12	600	3.20E-04	4.11E-02	4525.50
3	0.13	650	3.36E-04	4.31E-02	4745.49
4	0.14	700	3.50E-04	4.50E-02	4954.62
5	0.15	750	3.56E-04	4.58E-02	5032.78
6	0.16	800	3.86E-04	4.96E-02	5458.91
7	0.17	850	3.50E-04	4.50E-02	4954.77
8	0.18	900	3.54E-04	4.55E-02	5004.20
9	0.19	950	3.47E-04	4.46E-02	4910.69

Table 6.14 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS4*

The description of the calculations performed in order to create a benchmark (in this case the non-minimal dark sector benchmark NMDS4) continues with a description of the calculations performed for the $\phi\phi$ production process of benchmark NMDS4. In Table 6.15, the $\phi\phi$ process is shown. The author generated this one state with version 2.3.3 of MadGraph. The cross section was obtained from the resulting LHE file. The total cross section of the $\chi\chi$ process, the total cross section of the $\phi\chi$ process, and the total cross section of the $\phi\phi$ process were summed to provide a total cross section for the entire benchmark. As mentioned previously, relative cross sections were calculated by the author for the $\phi\phi$ process and each individual subprocess of the $\phi\chi$ and $\chi\chi$ processes. These relative cross sections were then used to ensure that each subprocess was generated with a number of events that reflects its relative weight.

$\phi\phi$: $\delta = 1$, $m_\phi = 1000$ TeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$				
m of χ_m	n of χ_n	σ (pb)	Relative σ (pb)	Total Events
0	0	4.30E-03	0.55	6.08E+04

Table 6.15 *Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS4*

Once all of subprocesses were generated with the correct number of events in

MadGraph, the author passed the resultant LHE files through special code developed by the non-minimal dark sector theorists to further decay the χ_m and χ_n in the case of the $\chi\chi$ process, or further decay the χ_m in the case of the $\phi\chi$, or further decay the results of the $\phi\phi$ decay. Once these further decays were executed, the final files were hadronized with Pythia. (The theorists sent their code in files directly to the author. No repository is available for citation.)

The mass spacing (or mass difference) between the different χ 's in the ensemble is dictated by the mass relation equation 6.1 and determines the p_T of the jets. A lower mass spacing yields a lower jet p_T spectrum. A very large mass spacing yields higher p_T signals. The coupling determines the number of interactions between various states of the ensemble. Stronger coupling yields more interactions, while weaker coupling yields less. The benchmark NMDS4 with its moderate mass spacing of 50 GeV yields moderate p_T signals. An ensemble has equal mass spacings between each ensemble constituent only when $\delta = 1$ due to the mass relation equation 6.1. Therefore, benchmark NMDS4 has equal mass spacings between each ensemble constituent.

This concludes the description of the procedure for generating a single non-minimal dark sector benchmark. This procedure was used to generate each of the 6 non-minimal dark sector benchmarks in Table 6.12. The rest of the section describes the parameters and properties of the remaining 5 benchmarks. For each benchmark, 3 tables are provided. There is one for each of the following production processes: $\phi\chi$ production, $\chi\chi$ production, and $\phi\phi$ production.

Tables 6.16, 6.17, and 6.18 show the parameter values and cross sections for all subprocesses of $\chi\chi$ production, all subprocesses of $\phi\chi$ production, and the $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS2. With a mass spacing identical to that of the NMDS4 benchmark, the NMDS2 benchmark similarly yields a signature with moderate p_T .

$\chi\chi$: $\delta = 1$, $m_\phi = 800$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$									
m of χ_m	n of χ_n	m_m	m_n	c_m	c_n	σ (pb)	Relative σ (pb)	Subprocess	Total number of Events
0	0	500	500	0.10	0.10	2.82E-06	1.04E-04	S1	12
1	0	550	500	0.11	0.10	3.00E-06	1.10E-04	S2	13
2	0	600	500	0.12	0.10	3.12E-06	1.15E-04	S3	13
3	0	650	500	0.13	0.10	3.21E-06	1.18E-04	S4	14
4	0	700	500	0.14	0.10	3.27E-06	1.20E-04	S5	14
5	0	750	500	0.15	0.10	3.28E-06	1.21E-04	S6	14
1	1	550	550	0.11	0.11	3.19E-06	1.18E-04	S7	13
2	1	600	550	0.12	0.11	3.34E-06	1.23E-04	S8	14
3	1	650	550	0.13	0.11	3.44E-06	1.27E-04	S9	14
4	1	700	550	0.14	0.11	3.51E-06	1.29E-04	S10	15
5	1	750	550	0.15	0.11	3.54E-06	1.30E-04	S11	15
2	2	600	600	0.12	0.12	3.52E-06	1.30E-04	S12	15
3	2	650	600	0.13	0.12	3.63E-06	1.34E-04	S13	15
4	2	700	600	0.14	0.12	3.70E-06	1.36E-04	S14	16
5	2	750	600	0.15	0.12	3.74E-06	1.38E-04	S15	16
3	3	650	650	0.13	0.13	3.77E-06	1.39E-04	S16	16
4	3	700	650	0.14	0.13	3.85E-06	1.42E-04	S17	16
5	3	750	650	0.15	0.13	3.90E-06	1.44E-04	S18	16
4	4	700	700	0.14	0.14	3.95E-06	1.46E-04	S19	17
5	4	750	700	0.15	0.14	4.01E-06	1.48E-04	S20	17
5	5	750	750	0.15	0.15	4.09E-06	1.51E-04	S21	17

Table 6.16 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS2*

$\phi\chi$: $\delta = 1$, $m_\phi = 800$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$					
m of χ_m	c_m	m_m	σ (pb)	Relative σ (pb)	Total Events
0	0.10	500	8.60E-04	3.17E-02	3485
1	0.11	550	9.18E-04	3.38E-02	3723
2	0.12	600	9.65E-04	3.56E-02	3911
3	0.13	650	9.97E-04	3.67E-02	4042
4	0.14	700	1.02E-03	3.76E-02	4131
5	0.15	750	1.03E-03	3.78E-02	4160

Table 6.17 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS2*

$\phi\phi$: $\delta = 1$: $m_\phi = 800$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$				
m of χ_m	n of χ_n	σ (pb)	Relative σ (pb)	Total Events
0	0	2.13E-02	7.84E-01	86251

Table 6.18 *Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS2*

Tables 6.19, 6.20, and 6.21 show the parameter values and cross sections for all subprocesses of $\chi\chi$ production, all subprocesses of $\phi\chi$ production, and the $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS3. The NMDS3 benchmark has a δ value of 0.5 which forces the mass spacing between each possible ensemble constituent to be inconsistent. The Δm value of 600 GeV is a remarkably large mass gap. The combination of a δ value of 0.5 and a Δm value of 600 GeV lends to huge mass spacing between the lowest ensemble consistents and then gradually

decreasing mass spacing between higher ensemble constituents. The smallest the mass spacing gets between ensemble constituents is 160.8 GeV, which is still relatively large compared to the 50 GeV mass spacing of benchmarks NMDS1, NMDS2, and NMDS4. The smallest mass spacing between ensemble constituents can be seen in Table 6.19 by looking at the mass, m_m , of the χ_m with $m = 3$ (1539.20 GeV) and the mass of the χ_m with $m = 4$ (1700 GeV). The largest mass spacing between ensemble constituents can be seen in Table 6.19 by looking at the mass of the χ_m with an m of 0 (500 GeV) and the mass of the χ_m with an m of 1 (1100 GeV).

The largest the mass spacing becomes is 600 GeV. These mass spacing values causes this benchmark to be characterized by very energetic decays.

$\chi\chi$: $\delta = 0.5$, $m_\phi = 1841.64$ GeV, $m_0 = 500$ GeV, $\Delta m = 600$ GeV, $\gamma = 1$, $c_0 = 0.1$									
m of χ_m	n of χ_n	m_m	m_n	c_m	c_n	σ (pb)	Relative σ (pb)	Subprocess	Total number of Events
0	0	500.00	500.00	0.10	0.10	3.07E-07	2.31E-03	S1	255
1	0	1100.00	500.00	0.22	0.10	1.53E-07	1.15E-03	S2	127
2	0	1348.50	500.00	0.27	0.10	3.99E-07	3.00E-03	S3	330
3	0	1539.20	500.00	0.31	0.10	3.33E-07	2.51E-03	S4	276
4	0	1700.00	500.00	0.34	0.10	3.14E-07	2.36E-03	S5	260
1	1	1100.00	1100.00	0.22	0.22	7.40E-07	5.56E-03	S6	612
2	1	1348.50	1100.00	0.27	0.22	6.93E-07	5.21E-03	S7	574
3	1	1539.20	1100.00	0.31	0.22	6.33E-07	4.76E-03	S8	524
4	1	1700.00	1100.00	0.34	0.22	5.70E-07	4.29E-03	S9	472
2	2	1348.50	1348.50	0.27	0.27	6.95E-07	5.22E-03	S10	575
3	2	1539.20	1348.50	0.31	0.27	6.15E-07	4.62E-03	S11	509
4	2	1700.00	1348.50	0.34	0.27	5.48E-07	4.12E-03	S12	454
3	3	1539.20	1539.20	0.31	0.31	5.47E-07	4.11E-03	S13	453
4	3	1700.00	1539.20	0.34	0.31	5.06E-07	3.81E-03	S14	419
4	4	1700.00	1700.00	0.34	0.34	4.60E-07	3.46E-03	S15	381

Table 6.19 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS3*

$\phi\chi$: $\delta = 1$, $m_\phi = 1841.64$ GeV, $m_0 = 500$ GeV, $\Delta m = 600$ GeV, $\gamma = 1$, $c_0 = 0.1$					
m of χ_m	c_m	m_m	σ (pb)	Relative σ (pb)	Total Events
0	0.10	500.00	1.47E-05	1.10E-01	12134
1	0.22	1100.00	2.48E-05	1.87E-01	20548
2	0.27	1348.50	2.31E-05	1.74E-01	19107
3	0.31	1539.20	2.07E-05	1.55E-01	17084
4	0.34	1700.00	1.82E-05	1.37E-01	15050

Table 6.20 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS3*

$\phi\phi$: $\delta = 1$, $m_\phi = 1841.64$ GeV, $m_0 = 500$ GeV, $\Delta m = 600$ GeV, $\gamma = 1$, $c_0 = 0.1$				
m of χ_m	n of χ_n	σ (pb)	Relative σ (pb)	Total Events
0	0	2.40719E-05	0.180921021	19901.31227

Table 6.21 *Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS3*

Tables 6.22, 6.23, and 6.24 show the parameter values and cross sections for all subprocesses of $\chi\chi$ production, all subprocesses of $\phi\chi$ production, and the $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS5. The NMDS benchmark NMDS5 has a very small Δm value of 10 GeV. The goal of this small Δm is to allow for a benchmark that probes the low p_T regime. The use of a higher γ value is an attempt to force stronger coupling between neighboring ensemble constituents and dissuade loose coupling that prompts direct decays to lower ensemble constituent and increases the mass spacing and p_T of resultant decays.

$\chi\chi$: $\delta = 1$, $m_\phi = 600$ GeV, $m_0 = 500$ GeV, $\Delta m = 10$ GeV, $\gamma = 3$, $c_0 = 0.1$									
m of χ_m	n of χ_n	m_m	m_n	c_m	c_n	σ (pb)	Relative σ (pb)	Subprocess	Total number of Events
0	0	500	500	0.10	0.10	4.61E-06	2.81E-05	S1	4
1	0	510	500	0.11	0.10	5.06E-06	3.08E-05	S2	4
2	0	520	500	0.11	0.10	5.48E-06	3.34E-05	S3	4
3	0	530	500	0.12	0.10	6.04E-06	3.68E-05	S4	5
4	0	540	500	0.13	0.10	6.51E-06	3.97E-05	S5	5
5	0	550	500	0.13	0.10	7.22E-06	4.40E-05	S6	5
6	0	560	500	0.14	0.10	7.77E-06	4.74E-05	S7	6
7	0	570	500	0.15	0.10	8.36E-06	5.10E-05	S8	6
8	0	580	500	0.16	0.10	9.14E-06	5.57E-05	S9	7
9	0	590	500	0.16	0.10	9.81E-06	5.98E-05	S10	7
1	1	510	510	0.11	0.11	5.59E-06	3.41E-05	S11	4
2	1	520	510	0.11	0.11	6.09E-06	3.71E-05	S12	5
3	1	530	510	0.12	0.11	6.71E-06	4.09E-05	S13	5
4	1	540	510	0.13	0.11	7.31E-06	4.46E-05	S14	5
5	1	550	510	0.13	0.11	7.82E-06	4.77E-05	S15	6
6	1	560	510	0.14	0.11	8.47E-06	5.17E-05	S16	6
7	1	570	510	0.15	0.11	9.23E-06	5.63E-05	S17	7
8	1	580	510	0.16	0.11	9.94E-06	6.06E-05	S18	7
9	1	590	510	0.16	0.11	1.07E-05	6.54E-05	S19	8
2	2	520	520	0.11	0.11	6.67E-06	4.07E-05	S20	5
3	2	530	520	0.12	0.11	7.35E-06	4.48E-05	S21	5
4	2	540	520	0.13	0.11	7.98E-06	4.87E-05	S22	6
5	2	550	520	0.13	0.11	8.61E-06	5.25E-05	S23	6
6	2	560	520	0.14	0.11	9.04E-06	5.51E-05	S24	7
7	2	570	520	0.15	0.11	9.81E-06	5.98E-05	S25	7
8	2	580	520	0.16	0.11	1.08E-05	6.61E-05	S26	8
9	2	590	520	0.16	0.11	1.16E-05	7.08E-05	S27	8
3	3	530	530	0.12	0.12	7.90E-06	4.82E-05	S28	6
4	3	540	530	0.13	0.12	8.48E-06	5.17E-05	S29	6
5	3	550	530	0.13	0.12	9.40E-06	5.73E-05	S30	7
6	3	560	530	0.14	0.12	1.01E-05	6.13E-05	S31	7
7	3	570	530	0.15	0.12	1.07E-05	6.54E-05	S32	8
8	3	580	530	0.16	0.12	1.19E-05	7.26E-05	S33	8
9	3	590	530	0.16	0.12	1.26E-05	7.68E-05	S34	9
4	4	540	540	0.13	0.13	9.46E-06	5.77E-05	S35	7
5	4	550	540	0.13	0.13	1.03E-05	6.27E-05	S36	7
6	4	560	540	0.14	0.13	1.10E-05	6.69E-05	S37	8
7	4	570	540	0.15	0.13	1.19E-05	7.27E-05	S38	8
8	4	580	540	0.16	0.13	1.27E-05	7.76E-05	S39	9
9	4	590	540	0.16	0.13	1.39E-05	8.50E-05	S40	10
5	5	550	550	0.13	0.13	1.09E-05	6.63E-05	S41	8
6	5	560	550	0.14	0.13	1.19E-05	7.24E-05	S42	8
7	5	570	550	0.15	0.13	1.29E-05	7.87E-05	S43	9
8	5	580	550	0.16	0.13	1.39E-05	8.49E-05	S44	10
9	5	590	550	0.16	0.13	1.50E-05	9.17E-05	S45	11
6	6	560	560	0.14	0.14	1.25E-05	7.65E-05	S46	9
7	6	570	560	0.15	0.14	1.39E-05	8.49E-05	S47	10
8	6	580	560	0.16	0.14	1.50E-05	9.13E-05	S48	11
9	6	590	560	0.16	0.14	2.57E-06	1.57E-05	S49	2
7	7	570	570	0.15	0.15	1.52E-05	9.27E-05	S50	11
8	7	580	570	0.16	0.15	1.64E-05	1.00E-04	S51	12
9	7	590	570	0.16	0.15	1.73E-05	1.05E-04	S52	12
8	8	900	900	0.18	0.18	2.46E-06	3.16E-04	S53	35
9	8	950	900	0.19	0.18	2.42E-06	3.11E-04	S54	35
9	9	950	950	0.19	0.19	2.58E-06	3.32E-04	S55	37

Table 6.22 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS5*

$\phi\chi$: $\delta = 1$, $m_\phi = 600$ GeV, $m_0 = 500$ GeV, $\Delta m = 10$ GeV, $\gamma = 3$, $c_0 = 0.1$						
m of χ_m	c_m	m_m	Subprocess	σ (pb)	Relative σ (pb)	Total Number of Events
0	0.10	500	S1	2.44E-03	1.49E-02	1635
1	0.11	510	S2	2.68E-03	1.63E-02	1797
2	0.11	520	S3	2.83E-03	1.72E-02	1897
3	0.12	530	S4	3.13E-03	1.91E-02	2099
4	0.13	540	S5	3.52E-03	2.14E-02	2359
5	0.13	550	S6	3.67E-03	2.24E-02	2462
6	0.14	560	S7	3.97E-03	2.42E-02	2660
7	0.15	570	S8	4.30E-03	2.62E-02	2883
8	0.16	580	S9	4.63E-03	2.83E-02	3109
9	0.16	590	S10	2.48E-03	1.51E-02	1661

Table 6.23 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS5*

$\phi\phi$: $\delta = 1$, $m_\phi = 600$ GeV, $m_0 = 500$ GeV, $\Delta m = 10$ GeV, $\gamma = 3$, $c_0 = 0.1$				
m of χ_m	n of χ_n	σ (pb)	Relative σ (pb)	Total Number of Events
0	0	1.30E-01	7.91E-01	87041

Table 6.24 *Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS5*

Tables 6.25, 6.26, and 6.27 show the parameter values and cross sections for all subprocesses of $\chi\chi$ production, all subprocesses of $\phi\chi$ production, and the $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS1. The NMDS1 benchmark has a Δm value of 50 GeV. While this mass gap may not seem small, the δ value of 0.5 causes the mass spacing to gradually decrease such that the mass gap between the lowest ensemble constituent and the second ensemble constituent is 50 GeV, while the mass gap between the highest possible ensemble constituent and second highest possible ensemble constituent is less than 10 GeV. In fact, aside from the gap of 50 GeV in mass between the lowest and second lowest possible ensemble constituent, the highest mass spacing between neighboring ensemble constituents is roughly 20 GeV. All other spacings are less than 15 GeV. This combination of Δm and δ allow for a benchmark that probes the low p_T regime.

$\chi\chi$: $\delta = 1/2$, $m_\phi = 650$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$									
m of χ_m	n of χ_n	m_m	m_n	c_m	c_n	σ (pb)	Relative σ (pb)	Subprocess	Total number of Events
0	0	500.00	500.00	0.10	0.10	3.39E-06	3.09E-05	S1	4
1	0	550.00	500.00	0.13	0.10	5.31E-06	4.83E-05	S2	6
2	0	570.70	500.00	0.15	0.10	5.99E-06	5.45E-05	S3	6
3	0	586.60	500.00	0.16	0.10	6.57E-06	5.98E-05	S4	7
4	0	600.00	500.00	0.17	0.10	7.59E-06	6.91E-05	S5	8
5	0	611.80	500.00	0.18	0.10	8.10E-06	7.38E-05	S6	9
6	0	622.50	500.00	0.19	0.10	8.76E-06	7.98E-05	S7	9
7	0	632.30	500.00	0.20	0.10	9.34E-06	8.51E-05	S8	10
8	0	641.40	500.00	0.21	0.10	1.05E-05	9.59E-05	S9	11
1	1	550.00	550.00	0.13	0.13	9.85E-06	8.97E-05	S10	10
2	1	570.70	550.00	0.15	0.13	9.84E-06	8.96E-05	S11	10
3	1	586.60	550.00	0.16	0.13	1.13E-05	1.03E-04	S12	12
4	1	600.00	550.00	0.17	0.13	1.15E-05	1.04E-04	S13	12
5	1	611.80	550.00	0.18	0.13	1.26E-05	1.15E-04	S14	13
6	1	622.50	550.00	0.19	0.13	1.36E-05	1.24E-04	S15	14
7	1	632.30	550.00	0.20	0.13	1.51E-05	1.38E-04	S16	16
8	1	641.40	550.00	0.21	0.13	1.66E-05	1.51E-04	S17	17
2	2	570.70	570.70	0.15	0.15	1.11E-05	1.01E-04	S18	12
3	2	586.60	570.70	0.16	0.15	1.26E-05	1.15E-04	S19	13
4	2	600.00	570.70	0.17	0.15	1.40E-05	1.28E-04	S20	15
5	2	611.80	570.70	0.18	0.15	1.49E-05	1.36E-04	S21	15
6	2	622.50	570.70	0.19	0.15	1.63E-05	1.48E-04	S22	17
7	2	632.30	570.70	0.20	0.15	1.81E-05	1.65E-04	S23	19
8	2	641.40	570.70	0.21	0.15	1.89E-05	1.72E-04	S24	19
3	3	586.60	586.60	0.16	0.16	1.41E-05	1.28E-04	S25	15
4	3	600.00	586.60	0.17	0.16	1.56E-05	1.42E-04	S26	16
5	3	611.80	586.60	0.18	0.16	1.70E-05	1.54E-04	S27	17
6	3	622.50	586.60	0.19	0.16	1.77E-05	1.61E-04	S28	18
7	3	632.30	586.60	0.20	0.16	2.00E-05	1.83E-04	S29	21
8	3	641.40	586.60	0.21	0.16	2.06E-05	1.87E-04	S30	21
4	4	600.00	600.00	0.17	0.17	1.74E-05	1.59E-04	S31	18
5	4	611.80	600.00	0.18	0.17	1.89E-05	1.72E-04	S32	19
6	4	622.50	600.00	0.19	0.17	2.04E-05	1.86E-04	S33	21
7	4	632.30	600.00	0.20	0.17	2.17E-05	1.98E-04	S34	22
8	4	641.40	600.00	0.21	0.17	2.45E-05	2.23E-04	S35	25
5	5	611.80	611.80	0.18	0.18	2.01E-05	1.83E-04	S36	21
6	5	622.50	611.80	0.19	0.18	2.18E-05	1.98E-04	S37	22
7	5	632.30	611.80	0.20	0.18	2.35E-05	2.14E-04	S38	24
8	5	641.40	611.80	0.21	0.18	2.50E-05	2.28E-04	S39	26
6	6	622.50	622.50	0.19	0.19	2.39E-05	2.17E-04	S40	24
7	6	632.30	622.50	0.20	0.19	2.54E-05	2.31E-04	S41	26
8	6	641.40	622.50	0.21	0.19	2.62E-05	2.39E-04	S42	27
7	7	632.30	632.30	0.20	0.20	2.63E-05	2.40E-04	S43	27
8	7	641.40	632.30	0.21	0.20	2.84E-05	2.59E-04	S44	29
8	8	641.40	641.40	0.21	0.21	3.03E-05	2.76E-04	S45	31

Table 6.25 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS1*

$\phi\chi$: $\delta = 1/2$, $m_\phi = 650$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$						
m of χ_m	c_m	m_m	Subprocess	σ (pb)	Relative σ (pb)	Total Number of Events
0	0.10	500.00	S1	1.52E-03	1.38E-02	1520
1	0.13	550.00	S2	2.32E-03	2.12E-02	2328
2	0.15	570.70	S3	2.74E-03	2.50E-02	2746
3	0.16	586.60	S4	3.11E-03	2.83E-02	3111
4	0.17	600.00	S5	3.47E-03	3.16E-02	3472
5	0.18	611.80	S6	3.73E-03	3.40E-02	3736
6	0.19	622.50	S7	3.96E-03	3.61E-02	3972
7	0.20	632.30	S8	4.24E-03	3.86E-02	4251
8	0.21	641.40	S9	4.59E-03	4.18E-02	4598

Table 6.26 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS1*

$\phi\phi$: $\delta = 1/2$, $m_\phi = 650$ GeV, $m_0 = 500$ GeV, $\Delta m = 50$ GeV, $\gamma = 1$, $c_0 = 0.1$					
m of χ_m	n of χ_n	Subprocess	σ (pb)	Relative σ (pb)	Total Number of Events
0	0	S1	7.94E-02	7.23E-01	79538

Table 6.27 *Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS1*

Tables 6.28, 6.29, and 6.30 show the parameter values and cross sections for all subprocesses of production, all subprocesses of production, and the production for Non-minimal Dark Sector Benchmark NMDS6. The NMDS6 benchmark has a Δm value of 20 GeV and a δ value of 0.5. This combination affords the smallest mass spacing among ensemble constituents out of any of the non-minimal dark sectors discussed. The mass spacing diminishes to as low as 6 GeV between the highest possible ensemble constituent and second higher ensemble constituent. The γ value of 2 attempts to strengthen the coupling to prevent direct decays to the lowest ensemble constituent. The NMDS6 benchmark marks a sample that probes the low p_T regime.

$\chi\chi$: $\delta = 1/2$: $m_\phi = 561.27$ GeV, $m_0 = 500$ GeV, $\Delta m = 25$ GeV, $\gamma = 2$, $c_0 = 0.1$									
m of χ_m	n of χ_n	m_m	m_n	c_m	c_n	σ (pb)	Relative σ (pb)	Subprocess	Total number of Events
0	0	500.00	500.00	0.10	0.10	4.19E-06	1.95E-05	S1	3
1	0	525.00	500.00	0.11	0.10	4.70E-06	2.19E-05	S2	3
2	0	535.35	500.00	0.11	0.10	4.97E-06	2.31E-05	S3	3
3	0	543.30	500.00	0.12	0.10	5.15E-06	2.40E-05	S4	3
4	0	550.00	500.00	0.12	0.10	6.63E-06	3.09E-05	S5	4
5	0	555.90	500.00	0.12	0.10	5.65E-06	2.63E-05	S6	3
1	1	525.00	525.00	0.11	0.11	5.13E-06	2.39E-05	S7	3
2	1	535.35	525.00	0.11	0.11	5.62E-06	2.61E-05	S8	3
3	1	543.30	525.00	0.12	0.11	5.61E-06	2.61E-05	S9	3
4	1	550.00	525.00	0.12	0.11	6.11E-06	2.84E-05	S10	4
5	1	555.90	525.00	0.12	0.11	6.13E-06	2.85E-05	S11	4
2	2	535.35	535.35	0.11	0.11	5.95E-06	2.77E-05	S12	4
3	2	543.30	535.35	0.12	0.11	6.34E-06	2.95E-05	S13	4
4	2	550.00	535.35	0.12	0.11	6.53E-06	3.04E-05	S14	4
5	2	555.90	535.35	0.12	0.11	6.46E-06	3.00E-05	S15	4
3	3	543.30	543.30	0.12	0.12	6.12E-06	2.85E-05	S16	4
4	3	550.00	543.30	0.12	0.12	6.51E-06	3.03E-05	S17	4
5	3	555.90	543.30	0.12	0.12	6.85E-06	3.19E-05	S18	4
4	4	550.00	550.00	0.12	0.12	6.52E-06	3.03E-05	S19	4
5	4	555.90	550.00	0.12	0.12	6.73E-06	3.13E-05	S20	4
5	5	555.90	555.90	0.12	0.12	7.17E-06	3.34E-05	S21	4

Table 6.28 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\chi\chi$ production for Non-minimal Dark Sector Benchmark NMDS6*

$\phi\chi$: $\delta = 1/2$: $m_\phi = 561.27$ GeV, $m_0 = 500$ GeV, $\Delta m = 25$ GeV, $\gamma = 2$, $c_0 = 0.1$						
m of χ_m	c_m	m	Subprocess	σ (pb)	Relative σ (pb)	Total Events
500	0.1	0	S1	0.00313	0.0146	1600
525	0.11025	1	S2	0.00351	0.0163	1790
535.35	0.1146	2	S3	0.00371	0.0173	1900
543.3012	0.11807	3	S4	0.00382	0.0178	1950
550	0.121	4	S5	0.00392	0.0183	2010
555.9	0.1236	5	S6	0.00405	0.0188	2070

Table 6.29 *Parameters and Cross Sections (σ) for individual subprocesses (possible decay states) of $\phi\chi$ production for Non-minimal Dark Sector Benchmark NMDS6*

$\phi\phi$: $\delta = 1/2$: $m_\phi = 561.27$ GeV, $m_0 = 500$ GeV, $\Delta m = 25$ GeV, $\gamma = 2$, $c_0 = 0.1$					
m of χ_m	n of χ_n	σ (pb)	Relative σ (pb)	Subprocess	Total Events
0	0	1.93E-01	8.97E-01	S1	98600

Table 6.30 *Parameters and Cross Sections (σ) for $\phi\phi$ production for Non-minimal Dark Sector Benchmark NMDS6*

6.3 Scalar Cascade Parameters and MC Samples

Tables 6.31, 6.32, 6.33, 6.34, and 6.35 show the calculations of masses, couplings, and cross sections of the 2 Body Stable Cascade decay benchmarks, 2 Body Unstable Cascade decay benchmarks, 3 Body Stable Cascade decay benchmarks, 3 Body Stable Chain decay benchmarks with up to 15 scalars, and Body Stable Chain decay benchmarks with up to 25 scalars, respectively. None of the benchmarks shown have been excluded at the LHC as their smaller cross sections have evaded exclusion limits. Some key, general trends are the following. As the number of scalars, ϕ 's, in the model increase, the Standard Model and hidden coupling strengths decrease. As the Standard Model coupling decreases, the cross section decreases. The 3-body Chain Decay has the most scalars, so the coupling is affected greater in these models than in others.

The tables in this section were produced using calculations performed by the author. The general procedure for producing these tables is described as follows. First, the author chose the number of scalars, N , to include in a benchmark. Then, the number of scalars was used to calculate the highest allowed values of Standard Model couplings λ^{SM} and hidden sector couplings λ^{HS} . The author then chose values for λ^{SM} and λ^{HS} . Subsequently, masses of each scalar were calculated. The motivations for these choices of the coupling constants and scalar masses for each class of scalar cascades is described in the passages below. The author used MadGraph to generate LHE files by assigning couplings and masses in UFO models provided

by theorists [57]. The author then hadronized the LHE files with Pythia. The cross section of the scalar cascade benchmark was obtained from the LHE files. (There is no citation available for the code for these UFO models as they were developed specifically for this analysis.)

The benchmarks for the 2 Body Stable Cascade Decay Class were designed by clinging as closely to the allowed limit of the Standard Model couplings, λ^{SM} , as possible. For the 2 Body Stable Cascade Decay Class, this necessarily meant using as close to, but less than, $\frac{4\pi}{N}$ as possible. N is the number of scalars, ϕ 's, in the model. Smaller values of λ^{SM} result in smaller cross sections, so the value of λ^{SM} was maximized in order to have as large a cross section as possible. As can be seen in Table 6.31, the resultant cross sections are small, so this was an important optimization step.

The mass spacing between ϕ 's necessarily was such that the mass of the original scalar is roughly greater than or equal to twice the mass of the subsequent scalar. The hidden sector coupling for the 2 Body Stable Cascade Decay Class was set such that $\lambda^{\text{HS}} \lesssim \frac{4\pi m_\phi}{N^{3/2}}$, where N is the number of scalars [57]. The tables in this section chart out, from left to right, the mass of each scalar in the model starting with the heaviest scalar and proceeding to the lightest scalar. The tables map out the number of stable particles (StaP) and Standard Model particles (SMP) present in each step of the cascade. The coupling of the hidden sector, λ^{HS} , and the coupling of the Standard Model, λ^{SM} , are provided for each benchmark. The cross section (σ) is listed in pb. Table 6.31 shows that models with 8 scalars have extremely small cross sections. Coupled with very small trigger efficiencies, 2 Body Stable Cascade Decay models with 8 scalars are not viable for discovery at 139 fb^{-1} with 13 TeV. The 2 Body Stable Cascade Decay models with 7 scalars, however, are of greater interest, as they afford slightly higher cross sections while still yielding high particle multiplicity with a total of 128 Standard Model particles produced by the end of the decay. In order to reach this high multiplicity while still abiding by the mass relations, the lowest scalar, the stable particle, had to be confined to a small mass. The lowest scalar, in the stable decay classes, is what manifests as missing energy. The goal was to maximize the mass of the lowest scalar in order to maximize the amount of missing energy and, thereby, allow for a robust triggering strategy. (If there is substantial missing energy in an event, the ATLAS missing energy trigger can be used to capture the event.)

However, for these higher multiplicity states, the mass of the lowest scalar necessarily

has to be small. However, as will be seen later in this chapter, this low scalar mass for the lightest scalar allows for an unexpectedly sufficient amount of missing energy for triggering for a fraction of events. Models with 4 or fewer scalars, with small cross sections and using only 139 fb^{-1} of luminosity, are not likely to have substantial discriminating power against background, as there will be only 16 or few Standard Model particles. Those models were, therefore, not considered. Scalar cascades with 5 scalars allow for higher missing energy, but have only 32 Standard Model particles. The kinematic distributions discussed later will show that these models have less spectacular kinematic features. From this 2 Body Stable Cascade Decay Class, benchmarks B10 and B34 were chosen to use in the subsequent truth-level studies. The B10 benchmark was chosen because of the staggering 128 Standard Model particles in the final state that it yields. Additionally, as will be seen later, this benchmark yields an unexpectedly sufficient amount of missing energy for triggering, despite the mass of the lowest scalar being small. This allows the interesting phase space of high particle multiplicity to be picked up by the detector. The B34 benchmark was chosen because, though it only has 32 Standard Model particles in the final state, it has the potential to have significant missing energy. This benchmark was chosen to study whether a benchmark with lower particle multiplicity has any distinctive kinematic features.

2 Body			4	8	16	32	64	128			
Stable			StaP	StaP	StaP	StaP	StaP	StaP			
Cascade	m_{ϕ_1} (GeV)	m_{ϕ_2} (GeV)	8 SMP m_{ϕ_3} (GeV)	16 SMP m_{ϕ_4} (GeV)	32 SMP m_{ϕ_5} (GeV)	64 SMP m_{ϕ_6} (GeV)	128 SMP m_{ϕ_7} (GeV)	256 SMP m_{ϕ_8} (GeV)	λ^{HS}	λ^{SM}	σ (pb)
B1	367	183	91	45	22	10			7	2	2.13E-04
B2	271	135	67	33	16	10			7	2	1.11E-03
B3	255	127	63	31	15	5			3	2	1.52E-03
B4	385	192	95	47	23	11	3		1	1.71	1.20E-04
B5	247	123	61	30	20				7	2.42	2.64E-03
B6	327	163	81	40	30				7	2.42	5.90E-04
B7	247	123	61	50					8	3.05	4.19E-03
B8	282	140	55	40					8	3.05	2.08E-03
B9	191	95	47	23	11	0.5			0.4	2	6.84E-03
B10	293	146	72	35	17	8	0.5		0.3	1.71	5.33E-04
B11	146	72	35	17	8	0.5			0.4	2	2.72E-02
B12	286	142	70	30	14	0.5			0.4	2	8.29E-04
B13	207	103	51	25	12	2			1.4	2	4.52E-03
B14	285	142	70	34	15	1			0.7	2	8.44E-04
B15	212	105	52	25	9	1			0.7	2	3.99E-03
B16	189	94	46	22	10	5			2.5	2	7.22E-03
B17	250	123	59	28	13	3			2	2	1.69E-03
B18	250	124	61	28	12	1			0.7	2	1.69E-03
B19	290	144	70	30	14	0.5			0.4	2	7.71E-04
B20	210	104	50	24	11	1			0.7	2	4.19E-03
B21	266	130	64	30	13	1			0.7	2	1.22E-03
B22	260	128	63	50					8	3.05	3.19E-03
B23	220	108	53	40					8	3.05	7.66E-03
B24	247	122	60	4					4	3.05	4.19E-03
B25	198	98	47	23	11	1			0.7	2	5.69E-03
B26	286	142	70	34	16				7	2.42	1.21E-03
B27	200	47	23	11	1				0.9	2.42	7.91E-03
B28	223	111	55	27	13	6	1		0.5	1.71	2.24E-03
B29	305	145	72	35	16	7	1		0.5	1.71	4.30E-04
B30	300	125	62	30	14	6	1		0.5	1.71	4.70E-04
B31	383	191	95	47	23	11	5	0.5	0.25	1.5	9.46E-05
B32	176	87	43	21	10	4			3	2	1.04E-02
B33	200	47	23	11	1				0.8	2.42	7.91E-03
B34	295	147	73	36	25				8	2.42	1.03E-03
B35	255	127	63	31	25				8	2.42	2.23E-03
B36	287	143	71	35	30				8	2.42	1.19E-03
B37	327	163	81	40	29				8	2.42	5.90E-04
B38	191	95	47	23	11	5	0.5		0.3	1.71	5.00E-03
B39	220	95	47	23	11	5	0.5		0.3	1.71	2.41E-03

Table 6.31 *Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of Standard Model particles (SMP), number of stable particles (StaP), and cross sections (σ) for 2 Body Stable Cascade Benchmarks. The benchmarks chosen for the truth-level analysis are highlighted.*

The benchmarks for the 2 Body Unstable Cascade Decay Class were designed by using values as close to the allowed limit of the Standard Model couplings, λ^{SM} , as possible. The coupling of the 2 Body Unstable Cascade Decay Class was restricted to less than $\frac{4\pi}{N}$, where N is the number of scalars, ϕ 's, in the model. The mass spacing between ϕ 's necessarily was such that the mass of the original scalar is roughly greater than or equal to twice the mass of the subsequent scalar. The hidden sector coupling for the 2 Body Stable Cascade Decay Class was set such that $\lambda^{\text{HS}} \lesssim \frac{4\pi m_\phi}{N^{3/2}}$, where N is the number of scalars. Table 6.32 maps out the number of Standard Model particles (SMP) present at each step of the cascade. Models in the 2 Body Unstable Cascade Decay Class with 5 scalars allow for larger cross sections than 6

scalar or 7 scalar models while still retaining an unusually high particle multiplicity of 64 Standard Model particles. However, the mass of the highest scalar is too low to garner the necessary kinematics for effective triggering. Models with 7 scalars have too small cross sections for any substantial yield when pairing with trigger efficiencies. Models with 6 scalars allow for 128 Standard Model particles to be generated. Depending on the mass, the cross sections are small, but not entirely unviable. Trigger efficiencies lessen the yield. The benchmark B6 from the 2 Body Unstable Cascade Decay class is taken for discussion in the subsequent truth-level analysis due to its high particle multiplicity of 128 Standard Model particles in the final state and potentially viable cross section.

2 Body Unstable	m_{ϕ_1} (GeV)	8 SMP m_{ϕ_2} (GeV)	16 SMP m_{ϕ_3} (GeV)	32 SMP m_{ϕ_4} (GeV)	64 SMP m_{ϕ_5} (GeV)	128 SMP m_{ϕ_6} (GeV)	256 SMP m_{ϕ_7} (GeV)	λ^{HS}	λ^{SM}	σ (pb)
B1	259	127	63	31	20			8	2.4	2.02E-03
B2	247	123	61	50				8	3.05	4.19E-03
B3	175	87	43	21	10			8	2.4	1.54E-02
B4	127	63	31	15	5			4	2.4	8.10E-02
B5	255	127	63	31	15	5		3	2	1.52E-03
B6	223	111	55	27	13	4		2.5	2	3.07E-03
B7	143	71	35	17	8			7	2.4	4.36E-02
B8	127	63	31	20				8	3.05	1.31E-01
B9	167	83	41	20				8	3.05	3.16E-02
B10	225	112	50	30				8	3.05	6.82E-03
B11	200	90	40	10				8	3.05	1.26E-02
B12	181	90	18	5				5	3.05	2.10E-02
B13	200	70	28	5				5	3.05	1.26E-02
B14	242	120	58	28	5			4	2.4	2.88E-03
B15	350	170	80	36	16	4		3	2	2.78E-04
B16	351	175	87	43	21	10		7	2	2.73E-04
B17	321	160	79	39	19	9	4	2	1.7	3.22E-04
B18	359	179	79	39	19	9	4	2	1.7	1.75E-04
B19	447	223	111	55	27	13	6	2.5	1.7	5.03E-05

Table 6.32 *Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of Standard Model particles (SMP), and cross sections (σ) for 2 Body Unstable Cascade Benchmarks. The benchmark chosen for the truth-level analysis is highlighted.*

The benchmarks for the 3 Body Stable Cascade Decay Class were designed by using values as close to the allowed limit of the Standard Model couplings, λ^{SM} , as possible. The coupling of the 3 Body Stable Cascade Decay Class was restricted to less than $\frac{4\pi}{N}$, where N is the number of scalars, ϕ 's, in the model. The mass spacing between ϕ 's necessarily was such that the mass of the original scalar is roughly greater than or equal to three times the mass of the subsequent scalar. The hidden sector coupling for the 3 Body Stable Cascade Decay Class was set such that $\lambda^{\text{HS}} = \lesssim \frac{16\pi^2}{N^2}$, where N is the number of scalars. Table 6.33 maps out the number of Standard Model particles (SMP) and stable particles (StaP) present at each step of the cascade. Models in the 3 Body Stable Cascade Decay Class with 6 scalars allow for extremely high particle

multiplicity, totalling a staggering 324 Standard Model particles. However, the cross sections are too low to be viable for a search at 13 TeV with 139 fb^{-1} . Models with 4 scalars allow for lightest scalars to have higher masses, thereby allowing for missing energy. However, the particle multiplicity for models with 4 scalars is confined to 36 Standard Model particles. Models with 5 scalars, however, allow for 108 Standard Model particles by the end of the decay. The lightest scalars necessarily needed to have small masses in order to allow for such high particle multiplicity. However, as seen in the 2 Body Stable Cascade Decay, this low scalar mass for the lightest scalar still allows for an unexpectedly sufficient amount of missing energy for triggering for a fraction of events. Benchmarks B26 and B21 are taken from the 3 Body Stable Cascade Decay class for the subsequent truth-level analysis discussion in order to highlight the kinematics of a mass point with a lightest scalar bearing considerable mass and with lower particle multiplicity and a mass point with a lightest scalar having very little mass and yielding high particle multiplicity, respectively.

3 Body Stable Benchmark	m_{ϕ_1} (GeV)	m_{ϕ_2} (GeV)	6 StaP 12 SMP m_{ϕ_3} (GeV)	18 StaP 36 SMP m_{ϕ_4} (GeV)	54 StaP 108 SMP m_{ϕ_5} (GeV)	162 StaP 324 SMP m_{ϕ_6} (GeV)	λ^{HS}	λ^{SM}	σ (pb)
B1	362	120	38	12	2		6	2.4	3.32E-04
B2	330	109	36	11	0.5		6	2.4	5.52E-04
B3	298	98	32	10.6	0.5		6	2.4	9.59E-04
B4	308	102	33	20			8	3	1.25E-03
B5	360	116	38	25			8	3	5.35E-04
B6	215	70	23	10			8	3	8.35E-03
B7	200	60	18	5			8	3	1.22E-02
B8	165	48	14	2			8	3	3.26E-02
B9	260	80	20	5			8	3	3.09E-03
B10	229	76	25	7	2		6	2.4	3.85E-03
B11	152	50	16	5	1		6	2.4	3.18E-02
B12	283	94	31	10	3		6	2.4	1.26E-03
B13	151	50	15	4			8	3	5.14E-02
B14	220	68	22	5			8	3	7.41E-03
B15	337	112	37	12	2		6	2.4	4.92E-04
B16	294	94	31	10	1		6	2.4	1.03E-03
B17	256	85	28	9	0.5		6	2.4	2.15E-03
B18	200	50	16	5			8	3	1.22E-02
B19	244	77	25	7	2		6	2.4	2.76E-03
B20	189	62	20	5	1		6	2.4	1.04E-02
B21	310	103	34	11	1		6	2.4	7.75E-04
B22	320	106	35	30			8	3	1.02E-03
B23	190	45	13	1			8	3	1.58E-02
B24	160	50	16	5	1		6	2.4	2.44E-02
B25	368	122	40	33			8	3	4.73E-04
B26	332	110	36	29			8	3	8.35E-04
B27	420	139	46	40			8	3	2.24E-04
B28	452	150	49	16	5	1	4	2	6.52E-05
B29	251	83	27	20			8	3	3.72E-03

Table 6.33 *Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of Standard Model particles (SMP), number of stable particles (StaP), and cross sections (σ) for 3 Body Stable Cascade Benchmarks. The benchmarks chosen for the truth-level analysis are highlighted.*

Both Tables 6.34 and 6.35 display benchmarks for the 3 Body Stable Chain Decay

Class. The benchmarks for the 3 Body Stable Chain Decay were designed by using values as close to the allowed limit of the Standard Model couplings, λ^{SM} , as possible. The coupling of the 3 Body Stable Chain Decay class was restricted to less than $\frac{4\pi}{N}$, where N is the number of scalars, ϕ 's, in the model. The mass spacing between ϕ 's necessarily was such that the mass of the original scalar is greater than the sum of the mass of the subsequent scalar and the masses of two Standard Model particles in order to allow for their generation. There is no hidden sector coupling for the 3 Body Stable Chain Decay Class, so $\lambda^{\text{HS}} = 0$. Tables 6.34 and 6.35 map out the number of Standard Model particles (SMP) and stable particles (StaP) present at each step of the cascade. Table 6.34 describes benchmarks containing less than 15 scalars, while Table 6.35 describes benchmarks containing less than 25 scalars. The 3-body Chain Decay Class has the most scalars out of any of the classes discussed so far, so the coupling is affected greater in these models than in others. Smaller couplings lend to smaller cross sections. In order to achieve the highest particular multiplicity without immediately going into territory where cross sections are not viable, the lightest scalar has to be given a smaller mass. The stable chain decays, unlike the stable cascade decays for 2 Body or 3 Body decays which allowed for even 50 stable particles, have only 2 stable scalars at the end of the decay. This means that, if the lightest scalar has a small mass, there is not a multiplicative factor of say 50 that makes the total expected missing energy 50 times larger than the lightest scalar. Instead, the expected missing energy is only two times the lightest scalar mass. Therefore, the tables also contain examples of benchmarks with lightest scalars containing high masses and then, subsequently, containing a smaller number of scalars in the whole benchmark. Benchmark B1 from Table 6.35 is taken for further discussion in the subsequent truth-level analysis due to its high particle multiplicity of 64 Standard Model particles and potentially viable cross section.

3 Body				4	8	12	16	20	24	28	32	36	40
Chain	λ^{SM}	σ (pb)	m_{ϕ_1} (GeV)	SMP m_{ϕ_2} (GeV)	SMP m_{ϕ_3} (GeV)	SMP m_{ϕ_4} (GeV)	SMP m_{ϕ_5} (GeV)	SMP m_{ϕ_6} (GeV)	SMP m_{ϕ_7} (GeV)	SMP m_{ϕ_8} (GeV)	SMP m_{ϕ_9} (GeV)	SMP $m_{\phi_{10}}$ (GeV)	SMP $m_{\phi_{11}}$ (GeV)
B1	1	9.50E-04	214	203	192	181	170	158	145	132	120	85	50
B2	1.53	1.40E-03	234	222	210	198	186	174	162	150			
B3	1.53	8.72E-04	256	243	230	218	206	194	182	170			
B4	1.53	8.54E-04	257	246	235	224	213	202	191	180			
B5	0.95	4.94E-04	238	226	213	201	190	177	165	150	138	125	112
B6	0.95	6.03E-04	229	218	207	195	180	165	150	138	125	113	94
B7	1.3	2.97E-03	190	177	165	150	138	125	112	100	80		
B8	1.2	1.37E-03	214	203	192	181	170	158	145	132	120	100	
B9	1.2	1.17E-04	340	330	320	310	300	290	280	270	260	250	
B10	1.7	2.35E-04	340	330	320	310	300	290	280				
B11	0.89	1.97E-05	420	400	390	380	370	360	350	340	330	324	313
B12	1.2	7.35E-05	370	360	350	340	330	324	313	302	291	280	
B13	0.89	2.72E-04	260	242	228	213	198	186	175	162	150	100	60
B14	0.95	3.10E-04	260	255	250	245	240	235	230	225	220	215	210
B15	0.95	4.94E-04	238	226	213	201	190	177	165	150	138	125	112
B16	0.95	6.03E-04	229	218	207	195	180	165	150	138	125	113	94
B17	1.3	2.97E-03	190	177	165	150	138	125	112	100	40		
B18	1.2	1.37E-03	214	203	192	181	170	158	145	132	120	116	
B19	1.7	2.51E-04	336	330	324	318	312	306	300				
B20	1.7	2.35E-04	340	335	330	325	320	315	310				
B21	1.7	6.29E-05	430	425	420	415	410	405	400				
B22	1.3	6.85E-04	252	243	234	225	216	207	198	189	180		
			44 SMP	48 SMP	52 SMP								
Continued			$m_{\phi_{12}}$ (GeV)	$m_{\phi_{13}}$ (GeV)	$m_{\phi_{14}}$ (GeV)								
B1			10										
B5			100	80									
B6			80	60									
B11			302	291	280								
B13			40	20	5								
B14			205	200									
B15			100	30									
B16			80	10									

Table 6.34 *Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), number of stable particles (StaP), and cross sections (σ) for 3 Body Stable Chain Benchmarks up to 15 Scalars. Each benchmark yields 2 stable scalars.*

Chain Decay 3 Body				4 SMP	8 SMP	12 SMP	16 SMP	20 SMP	24 SMP	28 SMP	32 SMP
Benchmark	λ^{SM}	σ (pb)	m_{ϕ_1} (GeV)	m_{ϕ_2} (GeV)	m_{ϕ_3} (GeV)	m_{ϕ_4} (GeV)	m_{ϕ_5} (GeV)	m_{ϕ_6} (GeV)	m_{ϕ_7} (GeV)	m_{ϕ_8} (GeV)	m_{ϕ_9} (GeV)
B1	0.73	1.20E-03	181	170	159	148	137	126	115	104	93
B2	0.45	1.83E-04	216	207	198	189	180	171	162	153	144
B3	0.45	8.55E-05	250	240	230	220	210	200	190	180	170
B4	0.6	8.11E-04	181	171	161	151	141	131	121	111	101
B5	0.73	3.48E-04	230	225	220	215	210	205	200	195	190
B6	0.73	2.19E-03	161	151	141	131	121	111	101	91	81
B7	0.59	1.90E-04	238	226	214	202	190	179	167	155	143
B8	0.45	1.01E-04	242	232	222	212	202	192	182	172	162
B9	0.45	3.69E-03	121	116	111	106	101	96	91	86	81
B10	0.57	5.00E-04	195	190	185	180	175	170	165	160	155
			36 SMP	40 SMP	44 SMP	48 SMP	52 SMP	56 SMP	60 SMP	64 SMP	
Continued			$m_{\phi_{10}}$ (GeV)	$m_{\phi_{11}}$ (GeV)	$m_{\phi_{12}}$ (GeV)	$m_{\phi_{13}}$ (GeV)	$m_{\phi_{14}}$ (GeV)	$m_{\phi_{15}}$ (GeV)	$m_{\phi_{16}}$ (GeV)	$m_{\phi_{17}}$ (GeV)	
B1			82	71	60	49	38	27	16	5	
B2			135	126	117	108	99	90	81	72	
B3			160	150	140	130	120	110	100	90	
B4			91	81	71	61	51	41	31	21	
B5			185	180	175	170	165	160	155	150	
B6			71	61	51	41	31	21	11	1	
B7			131	119	107	95	83	71	59	47	
B8			152	142	132	122	112	102	92	82	
B9			76	71	66	61	56	51	46	41	
B10			150	145	140	135	130	125	120	115	
			68 SMP	72 SMP	76 SMP	80 SMP	84 SMP	88 SMP	92 SMP	96 SMP	
Continued			$m_{\phi_{18}}$ (GeV)	$m_{\phi_{19}}$ (GeV)	$m_{\phi_{20}}$ (GeV)	$m_{\phi_{21}}$ (GeV)	$m_{\phi_{22}}$ (GeV)	$m_{\phi_{23}}$ (GeV)	$m_{\phi_{24}}$ (GeV)	$m_{\phi_{25}}$ (GeV)	
B2			63	54	45	36	27	18	9	0.5	
B3			80	70	60	50	40	30	20	10	
B4			11	1							
B7			35	23	12	1					
B8			72	62	52	42	32	22	12	2	
B9			36	31	26	21	16	11	6	1	
B10			110	105	100						

Table 6.35 *Masses, hidden sector couplings (λ^{HS}), Standard Model couplings (λ^{SM}), decay particle multiplicity, and cross sections (σ) for 3 Body Stable Chain Benchmarks up to 25 Scalars. Each benchmark yields 2 stable scalars. The benchmark chosen for the truth-level analysis is highlighted.*

The scalar cascade benchmarks that appear in the truth-level analysis in this thesis are shown in Figure 6.36, where they are provided simpler names that will be used throughout the remainder of this thesis. For example, Scalar Cascade 2 Body Stable Cascade Benchmark B10 from Table 6.31 will be referred to as SC1.

Signal Sample	Signal Sample Name
Scalar Cascade 2 Body Stable Cascade Benchmark B10 from Table 6.31	SC1
Scalar Cascade 3 Body Stable Cascade Benchmark B21 from Table 6.33	SC2
Scalar Cascade 3 Body Stable Chain Benchmark B1 from Table 6.35	SC3
Scalar Cascade 2 Body Stable Cascade Benchmark B34 from Table 6.31	SC4
Scalar Cascade Unstable Cascade Benchmark B6 from Table 6.32	SC5
Scalar Cascade 3 Body Stable Cascade Benchmark B26 from Table 6.33	SC6

Table 6.36 *Scalar Cascade Signal Samples Names*

6.4 Summary

Though no experiment has ever searched for noncommutative black holes, scalar cascades, or non-minimal dark sectors, this chapter took into account the model-independent limits from other ATLAS searches in order to determine the unexcluded and potentially reachable phase space for each model and then create a set of benchmarks for each model. Six noncommutative black hole benchmarks were designed. These were designed to both probe the effect that varying the noncommutative scale has on the transverse momentum spectrum of the model, as well as study the effect that varying the minimum black hole mass has on the kinematics of the model. Six non-minimal dark sector benchmarks were designed, where 2 benchmarks were designed to target models with a moderate transverse momentum spectrum, 3 models were designed to target phase space characterized by low transverse momentum, and 1 model was designed to target a phase space characterized by high transverse momentum. For the scalar cascade model, 6 benchmarks were designed in order to study the effect that particle multiplicity has on the kinematics on the model and in order to study the effect that the mass of the lightest stable scalar has on the amount of missing energy produced. The next chapter will study the characteristics of the selected benchmarks in more detail, providing studies that then define the entire analysis strategy.

Chapter 7

A jet-substructure-based analysis strategy and a novel jet-reconstruction algorithm

This chapter presents 2D representations (event displays) of what would appear in the ATLAS detector if a proton-proton collision generated a scalar cascade, noncommutative black hole or non-minimal dark sector. These provide depictions of how final state particles are spread throughout the detector and give insight into the number of jets each exotic model makes. The event displays show that each exotic model exhibits jet substructure. This chapter presents studies that demonstrate that the thresholds imposed by standard ATLAS reconstruction on large R jets are often insufficient to capture the exotic processes, and details the design and optimization of a novel large R jet reconstruction algorithm as a remedy.

7.1 Event Displays

The optimal analysis strategy can be derived by studying the 2D representations of the collisions, called event displays. The event displays in this chapter were produced by code developed by the author. They contain truth-level information. As described in the previous chapter, the author used Monte Carlo generators to simulate the production and decay of the scalar cascades, noncommutative black holes, and non-minimal dark sectors. These generators outputted LHE files, defined in the previous

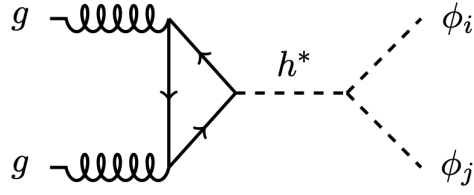
chapter, which the author then transformed, via standard ATLAS commands, into an ATLAS data format suitable for analyzing truth information. The event displays were then produced by running the author’s event display code over these files.

The next chapter will outline the basic cuts implemented for the analysis. The one cut that needs introduction in this chapter is the 200 GeV E_T^{miss} cut. A 200 GeV E_T^{miss} cut is applied when creating every event display. This is motivated by calculations performed by the author with code developed by the author to determine the efficiencies for various Run 2 triggers of each benchmark for scalar cascades, noncommutative black holes, and non-minimal dark sectors and, thereby, determine the optimal trigger strategy for each analysis. The E_T^{miss} trigger provided the best trigger efficiencies for each model. More details on the triggers used in the analysis will be provided in the next chapter. The trigger efficiencies for each signal sample appear in the Appendix.

This section will systematically go through the event displays for four of the six scalar cascade benchmarks. Then, the event displays for the two noncommutative black hole benchmarks and the three non-minimal dark sector benchmarks that are used for the reconstructed analysis will be shown. Though studied by the author, the event displays for the Scalar Cascade Benchmark SC5 and the Scalar Cascade Benchmark SC6 are not included here as these benchmarks pass the E_T^{miss} trigger at very low rates. The kinematics of the Scalar Cascade Benchmark SC5 and the Scalar Cascade Benchmark SC6 are included in the kinematics distributions, however.

The event displays illustrate the distribution of final state particles in the η - ϕ plane for a single collision event for a given model. Though each event display is a single event, hundreds of event displays were studied for each of selected benchmarks. Representative event displays are included in this chapter. Final state truth particles are tagged as originating from the decay of a given exotic process or from initial state radiation (ISR). ISR is the radiation of gluons before or after the hard scattering process of the collision.

As a reminder, all scalar cascade classes studied in this thesis involve the pair production of two new scalars ϕ_i and ϕ_j , as seen in Figure 7.1.



(a) Pair production $\sigma \sim (\lambda^{\text{SM}})^2$

Figure 7.1 Scalar cascade generation via pair production. Figure taken from [3]

All decay products from the scalar cascades process stem from either ϕ_i or ϕ_j . The event displays for scalar cascades in this chapter identify the final state particles as coming from ϕ_i or ϕ_j of the scalar cascade process. Figure 7.2 shows the legend for all event displays describing scalar cascade events.

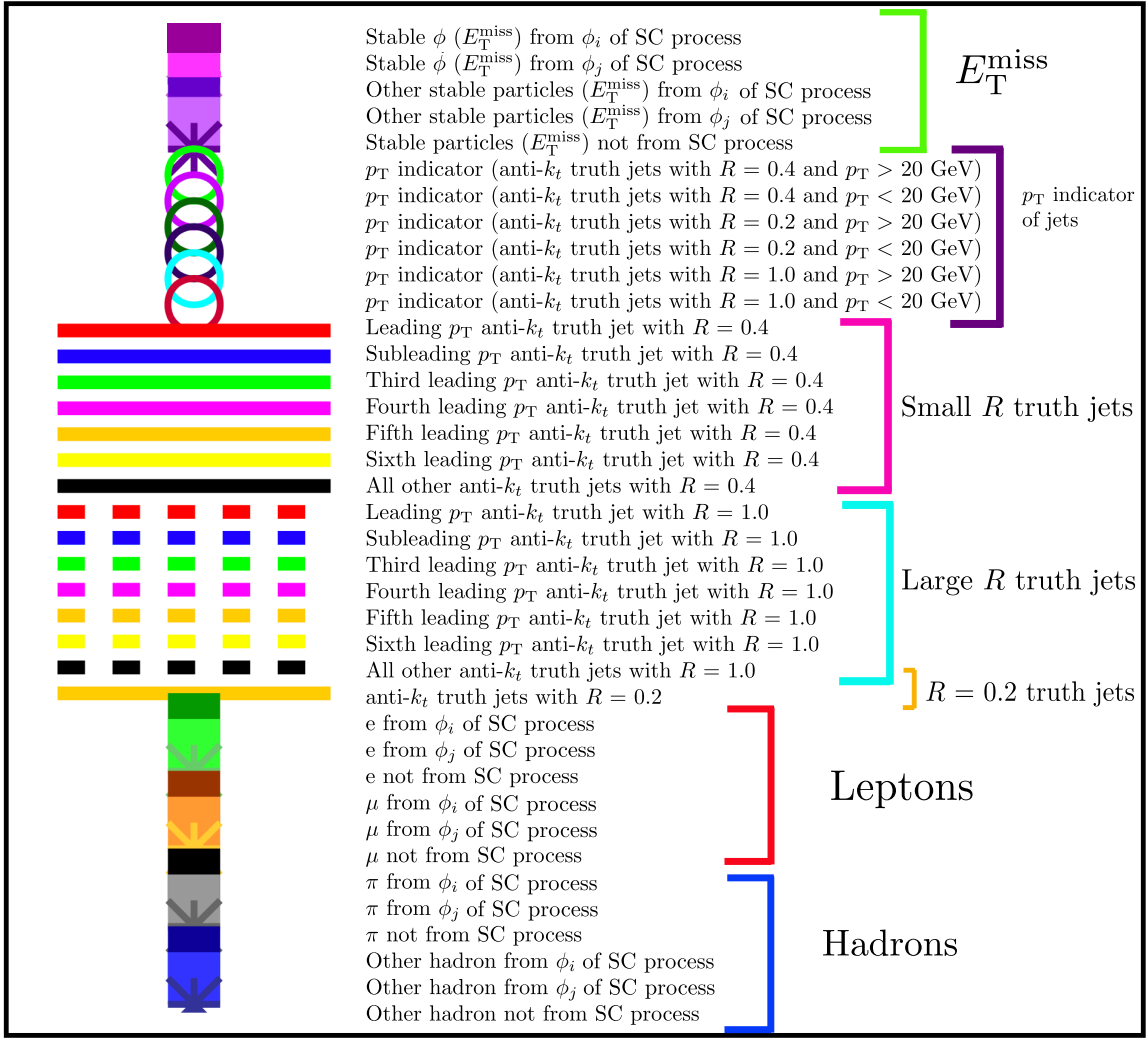


Figure 7.2 Legend for all Scalar Cascade Event Displays. The final state particles are tagged as coming from ϕ_i or ϕ_j of the scalar cascade model or as not coming from the scalar cascade process. The legend includes a breakdown of seven main categories: E_T^{miss} , leptons, hadrons, anti- k_t truth jets of $R = 0.2$, $R = 0.4$, and $R = 1.0$, and p_T indicators (the smallest colored circles in the display) that indicate whether the truth jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS. The anti- k_t truth jets are additionally color coded to indicate the jet with the highest p_T (leading jet), the jet with the second highest p_T (subleading jet), etc.

The event display legend identifies the types of final state particles produced and tags them as coming from either ϕ_i or ϕ_j of the scalar cascade model or as not coming from the scalar cascade model. The E_T^{miss} is sectioned into five different categories: the lightest stable scalar that originates from ϕ_i or ϕ_j of the scalar cascade model, other stable particles coming from ϕ_i or ϕ_j of the scalar cascade model, or stable particles not originating from the scalar cascade process. Anti- k_t truth jets of three

different radii are studied, with the leading through sixth leading jet color coded for anti- k_t truth jets of $R = 0.4$ (small R truth jets) and anti- k_t truth jets of $R = 1.0$ (large R truth jets). The p_T of each anti- k_t jet is indicated by a colored circle that states whether the jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS.

Figure 7.3 shows the event display for the Scalar Cascade Benchmark SC1 defined in Table 6.36.

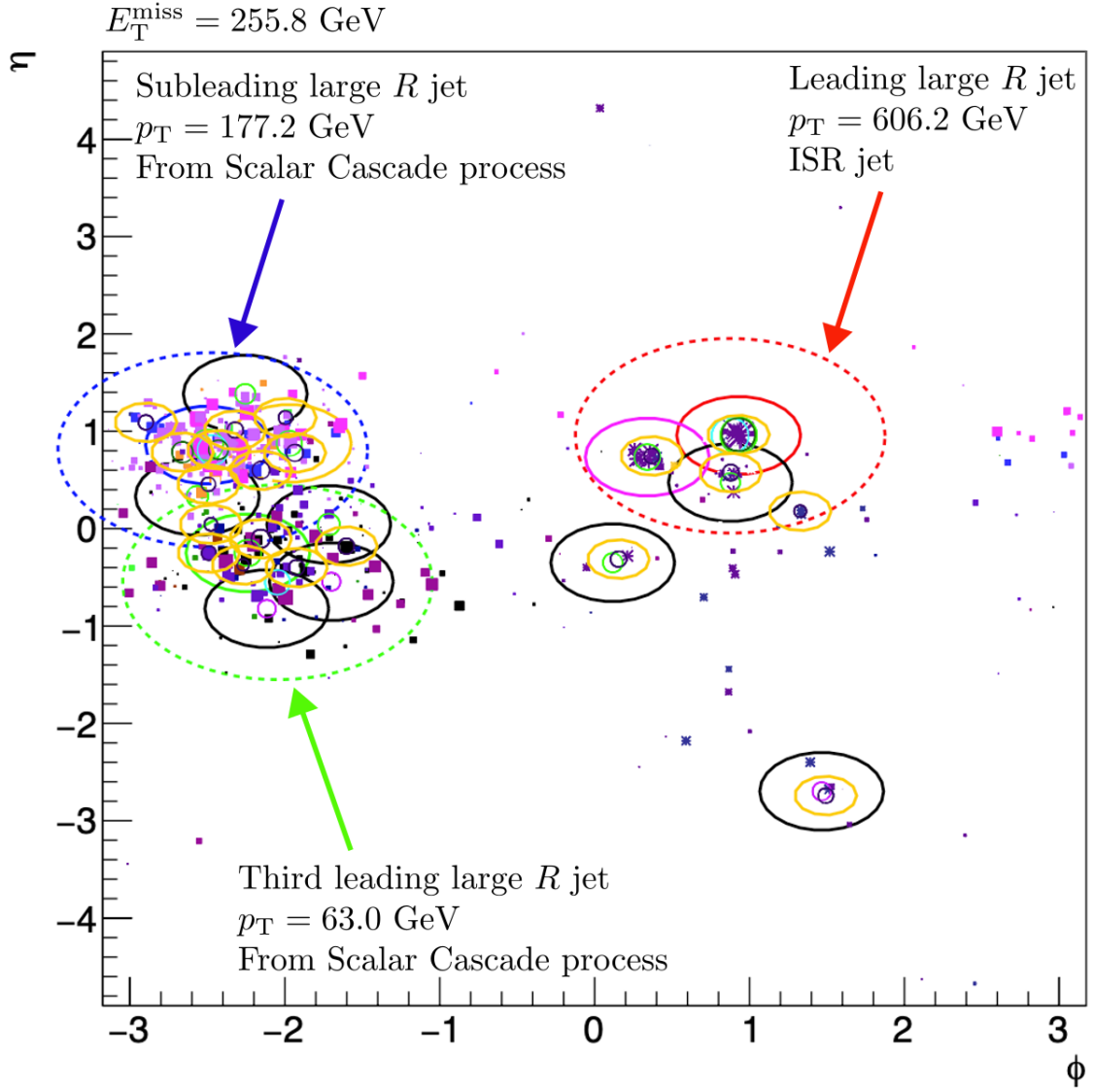


Figure 7.3 *Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC1 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.*

For the Scalar Cascade Benchmark SC1 (which contains a total of 128 standard model particles), the leading (highest p_T) large R jet with radius $R = 1.0$, drawn as the red ellipse, is shown as containing particles from initial state radiation (indicated by the asterisks), not the scalar cascade process. The second leading large R jet, the blue ellipse, contains the final state particles from one ϕ of the scalar cascades process, while the third leading large R jet, the green ellipse, contains the final state

particles from the other ϕ of the scalar cascades process. The large R jet from initial state radiation has a three prong structure, containing three small R jets. The two large R jets containing the exotics process both exhibit multi-prong structures, with the second leading large R jet being 4-pronged and the third leading large R jet being 2-pronged. At first glance, the third leading large R jet appears to be 4-pronged as well. However, due to two of the four small R jets in the third leading R jet being below the 20 GeV p_T threshold that ATLAS uses for small R jets, as indicated by the small magenta circles that serve as p_T indicators, the third leading large R jet is only 2-pronged. All large R jet in the event, therefore, exhibit jet substructure. As indicated by the annotated large R jet p_T values on the event display, the two large R jets containing the exotics process are both below the 200 GeV threshold that ATLAS requires for large R jet calibration. These two large R jets would not be usable if reliant on standard ATLAS large R jet reconstruction. This means that nearly all of the exotic process would be lost from the event. (As a reminder, the large R jets depicted in the event displays in this chapter are large R truth jets.) These two large R jets with radius $R = 1.0$ are very close together, even overlapping. There is a strikingly large amount of particles grouped together.

Figure 7.4 shows another event display for the Scalar Cascade Benchmark SC1, this time with there being only two large R jets: a leading large R jet coming from initial state radiation and a second leading large R jet containing nearly all of the decay products from the exotics process. While each event display represents a single event, large-scale studies at the end of the chapter for Scalar Cascade Benchmark SC1 show that the leading large R jet is predominantly from initial state radiation.

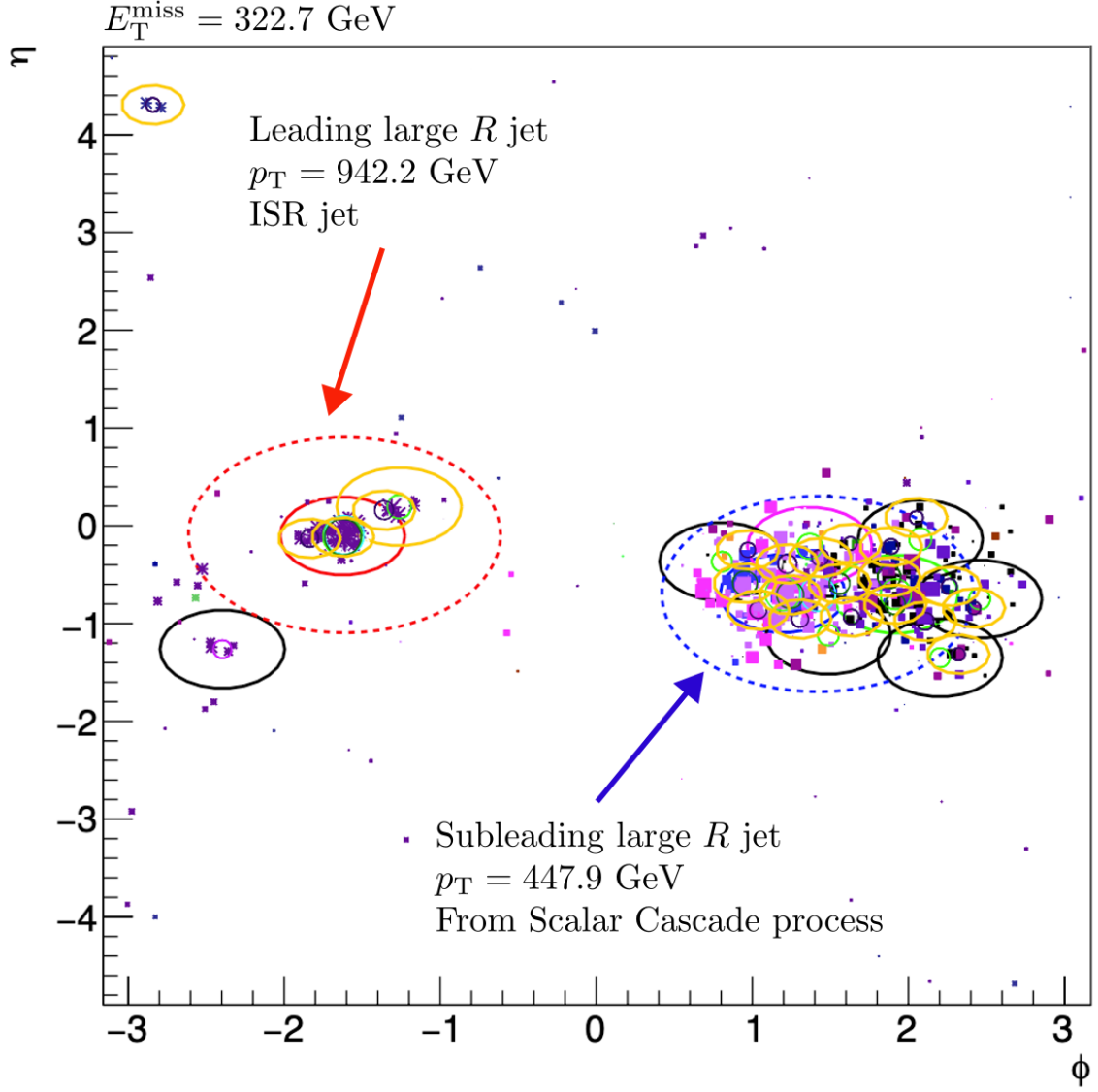


Figure 7.4 *Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC1 (defined in Table 6.36) for a single event. A $200 \text{ GeV } E_T^{\text{miss}}$ cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.*

The event topology yielded by the Scalar Cascade Benchmark SC1 is mirrored in the event displays of the Scalar Cascade Benchmark SC2 (defined in Table 6.36 and which contains 108 standard model particles), shown in Figure 7.5 and Figure 7.6, where the leading large R jet comes from initial state radiation and the exotics process is contained in the second leading and third leading large R jet (grouped closely together and overlapping) in Figure 7.5 and almost entirely in the second

leading large R jet in Figure 7.6. The large R jets containing the exotic process again fail the standard ATLAS large R reconstruction technique and exhibit both a strikingly large amount of particles grouped closely together and pronounced jet substructure.

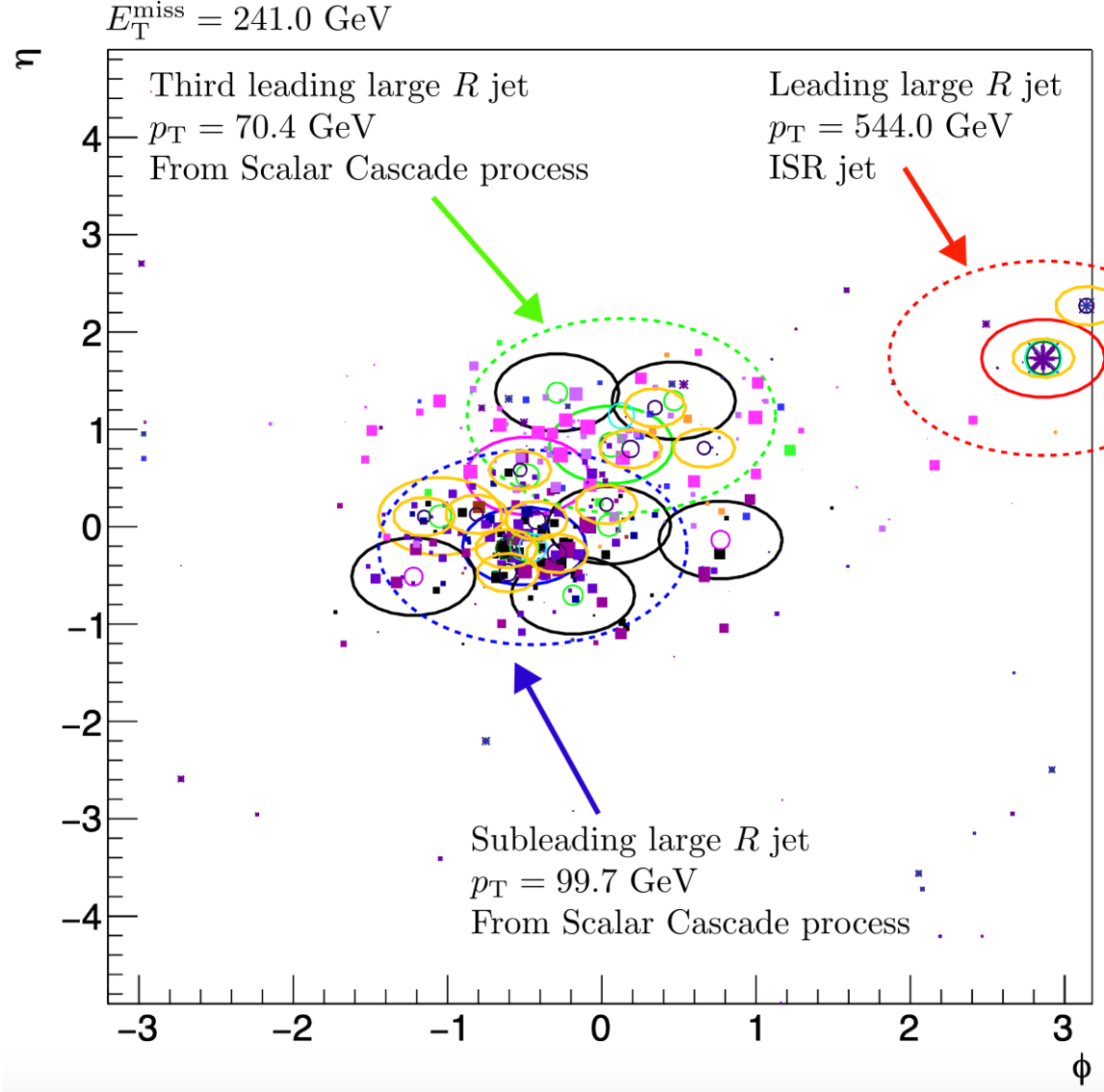


Figure 7.5 *Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC2 (defined in Table 6.36) for a single event. A $200 \text{ GeV } E_T^{\text{miss}}$ cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.*

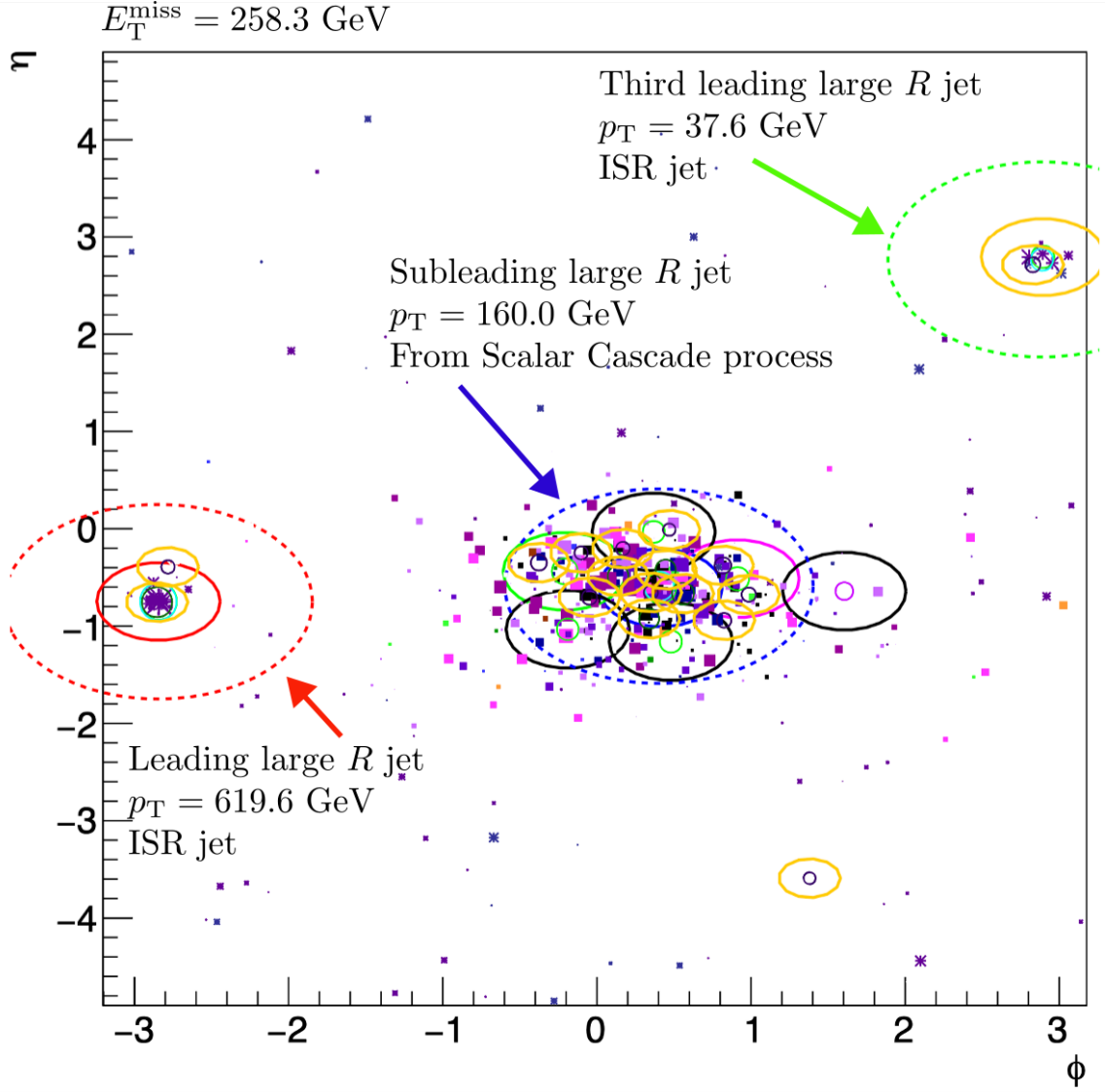


Figure 7.6 *Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC2 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.*

Jet substructure is observed in Figure 7.7 in the event display of the Scalar Cascade Benchmark SC3 (defined in Table 6.36 and which contains 64 Standard model particles), where the leading R jet stems from initial state radiation and the exotic process is contained in the second leading large R jet. This model is less striking than the Scalar Cascade Benchmark SC1 and the Scalar Cascade Benchmark SC2 in terms of sheer particle multiplicity.

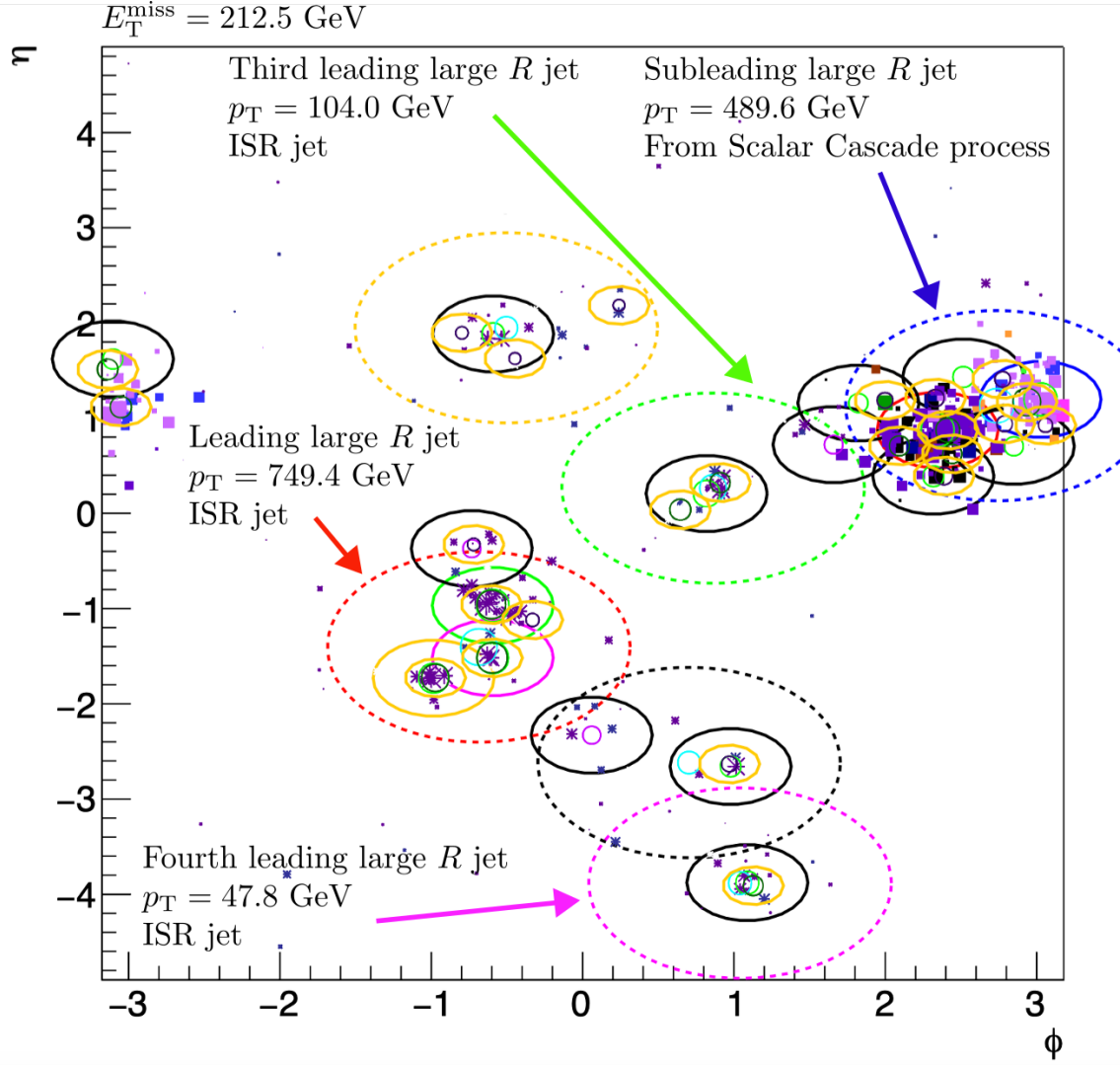


Figure 7.7 *Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC3 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.*

The Scalar Cascade Benchmark SC4 (defined in Table 6.36), with only 32 stable particles produced from the decay, exhibits jet substructure but not as striking particle multiplicity as the scalar cascade benchmarks mentioned above. Again, all but the leading large R jet fails ATLAS standard large R jet construction. The particles from the scalar cascade process are contained in the second and third leading large R jet, which are 2-pronged and 3-pronged, respectively.

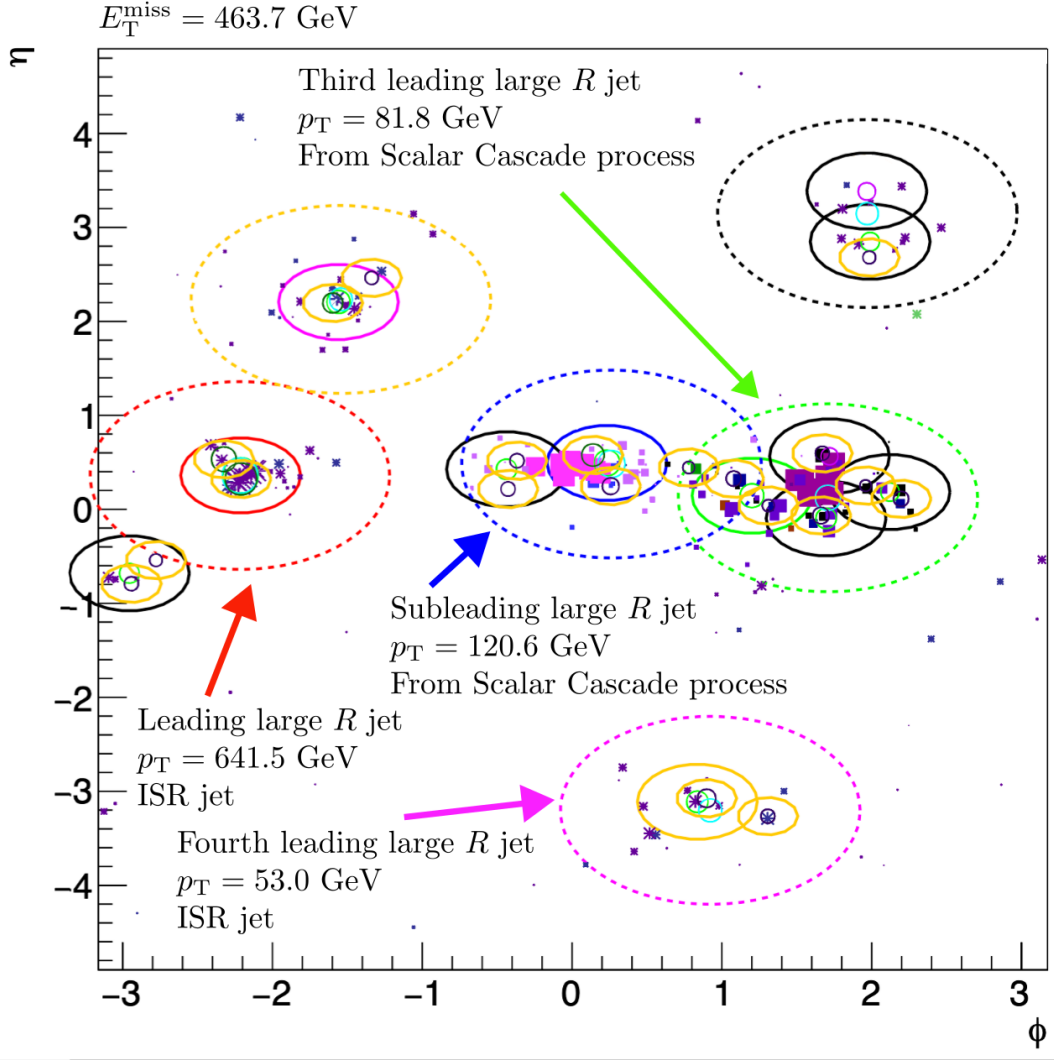


Figure 7.8 *Distribution of final state particles in the η - ϕ plane for Scalar Cascade Benchmark SC4 (defined in Table 6.36) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from ϕ_i or ϕ_j of the scalar cascade model. The corresponding legend for this event display is found in Figure 7.2.*

The next set of event displays are for the non-minimal dark sector models. As described earlier, there are three different production channels for non-minimal dark sectors. As a reminder, Figure 7.9 shows the $\phi\phi$ production process of non-minimal dark sectors. The decay products of one of the two ϕ 's and the decay products of the other ϕ will be referred to as having originated from “branch 1” and “branch 2”.

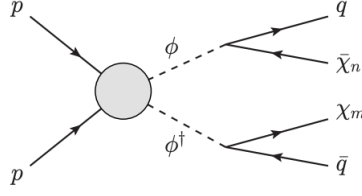


Figure 7.9 $\phi\phi$ Production Channel of Non-minimal Dark Sectors

Figure 7.10 shows the $\chi\phi$ production process of non-minimal dark sectors. The decay products of the χ and the decay products of the ϕ will be referred to as having originated from “branch 1” and “branch 2”.

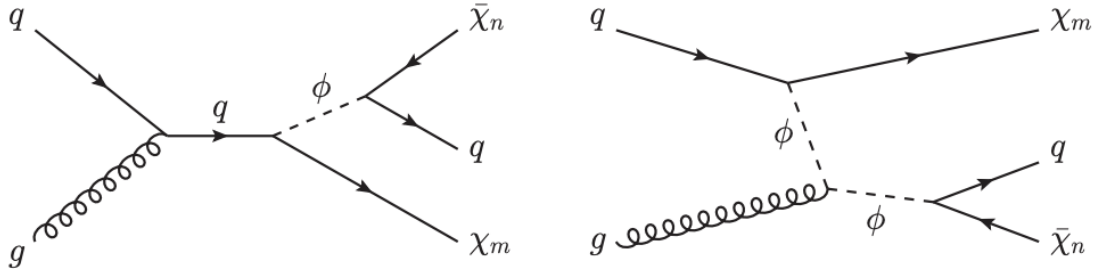


Figure 7.10 $\chi\phi$ Production Channel of Non-minimal Dark Sectors

Figure 7.11 shows the $\chi\chi$ production process of non-minimal dark sectors. The decay products of one of the χ 's and the decay products of the other χ will be referred to as having originated from “branch 1” and “branch 2”.

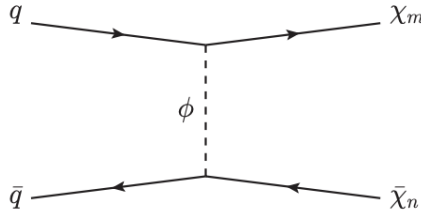


Figure 7.11 $\chi\chi$ Production Channel of Non-minimal Dark Sectors

Figure 7.12 shows the event display legend for final state particles coming from branch 1 or branch 2 for the non-minimal dark sector model.

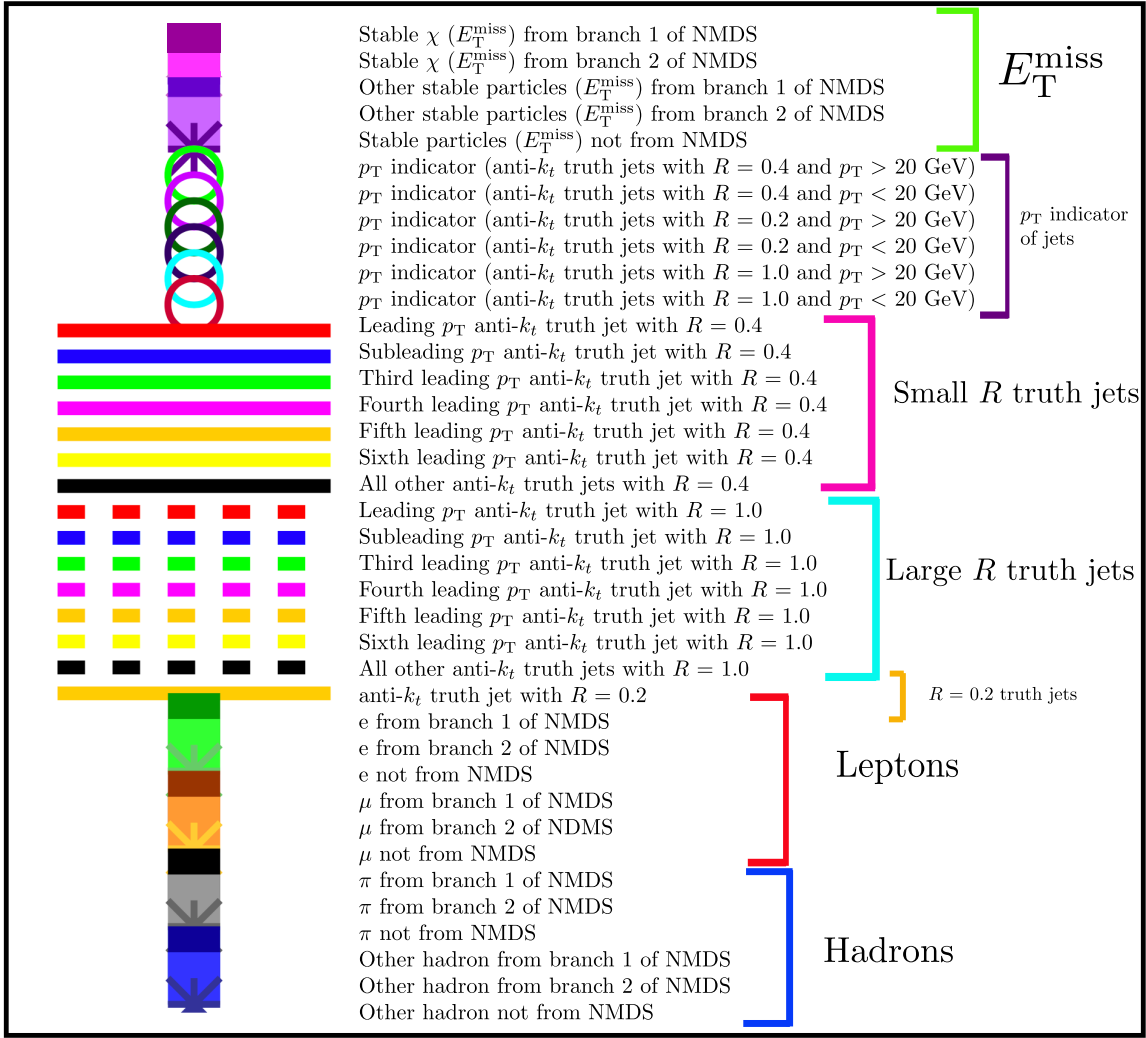


Figure 7.12 Legend for all NMDS Event Displays. The final state particles are tagged as coming from branch 1 or branch 2 of the NDMS model or as not coming from the NMDS process. The legend includes a breakdown of seven main categories: E_T^{miss} , leptons, hadrons, anti- k_t truth jets of $R = 0.2$, $R = 0.4$, and $R = 1.0$, and p_T indicators (the smallest colored circles in the display) that indicate whether the truth jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS. The anti- k_t truth jets are additionally color coded to indicate the jet with the highest p_T (leading jet), the jet with the second highest p_T (subleading jet), etc.

The event display legend identifies the types of final state particles produced and tags them as coming from either branch 1 or branch 2 of the non-minimal dark sector process or as not coming from the non-minimal dark sector process. The E_T^{miss} is sectioned into five different categories: the stable χ that originates from branch 1 or branch 2 of the non-minimal dark sector process, other stable particles coming from branch 1 or branch 2 of the non-minimal dark sector model, or stable particles

not originating from the non-minimal dark sector process. Anti- k_t truth jets of three different radii are studied, with the leading through sixth leading jet color coded for anti- k_t truth jets of $R = 0.4$ (small R truth jets) and anti- k_t truth jets of $R = 1.0$ (large R truth jets). The p_T of each anti- k_t jet is indicated by a colored circle that states whether the jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS.

Figure 7.13 shows the event display for final state particles coming from branch 1 and branch 2 of non-minimal dark sector process for benchmark NMDS1 (defined in Table 6.12). This is a benchmark with small mass spacings between ensemble constituents that lead to a final state with a low p_T spectrum. In the event display, the leading large R jet is from initial state radiation and exhibits jet substructure, while the decay products of the exotic process are spread among four additional jets, with the second leading jet exhibiting jet structure. All large R jets apart from the leading large R jet fail the standard ATLAS reconstruction technique.

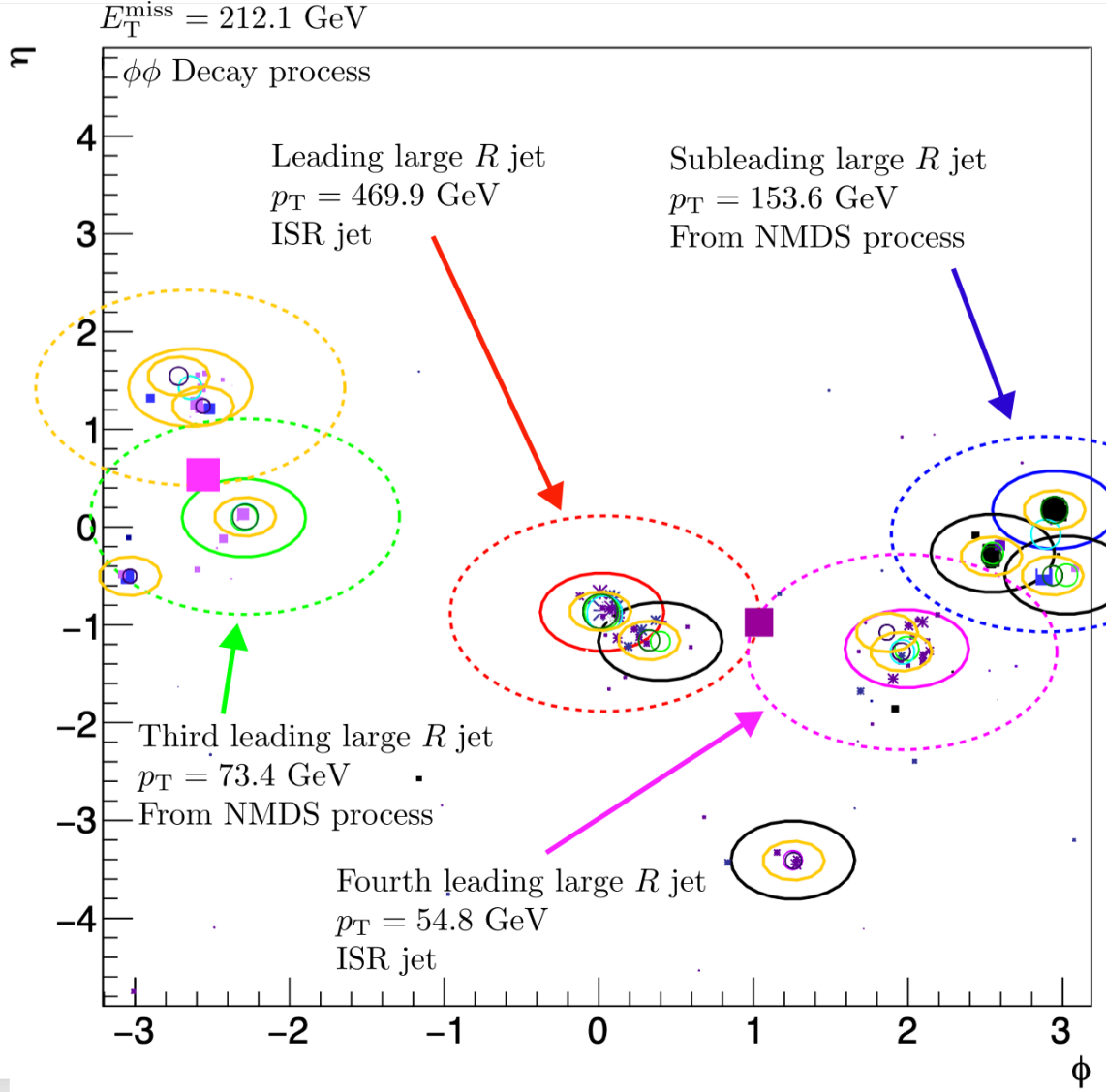


Figure 7.13 *Distribution of final state particles in the η - ϕ plane for NMDS Benchmark NMDS1 (defined in Table 6.12) for a single event. A $200 \text{ GeV } E_T^{\text{miss}}$ cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from branch 1 or branch 2 of the NMDS model. The corresponding legend for this event display is found in Figure 7.12.*

The event display of the non-minimal dark sector benchmark NMDS2 (defined in Table 6.12) is shown in Figure 7.14. This is a benchmark with mass spacings of 50 GeV between ensemble constituents that lead to a final state with a moderate p_T spectrum. Here is the first example of the leading large R jet not being from initial state radiation. The subleading and third leading large R jets all contain jet substructure. The exotic decay products from this model are scattered among four

large R jets.

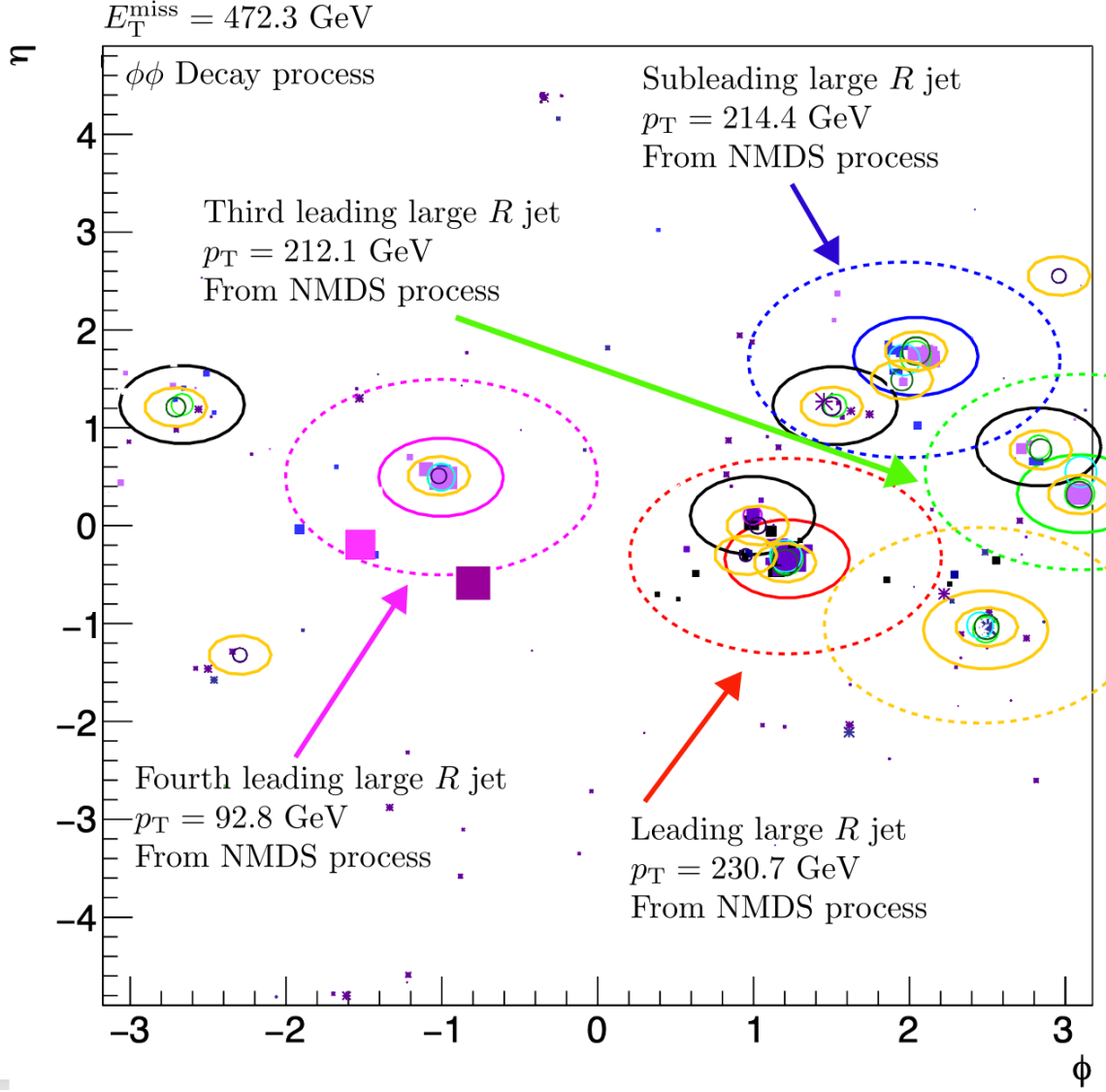


Figure 7.14 *Distribution of final state particles in the η - ϕ plane for NMDS Benchmark NMDS2 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from branch 1 or branch 2 of the NMDS model. The corresponding legend for this event display is found in Figure 7.12.*

The event display of the non-minimal dark sector benchmark NMDS3 (defined in Table 6.12) is shown in Figure 7.15. This is a benchmark with large mass spacings between ensemble constituents that lead to a final state with a high p_T spectrum. The decay products of the exotics process are spread over a large number of large R

jets, with a number exhibiting jet substructure.

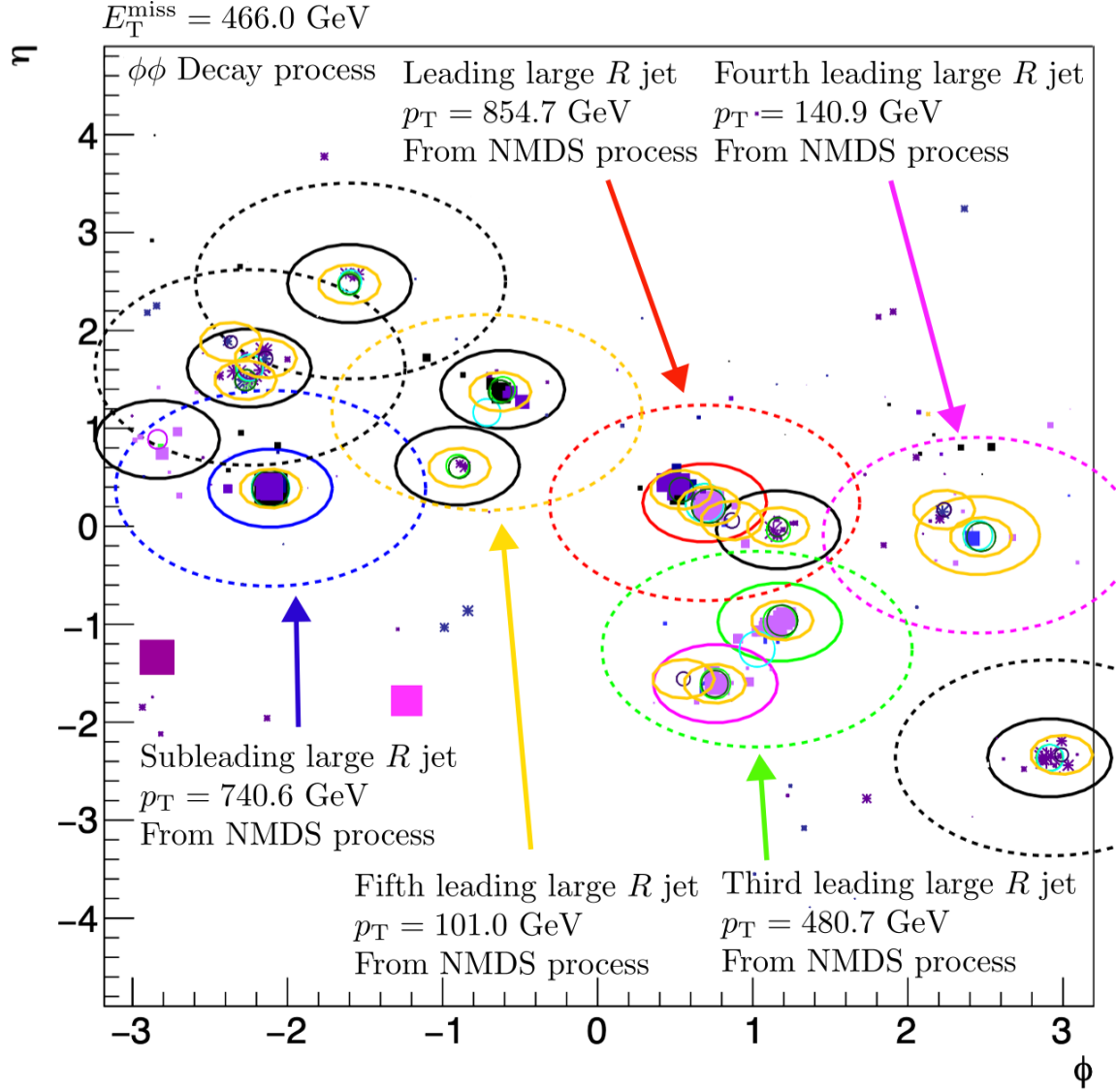


Figure 7.15 *Distribution of final state particles in the η - ϕ plane for NMDS Benchmark NMDS3 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from branch 1 or branch 2 of the NMDS model. The corresponding legend for this event display is found in Figure 7.12.*

Figure 7.16 shows the event display legend for final state particles coming from noncommutative black holes.

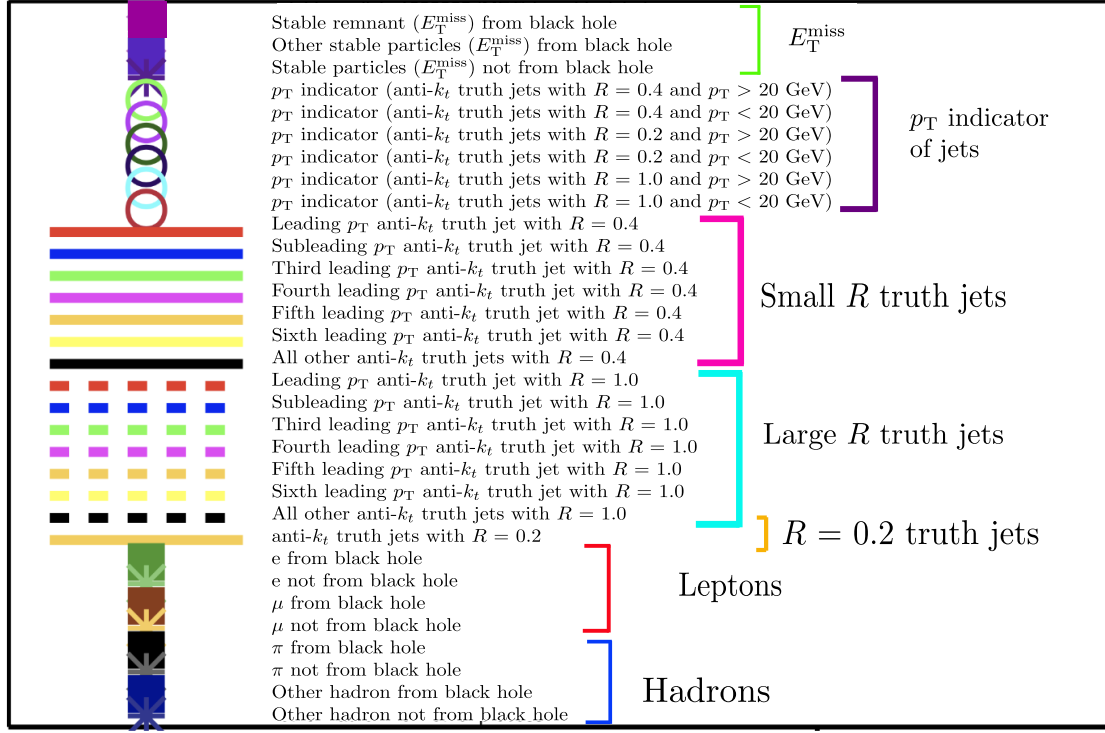


Figure 7.16 Legend for all NCBH Event Displays. The final state particles are tagged as either coming from the black hole or as not coming from the black hole. The legend includes a breakdown of seven main categories: E_T^{miss} , leptons, hadrons, anti- k_t truth jets of $R = 0.2$, $R = 0.4$, and $R = 1.0$, and p_T indicators (the smallest colored circles in the display) that indicate whether the truth jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS. The anti- k_t truth jets are additionally color coded to indicate the jet with the highest p_T (leading jet), the jet with the second highest p_T (subleading jet), etc.

The event display legend for noncommutative black holes identifies the types of final state particles produced and tags them as either coming from the black hole or not coming from the black hole. The E_T^{miss} is sectioned into three different categories: the stable remnant from the black hole, other stable particles from the black hole, or stable particles not from the black hole. Anti- k_t truth jets of three different radii are studied, with the leading through sixth leading jet color coded for anti- k_t truth jets of $R = 0.4$ (small R truth jets) and anti- k_t truth jets of $R = 1.0$ (large R truth jets). The p_T of each anti- k_t jet is indicated by a colored circle that states whether the jet is greater or less than 20 GeV, the minimum p_T value for calibrated small R jets in ATLAS.

Figure 7.17 shows an event display of Noncommutative Black Hole Benchmark NCBH1 (defined in Table 6.12) with a E_T^{miss} cut of 200 GeV with final state

particles identified as coming from the black hole. This is the benchmark with a noncommutative scale ($\sqrt{\theta}M_D$) of 0.55, which shifts the p_T spectrum of the benchmark to be higher than that of the NCBH4 benchmark (which has $\sqrt{\theta}M_D = 0.63$). All three large R jets exhibit jet substructure, displaying two prongs. The third leading large R jet fails the standard ATLAS large R jet reconstruction.

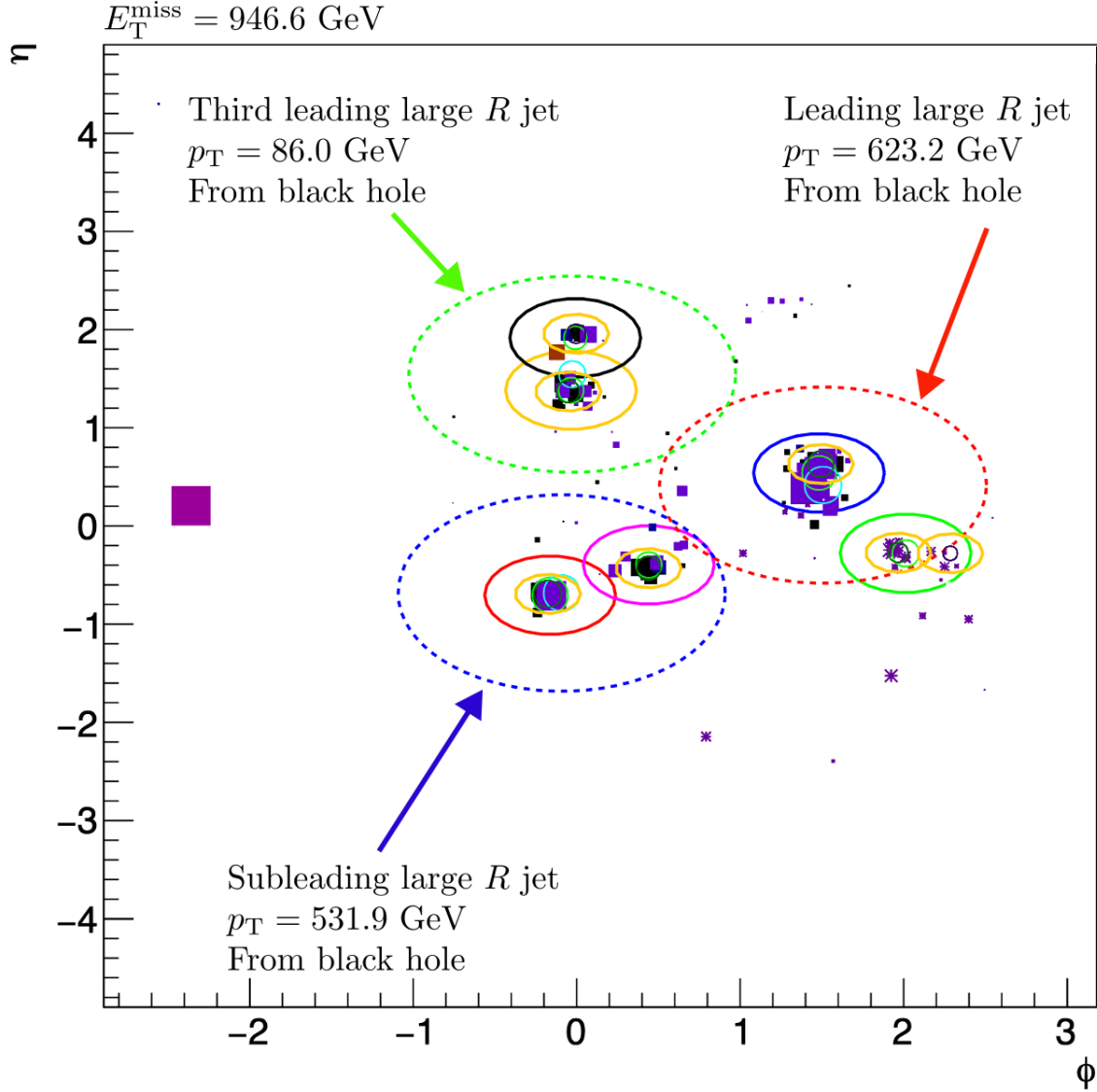


Figure 7.17 *Distribution of final state particles in the η - ϕ plane for NCBH Benchmark NCBH1 (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from the black hole. The corresponding legend for this event display is found in Figure 7.16.*

Figure 7.17 shows an event display of Noncommutative Black Hole Benchmark NCBH4 (defined in Table 6.12) with a E_T^{miss} cut of 200 GeV with final state particles identified as coming from the black hole. As noted above, this is the benchmark with $\sqrt{\theta}M_D = 0.63$, which yields a lower p_T spectrum than that of the NCBH1 benchmark. The decay products of the exotic process are spread among five large R jets, with the leading large R jet exhibiting a 3 prong structure and all of the large R jets except the leading jet failing the standard ATLAS jet reconstruction.

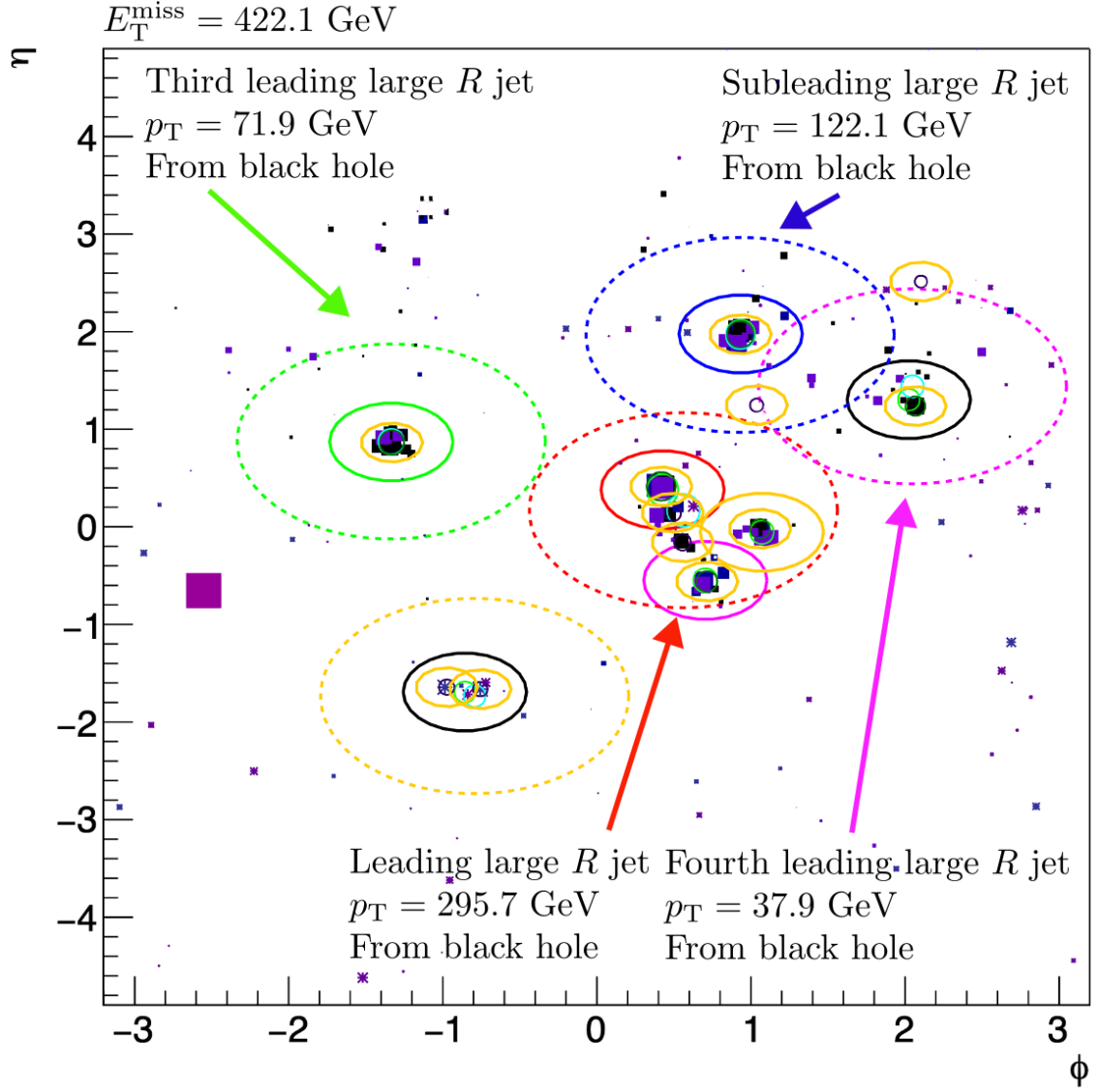


Figure 7.18 *Distribution of final state particles in the η - ϕ plane for NCBH Benchmark NCBH₄ (defined in Table 6.12) for a single event. A 200 GeV E_T^{miss} cut is applied. Large, dashed ellipses represents large R anti- k_t truth jets of $R = 1.0$ and smaller, solid ellipses represent small R anti- k_t truth jets of $R = 0.4$. Final state particles are indicated as colored dots and are tagged as originating from the black hole. The corresponding legend for this event display is found in Figure 7.16.*

The event displays show that scalar cascades, non-minimal dark sectors, and noncommutative black holes all exhibit jet substructure. The large R jets that contain the exotics process are often below the 200 GeV p_T threshold imposed by standard ATLAS reconstruction on large R jets. The event displays motivate an analysis strategy centered around jet substructure. The next sections seek to identify the best ways to target the substructure.

7.2 Large R jet Distributions

The event displays show that all three models - noncommutative black holes, non-minimal dark sectors, and scalar cascades - exhibit striking jet substructure. Additionally, the jet substructure is often too soft to be reconstructed with standard ATLAS jet reconstruction. Common variables that ATLAS uses to distinguish exotic processes that have a strong dependence on large R jets are mass and p_T , often employing cuts requiring high mass and high p_T . The mass and p_T of a jet are extracted from its four-vector. This section will provide large R jet kinematic plots that show that, in addition to largely failing the ATLAS standard reconstruction, these variables are not useful for noncommutative black holes, non-minimal dark sectors, and scalar cascades. Alternative strategies to elucidate the jet substructure from noncommutative black holes, non-minimal dark sectors, and scalar cascades have to be employed.

Many searches in ATLAS place a requirement on the sum of the mass of all large R jets in the event that have a p_T greater than 100 GeV, and typically within a certain η range. A p_T threshold of 100 GeV is used for these large R jets because the mass of the W boson and the mass of the Z boson are 80 GeV and 90 GeV, respectively, and one wants to consider at least a minimally boosted case in order to be able to use the identification of boosted bosons to distinguish background from signal. To increase the signal to background ratio, typical cuts used are, for example, a requirement of at least 500 GeV or at least 340 GeV for the sum of large R jets with a p_T greater than 100 GeV, as referenced earlier in the thesis in regards to a SUSY search [9]. (The SUSY search used two different cuts because they had more than one signal region.) However, as seen in Figure 7.19, the mass sum of all large R jets with p_T greater than 100 GeV for each of the three models emphasize how mass is not a useful quantity to apply a strong cut on for this analysis. A requirement on the sum of masses of 500 GeV for scalar cascades results in a loss of all signal and almost all signal for noncommutative black holes. All non-minimal dark sectors except the NMDS3 benchmark fail the cut almost entirely. For a 340 GeV sum of mass cut, a severe loss of signal would result for all but the NMDS3 benchmark.

Though not the primary focus of this section, it is worthwhile to comment on the distributions of various benchmarks for a given model in order to understand the effect that various parameters have on the phenomenology of the model. In Figure 7.19, the distributions of the mass sum for noncommutative black hole benchmarks NCBH4, NCBH5, and NCBH6 are largely aligned. These benchmarks

all live in 6 extra dimensions and have a $\sqrt{\theta}M_D$ value of 0.63. However, unlike benchmark NCBH4, which has a minimum black hole mass of 6.7 TeV, benchmarks NCBH5 and NCBH6 have minimum black hole masses of 7.8 TeV and 8.9 TeV, respectively. Since minimum black hole mass and M_D are related, benchmarks NCBH5 and NCBH6 also have different values of M_D (7 TeV and 8 TeV, respectively) than benchmark NCBH4, which has a M_D value of 6 TeV. The distributions for the mass sum of noncommutative black hole benchmarks NCBH1, NCBH2, and NCBH3 are similarly largely aligned and distinctly separate from the distributions for NCBH4, NCBH5, and NCBH6. Benchmarks NCBH1, NCBH2, and NCBH3 have a $\sqrt{\theta}M_D$ value of 0.55 and live in 5 extra dimensions. While the minimum black hole mass for benchmark NCBH1 is 6.8 TeV, the minimum black hole mass for benchmarks NCBH2 and NCBH3 are 7.8 TeV and 8.9 TeV, respectively. (The M_D value of benchmark NCBH1 is 6.5 TeV, while the M_D value of benchmarks NCBH2 and NCBH3 are 7.5 TeV and 8.5 TeV, respectively.) The noncommutative scale is seen to have a stronger effect on the phenomenology than the relatively negligible effect of varying the minimum black hole mass. The mass sum is larger for benchmarks with $\sqrt{\theta}M_D = 0.55$ than for benchmarks with 0.63. The non-minimal dark sector benchmarks with small mass spacings between ensemble constituents (NMDS6, NMDS5, NMDS1) and therefore lower p_T spectrums have mass sum distributions that are noticeably lower than the non-minimal dark sector benchmarks with moderate mass spacings of 50 GeV between ensemble constituents. The non-minimal dark sector benchmark with the highest mass spacings between ensemble constituents has the highest values for the mass sum. The scalar cascade benchmarks with the smallest number of Standard Model particles in the final state, SC4 and SC6, has lower mass sum distributions than the scalar cascade benchmarks with the highest number of Standard Model particles. The distributions for benchmarks SC3 and SC5 are barely, if not at all, noticable, due to the low rates at which they pass the 200 GeV E_T^{miss} cut. This low acceptance is due to benchmark SC5 having no stable scalars in the final state and benchmark SC3 having only 2 stable scalars with masses of 5 GeV.

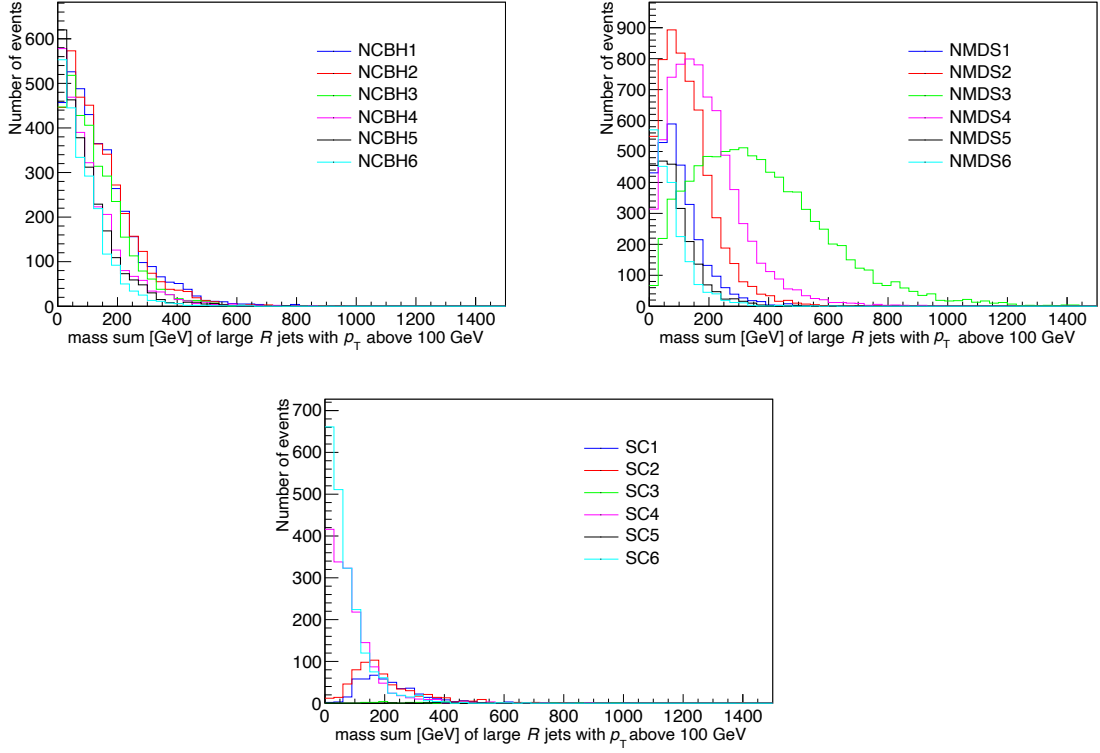


Figure 7.19 *Noncommutative Black Hole, Non-minimal Dark Sector, and Scalar Cascades Large R jet Sum of masses with a $200 \text{ GeV } E_T^{\text{miss}}$ cut applied.*

As seen in Figure 7.20, the p_T distributions of the second leading, third leading, and fourth leading large R jets for scalar cascade benchmarks all would suffer from the $200 \text{ GeV } p_T$ requirement from ATLAS for large R jets. As this is where the event displays showed the exotic signal to be, this would destroy the sensitivity to new physics. The leading large R jet p_T distributions would similarly not benefit from a requirement of at least one large R jet with a p_T of 460 GeV . (This is the threshold cut required to trigger on an event with an least one large R jet. It also serves as an example of a high p_T cut an analysis may impose on the leading large R jet to distinguish signal from background.) The leading large R jet p_T is, therefore, not a good variable to cut on for large R jets in the analysis.

As with the mass sum distributions, the scalar cascade benchmarks with the smallest number of Standard Model particles in the final state, SC4 and SC6, have softer p_T distributions for large R jets than the scalar cascade benchmarks with the highest number of Standard Model particles. The distributions for benchmarks SC3 and SC5 largely disappear again due to the low rates at which they pass the $200 \text{ GeV } E_T^{\text{miss}}$ cut.

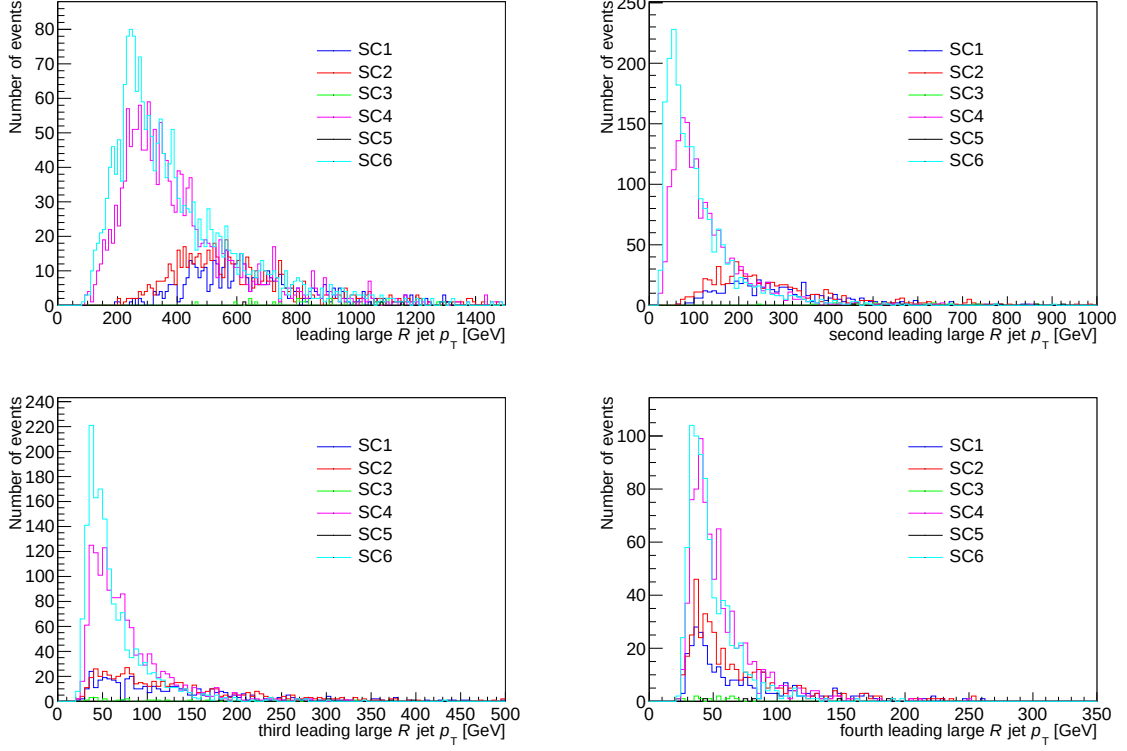


Figure 7.20 *Scalar Cascades Large R jet p_T distributions with a 200 GeV E_T^{miss} cut applied*

As seen in Figure 7.21, the p_T distributions of the third leading and fourth leading large R jets for noncommutative black holes all would suffer from the 200 GeV p_T requirement from ATLAS for large R jets. As this is where the event displays showed that exotic signal could lie in these jets for noncommutative black holes, that similarly would not allow for sensitivity to new physics. The leading large R jet p_T distributions lose roughly half of the signal if they were to undergo a 460 GeV cut in p_T . A 460 GeV cut on leading large R jet p_T is, therefore, not an optimal cut.

As with the mass sum distributions, the large R jet p_T distributions for noncommutative black hole benchmarks with a $\sqrt{\theta}M_D$ value of 0.63 (NCBH4, NCBH5, and NCBH6) are largely aligned and distinctly lower and separate from the large R jet p_T distributions for noncommutative black hole benchmarks with a $\sqrt{\theta}M_D$ value of 0.55 (NCBH1, NCBH2, and NCBH3). The noncommutative scale has a greater impact than the minimum black hole mass. These observations hold true for the leading and subleading large R jets. For the third and fourth leading large R jets, there is less of a distinction among the 6 models.

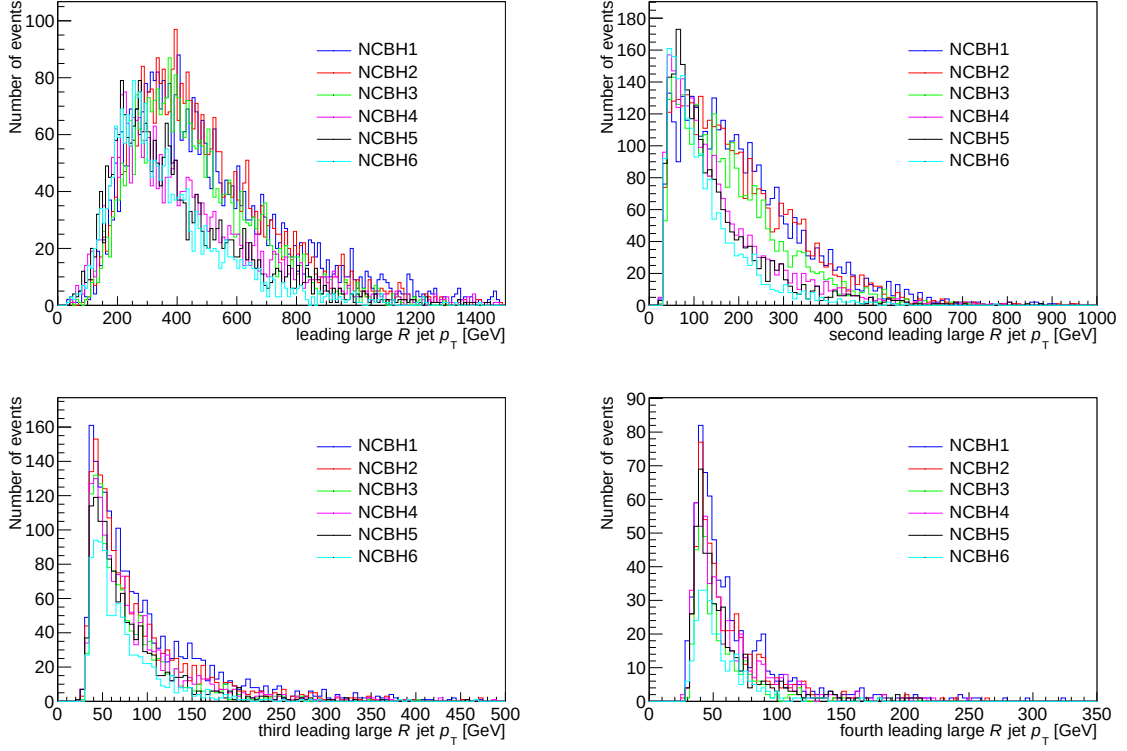


Figure 7.21 *Noncommutative Black Hole Large R jet p_T distributions with a 200 GeV E_T^{miss} cut applied*

As seen in Figure 7.22, the second leading large R jet would struggle with a 200 GeV p_T for all non-minimal dark sector models except the NMDS3 benchmark. Such a cut on the third and fourth leading large R jets would destroy sensitivity for all but the NMDS3 benchmark. All models except the NMDS3 benchmark and perhaps the NMDS2 benchmark would suffer from a 460 GeV p_T cut on the leading large R jet.

The non-minimal dark sector benchmarks with small mass spacings between ensemble constituents (NMDS6, NMDS5, NMDS1) have lower large R jet p_T spectrums than the non-minimal dark sector benchmarks with moderate mass spacings of 50 GeV between ensemble constituents. The non-minimal dark sector benchmark with the highest mass spacings between ensemble constituents has the highest values for the large R jet p_T spectrums.

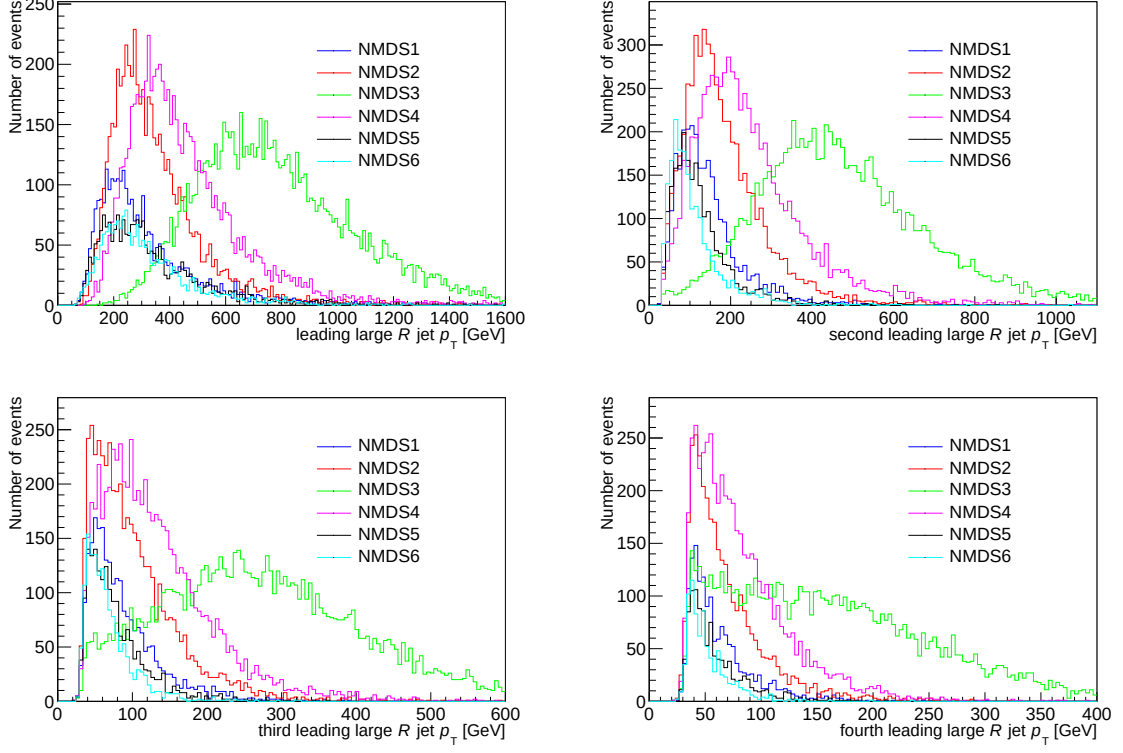


Figure 7.22 *Non-minimal Dark Sectors Large R jet p_T distributions with a 200 GeV E_T^{miss} cut applied*

This motivates the use of innovative designs to distinguish noncommutative black holes, non-minimal dark sectors, and scalar cascades from background.

7.3 Novel Large R jet Reconstruction Algorithm (Hybrid Jet Reclustering) vs. Standard Large R Jet Reclustering

From the event displays and large R jet distributions, it is clear that scalar cascades, non-minimal dark sectors, and noncommutative black holes all exhibit soft jet substructure. As the usual strategies employed by ATLAS searches are not well suited for this signature, a different strategy must be adopted. A novel large R jet reconstruction algorithm was designed specifically for this analysis. The ATLAS technique of jet reclustering allows for large R jets with p_T less than 200 GeV to be created from calibrated small R jets, as well as allowing for flexibility in the size of the jet radius. However, because the only inputs are small R jets, the standard

ATLAS jet reclustering only includes tracks that are within the small R jets in the reclustered large R jet. This results in a potentially significant loss of soft particles. To avoid this, a novel jet reconstruction technique, deemed hybrid jet reclustering, was developed to allow for the formation, from calibrated objects, of large R jets even below 200 GeV, while also allowing for a flexible radius of the new large R jet and the preservation of all tracks within the large R jet catchment area. Hybrid reclustering takes as input calibrated small R jets and all tracks within the full jet catchment area of the large R jet.

On the left side of Figure [95], a standard ATLAS large R jet (without reclustering) and all of its associated truth particles in the full jet catchment area of the large R jet are shown. On the right side of Figure 7.23, the reclustered large R jet is shown with all of the truth particles associated with a small R jet. It can be seen that reclustering causes all truth particles not associated with a small R jet in the large R jet to be lost. (Though Figure 7.23 was produced using truth particles, the same conclusion holds for tracks.) Hybrid jet reclustering eliminates this issue while also allowing for variable radius and lower p_T , unlike standard ATLAS large R jets. Hybrid jet reclustering can have broad applications for many other physics searches.

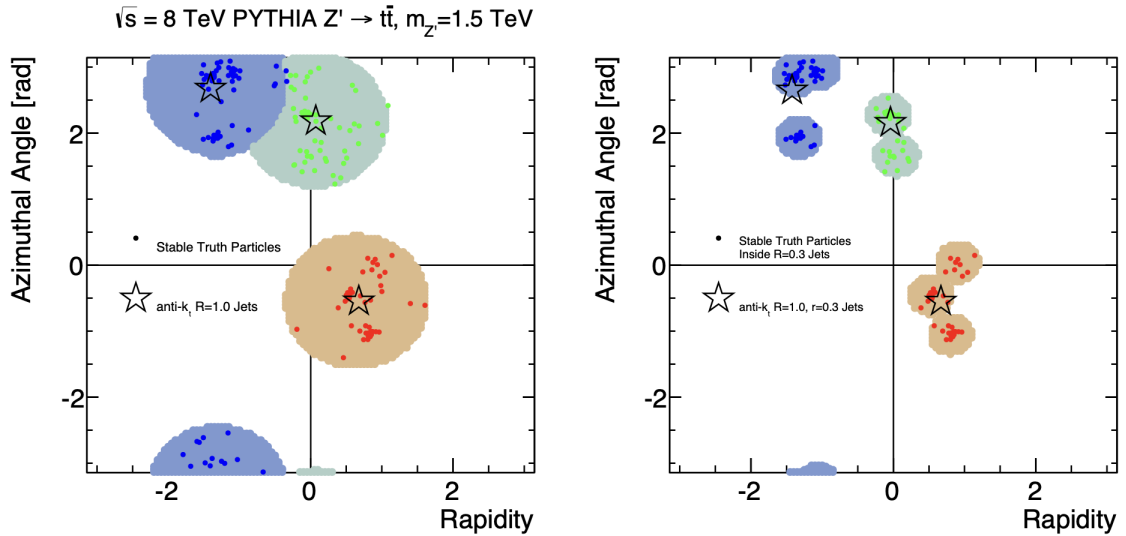


Figure 7.23 *Left plot: A visual of large R jet reclustering showing (a) Left: the large R jets and all of its associated truth particles (b) Right: all of the small R jets associated with the large R jet. Figure taken from [95]*

7.3.1 Technical details of Hybrid Jet Reclustering

The author implemented the hybrid jet reclustering algorithm in an analysis framework that the author created for this analysis. The analysis framework connects to common ATLAS functionalities, such as object calibration and object isolation (principles described earlier in the thesis) via the xAODAnahelpers package [96].

The specific technical details of the hybrid jet reclustering algorithm are the following. First, the anti- k_t algorithm is run over all calibrated small R jets with $p_T > 20$ GeV (and that pass the rest of the requirements set forth by the analysis for jet selection, which are outlined in the next chapter). The radius of the large R jet must be specified when running the anti- k_t algorithm. (In this analysis, for reasons that will be evident later in the chapter, the radius was specified to be $R = 1.5$.) This creates a Standard ATLAS reclustered large R jet. The constituents, or components, of this jet are small R jets.

This next step is where the hybrid jet reclustering algorithm makes a departure from Standard ATLAS jet reclustering. A list of all of the tracks in the event (that pass the requirements set forth by the analysis for track selection, which are outlined in the next chapter) is created. This list will be referred to as the list of original tracks. This list of original tracks is then copied to a new list which will be referred to as the list of “ghost tracks”, a term that will be explained shortly. This list of ghost tracks contains identical copies of the original tracks with the exception that the p_T of the track has been set to an infinitesimally small value. The anti- k_t algorithm is then run (with the same designated radius as before) over all calibrated small R jets with $p_T > 20$ GeV and the list of ghost tracks in the event. Since the ghost tracks have infinitesimally small p_T , they do not, in any way, bias or affect the reclustering algorithm. This is why these tracks are referred to as ghost tracks. It is as though each track is invisible, or a ghost. At this point, it is known which reclustered jet a ghost track is in. For the analysis, however, only real tracks are useful. Therefore, each ghost track in the list of ghost tracks is mapped back to each track in the list of original tracks. This is easy to do, of course, as the ordering of the list of original tracks is the same as the ordering of the copied list of ghost tracks. Now, it is known which tracks in the original list of tracks, the real tracks, are associated to which large R in the event. Additionally, it is known which small R jet is associated to which large R in the event. There are now two views of the large R jet: a track view in which the constituents, or components, of the large R jet are tracks and a small R jet view in which the constituents of the large R jet are small R jets. With this

novel technique, the large R jet captures even tracks that are outside the small R jets and, therefore, not previously available. This is a hybrid jet.

This thesis uses the anti- k_t algorithm to form the hybrid jets. This is because, prior to developing the novel hybrid jet reclustering technique, the author conducted studies with generic large R jet reclustering and found that the anti- k_t algorithm was more efficient at capturing the final state particles originating from the exotics process (for example, the scalar cascade process) than the k_t and Cambridge-Aachen algorithms. The anti- k_t algorithm was subsequently used for the hybrid jet reclustering. The next subsection details the decision to use a radius of $R = 1.5$ for the hybrid jet reclustering algorithm. Though not shown below, the hybrid jet reclustering was additionally tested for radius values of $R = 1.0$, $R = 1.2$, $R = 1.4$, $R = 1.5$, $R = 1.6$, $R = 1.8$, and $R = 2.0$. A radius of $R = 1.5$ was shown to be more efficient at capturing the final state particles originating from the exotics process than a radius of $R < 1.5$, while also not being much less efficient than large values of R . As larger values of R could potentially cause signal to be less distinctive from background, a radius of $R = 1.5$ was chosen.

Afterwards, studies by the author are shown that were used to validate the hybrid jet reclustering technique. Specifically, the performance of the hybrid jet reclustering technique at capturing the final state particles originating from the exotics process was compared to that of standard ATLAS jet reclustering, showing that the hybrid jet reclustering technique outperformed the standard ATLAS jet reclustering algorithm.

7.3.2 Determining a radius for Hybrid Jet Reclustering

To determine the radius of the hybrid jet used in the analysis, truth level studies of ΔR quantities by the author, in addition to the studies mentioned earlier, were used. In Figure 7.24, when applying a 200 GeV E_T^{miss} cut, the ΔR (distance in η - ϕ space) between the ϕ_i of the scalar cascades process and the final state particles from the scalar cascades process was determined to be almost exclusively below a value of 2 and predominantly below a value of 1.5. As noted above, a radius of $R = 1.5$ was chosen as larger values of R could potentially weaken the ability to separate signal from background. The 200 GeV E_T^{miss} cut shifted the small R jet p_T of the scalar cascades to be almost exclusively above 200 GeV, motivating a 200 GeV p_T cut for the analysis. A point to note is that, due to the 200 GeV E_T^{miss} cut applied in the bottom plot of Figure 7.24, there are considerably fewer events than in the two top plots of Figure 7.24. Though not marked as such, the scale of the z axis in the top

plots and the bottom plot are different, so it is not such that the light blue in the bottom plot indicates more events than the dark blue in the top two plots.

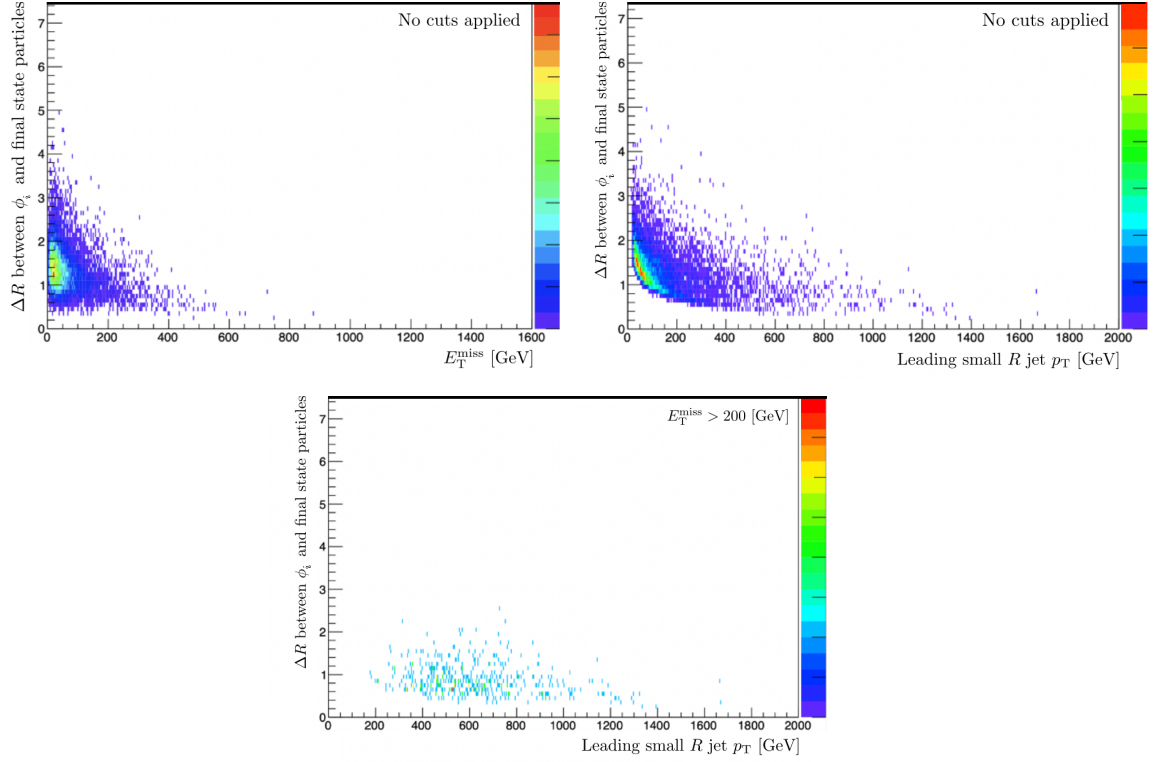


Figure 7.24 ΔR between final state particles and ϕ_i of the scalar cascade process as a function of (top left) E_T^{miss} without any cuts applied, (top right) leading small R jet p_T without any cuts applied, (bottom) leading small R jet p_T with a 200 GeV E_T^{miss} cut applied. This is for the Scalar Cascades Benchmark SC1.

7.3.3 Validating the Hybrid Jet Reclustering Algorithm

The performance gain of using the novel hybrid jet reclustering technique with a radius of 1.5 over the standard ATLAS reclustering technique with a radius of 1.5 can be seen in Figure 7.25.

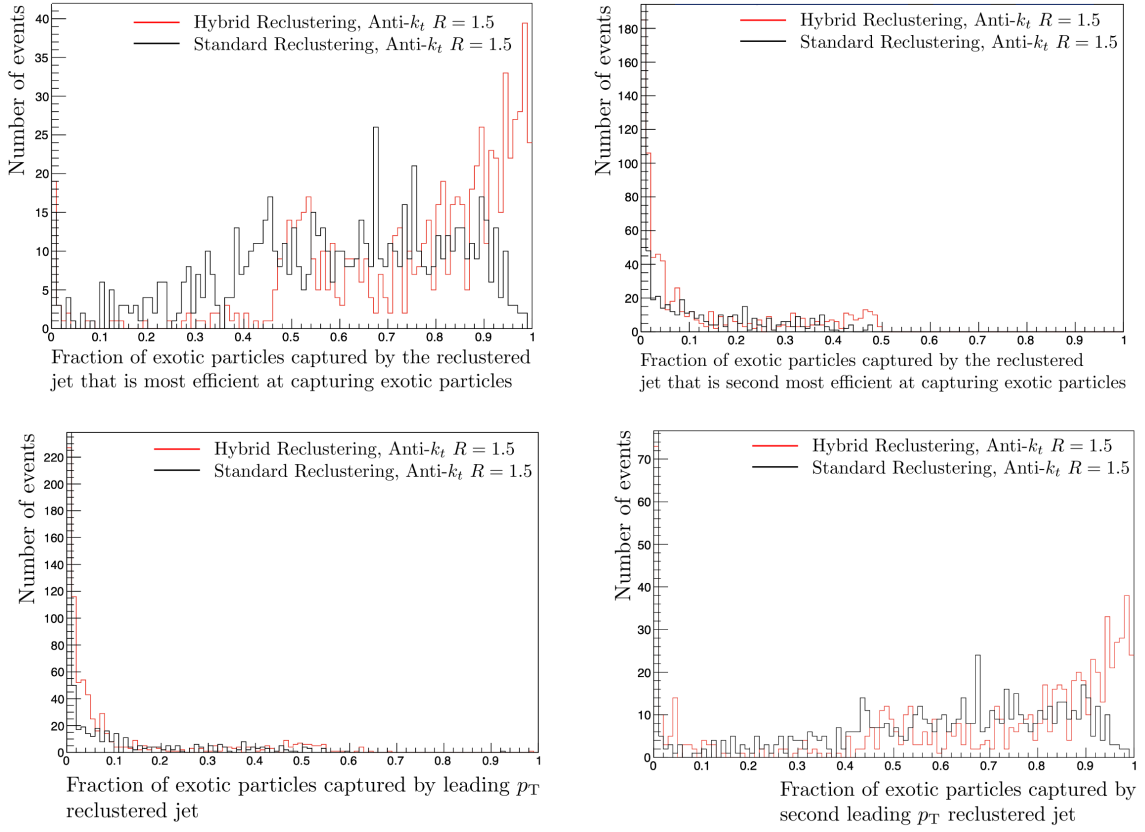


Figure 7.25 *Hybrid Jet Reclustering vs. Standard Large R jet Reclustering performance comparison for Scalar Cascade Benchmark SC1: Fractions of exotic particles captured by (Top left) the reclustered jet that is most efficient at capturing exotic particles (Top right) the reclustered jet that is second most efficient at capturing exotic particles (Bottom left) the leading p_T reclustered jet (Bottom right) second leading p_T reclustered jet. A $200 \text{ GeV } E_T^{\text{miss}}$ cut applied.*

The top left plot shows, for the Scalar Cascade Benchmark SC1, the fraction exotic particles from the scalar cascade process that is captured by the reclustered jet that is determined to be the most efficient at capturing the exotic particles for both the hybrid reclustering technique and the standard reclustering technique. The hybrid reclustering technique captures more exotic particles than the standard ATLAS reclustering. The bottom two plots show that the second leading reclustered jet is the jet that best captures the exotic particles from the scalar cascade decay. The bottom right plot shows that the hybrid reclustering technique is more efficient at capturing the particles from the exotic process than the standard reclustering technique.

This novel reclustering technique is used in the analysis to form reclustered jets from which a neural network takes track information to learn jet substructure.

7.4 Summary

Event displays show that scalar cascades, noncommutative black holes, and non-minimal dark sectors all exhibit jet substructure. Additionally, the jet substructure is often soft. The large R jets containing the exotics process are often below the 200 GeV p_T threshold that ATLAS requires for large R jet calibration. Due to the soft nature of the large R jets, the standard ways to differentiate exotic signal from background via means of the large R jet p_T and large R jet mass are not viable. A novel jet reclustering technique, called hybrid jet reclustering, was developed to allow the creation of jets that are below the 200 GeV p_T threshold and capture as many tracks as possible, including those inaccessible by existing algorithms. Studies were performed to demonstrate that the hybrid jet reclustering algorithm captures a higher fraction of the decay products of the exotic processes than the standard ATLAS reclustering algorithm. Additionally, optimization studies were performed to choose a radius of $R = 1.5$ for the hybrid jet reclustering algorithm for use in the analyses. The hybrid jet reclustering algorithm is used in the subsequent chapters to reconstruct large R jets whose track information is then fed into a neural network to learn the jet substructure in order to distinguish signal from background. This is applicable to all three models. Now that the analysis strategy and basic analysis objects have been established, the next chapter details the event selection for the analyses.

Chapter 8

Datasets and Event Selection

This chapter details the preliminary event selection (termed “preselection”) for each search. It also describes the simulated Monte Carlo samples used to model Standard Model background and exotic processes and lists the datasets used to analyze Run 2 data. It quantifies the background rejection afforded by the initial cuts.

8.1 Datasets

This analysis uses the full Run 2 datasets. The LHC divides data into roughly 1 minute long segments called lumiblocks, marking them as good or bad depending on whether the LHC beam was in good condition during the time of data-taking. The list of good lumiblocks for each year are collected in the following good run lists (GRLs), shown in Table 8.1:

Year	GRL
2015	data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS.StandardGRL_All.Good.25ns
2016	data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS.StandardGRL_All.Good.25ns
2017	data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS.StandardGRL_All.Good.25ns.Triggerno17e33prim
2018	data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS.StandardGRL_All.Good.25ns.Triggerno17e33prim

Table 8.1 *Good Run Lists for Run 2 data-taking*

This equates to 139 fb^{-1} of good quality data, with 3.21 fb^{-1} being attributed to 2015, 32.99 fb^{-1} being attributed to 2016, 44.31 fb^{-1} coming from 2017, and 58.45 fb^{-1} coming from 2018. Events are rejected if the events are not complete, if a subsystem was not working properly during the recording of the event, or if the event is corrupted.

Events are selected via data filtering algorithms called triggers. This analysis selects events based on the presence of substantial missing energy (E_T^{miss}) via E_T^{miss} triggers. The E_T^{miss} triggers used in this analysis are documented, by year, in Table 8.2.

Year	Trigger
2015	HLT_xe70_mht
2016	HLT_xe90_mht_L1XE50, HLT_xe110_mht_L1XE50
2017	HLT_xe90_pufit_L1XE50, HLT_xe110_pufit_L1XE55
2018	HLT_xe110_pufit_xe70_L1XE50, HLT_xe120_pufit_L1XE50, HLT_xe110_pufit_xe65_L1XE50

Table 8.2 E_T^{miss} triggers by year of Run 2 data-taking

The E_T^{miss} triggers are known to start becoming 100% efficient somewhere in the range of 200 GeV to 250 GeV, depending on the given ATLAS analysis [97]. The analysis therefore applies a basic 200 GeV E_T^{miss} cut for data and Monte Carlo background and signal samples. (The author applied this cut directly to the samples.) For two of the three searches (non-minimal dark sectors and noncommutative black holes), a higher E_T^{miss} cut is ultimately used in the final analysis result. The E_T^{miss} cut in the context of scalar cascades will be revisited in the discussion of experimental uncertainties. Signal Monte Carlo trigger efficiencies appear in the appendix.

8.1.1 Background Processes

An assortment of Monte Carlo generators are used at the LHC to generate simulated events, whether they are Standard Model background processes or exotics signal processes. The hard scatter process from the proton-proton collision is first simulated and the resulting final state particles are then showered and decayed. An intricate detector simulation then processes the entire event information. This simulation is based on GEANT4 [98]. The average number of bunch crossings per interaction crossing observed in ATLAS data plays an important role, as Monte Carlo simulation is then weighted to this distribution via three different MC Campaigns (MC16a, MC16d, and MC16e). MC16a corresponds to the 2015 and 2016 data-taking period, while MC16d corresponds to the 2017 data-taking period, and MC16e corresponds to the 2018 data-taking period.

As scalar cascades, noncommutative black holes, and non-minimal dark sectors all have multiple jets in the final state, the main background contributions come from Standard Model background processes that yield jets. Therefore, the dominant background processes that the analysis has to distinguish signal from are W/Z + jets, dibosons, top quark pair production, single top production, and QCD multijets.

(QCD multijets refers to jets arising from Quantum Chromodynamics processes.) The background Monte Carlo samples used are summarized, according to their DSID (data set identifier) in Table 8.3 with full sample names included in the appendix.

DSID	Background Process
700323 - 700325	Sh_2211_Zmumu_*
700326 - 700334	Sh_2211_Ztautau_*
700335 - 700337	Sh_2211_Znunu_*
700338 - 700340	Sh_2211_Wenu_*
700341 - 700343	Sh_2211_Wmunu_*
700344 - 700349	Sh_2211_Wtaunu_*
700439	Sh_2211_Zqq_ptZ_100_ECMS
700441	Sh_2211_Wqq_ptW_200_ECMS
410470	PhPy8EG_A14_ttbar_hdamp258p75_nonallhad
410471	PhPy8EG_A14_ttbar_hdamp258p75_allhad
410644	PowhegPythia8EvtGen_A14_singletop_schan_lept_top
410645	PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop
410646	PowhegPythia8EvtGen_A14_Wt_DR_inclusive_top
410647	PowhegPythia8EvtGen_A14_Wt_DR_inclusive_antitop
410658	PhPy8EG_A14_tchan_BW50_lept_top
410659	PhPy8EG_A14_tchan_BW50_lept_antitop
364702 - 364712	Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ*WithSW
361106	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee
363355	Sherpa_221_NNPDF30NNLO_ZqqZvv
363357	Sherpa_221_NNPDF30NNLO_WqqZvv
363360	Sherpa_221_NNPDF30NNLO_WplvWmqq
363489	Sherpa_221_NNPDF30NNLO_WlvZqq
363359	Sherpa_221_NNPDF30NNLO_WpqqWmlv

Table 8.3 *Standard Model MC Background Processes used in these searches and their corresponding DSIDs*

The production of a W boson in association with a quark is depicted in Figure 8.1. W/Z + jet processes are simulated and hadronized with the Sherpa 2.2.11 generator.

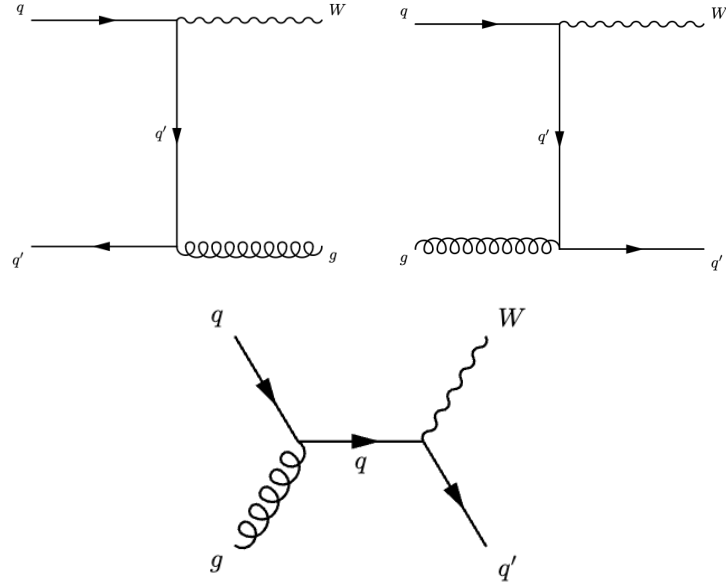


Figure 8.1 *W production in association with a jet*

Diboson production is depicted in Figure 8.2. Diboson processes are simulated and hadronized with the Sherpa 2.2.11 generator [99].

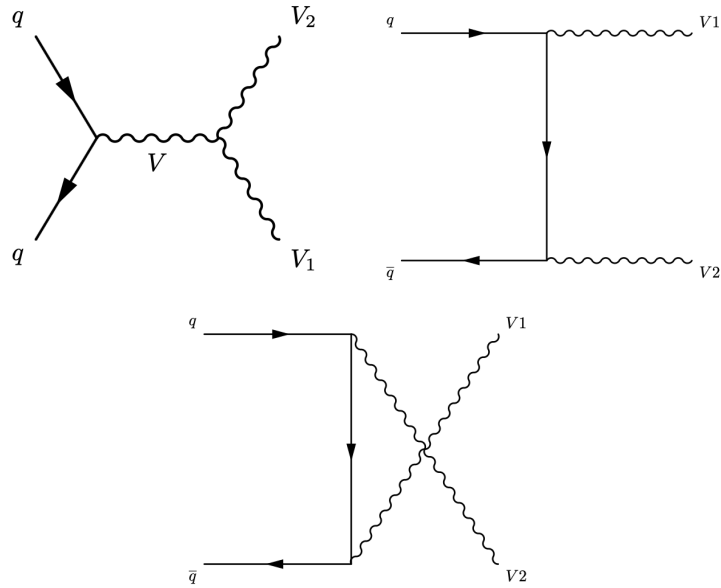


Figure 8.2 *Diboson production, where V can be a W boson or Z boson*

Top quark pair production is shown in Figure 8.3. Top quark pair production is simulated with Powheg Box 2 [100] and hadronized with [101].

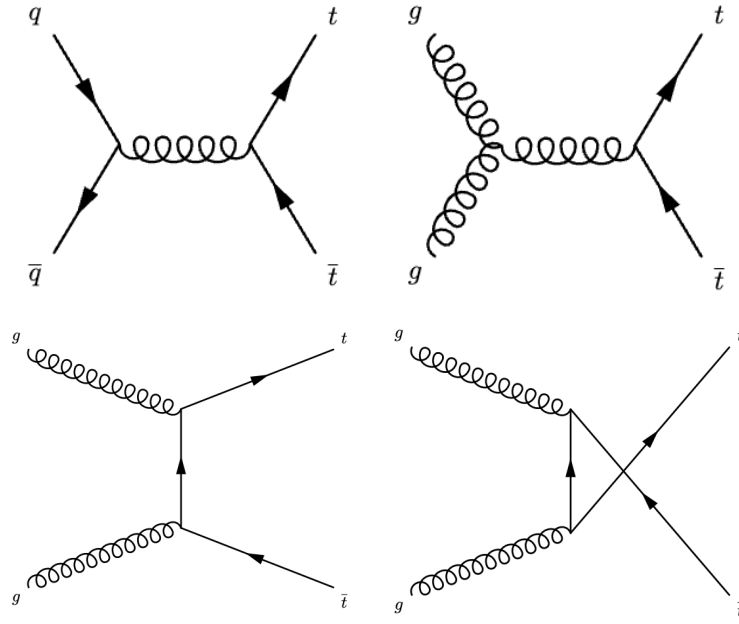


Figure 8.3 *Top quark pair production*

Single top production is shown in Figure 8.4. Single top production is simulated with Powheg Box 2 and hadronized with Pythia.

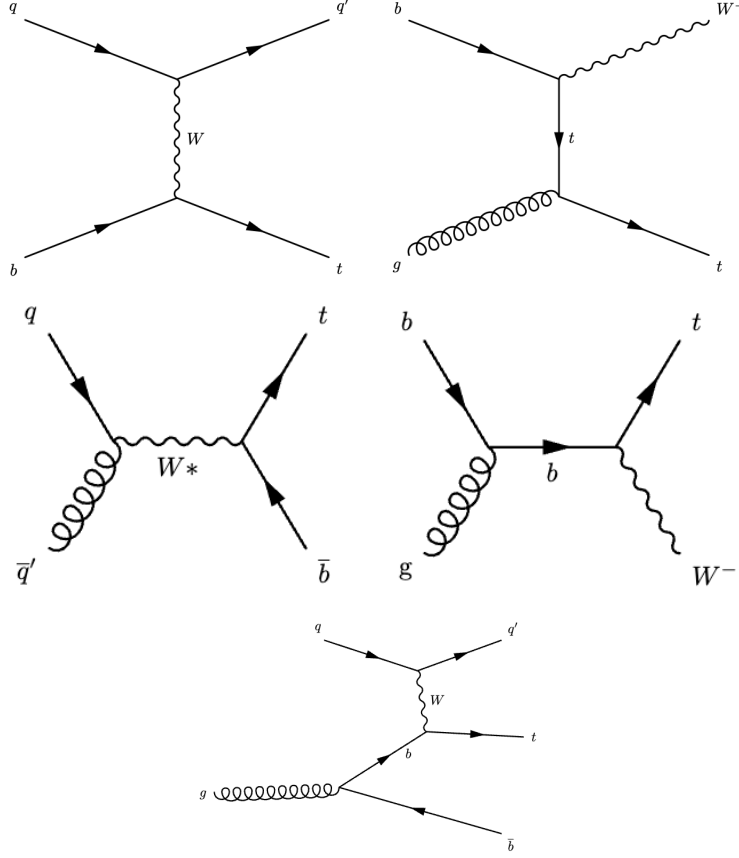


Figure 8.4 *Single Top Production*

QCD multijet production is simulated with Pythia8.230 using the A14 tune and NNPDF2.3 LO.

8.1.2 Object Definition and Selection

Primary Vertex (PV) The vertex in which the scalar sum of transverse momenta of all tracks associated to it is the highest is the primary vertex. This analysis requires the selection of events in which the primary vertex contains at least 2 tracks.

Tracks This analysis reduces contribution from pileup by requiring tracks to adhere to a tight reconstruction working point [102], as well as tight track-to-vertex association requirements [102]. Specifically, tracks are required to have track $p_T > 500$ MeV. The impact parameter d_0 is defined as the radial distance between the point of closest approach of a track to the primary vertex. Tracks are required to have $d_0 < 0.5$ mm. The impact parameter $\Delta z_0 \sin \theta$ is defined as the longitudinal distance between the point of closest approach of a track to the primary vertex.

Tracks are additionally required to have $\Delta z_0 \sin \theta < 0.5$ mm. Tracks are restricted to the full range of the Inner Detector: $-2.5 < \eta < 2.5$.

Muons

This analysis vetos both muons and electrons in order to enhance background rejection as signal typically does not contain leptons. However, muons and electrons are first both used in the overlap removal algorithm (to be discussed shortly) and in the algorithm used to reconstruct missing energy (to be discussed shortly). Muons are selected such that the minimum p_T is greater than 20 GeV and $-2.5 < \eta < 2.5$. The medium quality cut [103] and tight isolation criteria [103] are applied and the muons are calibrated. The muons are subsequently vetoed later in the preselection.

Electrons Electrons are required to meet the following criteria. Electrons are selected such that the minimum p_T is greater than 20 GeV and $-2.7 < \eta < 2.7$. The tight identification criteria [104] and tight isolated criteria [104] are applied and the electrons are calibrated. Electrons are subsequently vetoed later in the preselection.

Jets

ATLAS small R jets of radius $R = 0.4$ are reconstructed with the ParticleFlow algorithm. (As a reminder, the Particle Flow algorithm is an algorithm that runs the anti- k_t jet reclustering algorithm with both topo-clusters and tracks provided as input [84]). These jets are calibrated and have $p_T < 20$ GeV and $-3.2 < \eta < 3.2$. The JVT tight working point is applied [83]. The analysis applies two different types of cleaning criteria to jets: **Tight** cleaning and **Loose** cleaning. The **Tight** cleaning criteria comprise a stricter set of quality requirements than the **Loose** cleaning criteria. The **Loose** cleaning criteria [105] is applied to all small R jets in the event above 20 GeV. Additionally, the leading small R jet in the event is required to pass the **Tight** cleaning criteria [105].

Missing Energy

The missing energy is constructed using the electrons, muons, and jets that fulfill the criteria above. The missing energy construction is performed before overlap removal. As noted in the Reconstruction chapter, the definition of the soft term used in this analysis is the track-based soft term. (As a reminder, this means that the soft term consists of tracks unassociated to electrons, muons, and jets).

Overlap Removal

To avoid particles being reconstructed as different objects, a procedure called “Overlap Removal” is performed. In the instance where a jet and electron are separated by a radius $R < 0.4$, the electron is removed and the jet is kept. Similarly, in the instance where a jet and muon are separated by a radius $R < 0.4$, the muon is removed and the jet is kept.

Reclustered Large R jets The reclustered large R jets are created with the anti- k_t algorithm with radius $R = 1.5$. The reclustered jets take as input the calibrated small R jets passing the criteria outlined above, as well as the tracks passing the criteria outlined above.

8.1.3 Cuts Specific to Background and Cuts Specific to Data

QCD multijet samples are sliced by transverse momentum into 13 individual samples in order to ensure sufficient statistics. The naming scheme of these samples are Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ*WithSW, where the JZ0W slice corresponds to the QCD multijet sample with the lowest transverse momentum and JZ12 corresponds to the sample with the highest transverse momentum. This analysis ignores QCD multijet samples JZ0 and JZ1 because the analysis selects events with leading small R jets with $p_T > 200$ GeV and the JZ0 and JZ1 samples correspond to events with leading small R jet $p_T < 200$ GeV. QCD multijet slices JZ2W, JZ3W, and JZ4W are additionally treated to account for pileup overlay. Both the JZ2W and JZ3W QCD multijet samples undergo a cut that requires the leading small R jet $p_T < 1500$ GeV, while the JZ4W QCD multijet sample undergoes a cut that requires the leading small R jet $p_T < 1000$ GeV. These cuts are applied only to QCD multijet samples because QCD multijet samples are sliced by p_T . This is a standard procedure in ATLAS for physics analyses dealing with QCD jets. As noted earlier, in ATLAS Monte Carlo generation, pileup is added on top of the hard scatter. By definition, pileup is supposed to be softer than the hard scatter. In QCD multijet slices with lower p_T , such as JZ2W, JZ3W, and JZ4W, the hard scatter may occasionally be softer than pileup. Since lower p_T QCD multijet slices have higher cross sections than higher p_T QCD multijet slices, a pileup event with high p_T from a JZ2W slice will appear as an anomalously weighted event in the full jet p_T spectrum. These cuts remove anomalously weighted events.

Data events are vetoed if any of the small R jets in the event that pass the selection criteria for jets points to a bad tile module. The effect of the bad tile module veto and tight cleaning cuts on individual years can be found in the appendix. Both

the bad tile module event veto and the tight cleaning requirement for the leading small R jet offer great improvements to the $E_T^{\text{miss}} \phi$ distribution, resulting in the elimination of peaks in the distribution when both are applied. While the tight cleaning requirement is applied to Monte Carlo, the bad tile module veto can only be applied to data. This is because the ATLAS Collaboration does not simulate the bad tile modules in Monte Carlo, so there is no need to correct for them. The ATLAS detector description in Monte Carlo is reasonably close to an ideal detector, so random bits of the detector are not turned off. An efficiency cut is therefore applied to Monte Carlo to reflect the reduction in data due to rejecting events in which a jet points to a bad tile module. This efficiency factor is calculated for the selection cuts for each model prior to machine learning and will be quoted at the relevant parts of this thesis.

After applying the bad tile module veto, the log plot of the $E_T^{\text{miss}} \phi$ distribution in data (calculated by the author and shown in Figure 8.5) looks as expected.

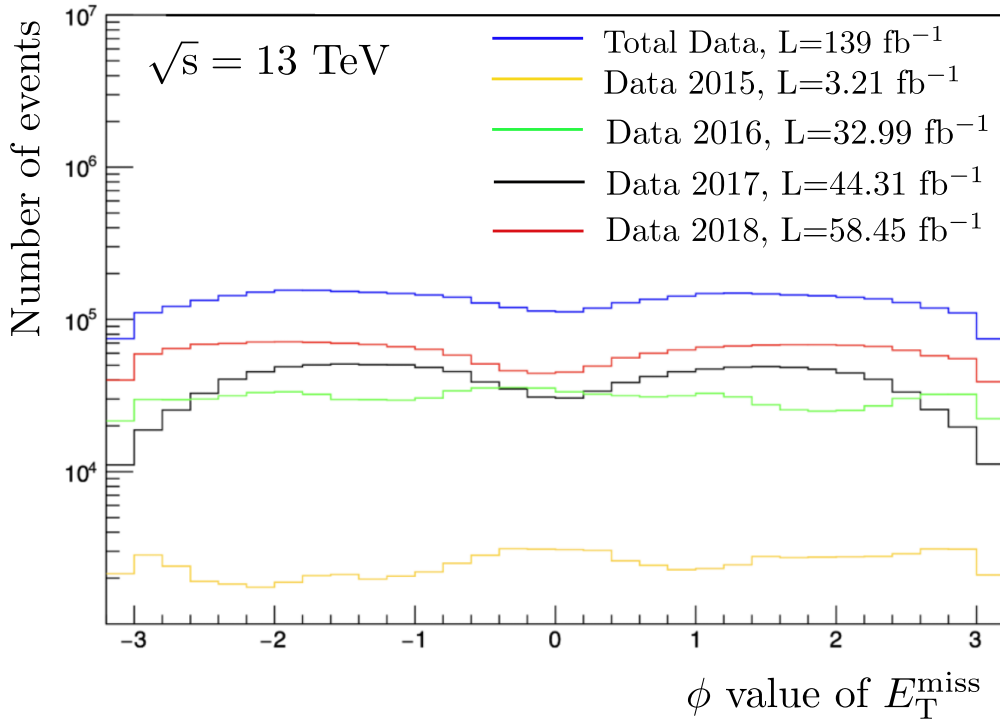


Figure 8.5 *Log of $E_T^{\text{miss}} \phi$ distribution after bad tile module event veto and after applying tight cleaning. Full Run 2 dataset is used.*

8.1.4 Scalar Cascade Cuts

As mentioned previously, the ATLAS collaboration limited the scalar cascade analysis to 1 benchmark due to the amount of Monte Carlo events required to generate a single reconstructed sample for use in machine learning. Scalar Cascade Benchmark SC1 (defined in Table 6.36) was chosen because, out of all of the benchmarks, it provided the most favorable combination of number of Standard Model particles in the final state (a staggering 128 particles), trigger efficiencies for the E_T^{miss} trigger, and cross section. The scalar cascade selection cuts prior to machine learning are summarized in the Table 8.4. These include cuts on the number of reclustered large R jets, the η values and mass values of the reclustered jets, the number of small R jets in the event, and the minimum angle between the closest large R jet and E_T^{miss} , almost entirely dictated by the event displays.

A cut of at least 50 GeV is applied to the second leading small R jet for cleaning purposes and background rejection. At least 5 small R jets of at least 20 GeV are required as an efficient way to reduce background. As seen in Figure 8.6, the distribution of the number of small R jets in the scalar cascade model peaks at 8 small R jets, while the distribution for background peaks at 5 small R jets. The ϕ value of the E_T^{miss} is largely aligned with the ϕ value of the closest leading large R jet, as in seen in Figure 8.6.

The angle between the closest large R jet to the E_T^{miss} is defined as $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, where

$$\Delta\phi_{\min}(J, E_T^{\text{miss}}) = \phi_{J_{\min}} - \phi_{E_T^{\text{miss}}}. \quad (8.1)$$

The angle between the closest small R jet to the E_T^{miss} is defined as $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, where

$$\Delta\phi_{\min}(j, E_T^{\text{miss}}) = \phi_{j_{\min}} - \phi_{E_T^{\text{miss}}}. \quad (8.2)$$

In Figure 8.6, $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ can be seen peaking sharply at 0 and the vast majority of signal residing below values of 0.5. This invokes a requirement that $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is less than 1. While the analysis places a cut on $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, it is worthwhile to observe the distribution of $\Delta\phi_{\min}(j, E_T^{\text{miss}})$.

Many dark matter searches, whether they are monojet searches or multijet searches, reduce QCD background by requiring $\Delta\phi_{\min}(j, E_T^{\text{miss}})$ or $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ to be greater than 0.5. QCD does not typically produce E_T^{miss} . Measurable E_T^{miss} in QCD events comes typically from mis-reconstruction of the event. Jets are dense objects, and there is some inefficiency in reconstructing all of the particles in them. Since QCD is a process that yields jets, it is subject to this reconstruction inefficiency. The E_T^{miss} in QCD is aligned with the jet because the E_T^{miss} comes from the component of the jet that is not reconstructed and, therefore, appears as E_T^{miss} .

As $\Delta\phi_{\min}(j, E_T^{\text{miss}})$ can be seen in Figure 8.6 to peak at 0 and lie almost entirely below 0.5, this immediately places this search outside the bounds of phase space typically explored. In recent years, both ATLAS [53] and CMS [54] have undertaken searches for semivisible jets, which has a distinctive signature of E_T^{miss} aligned with a jet. Scalar cascades, however evade these search strategies due to the soft nature of the scalar cascades, wherein the scalar cascades either fail the high p_T triggers employed or fail the high jet mass requirements. (The semivisible jet model has E_T^{miss} aligned with a jet because its decay products are stable particles that appear as E_T^{miss} and unstable particles, all of which are captured in a jet. Scalar Cascade Benchmark SC1 produces tracks that are too soft to reconstruct, resulting in measurable E_T^{miss} which is large enough to pass the E_T^{miss} trigger. This could then appear as E_T^{miss} aligned with the jet.) The unique phase space of scalar cascades is further highlighted in the strikingly high track multiplicity in the second leading reclustered jet, as seen in Figure 8.6. With the number of tracks in the second leading reclustered jet peaking at roughly 100 tracks, the scalar cascades clear the vast majority of background. (The number of tracks in the second leading reclustered jet is related to the number of Standard Model particles in a given scalar cascade benchmark, which is related to the number of scalars in the model. The peak at 100 tracks corresponds to the Scalar Cascade Benchmark SC1. However, this can describe other scalar cascade benchmarks that yield high Standard Model particle multiplicities).

As the particles from the exotics process are contained largely in the second leading reclustered large R jet of $R = 1.5$, at least two reclustered large R jets are required. Both the leading reclustered jet and the subleading reclustered jet are required to have greater than 0 tracks in order for tracks to be inputted in a machine learning algorithm focused on learning the jet substructure of the event. Due to the strikingly high track multiplicity, the second leading reclustered jet is additionally required to have at least 30 tracks.

The plots in Figure 8.6 were produced by scaling the simulated signal events to

a luminosity reflecting the full run 2 dataset and scaling by the total predicted cross section, then normalizing to one. The background distributions were similarly created by scaling the simulated background events to the full Run 2 dataset and scaling by cross section, then normalizing each individual background process in a manner such that the total background process was normalized to 1. These plots, and all similar plots appearing in the thesis, were produced via calculations performed by the author with an analysis framework created by the author (and referred to earlier).

Distributions given in Figure 8.6 have arbitrary units. Even though the integrated luminosity was previously defined as $\int \mathcal{L}(t)dt$, in the following plots it is marked with the label “L”. This holds true for all plots that carry this label. Additionally, a point to note is that $L = 139 \text{ fb}^{-1}$ is included even on plots that have been normalized. It is helpful to include the luminosity used even on normalized plots with arbitrary units because it indicates what data the MC background samples correspond to and also how the Monte Carlo samples were weighted. (For example, MC16a was weighted to a luminosity of 36.2 fb^{-1} , while MC16d was weighted to a luminosity of 44.31 fb^{-1} , and MC16e was weighted to a luminosity of 58.45 fb^{-1} , as opposed to being evenly weighted throughout the different MC campaigns).

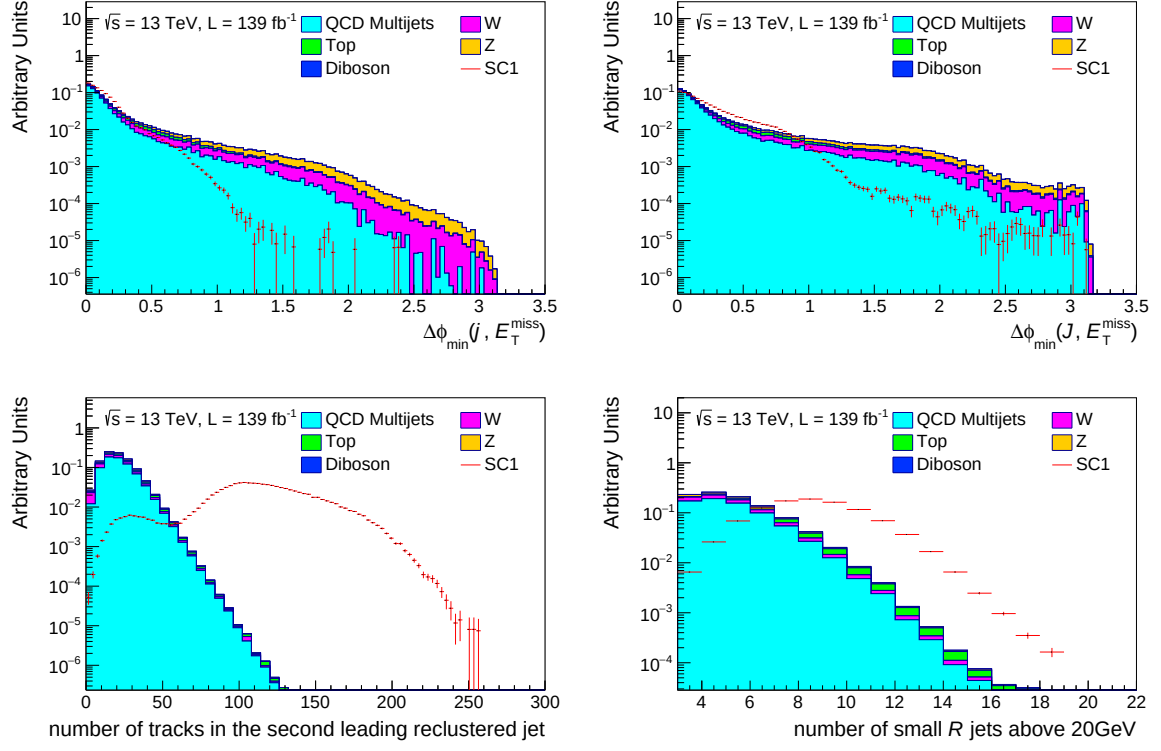


Figure 8.6 *Discriminating Variables of Scalar Cascade Benchmark SC1. Depicted here are $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, and the number of tracks in the second leading reclustered jet.*

Quantity	Requirement
Reclustered Jets	At least 2 reclustered jets of $R = 1.5$ number of small R jets in leading reclustered jet > 0 number of small R jets in subleading reclustered jet > 0 leading reclustered jet $ \eta < 2$ subleading reclustered jet $ \eta < 2$
Small R jets	> 4 small R jets p_T of at least 20 GeV Second leading small R jet $p_T > 50$ GeV
E_T^{miss} alignment	$\Delta\phi_{\min}(J, E_T^{\text{miss}}) < 1$
Tracks	Up to 250 tracks in leading reclustered jet Up to 250 tracks in subleading reclustered jet Greater than 0 tracks in leading reclustered jet Greater than 30 tracks in subleading reclustered jet
Lepton Veto	No electrons > 20 GeV No muons > 20 GeV
Trigger	lowest unrescaled E_T^{miss} triggers
E_T^{miss}	> 200 GeV
Leading small R jet p_T	> 200 GeV

Table 8.4 *Scalar Cascades cuts prior to Neural Network Cuts*

The effects of the Scalar Cascade cuts are seen in Table 8.5. Due to the relatively low cross section, after the selection, only an expected 2.7 signal events survive. The significance, defined as

$$\text{significance} = \frac{\text{number of signal events}}{\sqrt{\text{number of background events}}}, \quad (8.3)$$

for Scalar Cascade Benchmark SC1, is 0.0053. In the remaining chapters, the term significance when used refers to Equation 8.3.

QCD is the primary background, contributing 76.7%, then followed by Z , W , and Top processes.

Process	SC1	Total Background	QCD	Z	W	Top
Event Yields	2.7	256845	197086.6	9938.6	18153.3	30396.3
Background Composition %			76.73%	3.87	7.07%	11.83%
Tile Efficiency Correction	51.84%					
Significance: signal/sqrt(background)	0.0053					
Increase in significance from 200 E_T^{miss} cut	265%					

Table 8.5 *MC Event Yields for Scalar Cascades cuts prior to the SC Neural Network Cuts defined in Table 8.4. Background composition, signal significance, and Tile Efficiency Correction resulting from these cuts are also included.*

8.1.5 Noncommutative Black Hole and Non-minimal Dark Sector Cuts

The noncommutative black hole and non-minimal dark sector selection cuts prior to machine learning are summarized in the table below. Though noncommutative black holes and non-minimal dark sectors contain many jets in the final state, there is a modest requirement of a minimum of 3 small R jets for noncommutative black holes and non-minimal dark sectors. A requirement for a larger number of small R jets for noncommutative black holes and non-minimal dark sectors was found to bias the surviving background events to those exhibiting jet substructure and, thereby, decreased the performance of the neutral network. Scalar cascades were immune to this effect because of their high track multiplicity, which distinguishes them from background. Additionally, noncommutative black holes and non-minimal dark sectors do not have E_T^{miss} aligned with the closest small R jet or closest large R jet, so there is no requirement that $\Delta\phi_{\min}(J, E_T^{\text{miss}}) < 1$. Detailed studies of the kinematic distributions of noncommutative black holes and non-minimal dark sectors appear in the next chapter.

Quantity	Requirement
Reclustered Jets	At least 2 reclustered jets of $R = 1.5$ number of small R jets in leading reclustered jet > 0 number of small R jets in subleading reclustered jet > 0 leading reclustered jet $ \eta < 2$ subleading reclustered jet $ \eta < 2$
Small R jets	> 2 small R jets p_T of at least 20 GeV Second leading small R jet $p_T > 50$ GeV Third leading small R jet $p_T > 20$ GeV
Tracks	Up to 250 tracks in leading reclustered jet Up to 250 tracks in subleading reclustered jet Greater than 0 tracks in leading reclustered jet Greater than 0 tracks in subleading reclustered jet
Lepton Veto	No electrons > 20 GeV No muons > 20 GeV
Trigger	lowest unrescaled E_T^{miss} triggers
E_T^{miss}	> 200 GeV
Leading small R jet p_T	> 200 GeV

Table 8.6 *Most basic set of cuts for NCBH and NMDS prior to Neural Network Cuts. Some NCBH and NMDS benchmarks have additional cuts prior to the NN, which will be discussed in the next chapter*

As with the scalar cascades, in addition to requiring the E_T^{miss} to be greater than 200 GeV, the leading small R jet p_T to be greater than 200 GeV, and the subleading small R jet p_T to be greater than 50 GeV, the third leading small R jet p_T is required to be greater than 20 GeV.

As seen in Figure 8.7, apart from the NMDS benchmark NMDS3 which has large mass spacings between ensemble constituents which lead to energetic decays and the

scalar cascades (whose lead jet is always initial state radiation and not the exotic process), all NCBH and NMDS benchmarks peak around 200 GeV in lead small R jet p_T (when the 200 GeV requirement is set for small R jets).

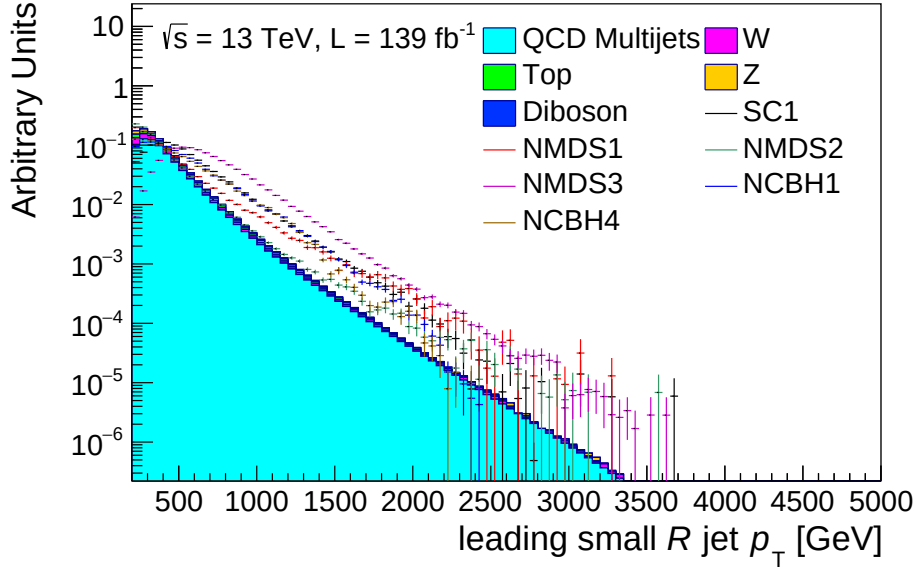


Figure 8.7 *Lead small R jet p_T with basic 200 GeV cuts applied.*

The effects of these basic cuts are seen in Table 8.7. QCD is the primary background, contributing to 72% of the background, while Z processes, W processes, and Top processes make up 7.60%, 12.42%, and 7.27% of the background. The significance of the Scalar Cascade Benchmark SC1 is 0.002, while the significance of the NMDS benchmarks NMDS1, NMDS2, and NMDS3 (defined in Table 6.12) are 0.8, 0.51, and 0.006. The NCBH benchmarks NCBH1 and NCBH4 (both defined in Table 6.11) have significances of 0.78 and 0.95. All of these require optimization to have decent sensitivity, and those optimizations occur in the subsequent chapters. The Scalar Cascade cuts imposed in Table 8.4 yield a significance that it is 2.65 times improved compared to the one obtained with the cuts in Table 8.6.

Background Process	Total Background	QCD	Z Processes	W Processes	Top Processes	
Background Yield	3354078.5	2415577.88	256004.5	416688.43	243832.46	
Background Composition		72%	7.60%	12.42%	7.27%	
Signal Process	SC1	NMDS1	NMDS2	NMDS3	NCBH1	NCBH4
Signal Yield	3.90	1462.91	933.58	10.9	979.1	1194.85
Significance	0.002	0.8	0.51	0.006	0.78	0.95
Tile Efficiency Correction	68.6%					

Table 8.7 *MC Event Yields for Noncommutative Black Holes, Nonminimal Dark Sectors, and Scalar Cascades and Background processes after applying the basic NCBH and NMDS cuts defined in Table 8.6. Background composition, signal significance, and tile efficiency correction resulting from these cuts are also included.*

8.2 Summary

This chapter described the simulated Monte Carlo samples used in the analyses to model Standard Model background processes and the exotic processes and lists the datasets used to analyze Run 2 data. It detailed the preliminary event selection for each search. High track multiplicity in the second leading reclustered jet, high jet multiplicity, and small values of $\Delta\phi_{\min}(J, E_{\text{T}}^{\text{miss}})$ were shown to be defining features of Scalar Cascade Benchmark SC1, motivating cuts on these variables. As will be seen in the next chapter, noncommutative black holes and non-minimal dark sectors are less distinct from background than scalar cascades. A looser set of preselection cuts (the most important being a requirement of at least 3 small R jets with $p_{\text{T}} > 20$ GeV in the event) were applied to noncommutative black holes and non-minimal dark sectors in order to avoid biasing the surviving background events to be rich in jet substructure and thereby degrade the neural network performance. The significance obtained for each model after initial cuts are all below values of 1. These significance values are all generally low, which shows that further optimization is required in order to increase background rejection while preserving high signal efficiency. The next chapter describes how this optimization is performed via machine learning. The position and transverse momentum of tracks in the leading reclustered jet, tracks in the subleading reclustered jet, and all small R jets in the event are used to learn the jet substructure of the event and distinguish signal from background, thereby improving the sensitivity of each analysis.

Chapter 9

Machine Learning for Signal Region Definition and Background Estimation

Machine learning is a powerful tool that is used in many different areas of high energy physics. This chapter introduces machine learning and neural networks and describes how machine learning is used for signal region definition and background estimation in each search. It introduces the machine learning architecture, which relies on permutation invariant neural networks called particle flow networks. Track information from the most energetic and second most energetic large R jets is provided to the neural network so that the jet substructure can be learned and used to distinguish signal from background. Event level variables and the kinematic information of all small R jets in the event also serve as inputs to the neural network (NN). Two decorrelated variables, the NN output and another variable that discriminates signal from background, are used for the data-driven estimation of the Standard Model background processes. Monte Carlo is used to train the neural networks and for cross-checks in the background estimation, so this chapter presents comparisons between Monte Carlo and data to check that things are behaving as expected prior to the use of machine learning. However, as the background estimation is data-driven, there does not need to be perfect agreement between the two. The performance of the neural network for each analysis is evaluated to show that they are performing well.

9.1 Principles of Machine Learning and Neural Networks

Machine learning (ML) is a branch of artificial intelligence (AI) that allows algorithms to learn patterns without being explicitly programmed. High energy physics at the LHC employs machine learning in a multitude of ways to distinguish interesting events from uninteresting events. Machine learning is used in certain trigger designs, allowing a better selection of interesting events, in object reconstruction to more accurately reconstruct objects, and in event selection to distinguish signal from background.

There are four types of machine learning: supervised learning, unsupervised learning, semi-supervised learning, and reinforced learning. Supervised learning trains a model on labeled datasets, while unsupervised learning trains on unlabeled datasets. Semi-supervised learning trains on a small number of labeled datasets and a larger number of unlabeled datasets. Reinforcement learning allows for learning through trial and error, using rewards to signal positive behavior and punishments to signal negative behavior.

This thesis employs supervised learning. There are two types of supervised learning: classification and regression. The distinction between classification supervised learning and regression supervised learning is that the output of classification supervised learning is categorical, providing concrete values, while the output of regression supervised learning are continuous values. Classification supervised learning can further be broken down into four types: binary classification, multiclass classification, multilabel classification, and imbalanced classification. Binary classification allows for output data to have two mutually exclusive labels, while multiclass classification allows for output to be of three or more classes. Multilabel classification allows for output to have multiple mutually non-exclusive classes. Imbalanced classification is when the training dataset has an unequal distribution of classes. This thesis uses machine learning to distinguish signal from background via binary classification supervised learning through the use of neural networks.

Neural networks are a subset of machine learning whose structure is inspired by the human brain. A neuron, depicted in Figure 9.1, is comprised of four key parts. The dendrite receives signals from other neurons. The neuron body, termed the soma, sends signals to the axon if the signals received from the dendrite exceed a certain

threshold voltage. The axon sends signals from the soma to other neurons. The axons are connected to other neurons via synapses, the fourth key part of a neuron. The dendrite can be considered the input layer of the neuron and the axon can be considered the output layer of the neuron.

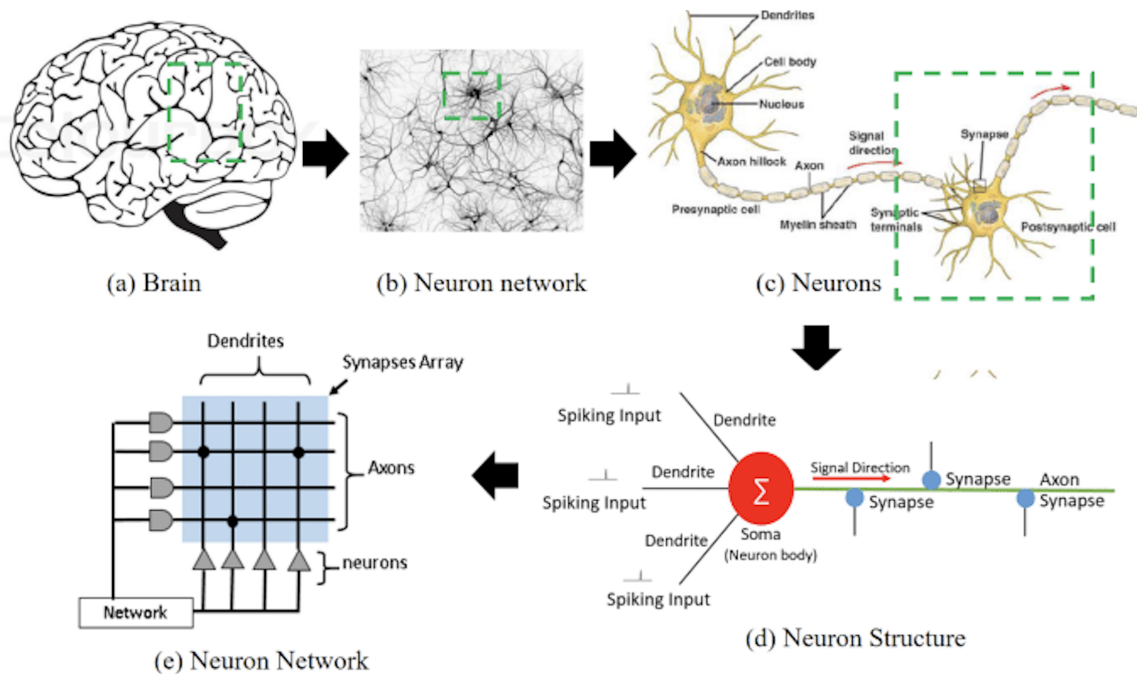


Figure 9.1 *Neuron structure Reference [106]*

Similarly, an artificial neural network, seen in Figure 9.2, has a structure in which there is an input layer and an output layer. The input layer is connected to the output layer via a middle layer called the hidden layer. There can be one or more hidden layers. Dense neural networks containing two or more hidden layers are called deep neural networks.

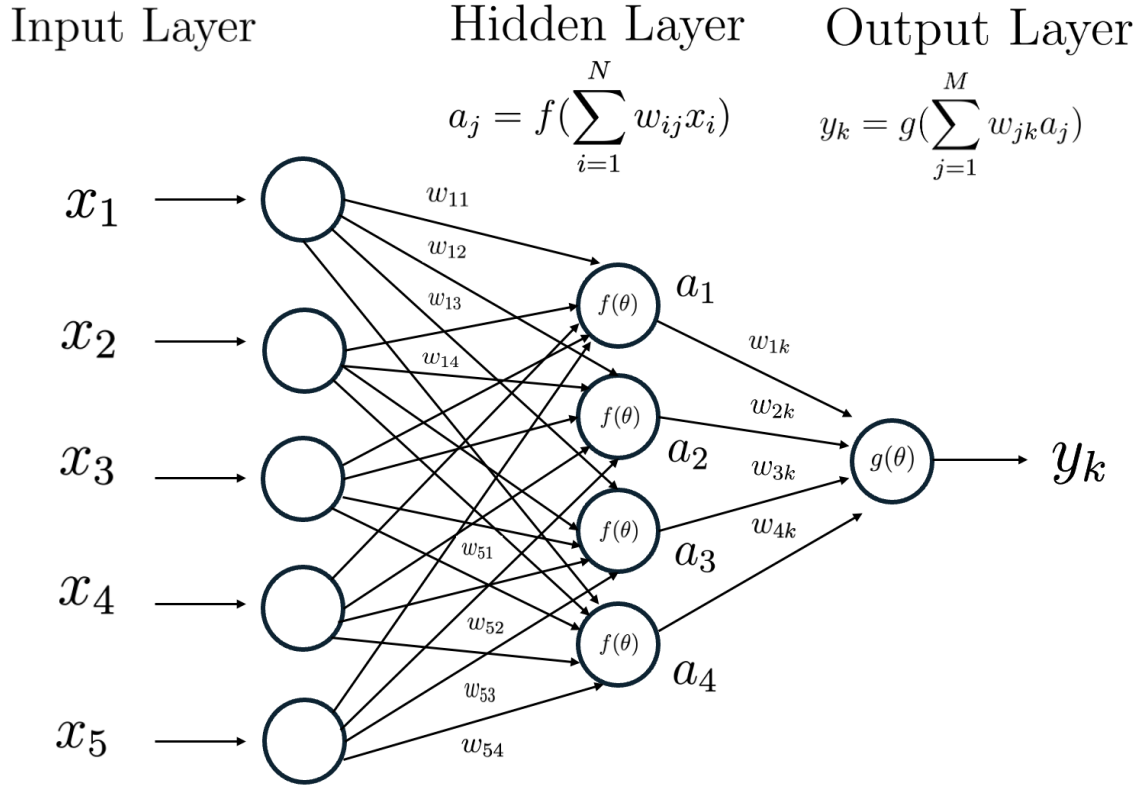


Figure 9.2 *Architecture of a neural network with one or multiple hidden layers. This figure is a personal remake of a figure from Reference [107]*

Figure 9.2 depicts a dense neural network with a single hidden layer. Each of the nodes in the hidden layer is connected to each of the nodes in the input layer. Similarly, each of the nodes in the hidden layer is connected to each of the nodes in the output layer. In the case of multiple hidden layers, not depicted in Figure 9.2, each node of the input layer would be connected to each node of the first hidden layer. Each node of the first hidden layer would be connected to each node of the second hidden layer. If a third hidden layer exists, each node of the second hidden layer would be connected to each node of the third hidden layer. In the absence of a third hidden layer, each node of the second hidden layer would be connected to each node of the output layer.

The nodes of the input layer are depicted in Figure 9.2 as x_1 , x_2 , x_3 , x_4 , and x_5 . The nodes of the hidden layer are denoted in the figure as a_1 , a_2 , a_3 , a_4 , a_5 . Each connection has a weight. The input node x_1 is connected to the input node a_1 with the weight w_{11} , while input node x_1 is connected to the input node a_3 with the weight w_{13} and the input node x_4 is connected to the input node a_2 with the weight w_{42} . The general formula to compute the value of given node, a_j , in the hidden layer is

$$a_j = f\left(\sum_{i=1}^N w_{ij}x_i\right). \quad (9.1)$$

The value of each input node, x_i , is multiplied by the value of each weight, w_{ij} , of the connection between x_i and a_j and summed together. The sum is then passed through an activation function, f , to give the value of a given node, a_j . There are a number of different activation functions used in neural networks. The most popular functions can be seen in Figure 9.3

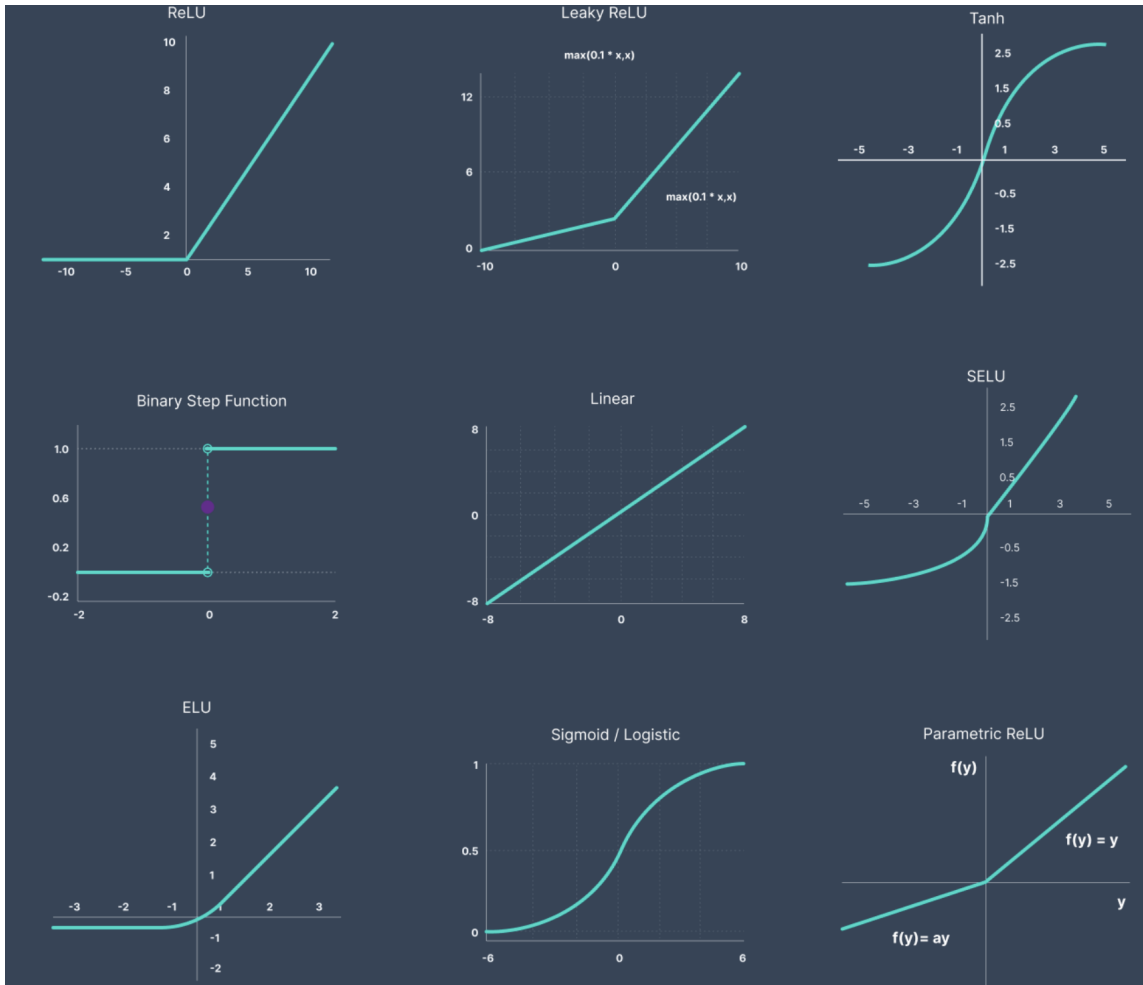


Figure 9.3 Activation functions in ML. Taken from Reference [108]

The rectified linear unit (ReLU) is an activation function that is commonly used for hidden layers of neural networks.

The value of the node, y_k , in the output layer is calculated with the following function

$$y_k = g\left(\sum_{j=1}^M w_{jk} a_j\right). \quad (9.2)$$

The value of each node, a_j , in the hidden layer is multiplied by the value of each weight, w_{jk} , of the connection between a_j and y_k and summed together. The sum is then passed through an activation function to give the value of the output node, y_k . The sigmoid activation function is commonly used in the output layer of binary classification, as the outputted values of the function are either 0 or 1.

Neural networks are trained by minimizing the difference between the function output of the neural network and the desired output of the neural network. As noted, in binary classification supervised learning, the desired output of a neural network is either 0 or 1, where 0 is typically used to denote background in a physics analysis context and 1 is typically used to denote signal. A loss function calculates the difference between the function output of the neural network and the desired output. It is, therefore, the loss function that must be minimized. There are a variety of loss functions that can be used for neural networks. A loss function that is commonly used due to the quality of its performance is the binary cross entropy loss function.

Binary cross entropy [109] is described through the following equation:

$$BCE = -\frac{1}{N} \left[\sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \right]. \quad (9.3)$$

The term y_i is the desired output (label) of a neural network, being either 0 or 1 for binary classification, N is the number of data samples considered during the loss calculation, and $p(y_i)$ is the function output (prediction) of the neural network.

The training of the neural network occurs through a process called backpropagation. A subset of the full dataset is randomly selected to form a training dataset. Data in this context means labeled simulation.

Several different techniques can be used to update the neural network weights during backpropagation. Some options for the binary cross entropy loss function, being a differentiable function, are batch gradient descent, mini-batch gradient descent, and stochastic gradient descent.

In each of these three descent methods, the following general procedure is followed.

The neural network weights of this training datasets are randomly initialized. The neural network is run over the training dataset. The loss function calculates the error from the predicted output and the desired output (labels). The neural network weights are then adjusted to minimize the loss function. The process is repeated for either a set number of epochs, where an epoch is one iteration of this process, or, in the case of the technique of “early stopping”, when the training stops improving. The difference among the three gradient descent methods is the amount of data that is used to minimize the loss function in a single epoch.

In batch gradient descent [110], the neural network weights are updated according to the following relation

$$\theta = \theta - \eta \cdot \nabla_{\theta} J\theta, \quad (9.4)$$

where $\nabla_{\theta} J\theta$ is the gradient of the loss function with respect to θ for the entire training dataset and η is the learning rate, or rate by which to approach the minima of the loss function. Here, the weights are updated once in a single epoch for the entire training dataset.

In stochastic gradient descent [110], the neural network weights are updated according to the following relation

$$\theta = \theta - \eta \cdot \nabla_{\theta} J\theta; x^{(i)}; y^{(i)}, \quad (9.5)$$

where the weights are updated for each event, $x^{(i)}$, and label, $y^{(i)}$

In mini-batch gradient descent [110], the neural network weights are updated according to the following relation

$$\theta = \theta - \eta \cdot \nabla_{\theta} J\theta; x^{(i:i+n)}; y^{(i:i+n)}, \quad (9.6)$$

where the weights are updated for each mini-batch of n training samples in the training dataset.

This thesis employs mini-batch gradient descent for backpropagation. Any of the gradient descent methods described have been acceptable to use. Mini-batch gradient descent for backpropagation was chosen simply out of necessity of having to chose one.

9.2 Deep Sets Based Architectures

Now that the basic principles of generic neural networks have been explained, more complex neural networks can be discussed. The Deep Sets architecture treats inputs as an unordered, variable length set. While a sequence is a list of elements where order matters, a set is a list of elements where order does not matter. In machine learning, the term latent space refers to a lower dimensional space, a space with reduced features, that encodes the most important features of higher-dimensional input data. The fundamental principle applied in [111] is that a permutation invariant function of variable length inputs can be constructed by mapping the inputs into a latent space, summing the latent space representations of each of those inputs, and feeding the summation into a reverse mapping from the latent space to the results of the target function that are desired. The key to this recipe is that summation is a permutation invariant operation.

Mathematically, this can be depicted [112] as

$$f(\{x_i, \dots, x_M\}) = F\left(\sum_{i=1}^M \Phi(x_i)\right), \quad (9.7)$$

where Φ is the function that maps each element, $x_i, \dots, x_M$ individually into the latent space and F is the function that acts on a single entity, the sum of the latent space representations of all of the individual elements, to provide a mapping out of the latent space to a single output value.

The allowance of the Deep Sets architecture for the permutation invariance and variable length of the inputs is important in a collider physics context because each event contains a variable length, unordered list of particles. The Deep Sets architecture described in [111] has been adapted and applied to particle physics in [112]. There, an observable, $\mathcal{O}$, is defined as a function of particles, such that

$$\mathcal{O}(\{p_1, \dots, p_M\}) = F\left(\sum_{i=1}^M \Phi(p_i)\right). \quad (9.8)$$

The $p_1, \dots, p_M$ denotes the attributes (charge, momentum, flavor, etc.) of M particles in an event. The Φ is the function that maps each particle into the latent space.

F is the function that acts on the sum of the latent space representations of each particle, providing a single output.

These operations are depicted pictorially in Figure 9.4

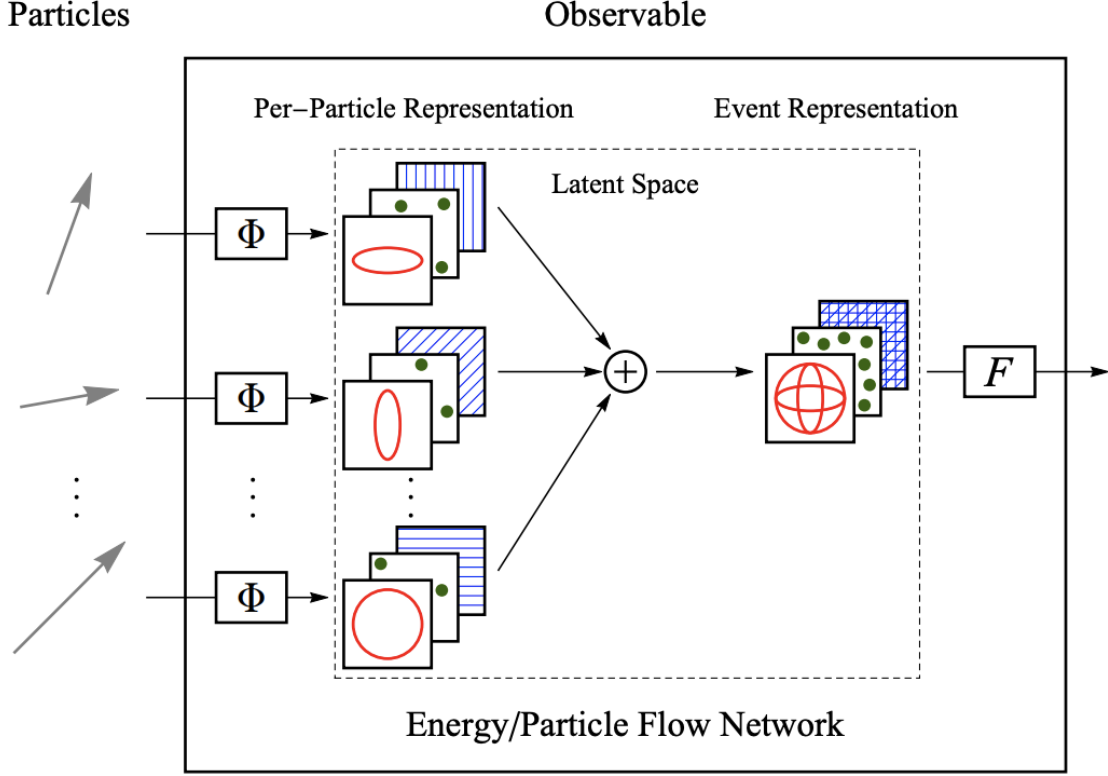


Figure 9.4 *Illustration of Φ providing a per particle mapping into latent space and F acting on the sum of the latent space representations to map out of latent space and provide a single output. Figure taken from Reference [112]*

Equation 9.8 can be modified to allow for infrared-collinear safe observables. Infrared-collinear safe observables are observables that collinear safe (robust against collinear splittings of a particle) and infrared safe (robust against the addition of a soft particle). Collinear safety and infrared safety were described in more detail in Chapter 4. To achieve this, a weight can be applied to each particle and the function Φ can be made to only depend on the geometry, $\hat{p}_i$, of the particle. The geometry of the particle, $\hat{p}_i$, is the rapidity, y_i , and ϕ_i of the particle. The values are calculated using the four-vector information of the particle. The weight applied would correspond to energy and be either $z_i = \frac{E_i}{\sum_j E_j}$ or $z_i = \frac{p_{T,i}}{\sum_j p_{T,j}}$, where E_i is the energy.

The infrared-collinear safe equation can be written as

$$\mathcal{O}(\{p_1, \dots, p_M\}) = F\left(\sum_{i=1}^M z_i \Phi(\hat{p}_i)\right). \quad (9.9)$$

Reference [112] calls networks described by Equation 9.8 particle flow networks, while calling networks described by Equation 9.9 energy flow networks. Reference [112] shows that particle flow networks perform better than energy flow networks in a number of tests regarding differentiating between gluon jets and light quark jets. While it is not necessary to distinguish gluon jets from quark jets for this analysis, what is needed is to distinguish signal from background and background will have jets from gluons. This chapter will detail how particle flow networks are used to distinguish signal from background in the analyses.

9.3 Distance Correlation (DisCo)

The background estimation for the analysis, the so-called ABCD method and which will be discussed in detail later, is contingent on the employment of two independent variables. Often, analyses at the LHC determine these variables by creating an extensive series of 2D histograms for a large number of pairings of different variables, hoping to find variables that are inherently independent. Another tactic, and that which is used in this analysis, is to force two variables to be decorrelated via the use of machine learning. To perform this automated and forced decorrelation, a function is needed that provides a way to measure the nonlinear dependence of distributions. Additionally, in order to use the function as a loss function for machine learning, the function needs to be differentiable and to be able to be calculated from the sample. A function that meets these demands is the distance correlation (DisCo) function. The distance correlation function can be understood best by first explaining distance covariance, from which it is derived.

Distance covariance is defined as

$$\text{dCov}(X, Y)^2 = \int d^p s d^q t |f_{X,Y}(s, t) - f_X(s)f_Y(t)|^2 w(s, t), \quad (9.10)$$

where X and Y are random variables, X is contained within all real numbers in p dimensions and Y is contained within all real numbers in q dimensions. The functions

f_X and f_Y are characteristic of X and Y , respectively. The joint characteristic function of X and Y is $f_{X,Y}$. The relation $f_{X,Y} = f_X f_Y$ upholds if and only if X and Y are independent variables [113].

The weight function, $w(s, t)$ from 9.10 is defined as as

$$w(s, t) \propto |s|^{-(p+1)} |t|^{-(q+1)}. \quad (9.11)$$

The distance covariance equation can be rewritten as

$$\text{dCov}(X, Y)^2 = \langle |X - X'| |Y - Y'| \rangle + \langle |X - X'| \rangle \langle |Y - Y'| \rangle - 2 \langle |X - X'| |Y - Y''| \rangle, \quad (9.12)$$

where $|X - X'|$, $|Y - Y'|$, and $|Y - Y''|$ are Euclidean vector norms and (X', Y') , (X, Y) , and (X'', Y'') are independent and identically distributed random variables from the joint distribution of (X, Y)

Distance correlation is obtained by normalizing 9.12 by the individual distance variances, such that

$$\text{dCorr}(X, Y)^2 = \frac{\text{dCov}(X, Y)^2}{\text{dCov}(X, X) \text{dCov}(Y, Y)}. \quad (9.13)$$

Distance correlation ranges from 0 to 1 and is 0 if and only if X and Y are independent [113].

The distance correlation term can be used to form the following loss function [113], called the DisCo loss function:

$$L = L_{\text{classifier}}(\vec{y}, \overrightarrow{y_{\text{true}}}) + \lambda \text{dCorr}_{y_{\text{true}}=0}^2(\vec{m}, \vec{y}). \quad (9.14)$$

This loss function has two parts. $L_{\text{classifier}}$ is the generic loss function that compares the output from the neural network and the truth label, y_{true} . In this analysis, $L_{\text{classifier}}$ is the binary cross entropy loss function. The term $\lambda \text{dCorr}_{y_{\text{true}}=0}^2(\vec{m}, \vec{y})$ is the term that incorporates distance correlation. The λ is the hyperparameter that determines the strength of the decorrelation. The NN output is denoted as $\vec{y}$. The

variable that is to be forced to be decorrelated from the NN output is denoted here as $\vec{m}$. The term $y_{\text{true}} = 0$ indicates that only the background, which has a truth label of 0, is forced to be decorrelated. For reasons related to the background determination (as will be explained in Chapter), while it is important for these two variables to be decorrelated for background, it is also important for the two variables to be highly correlated for signal. Therefore, the distance correlation is not applied to signal, $y_{\text{true}} = 1$.

9.4 Particle Flow Network - DisCo Architecture

The analysis distinguishes signal from background and creates two decorrelated variables for background estimation via a neural network that employs particle flow networks and a DisCo loss function. The neural network was created by the author and implemented in a machine learning framework created by the author. The definition of the distance correlation was taken from Reference [114] The architecture of the neural network is depicted in Figure 9.5.

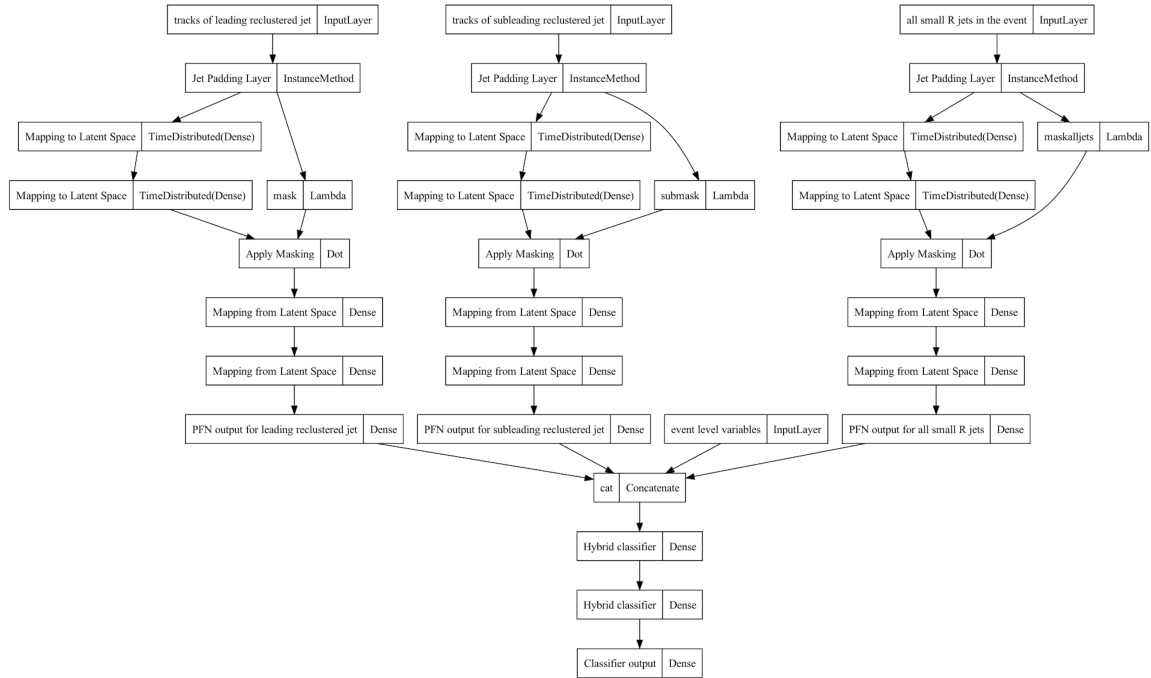


Figure 9.5 *Particle Flow Network - DisCo NN Architecture*

As shown in previous chapters, jet substructure is key to the analysis strategy. The neural network learns the jet substructure of the event by taking as input the p_T , rapidity, and ϕ of all tracks in the leading reclustered jet and, additionally, the p_T ,

rapidity, and ϕ of all tracks in the subleading reclustered jet. To account for there being multiple small R jets in the event, the p_T , rapidity, and ϕ of all small R jets in the event are also taken as input. This kinematic information serves to form three separate input layers, each supplied to a separate particle flow network (PFN). As a preprocessing step, the p_T of each track in the leading reclustered jet is divided by the total p_T of all of the tracks in the leading reclustered jet. The same normalization process is applied to the tracks in the subleading reclustered jet, and to the small R jets. This normalization procedure forces the neural network to learn the internal structure of the reclustered jet or event, instead of recognizing specific p_T values.

The average ϕ of all of the tracks in the leading reclustered jet is subtracted from the ϕ of each track in the leading reclustered jet. The ϕ values of tracks in the subleading reclustered jet, and the small R jets in the event, also have their respective averages subtracted. A similar procedure is undertaken for the rapidity values. This processing means each track in the leading reclustered jet has its position expressed in relation to the centre of the jet instead of in terms of absolute coordinates of the detector. This holds true for the tracks in the subleading reclustered jet and the small R jets in the event. This prevents the neural network from identifying events as signal or background based on what part of the detector they hit and, instead, forces the neural network to focus on the important information - the location of particles with respect to each other.

As the number of tracks in the leading reclustered jet, the number of tracks in the subleading reclustered jet, and the number of small R jets in the event are not known in advance, each layer is passed to a corresponding padding layer, in which a set maximum number is designated for, depending on the input layer, the number of tracks in the leading reclustered jet, the number of tracks in the second leading reclustered jet, or the number of small R jets in the event. This padding layer takes a maximum value of 250, 250, and 25 for the number of tracks in the leading reclustered jet, the number of tracks in the second leading reclustered jet, or the number of small R jets in the event, respectively. If an event contains less than the maximum number of particles, the event is padded with zeros in lieu of the missing particles, such that each event has the maximum number of entries populated. If, for example, an event contains 250 tracks in the leading reclustered jet, then only the first 250 tracks in the inputted list of tracks in the leading reclustered jet are taken.

This padded layer is then mapped to latent space with two sequential time distributed layers. The time distributed layers are what implement the deep sets theorem

described earlier. Each time distributed layer applies in a permutation invariant manner a function Φ , which is a dense layer with 128 nodes and a rectified linear unit (ReLU) activation function, to each track in either the first or second leading reclustered jet or to each small R jet in the event. This is the per-particle mapping described earlier. The output of each of the time distributed layers is then passed to a dot layer in which a mask is applied such that the entries in each event contain a p_T of zero, the padded entries are ignored. The output is then mapped out of latent space with two sequential dense layers, acting as the function F described earlier, each containing 128 nodes and utilizing a ReLU activation function. The output of the last layer is then passed to a dense layer of 4 nodes and a ReLU activation function.

At this point, as depicted in Figure 9.5, there are three separate intermediate output layers: one for the particle flow network of the leading reclustered jet, one for the particle flow network of the second leading reclustered jet, and one for the particle flow network of all small R jets in the event. Before concatenating these layers, a separate, additional layer is formed containing some select event level variables that help distinguish signal from background. These four layers are then concatenated into a single input that is then passed into a dense layer with 128 nodes with a ReLU activation function. This is the first of several dense layers that are operating as a standard neural network classifier, taking as input the information from the particle flow networks and event level variables and providing a classifier output. The output of the last dense layer is then passed into another dense layer with 128 nodes with a ReLU activation function. The output is then passed into a final dense layer with a single node with a sigmoid activation function. The sigmoid activation function allows for a single output result with values ranging from 0 to 1. This classifier result is obtained by employing a DisCo loss function 9.14. The studies below will show the advantage of using a DisCo loss function instead of a binary cross entropy loss function. The optimizer used to perform the mini-batch gradient descent is the Adam optimizer. The “learning rate” controls the step size at each iteration of the Adam Optimizer as it minimizes the loss function. The Adam Optimizer is run with a learning rate of 0.001. This is a default choice commonly used in machine learning. When minimizing the loss function, the impact of the local gradient is combined with a set of historical gradients from previous iterations, weighted such that their influence decays in future iterations. This decay rate is controlled by the parameter beta, which takes a value between 0 to 1. The Adam Optimizer is run with a beta value of 0.9, which, again, is a default choice commonly used in machine learning.

9.5 Introduction to event shape variables and other event level characterizations

Event shape variables are variables used to describe the geometry of an event. Five event shape variables commonly used in high energy physics analyses are sphericity, transverse sphericity, aplanarity, event shape variable C , and event shape variable D . The variables can be constructed from the eigen values, λ_1 , λ_2 , and λ_3 , of the linearized sphericity tensor, M_{xyz} . The linearized sphericity tensor M_{xyz} is defined as

$$M_{xyz} = \frac{1}{\sum_i |\vec{p}_i|} \sum_i \frac{1}{|\vec{p}_i|} \begin{pmatrix} p_{x,i}^2 & p_{x,i}p_{y,i} & p_{x,i}p_{z,i} \\ p_{y,i}p_{x,i} & p_{y,i}^2 & p_{y,i}p_{z,i} \\ p_{z,i}p_{x,i} & p_{z,i}p_{y,i} & p_{z,i}^2 \end{pmatrix}. \quad (9.15)$$

where $p_{x,i}$, $p_{y,i}$, and $p_{z,i}$ are the x , y , and z components of the i^{th} input particle's momentum in three dimensions. The inputs can be track, jets, constituents of jets, etc. The eigen values of this tensor must meet the criterion that

$$\sum_k \lambda_k = 1. \quad (9.16)$$

and

$$\lambda_1 > \lambda_2 > \lambda_3. \quad (9.17)$$

Sphericity, S , is defined as

$$S = \frac{3}{2} (\lambda_2 + \lambda_3). \quad (9.18)$$

Sphericity indicates the homogeneity of the final state energy distribution. A high sphericity value indicates an isotropic event, while a low sphericity value indicates a dijet event.

Aplanarity, A , is defined as

$$A = \frac{3}{2}\lambda_3. \quad (9.19)$$

Aplanarity values of 0 indicate a planar event, while a isotropic event has an aplanarity value of 0.5. Since aplanarity and sphericity are both calculated from the eigen values of the linearized sphericity tensor, they are correlated.

The event shape variable C is defined as

$$C = 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3). \quad (9.20)$$

The event shape variable C ranges from 0 to 1. A small value indicates a dijet event, while a large value indicates a 3 jet event.

The event shape variable D is defined as

$$D = 27(\lambda_1\lambda_2\lambda_3). \quad (9.21)$$

The event shape variable D ranges from 0 to 1. A small value indicates a dijet event, while a large value indicated a four jet event.

The transverse linearized sphericity tensor is defined as

$$M_{xy} = \frac{1}{\sum_i |\vec{p}|} \sum_i \frac{1}{|\vec{p}|} \begin{pmatrix} p_{x,i}^2 & p_{x,i}p_{y,i} \\ p_{y,i}p_{x,i} & p_{y,i}^2 \end{pmatrix}. \quad (9.22)$$

The transverse sphericity is defined as

$$S_{\perp} = \frac{2\lambda_2}{\lambda_1 + \lambda_2}. \quad (9.23)$$

A high transverse sphericity value indicates an isotropic event, while a low transverse sphericity value indicates a dijet event.

Other ways to describe events involved studying the momentum imbalance in the event and studying the placement of the missing energy with respect to the various jets in the event.

The p_T balance of the leading and subleading large R jet, $p_T^{\text{bal}}(J_1, J_2)$, is defined as

$$p_T^{\text{bal}}(J_1, J_2) = \frac{|\vec{p}_T(J_1) + \vec{p}_T(J_2)|}{|\vec{p}_T(J_1)| + |\vec{p}_T(J_2)|}, \quad (9.24)$$

where the leading large R jet is denoted as J_1 and the subleading large R jet is denoted as J_2 .

The p_T balance of the closest small R jet to the E_T^{miss} and the farthest small R jet to the E_T^{miss} is defined as

$$p_T^{\text{bal}}(j_{\min}, j_{\max}) = \frac{|\vec{p}_T(j_{\min}) + \vec{p}_T(j_{\max})|}{|\vec{p}_T(j_{\min})| + |\vec{p}_T(j_{\max})|}, \quad (9.25)$$

where the closest small R jet to the E_T^{miss} is denoted as $j_{\min}$ and the farthest small R jet to the E_T^{miss} is denoted as $j_{\max}$.

The angle between the closest large R jet to the E_T^{miss} is defined as $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, where

$$\Delta\phi_{\min}(J, E_T^{\text{miss}}) = \phi_{J_{\min}} - \phi_{E_T^{\text{miss}}}. \quad (9.26)$$

The angle between the farthest large R jet to the E_T^{miss} and the E_T^{miss} is defined as $\Delta\phi_{\max}(J, E_T^{\text{miss}})$, where

$$\Delta\phi_{\max}(J, E_T^{\text{miss}}) = \phi_{J_{\max}} - \phi_{E_T^{\text{miss}}}. \quad (9.27)$$

The difference between $\Delta\phi_{\max}(J, E_T^{\text{miss}})$ and $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is $\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}})$, defined as

$$\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}}) = \Delta\phi_{\max}(J, E_T^{\text{miss}}) - \Delta\phi_{\min}(J, E_T^{\text{miss}}). \quad (9.28)$$

These variables can similarly be defined for small R jets.

The angle between the closest small R jet to the E_T^{miss} is defined as $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, where

$$\Delta\phi_{\min}(j, E_T^{\text{miss}}) = \phi_{j_{\min}} - \phi_{E_T^{\text{miss}}}. \quad (9.29)$$

The angle between the farthest small R jet to the E_T^{miss} and the E_T^{miss} is defined as $\Delta\phi_{\max}(j, E_T^{\text{miss}})$, where

$$\Delta\phi_{\max}(j, E_T^{\text{miss}}) = \phi_{j_{\max}} - \phi_{E_T^{\text{miss}}}. \quad (9.30)$$

The difference between $\Delta\phi_{\max}(j, E_T^{\text{miss}})$ and $\Delta\phi_{\min}(j, E_T^{\text{miss}})$ is defined as

$$\Delta\phi_{\text{maxmin}}(j, E_T^{\text{miss}}) = \Delta\phi_{\max}(j, E_T^{\text{miss}}) - \Delta\phi_{\min}(j, E_T^{\text{miss}}). \quad (9.31)$$

The average distance in the $\eta - \phi$ plane between all small R jets in the event is termed $|\Delta R_{\text{avg}}(j)|$. The distance in the $\eta - \phi$ plane between the leading small R jet and subleading small R is termed $|\Delta R(j_1, j_2)|$.

9.6 Scalar Cascades Neural Network Results

As a reminder, the Scalar Cascades analysis is undertaken for Scalar Cascade Benchmark SC1, defined in Table 8.4. As noted previously, the neural network for all analyses in this thesis contains four parts: a particle flow network for the leading reclustered jet, a particle flow network for the second leading reclustered jet, a particle flow network for all small R jets in the event, and an event level tagger. The sole variables inputted to the event level tagger of the neural network for scalar cascades are the number of tracks in the second leading reclustered jet, the event shape variable C , and the mass of the second leading reclustered jet. As seen in Figure 9.6, and as noted earlier, the track multiplicity in the second leading reclustered jet is by far the strongest discriminating variable for scalar cascades, with its peak of 100 tracks clearing the bulk of background.

It is worthwhile to note that the total number of scalar cascade events present in the plots in Figure 9.6 is 2.72 events. It is not such, therefore, that a cut on high track multiplicity would yield a plethora of signal events. The small cross section combined with the low trigger efficiency of the scalar cascades are factors that make this challenging analysis even more challenging.

As seen in Figure 9.6, scalar cascades exhibit a peak at an event shape variable C value of 0.03, a very low value indicating a strong dijet-like structure to the event. Background exhibits more of a plateau-like distribution predominantly from event variable C values ranging from 0.08 to 0.75. Adding the mass of the second leading reclustered jet to the event level variable tagger is useful because the exceptional background events with high track multiplicity will likely be a jet with a mass at the tail end of the mass spectrum for background, whereas a scalar cascades have high track multiplicity and are predominantly peaked around masses of 100 GeV. (These signal distributions are plotted for Scalar Cascade Benchmark SC1, but they can be applicable to other scalar cascade benchmarks depending on the specific parameters of the model, such as the number of final state particles in the model and the number of scalars in the model.)

In order to determine that the number of tracks in the second leading reclustered jet is the most discriminating variable for scalar cascades, the author calculated and plotted, via the analysis framework created by the author and described in 7.3.1, hundreds of potential analysis variables to compare the difference in shape between signal and background. The number of tracks in the second leading reclustered jet was the most distinctive variable. The event shape variable C , though not nearly as powerful, was another strong discriminating variable. The reason only three variables appear in the event level tagger for scalar cascades is because adding more variables diluted the strength of the neural network because no variable is as discriminating as the number of tracks in the second leading reclustered jet. Three variables, instead of one variable, were included because the variable that the neural network was decorrelating against was the number of tracks in the second leading reclustered jet. Adding two additional variables to the event level tagger acted as a way to make the task of decorrelating the neural network simpler.

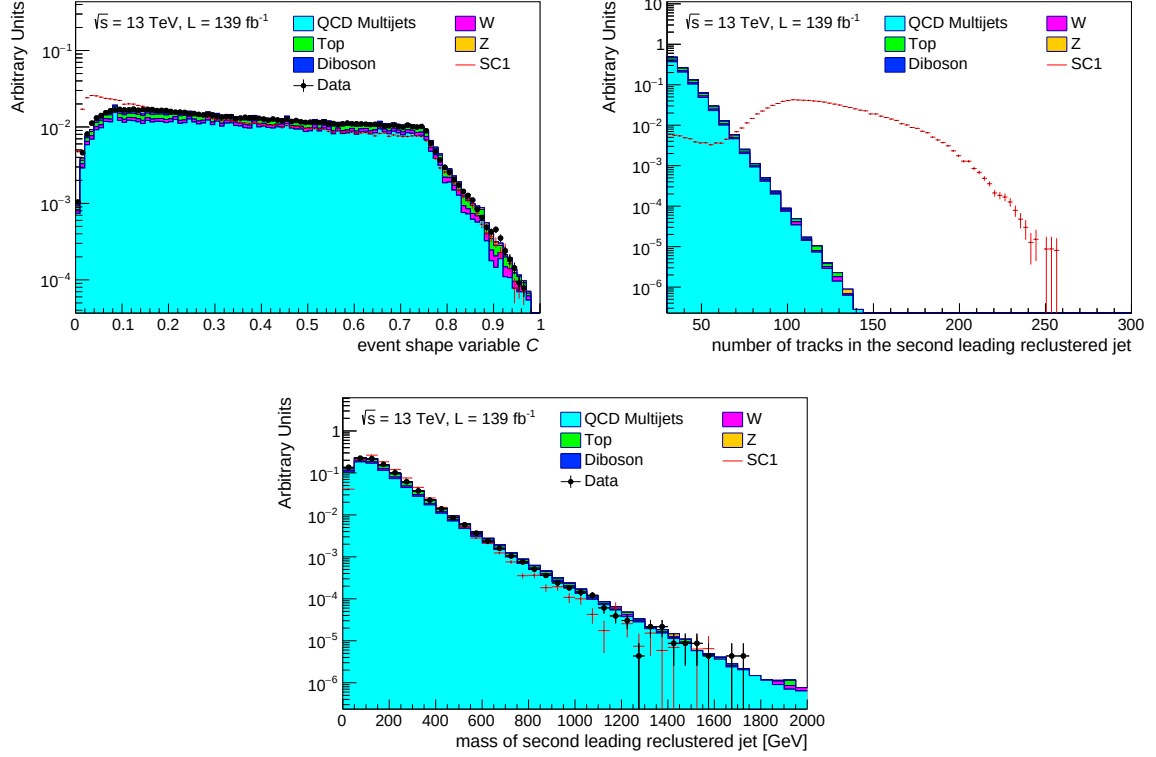


Figure 9.6 *Variables inputted into event level tagger of NN for Scalar Cascades. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied to these variables. Data and Monte Carlo distributions are overlaid for the distributions of event shape variable C and the mass of the second leading reclustered jet.*

This analysis employs a data-driven background estimation technique. However, background Monte Carlo is used in the training of the neural network, as well as cross-checks in the background estimation. Therefore, comparisons of Monte Carlo to data are performed prior to the use of neural networks to show reasonable agreement of MC to data, such that behavior after the use of the neural network can sensibly be compared. As this is a data-driven background estimation technique, the agreement does not have to be perfect. Qualitatively, Figure 9.6 shows that the shape of data is broadly consistent with background Monte Carlo for the event shape variable C and the mass of the second leading reclustered jet for the scalar cascade model. The inclusive distribution of the number of tracks in the second leading reclustered jet in Figure 9.6 was not overlaid with data because the number of tracks in the second leading reclustered jet is used as the decorrelated variable in the training of the neural network for scalar cascades. The next chapter will discuss the choice to require greater than 80 tracks for signal region definition. In order to look at the comparisons of Monte Carlo and data outside of the signal region, the distributions

of the number of tracks in the second leading reclustered jet for data and Monte Carlo were compared for track multiplicities less than 80 tracks in the second leading reclustered jet, as seen in Figure 9.7.

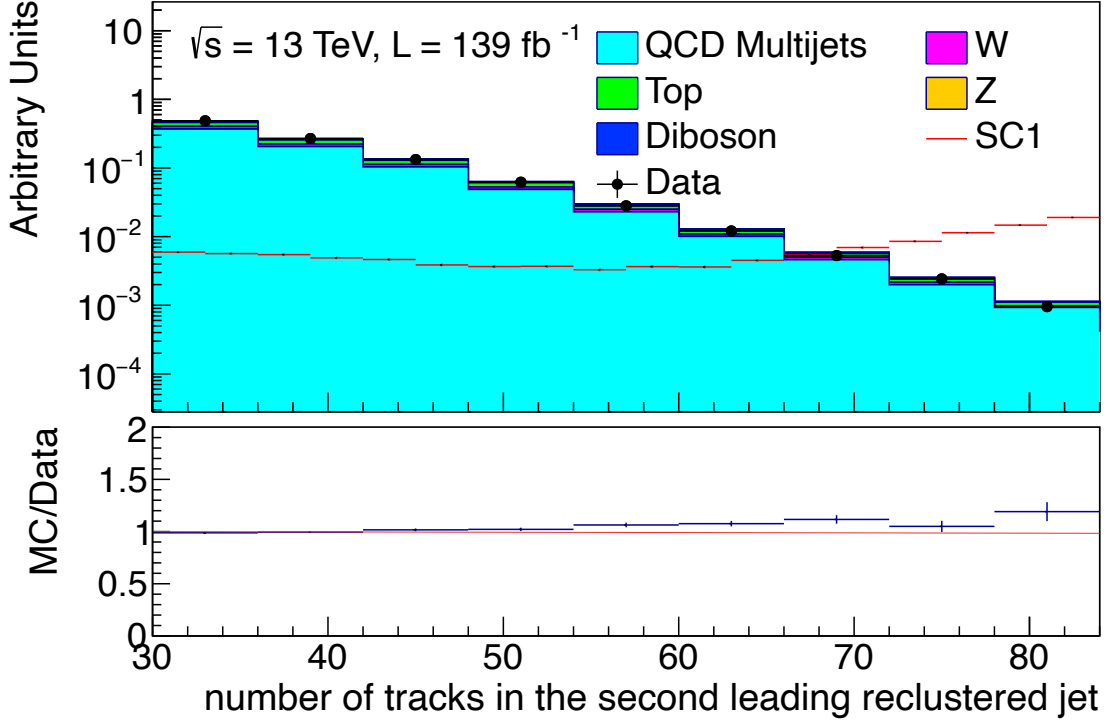


Figure 9.7 *Top plot: Comparison of Monte Carlo and data distributions of number of tracks in the second leading reclustered jet for Scalar Cascades. The comparisons are made for track multiplicities smaller than 80 tracks in order to avoid the signal region. This distribution is normalized. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied. Bottom plot: The Monte Carlo to Data ratio of the number of tracks in the second leading reclustered jet. This distribution was created from the normalized distributions in the top plot.*

The Monte Carlo to data ratio stays at a value of 1 within errors until greater than 55 tracks in the second leading reclustered jet. Beyond 55 tracks, the Monte Carlo and data distributions stay within roughly 10% of each other, with the exception being in the last bin, where the discrepancy is slightly more than 2σ (or about 20%). This small discrepancy is not a concern, however, as the background estimation is data-driven.

Another cross-check is the comparison of Monte Carlo and data of the η distributions of the leading and subleading reclustered jets, as seen in Figure 9.8 and 9.9.

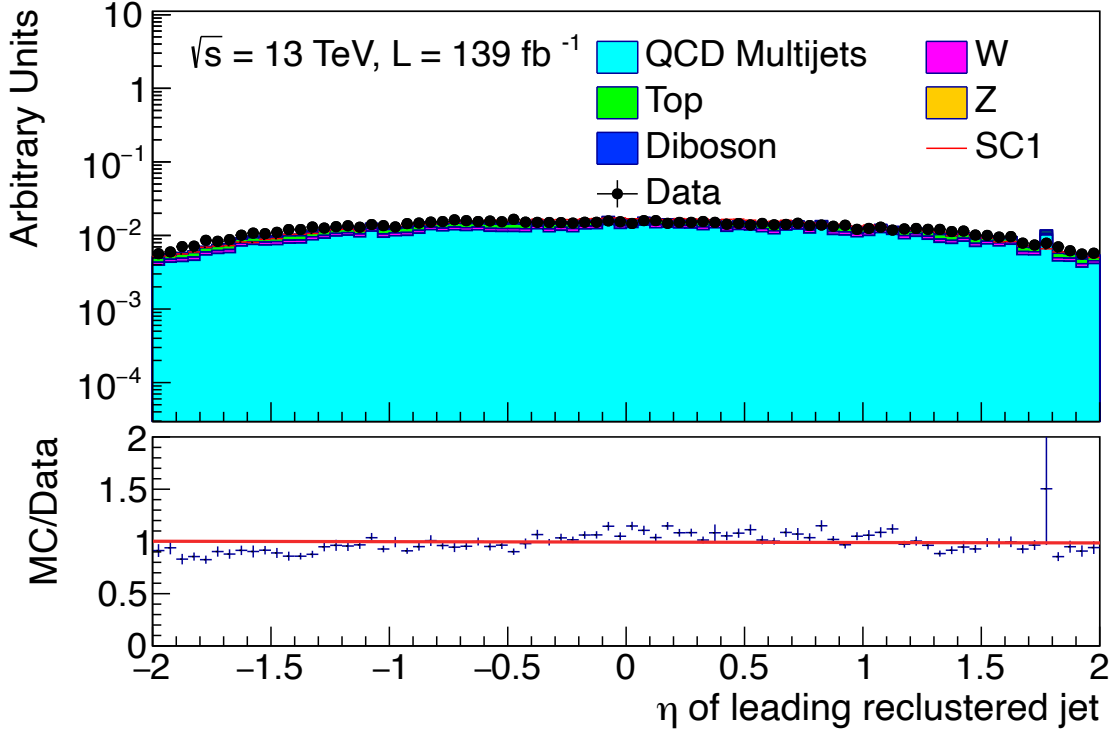


Figure 9.8 *Top plot: Data and Monte Carlo comparison for the η distributions of the leading reclustered jets for Scalar Cascades. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the η of the leading reclustered jets for Scalar Cascades. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied to these variables.*

The distributions of Monte Carlo and data for the η of the leading reclustered jet are within either 2σ or 20% for all bins. This is similarly true for the distributions of Monte Carlo and data for the η of the subleading reclustered jet. This is an acceptable level of agreement. The Monte Carlo to data ratio for the η of the leading reclustered jet has a spike at roughly $\eta = 1.78$, while the ratio for the η of the second leading reclustered jet has a spike at roughly $\eta = -0.98$. Each of these two spikes can be attributed to an anomalously weighted Monte Carlo event. As these spikes still yield ratio values of 1 within 1σ , they can safely be ignored.

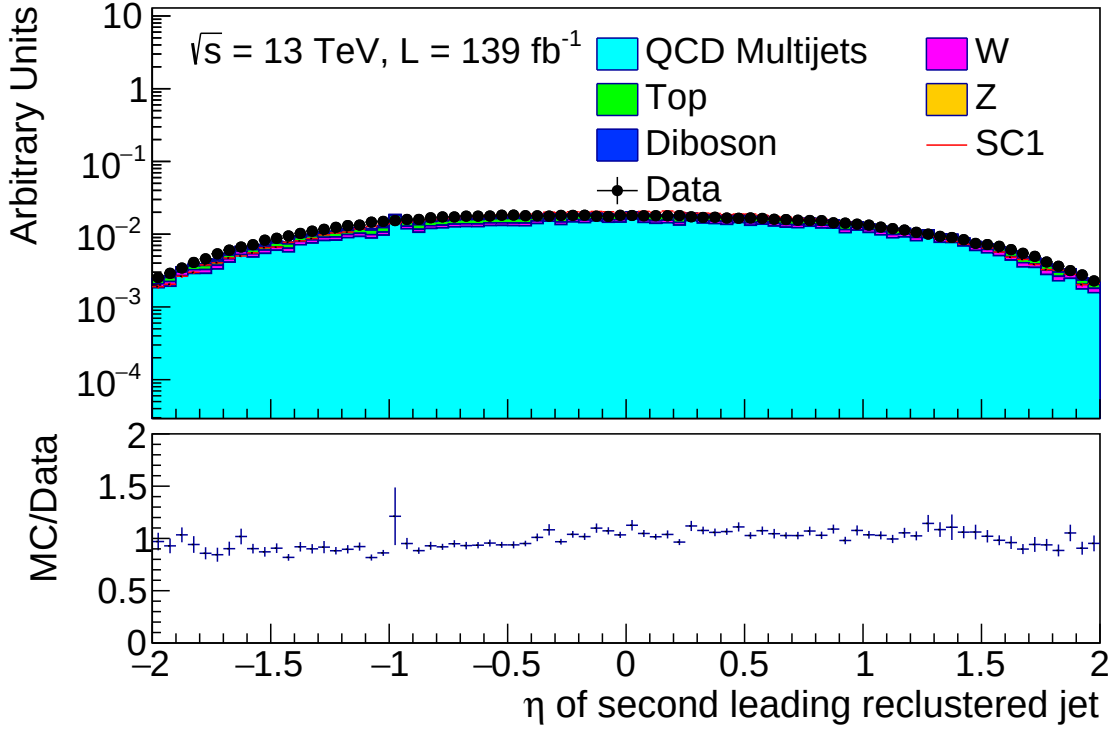


Figure 9.9 *Top plot: Data and Monte Carlo comparison for the η distributions of the second leading reclustered jets for Scalar Cascades. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the η of the second leading reclustered jets for Scalar Cascades. The Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 have been applied to these variables.*

As discussed previously, the p_T , rapidity, and ϕ of all of the tracks in the leading reclustered jet were provided as inputs to the leading reclustered jet particle flow network, while the p_T , rapidity, and ϕ of all of the tracks in the second reclustered jet were provided as inputs to the second leading reclustered jet particle flow network, and the p_T , rapidity, and ϕ of all the small R jets in the event were provided as inputs to the particle flow network of all small R jets in the event. Shown in Figure 9.10 and Figure 9.11, these distributions look as expected before the application of machine learning. The p_T distribution of track particles in the second leading reclustered jet is more sharply peaked under 20 GeV for signal than background. Similarly, the p_T distribution of all small R jets in the event is more sharply peaked at 20 GeV for signal than background. The slight step at 200 GeV in the distribution of all small R jets in the event is due to the requirement that the leading small R jet be at least 200 GeV. The rapidity of the track particles in the second leading reclustered jet for signal is more sharply peaked at central values of rapidity for signal than background.

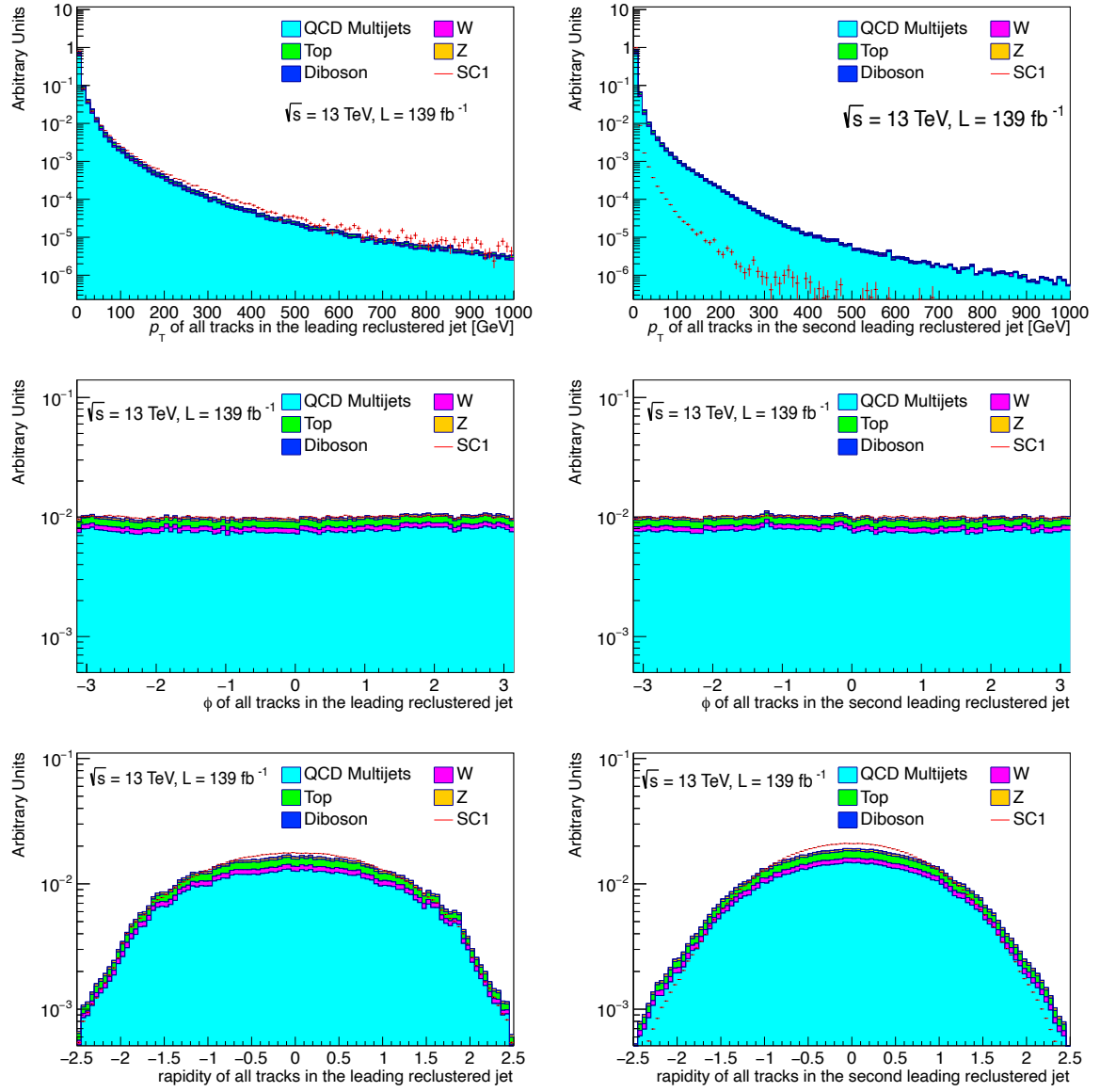


Figure 9.10 *Kinematics of all tracks within the first and second leading reclustered jets for Scalar Cascades and Background with Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 applied*

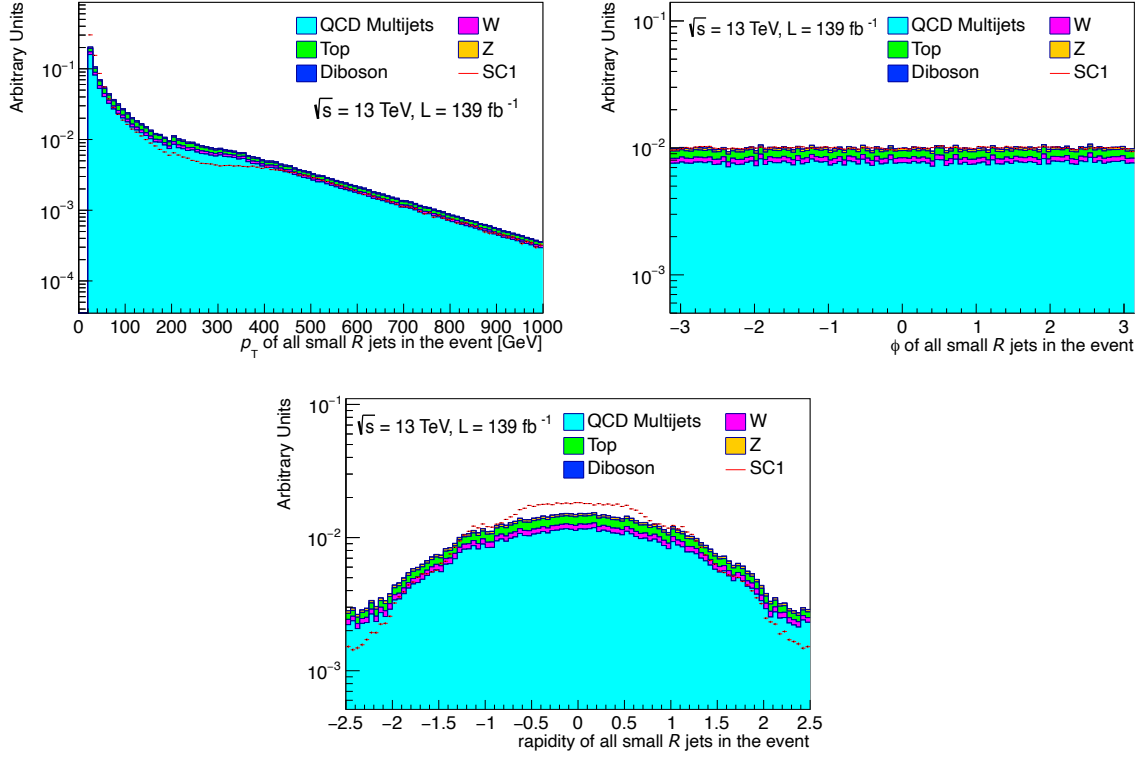


Figure 9.11 *Kinematics of all small R jets in the event for Scalar Cascades and Background with Pre-ML Scalar Cascade Analysis Cuts from Table 8.4 applied*

The rapidity of the small R jets in the second leading reclustered jet are similarly more sharply peaked at central values of rapidity than background. Otherwise the remaining 6 of these 9 distributions themselves do not provide any significant signal versus background discrimination. However, the power of particle flow networks is the ability to examine the structure within each jet or - in the case of the small R jet particle flow network, within the jet - rather than looking at each of these low level kinematic inputs as event level observables.

Below is a presentation and discussion of the results of the neural network with running with a DisCo λ parameter of 50. The next chapter will provide a discussion regarding the choice to use 50 for the DisCo λ parameter. The performance of a neural network can be assessed with a receiver operating characteristic (ROC) curve. A ROC curve is created using two quantities: the true positive rate and the false positive rate. The true positive rate (tpr) is the fraction of the total number of events in which signal events are correctly identified as signal events. The false positive rate (fpr) is the fraction out of the total number of events in which background events are incorrectly marked as signal events. In Figure 9.12, tpr is plotted on the x

axis, while $1 - \text{fpr}$ is plotted on the y axis. Extremely high background rejection is achieved for roughly 80% signal retention and staying well above 99% background rejection through over 95% signal retention. This points to the robustness of the neural network.

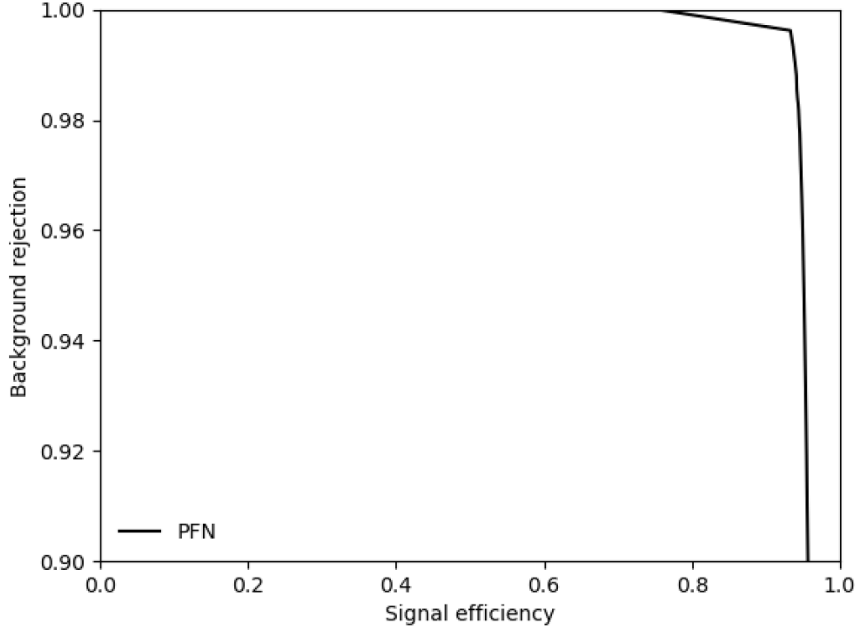


Figure 9.12 *ROC Curve of Background Rejection vs. Signal Efficiency for the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet.*

Two metrics against which to access the stability of the neural network are the evolution of the accuracy and the loss of the neural network. For these metrics, it is customary to compare the accuracy or loss of the training sample with the respective quantity of the validation samples. In this case (and is the case for the neural networks for noncommutative black holes and non-minimal dark sectors), the training sample consisted of roughly 100k signal events and 100k background events. Each signal event was weighted according to its total predicted cross section and scaled to a luminosity reflecting the full run 2 dataset. The weight of each signal event was then summed to create a total signal weight for the signal sample. The weight of each signal event was then divided by the total signal weight for the signal sample. To generate the 100k sample of background events for the training sample, each Monte Carlo background dataset for the analysis (from Table 8.3) was slimmed by a factor of 100 and the weight of each event was then multiplied by 100. Each background event was then weighted according to its total predicted cross section and

scaled to a luminosity reflecting the full run 2 dataset. Then, the same normalization procedure that was applied to signal was applied to background events, such that the total weight of background events is normalized to 1. With this construction, the signal sample and background sample are of equivalent significance to the training. This normalization procedure was used for all neural networks in this thesis. As a reminder, for all neural networks in this thesis, the training was done through backpropagation, specifically mini-batch gradient descent, which was described in 9.1.

The validation sample for this neural network was millions of MC simulated events from all background processes properly weighted by cross section and scaled to the full Run 2 luminosity. The purpose for having different samples to train and validate the neural network is to ensure that the neural network does not generalize poorly when running on different data than that which it was trained on. Overtraining occurs when the neural network learns trends specific to the training sample and not the whole data sample. Undertraining occurs when the neural network performs poorly on the training dataset. In Figure 9.13, it is positive to see that the accuracy of the neural network for both the training and validation datasets improves as the number of epochs that the neural network is trained for increases. The NN was trained for 50 epochs. The accuracy of over 90% suggests a strong ability to distinguish signal from background.

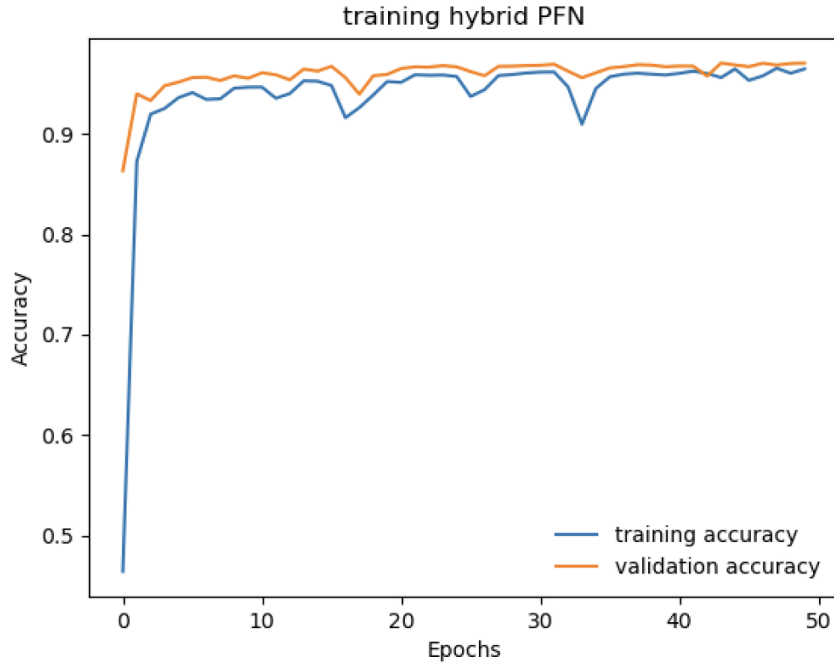


Figure 9.13 *Evolution of NN validation accuracy and training accuracy as a function of epochs for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet.*

In Figure 9.14, the loss of the neural network decreases over time for both the training sample and the validation sample, which is the desired behavior, as the goal in training is to minimize the loss.

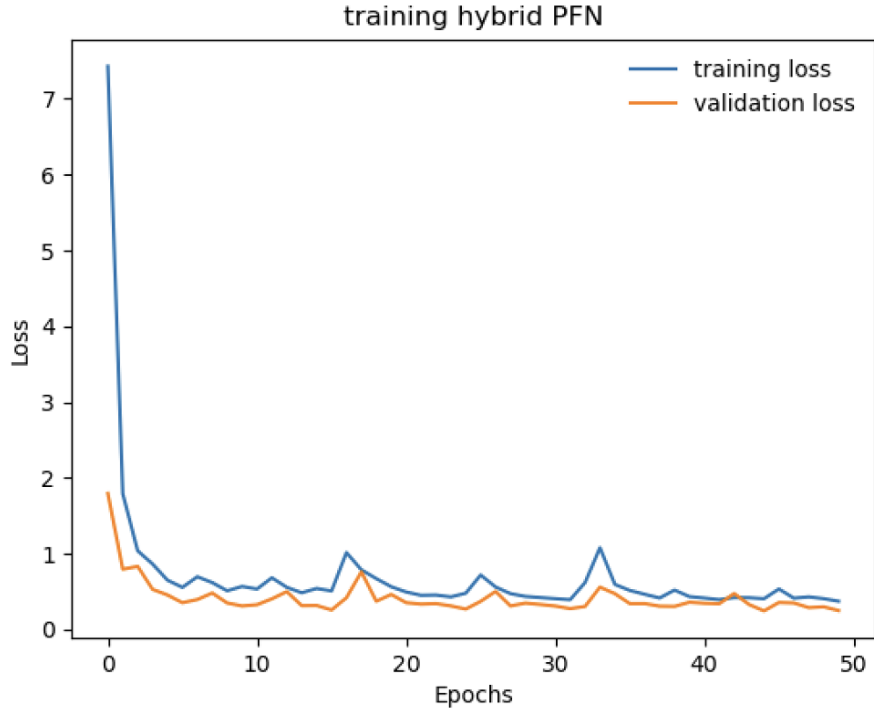


Figure 9.14 *Evolution of NN validation loss and training loss as a function of epochs for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet.*

The ability of the neural network to correctly identify signal from background can be seen by plotting the output labels, 0 (background) or 1 (signal), for signal Monte Carlo simulated events. In Figure 9.15, it can be seen that almost all of the signal is peaked in the bin ranging from 0.99 neural network output to 1.0 neural network output. (A green circle highlights this bin.) This suggests that the neural network has a very strong ability to differentiate signal from background.

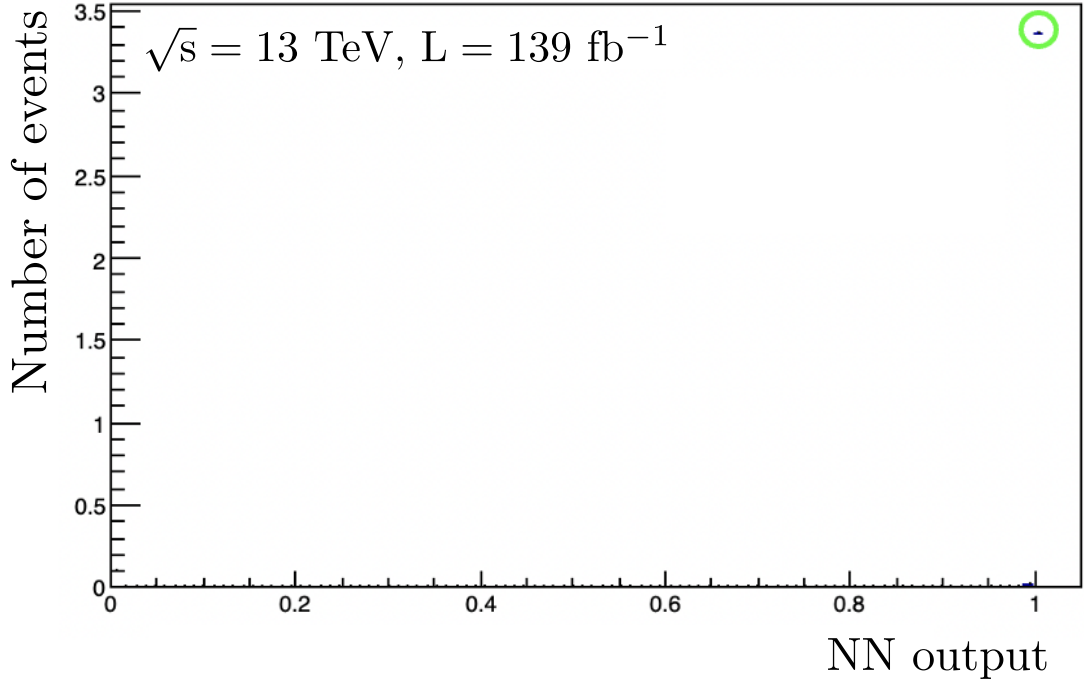


Figure 9.15 *NN prediction on signal events for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet. (A green circle highlights the bin containing the majority of the signal.)*

Similarly, looking at the plot of output labels for background Monte Carlo simulated events in Figure 9.16, it can be seen that the vast majority of background events are encapsulated in the bin of the neural network that ranges from 0 to 0.01. (A green circle highlights this bin.) This again suggest a robust neural network. Notably, the amount of background in the bin ranging from a neural network output of 0.99 to a neural network output of 1.0 is larger than the amount of signal in that bin. This necessitates cutting on an additional variable. This additional variable will be shown in the next chapter to be the number of tracks in the second leading reclustered jet.

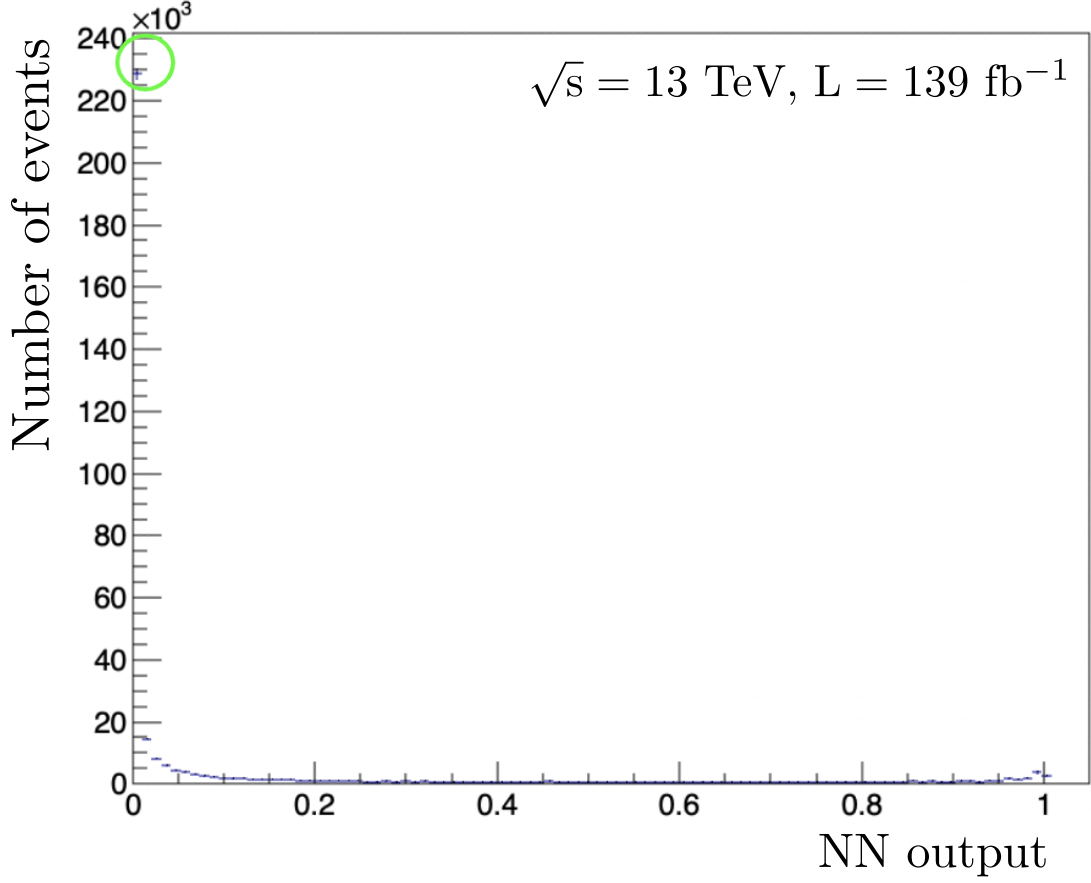


Figure 9.16 *NN prediction on background events for NN using the Scalar Cascade Benchmark SC1 with DisCo hyperparameter λ value of 50 and SC cuts from Table 8.4 applied. The decorrelated variable is the number of tracks in the second leading reclustered jet. (A green circle highlights the bin containing the majority of the background.)*

9.7 Noncommutative Black Holes and Non-minimal Dark Sector Neural Networks

The variables inputted to the event level tagger of the neural network for non-commutative black holes and non-minimal dark sectors are the five event shapes sphericity, aplanarity, event shape variable C , event shape variable D , and transverse sphericity, the number of small R jets above 20 GeV, $p_T^{\text{bal}}(j_{\min}, j_{\max})$, the $p_T^{\text{bal}}(J_1, J_2)$, $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(J, E_T^{\text{miss}})$, $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, η of leading reclustered jet, η of second leading reclustered jet, number of tracks in the leading reclustered jet, number of tracks in the second leading reclustered jet, number of small R jets in the leading reclustered jet, number of small R jets in the second

leading reclustered jet, $|\Delta R_{\text{avg}}(j)|$, and $|\Delta R(j_1, j_2)|$. In order to select these variables, the author calculated and plotted, via the analysis framework created by the author and described in 7.3.1, hundreds of potential analysis variables. The difference in shape between signal and background were compared and variables offering the most pronounced difference in shape were chosen to be included in the event level tagger.

This section describes the NN taggers for noncommutative black holes and non-minimal dark sectors at 200 GeV and then, for some of the benchmarks at additional cuts. This serves as a useful way to understand the general behavior of the models at fewer cuts, how the behavior changes with additional cuts, discuss the challenges faced by various models, and then segue into the final analysis results obtained for the benchmarks that undergo additional cuts.

9.7.1 Noncommutative Black Holes and Non-minimal Dark Sector Neural Networks - 200 GeV E_T^{miss} cut

In Figure 9.17, noncommutative black holes and nonminimal dark sectors exhibit a higher fraction of events with mild (around 0.4) to high (greater than 0.7) sphericity values than background, indicating that the noncommutative black holes and nonminimal dark sectors have a higher fraction of more isotropic events than background. This holds true for the transverse sphericity spectrum, in which there is an even more pronounced difference between the fraction of noncommutative black holes and nonminimal dark sectors and background with mild to high transverse sphericity values, indicating a higher fraction of isotropic events than background. The background aplanarity falls off much more sharply after peaking in the lowest aplanarity bin of 0 to 0.2 than the noncommutative black hole and nonminimal dark sector aplanarity distributions do, despite them also peaking in the lowest bin. The nonminimal dark sector benchmark NMDS3, which is the non-minimal dark sector benchmark with the largest mass spacing between ensemble constituents and, therefore, the most energetic decays, notably plateaus, rather than peaks, from 0 to 0.5 in aplanarity. This suggests that a higher fraction of background events exhibit a dijet-like behavior, as denoted by low aplanarity values, than noncommutative black hole and non-minimal dark sectors. Noncommutative black holes and nonminimal dark sectors exhibit a higher fraction of events with higher event shape variable C , event shape variable D , and transverse sphericity values, indicating a more three jet-like structure than background, a more four jet-like structure than background, and a more isotropic distribution than background, respectively. Qualitatively, the

shape of Monte Carlo and data is observed to be broadly consistent for aplanarity, event shape variable C , event shape variable D , sphericity, and the number of small R jets above 20 GeV. Data is not overlaid in the inclusive transverse sphericity distribution, as transverse sphericity will serve as a decorrelated variable for one of the nonminimal dark sector benchmarks. (As a reminder, in addition to QCD multijets, the background processes used in the analyses in this thesis include both the hadronic and leptonic decays of Z bosons, W boson, and top quarks. The analyses veto leptons and muons above 20 GeV, but the leptonic decays of background processes are included to account for events which do not fail the lepton veto. The list of background samples used can be found in the Appendix, as well as in Table 8.3).

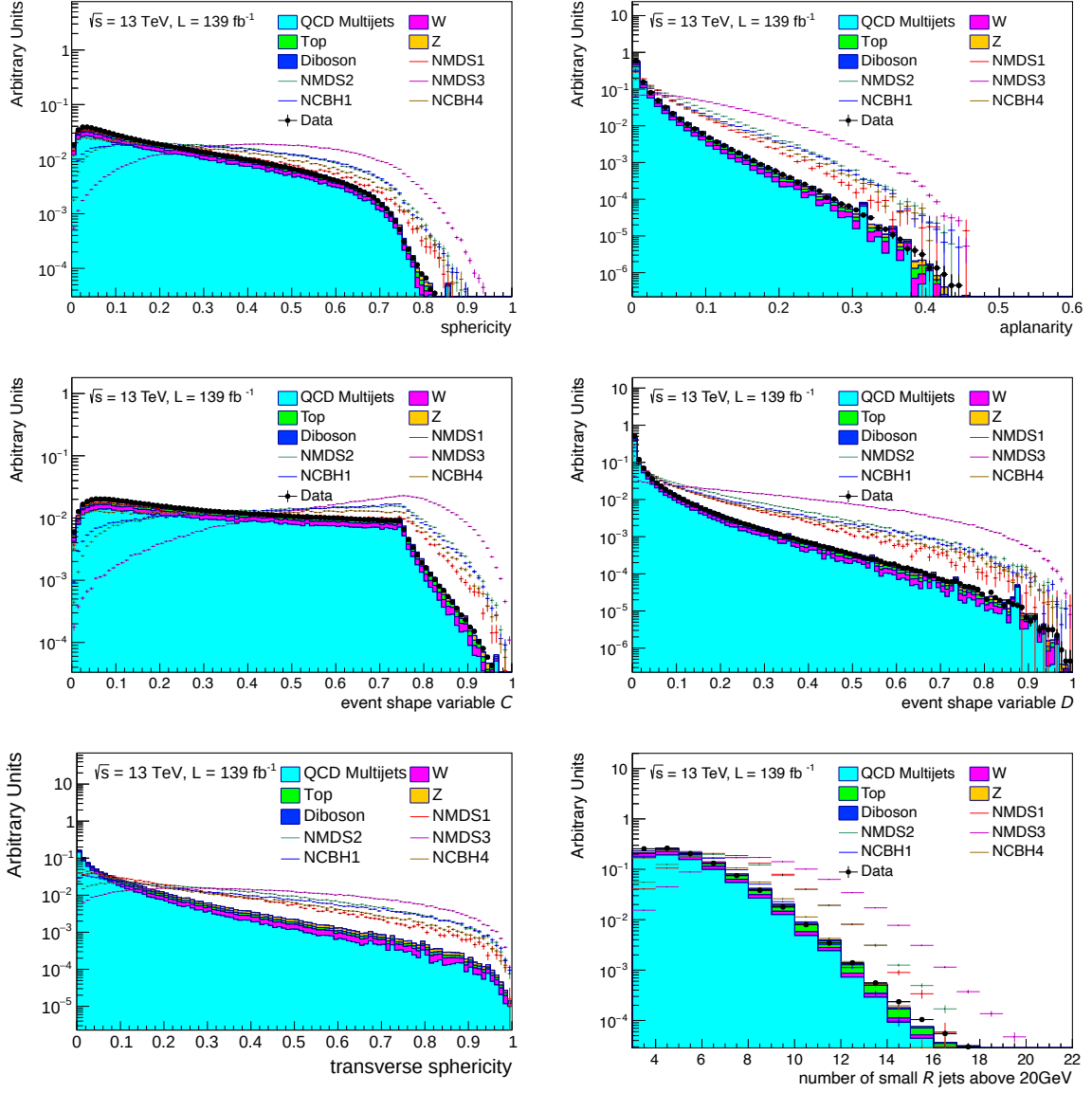


Figure 9.17 *Discriminating Variables for Noncommutative Black Hole and Nonminimal Dark Sector with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. Monte Carlo and Data distributions are overlaid for all variables except transverse sphericity, which serves as a decorrelated variable for a non-minimal dark sector benchmark.*

Both $p_T^{\text{bal}}(j_{\min}, j_{\max})$ and $p_T^{\text{bal}}(J_1, J_2)$ offer discriminatory power as noncommutative black holes and nonminimal dark sectors peak at higher values than background. Both the distribution of $\Delta\phi_{\min}(j, E_T^{\text{miss}})$ and the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ indicate that a higher fraction of events for noncommutative black holes and non-minimal dark sectors have a greater separation between the E_T^{miss} and the closest jet, be it small R jets or large R jets.

From the distribution of $\Delta\phi_{\text{maxmin}}(j, E_T^{\text{miss}})$ and the distribution of $\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}})$, it can be seen that noncommutative black holes and non-minimal dark sectors typically have smaller values of $\Delta\phi_{\text{maxmin}}(j, E_T^{\text{miss}})$ and $\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}})$ than background.

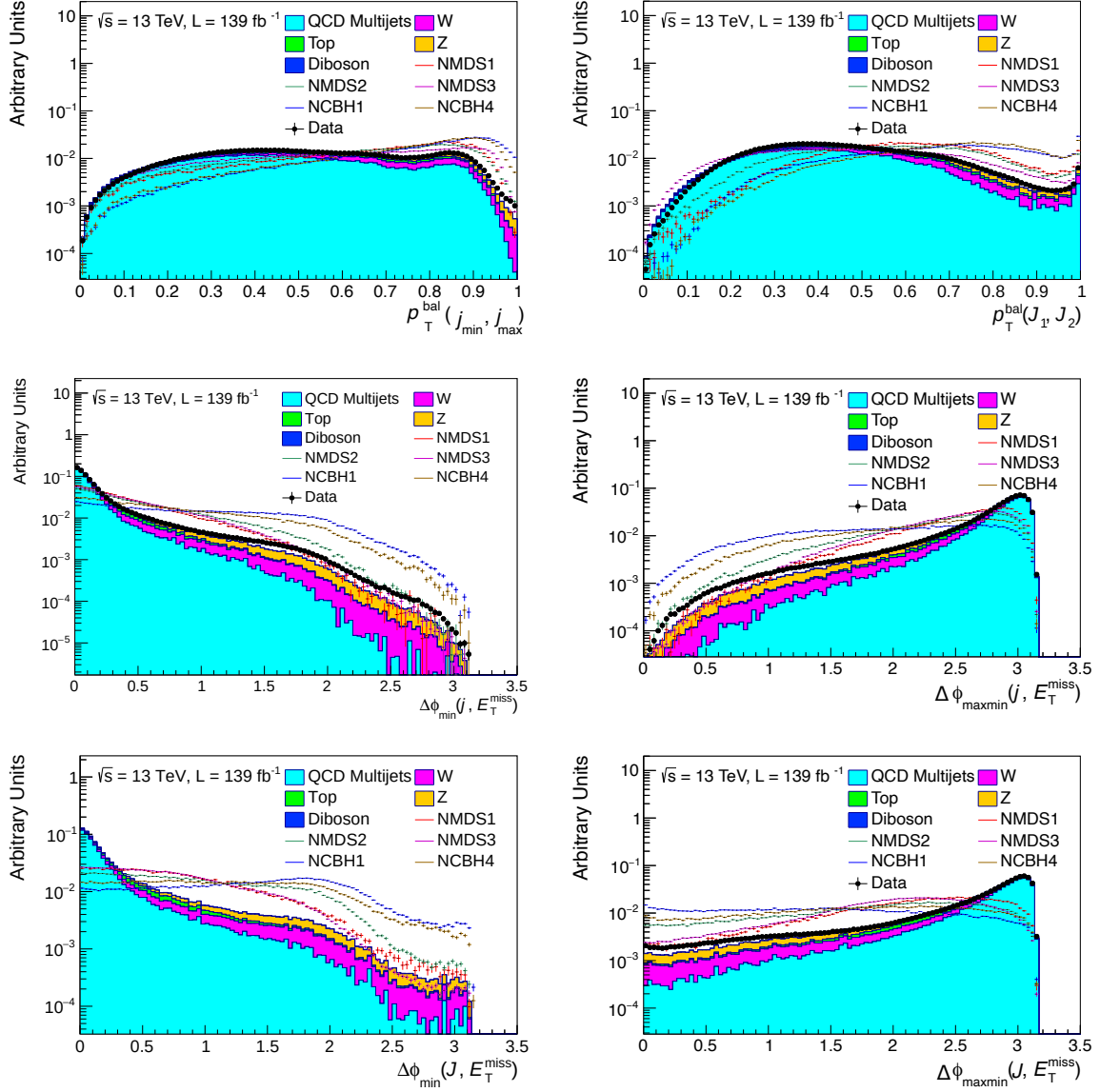


Figure 9.18 *Discriminating Variables for Noncommutative Black Hole and Nonminimal Dark Sector with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. Monte Carlo and Data distributions are overlaid for all variables except $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$, which serves as a decorrelated variable for noncommutative black holes.*

Qualitatively, data and Monte Carlo distributions in Figure 9.18 are broadly consistent in shape, with any deviations being restricted to the tails of the

distributions (as seen in the distributions of $\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}})$, $\Delta\phi_{\text{maxmin}}(j, E_T^{\text{miss}})$, and $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$).

The $|\Delta R(j_1, j_2)|$ peaks at lower values for noncommutative black holes and scalar cascades, suggesting that the leading and subleading small R jets are spaced closer together for these exotic signals than for background. Noncommutative black holes and non-minimal dark sectors have a higher fraction of events with lower values of $|\Delta R_{\text{avg}}(j)|$ than background, suggesting that the small R jets are more closely spaced together than for background. The η of the leading reclustered jet and the subleading reclustered jet offer discriminating power as the leading and subleading jets for noncommutative black holes and non-minimal dark sectors peak more sharply at central values than background. A higher fraction of events have higher track multiplicity values in the leading and subleading reclustered jets for noncommutative black holes and non-minimal dark sectors than background. Similarly, a higher fraction of events have a higher number of small R jets in the leading and subleading reclustered jets for noncommutative black holes and non-minimal dark sectors than background. Qualitatively, data and Monte carlo distributions in Figure 9.19 are broadly consistent in shape.

Figures 9.17, 9.18, and 9.19 show that all benchmarks for noncommutative black holes and non-minimal dark sectors have powerful discriminating power from background in variables that are sensitive to high jet multiplicity, such as event shapes, as well as variables that describe the momentum imbalance in the event and variables that describe the placement of the missing transverse momentum with respect to various jets in the event. This is expected as both noncommutative black holes and non-minimal dark sectors are characterized by high jet multiplicity. Noncommutative black holes yield a stable remant, which appears as E_T^{miss} while non-minimal dark sectors yield two stable fermions that appear as E_T^{miss} . Having the E_T^{miss} come primarily from one object is not the same phenomenological scenario as having the E_T^{miss} come primarily from two objects. However, these scenarios are a lot more similar than having the E_T^{miss} come from a higher multiplicity of particles. This explains why noncommutative black holes and non-minimal dark sectors distributions in variables related to the placement of the E_T^{miss} and the jets in the event are distinct, though broadly share similar features.

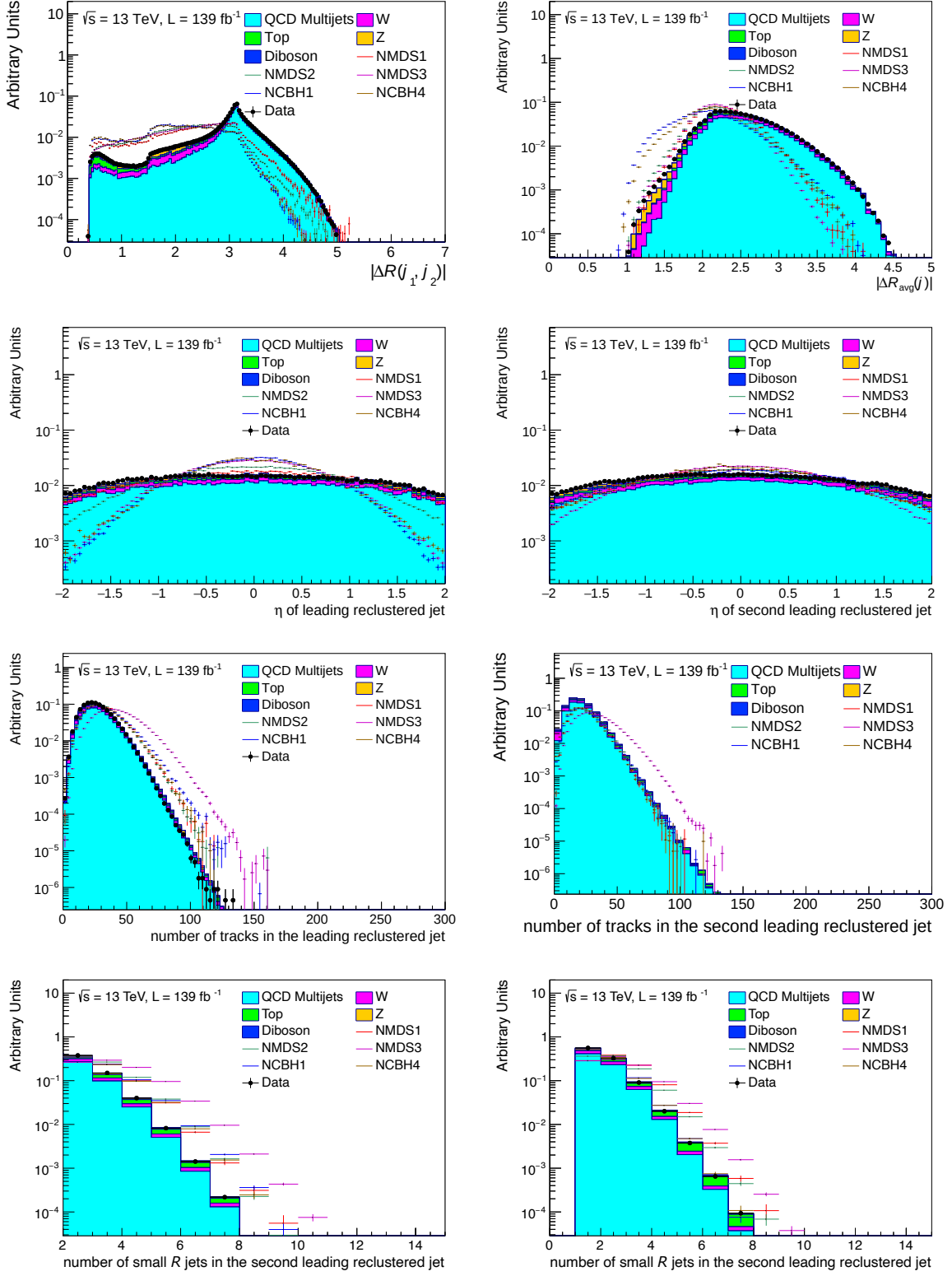


Figure 9.19 *Discriminating Variables for Noncommutative Black Hole and Nonminimal Dark Sector with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied*

The agreement between Monte Carlo and data can be examined quantitatively in control regions for the variables that will serve as the decorrelated variables for the various models. As discussed earlier, the control region for the number of tracks in the second leading reclustered jet is track multiplicity less than 80. The Monte Carlo to data comparisons for the track multiplicity in the second leading reclustered jet, with the basic noncommutative black hole and nonminimal dark sector cuts from Table 8.6 applied, in Figure 9.20, show agreement within 20% through 70 tracks, with the Monte Carlo to data ratio increasing slightly towards higher values of track multiplicity, as expected.

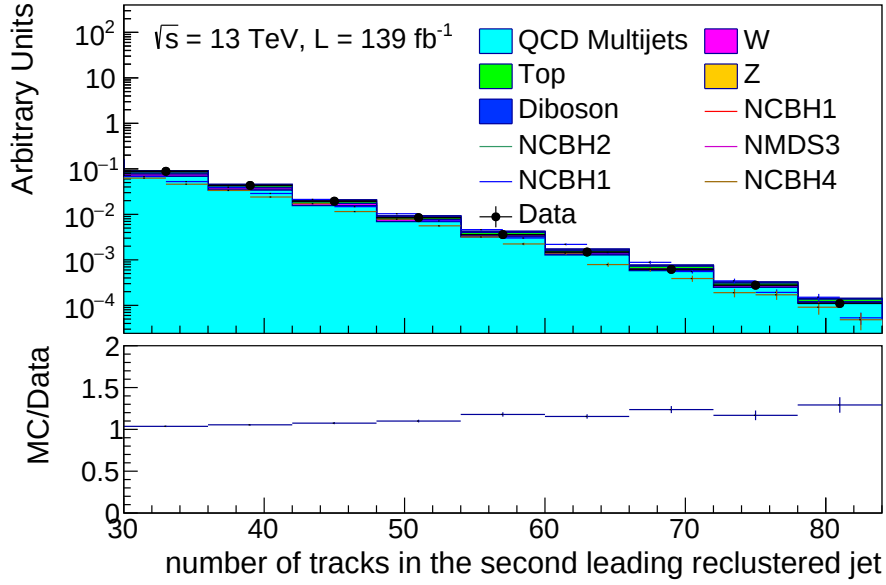


Figure 9.20 *Top plot: Data and Monte Carlo comparisons for the track multiplicity in the second leading reclustered jet with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the number of tracks in the second leading reclustered jet.*

The decorrelated variable for noncommutative black holes is $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. As will be discussed in the next chapter, the signal depleted region of phase space for which it is safe to compare Monte Carlo and data is $\Delta\phi_{\min}(J, E_T^{\text{miss}}) < 0.75$. Figure 9.21 shows that the Monte Carlo and data distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ agree within 5% for all but 4 bins, which agree within 10%.

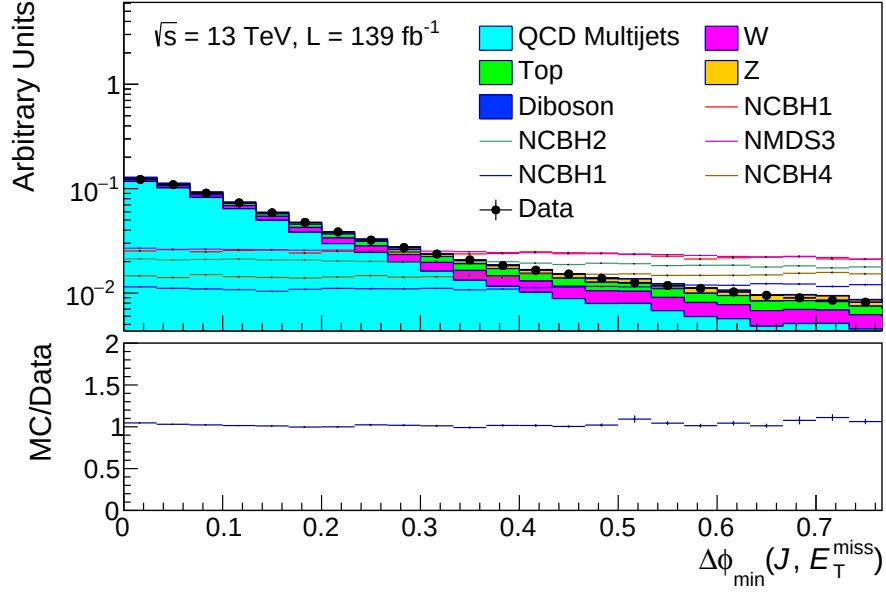


Figure 9.21 *Top plot: Data and Monte Carlo comparisons for $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$.*

The decorrelated variable for one of the nonminimal dark sector benchmarks, as will be discussed, is the transverse sphericity. In the next chapter, it will be shown that the control region for this variable is transverse sphericity values less than 0.2. Figure 9.22 shows that the Monte Carlo and data distributions for transverse sphericity agree within 10%.

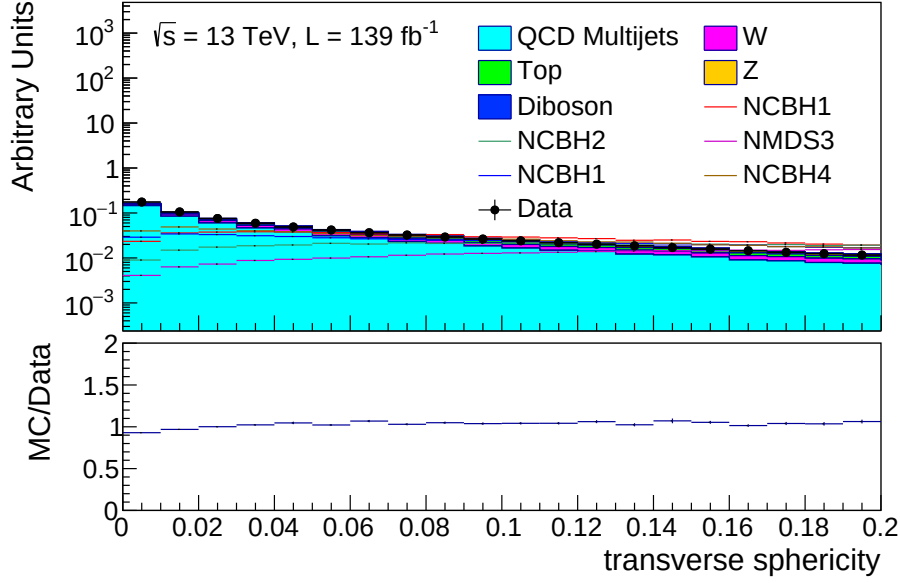


Figure 9.22 *Top plot: Data and Monte Carlo comparisons for the transverse sphericity with Basic Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts from Table 8.6 Applied. This distribution is normalized. Bottom plot: Monte Carlo to data ratio of the normalized distribution of the transverse sphericity.*

As with the scalar cascades, the p_T , rapidity, and ϕ of all of the tracks in the leading reclustered jet for noncommutative black holes and non-minimal dark sectors are provided as inputs to the leading reclustered jet particle flow network, while the p_T , rapidity, and ϕ of all of the tracks in the second reclustered jet are provided as inputs to the second leading reclustered jet particle flow network, and the p_T , rapidity, and ϕ of all the small R jets in the event are provided as inputs to the particle flow network of all small R jets in the event. Shown in Figure 9.23 and Figure 9.24, these distributions look as expected before the application of machine learning. The rapidity of the track particles in the leading and second leading reclustered jet are more sharply peaked at central values of rapidity than background. The reduction of events in these two distributions below rapidity = -2 and above rapidity = 2 is due to the analysis cut that restricts the η of the leading and subleading reclustered jets to $-2 < \eta < 2$.

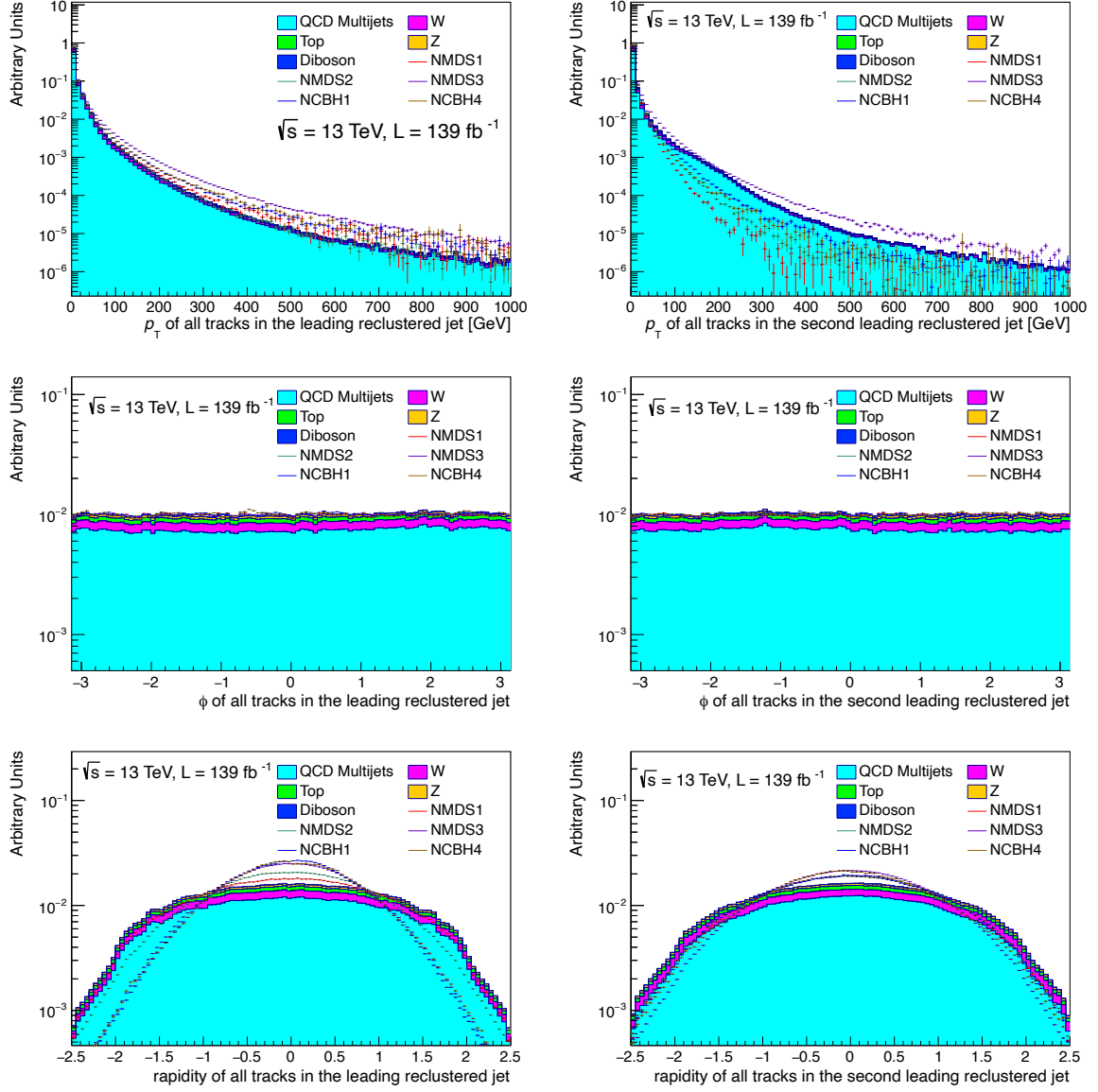


Figure 9.23 *Kinematics of all tracks within the first and second leading reclustered jets for Noncommutative Black Hole and Nonminimal Dark Sector with Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts Applied*

The rapidity of all small R jets in the event are similarly more sharply peaked at central values of rapidity than background. The reduction of events in these two distributions below rapidity = -2 and above rapidity = 2 is due to the analysis cut that restricts the η of the leading and subleading reclustered jets to $-2 < \eta < 2$. The step in the p_T distribution of all small R jets in the event is due to the analysis cut that requires the leading small R jet $p_T > 200$ GeV.

Otherwise the remaining 5 of these 9 distributions themselves do not provide any

notable signal versus background discrimination. However, as mentioned previously, the power of particle flow networks is the ability to examine the structure within each jet or - in the case of the small R jet particle flow network, within the jet - rather than looking at each of these low level kinematic inputs as event level observables.

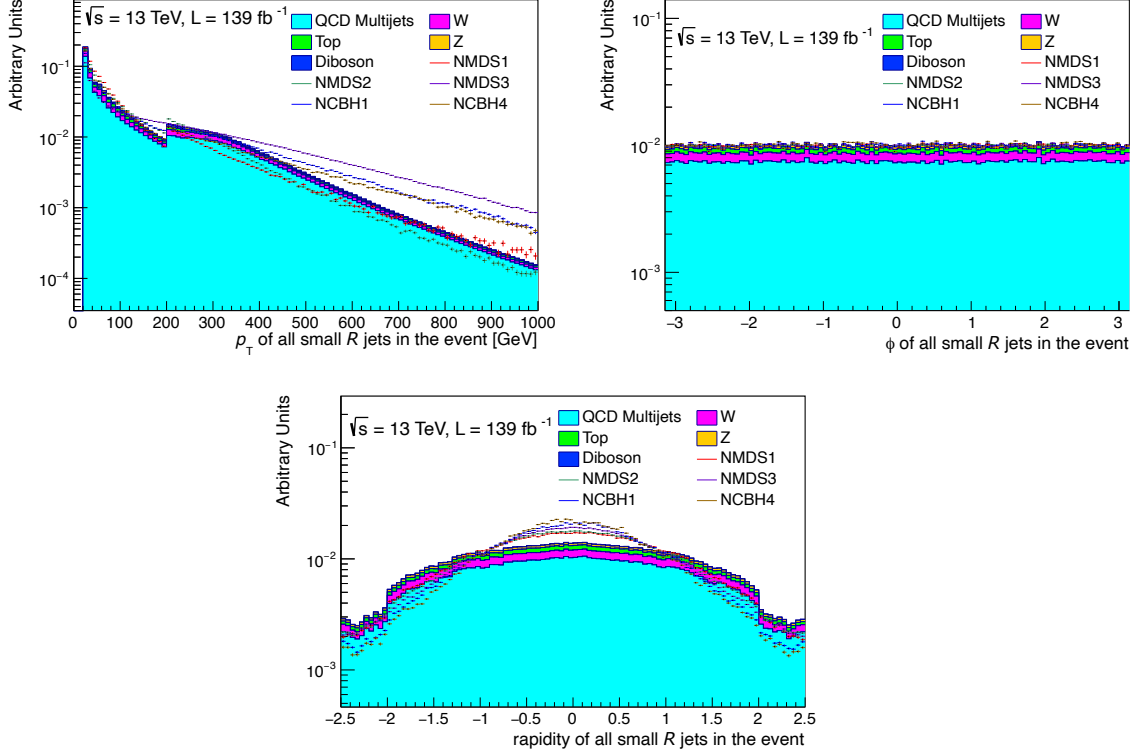


Figure 9.24 *Kinematics of all tracks within the first and second leading reclustered jets for Noncommutative Black Hole and Nonminimal Dark Sector with Noncommutative Black Hole and Nonminimal Dark Sector Pre-ML Analysis Cuts Applied*

9.7.2 Noncommutative Black Holes Networks - 500 GeV E_T^{miss} cut

Of the two noncommutative black holes used in the reconstructed analysis, the NCBH Benchmark NCBH1 is the one for which final search results have been produced. Instead of a 200 GeV E_T^{miss} cut, a 500 GeV E_T^{miss} cut was used for the final results of the NCBH Benchmark NCBH1. This provided a handle on reducing background events in the signal region beyond the reduction provided by the NN. As seen in Figure 9.25, both noncommutative black hole benchmarks and NMDS benchmark NMDS3 have substantially more E_T^{miss} than scalar cascades and the lower mass non-minimal dark sector models.

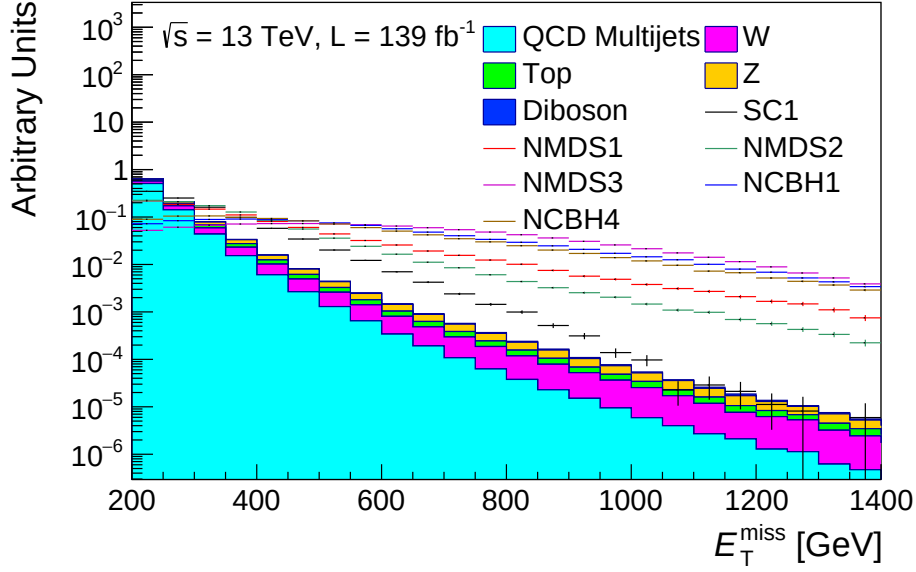


Figure 9.25 E_T^{miss} distribution of noncommutative black holes and non-minimal dark sectors with basic 200 GeV cuts applied.

Preliminary noncommutative black hole analysis results (not included in this thesis) at 200 GeV for both black hole benchmarks showed a large number (hundreds of thousands) of background events surviving the analysis cuts. The analysis became sensitive to signal cross sections ten times smaller by increasing to a 500 GeV E_T^{miss} cut. A requirement of $E_T^{\text{miss}} > 500$ GeV was, therefore, imposed. This cut has the effect of making the NN appear to perform worse. However, in reality, the NN simply has a harder task to perform because the background events that are easier to reject have already been removed by the 500 GeV E_T^{miss} cut.

The selection cuts for the analysis for the NCBH Benchmark NCBH1 can be seen in Table 9.1. The noncommutative black hole benchmark NCBH4 struggled more with the statistical fit at 500 GeV, so it is not included in the final search results. However, the NCBH benchmark NCBH4 is included in the distributions of signal vs. MC background at 500 GeV, shown below, to demonstrate it still holds discriminatory power at this cut. However, placing such a high E_T^{miss} cut proved to be suboptimal and a lesser E_T^{miss} cut would likely be advantageous. Optimization is underway (by the author, for publication) for NCBH Benchmark NCBH4.

Quantity	Requirement
Reclustered Jets	At least 2 reclustered jets of $R = 1.5$ number of small R jets in leading reclustered jet > 0 number of small R jets in subleading reclustered jet > 0 leading reclustered jet $ \eta < 2$ subleading reclustered jet $ \eta < 2$
Small R jets	> 2 small R jets p_T of at least 20 GeV Second leading small R jet $p_T > 50$ GeV Third leading small R jet $p_T > 20$ GeV
Tracks	Up to 250 tracks in leading reclustered jet Up to 250 tracks in subleading reclustered jet Greater than 0 tracks in leading reclustered jet Greater than 0 tracks in subleading reclustered jet
Lepton Veto	No electrons > 20 GeV No muons > 20 GeV
Trigger	lowest unscaled E_T^{miss} triggers
E_T^{miss}	> 500 GeV
Leading small R jet p_T	> 200 GeV

Table 9.1 *Selection cuts for analysis of Noncommutative Black Hole Benchmark NCBH1 with a 500 GeV E_T^{miss} cut*

The background rejection offered to both noncommutative black hole benchmarks from the 500 GeV E_T^{miss} cut is shown in Table 9.2.

A tile efficiency correction of 0.6305 is applied under the 500 GeV E_T^{miss} cut. The primary background now stems from W processes instead of QCD, leading at 30.54%. Benchmark NCBH1 has a significance of 2.4, which is lower than the 2.52 significance of benchmark NCBH4 due to benchmark NCBH4 having a higher cross section. The significance (defined in Equation 8.3) of the NCBH1 benchmark under the 500 GeV E_T^{miss} cut is 3.07 times improved compared to the benchmark's significance under

Background Process	Total Background	QCD	Z Processes	W Processes	Top Processes
Background Yield	3354078.502	2415577.88	256004.5	416688.43	243832.46
Background Composition		24.77%	26.14%	30.54%	15.35%
Signal Process	NCBH1	NCBH4			
Signal Yield	443.64	466.89			
Significance: Signal/sqrt(background)	2.4	2.52			
Increase in significance from 200 E_T^{miss} cut	3.07	2.65			
Tile Efficiency Correction	0.6305				

Table 9.2 *Background yield, signal yield, background composition, significance increases, and tile efficiency correction at 500 GeV E_T^{miss} cut for Noncommutative Black Holes*

the 200 GeV E_T^{miss} cut (from Table 8.7). The significance of the NCBH4 benchmark under the 500 GeV E_T^{miss} cut is 2.65 times improved compared to the benchmark's significance under the 200 E_T^{miss} GeV cut (from Table 8.7).

The variables inputted to the event level tagger of the neural network for noncommutative black holes with a 500 GeV E_T^{miss} cut outlined in Table 9.1 are the same variables inputted into the neural network for noncommutative black holes and non-minimal dark sectors at 200 GeV E_T^{miss} : sphericity, aplanarity, event shape variable C , event shape variable D , and transverse sphericity, the number of small R jets above 20 GeV, $p_T^{\text{bal}}(j_{\min}, j_{\max})$, the $p_T^{\text{bal}}(J_1, J_2)$, $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(J, E_T^{\text{miss}})$, $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, η of leading reclustered jet, η of second leading reclustered jet, number of tracks in the leading reclustered jet, number of tracks in the second leading reclustered jet, number of small R jets in the leading reclustered jet, number of small R jets in the second leading reclustered jet, $|\Delta R_{\text{avg}}(j)|$, and $|\Delta R(j_1, j_2)|$.

While the increase in the E_T^{miss} cut noticeably reduces the discriminatory power of the variables (in Figures 9.26, 9.27, and 9.28 compared to Figures 9.17, 9.18, and 9.19 for the 200 GeV E_T^{miss} cut) - some more than others - there are still visible shape differences in most of the variables. All variables remain in the tagger because, even if, for example, the number of small R jets above 20 GeV now more closely resembles that of background, being able to link the occurrence of an event with a high value of p_T balance of the leading large R jet with a given jet multiplicity may provide a handle as to whether an event is a signal event or a background event. The advantage of machine learning is the ability to derive connections that are not readily visible to the human eye.

As can be seen in Figure 9.26, under the 500 GeV E_T^{miss} cut, the noncommutative black holes exhibit a smaller fraction of events that contain high values of sphericity, aplanarity, event shape variable C , event shape variable D , and transverse sphericity. However, the fraction of events that contain high values of those quantities are

still greater than that of background, still providing discriminatory power against background. This means that the isotropic nature of the noncommutative black holes is still more distinguished than that of background under the harsher E_T^{miss} cut. It also indicates, however, that the E_T^{miss} cut has biased the distributions to be more dijet-like than before. Figure 9.26 shows, qualitatively, that data and Monte Carlo are broadly consistent in shape, with any deviations being restricted to the tails of the distributions.

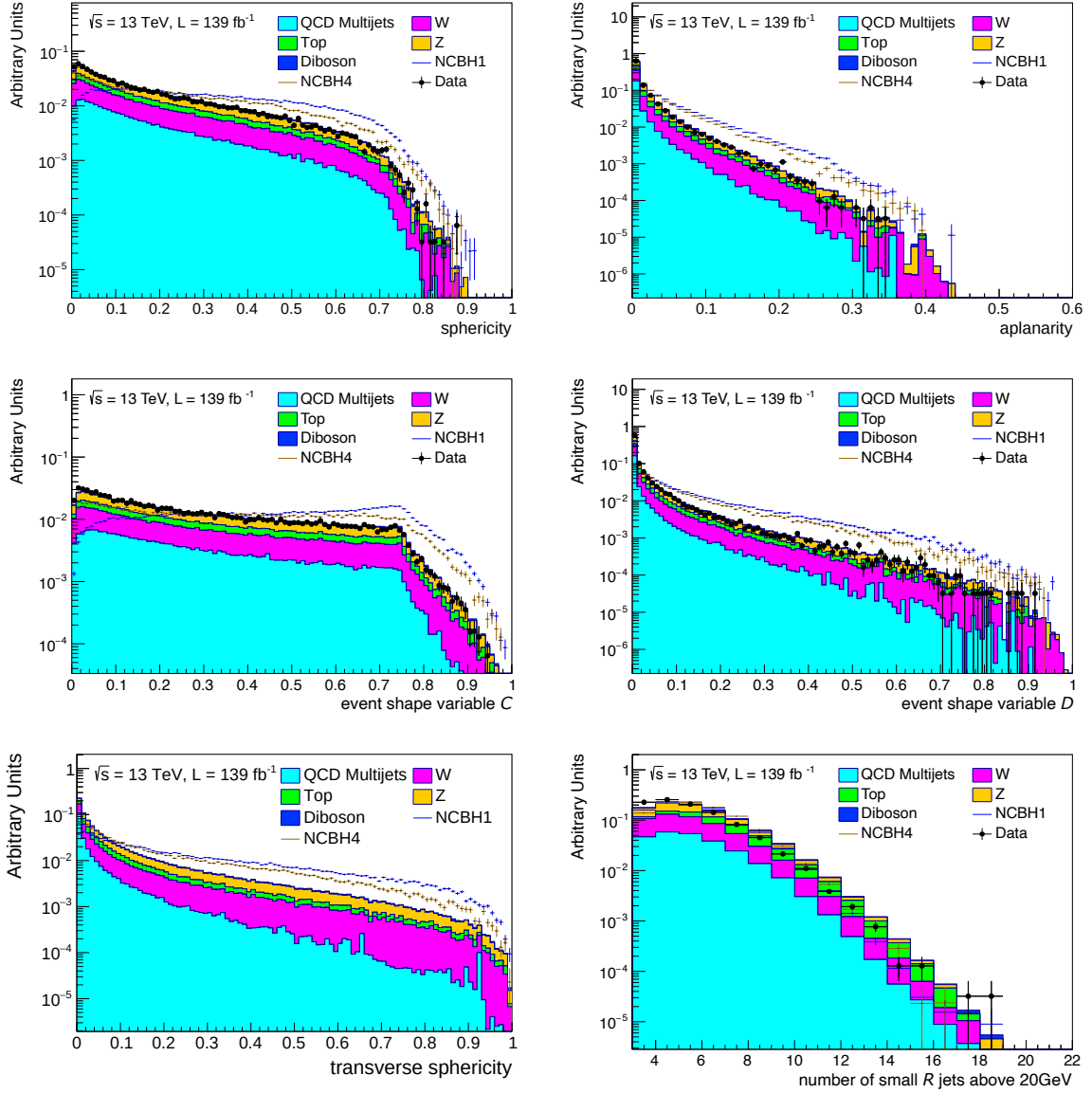


Figure 9.26 *Discriminating Variables for Noncommutative Black Hole with 500 GeV E_T^{miss} cut from Table 9.1 Applied.*

In Figure 9.27, under the 500 GeV E_T^{miss} cut, the noncommutative black holes retain discriminatory power against background, though to a significantly lesser extent

than under the 200 GeV E_T^{miss} cut, in $p_T^{\text{bal}}(J_1, J_2)$, $\Delta\phi_{\text{min}}(j, E_T^{\text{miss}})$, $\Delta\phi_{\text{maxmin}}(j, E_T^{\text{miss}})$, $\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}})$, $p_T^{\text{bal}}(j_{\text{min}}, j_{\text{max}})$, and $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$. Compared to background, there remain a higher fraction of events for noncommutative black holes and non-minimal dark sectors that have a greater separation between the E_T^{miss} and the closest jet, be it small R jets or large R jets. Figure 9.27 shows that data and Monte Carlo are, qualitatively, broadly consistent in shape, with deviations existing in the tails of the distributions of $\Delta\phi_{\text{maxmin}}(J, E_T^{\text{miss}})$, $\Delta\phi_{\text{min}}(j, E_T^{\text{miss}})$, $\Delta\phi_{\text{min}}(J, E_T^{\text{miss}})$, and $p_T^{\text{bal}}(J_1, J_2)$. This section will present ratio plots of data and Monte Carlo to quantify the agreement of Monte Carlo and data under these selection cuts.

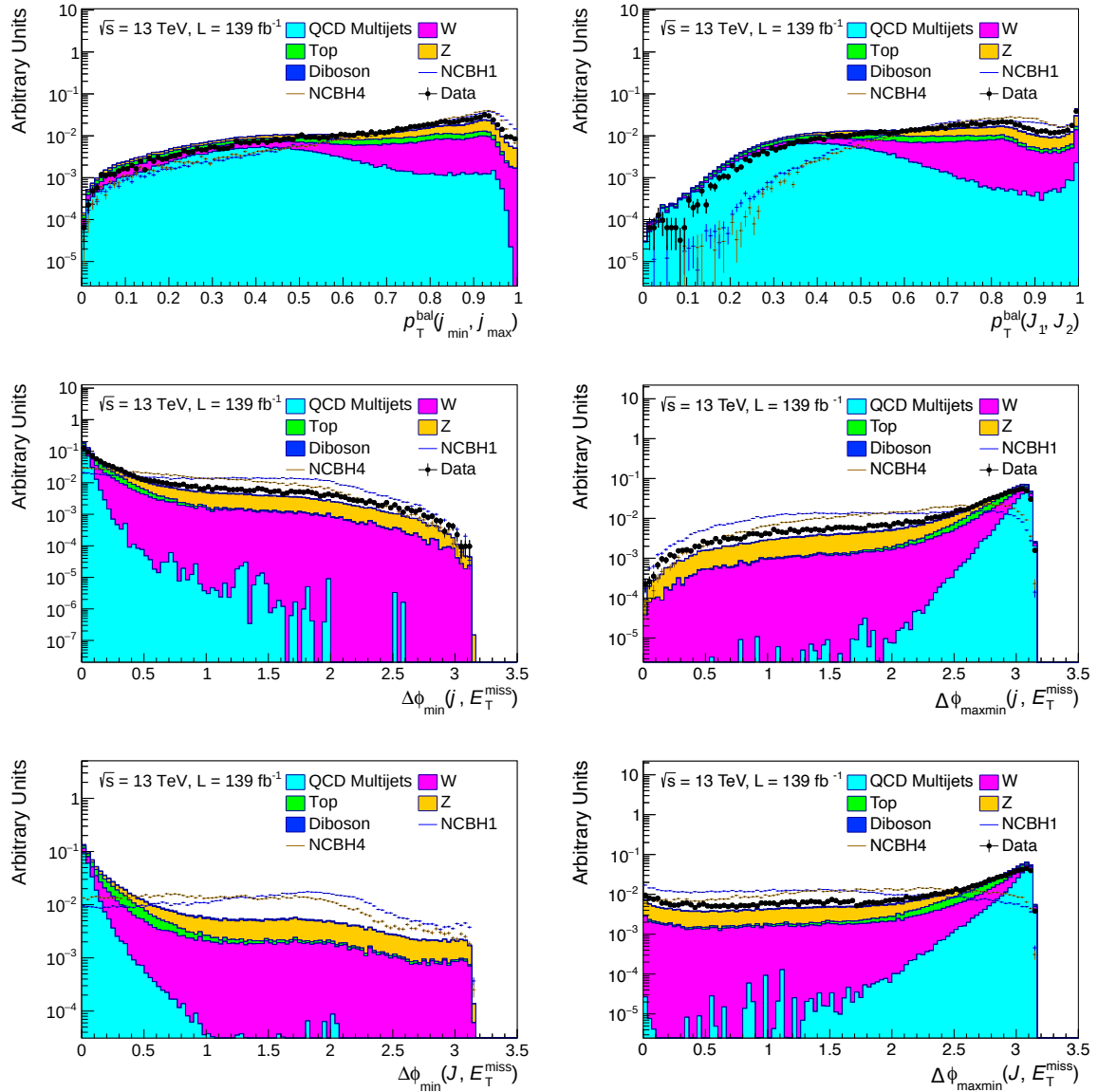


Figure 9.27 *Discriminating Variables for Noncommutative Black Hole with 500 GeV E_T^{miss} cut from Table 9.1 Applied.*

In Figure 9.28, under the 500 GeV E_T^{miss} cut, noncommutative black holes retain some discriminatory power in the η distribution of leading reclustered jet, though to a lesser extent than under the 200 GeV E_T^{miss} cut. No discriminatory power remains in the η distribution of the subleading reclustered jet. This holds true as well for the number of small R jets in the second leading reclustered jet and the number of tracks in the second leading reclustered jet. The observables $|\Delta R(j_1, j_2)|$ and $|\Delta R_{\text{avg}}(j)|$ retain discriminatory power, though to a lesser extent than under the 200 GeV E_T^{miss} cut. Figure 9.28 shows that the data and Monte Carlo distributions are, qualitatively, broadly consistent in shape.

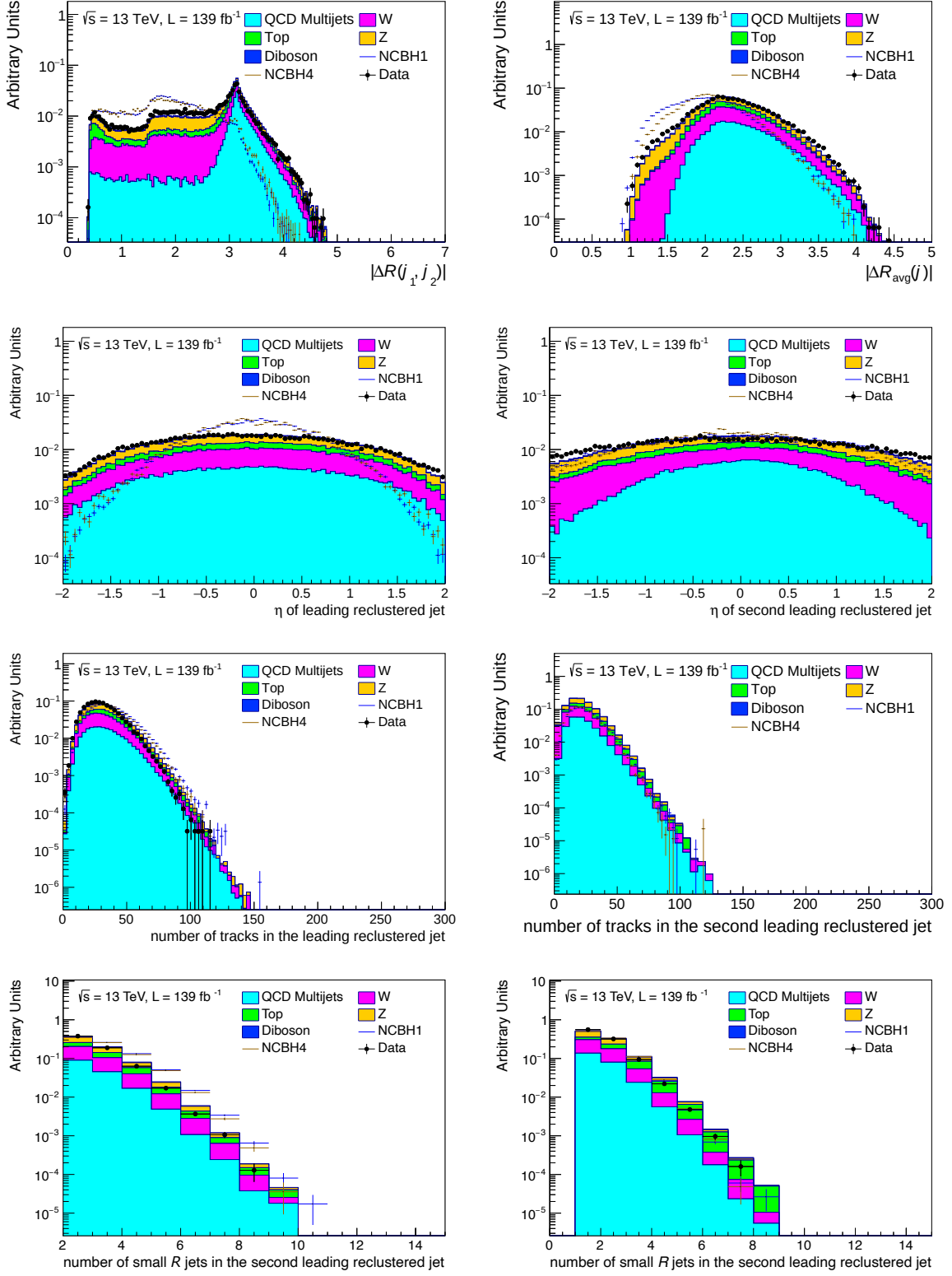


Figure 9.28 *Discriminating Variables for Noncommutative Black Hole with 500 GeV E_T^{miss} cut from Table 9.1 Applied.*

Figure 9.29 shows that the distributions of Monte Carlo and data in event shape

variable C agree within either 20% or 2σ .

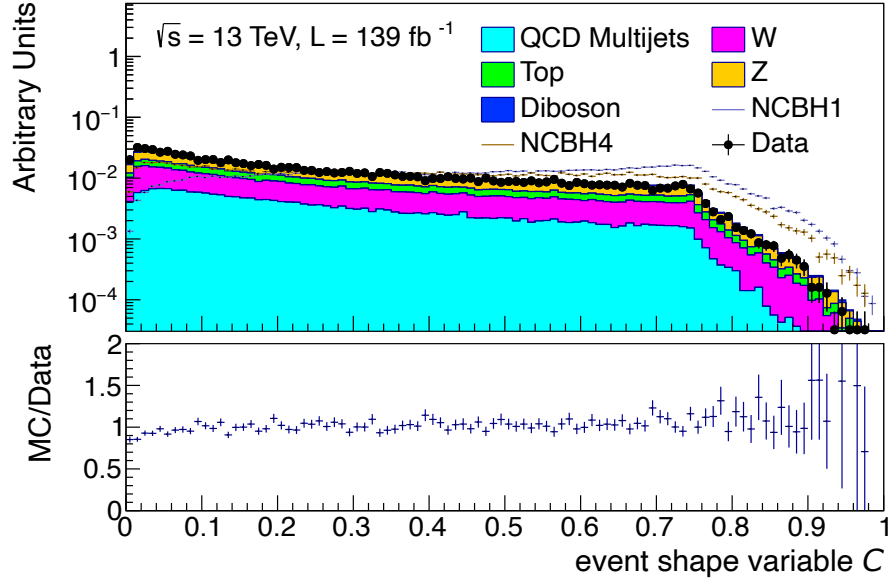


Figure 9.29 *Top plot: Comparison of Monte Carlo and data distributions for event shape variable C . This distribution is normalized. The 500 GeV Noncommutative Black Hole E_T^{miss} cuts from Table 9.1 have been applied. Bottom plot: The Monte Carlo to Data ratio of event shape variable C . This distribution was created from the normalized distributions in the top plot.*

Figure 9.30 shows that the Monte Carlo and data distributions for the η of the leading reclustered jet agree within 20%.

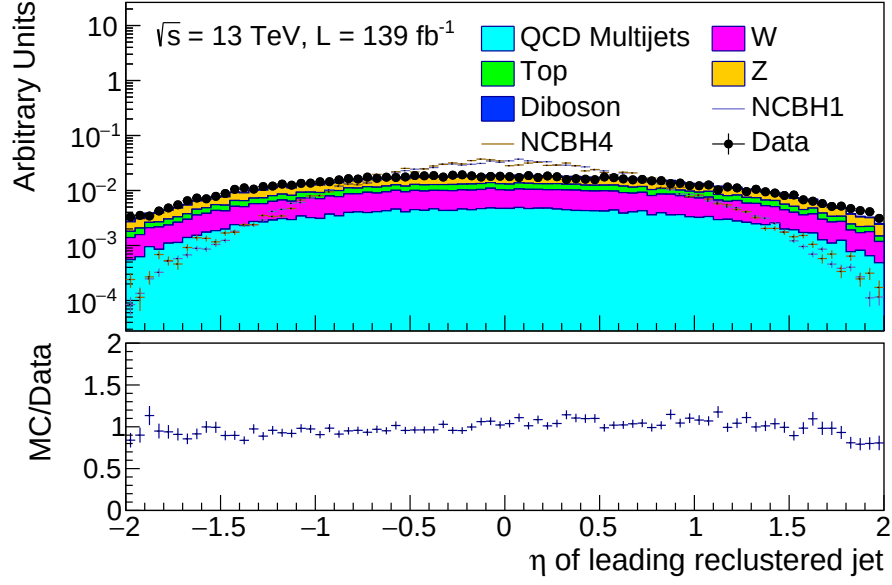


Figure 9.30 *Top plot: Comparison of Monte Carlo and data distributions for the η distribution of the leading reclustered jet. This distribution is normalized. The 500 GeV Noncommutative Black Hole E_T^{miss} cuts from Table 9.1 have been applied. Bottom plot: The Monte Carlo to Data ratio of the η distribution of the leading reclustered jet. This distribution was created from the normalized distributions in the top plot.*

The decorrelated variable for noncommutative black holes is $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. As will be discussed in the next chapter, the signal depleted region of phase space for which it is safe to compare Monte Carlo and data is $\Delta\phi_{\min}(J, E_T^{\text{miss}}) < 1.2$. Figure 9.31 shows largely good agreement between Monte Carlo in the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. The ratio stays largely flat for the entire distribution except for $\Delta\phi_{\min}(J, E_T^{\text{miss}}) < 0.2$, which is far from the signal region of $\Delta\phi_{\min}(J, E_T^{\text{miss}}) > 1.2$. As will be seen in the next chapter, everything below a $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ value of 0.75 is outside the range of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ used in the background estimation and so not taken into account in the estimation background in the signal region.

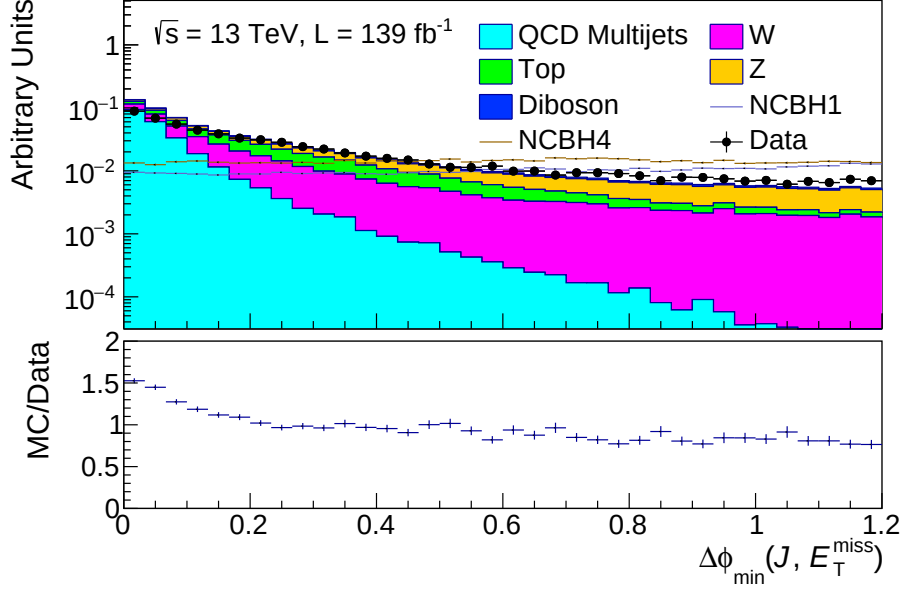


Figure 9.31 *Top plot: Comparison of Monte Carlo and data distributions for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This distribution is normalized. The 500 GeV Noncommutative Black Hole E_T^{miss} cuts from Table 9.1 have been applied. Bottom plot: The Monte Carlo to Data ratio of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This distribution was created from the normalized distributions in the top plot.*

9.7.3 Non-minimal Dark Sector Networks - 500 GeV E_T^{miss} cut with 1400 GeV H_T cut

Of the three non-minimal dark sector benchmarks used in the reconstructed analysis, the NMDS Benchmark NMDS3 is the one for which final search results have been produced. The reason for this will be explained in the subsequent passages. As a reminder, NMDS3 is the benchmark that has the highest mass spacing between ensemble constituents out of all of the non-minimal dark sector benchmarks. It, therefore, has the most energetic decays. As seen in Figure 9.25 and Figure 9.32, the NMDS3 has substantial E_T^{miss} and H_T . In this analysis, H_T is defined as the sum of transverse momenta of all jets in the event about 20 GeV.

A requirement of at least 500 GeV in E_T^{miss} and at least 1400 GeV in H_T was imposed. This set of cuts was chosen because it greatly reduces MC background while still allowing signal to be distinguishable from background. Cutting higher on E_T^{miss} or H_T would have made the signal less distinct against background. The cuts applied are shown in Table 9.3.

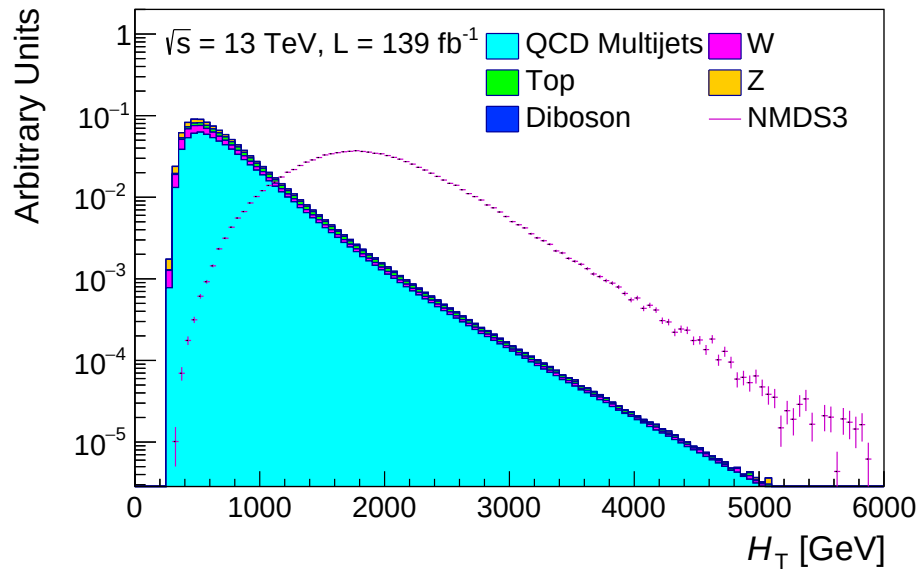


Figure 9.32 H_T distribution of NMDS benchmark NMDS3

Quantity	Requirement
Reclustered Jets	At least 2 reclustered jets of $R = 1.5$ number of small R jets in leading reclustered jet > 0 number of small R jets in subleading reclustered jet > 0 leading reclustered jet $ \eta < 2$ subleading reclustered jet $ \eta < 2$
Small R jets	> 4 small R jets p_T of at least 20 GeV Second leading small R jet $p_T > 50$ GeV
Tracks	Up to 250 tracks in leading reclustered jet Up to 250 tracks in subleading reclustered jet Greater than 0 tracks in leading reclustered jet Greater than 0 tracks in subleading reclustered jet
Lepton Veto	No electrons > 20 GeV No muons > 20 GeV
Trigger	lowest unscaled E_T^{miss} triggers
E_T^{miss}	> 500 GeV
H_T	> 1400 GeV
Leading small R jet p_T	> 200 GeV

Table 9.3 *Selection cuts for analysis of Non-minimal Dark Sector benchmark NMDS3 with a 500 GeV E_T^{miss} and 1400 GeV H_T cut.*

The effects of the cuts are seen in Table 9.4. QCD remains the primary background, contributing 44.95%. The significance (defined in Equation 8.3) with a 500 GeV E_T^{miss} and 1400 GeV H_T cut is 0.05, which is 8.4 times improved compared to the significance obtained from the basic 200 GeV E_T^{miss} cuts. With the basic 200 GeV E_T^{miss} cut, the NMDS benchmark NMDS3 was overtaken by background and the final analysis statistical fit failed. The 500 GeV E_T^{miss} cut and H_T provided a remedy, though the extremely small cross section will be shown to limit the sensitivity. The low p_T NMDS benchmark NMDS1 and the moderate p_T NMDS benchmark struggled with the 200 GeV basic cuts, failing the statistical fits. This was due to the very large

number of background events in the signal region being coupled with a nonclosure of 5%. Those benchmarks are excluded from the final search results as optimizations are underway, such as the use of more discriminating decorrelated variables. (These optimizations are being performed by the author to include in the publication.)

Background Process	Total Background	QCD	Z Processes	W Processes	Top Processes
Background Yield	5145.74	2313.1	591.94	1990.05	1173.83
Background Composition		44.95%	11.50%	38.70%	22.80%
Signal Process	NMDS3				
Signal Yield	3.58				
Significance: Signal/sqrt(background)	0.05				
Increase in significance from 200 E_T^{miss} cut	8.4				
Tile Efficiency Correction	0.47145				

Table 9.4 *MC background yields, MC background composition, Signal Yields, Significances, Tile Efficiency Correction with 500 GeV E_T^{miss} and 1400 GeV H_T cut.*

The variables inputted to the event level tagger of the neural network for non-commutative black holes and non-minimal dark sectors are the five event shapes sphericity, aplanarity, event shape variable C , event shape variable D , and transverse sphericity, the number of small R jets above 20 GeV, $p_T^{\text{bal}}(j_{\min}, j_{\max})$, the $p_T^{\text{bal}}(J_1, J_2)$, $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(J, E_T^{\text{miss}})$, $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, η of leading reclustered jet, η of second leading reclustered jet, number of tracks in the leading reclustered jet, number of tracks in the second leading reclustered jet, number of small R jets in the leading reclustered jet, number of small R jets in the second leading reclustered jet, $|\Delta R_{\text{avg}}(j)|$, and $|\Delta R(j_1, j_2)|$.

As can be seen in Figure 9.33 (compared to the distributions for the 200 GeV E_T^{miss} cut in Figure 9.17), the non-minimal dark sector benchmark NMDS3 retains powerful discriminatory power against MC background in higher values of the event shape variables sphericity, aplanarity, event shape variable C , event shape variable D , and transverse sphericity, indicating that requirement of at least 500 GeV E_T^{miss} and at least 1400 GeV in H_T has not tarnished the isotropic nature of the events. There remains a visible shape difference in the number of small R jets above 20 GeV for the MC background and non-minimal dark sector benchmark NMDS3, though this variable in particular has less discriminatory power than with just a 200 GeV E_T^{miss} cut. Figure 9.33 shows that the distributions of data and Monte Carlo are, qualitatively broadly consistent in shape in these variables, with any deviations being restricted to the tails. Transverse sphericity serves as the decorrelated variable for the NN for the non-minimal dark sector benchmark NMDS3, so data is not overlaid in the plot here and instead will be examined in a control region shortly.

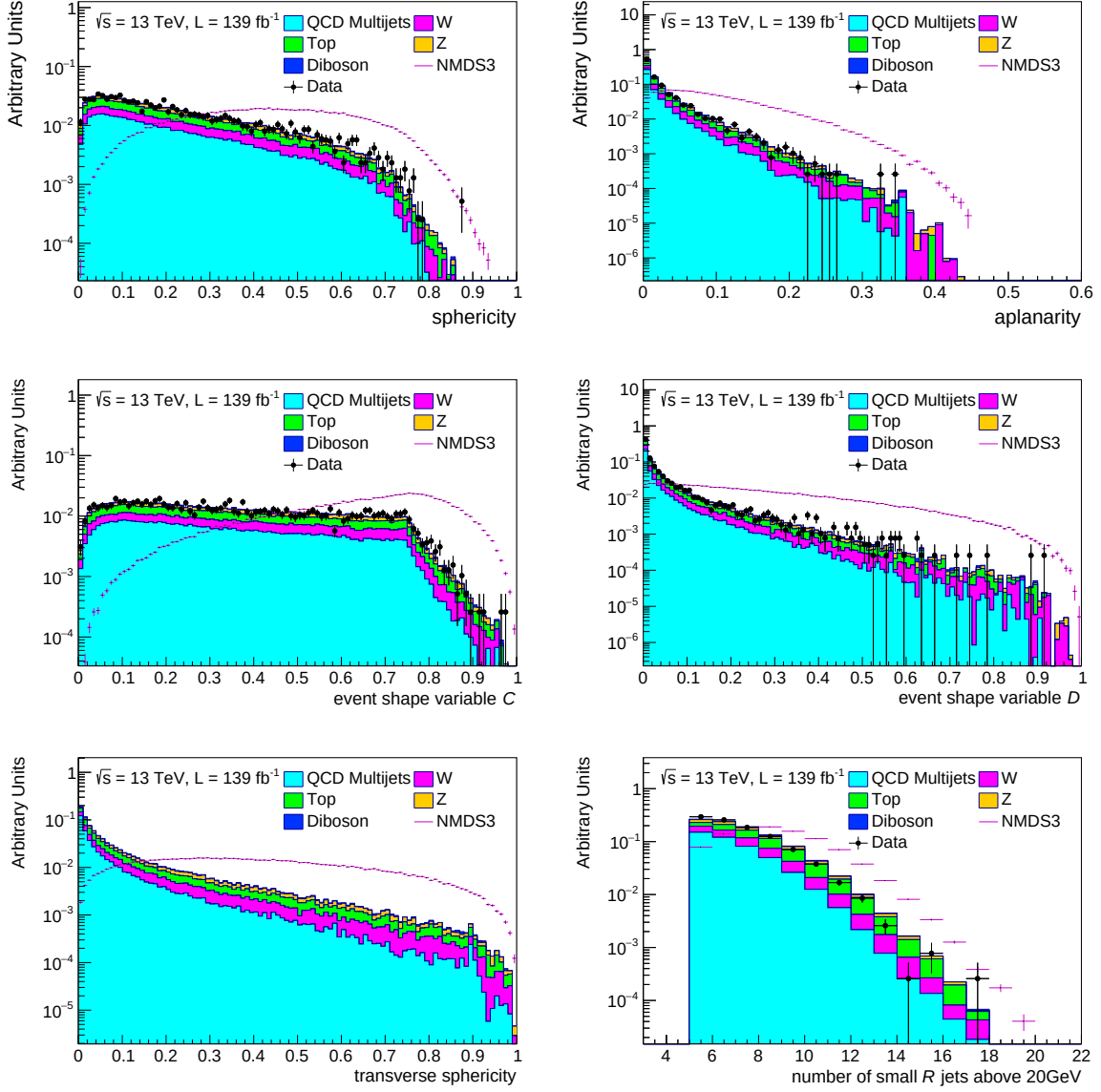


Figure 9.33 *Discriminating Variables for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied.*

As can be seen in Figure 9.34, though slightly less pronounced than with the 200 GeV E_T^{miss} cut (seen in Figure 9.18), the non-minimal dark sector benchmark NMDS3 retains strong discriminatory power against MC background in $p_T^{\text{bal}}(J_1, J_2)$, $\Delta\phi_{\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(j, E_T^{\text{miss}})$, $\Delta\phi_{\max\min}(J, E_T^{\text{miss}})$, $p_T^{\text{bal}}(j_{\min}, j_{\max})$, and $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. There remain a higher fraction of events for non-minimal dark sector benchmark NMDS3 that have a greater separation between the E_T^{miss} and the closest jet, be it small R jets or large R jets. Figure 9.34 shows that the distributions of data and Monte Carlo are broadly consistent in shape. Deviations appear in the

tail.

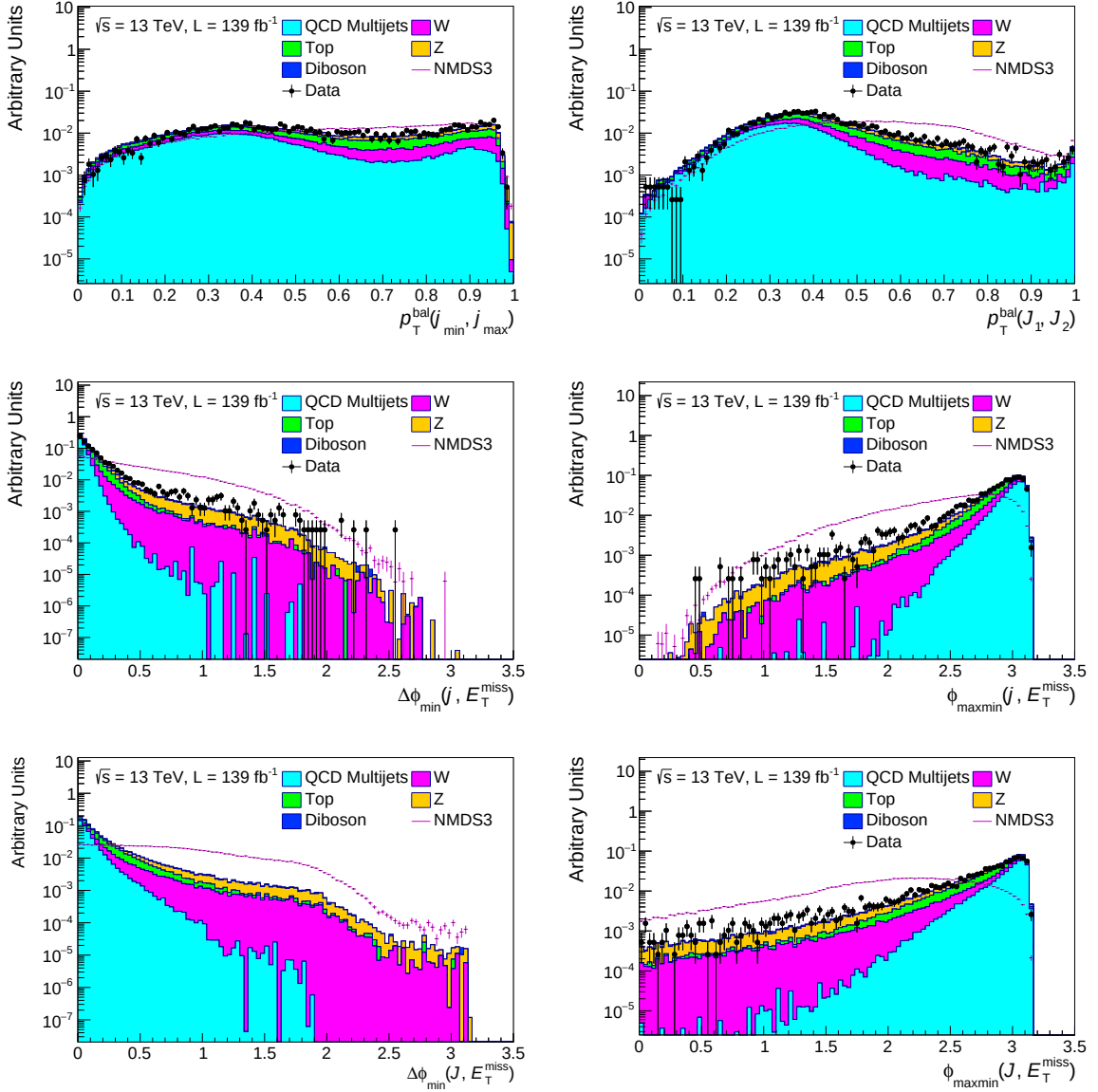


Figure 9.34 *Discriminating Variables for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied.*

Figure 9.35 shows that, under the 500 GeV E_T^{miss} cut and 1400 GeV H_T cut, the Non-minimal dark sector benchmark NMDS3 has lost almost all of its discriminatory power in the number of tracks in the second leading reclustered jet, the η of the second leading reclustered jet, and the number of small R jets in the second leading reclustered jet (compared to the 200 GeV E_T^{miss} cut in Figure 9.19). Though less powerful than before, the benchmark retains discriminatory power in the number of tracks in the leading reclustered jet, peaking at 50 tracks, while MC background

peaks at 30 tracks. The benchmark still holds slight shape differences in the η of the leading reclustered jet, having more central values of η , and in $|\Delta R_{\text{avg}}(j)|$, having a higher fraction of events with lower values than MC background. Strong discriminatory power remains in the $|\Delta R(j_1, j_2)|$, with signal having a higher fraction of events with the leading small R jet close to the second leading small R jet. Figure 9.35 shows that the Monte Carlo and data distributions are, qualitatively, broadly consistent in shape.

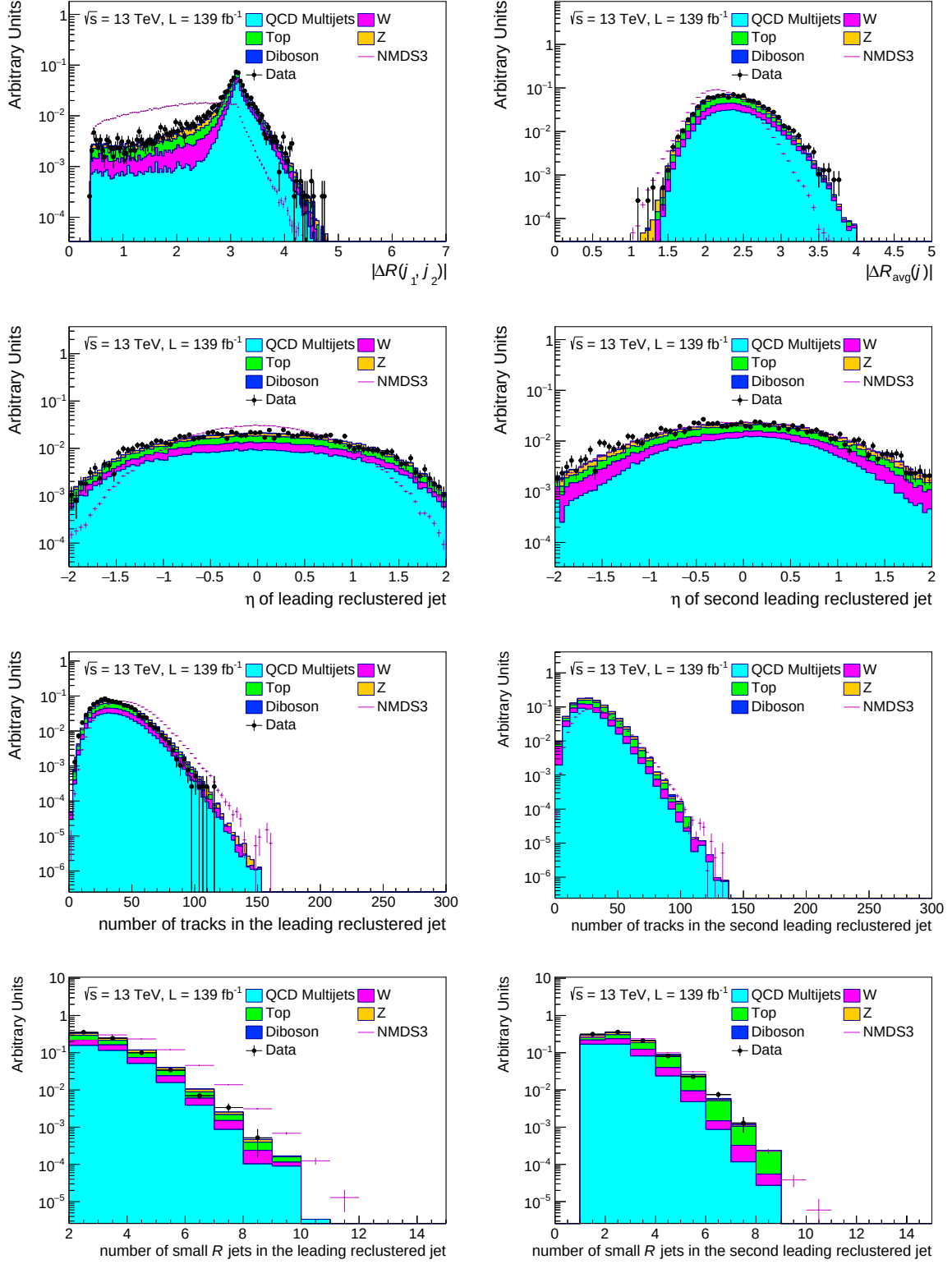


Figure 9.35 *Discriminating Variables for Non-minimal dark sector benchmark NMDS3 with $500 \text{ GeV } E_T^{\text{miss}}$ cut and $1400 \text{ GeV } H_T$ cut from Table 9.3 Applied.*

Figure 9.36 shows that, due to the large uncertainties, the ratio of the Monte Carlo and data distributions in event shape variable C is largely within 1.

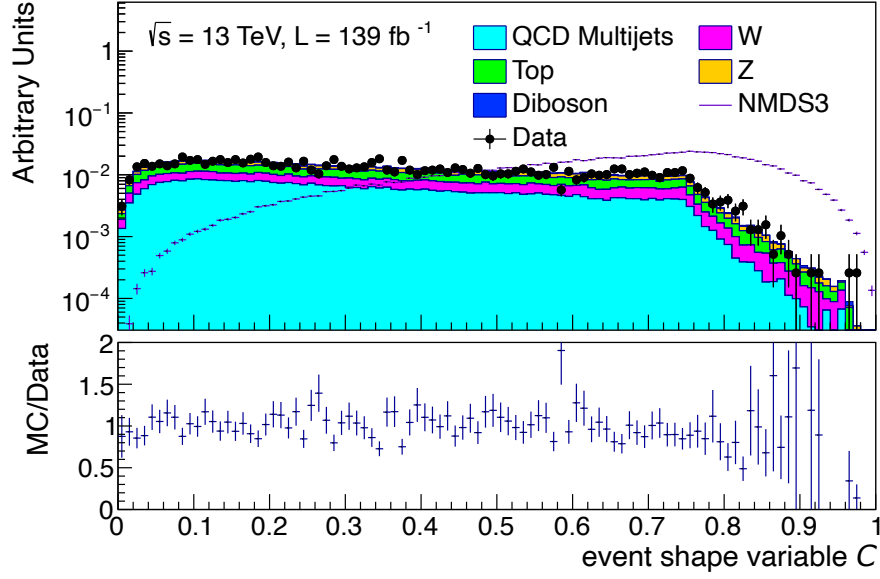


Figure 9.36 *Top plot: Comparison of Monte Carlo and data distributions for event shape variable C for Non-minimal dark sector benchmark NMDS3 with $500 \text{ GeV } E_T^{\text{miss}}$ cut and $1400 \text{ GeV } H_T$ cut from Table 9.3 Applied. This distribution is normalized. Bottom plot: The Monte Carlo to Data ratio of event shape variable C . This distribution was created from the normalized distributions in the top plot.*

As seen in Figure 9.37, the Monte Carlo to data ratio for the η distribution of the leading reclustered jet stays at 1 within a 2σ uncertainty for almost all bins.

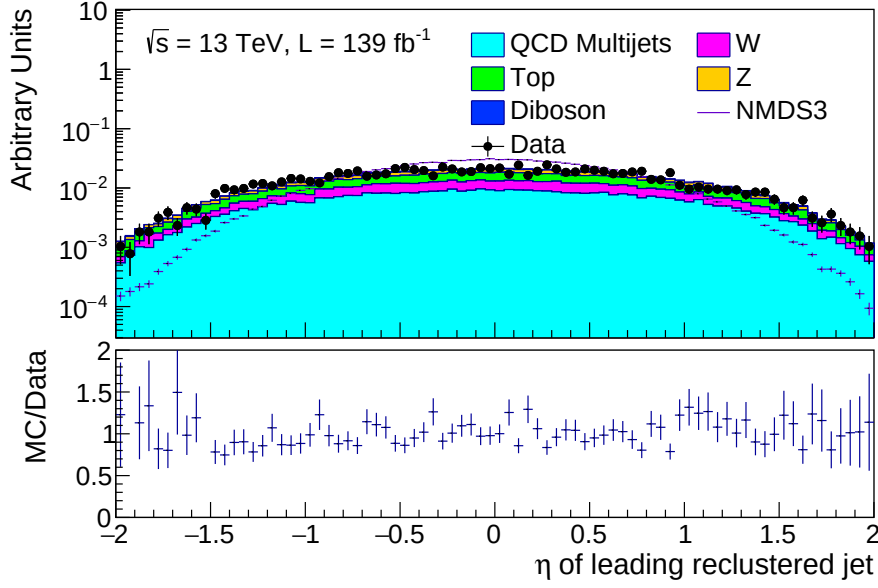


Figure 9.37 *Top plot: Comparison of Monte Carlo and data distributions for the η distribution of the leading reclustered jet for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied. This distribution is normalized. Bottom plot: The Monte Carlo to Data ratio of the η distribution of the leading reclustered jet. This distribution was created from the normalized distributions in the top plot.*

The decorrelated variable for Non-minimal dark sector benchmark NMDS3 is the transverse sphericity. As will be discussed in the next chapter, the signal depleted region of phase space for which it is safe to compare Monte Carlo and data is transverse sphericity values less than 0.2. Figure 9.3 shows good agreement between data and Monte Carlo in the distribution of the transverse sphericity. The ratio stays largely at 1 within error bars for the entire distribution.

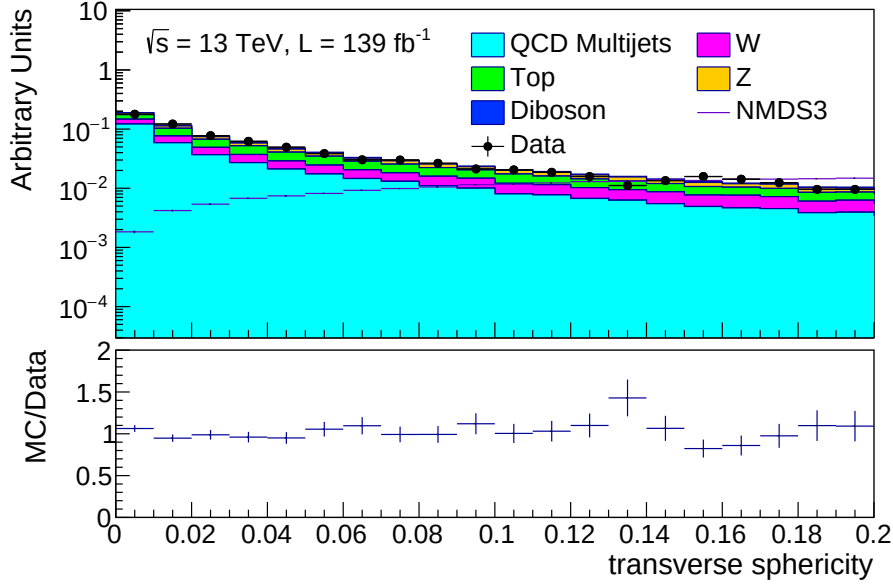


Figure 9.38 *Top plot: Comparison of Monte Carlo and data distributions for the η distribution of transverse sphericity for Non-minimal dark sector benchmark NMDS3 with 500 GeV E_T^{miss} cut and 1400 GeV H_T cut from Table 9.3 Applied. This distribution is normalized. Bottom plot: The Monte Carlo to Data ratio of the transverse sphericity. This distribution was created from the normalized distributions in the top plot.*

9.7.4 Noncommutative Black Holes Neural Network Results

The NN results of both the NCBH Benchmark NCBH1 and NCBH benchmark NCBH4 with a 200 GeV E_T^{miss} cut are compared to the NN results of NCBH Benchmark NCBH1 at a 200 GeV E_T^{miss} to illustrate the effect the 500 GeV E_T^{miss} cut has on the tagger. The general procedure for how all neural networks in this thesis were trained was described in 9.6. All noncommutative black hole neural networks were trained for 100 epochs.

Figure 9.39 shows the results of the NN for NCBH benchmark NCBH4 when running with a DisCo λ parameter of 2 and with the p_T balance of the closest and furthest small R jet to the E_T^{miss} . The ROC curve in Figure 9.39 shows a background rejection of 90% for a roughly 65% signal retention. Noncommutative black holes are inherently less striking against background compared to the scalar cascades, so it makes sense that the neural network of the noncommutative black holes is not as robust as that of the scalar cascades. Nonetheless, the background rejection indicates that the network is still performing well.

It is worthwhile to note that the DisCo λ parameter for this neural network and all noncommutative black hole networks in this thesis is a value of 2, which is far less than the DisCo λ parameter of 50 that the scalar cascades use. This is because higher values of the DisCo λ parameter degrade the performance of noncommutative black holes in a way that is not seen for the striking signature of scalar cascades. Inherently, the decorrelation provided by DisCo comes with the penalty of decreased NN performance. Since the noncommutative black holes are not as distinctive of a signature as scalar cascades, this penalty is more pronounced and a higher value for the DisCo λ parameter cannot be afforded for noncommutative black holes.

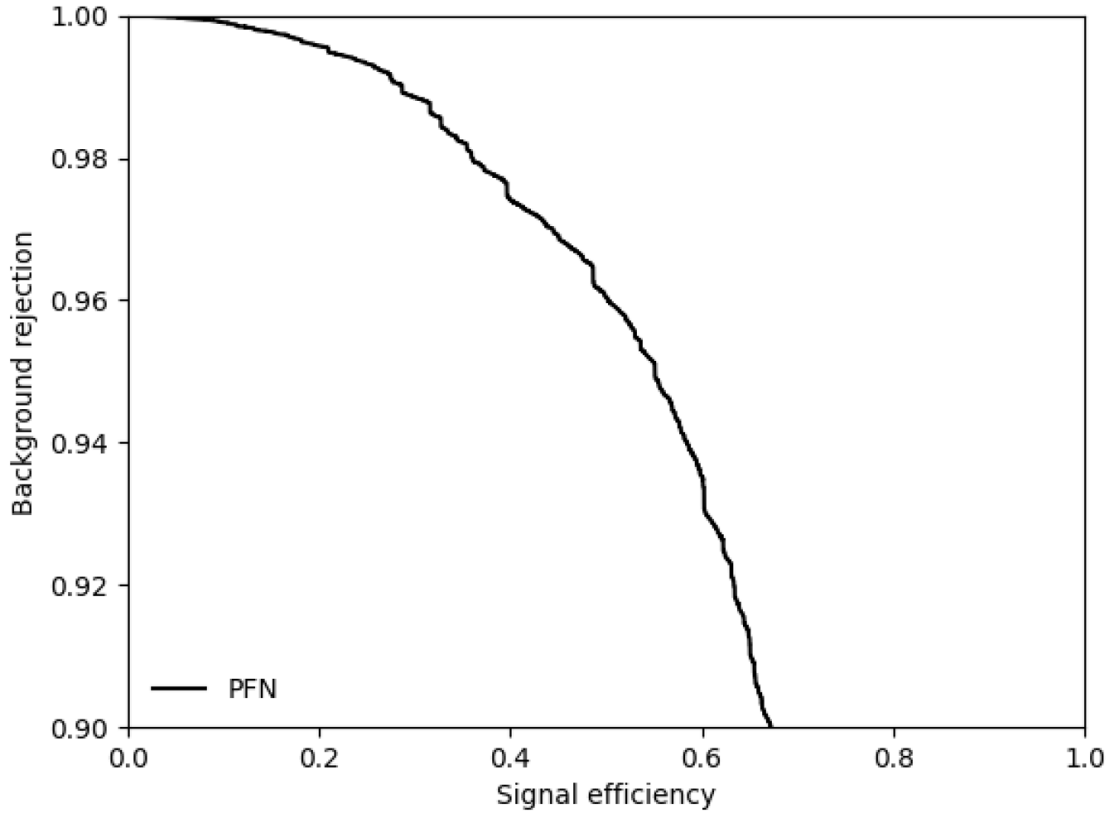


Figure 9.39 *ROC Curve of Background Rejection vs. Signal Efficiency for NCBH benchmark NCBH₄ with DisCo hyperparameter $\lambda = 2$ and $200 E_T^{\text{miss}}$ cut and basic cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} .*

The ROC curve in Figure 9.40, which shows the NN results for NCBH benchmark NCBH1, shows a background rejection of 90% for a roughly 73% signal retention. The NN is trained using a DisCo hyperparameter λ value of 2 and a decorrelated value of the p_T balance of the closest and further small R jet to the E_T^{miss} . Again, though less striking against background compared to the scalar cascades, the background rejection indicates that the network is still performing well.

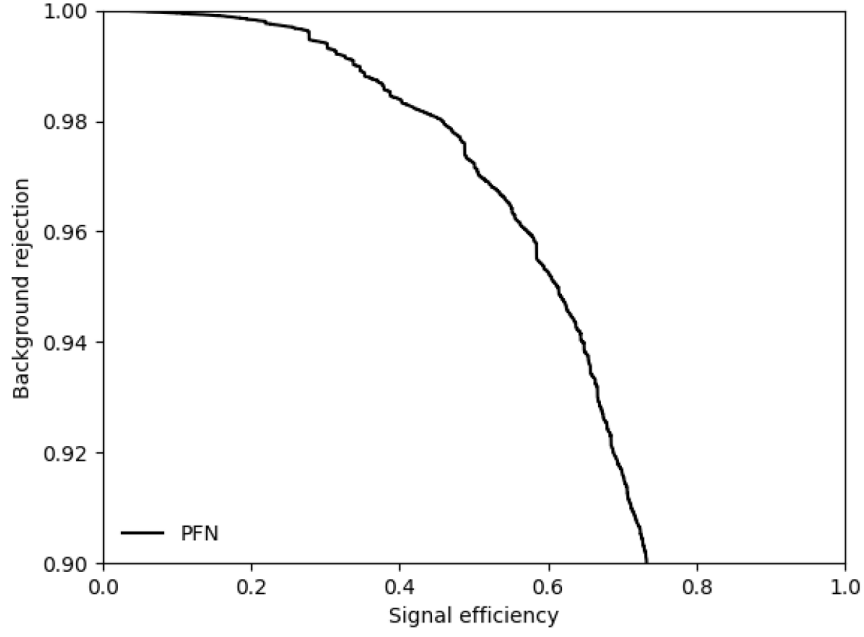


Figure 9.40 *ROC Curve of Background Rejection vs. Signal Efficiency for NCBH benchmark NCBH1 with DisCo hyperparameter λ value of 2 and 200 GeV E_T^{miss} cut and basic cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} .*

The output labels in Figure 9.41 for background Monte Carlo simulated events for this model and training show that background peaks at NN output values of 0. There is a shallow shoulder, however, that extends to higher NN output values.

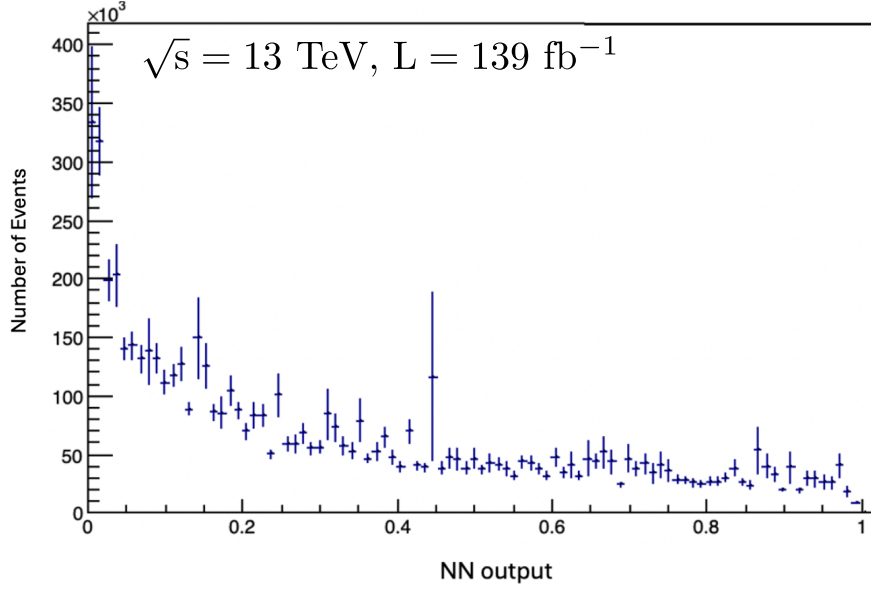


Figure 9.41 *NN prediction for background events using NN for NCBH benchmark NCBH1 with a DisCo hyperparameter λ value of 2 and applying the 200 GeV E_T^{miss} cut and basic cuts from Table 8.6. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} .*

As seen in Figure 9.42, the signal neural network output for the NCBH benchmark NCBH1 is peaked at a NN output of around 1 in the figure below, with a broad shoulder extending to lower NN output values. Together with the NN prediction for background events, this would be consistent with a background rejection of 90% for a signal retention of 73%.

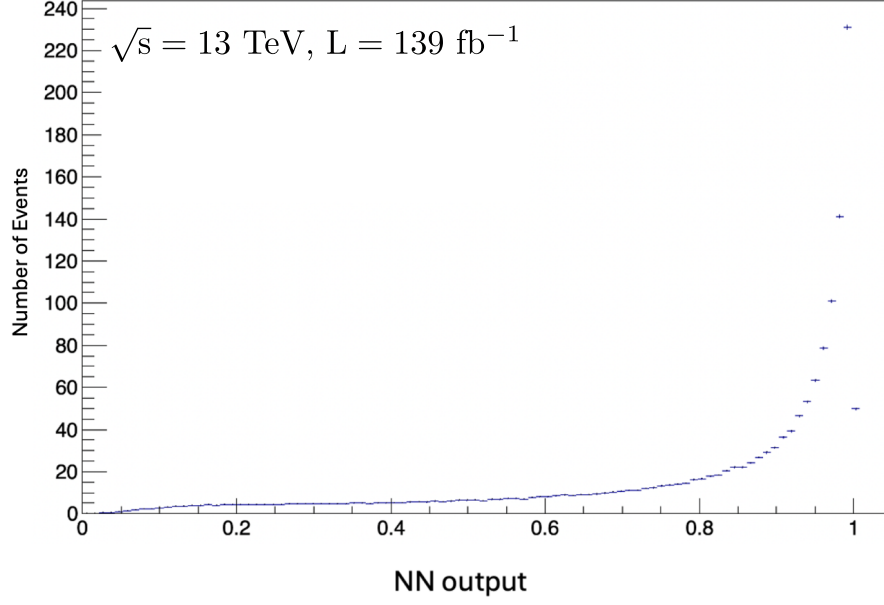


Figure 9.42 *NN prediction for signal events using NN for NCBH benchmark NCBH1 with a DisCo hyperparameter λ value of 2 and applying the 200 GeV E_T^{miss} cut and basic cuts from Table 8.6. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} .*

Below is a presentation and discussion of the results of the NN for NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} when running with a DisCo λ parameter of 2. The decorrelated variable is the minimum ϕ between the closest and furthest small R jet to the E_T^{miss} . The ROC curve in Figure 9.43 shows a background rejection of 90% for a roughly 47% signal retention. Though not the 65% or 73% signal rejection seen by the NCBH benchmarks at 200 GeV, the NN for the NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} cut still offers discrimination which, combined with the E_T^{miss} cut yields a strong handle against background.

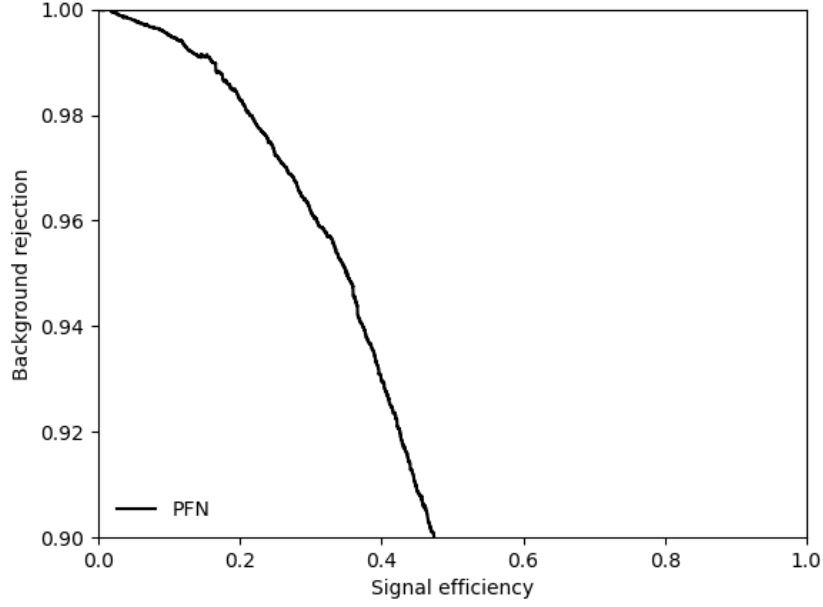


Figure 9.43 *ROC Curve of Background Rejection vs. Signal Efficiency using NN for NCBH benchmark NCBH1 with 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.*

The accuracy of this neural network for both the training and validation datasets improves as the number of epochs that the neural network is trained for increases, which is the desired behavior. The somewhat ragged nature of the curve is not a concern as this is commonly observed behavior in machine learning training. The important point is that the accuracy improves for both the training and validation datasets.

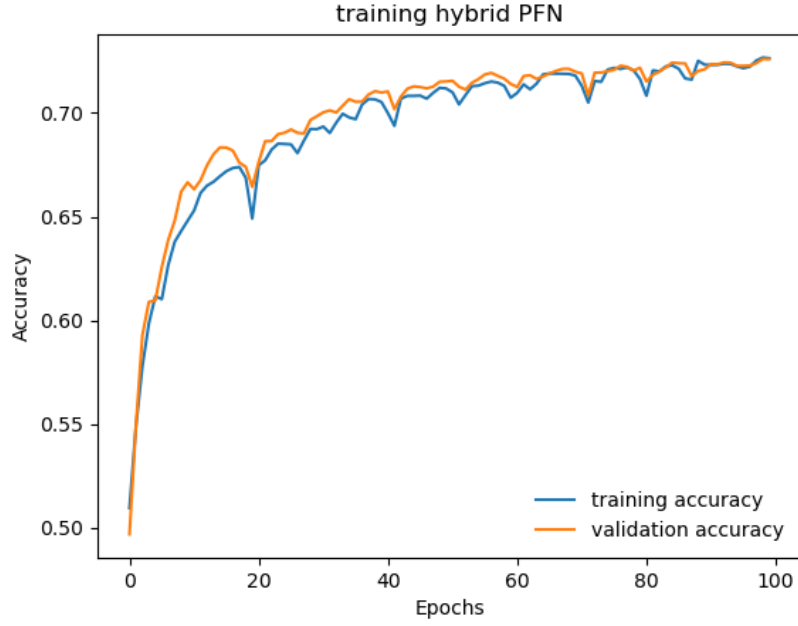


Figure 9.44 *Evolution of accuracy of NN validation accuracy and training accuracy as a function of epochs for NN using NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.*

The loss of the neural network decreases over time for both the training sample and the validation sample, which is the desired behavior, as the goal in training is to minimize the loss.

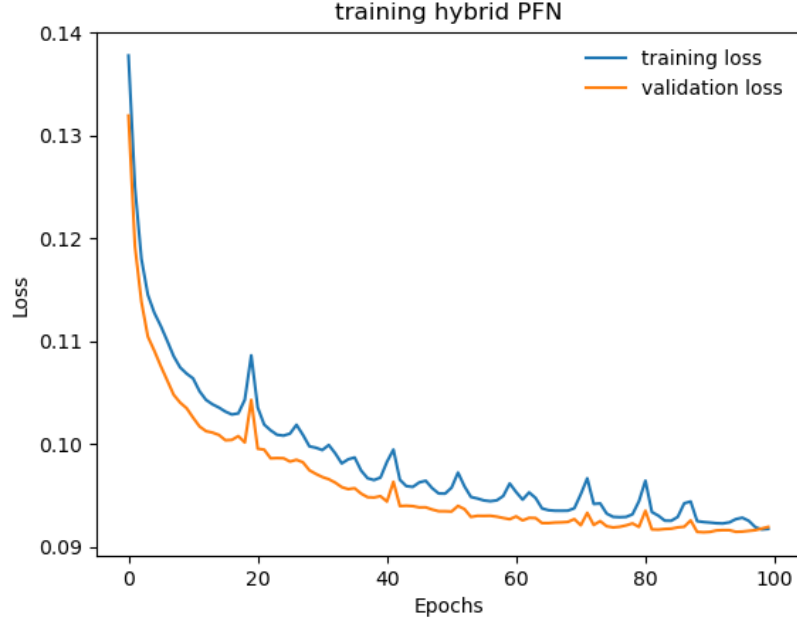


Figure 9.45 *Evolution of NN validation loss and training loss as a function of epochs for NN using NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.*

The plot of output labels for background Monte Carlo simulated events in Figure 9.46 for the NN using NCBH benchmark NCBH1 with a 500 GeV E_T^{miss} cut shows a very broad shoulder stemming from a NN output value of 0 and cutting nearly diagonally across the NN output values. Here an obvious difference can be seen between the NN performance of the noncommutative black holes with a 200 GeV E_T^{miss} cut and a 500 GeV E_T^{miss} cut, as the background NN output is far less sharply peaked around NN values of 0 when a 500 GeV E_T^{miss} cut applied.

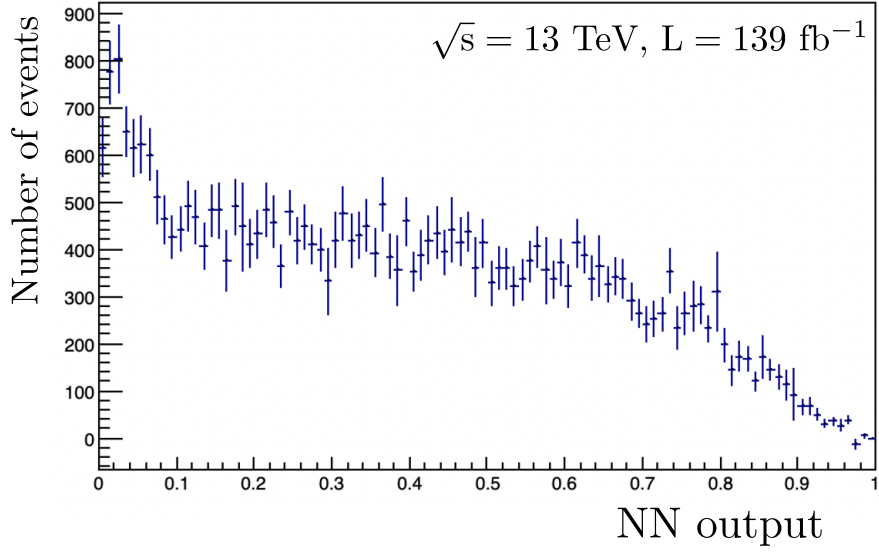


Figure 9.46 *NN prediction for background events using NN for NCBH benchmark NCBH1 with 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.*

The signal neural network output in Figure 9.47 is peaked at a NN output of around 0.88, with a broad shoulder extending to lower NN output values. The distribution is noticeably less peaked than the signal NN predictions of noncommutative black holes with a 200 GeV E_T^{miss} cut.

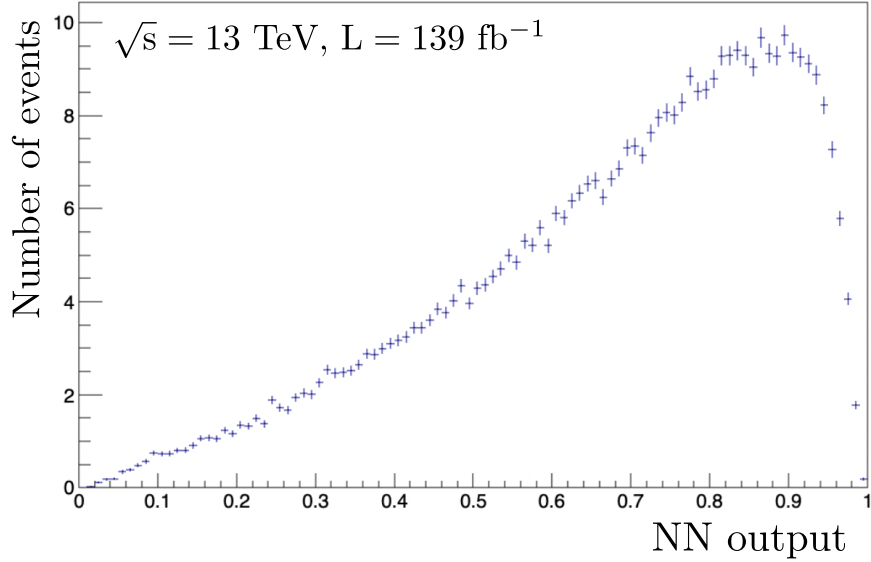


Figure 9.47 *NN prediction for signal events using NN for NCBH benchmark NCBH1 with 500 GeV E_T^{miss} cut and cuts from Table 9.1 applied. The decorrelated variable is the minimum ϕ between the closest large R jet and E_T^{miss} . The DisCo hyperparameter λ value is 2.*

9.7.5 Non-minimal Dark Sector Neural Network Results

Below is a presentation and discussion of the results of the non-minimal dark sector neural network when running with a DisCo λ parameter of 2. As noted, only the NMDS3 benchmark proceeds to final analysis results. However, it is instructive to also look at the behavior of the NMDS2 NN result at the 200 GeV E_T^{miss} cut to see how a non-minimal dark sector with a mass spacing of 50 GeV between ensemble constituents (which gives rise to a lower p_T spectrum and a less boosted final state than that of the NMDS3 benchmark) behaves. The general procedure for how all neural networks in this thesis were trained was described in 9.6. All non-minimal dark sector models were trained for 100 epochs.

The ROC curve in Figure 9.48 shows a background rejection of 90% for a roughly 63% signal retention. Just as the noncommutative black holes are less striking against background compared to the scalar cascades, the non-minimal dark sectors are less striking against background compared to scalar cascades, so it similarly makes sense that the neural network of the non-minimal dark sectors is not as robust as that of the scalar cascades. Nonetheless, the background rejection indicates that the non-minimal dark sector neural network is still performing well. As with the noncommutative black holes, the DisCo λ parameter is set to 2 for all non-minimal

dark sector neural networks in this thesis due to the degradation in performance observed at higher λ values.

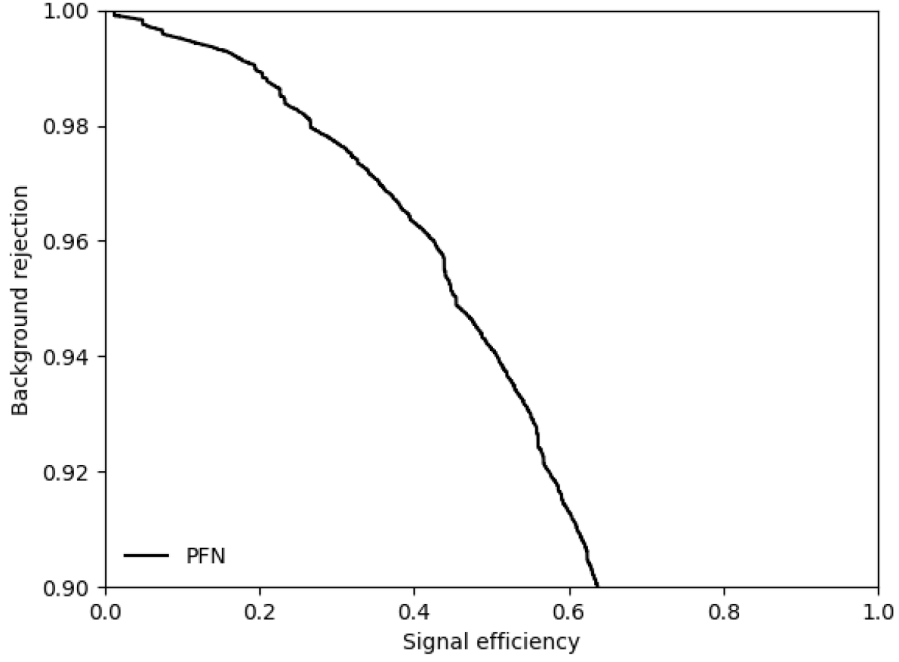


Figure 9.48 *ROC Curve of Background Rejection vs. Signal Efficiency using NN for NMDS benchmark NMDS2 with 200 GeV E_T^{miss} cut and basic cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} . The DisCo hyperparameter λ value is 2.*

From the plot in Figure 9.49 of output labels for background Monte Carlo simulated events for the non-minimal dark sector neural network, it can be seen that the background peaks at NN output values of roughly 0. There is a shallow shoulder, however, that extends to higher NN output values.

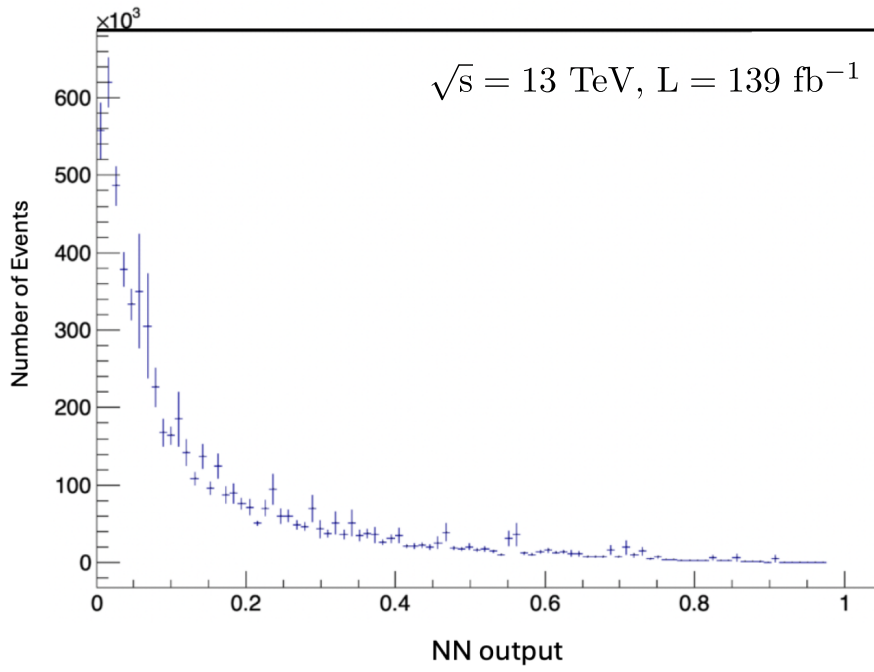


Figure 9.49 *NN prediction for background events using NN for NMDS benchmark NMDS2 with a 200 GeV E_T^{miss} cut and cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} . The DisCo hyperparameter λ value is 2.*

The signal NN output in Figure 9.50 is peaked at roughly 0.7, with a broad shoulder extending to lower NN output values. Together with the background NN output prediction, this would be consistent with a background rejection of 90% for a signal retention of 63%.

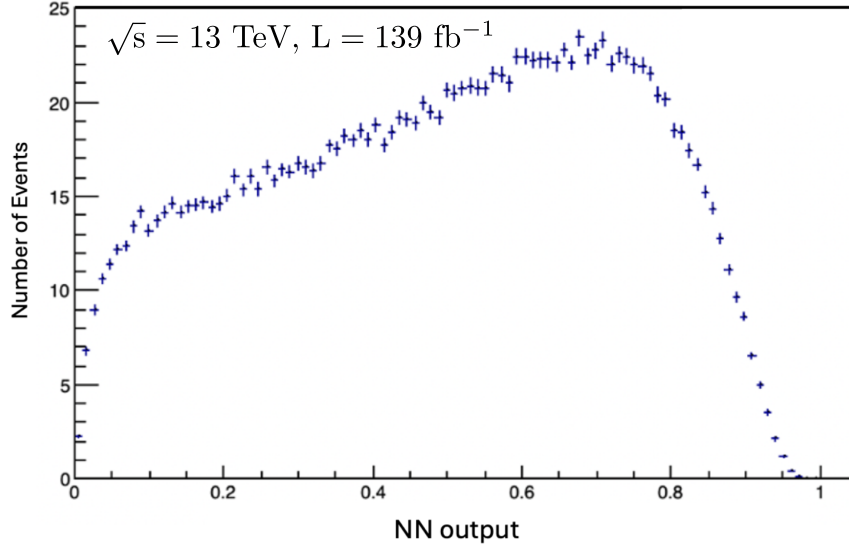


Figure 9.50 *NN prediction for signal events using NN for NMDS benchmark NMDS2 with a 200 GeV E_T^{miss} cut and cuts from Table 8.6 applied. The decorrelated variable is the p_T balance of the closest and further small R jet to the E_T^{miss} . The DisCo hyperparameter λ value is 2.*

However, even though the NN has the ability to reject background at a rate of 90% while retaining 63%, the NMDS benchmark NMDS2 struggled with the combination of large background events remaining the 200 GeV background cut coupled with the background uncertainty. Additional cuts degraded the performance of the NN.

Below is a presentation and discussion of the results of the NMDS benchmark NMDS3 (with the 500 GeV E_T^{miss} cut and 1400 GeV HT cut from Table 9.3) applied) when running with a DisCo λ parameter of 2. The ROC curve in Figure 9.51 shows a background rejection of 90% for a roughly 65% signal retention. Still less striking against background than scalar cascades, so it good to see that the NN performance is still performing well despite the harsher cuts.

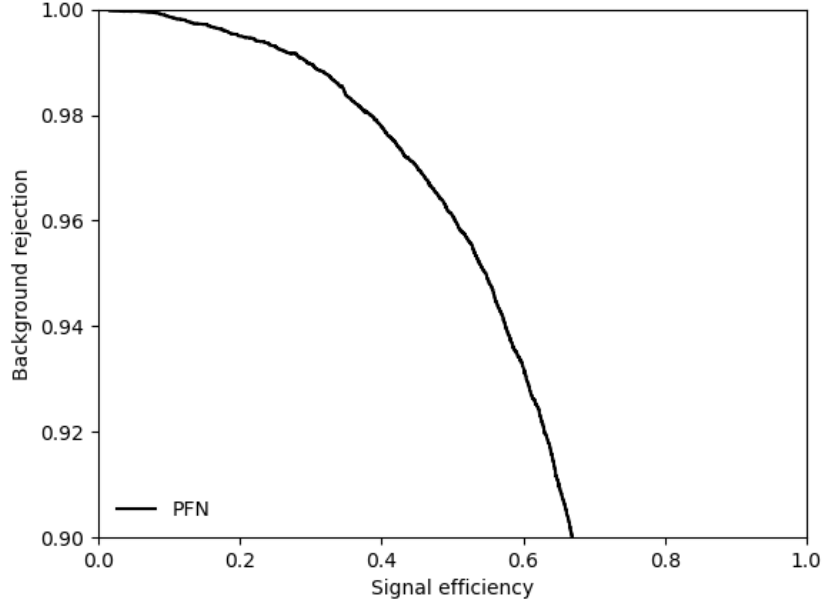


Figure 9.51 *ROC Curve of Background Rejection vs. Signal Efficiency using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.*

From Figure 9.52, it is positive to see that the accuracy of the neural network for both the training and validation datasets improves as the number of epochs that the neural network is trained for increases. The seemingly periodic structure in the accuracy curve is not a concern as this something that can be observed in machine learning training. The important thing is that the accuracy is improving for both datasets.

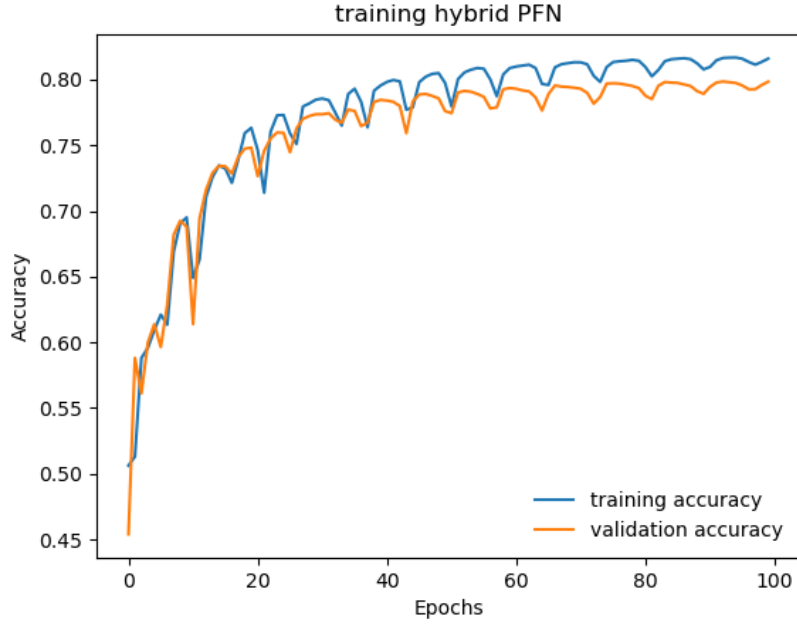


Figure 9.52 *Evolution of NN validation accuracy and NN training accuracy as a function of epochs using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.*

The loss of the neural network, seen in Figure 9.53, decreases over time for both the training sample and the validation sample, which is the desired behavior, as the goal in training is to minimize the loss. Again, the seemingly periodic structure in this curve is not a concern as this is something that can be observed in machine learning training. The important thing is that the loss is decreasing as a function of the number of epochs for both datasets.

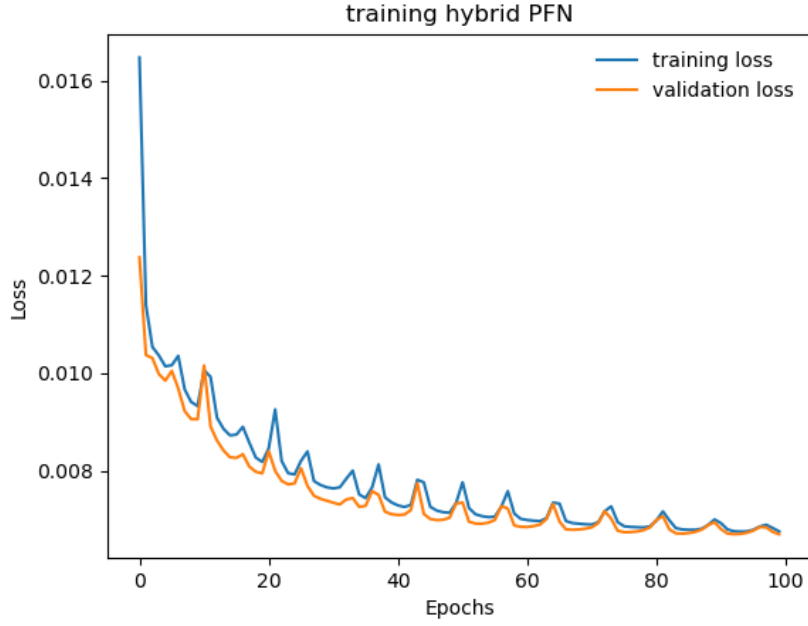


Figure 9.53 *Evolution of accuracy of NN validation loss and NN training loss as a function of epochs using NN for NMDs benchmark NMDs3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.*

Looking at the plot of output labels, in Figure 9.54, for background Monte Carlo simulated events for benchmark NMDs3, it can be seen that the background peaks at NN output values of 0.05. There is a shallow shoulder, however, that extends to higher NN output values.

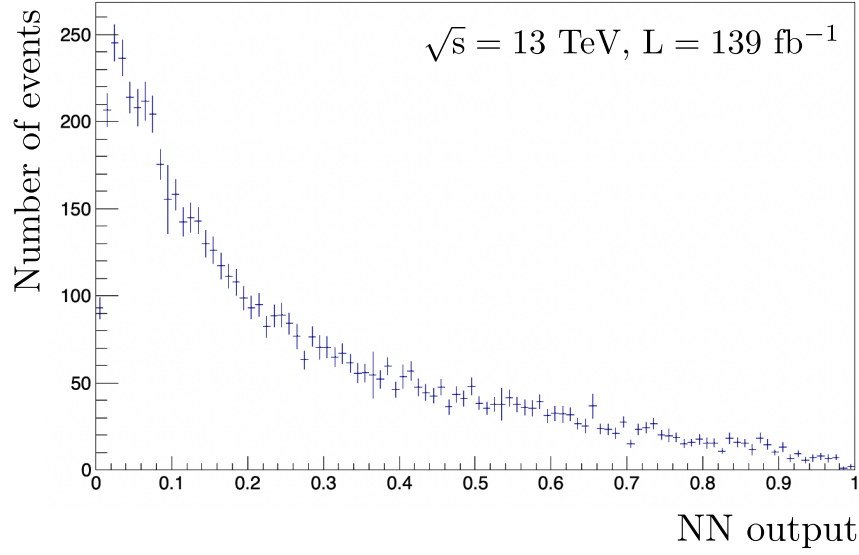


Figure 9.54 *NN prediction for background events using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.*

The signal neural network output in Figure 9.55 is peaked at a NN output of around 0.98, with a shoulder extending to lower NN output values. Together with the plot of the background NN output prediction, this would be consistent with a background rejection of 90% for a signal retention of 65%.

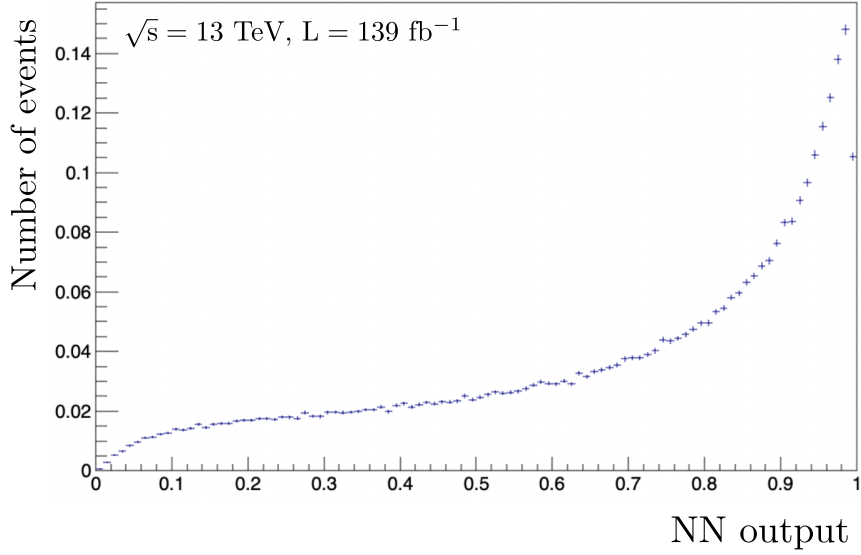


Figure 9.55 *NN prediction for signal events using NN for NMDS benchmark NMDS3 with a 500 GeV E_T^{miss} cut, 1400 GeV HT cut, and basic cuts from Table 9.3 applied. The decorrelated variable is the transverse sphericity. The DisCo hyperparameter λ value is 2.*

9.8 Summary

This chapter described how machine learning is used for signal region definition and background estimation in each search. A machine learning architecture was developed to use track information from the leading and subleading reclustered jet to learn the jet substructure of each model. Event level variables and the kinematic information of all small R jets in the event also served as input to the neural network. The machine learning architecture was used to create two decorrelated variables, the neural network output and a second variable that discriminates signal from background, for each analysis for use in the data-driven estimation of the Standard Model background processes that will be described in the next chapter. The decorrelated variables for scalar cascade benchmark SC1, non-minimal dark sector benchmark NMDS3, and noncommutative black hole benchmark NCBH1 are the number of tracks in the second leading reclustered jet, the $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, and the transverse sphericity, respectively. Monte Carlo and data comparisons were performed and showed an overall satisfactory agreement prior to the use of machine learning.

The performance of the neural network for each analysis was evaluated to show

that the neural networks are performing well. Out of the three models, the neural network for scalar cascade benchmark SC1 was the most robust, providing a background rejection of 99% through a signal retention of 95%. The robustness of the neural network is due to the strikingly high track multiplicity in the second leading reclustered jet, which peaked at roughly 100 tracks. Non-minimal dark sectors and noncommutative black holes are less distinguishable from background than scalar cascades and, as a result, their neural networks did not perform as well as that of SC1. The neural network for the NMDS benchmark NDMS3 performed second best with a 90% background retention for a 65% signal retention. The neural network for the noncommutative black hole benchmark NCBH1 provided a 90% background retention for a 47%. Due to its low cross section and trigger efficiencies, the SC1 benchmark has the lowest significance. The significance of scalar cascade benchmark SC1 prior to defining the signal region is 0.0053, while the significance of noncommutative black hole benchmark NCBH1 is 2.4, and the significance of nonminimal dark sector benchmark NMDS3 is 0.05. As a general observation, optimization is needed to improve the analysis sensitivity. The next chapter will describe how this is achieved for each analysis via a cut on the neural network output and a cut on the decorrelated variable.

Chapter 10

Background Estimation and Statistical Analysis

This chapter describes the derivation of the data-driven background estimation via the ABCD background estimation method. Monte Carlo simulation does not precisely model hadronic signatures and tends to have large uncertainties on the multijet Monte Carlo modeling. As a solution, the ABCD background estimation method uses data to estimate the background estimation in the signal region. The last chapter showed how the analyses in this thesis employ machine learning to construct two decorrelated variables. This chapter describes how these two variables are used in the ABCD background estimation method. The ABCD method creates a 2D plane, called an ABCD plane, where one decorrelated variable forms the y axis and the second decorrelated variable forms the x . The ABCD plane is split into four quadrants, wherein the background contribution in the signal region is determined from the data counts in the other 3 quadrants. Signal is largely localized in one region. The other 3 quadrants are referred to as the control regions. The Monte Carlo background and data distributions are compared in the control regions to ensure that the neural network is generalizing well and behaving as expected. Since this is a data-driven approach, the agreement between Monte Carlo and data does not have to be perfect. Checks are performed to ensure that the two variables are decorrelated so that they can be used to make the prediction of background estimation in the signal region. A nonclosure uncertainty is applied to account for any remaining correlation. The statistical uncertainties and systematic uncertainties on the background estimation are also calculated and shown, providing a total background yield in the signal region of the ABCD plane.

10.1 Background Estimation

As mentioned in the previous chapter, the background estimation for the analysis is reliant on the identification of two independent variables. Distance Correlation (DisCo) was introduced as a means to force a variable to be decorrelated with respect to the neural network output. The two variables used in the ABCD background estimation for this analysis are, therefore, the neural network output and the variable that DisCo forces to be decorrelated from the neural network output.

In the case of scalar cascades, the variable that DisCo forces to be decorrelated with respect to the neural network output is the number of tracks in the second leading reclustered jet. In the ABCD background estimation, a plane, such as the one depicted in Figure 10.1, is constructed by plotting one independent variable on the x axis and the second independent variable on the y axis.

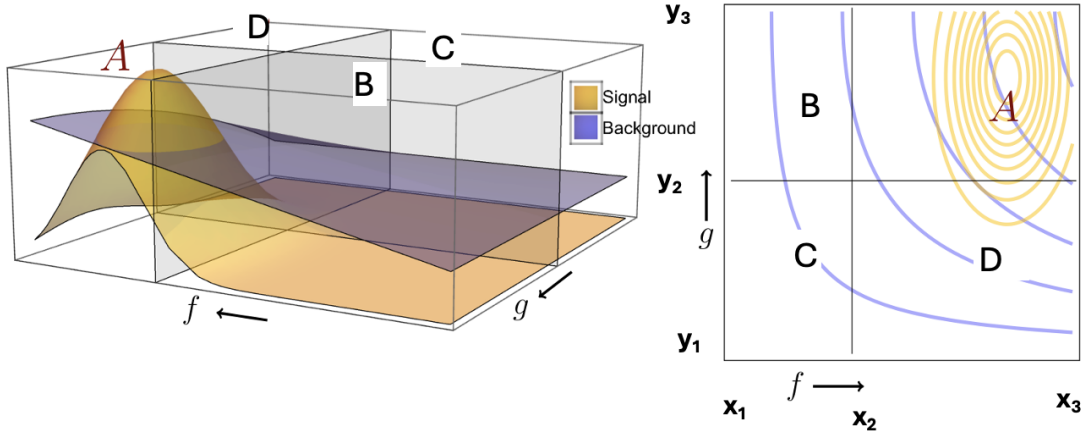


Figure 10.1 *Example of an ABCD plane. Figure modified from Reference [115]*

This plane is divided into four different regions in a manner such that a single region is created that contains the vast majority of the signal. This signal region will be referred to as region A, while the other regions are called region B, region C, and region D. In Figure 10.1, region A extends from x_2 to x_3 and y_2 to y_3 . Region D extends from x_2 to x_3 and y_1 to y_2 , while region B extends from x_1 to x_2 and y_2 to y_3 , and region C extends from x_1 to x_2 and y_1 to y_2 .

The ABCD background estimation method is a data-driven background approach. Monte Carlo simulation does not precisely model hadronic signatures and tends to have large uncertainties on the multijet Monte Carlo modeling. As a remedy, the ABCD background estimation uses data to estimate the background in the signal

region. It is customary for physics analyses at the LHC to not look at the data in the signal region until all aspects of the analysis have been thoroughly validated by the collaborations. This is to ensure that analyses do not bias the experimental result by optimizing the analysis based on observed trends in the data in the signal region. This practice of not looking at the data in the signal region is called blinding the data. (The signal region is then referred to as blinded.) Unblinding is the term given to the act of looking at the data in the signal region once all aspects of the analysis have been validated. Signal-free regions in which data can be looked at prior to unblinding are called the control regions. Regions B, C, and D will serve as the control regions in the ABCD plane. The control regions are used to validate the various aspects of the analysis. In the case of the ABCD background estimation, the control regions are used to estimate the background contribution in signal region A. Control regions are additionally used to make comparisons between data and Monte Carlo.

In the case there is not signal contamination in regions B, C, and, D, the number of background events in region A can be determined with the following relation

$$N_A^{\text{exp}} = \frac{N_D N_B}{N_C}. \quad (10.1)$$

This expression uses the principles of proportionality and that it assumes negligible correlation between x and y.

The statistical uncertainty on the expected number of events in the signal region is a Poisson, which is the typical uncertainty for counting numbers and which can be approximated by a Gaussian standard deviation, $\sigma_{\text{stat}}(N_A)^{\text{exp}}$, defined as

$$\sigma_{\text{stat}}(N_A)^{\text{exp}} = \sqrt{N_A^{\text{exp}}}. \quad (10.2)$$

The relation is approximately valid for large N. For this analysis, the author later demonstrates that the approximation is valid for $N = 6.6$ events. The statistical uncertainty on the expected number of events in any given region of the ABCD plane can be described as

$$\sigma(N_i)_{\text{stat}} = \sqrt{N_i}. \quad (10.3)$$

The statistical uncertainties on background regions B, C, and D can be added in quadrature to predict the systematic uncertainty on signal region A, such that

$$\sigma_{\text{sys}}(N_A^{\text{exp}}) = \sqrt{\sigma(N_B)_{\text{stat}}^2 + \sigma(N_C)_{\text{stat}}^2 + \sigma(N_D)_{\text{stat}}^2}. \quad (10.4)$$

A more meticulous way to write this is multiply the systematic uncertainty on each region by the gradient of the number of events in signal region A with respect to the number of events in the given region. The previous equation then becomes

$$\sigma_{\text{sys}}(N_A^{\text{exp}}) = \sqrt{\left(\frac{\partial N_A^{\text{exp}}}{\partial N_B}\right)^2 \sigma(N_B)_{\text{stat}}^2 + \left(\frac{\partial N_A^{\text{exp}}}{\partial N_C}\right)^2 \sigma(N_C)_{\text{stat}}^2 + \left(\frac{\partial N_A^{\text{exp}}}{\partial N_D}\right)^2 \sigma(N_D)_{\text{stat}}^2}.$$

The use of the gradient, ie: $\frac{\partial N_A^{\text{exp}}}{\partial N_i}$, where i is either B, C, or D, allows for the influence of the uncertainties on the number of events in a given region to be weighted by the influence of the number of events on the signal region. If the change in the number of events in, for example, region B does not affect the number of events in region A, then the uncertainty on the number of events in region B is irrelevant. The gradient would be 0 and, therefore, given the multiplication, the contribution to the systematic uncertainty of the ABCD plane from that region on the signal region would be zero. Conversely, if a small change in, for example, the number of events in region B has a large effect on the number of events in the signal region, then the uncertainty from region B on the signal region is very important.

Using equation 10.1, this expression can then be reduced to

$$\sigma_{\text{sys}}(N_A^{\text{exp}}) = \sqrt{\frac{N_B^2 N_D^2}{N_C^3} + \frac{N_D^2 N_B}{N_C^2} + \frac{N_B^2 N_D}{N_C^2}}. \quad (10.5)$$

To validate the ABCD background estimation, the ABCD technique is applied and tested outside of the signal region in the ABCD plane. There are a number of different ways to do this validation. Figure 10.2 illustrates one such set up.

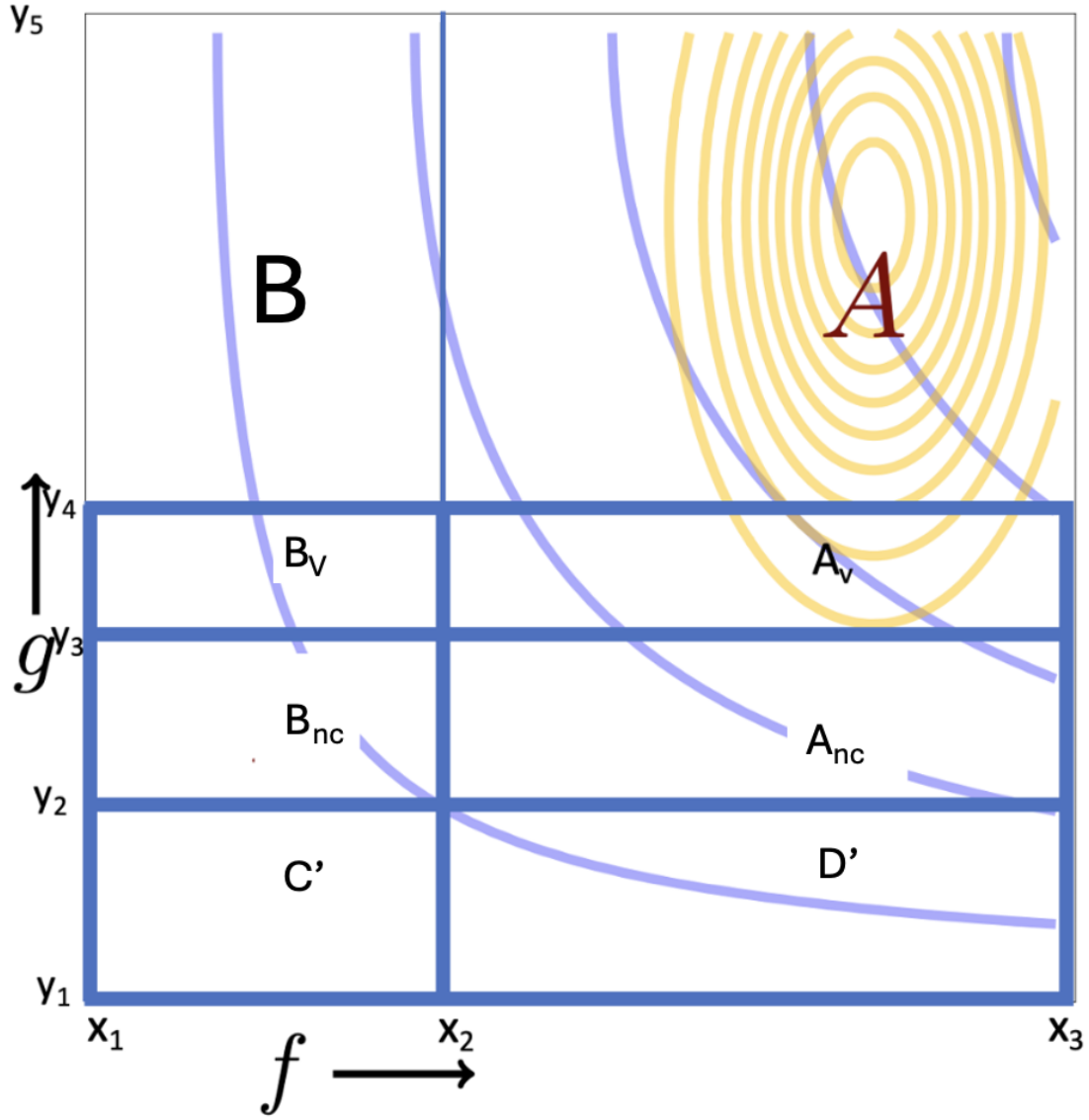


Figure 10.2 *Example of splitting an ABCD plane into additional regions for nonclosure and validation studies. Figure modified from Reference [115]*

Here, a second ABCD region is constructed within the original ABCD plane. The region formed from combining regions A_V and A_{nc} is the region for which the background will be estimated via the ABCD method. The “V” in A_V stands for validation and the “nc” in A_{nc} stands for nonclosure. This region spans from x_2 to x_3 and y_2 to y_4 and will be called region A' . The region A' is a validation region and, necessarily, is not signal enriched. Region D' , defined as the region spanning from x_2 to x_3 and y_1 to y_2 , will serve as a control region. Similarly, region B' , defined as the region spanning from x_1 to x_2 and y_2 to y_4 . Region C' , spanning from x_1 to x_2 and y_1 to y_2 , also serves as a control region. The number of events in region A' can

be found by the ABCD relation

$$N_{A'} = \frac{N_{B'}N_{D'}}{N_{C'}}. \quad (10.6)$$

The objective of creating this second ABCD plane is to determine whether the predicted number of background events in region A' is the same, within a given uncertainty, as the actual number of events in region A' . Adding in quadrature the statistical uncertainty of the number of events in region A' and the systematic uncertainty from the ABCD plane on region A' yields the following total uncertainty on the number of background events in region A'

$$\sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2}. \quad (10.7)$$

If the absolute value of the difference between the number of expected background events in region A' , $\frac{N_{B'}N_{D'}}{N_{C'}}$ and the actual number of background events in region A' is greater than Equation 10.7, then an uncertainty, called the nonclosure uncertainty, must be calculated in order to account for the discrepancy. In other words, the nonclosure uncertainty accounts for any correlation in the ABCD plane.

In order to determine the nonclosure uncertainty, another ABCD plane must be created and tested. To do so, regions D' , C' , B_{nc} , and A_{nc} are used. The number of events in A_{nc} is then calculated as

$$N_{A_{\text{nc}}} = \frac{N_{D'}N_{B_{\text{nc}}}}{N_{C'}}. \quad (10.8)$$

The relative nonclosure uncertainty, σ_{nc}^r , is then determined as

$$\sigma_{\text{nc}}^r = \frac{|N_{A_{\text{nc}}} - N_{A_{\text{nc}}}^{\text{exp}}|}{N_{A_{\text{nc}}}^{\text{exp}}}. \quad (10.9)$$

The nonclosure uncertainty on region A_{nc} , $\sigma_{\text{nc}}(N_{A_{\text{nc}}}^{\text{exp}})$, is obtained by multiplying the relative nonclosure uncertainty by the number of expected events in region A_{nc} such that

$$\sigma_{\text{nc}}(N_{A_{\text{nc}}}^{\text{exp}}) = \sigma_{\text{nc}}^r N_{A_{\text{nc}}}^{\text{exp}}. \quad (10.10)$$

Another ABCD plane is created to test this nonclosure uncertainty. The new ABCD plane is formed from the following regions: region A_V , region B_V , region D'' , C'' , which is a region formed by merging region C and region B_{nc} , and D'' , which is a region formed by merging region D and A_{nc} . The expected number of events in region A_V is calculated via the ABCD relation such that

$$N_{A_V} = \frac{N_{B''} N_{D''}}{N_{C''}}. \quad (10.11)$$

The nonclosure uncertainty on the number of events in region A_V is obtained by multiplying the relative nonclosure determined earlier by the expected number of events in region A_V , such that

$$\sigma_{\text{nc}}(N_{A_V}^{\text{exp}}) = \sigma_{\text{nc}}^r N_{A_V}^{\text{exp}}. \quad (10.12)$$

If $|N_{A_V, \text{observed}} - N_{A_V, \text{expected}}| < \sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2 + \sigma_{\text{nc}}(N_{A_V}^{\text{exp}})^2}$, the ABCD method is considered fully validated.

10.2 ABCD Analysis for Scalar Cascades

As noted earlier, the ABCD plane for the scalar cascades analysis is obtained by plotting the neural network output on the y axis and the number of tracks in the second leading reclustered jet on the x axis. The author produced all subsequent plots with the machine learning framework created by the author. After running the neural network prediction, the author defined the ABCD plane by applying cuts on the NN output and the second decorrelated variable.

The performance optimization of the neural network at decorrelating the number of tracks in the second leading reclustered jet from the neural network output was carried out by the author for a scan of DisCo hyperparameter λ values from 0 to 100 in increments of 5. A DisCo λ value of 0 corresponds to a standard binary

cross entropy loss function in which no DisCo loss has been added. The neural network that ran with a DisCo λ value of 50 provided an ABCD plane with a strong decorrelation of variables x and y .

The decorrelation between the x variable of the ABCD plane and the y variable of the ABCD plane can be obtained using the Pearson correlation coefficient calculation

$$r = \frac{\Sigma(x_i - \bar{x})\Sigma(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \Sigma(y_i - \bar{y})^2}}. \quad (10.13)$$

Two variables that are perfectly decorrelated will have a Pearson correlation coefficient of 0, while two variables that are perfectly correlated with have a Pearson correlation coefficient of 1, and two variables that perfectly anti-correlated will have a Pearson correlation coefficient of -1.

For the 2D plane ranging from number of tracks 0 to 250 on the x axis and NN output 0 to 1 on y axis and produced for a neural network run with DisCo λ parameter 50 (Figure 10.3), the Pearson correlation coefficient is 0.10, which corresponds to a correlation of 10%. This indicates, as seen in Figure 10.3, that the plane is very largely uncorrelated. A small fraction of background extends to higher track multiplicity at lowest neural network outputs and an even smaller fraction extends to higher track multiplicity at higher track multiplicity values. Otherwise, the background ABCD plane, seen on the left hand side of Figure 10.3, appears very largely uncorrelated.

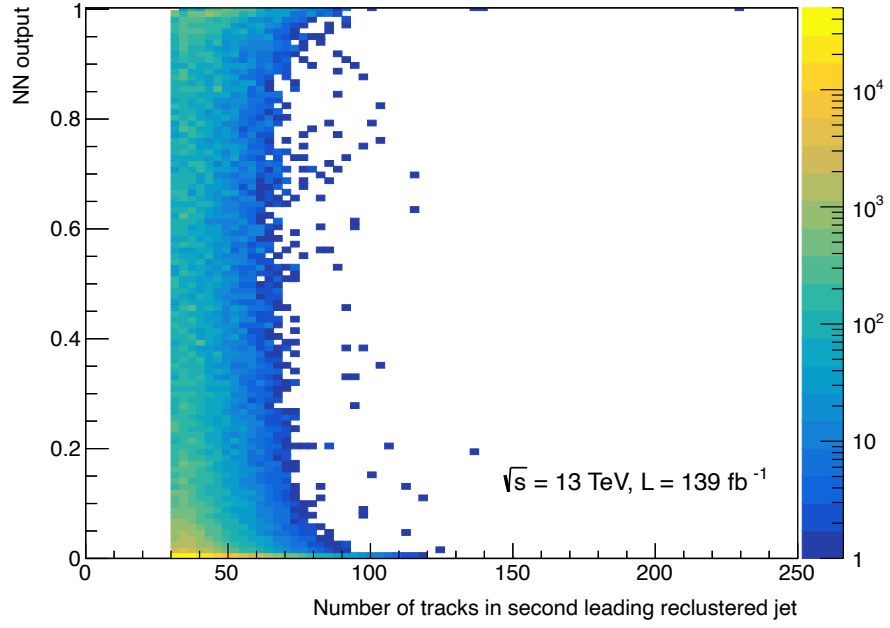


Figure 10.3 *MC Background ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis.*

The signal ABCD plane, Figure 10.4 can be seen to have the majority of events above a neural network output of 0.95 and above 80 tracks in the second leading reclustered jet.

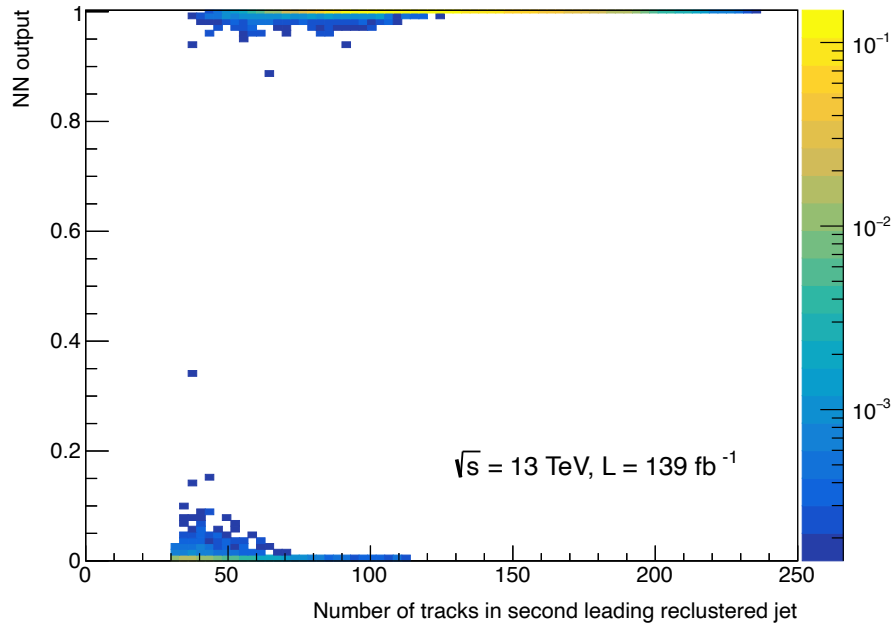


Figure 10.4 *MC Signal ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis.*

Each region of the ABCD plane for the scalar cascades is defined in Table 10.1.

Region	Range of Number of tracks in second leading reclustered jet (x axis)	Range of NN output (y axis)
A	80 to 250	0.95 to 1
B	0 to 80	0.95 to 1
C	0 to 80	0 to 0.95
D	80 to 250	0 to 0.95

Table 10.1 *Definitions of regions in the ABCD plane for the Scalar Cascade Analysis*

These regions are drawn on the MC background ABCD plane in Figure 10.5.

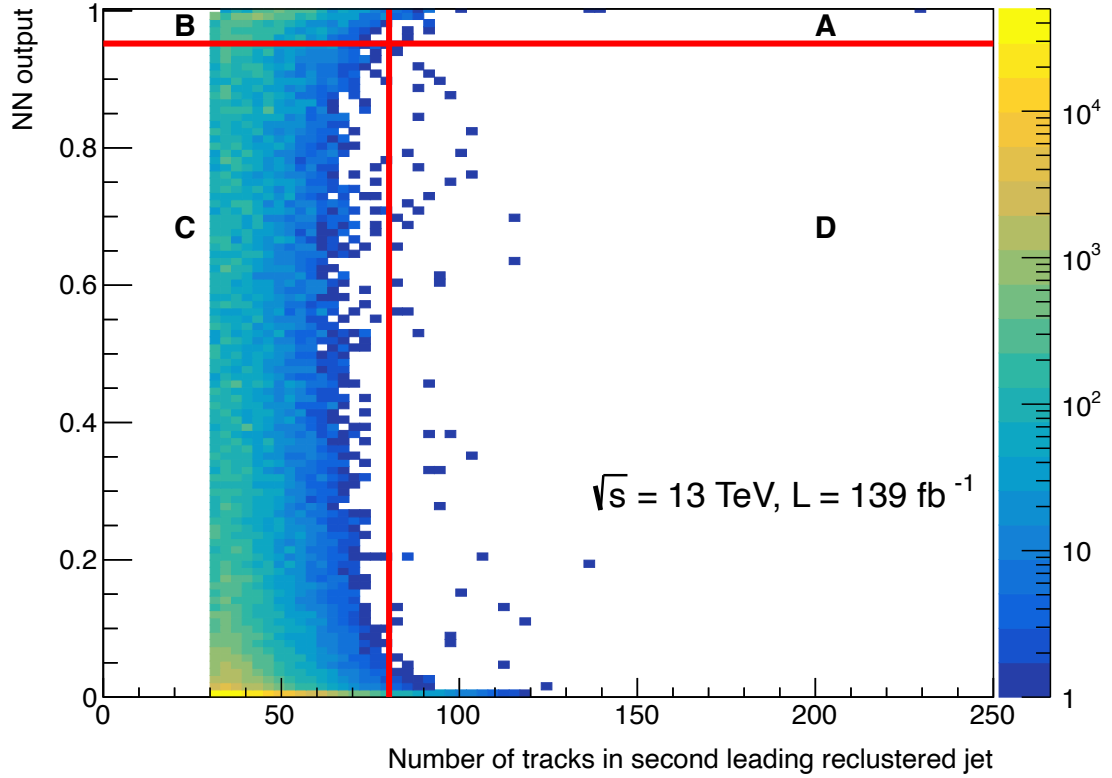


Figure 10.5 *MC Background ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis. ABCD region definitions are drawn on plane.*

The regions are drawn on the signal ABCD plane in Figure 10.6.

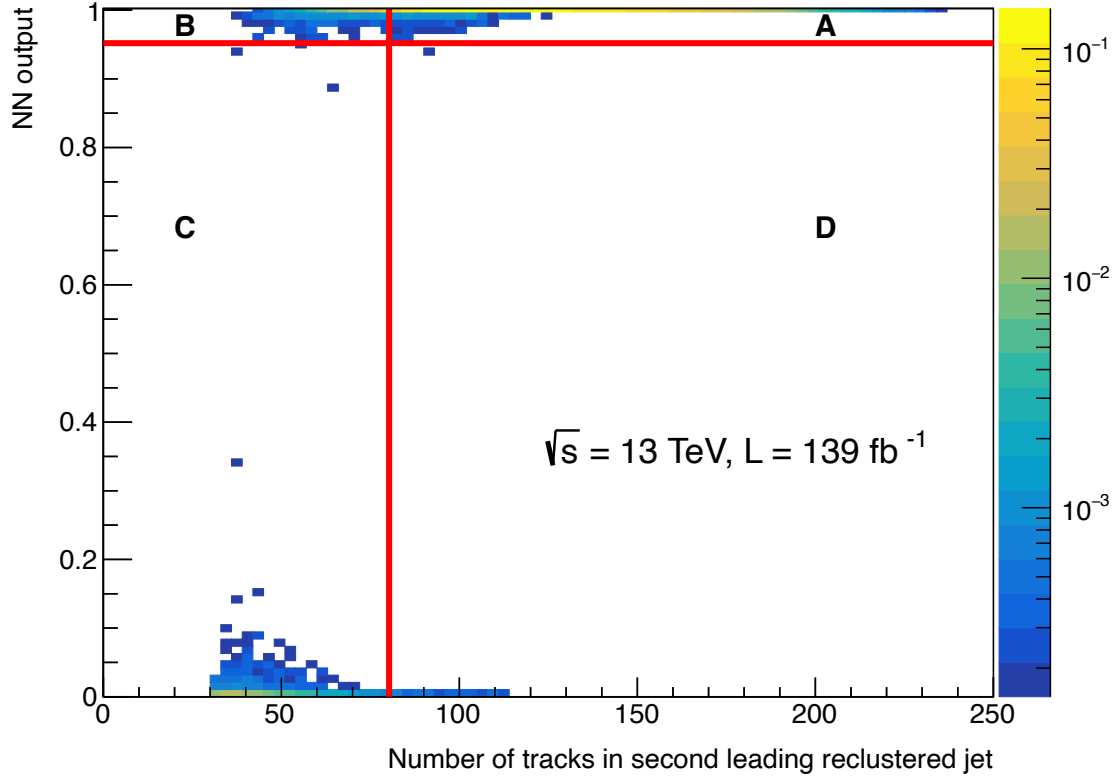


Figure 10.6 *MC Signal ABCD Plane with DisCo hyperparameter $\lambda = 50$ for Scalar Cascade Analysis. ABCD region definitions are drawn on plane.*

The signal, MC background, and data yields in each region of the ABCD plane for scalar cascades are outlined in Table 10.2. (All studies in this chapter and all analysis chapters use the full Run 2 data of 139 fb^{-1} . The data samples are defined in Table 8.1.)

Region	Signal yields	MC yields	Data Yields	% of Total Signal	$100 \times \frac{\text{Signal}}{\text{Background}}$
A: $80 < T < 250, 0.95 < NN < 1.0$	2.43	39.01	Blinded	90%	N/A
B: $0 < T < 80, 0.95 < NN < 1.0$	0.13	8895	5147	4.81%	0.00146%
C: $0 < T < 80, 0 < NN < 0.95$	0.12	247500	201900	4.44%	0.0000484%
D: $80 < T < 250, 0 < NN < 0.95$	0.017	411	259	0.0063%	0.0041%

Table 10.2 *Monte Carlo Yields, Data Yields, and Signal Yields for the A, B, C, and D Regions of the ABCD plane for the Scalar Cascade Analysis. Column “% of Total Signal” gives the percent of total signal events in the ABCD plane that are contained in a given region. The column “ $100 \times \frac{\text{Signal}}{\text{Background}}$ ” turns into a percent the fraction of signal events in a given region with respect to the number of background events in the region. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Table 10.2 shows that the vast majority of the signal lies in the signal region and, consequently, the validation regions have safely been formed in regions where there is very little signal compared to the total number of signal events in the plane. Additionally, the amount of signal in the validation regions is very small compared to the amount of background in the validation regions. This is the desired behavior. Background is also behaving as expected as the majority of it lies outside of the signal region.

The power of the neural network at discriminating signal from background can be seen by comparing the number of background events that survive all cuts prior to the machine learning to the number of background events in the signal region, after the application of machine learning. The original 256845 simulated background events were reduced to 39.01 events in the signal region. The significance (defined in Equation 8.3) obtained is 0.39. The significance prior to the NN and ABDC plane was 0.0053. With a 73.4 times improved significance, these results suggest a robust means for reducing background from the scalar cascade signal. While this significance is too small to set a limit on the cross section of the scalar cascade sample used, a limit can be set for higher cross sections. A powerful comparison to make is that of that significance at the 200 GeV E_T^{miss} and basic cuts from Table 8.6. Compared to an original significance of 0.002, the cuts from Table 8.4 and the cuts from the ABCD cuts on NN output and number of tracks in the second leading reclustered jet, allowed for a 182.6 times improved significance. This is impressive given the scalar cascades analysis only had 3.9 scalar cascade events from the most basic event selections in Table 8.6 and managed to preserve 2.70 signal events while providing significant reduction to background (as evidenced by the increase in significance). This serves as a testament to how striking this phase space is.

The passages below show the validation of the ABCD plane and results of the statistical analysis performed on data. The behavior of data was checked against the behavior of Monte Carlo in the control regions of the ABCD plane (regions B, C, and D). As seen in Figure 10.7, the normalized distributions of background and data are within 2σ uncertainties for all but one bin in the control region B.

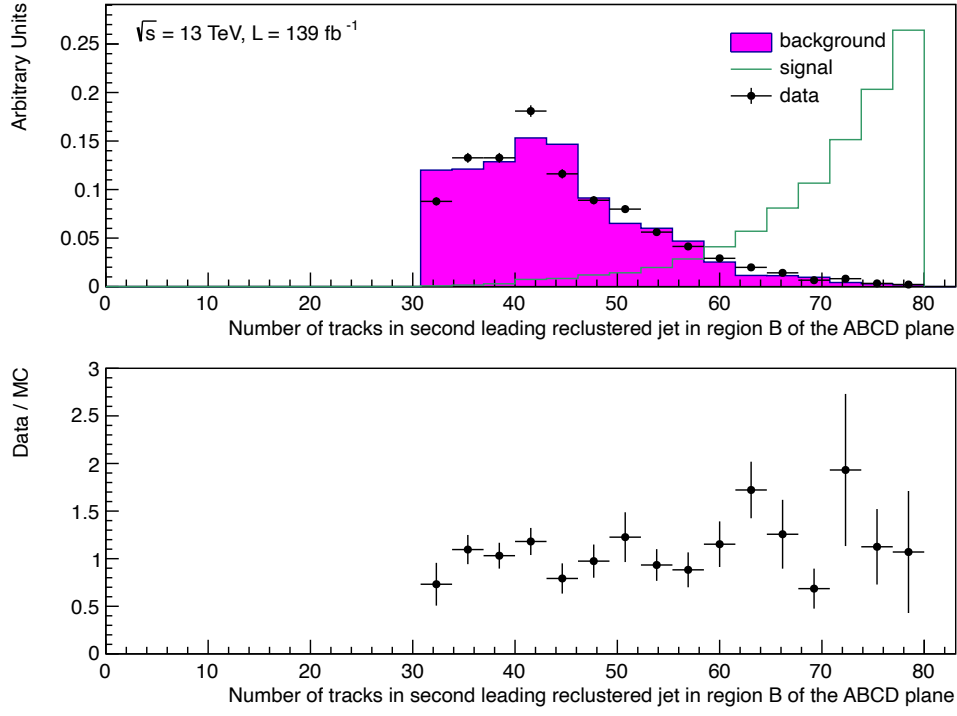


Figure 10.7 *Top plot: Number of tracks in the second leading reclustered jet in Control Region B of the ABCD plane for the Scalar Cascade Analysis. Distribution is normalized. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in Control Region B of the ABCD plane. This ratio is created from the normalized distribution shown in the top plot.*

Similarly, as seen in Figure 10.8, data and Monte Carlo are within either 2σ uncertainties or %10 in control region C.

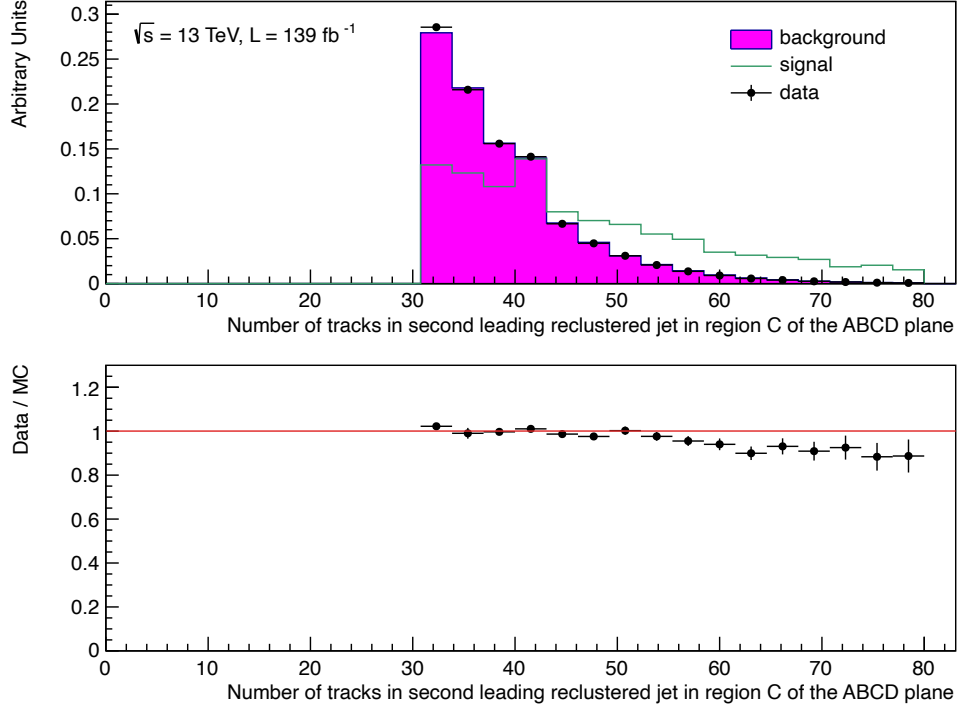


Figure 10.8 *Top plot: Number of tracks in the second leading reclustered jet in Control Region C of the ABCD plane for the Scalar Cascade Analysis. Distribution is normalized. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in Control Region C of the ABCD plane. This ratio is created from the normalized distribution shown in the top plot.*

Data and Monte Carlo are within 2σ uncertainties for all but two bins in control region D, seen in Figure 10.9. The anomalously weighted bin can safely be ignored as the error bars encompass a MC and data ratio value of 1.

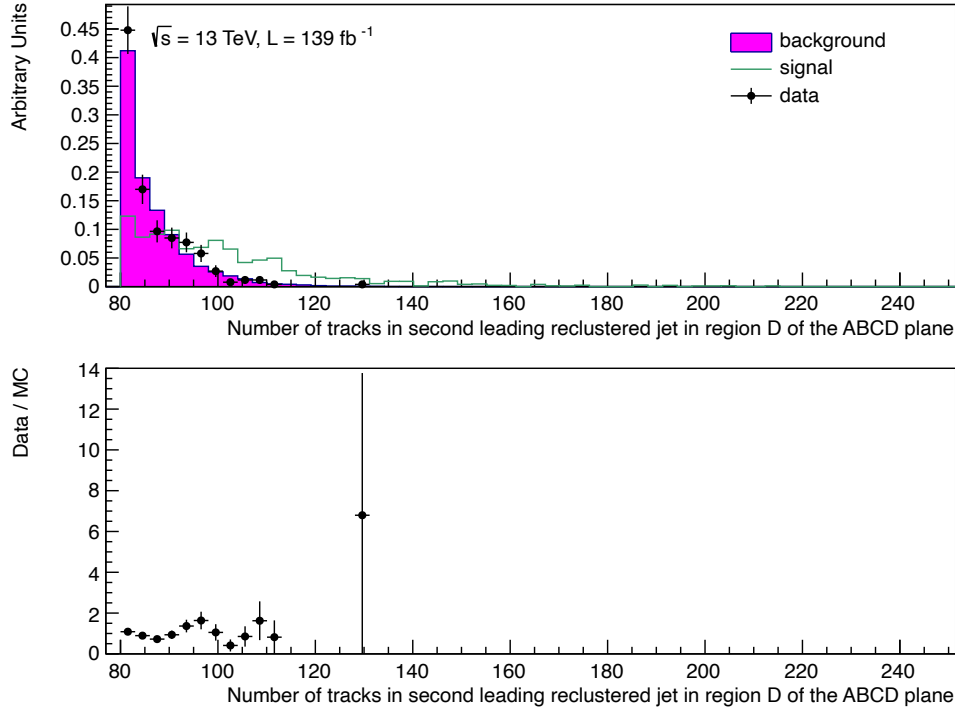


Figure 10.9 *Top plot: Number of tracks in the second leading reclustered jet in Control Region D of the ABCD plane for the Scalar Cascade Analysis. Distribution is normalized. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in Control Region D of the ABCD plane. This ratio is created from the normalized distribution shown in the top plot.*

In the signal region A, data is blinded, but the difference in shape between background and signal can be seen below.

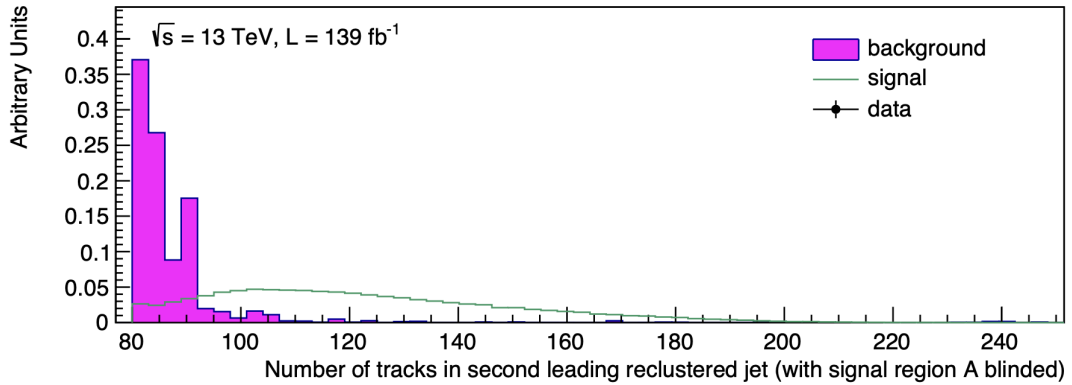


Figure 10.10 *Number of tracks in the second leading reclustered jet in Signal Region (Region A) of the ABCD plane for the Scalar Cascade Analysis. Comparison of distributions for MC background and MC signal. Data is absent. Distributions are normalized.*

When combining all four regions of the ABCD plane for MC and comparing the three control regions of data, still blinded in the signal region, MC and data are shown to largely agree within uncertainties, as seen in Figure 10.11. Though comparing four regions (3 control region and 1 signal region) of Monte Carlo background yields to only the 3 control regions of data, this is a reasonable comparison to make due to the low expected background yield (only 39 events) in the signal region.

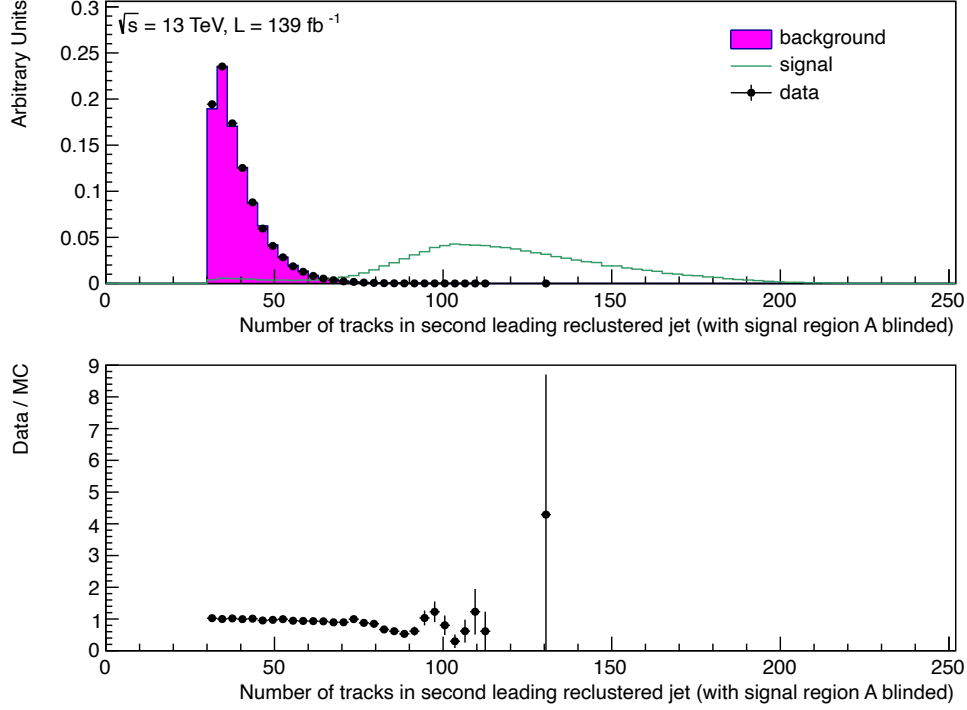


Figure 10.11 *Top plot: number of tracks in the second leading reclustered jet for all regions of the ABCD plane (Region A, B, C, and D) combined for Monte Carlo. This is compared to the number of tracks in the second leading reclustered jet for the control regions (Regions B, C, and D) for data. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in the four regions of the ABCD plane for Monte Carlo vs. the three control regions of the ABCD plane for data. This ratio is created from the normalized distribution shown in the top plot. This is for the Scalar Cascade Analysis.*

Finally, the distribution of the NN output from all four regions of the ABCD plane combined for Monte Carlo are within uncertainties of the distribution of the NN output from the three control regions for data. Again, this a reasonable comparison to make due to the low expected background yield (only 39 events) in the signal region.

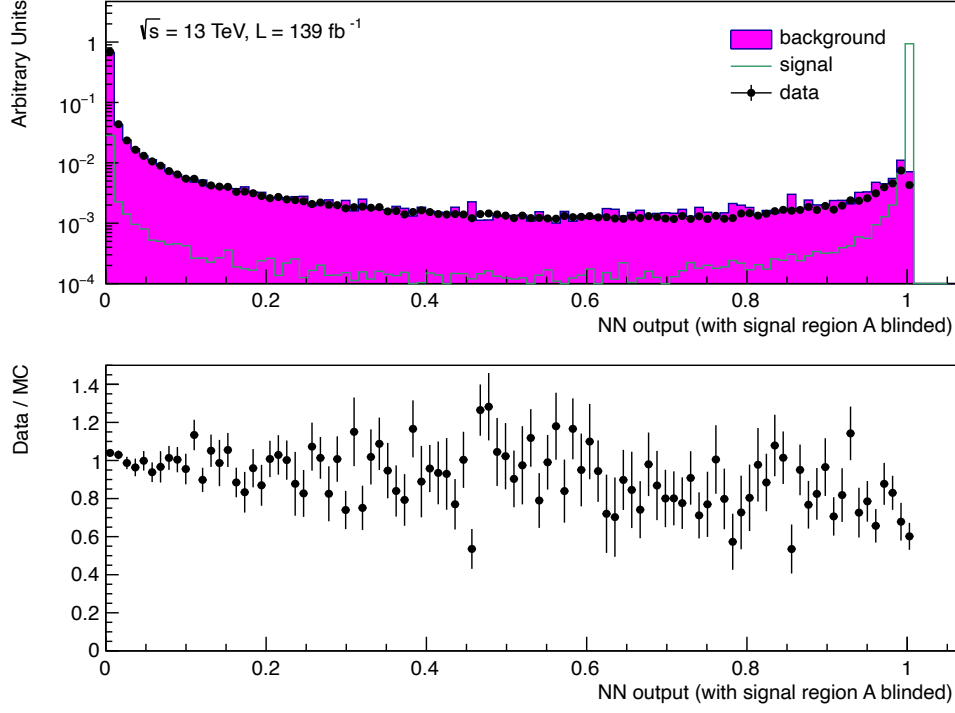


Figure 10.12 *Top plot: NN output for all regions of the ABCD plane (Region A, B, C, and D) combined for Monte Carlo. This is compared to all control regions (Regions B, C, and D) for data. Bottom plot: Ratio of MC to data for the distribution of the number of tracks in the second leading reclustered jet in the four regions of the ABCD plane for Monte Carlo vs. the three control regions of the ABCD plane for data. This ratio is created from the normalized distribution shown in the top plot. This is for the Scalar Cascade Analysis.*

The ABCD plane of data events blinded in the signal region is shown in Figure 11.2, closely resembling that of the MC ABCD plane.

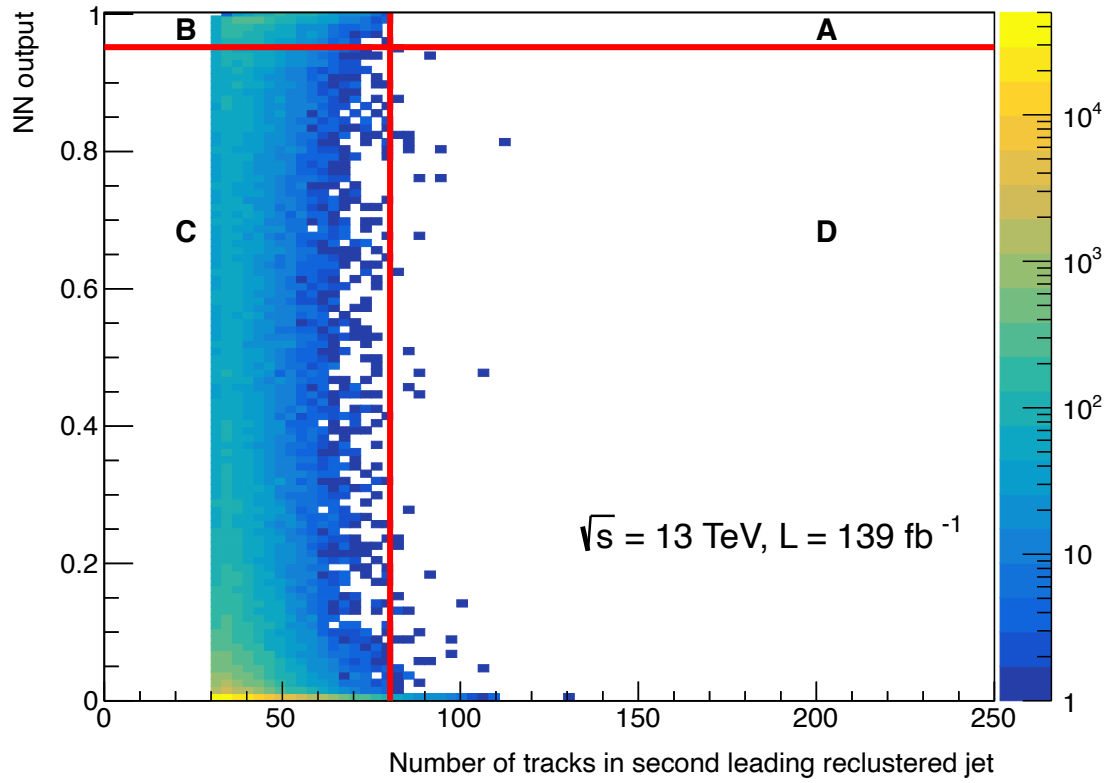


Figure 10.13 *Data ABCD plane with the signal Region blinded. This is for the Scalar Cascade Analysis. Full Run 2 dataset of 139 fb^{-1} used.*

Following the procedure outlined in the previous chapter, two smaller ABCD planes were formed within the ABCD plane specified above in order to determine a nonclosure uncertainty and in order to validate the ABCD plane method.

The validation regions for MC background for the Scalar Cascade analysis are shown in Figure 10.14.

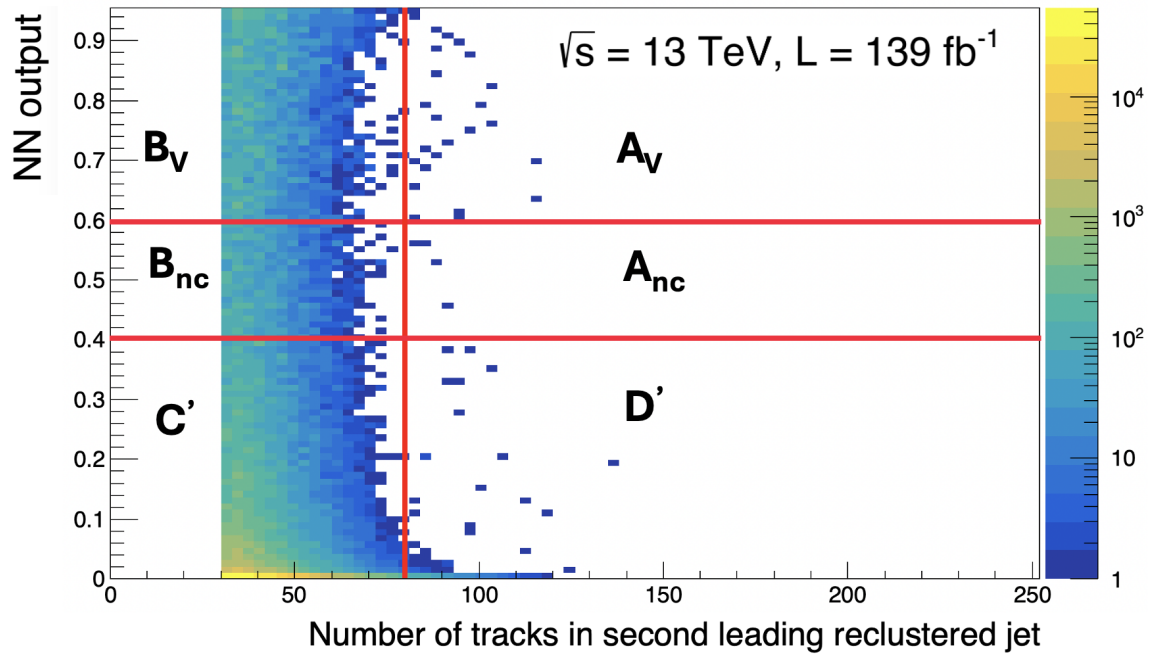


Figure 10.14 *Nonclosure and Validation Regions of ABCD plane for MC background for Scalar Cascade analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.*

The validation regions for data for the Scalar Cascade analysis are shown in Figure 10.15.

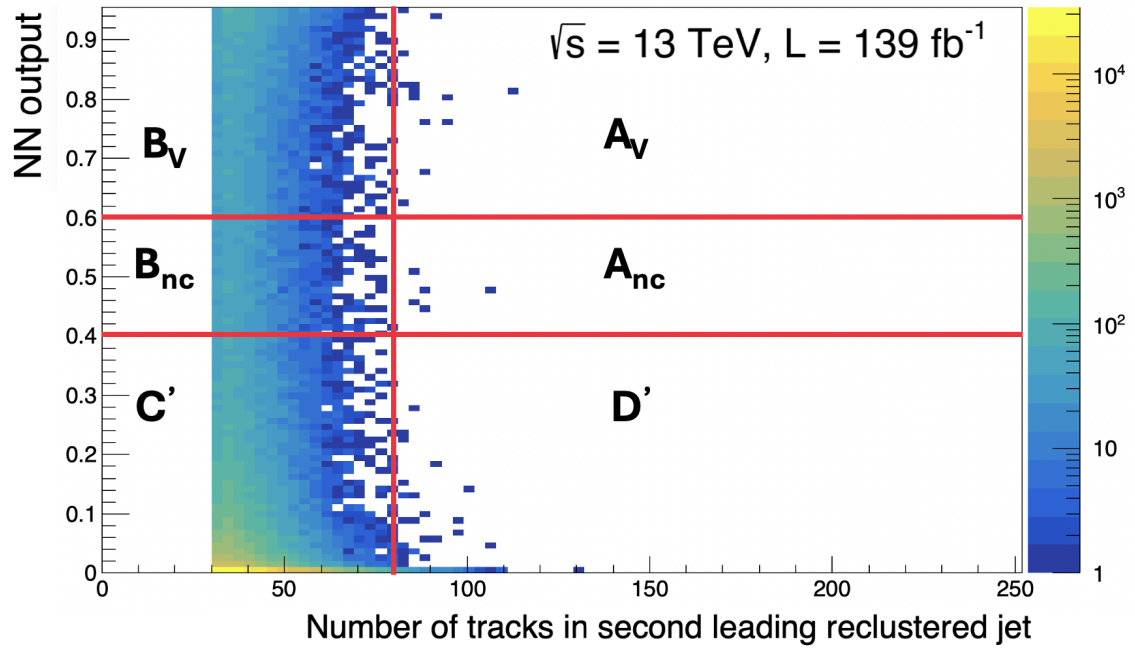


Figure 10.15 *Nonclosure and Validation Regions of ABCD plane for data for Scalar Cascade analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.*

The region of the ABCD plane that is used to determine the nonclosure uncertainty spans from a NN output of 0 to 0.6, where the division into different regions on the y axis occurs at 0.4, and a range of 0 to 250 tracks in the second leading reclustered jet, where the x axis is divided at 80. Table 10.3 summarizes the MC and data yields in the various regions of this plane. Both data and MC have small Pearson correlation coefficients indicating a high level of decorrelation. The MC nonclosure uncertainty is 45.82 %, while the data nonclosure uncertainty is 25.78 %. As a reminder, the nonclosure uncertainty is calculated via Equation 10.9 and accounts for any correlation in the ABCD plane.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A_{nc} : $80 < T < 250, 0.4 < NN < 0.6$	5.87	5	10.83	6.74
B_{nc} : $0 < T < 80, 0.4 < NN < 0.6$	6382.5	5235		
C' : $0 < T < 80, 0 < NN < 0.4$	189075	186500		
D' : $80 < T < 250, 0 < NN < 0.4$	320.93	240		
MC NC Uncertainty	45.82%			
Data NC Uncertainty	25.78%			
MC Corr. Coeff	-0.011			
Data Corr. Coeff	0.016			

Table 10.3 *Monte Carlo Yields and Data Yields in subvalidation regions A_{nc} , B_{nc} , C' , D' used to test for nonclosure. The definitions of A_{nc} , B_{nc} , C' , D' are found in Figure 10.2. Monte Carlo and Data Nonclosure Uncertainties (MC NC Uncertainty, Data NC Uncertainty), Pearson Correlation Coefficients for MC (MC Corr. Coeff) and Data (Data Corr. Coeff), expected MC and Data event yields (Exp MC events and Exp Data yields) are also included. The number of tracks in the second leading reclustered jet is denoted as ‘ T ’ and the NN output is denoted as NN . This is for the NCBH Analysis.*

The region of the ABCD plane that is used to apply the nonclosure uncertainty (Equation 10.9) and test for closure spans from a NN output of 0 to 0.95, where the division into different regions on the y axis occurs at 0.6, and a range of 0 to 250 tracks in the second leading reclustered jet, where the x axis is divided at 80. Table 10.4 summarizes the MC and data yields in the various regions of this plane. Both data and MC have small Pearson correlation coefficients indicating a relatively high level of decorrelation. The MC nonclosure uncertainty is 25.06%, while the data nonclosure uncertainty is 7.18%. The nonclosure of 25.8 % from the previous plane provides closure to this validation plane.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A_V : $80 < T < 250, 0.6 < NN < 0.95$	19.60	14	26.15	13.06
B_V : $0 < T < 80, 0.6 < NN < 0.95$	15514.22	10220		
$B_{nc} + C'$: $0 < T < 80, 0 < NN < 0.6$	234413.53	191700		
$A_{nc} + D'$: $80 < T < 250, 0 < NN < 0.6$	395.07	245		
MC NC Uncertainty	-25.06%			
Data NC Uncertainty	7.18%			
MC Corr. Coeff	0.026			
Data Corr. Coeff	0.047			

Table 10.4 *Monte Carlo Yields and Data Yields in subvalidation regions A_V , B_V , $A_{nc} + D'$, and $B_{nc} + C'$ used to validate nonclosure uncertainty. The definitions of A_V , B_V , $A_{nc} + D'$, and $B_{nc} + C'$ are found in Figure 10.2. Monte Carlo and Data Nonclosure Uncertainties (MC NC Uncertainty, Data NC Uncertainty), Pearson Correlation Coefficients for MC (MC Corr. Coeff) and Data (Data Corr. Coeff), expected MC and Data event yields (Exp MC events and Exp Data events) are also included. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘ T ’ and the NN output is denoted as NN.*

To reduce verbiage in the captions of event yields in the remainder of this caption, let MC Corr. Coeff be known as the Pearson Correlation Coefficient for MC, Data Corr. Coeff be known as the Pearson Correlation Coefficient for data, Data Yields be the Observed Data Yields, MC yields be the observed MC yields, Exp MC events be the expected MC event yield, Exp Data events be the expected data event yield, MC NC Uncertainty be the MC nonclosure uncertainty, and Data NC Uncertainty be the data nonclosure uncertainty.

However, it is not sufficient to say the nonclosure of 25.8 % validates that plane, as the MC in the signal region of the ABCD plane that is used in the actual scalar cascades analysis is 164 % and the discrepancy between 25.8 % and 164 % is large. Table 10.5 shows the breakdown of the yields, Pearson correlation coefficients, and uncertainties in the full ABCD plane spanning a NN output of 0 to 1 and number of tracks in the second leading reclustered jet from 0 to 250, where the signal region is defined as 80 to 250 tracks and 0.95 NN output to a NN output of 1.

Region	MC yields	Data Yields	Exp MC events	Exp Data events	Signal Yields
A	39.01	Blinded	14.77	6.60	2.43
B	8895	5147			0.13
C	247500	201900			0.12
D	411	259			0.017
MC NC Uncertainty	164.08%				
Data NC Uncertainty	Blinded				
MC Corr. Coeff	0.098				
Data Corr. Coeff	Blinded				

Table 10.5 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D in the ABCD plane used for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

The MC exhibits a Pearson Correlation Coefficient of 0.10. As there are few events in the signal region, and as a slight arm extended to higher track multiplicities in higher NN output, the 164% nonclosure in the signal region is understandable.

The next passages detail the procedure undertaken to derive the nonclosure uncertainty for the scalar cascades analysis. This involves studying the difference between data and Monte Carlo nonclosure uncertainties for alternative sets of validation regions. Tables 10.6, 10.7, 10.8, 10.9, 10.10, 10.11, 10.12, 10.13, and 10.14 show various alternative validation regions. For simplicity, the regions in the alternative validation regions are labeled in the tables as A, B, C, and D. Each of these regions represent some set of phase space with NN outputs below the signal region NN output of 0.95. Different combinations of cuts along the y axis, all under NN output values of 0.95, are made so as to study the Pearson correlation coefficients and nonclosure uncertainties as y varies.

For example, Tables 10.6 shows an alternative ABCD plane in which the NN output spans from 0 to 0.6, with the y axis dividing the ABCD regions at a NN output of 0.3. In Table 10.7, the NN output spans from 0 to 0.5, with the y axis diving at 0.3. In Table 10.13, NN output spans from 0 to 0.7, with the y axis diving at 0.4. The goal is to vary, in regions strictly below the signal region, either the range of the NN output used to define the ABCD plane, the point at which the NN is cut to form different regions in the ABCD, or both.

The Pearson correlation coefficients for the validation regions are all below 5% and the difference between signal and data nonclosure uncertainty (in global uncertainties, not relative uncertainties) is never more than 20%. In the instance

where there is a 20% difference, MC is higher than data. Taking the highest instance in which data is greater than MC, there is a global difference of 16%. To derive the nonclosure uncertainty in the signal region for data, the nonclosure uncertainty in the signal region for MC (164%) is taken (from Table 10.5) and the highest difference between data and MC (from Tables 10.6, 10.7, 10.8, 10.9, 10.10, 10.11, 10.12, 10.13, and 10.14) is added. The 20% difference (where MC is higher than data) is used to be conservative. This yields a conservative nonclosure of 184%. The nominal expected value of data events in the signal region, as predicted by the ABCD method, is 7 (6.6) events. The statistical uncertainty is 3 events, while the nonclosure uncertainty is 12 events, and the systematic uncertainty is 0.4 events. The errors are added in quadrature to yield a total uncertainty of 12 events.

NN range 0 to 0.3 to 0.6				
Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.3 < NN < 0.6$	13.87	10.40	23.59	17.69
B: $0 < T < 80, 0.3 < NN < 0.6$	13800	10350		
C: $0 < T < 80, 0 < NN < 0.3$	246800	185100		
D: $80 < T < 250, 0 < NN < 0.3$	421.9	316.43		
MC NC Uncertainty	-41.21%			
Data NC Uncertainty	-55.32%			
MC Corr. Coeff	-0.011			
Data Corr. Coeff	0.016			

Table 10.6 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.3 < NN < 0.5$	7.14	4	12.65	7.95
B: $0 < T < 80, 0.3 < NN < 0.5$	7401	6066		
C: $0 < T < 80, 0 < NN < 0.3$	185100	183200		
D: $80 < T < 250, 0 < NN < 0.3$	316.43	240		
MC NC Uncertainty	43.56%			
Data NC Uncertainty	49.66%			
MC Corr. Coeff	-0.015			
Data Corr. Coeff	0.013			

Table 10.7 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.3 < NN < 0.8$	17.265	9	29.732	17.65938
B: $0 < T < 80, 0.3 < NN < 0.8$	17392.5	13480		
C: $0 < T < 80, 0 < NN < 0.3$	185100	183200		
D: $80 < T < 250, 0 < NN < 0.3$	316.425	240		
MC NC Uncertainty	41.93%			
Data NC Uncertainty	49.04%			
MC Corr. Coeff	-0.0011			
Data Corr. Coeff	0.027			

Table 10.8 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.3 < NN < 0.7$	12.94	7	23.36	14.41
B: $0 < T < 80, 0.3 < NN < 0.7$	13665	11000		
C: $0 < T < 80, 0 < NN < 0.3$	185100	183200		
D: $80 < T < 250, 0 < NN < 0.3$	316.43	240		
MC NC Uncertainty	-44.62%			
Data NC Uncertainty	-51.42%			
MC Corr. Coeff	-0.011			
Data Corr. Coeff	0.022			

Table 10.9 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.3 < NN < 0.95$	26.6175	19	42.373769	24.5764
B: $0 < T < 80, 0.3 < NN < 0.95$	24787.5	18760		
C: $0 < T < 80, 0 < NN < 0.3$	185100	183200		
D: $80 < T < 250, 0 < NN < 0.3$	316.425	240		
MC NC Uncertainty	-37.18%			
Data NC Uncertainty	-22.69%			
MC Corr. Coeff	0.023			
Data Corr. Coeff	0.047			

Table 10.10 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.7 < NN < 0.95$	16.60	12	19.92	9.87
B: $0 < T < 80, 0.7 < NN < 0.95$	11909.60	7762		
C: $0 < T < 80, 0 < NN < 0.7$	238040.82	194200		
D: $80 < T < 250, 0 < NN < 0.7$	398.09	247		
MC NC Uncertainty	-16.68%			
Data NC Uncertainty	21.55%			
MC Corr. Coeff	0.026			
Data Corr. Coeff	0.047			

Table 10.11 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.5 < NN < 0.95$	23.58	15	31.64	16.38
B: $0 < T < 80, 0.5 < NN < 0.95$	18710.76	12700		
C: $0 < T < 80, 0 < NN < 0.5$	231239.65	189200		
D: $80 < T < 250, 0 < NN < 0.5$	391.07	244		
MC NC Uncertainty	-25.49%			
Data NC Uncertainty	-8.42%			
MC Corr. Coeff	0.026			
Data Corr. Coeff	0.047			

Table 10.12 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.4 < NN < 0.7$	8.4	7	16.45	9.90
B: $0 < T < 80, 0.4 < NN < 0.7$	9690	7692		
C: $0 < T < 80, 0 < NN < 0.4$	189075	186500		
D: $80 < T < 250, 0 < NN < 0.4$	320.93	240		
MC NC Uncertainty	49%			
Data NC Uncertainty	29%			
MC Corr. Coeff				
Data Corr. Coeff				

Table 10.13 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $80 < T < 250, 0.2 < NN < 0.95$	31.04	22	52.54	30.85
B: $0 < T < 80, 0.2 < S < 0.95$	30255	23260		
C: $0 < T < 80, 0 < S < 0.2$	179625	178700		
D: $80 < T < 250, 0 < S < 0.2$	311.93	237		
MC NC Uncertainty	40.92%			
Data NC Uncertainty	28.68%			
MC Corr. Coeff	0.023			
Data Corr. Coeff	0.047			

Table 10.14 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D for an alternative ABCD plane used for validation. This is for the Scalar Cascades Analysis. The number of tracks in the second leading reclustered jet is denoted as ‘T’ and the NN output is denoted as NN.*

10.3 NCBH ABCD Analysis

As described earlier, the DisCo loss function is used to decorrelate two variables, where one of them is the neural network output. These decorrelated variables form the two axes of the ABCD plane, which is used to estimate the background for the analysis. The ABCD plane for the noncommutative black hole benchmark NCBH1 is defined by plotting the minimum angle between the missing transverse energy and the closest large R jet ($\Delta\phi_{\min}(J, E_T^{\text{miss}})$) on the x axis, as the decorrelated variable, and the NN output on the y axis.

The MC background plane spanning the full NN range of 0 to 1 and the full range of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ can be found below in Figure 10.16

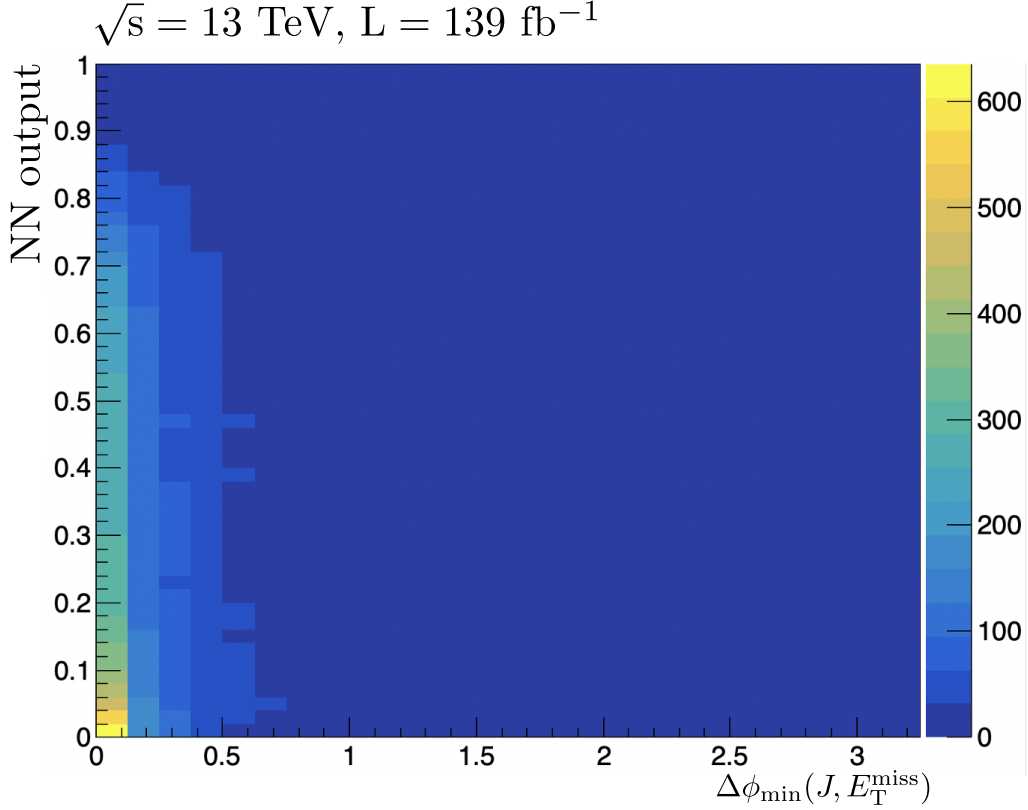


Figure 10.16 Plane showing NN output vs. $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for MC background for full range of NN output and full range of ϕ for NCBH Analysis of NCBH benchmark NCBH1

The MC background is seen to be predominantly below $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ values of 0.75. Creating a smaller ABCD plane by applying a cut that removes $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ values below 0.75 would both significantly reduce the MC background in the ABCD plane and remove the clustering of MC background events at lower NN values and low values of ϕ , thereby leading to better decorrelation.

The NCBH signal plane spanning the full NN range of 0 to 1 and the full range of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is shown in Figure 10.17

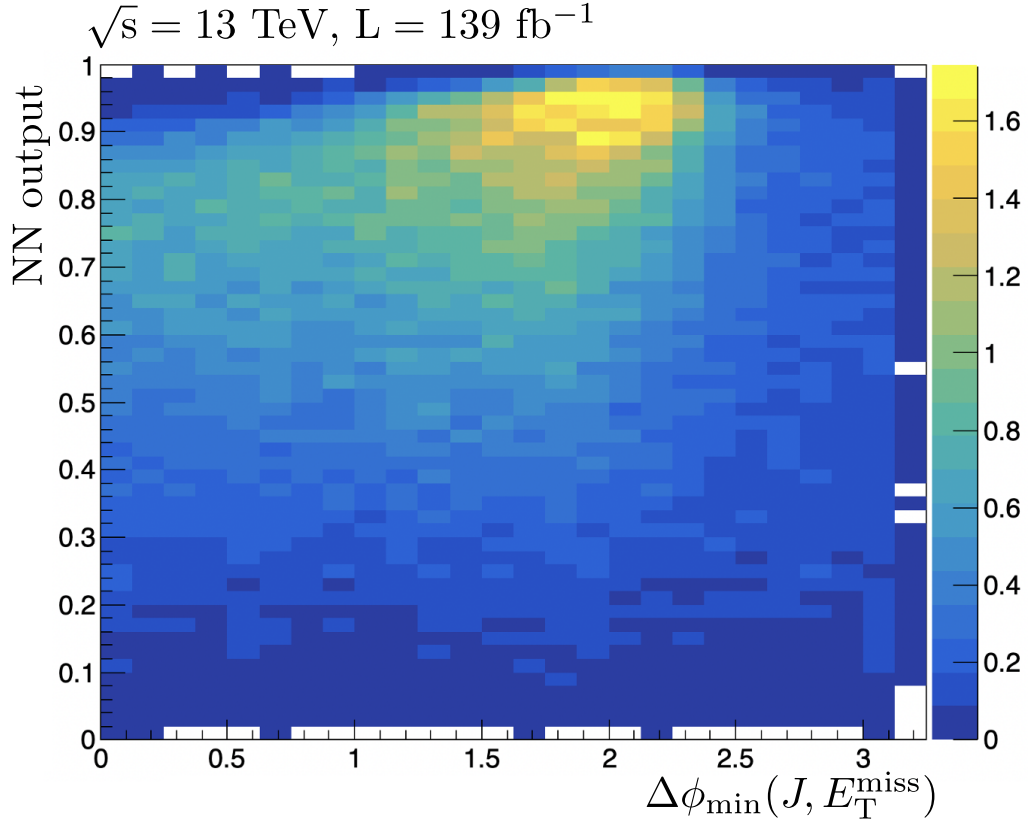


Figure 10.17 Plane showing NN output vs. $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for full range of NN output and range of ϕ . This is for the NCBH Analysis of NCBH benchmark NCBH1

The signal is contained predominantly in higher NN output values and higher values of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. However, signal does leak into lower values of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ and lower NN output values.

The 1D distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for background and signal can be found in Figure 10.18 and Figure 10.19, respectively.

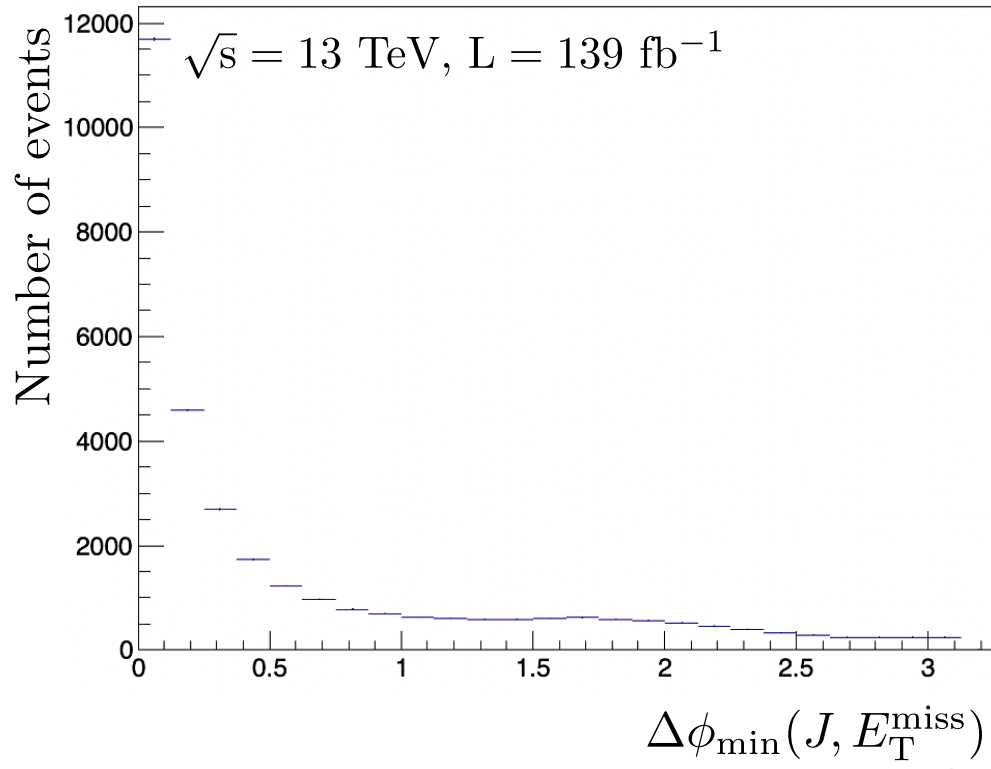


Figure 10.18 *1D Distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for MC background. This is for the NCBH Analysis of NCBH benchmark NCBH1*

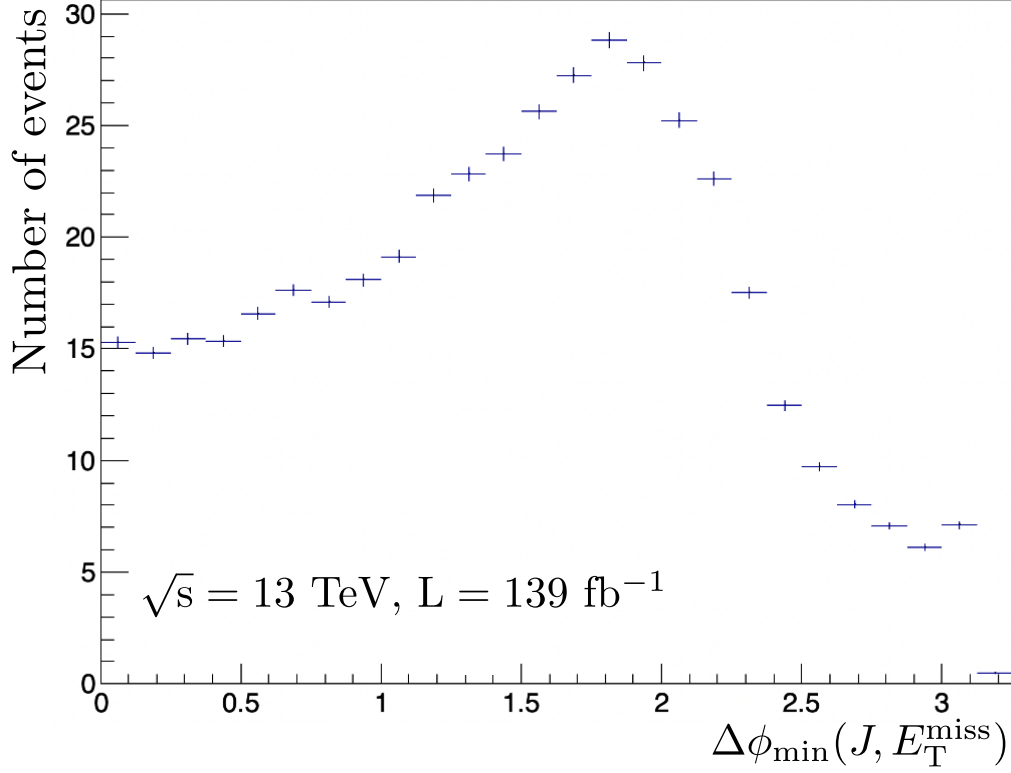


Figure 10.19 *1D Distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ for MC signal. This is for the NCBH Analysis of NCBH benchmark NCBH1*

In order to make a smaller, optimized ABCD plane, care must be taken to ensure that cuts are made to minimize the amount of signal contamination in the control regions. After performing a series of optimization studies, the author determined that applying a cut of $\Delta\phi_{\min}(J, E_T^{\text{miss}}) < 1.2$ would strongly reduce the MC background while also minimizing the signal contamination. This motivates having the x axis of the ABCD plane, the $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ values, span from 0.75 to 3.25, with the signal region containing ϕ values above 1.2. To mitigate signal contamination, the plane is confined to NN output values of 0.3 to 1, with the signal region encompassing values of NN output greater than 0.5.

The region definitions of the ABCD plane, as well as the Monte Carlo and Data Yields, expected number of events, and Pearson Correlation Coefficients are shown in Table 10.15.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.5 < NN < 1.0$	3134	Blinded	2697.2	3275.6
B: $0.75 < \phi < 1.2, 0.5 < NN < 1.0$	1058	1187		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.5$	678.6	782		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.5$	1730	2158		
MC NC Uncertainty	16.19%			
Data NC Uncertainty	Blinded			
MC Corr. Coeff	0.0236			
Data Corr. Coeff	Blinded			

Table 10.15 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions A, B, C, and D of the ABCD plane used for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is abbreviated as ‘ ϕ ’ and the NN output is denoted as NN.*

The final MC background ABCD plane for the NCBH benchmark NCBH1 is shown in Figure 10.20, appearing very largely uncorrelated.

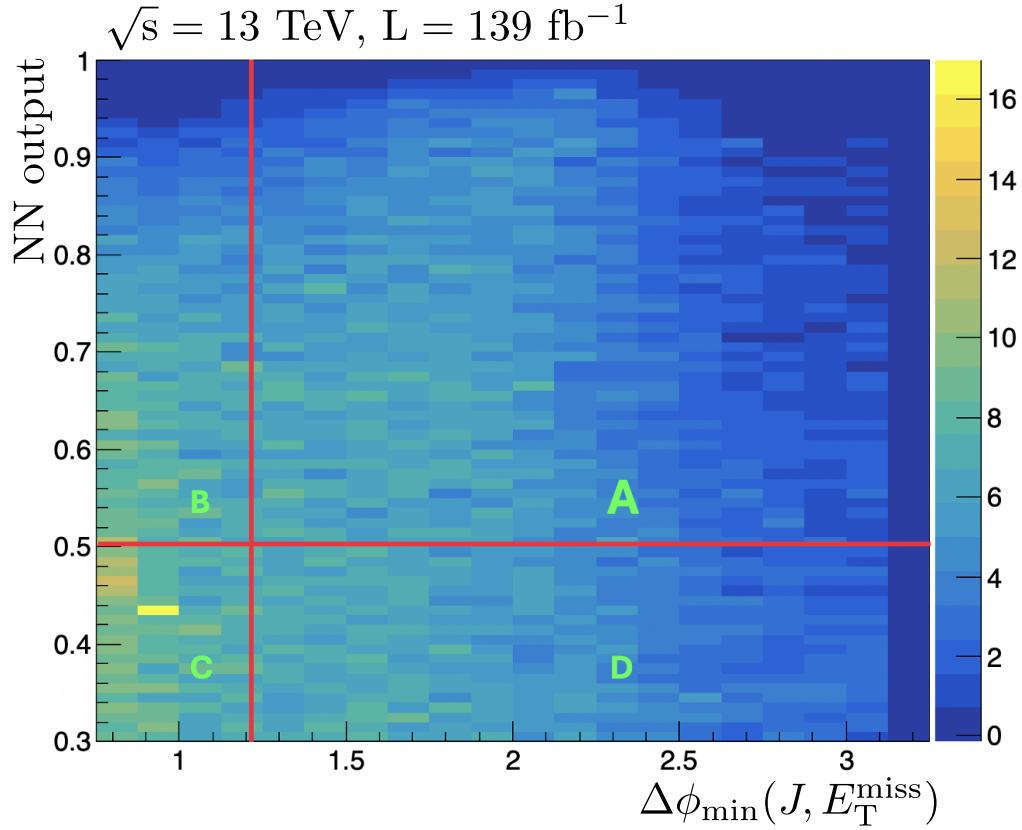


Figure 10.20 *MC background ABCD plane as a function of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ and NN output. This is for the NCBH Analysis of NCBH benchmark NCBH1.*

As seen in Table 10.15, there is relatively high decorrelation between the two

variables, as indicated by the Pearson Correlation Coefficient of 0.0236.

The signal ABCD plane for the NCBH benchmark NCBH1 is shown in Figure 10.21.

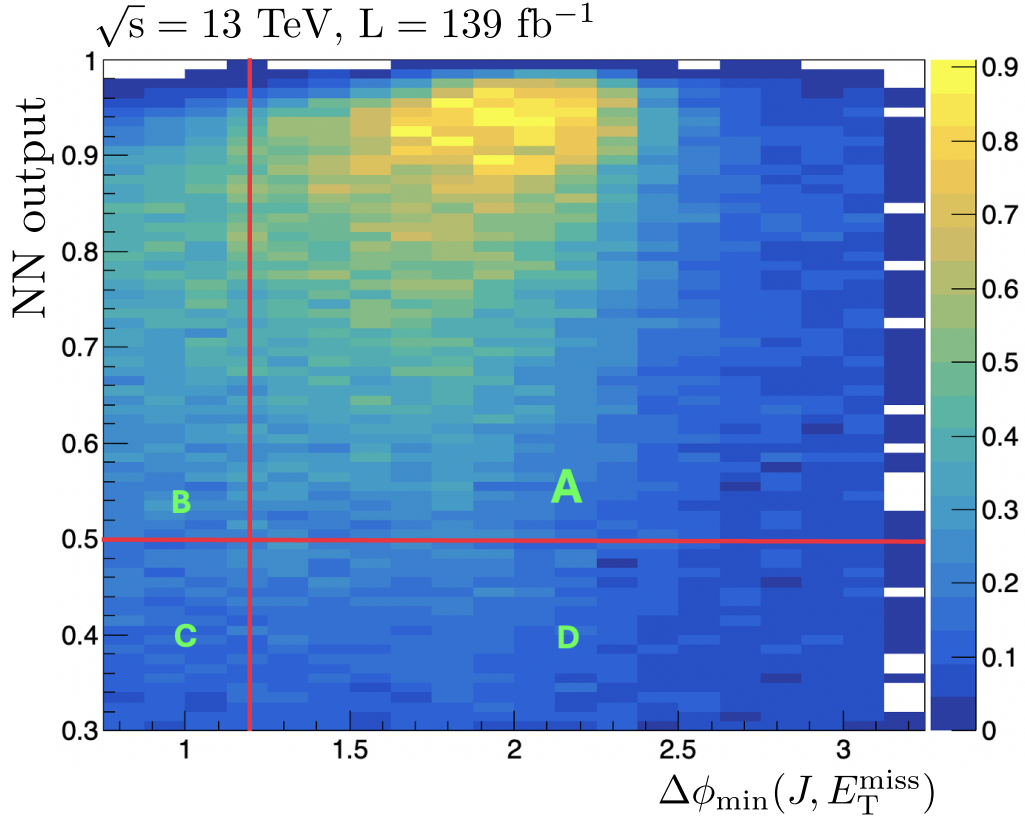


Figure 10.21 *Signal ABCD plane as a function of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ and NN output. This is for the NCBH Analysis of NCBH benchmark NCBH1.*

The signal yields for each region of the ABCD plane are defined in Figure 10.16.

Region	Signal yields	MC yields	% of Total Signal	$100 \times \frac{\text{Signal}}{\text{Background}}$
A: $1.2 < \phi < 3.25, 0.5 < \text{NN} < 1.0$	226	3134	69.3%	N/A
B: $0.75 < \phi < 1.2, 0.5 < \text{NN} < 1.0$	52.91	1058	16.2%	4.5%
C: $0.75 < \phi < 1.2, 0.3 < \text{NN} < 0.5$	10.41	678.6	3.2%	1.3%
D: $1.2 < \phi < 3.25, 0.3 < \text{NN} < 0.5$	36.95	1730	11.3%	2.2%

Table 10.16 *Monte Carlo Yields and Signal Yields for the A, B, C, and D Regions of the ABCD plane for the NCBH Analysis of NCBH benchmark NCBH1. Column “% of Total Signal” gives the percent of total signal events in the ABCD plane that are contained in a given region. The column “ $100 \times \frac{\text{Signal}}{\text{Background}}$ ” turns into a percent the fraction of signal events in a given region with respect to the number of background events in the region. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is abbreviated as ‘ ϕ ’ and the NN output is denoted as NN.*

Table 10.16 shows that the majority of the signal is largely contained in the signal region. Additionally, the amount of signal in the validation regions is small compared to the amount of background in the validation regions. This is the desired behavior.

The significance for the NCBH benchmark NCBH1 after the NN and ABCD plane cuts is 4. This significance is 1.68 times improved compared to the significance evaluated at the 500 GeV E_T^{miss} cut. Compared to the basic cuts at 200 GeV, the significance after the NN is 5.13 times improved. Though having a higher cross section than scalar cascades, noncommutative black holes are more difficult to distinguish from background. Still, the increase in significance provided by analysis cuts has greatly helped distinguish signal from background - just not to the extent that those of the Scalar Cascades analysis allowed for.

The behavior of data was checked against the behavior of Monte Carlo in the control regions of the ABCD plane (regions B, C, and D). As seen in Figure 10.22, the normalized distributions of background and data are within 2σ uncertainties.

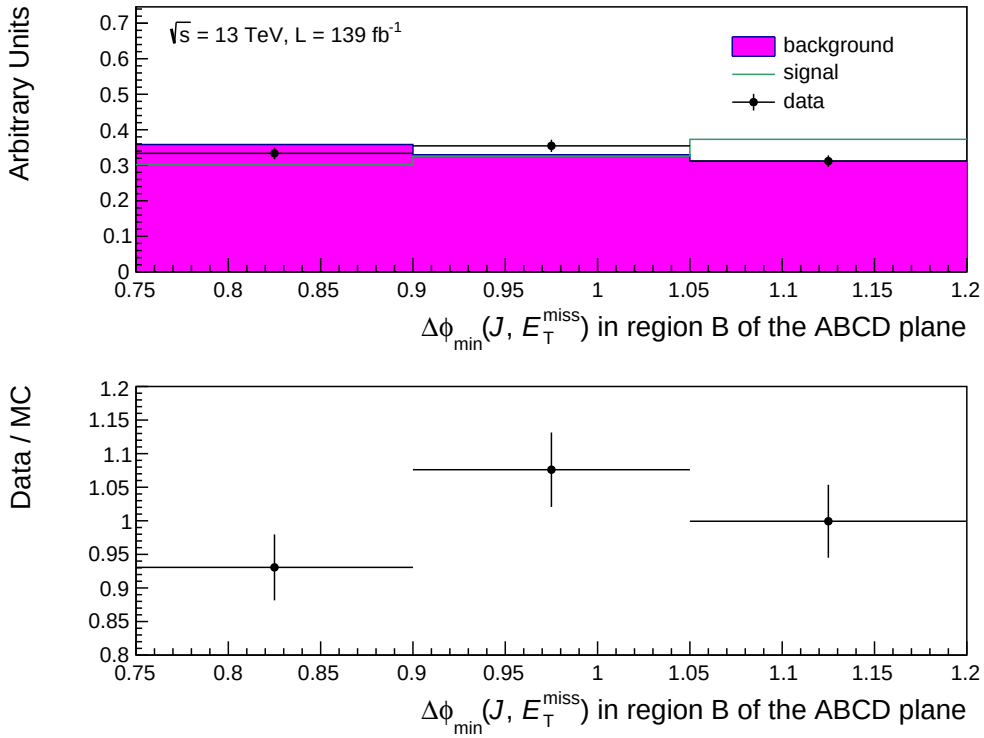


Figure 10.22 *Top plot: Normalized distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ in Control Region B with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This ratio is created from the normalized distribution shown in the top plot. This is for the NCBH Analysis of NCBH benchmark NCBH1.*

Similarly, the normalized distributions of background and data are within 2σ uncertainties in control region C in Figure 10.23.

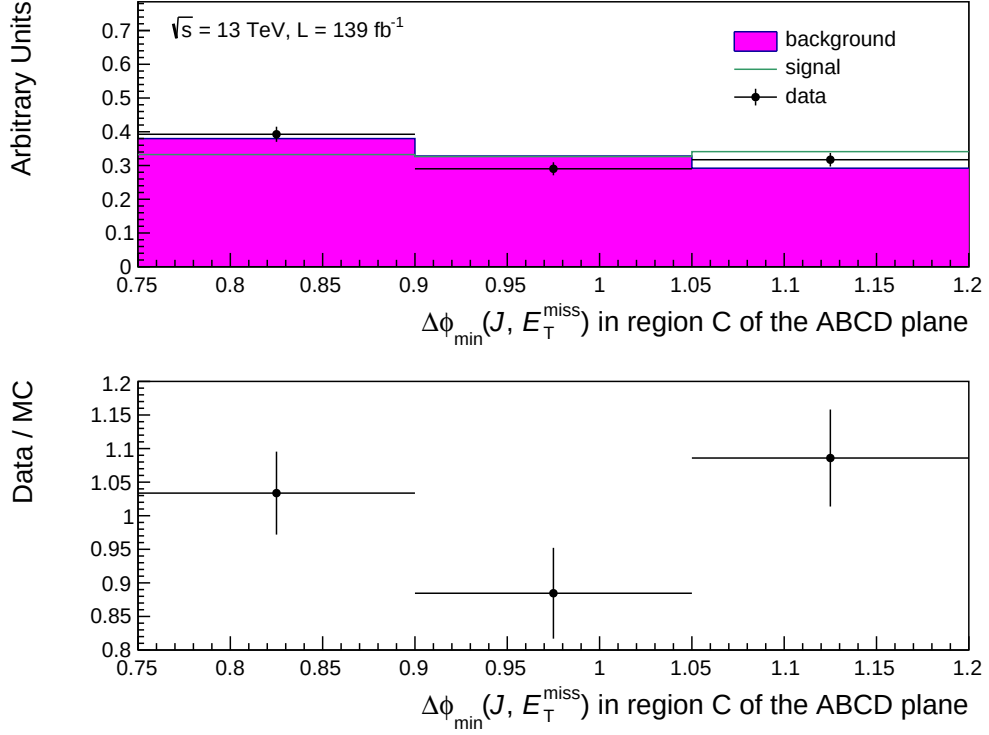


Figure 10.23 *Top plot: Normalized distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ in Control Region C with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This ratio is created from the normalized distribution shown in the top plot. This is for the NCBH Analysis of NCBH benchmark NCBH1.*

The normalized distributions of background and data are within 2σ uncertainties for all but two bins in control region D in Figure 10.24

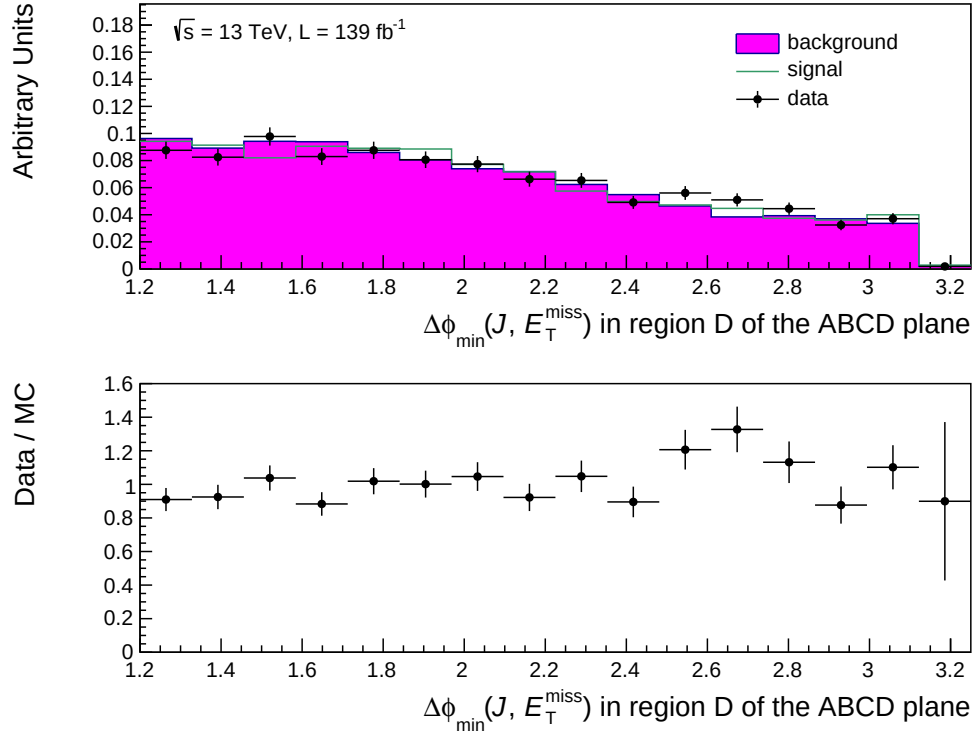


Figure 10.24 *Top plot: Normalized distributions of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ in Control Region D with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of $\Delta\phi_{\min}(J, E_T^{\text{miss}})$. This ratio is created from the normalized distribution shown in the top plot. This is for the NCBH Analysis of NCBH benchmark NCBH1.*

The validation regions for background are shown in Figure 10.25.

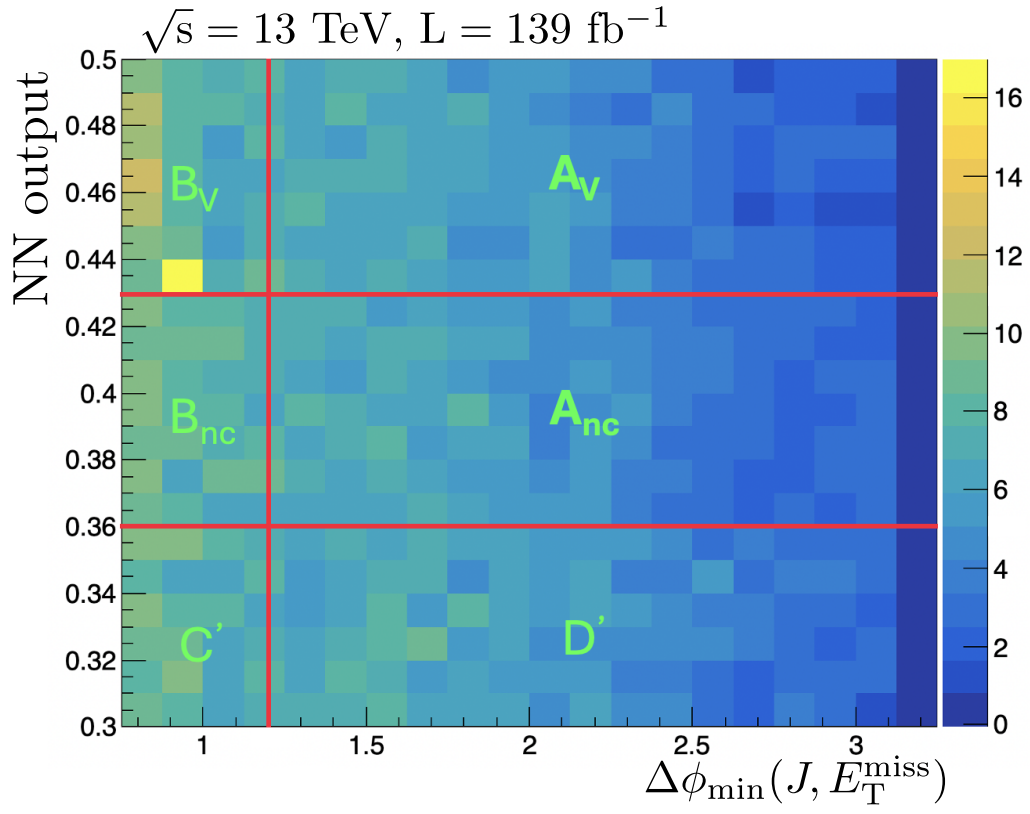


Figure 10.25 *Nonclosure and Validation Regions of ABCD plane for MC background for NCBH analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2*

The validation regions for data are shown in Figure 10.26.

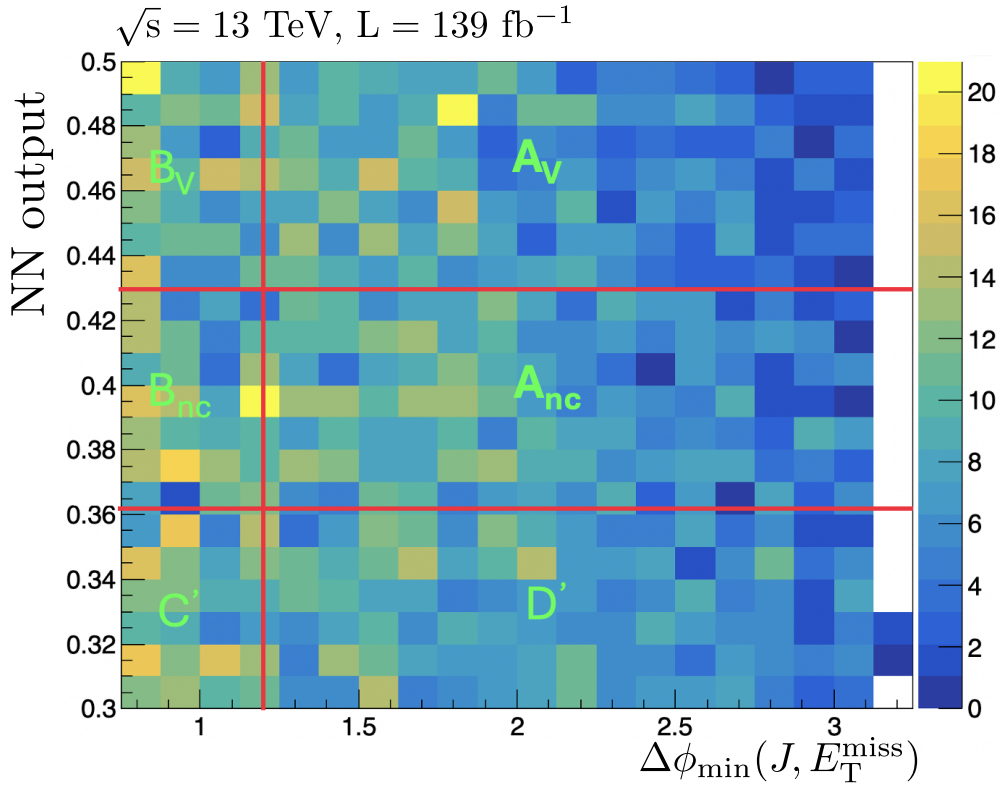


Figure 10.26 *Nonclosure and Validation Regions of ABCD plane for data for NCBH analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2*

The region of the ABCD plane that is used to determine the nonclosure uncertainty spans from a NN output of 0.3 to 0.43, where the division into different regions on the y axis occurs at 0.36, and a range of 0.75 to 3.25 in $\Delta\phi_{\min}(J, E_T^{\text{miss}})$, where the x axis is divided at 1.2. Table 10.17 summarizes the MC and data yields in the various regions of this plane. Both data and MC have small Pearson correlation coefficients indicating a high level of decorrelation. The MC nonclosure uncertainty is 7.51%, while the data nonclosure uncertainty is 4.12%.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.36 < NN < 0.43$	600.2	780	648.75	749.11
B: $0.75 < \phi < 1.2, 0.36 < NN < 0.43$	234.9	272		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.36$	196.9	244		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.36$	543.8	672		
MC NC Uncertainty	7.51%			
Data NC Uncertainty	4.12%			
MC Corr. Coeff	-0.0023			
Data Corr. Coeff	-0.0029			

Table 10.17 *Monte Carlo Yields, Data Yields, Monte Carlo and Data Nonclosure Uncertainties (MC NC Uncertainty, Data NC Uncertainty), Pearson Correlation Coefficients for MC (MC Corr. Coeff) and Data (Data Corr. Coeff), expected MC and Data event yields (Exp MC events and Exp Data yields) for regions of an ABCD plane used to calculate the nonclosure uncertainty for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.*

The region of the ABCD plane that is used to apply the nonclosure uncertainty and test for closure spans from a NN output of 0.3 to 0.5, where the division into different regions on the y axis occurs at 0.43. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is from 0.75 to 1.2 to 3.25. The table below summarizes the MC and data yields in the various regions of this plane. Both data and MC have small Pearson correlation coefficients indicating a relatively high level of decorrelation. The MC nonclosure uncertainty is 10.38%, while the data nonclosure uncertainty is 5.68%. The nonclosure of 4.12% from the previous plane provides closure to this validation plane as $|N_{A_V, \text{observed}} - N_{A_V, \text{expected}}| < \sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2 + \sigma_{\text{nc}}(N_{A_V}^{\text{exp}})^2}$.

NN range 0.3 to 0.43 to 0.5				
Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.43 < NN < 0.5$	586	706	653.9	748.5
B: $0.75 < \phi < 1.2, 0.43 < NN < 0.5$	246.8	266		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.43$	431.8	516		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.43$	1144	1452		
MC NC Uncertainty	-10.38%			
Data NC Uncertainty	-5.68%			
MC Corr. Coeff	-0.038			
Data Corr. Coeff	-0.059			

Table 10.18 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions of an ABCD plane used to validate the nonclosure uncertainty for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.*

Tables 10.19, 10.20, 10.21, and 10.22 show various alternative validation regions. For simplicity, the regions in the alternative validation regions are labeled in the tables as A, B, C, and D. Each of these regions represent some set of phase space with NN outputs below the signal region NN output of 0.5. Different combinations of cuts along the y axis, all under NN output values of 0.5, are made so as to study the Pearson correlation coefficients as y varies. These various regions all have Pearson correlation coefficients below 6%, showing an acceptable level of decorrelation. The highest nonclosure in data is 9.35%. The value of 3.58% derived from the subvalidation plane in Table 10.17 provides closure to all of these alternative validation regions, including the plane with the highest nonclosure uncertainty in data of 9.35%. The MC nonclosure in the signal region of the actual ABCD plane used for the NCBH analysis (NN output 0.3 to 0.5 to 1 and minimum ϕ 0.75 to 1.2 to 3.25) is 16.19%. As this is not drastically different from the data nonclosure uncertainty of 4.12% derived in the subvalidation region in the way seen for Scalar Cascades (164% vs. 25.8%), we can view the MC as a cross-check and safely use the fully data-driven approach, taking 4.12% as the nonclosure uncertainty.

The nominal expected value of data from the plane is 3276. The nonclosure uncertainty is 135 events, the total systematic error on the background is 167 events, and the total statistical error is 57 events. The total uncertainty of the background, when summing in quadrature the nonclosure uncertainty, total systematic error, and total statistical error, is 222 events.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.36 < NN < 0.5$	1186	1486	1330.362	1481.7049
B: $0.75 < \phi < 1.2, 0.36 < NN < 0.5$	481.7	538		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.36$	196.9	244		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.36$	543.8	672		
MC NC Uncertainty	-10.85%			
Data NC Uncertainty	0.03%			
MC Corr. Coeff	-0.038			
Data Corr. Coeff	-0.059			

Table 10.19 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ' ϕ ' and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.47 < NN < 0.5$	247.6	297	263.98	317.60
B: $0.75 < \phi < 1.2, 0.47 < NN < 0.5$	102.6	114		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.47$	576	668		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.47$	1482	1861		
MC NC Uncertainty	-6.21%			
Data NC Uncertainty	-6.48%			
MC Corr. Coeff	-0.038			
Data Corr. Coeff	-0.059			

Table 10.20 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.36 < NN < 0.47$	938.6	1189	1047.0	1167.74
B: $0.75 < \phi < 1.2, 0.36 < NN < 0.47$	379.1	424		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.36$	196.9	244		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.36$	543.8	672		
MC NC Uncertainty	-10.35%			
Data NC Uncertainty	1.82%			
MC Corr. Coeff	-0.037			
Data Corr. Coeff	-0.048			

Table 10.21 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ‘ ϕ ’ and the NN output is denoted as NN.*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $1.2 < \phi < 3.25, 0.33 < NN < 0.5$	1463	1833	1592.31	1676.18
B: $0.75 < \phi < 1.2, 0.33 < NN < 0.5$	581.2	655		
C: $0.75 < \phi < 1.2, 0.3 < NN < 0.33$	97.31	127		
D: $1.2 < \phi < 3.25, 0.3 < NN < 0.33$	266.6	325		
MC NC Uncertainty	-8.12%			
Data NC Uncertainty	9.35%			
MC Corr. Coeff	-0.038			
Data Corr. Coeff	-0.059			

Table 10.22 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NCBH Analysis of NCBH benchmark NCBH1. The $\Delta\phi_{\min}(J, E_T^{\text{miss}})$ is denoted as ' ϕ ' and the NN output is denoted as NN.*

10.4 NMDS ABCD Plane

The ABCD plane for the non-minimal dark sector benchmark NMDS3 is defined by plotting the transverse sphericity on the x axis and the NN output on the y axis. Figure 10.27 shows the MC background plane spanning the full NN output of 0 to 1 and the full range of transverse sphericity. The MC background is seen to be predominantly below NN output values of 0.4 and transverse sphericity values of 0.15. Creating a smaller, optimized ABCD plane spanning NN output values of 0.4 to 1.0 and transverse sphericity values of 0.15 to 1 greatly reduces the background, while also improving the decorrelation.

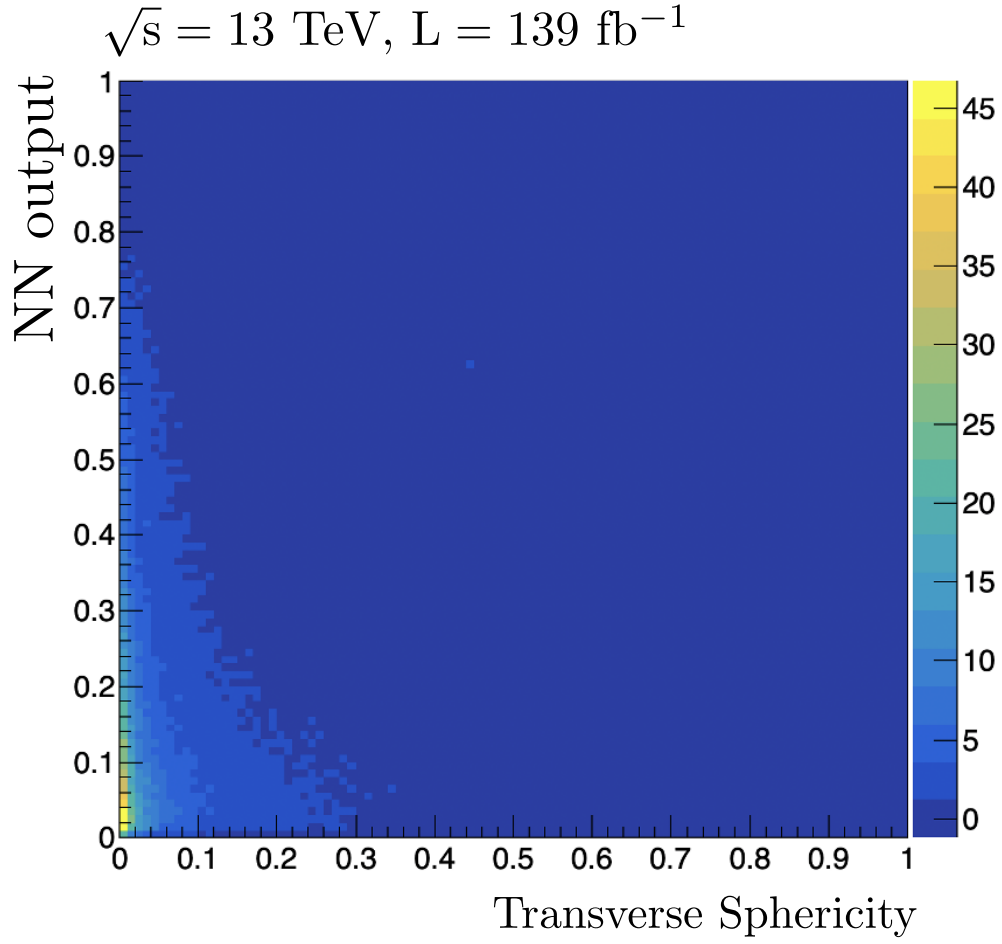


Figure 10.27 *NN output vs. transverse sphericity for MC background for full NN output range and full transverse sphericity range for NMDS analysis of benchmark NMDS3.*

Figure 10.28 shows the NN output vs. transverse sphericity distribution for the non-minimal dark sector NMDS3 benchmark. The signal most densely lies in high NN values and transverse sphericity values above 0.2. However, signal does spread throughout the plane, so care must be to ensure that regions are defined such that the signal region contains the majority of the signal.

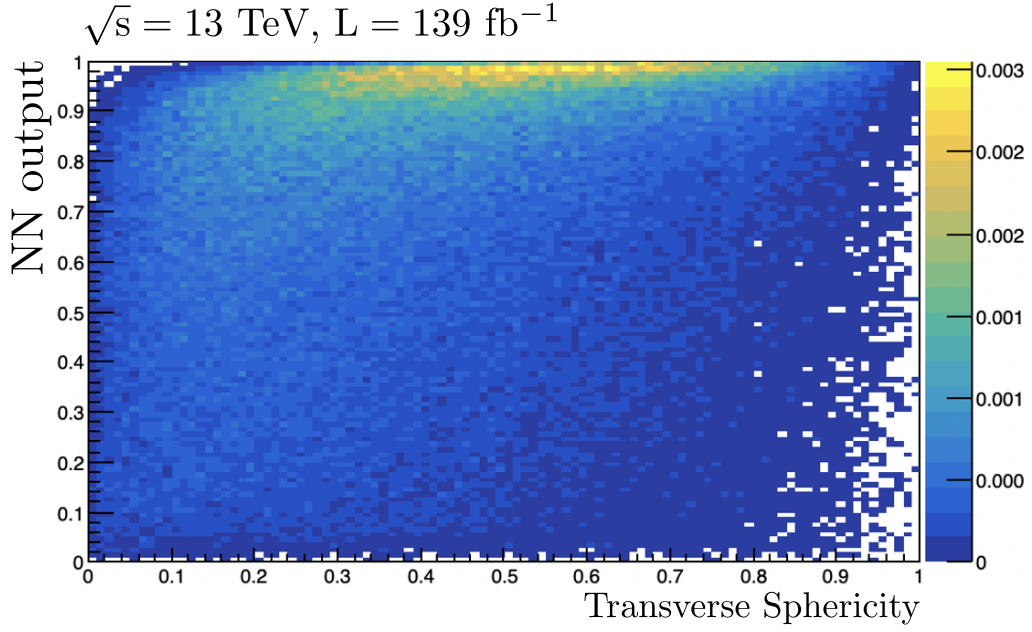


Figure 10.28 *NN output vs. transverse sphericity for MC signal for full NN output range and full transverse sphericity range for NMDS analysis of benchmark NMDS3.*

The regions of the ABCD plane for the non-minimal dark sector NMDS3 benchmark are resultingly defined below.

Region	MC yields	Data Yields	Exp MC events	Exp Data events	Signal Events
A	138.1	Blinded	91.34	116.09	1.83
B	24.94	31			0.12
C	52.56	47			0.077
D	192.5	176			0.59
MC NC Uncertainty	51.19%				
Data NC Uncertainty	N/A				
MC Corr. Coeff	0.12				
Data Corr. Coeff	N/A				

Table 10.23 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of the ABCD plane for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN.*

The ABCD plane for MC background is shown in Figure 10.29

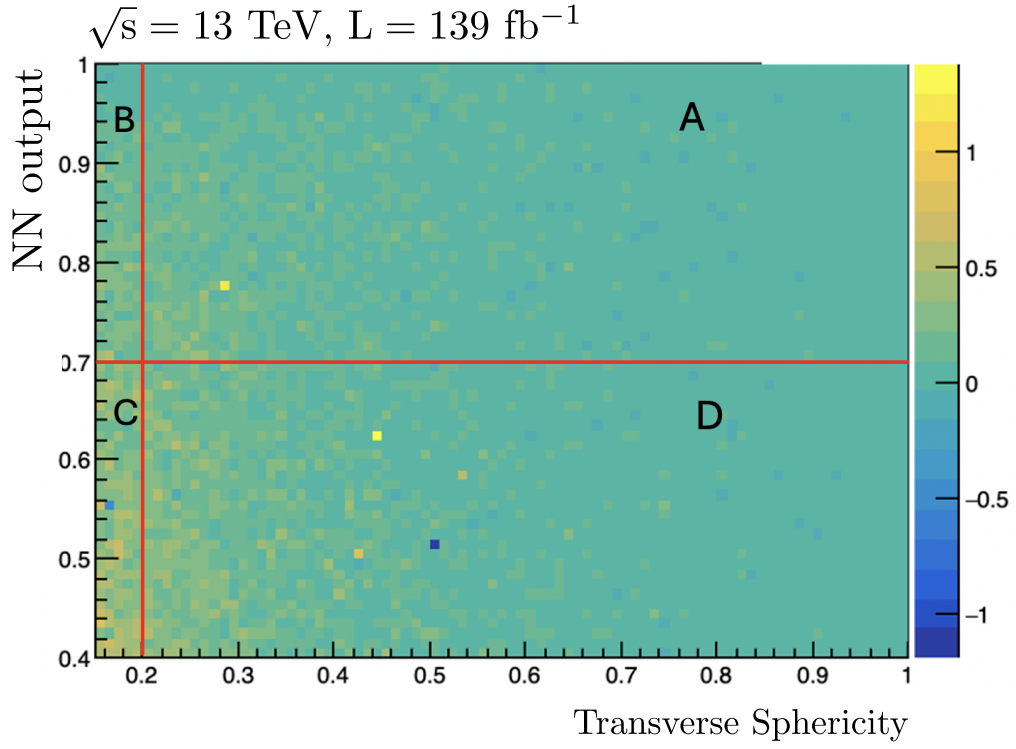


Figure 10.29 *MC background ABCD plane as a function of transverse sphericity and NN output for NMDS analysis of NMDS benchmark NMDS3.*

Table 10.23 shows that the Monte Carlo has a Pearson Correlation coefficient of 0.12, indicating that the plane has an acceptable level of decorrelation. Figure 10.29 echos this.

The ABCD plane for signal is shown in Figure 10.30

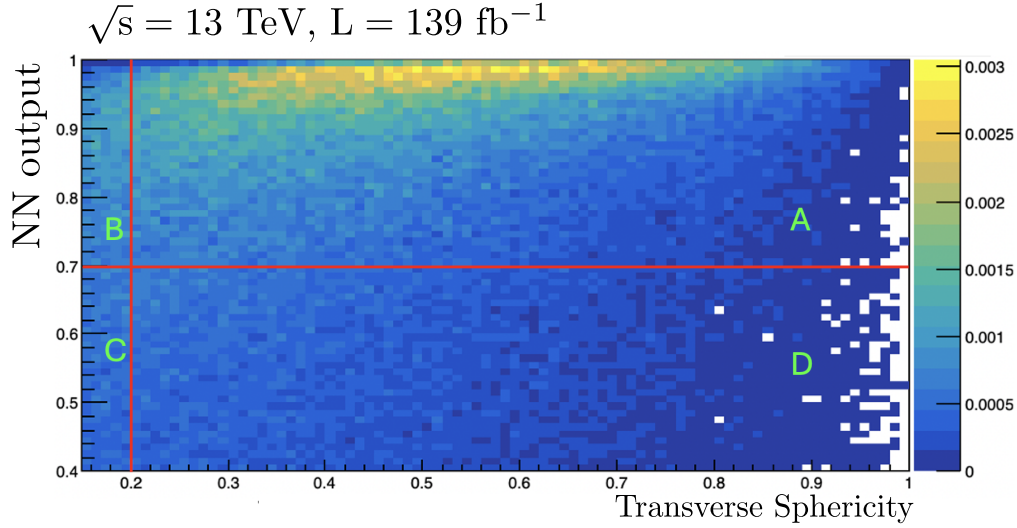


Figure 10.30 *MC signal ABCD plane as a function of transverse sphericity and NN output for NMDS analysis of NMDS benchmark NMDS3.*

The signal yields in each region of the ABCD plane are shown in Table 10.24.

Region	Signal yields	MC yields	% of Total Signal	$100 \times \frac{\text{Signal}}{\text{Background}}$
A: $0.2 < S < 1, 0.7 < NN < 1.0$	1.83	138.1	70%	N/A
B: $0.15 < S < 0.2, 0.7 < NN < 1.0$	0.12	24.94	4.7%	0.4%
C: $0.15 < S < 0.2, 0.4 < NN < 0.7$	0.077	52.56	2.9%	0.16%
D: $0.2 < S < 1, 0.4 < NN < 0.7$	0.59	192.5	22.4%	0.33%

Table 10.24 *Monte Carlo Yields and Signal Yields for the A, B, C, and D Regions of the ABCD plane for the NMDS Analysis of NMDS benchmark NMDS3. Column “% of Total Signal” gives the percent of total signal events in the ABCD plane that are contained in a given region. The column “ $100 \times \frac{\text{Signal}}{\text{Background}}$ ” turns into a percent the fraction of signal events in a given region with respect to the number of background events in the region. The transverse sphericity is abbreviated as ‘S’ and the NN output is denoted as NN.*

Table 10.24 shows that the majority of the signal is contained in the signal region. Additionally, the amount of signal compared to background in each region is small.

The significance (defined in Equation 8.3) after applying the cuts to form the signal region (defined as a transverse sphericity value greater than 0.2 and a NN output greater than 0.7) is 0.16. The significance is 3.12 times improved compared to the significance obtained upon applying a 500 GeV E_T^{miss} and 1400 GeV H_T cut and basic cuts. In total, the final significance is 26.2 times improved compared to the original

significance of the signal from applying a 200 GeV E_T^{miss} cut and basic cuts. Of the three searches, the non-minimal dark sector search has the second highest increase in significance, preceded by the scalar cascades and followed by the noncommutative black holes.

The behavior of data was checked against that of Monte Carlo in the control regions of the ABCD plane (regions B, C, and D). As seen in Figure 10.31, the normalized distributions of background and data are largely in agreement in the control region B. All but two bins have error bars that capture a MC to Data ratio of 1.

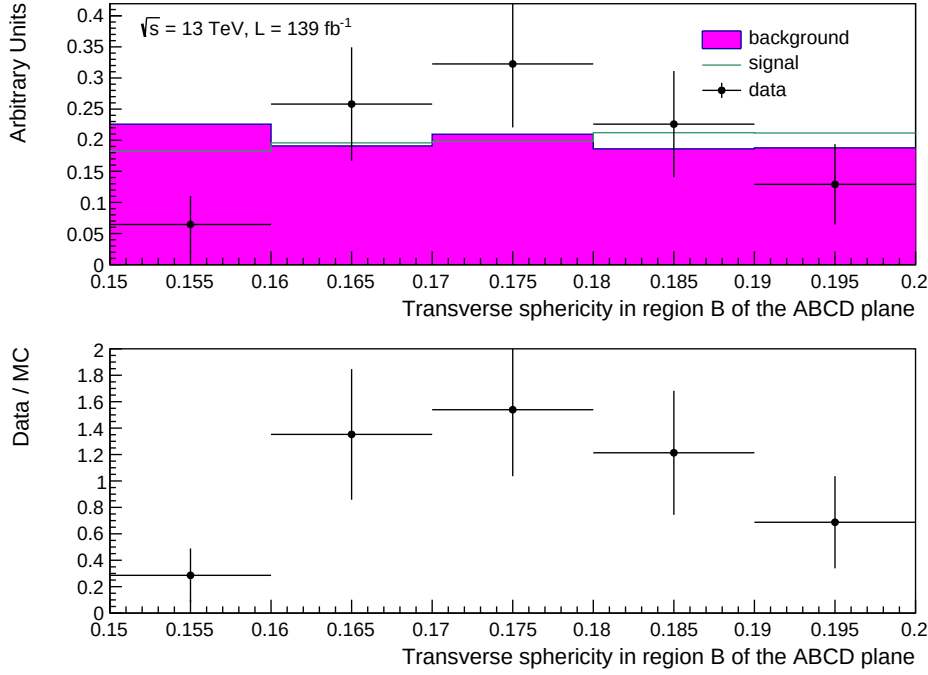


Figure 10.31 *Top plot: Normalized distributions of transverse sphericity in Control Region B of NMDS ABCD plane with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of transverse sphericity in control Region B. This ratio is created from the normalized distribution shown in the top plot. This is for the NMDS Analysis of NMDS benchmark NMDS3*

In Figure 10.32, the normalized distributions of background and data are largely in agreement in the control region C, with only 1 bin falling outside σ of a ratio of 1.

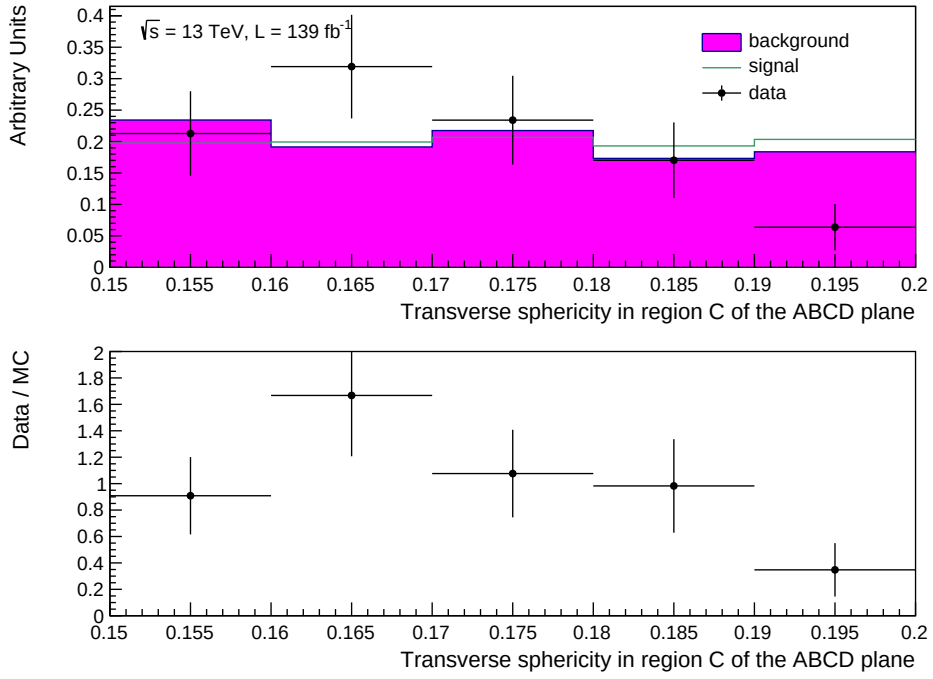


Figure 10.32 *Normalized distributions of transverse sphericity in Control Region C of NMDS ABCD plane with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of transverse sphericity in control Region C. This ratio is created from the normalized distribution shown in the top plot. This is for the NMDS Analysis of NMDS benchmark NMDS3*

In Figure 10.33, the normalized distributions of background and data in the control region D are largely in agreement due to large error bars, reflective of the small statistics.

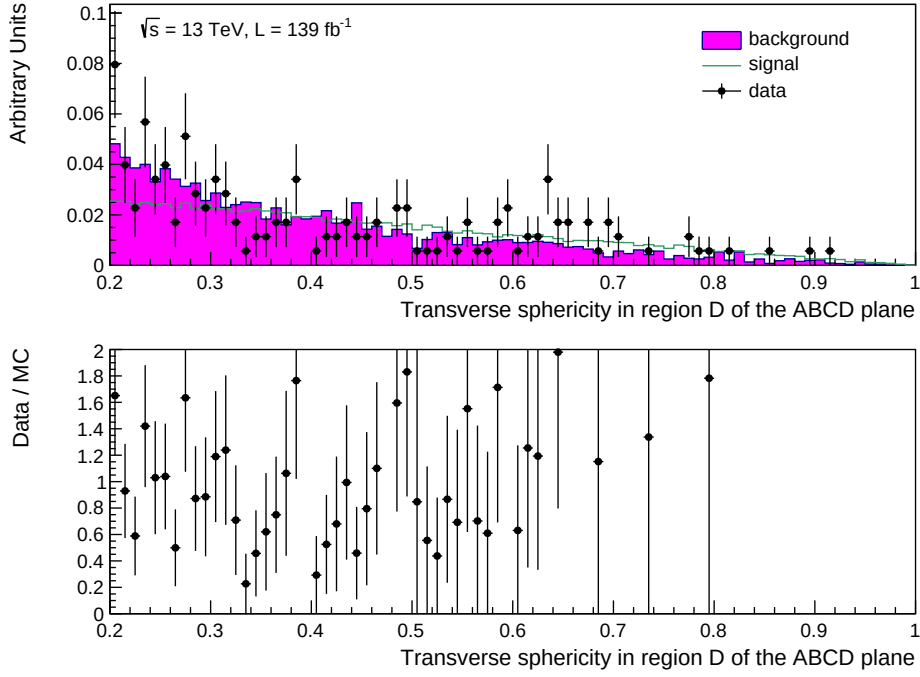


Figure 10.33 *Normalized distributions of transverse sphericity in Control Region D of NMDS ABCD plane with data, background MC, and signal overlaid. Bottom plot: Ratio of MC to data for the distribution of transverse sphericity in control Region D. This ratio is created from the normalized distribution shown in the top plot. This is for the NMDS Analysis of NMDS benchmark NMDS3*

Following the procedure outlined in the previous chapter, two smaller ABCD planes were formed within the ABCD plane specified above in order to determine a nonclosure uncertainty and in order to validate the ABCD plane.

These regions for MC background appear in Figure 10.34

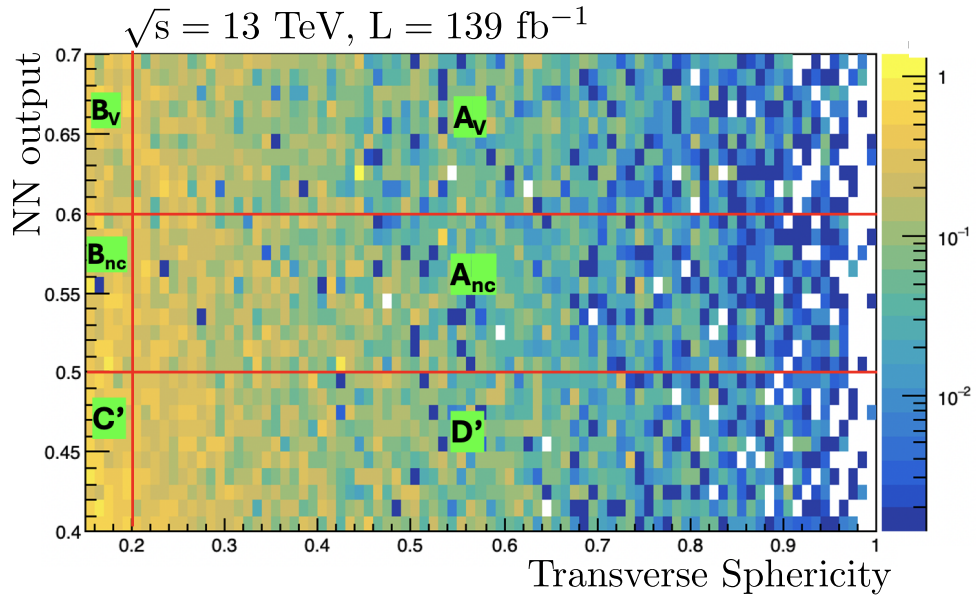


Figure 10.34 *Nonclosure and Validation Regions of ABCD plane for MC background for NMDS analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.*

These regions for data appear in Figure 10.35

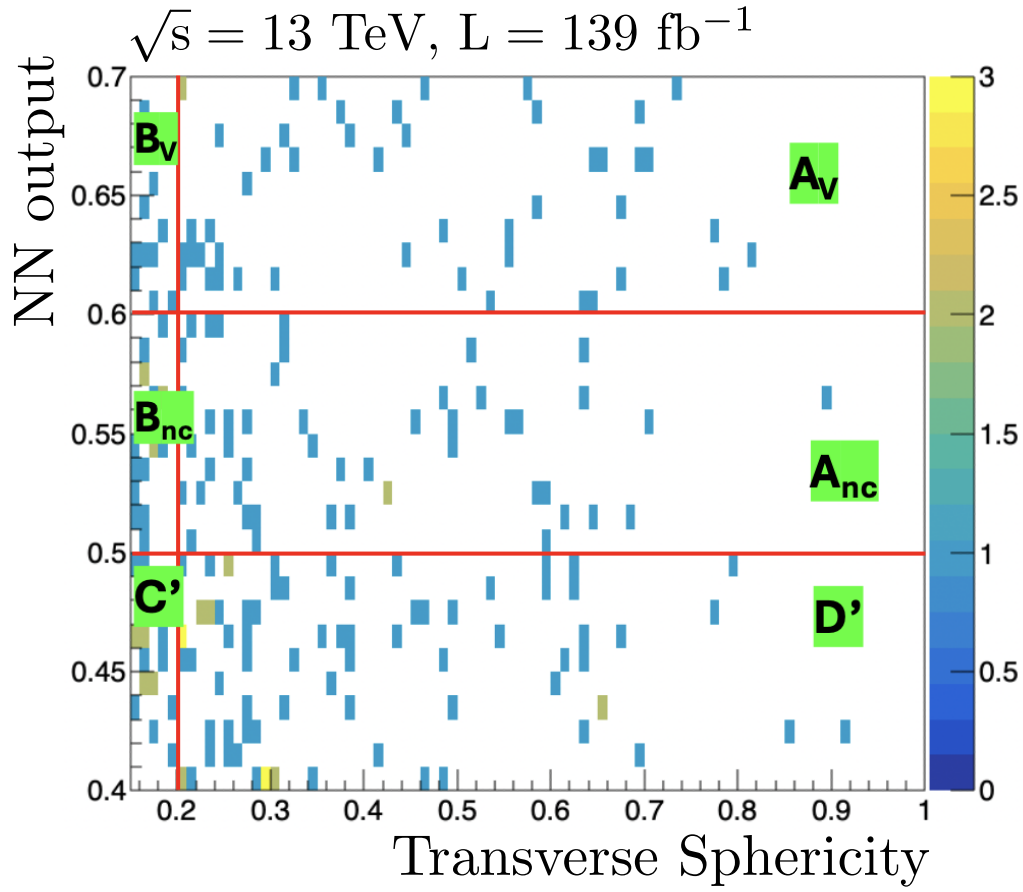


Figure 10.35 *Nonclosure and Validation Regions of ABCD plane for data for NMDS analysis. Regions A_V , A_{nc} , B_V , B_{nc} , C' , and D' are defined in Figure 10.2.*

The region of the ABCD plane that is used to determine the nonclosure uncertainty spans from a NN output of 0.4 to 0.6, where the division into different regions on the y axis occurs at 0.5, and a range of 0.15 to 1 in the transverse sphericity, where the x axis is divided at 0.2. The table below summarizes the MC and data yields in the various regions of this plane. Both data and MC have small Pearson correlation coefficients indicating an acceptable level of decorrelation. The MC nonclosure uncertainty is 6%, while the data nonclosure uncertainty is 37%. The MC Pearson Correlation Coefficient is 0.018, while the data Pearson correlation coefficient is 0.10. Both suggest an acceptable level of decorrelation.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A_{nc} : $0.2 < S < 1.0$, $0.5 < NN < 0.6$	61.89	48	58.20	75.79
B_{nc} : $0.15 < S < 0.2$, $0.5 < NN < 0.6$	16.82	18		
C' : $0.15 < S < 0.2$, $0.4 < NN < 0.5$	21.21	19		
D' : $0.2 < S < 1.0$, $0.4 < NN < 0.5$	73.39	80		
MC NC Uncertainty	6%			
Data NC Uncertainty	-37%			
MC Corr. Coeff	0.018			
Data Corr. Coeff	0.10			

Table 10.25 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions in the ABCD plane used calculate the nonclosure uncertainty. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN.*

The region of the ABCD plane that is used to apply the nonclosure uncertainty and test for closure spans from a NN output of 0.4 to 0.7, where the division into different regions on the y axis occurs at 0.6. The transverse sphericity ranges from 0.15 to 1, where the x axis occurs at 0.2. The table below summarizes the MC and data yields in the various regions of this plane. Both data and MC have small Pearson correlation coefficients indicating an acceptable level of decorrelation. The MC nonclosure uncertainty is 11%, while the data nonclosure uncertainty is 39%. The nonclosure of 37% from the previous plane provides closure to this validation plane within the uncertainty of the current plane, as $|N_{A_V, \text{observed}} - N_{A_V, \text{expected}}| < \sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2 + \sigma_{nc}(N_{A_V}^{\text{exp}})^2}$.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A_V : $0.2 < S < 1.0$, $0.6 < NN < 0.7$	57.26	48	51.71	34.59
B_V : $0.15 < S < 0.2$, $0.6 < NN < 0.7$	14.53	10		
$A_{nc} + D'$: $0.15 < S < 0.2$, $0.4 < NN < 0.6$	38.02	37		
$B_{nc} + C'$: $0.2 < S < 1.0$, $0.4 < NN < 0.6$	135.3	128		
MC NC Uncertainty	11%			
Data NC Uncertainty	39%			
MC Corr. Coeff	0.018			
Data Corr. Coeff	0.10			

Table 10.26 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for regions in the ABCD plane used to validate the nonclosure uncertainty. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN*

The tables below show various alternative validation regions. These various regions

all have Pearson correlation coefficients below 10%, showing an acceptable level of decorrelation. The highest nonclosure in the alternative validation regions is 52%. The MC nonclosure in the signal region of the ABCD plane that is actually used in the NMDS analysis (NN output 0.4 to 0.7 to 0.1 and transverse sphericity of 0.15 to 0.2 to 1.0) is 51.2%. The nonclosure of 37% provides closure to the data validation region containing a 52% nonclosure uncertainty. The nonclosure is, therefore taken to be 37%.

The nominal expected value of data from the plane is 116 events. The total systematic uncertainty is 28 events, the nonclosure uncertainty is 43 events, the total statistical uncertainty is 11 events. Adding the statistical uncertainty, systematic uncertainty, and nonclosure uncertainty in quadrature, the total uncertainty on the background is 53 events.

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $0.2 < S < 1.0, 0.5 < NN < 0.7$	119.2	96	108.48	117.89
B: $0.15 < S < 0.2, 0.5 < NN < 0.7$	31.35	28		
C: $0.15 < S < 0.2, 0.4 < NN < 0.5$	21.21	19		
D: $0.2 < S < 1.0, 0.4 < NN < 0.5$	73.39	80		
MC NC Uncertainty	9.9%			
Data NC Uncertainty	-19%			
MC Corr. Coeff	0.018			
Data Corr. Coeff	0.10			

Table 10.27 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $0.2 < S < 1.0, 0.55 < NN < 0.65$	59.72	48	48.19	57.93
B: $0.15 < S < 0.2, 0.55 < NN < 0.65$	13.94	16		
C: $0.15 < S < 0.2, 0.4 < NN < 0.55$	30.4	29		
D: $0.2 < S < 1.0, 0.4 < NN < 0.55$	105.1	105		
MC NC Uncertainty	24%			
Data NC Uncertainty	-17%			
MC Corr. Coeff	0.010			
Data Corr. Coeff	0.032			

Table 10.28 *Monte Carlo and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $0.2 < S < 1.0$, $0.53 < NN < 0.62$	53.63	42	48.26	71.70
B: $0.15 < S < 0.2$, $0.53 < NN < 0.62$	13.87	17		
C: $0.15 < S < 0.2$, $0.4 < NN < 0.53$	26.75	23		
D: $0.2 < S < 1.0$, $0.4 < NN < 0.53$	93.08	97		
MC NC Uncertainty	11%			
Data NC Uncertainty	-41%			
MC Corr. Coeff	0.00038			
Data Corr. Coeff	0.014			

Table 10.29 *MC and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN*

Region	MC yields	Data Yields	Exp MC events	Exp Data events
A: $0.2 < S < 1.0$, $0.62 < NN < 0.7$	45.83	37	43.12	24.33
B: $0.15 < S < 0.2$, $0.62 < NN < 0.7$	11.94	7		
C: $0.15 < S < 0.2$, $0.4 < NN < 0.62$	40.62	40		
D: $0.2 < S < 1.0$, $0.4 < NN < 0.62$	146.7	139		
MC NC Uncertainty	6.3%			
Data NC Uncertainty	52%			
MC Corr. Coeff	0.018			
Data Corr. Coeff	0.10			

Table 10.30 *MC and Data Observed Yields, Expected Yields, Nonclosure Uncertainties, and Pearson Correlation Coefficients for Regions A, B, C, and D of an alternative ABCD plane used for validation. This is for the NMDS Analysis of NMDS benchmark NMDS3. The transverse sphericity is denoted as ‘S’ and the NN output is denoted as NN*

10.5 Summary

This chapter described how a data-driven ABCD background estimation was used in the signal region for each analysis. Comparisons between Monte Carlo background and data in the control regions of the plane show that the neural networks for all analyses generalize well and behave as expected. Checks were performed to ensure that the two variables forming the ABCD plane were decorrelated. A nonclosure uncertainty was derived for each analysis to account for any remaining correlation. The scalar cascade analysis had the highest nonclosure uncertainty out of the three models, with a value of 184% (the relatively large uncertainty due to the small number of background events in the signal region), while the non-minimal dark sector analysis had a nonclosure uncertainty of 37% and the noncommutative black

hole analysis had a nonclosure uncertainty of 4.12%. The significance for scalar cascades, noncommutative black holes, and non-minimal dark sectors after defining a signal region were 0.39, 4, and 0.16, respectively. The significance of the scalar cascade analysis was 73.4 times improved compared to prior to the NN, while the significance of the noncommutative black hole analysis was 1.68 times improved, and the significance of the non-minimal dark sector analysis was 3.12 times improved. This is as expected, due to the scalar cascades having the best neural network performance out of the three models, as well as having a second decorrelated variable with the most discriminatory power out of any of the models. The significance of the non-minimal dark sector model is low due to its small cross section while also not having as distinctive characteristics as scalar cascades. The noncommutative black hole model benefited from having a higher cross section than the other models, which allowed for a higher significance. The data counts and uncertainties derived via the ABCD method serve as inputs to the statistical analysis performed to estimate the search sensitivity on the exotic models and that will be described in the next chapter.

Chapter 11

Statistical Analysis

The last chapter defined a data-driven background estimate technique that created a 2D plane consisting of four quadrants: three control regions and one signal region. The data count in the signal region was derived from the data count in the control regions. This chapter takes as input the number of data and signal events in each region, together with their associated uncertainties, and derives the ABCD background prediction for the signal region using a likelihood fit. Potential signal contamination is taken into account in the estimate by expressing each number of events as the sum of a background contribution and a signal contribution, modified with a signal strength parameter.

The likelihood fit is used to determine the probability to obtain n events given the parameters provided, including the systematic uncertainties. A description is given of the Frequentist Approach, deriving a test statistic to test compatibility of data with a background-only or a signal-plus-background hypothesis, and arriving at the p -value to quantify the probability of such hypotheses. Exclusion limits can finally be derived in function of model parameters, indicating at what signal strength a given hypothesis can be ruled out.

11.1 Signal Contamination

Up until now, no consideration has been given to signal contamination in the calculation of the number of expected background events in a given region. If signal were to appear outside of the signal region, region A, the number of signal events in a

given region could be subtracted from the number of events observed in a given region to yield the number of background events in that region. However, this assumes that the number of signal events in each region is known precisely. Given that signal comes with theoretical and experimental uncertainties, it is only possible to state that, in a given region of the ABCD plane, the actual number of signal events is given by a range spanning the expected number of signal events plus or minus the total uncertainty on the signal in that region - unless, however, a statistical fit is performed in which the uncertainties and the signal strength are treated as parameters to be fit to the data. If a statistical fit is performed, the value of each parameter in the fit can be determined by seeing which value allows for the best description of the data. When dealing with Monte Carlo, the signal strength of the signal is normalized to 1. Monte Carlo simulation, when weighting each event by the cross section of the signal process and the relevant integrated luminosity, provides the nominal number of signal events expected. A signal strength of 1 in collected data corresponds to the signal plus background hypothesis, confirming the expectation of the signal at the expected rate. If the signal process is absent in the collected data, the signal strength is zero. A signal strength of zero corresponds to the null hypothesis, or scenario in which only background events are expected in the signal region. The signal strength can be a continuous variable. A negative signal strength indicates a deficit of events, ie. below the average background events expected.

The total number of events observed in each region of the ABCD plane can be described by the following relation:

$$\Gamma[n_i] = \mu s_i(\mathbf{\Lambda}) + b_i(\mathbf{\Lambda}, \mathbf{\Omega}). \quad (11.1)$$

The contribution from signal to the total number of events (either observed or expected) in a given region is expressed as the signal strength μ multiplied by the number of signal events in the given region. The signal strength, μ , is a parameter that can be fit to the data via a statistical fit. The contribution from background to the total number of events in a given region is expressed as $b_i(\mathbf{\Lambda}, \mathbf{\Omega})$, where b_i is the number of background events and the full expression shows that b_i depends on uncertainties $\mathbf{\Lambda}$ and normalization parameters $\mathbf{\Omega}$. The contribution from background to the total number of events in a given region can be described in terms of normalized parameters, instead of counts, that can be fit to data with a statistical fit.

The number of background events in region B can be defined as the parameterized value μ_B such that

$$b_B = \mu_B. \quad (11.2)$$

The number of background events in region C can then be described in terms of μ_B and the relative ratio, ϵ_{BC} , of the number of background events in B to the number of background events in region C. Like μ_B , ϵ_{BC} is then a parameter than can be fit to data via the statistical fit and can be defined as

$$b_C = \mu_B \epsilon_{BC}. \quad (11.3)$$

Similarly, the number of background events in region D can then be described in terms of μ_B and the relative ratio, ϵ_{BD} , of the number of background events in B to the number of background events in region D. The normalized parameter ϵ_{BD} can then be fit to data via the statistical fit

$$b_D = \mu_B \epsilon_{BD}. \quad (11.4)$$

The number of background events in region A can be described in terms of μ_B , ϵ_{BC} , and ϵ_{BD} , such that

$$b_A = \mu_B \epsilon_{BC} \epsilon_{BD}. \quad (11.5)$$

These normalized background contributions can be added back into equation, such that expected rate of events in region A is

$$\Gamma[n_A] = \mu s_A(\mathbf{\Lambda}) + b_A, \quad (11.6)$$

while the expected rate in region B is

$$\Gamma[n_B] = \mu s_B(\mathbf{\Lambda}) + b_B. \quad (11.7)$$

The expect rate in region C is

$$\Gamma[n_C] = \mu s_C(\mathbf{\Lambda}) + b_C, \quad (11.8)$$

while the expected rate in region D is

$$\Gamma[n_D] = \mu s_D(\mathbf{\Lambda}) + b_D. \quad (11.9)$$

11.2 Likelihood Fit

The statistical fit used in this analysis to fit free parameters to data is the likelihood fit, depicted in Equation 11.10 [116] as

$$\mathcal{L}(n \mid \mathbf{\Lambda}, \mathbf{\Omega}) = \prod_{i=A,B,C,D} \text{Pois}(n_i \mid \mu s_i(\mathbf{\Lambda}) + b_i(\mathbf{\Lambda}, \mathbf{\Omega})) \prod_{\Lambda \in \mathbf{\Lambda}} f_{\Lambda}(\Lambda_{\Lambda} \mid \Lambda). \quad (11.10)$$

The likelihood function provides the probability to obtain n observed events given the parameters provided. The goal is to maximize the likelihood fit.

The likelihood function can be broken down into two parts. The first term, $\prod_{i=A,B,C,D} \text{Pois}(n_i \mid \mu s_i(\mathbf{\Lambda}) + b_i(\mathbf{\Lambda}, \mathbf{\Omega}))$, is the product of poisson distributions. In this analysis, $\mathbf{\Omega}$ are the normalized parameters from the ABCD plane derived earlier. The vector $\mathbf{\Lambda}$ is the set of nuisance parameters. There are many different nuisance parameters in this analysis. Experimental uncertainties on a number of different reconstructed objects, such as jets, E_T^{miss} , tracks, muons, and electrons, are accounted for in this analysis as nuisance parameters. These nuisance parameters represent systematic uncertainties. The distributions of events in each region of the ABCD plane are assumed in this equation to be described by a Poisson distribution as the events are independent and random. The Poisson distribution for each region of the ABCD plane takes the following form

$$P(n_i \mid \lambda) = \frac{e^{-\lambda} \lambda^{n_i}}{n_i!}, \quad (11.11)$$

where $i = A, B, C$, or D represents the region of the ABCD plane, λ is the sum of the signal and background events predicted in that particular region of the ABCD

plane and $P(n_i | \lambda)$ is the probability of observing n_i events in data in that particular region of the ABCD plane.

The second term, $\prod_{\Lambda \in \Lambda} f_{\Lambda}(\Lambda_{\Lambda} | \Lambda)$, is the product, over all nuisance parameters, of the constraint functions f , which restrict the values the nuisance parameters can take. Λ_{Λ} is the value of the nuisance parameter before the fit.

There are several options for which form the constraint functions can take. In this analysis, the constraint functions are described by Gaussian distributions:

$$f(\Lambda) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\Lambda_{\Lambda} - \Lambda)^2}{2\sigma^2}}. \quad (11.12)$$

The width of the Gaussian is σ .

In this analysis, the Poisson distributions in Equation 11.10 describing the number of expected events in each region of the ABCD plane are approximated as Gaussian distributions. In regions B, C, and D, the width of Gaussian is the statistical error on the number of expected events in the given region. In the signal region, region A, the width of the Gaussian is the sum in quadrature of the nonclosure uncertainty and the statistical error on the number of expected events in the signal region. The motivation for approximating the Poisson distributions describing the number of events in each region of the ABCD plane by Gaussian distributions was to allow for a less constrained fit. Fitting the Gaussian with the sum in quadrature of the statistical error and nonclosure uncertainty allowed for a well behaved fit. The use of this approximation is justified because the values of the expected number of events in Regions A, B, C, and D are large for noncommutative black holes and non-minimal dark sectors.

Similarly, the use of this approximation is justified for scalar cascades because the number of expected number of events (6.6) in Region A, which is derived via $\frac{N_B N_D}{N_C}$, can be seen in Figure 11.1 to have a Gaussian nature.

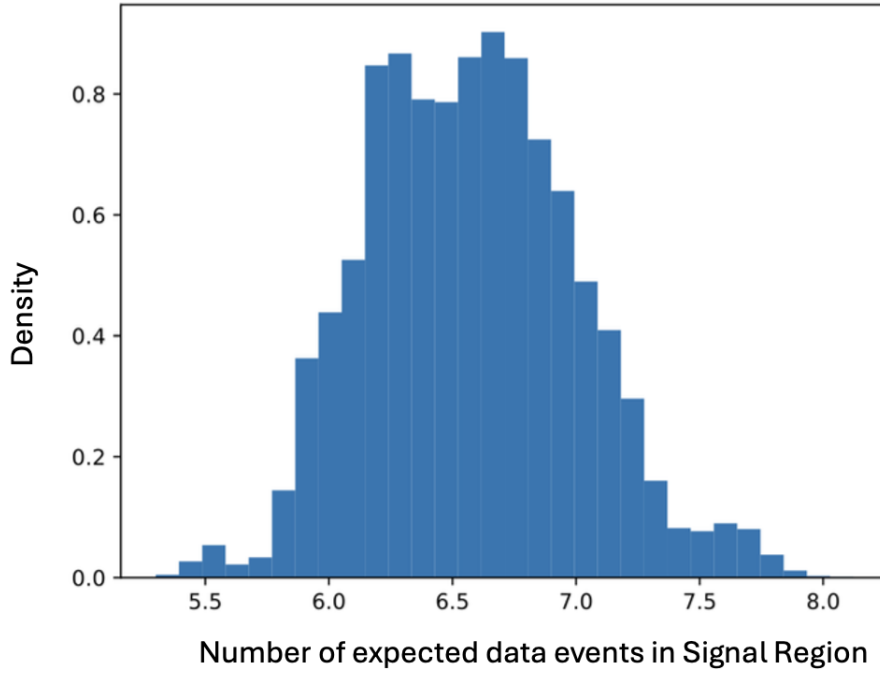


Figure 11.1 *Distribution displaying the Gaussian nature the expected (6.6) data events in the signal region of the Scalar Cascades ABCD plane. This serves as sufficient evidence for the number of events in the signal region to be approximated as a Gaussian instead of a Poisson distribution. This plot was created by drawing random values for the number of events in regions B, C, and D assuming Gaussian distributions, calculating the resulting value of the number of expected events in region A, and producing the probability density distribution of the number of expected events in region A*

11.3 Frequentist Approach

The result of the likelihood fit can be evaluated with the frequentist approach [117]. A quantity called a p -value is used to quantify the probability of observing a certain outcome under a given hypothesis, H . Two hypotheses are tested in this analysis: the background only hypothesis, H_b , in which the assumption is that data is consistent with the Standard Model, and the signal plus background hypothesis, H_{s+b} , which assumes the presence of a signal. The p -value is defined as

$$p = \int_{q^{obs}}^{\infty} f(q | H) dq, \quad (11.13)$$

where q is a test statistic that is a single-valued function of the data observed. The

q value is obtained by use of the negative log-likelihood ratio,

$$\lambda(\mu) = -2\ln \frac{L(\mu, \hat{\mathbf{\Lambda}}(\mu))}{L(\hat{\mu}, \hat{\mathbf{\Lambda}})}. \quad (11.14)$$

The numerator is the conditional likelihood which is maximized for a given signal strength μ . The denominator is the unconditional likelihood which is the maximum fitted value for any μ and $\mathbf{\Lambda}$. The ratio of the two likelihoods is bounded between 0 and 1, as the conditional likelihood cannot be larger than the unconditional likelihood. A value of 1 indicates good agreement between the data and the hypothesis, whereas lower values (and thus higher values of $\lambda(\mu)$) indicate worse agreement.

When using the background only hypothesis, the test statistic, q_θ , is defined such that the background only hypothesis is not rejected due to fluctuations from values of $\hat{\mu}$, described earlier as a deficit of events with respect to the background prediction:

$$q_\theta = \begin{cases} \lambda(0), & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0 \end{cases}. \quad (11.15)$$

Similarly, when using the signal plus background hypothesis, the test statistic, q_μ , is defined such that the background plus signal hypothesis is not rejected due to a surplus with respect to the background plus signal hypothesis.

$$q_\mu = \begin{cases} \lambda(\mu), & \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu \end{cases}. \quad (11.16)$$

Random pseudodata are used to determine the probability density functions of the test statistic so that the p -value can be calculated. Once the p -value is calculated, a significance, Z , can be derived. The significance is calculated by taking the inverse cumulative distribution function, Φ of a Gaussian, such that

$$Z = \Phi^{-1}(1 - p). \quad (11.17)$$

Models can be excluded at the LHC if a significance of $Z = 1.64$ or greater is obtained. A Z value of 1.64 corresponds to a p -value of 0.05. A p -value of 0.05 means that a

model can be excluded with a 95% confidence level. Evidence of new physics can be claimed with a Z value of 3. The threshold for discovery of new physics is set higher with a Z value of 5.

Exclusion limits, as those seen in Figure 11.2, show the derived p-values as a function of the signal strength in the signal plus background model. Expected exclusion limit plots enables one to answer whether they have the ability to identify the presence of a signal over background - whether the expected signal strength is detectable within the uncertainty on the background prediction. The expected yields in the ABCD plane from signal Monte Carlo correspond to $\mu = 1$. The exclusion limit plot identifies at which signal strength, μ , should one be able to exclude the μ -scaled signal. The intersection of the $p=0.05$ horizontal with the observed exclusion limit gives the signal strength μ value above which the model is excluded at 95% C.L or better. It is not possible to exclude signals with a smaller μ value because the signal may just be hiding behind background fluctuations. The yellow and green bands represent the one and two σ uncertainties on those expected limits. The points at which the yellow and green lines intersect the red line provides the lower and upper limit on the range of μ values that one is sensitive to with a range of uncertainties.

Expected and observed exclusion limits were derived by the author for all analyses using a general framework [118] that implemented the frequentist approach discussed above. The inputs to this framework were the data and signal yields in each region of the ABCD plane, as well as systematic uncertainties.

The author derived systematic uncertainties using an analysis framework created by the author. The framework links to a package [96] that implements the methods and settings recommended by the ATLAS Combined Performance groups (the groups responsible for reconstruction, identification, and measurement of physics objects).

The result for scalar cascades when using a nonclosure uncertainty of 184% (derived in the previous chapter) is shown in Figure 11.2.

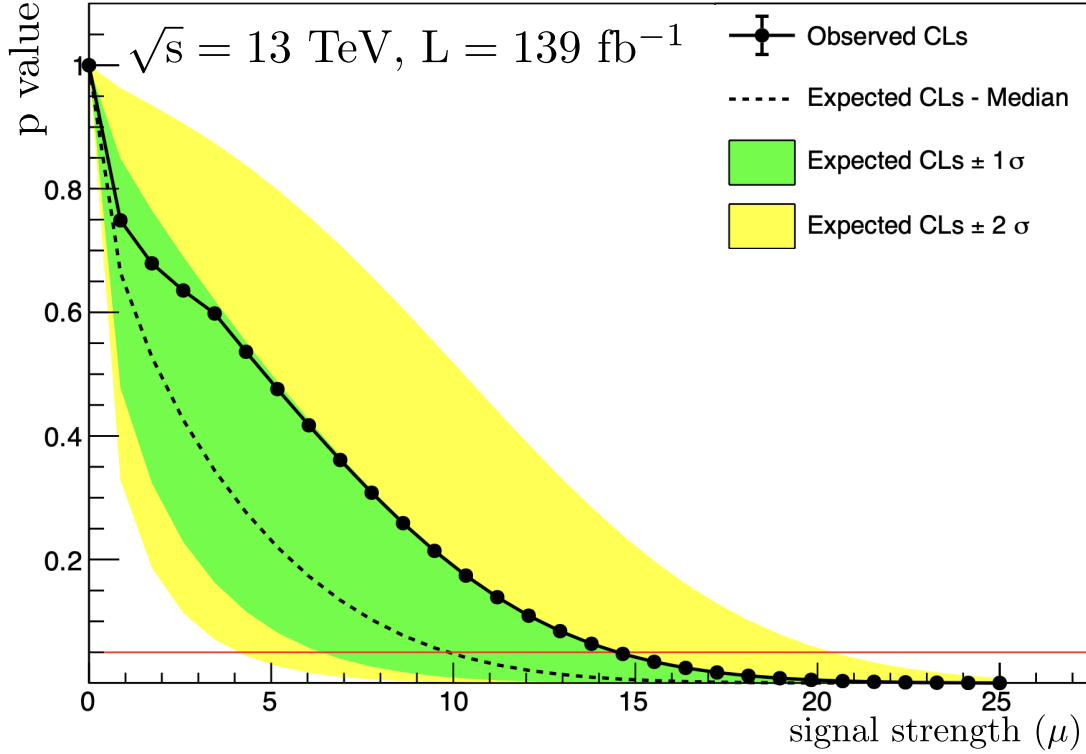


Figure 11.2 *Scalar Cascade Exclusion Limit. Result obtained via use of the Frequentist method.*

This plot was obtained by incorporating the full statistical uncertainty from the ABCD plane, a 0.26% experimental signal uncertainty, a 184% nonclosure uncertainty, a conservative 5% signal theory uncertainty, a 5% uncertainty to account for any inefficiency in the E_T^{miss} trigger due to the 200 GeV E_T^{miss} cut (as the E_T^{miss} trigger is known to start becoming 100% efficient somewhere in the range of 200 GeV to 250 GeV depending on the given analysis) [97], and a luminosity uncertainty of 1.7% [119]. The experimental signal uncertainty was derived by adding in quadrature the full set of experimental systematic uncertainties for signal. These signal systematic uncertainties include the E_T^{miss} systematic uncertainties, jet systematic uncertainties, and track systematic uncertainties (as well as electron and muon systematic uncertainties). The break down of all of these systematics (for all three searches) are included in the appendix. The jet systematic uncertainties are the leading systematic uncertainties, followed by the E_T^{miss} systematics, track systematics, and zero valued electron and muon systematic. (This breakdown is true for all three searches.)

With a nonclosure uncertainty of 184%, the expected exclusion limit for the scalar cascade benchmark is a μ value of 9.9. This means that the exclusion limit is set

at a cross section of 9.9 times the nominal scalar cascade cross section. Since the benchmark of the scalar cascade cross section is 5.33×10^{-4} pb, this means that a limit can be set on a cross section of 5.28×10^{-3} pb. The number of data events observed was 23 events, so the observed limit becomes a signal strength of 14.52, which translates to a cross section of 7.7×10^{-3} pb. The exclusion plot shows that the observed limit exhibits a roughly one sigma excess above the background hypothesis. This is not, however, a significant excess.

11.4 NCBH Benchmark NCBH1 Exclusion Limit

The total data events observed in the signal region is 3703 events. The expected exclusion limit for NCBH Benchmark NCBH1 when incorporating the full statistical uncertainty from the ABCD plane, a luminosity uncertainty of 1.7%, an experimental signal uncertainty of 0.65%, a conservative 5% signal theory uncertainty, and a nonclosure uncertainty of 4.12% is a μ value of 6.58. This means that the exclusion limit is set at a cross section of 6.58 times the nominal noncommutative black hole cross section. The cross section of the NCBH Benchmark NCBH1 is 0.022 pb, meaning that an expected exclusion limit can be placed on a cross section of 0.14 pb. The observed exclusion limit is 10.49, meaning that an observed exclusion limit can be placed on a cross section of 0.23 pb. These limits can be seen in Figure 11.3. The exclusion plot shows that the observed limit exhibits a roughly two sigma excess above the background hypothesis. This is not, however, a significant excess.

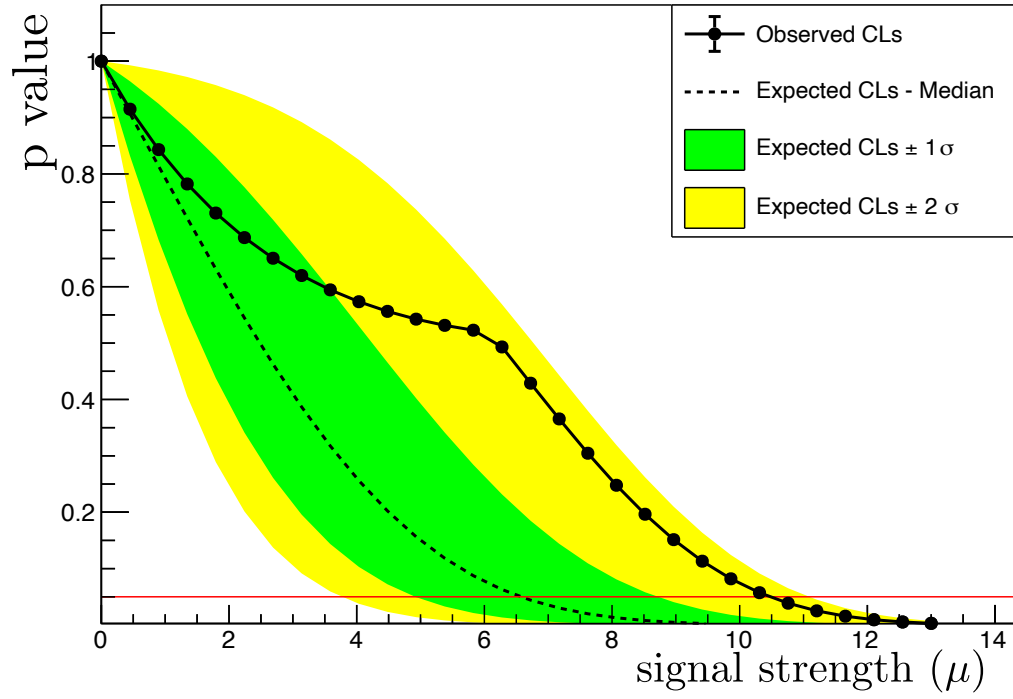


Figure 11.3 *NCBH Benchmark NCBH1 exclusion limit. Result obtained via use of the Frequentist method.*

11.5 NMDS Exclusion Limit

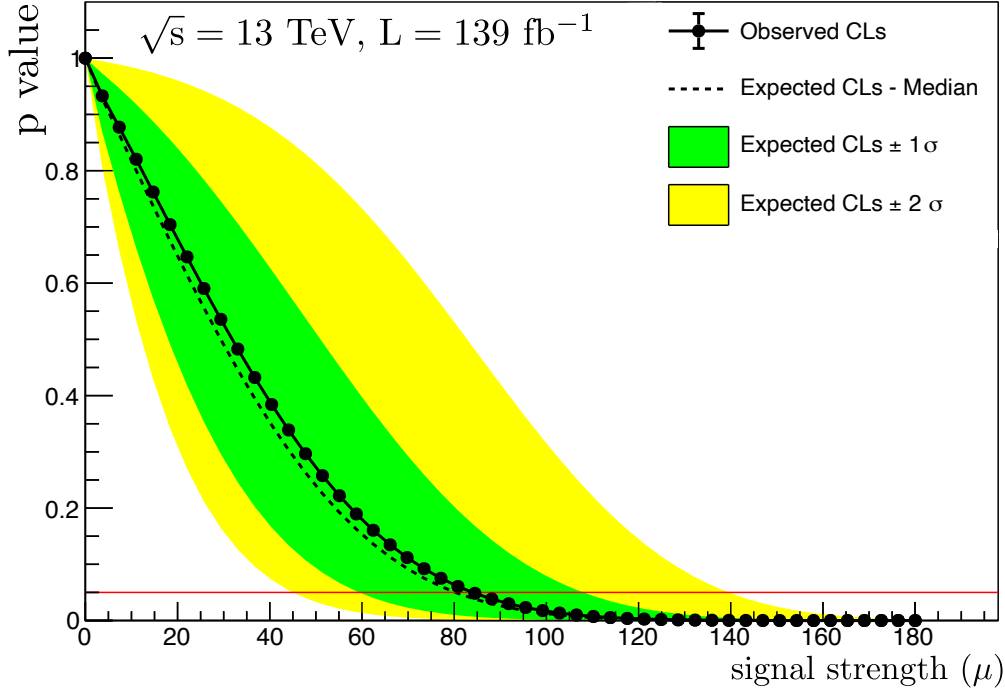


Figure 11.4 *NMDS Benchmark NMDS3 exclusion limit. Result obtained via use of the Frequentist method.*

The total data events observed in the signal region is 124 events. The expected exclusion limit for NMDS Benchmark NMDS3 when incorporating the full statistical uncertainty from the ABCD plane, a luminosity uncertainty of 1.7%, an experimental signal uncertainty of 1.7%, a conservative theoretical signal uncertainty of 5%, and a nonclosure uncertainty of 37% is a μ value of 80.3. This means that the exclusion limit is set at a cross section of 80.3 times the nominal non-minimal dark sector cross section. The cross section of the NCBH Benchmark NMDS3 is 1.33×10^{-4} pb, meaning that an expected exclusion limit can be placed on a cross section of 0.0107 pb. The observed exclusion limit is a μ of 84.03, meaning that an observed exclusion limit can be placed on a cross section of 0.011 pb. These limits can be seen in Figure 11.4.

11.6 Summary

This chapter discussed the statistical method used to quantify the probability of various model hypotheses, taking into account uncertainties, and finally deriving expected and observed exclusion limits. The method was then applied to all three models, and the outcome was briefly discussed. No significant excess was observed in any of the three searches. No evidence of string theory has been reported. The next chapter concludes the thesis and gives context to these results.

Chapter 12

Conclusion

The existence of extra dimensions could lower the fundamental Planck scale to the low TeV scale and very excitingly allow string theory to be probed at the LHC. This thesis presents novel searches for signatures of string theory via three models that have never been searched for by any experiment: scalar cascades, noncommutative black holes, and non-minimal dark sectors. This thesis includes the first time exploration of a novel final state characterized by a large multiplicity of soft jets and E_T^{miss} . These searches are performed on 139 fb^{-1} of proton-proton collision data at a center of mass energy $\sqrt{s} = 13 \text{ TeV}$. Deep learning is employed through the use of particle flow networks to learn jet substructure and design analysis regions. The background estimation is automated through the implementation of distance correlation in the neural network loss function. A novel jet reclustering algorithm, called “hybrid jet reclustering”, has been developed to overcome the challenges associated with a soft final state.

No significant excesses are observed with respect to the expected background processes in any of the three searches. For scalar cascades, considering benchmark SC1 (defined in Table 6.36) with a cross section of $5.33 \times 10^{-4} \text{ pb}$, observed upper limits are set at the 95% C.L. at a cross section of $7.7 \times 10^{-3} \text{ pb}$. For noncommutative black holes, considering benchmark NCBH1 (defined in Table 6.11) with a cross section of 0.022 pb for $n=5$ extra dimensions and $M_D = 6.5 \text{ TeV}$, observed upper limits are set at the 95% C.L. at a cross section of 0.23 pb . For non-minimal dark sectors, considering benchmark NMDS3 (defined in Table 6.12) with a cross section of 0.000133 pb , observed upper limits are set at the 95% C.L. at a cross section of 0.011 pb .

As no search existed from any experiment for any of these models, the analysis strategy for each model was designed from scratch. Challenges posed by each of these models were experienced for the first time. Hybrid jet reclustering has broad applications for many other physics searches. The neural network can similarly be employed in other searches. These searches serve as trailblazers for probing some of the signatures of string theory. With no significant excesses observed, we are yet to discover experimental evidence for string theory. These searches provide a roadmap for future searches, highlighting the challenges and solutions for complicated final states. The potential for LHC experiments to discover evidence of string theory is still immense. By continuing to develop novel search strategies and novel reconstruction techniques, target new models, improve the performance of the detector, and remain relentless in the absence of significance excesses, the LHC experiments may still one day be the first experiments to discover evidence of string theory.

Chapter 13

Appendix

13.1 Effects of tight cleaning and bad tile event veto on data

The effect of the bad tile module veto and tight cleaning cuts can be seen in Figure 13.1 for data 2017. The application of both the bad tile module veto and tight cleaning allowed for the elimination of peaks in the distribution of the ϕ_i value of E_T^{miss} in all data years.

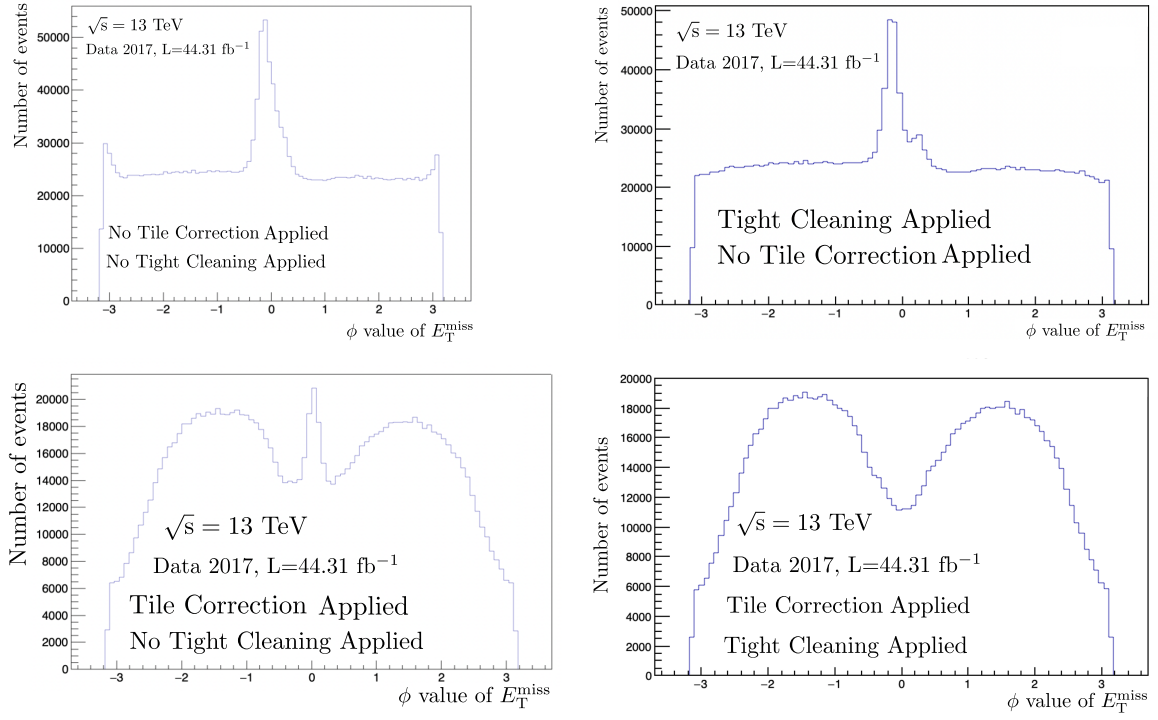


Figure 13.1 *Data 2017. Distribution of the ϕ_i value of E_T^{miss} in 4 cases: (top left) without the bad tile module event veto and without tight cleaning, (top right) without the bad tile module event veto and with tight cleaning, (bottom left) with the tile module event veto and without tight cleaning, and (bottom right) with the bad tile module event veto and with tight cleaning*

13.2 Systematic Uncertainties for Signal Monte Carlo Models

Systematic Uncertainty	Value
JET_Flavor_Response__1down	0.997
JET_Flavor_Response__1up	0.987
JET_Flavor_Composition__1down	0.982
JET_Flavor_Composition__1up	0.932
JET_Pileup_RhoTopology__1up	0.763
TRK_FAKE_RATE_TIGHT	0.691
JET_EffectiveNP_Modelling1__1down	0.478
JET_EffectiveNP_Modelling1__1up	0.476
JET_Pileup_OffsetNPV__1down	0.311
JET_Pileup_OffsetNPV__1up	0.280
JET_EtaIntercalibration_Modelling__1down	0.259
JET_EtaIntercalibration_Modelling__1up	0.258
JET_Pileup_OffsetMu__1up	0.194
JET_Pileup_OffsetMu__1down	0.188
TRK_EFF_TIGHT_GLOBAL	0.165
JET_EffectiveNP_Mixed1__1up	0.128
JET_EffectiveNP_Mixed2__1down	0.128
JET_Pileup_PtTerm__1down	0.125
JET_EffectiveNP_Mixed2__1up	0.123
JET_Pileup_PtTerm__1up	0.112
JET_EtaIntercalibration_NonClosure.2018data__1up	0.111
JET_EtaIntercalibration_NonClosure.2018data__1down	0.101
JET_EffectiveNP_Mixed1__1down	0.0952
TRK_EFF_TIGHT_IBL	0.0934
JET_JER_EffectiveNP_4__1down	0.0864
TRK_EFF_TIGHT_PHYSMODEL	0.0823
JET_JER_EffectiveNP_2__1down	0.0730
JET_EtaIntercalibration_TotalStat__1down	0.0730
JET_EtaIntercalibration_TotalStat__1up	0.0578
JET_EffectiveNP_Statistical3__1down	0.0537
JET_JER_EffectiveNP_9__1down	0.0461
JET_EffectiveNP_Statistical3__1up	0.0450
JET_JER_EffectiveNP_1__1up	0.0438
JET_JER_EffectiveNP_6__1up	0.0397
JET_JER_EffectiveNP_2__1up	0.0385
JET_EffectiveNP_Detector1__1up	0.0385
JET_EffectiveNP_Modelling3__1down	0.0374
JET_EffectiveNP_Modelling3__1up	0.0339

Table 13.1 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for non-minimal dark sector benchmark NMDS3 - Part 1 of 3 Tables*

Systematic Uncertainty	Value
JET_EffectiveNP_Detector1__1down	0.0334
JET_JER_EffectiveNP_4__1up	0.0321
JET_EffectiveNP_Modelling4__1up	0.0251
JET_JER_EffectiveNP_5__1down	0.0245
JET_EffectiveNP_Statistical4__1down	0.0234
JET_EffectiveNP_Statistical2__1up	0.0234
JET_EffectiveNP_Statistical2__1down	0.0210
JET_JER_EffectiveNP_6__1average	0.0207
JET_JER_EffectiveNP_8__1up	0.0199
JET_EffectiveNP_Modelling4__1down	0.0199
JET_EffectiveNP_Statistical4__1up	0.0193
JET_EffectiveNP_Modelling2__1down	0.0193
JET_JER_EffectiveNP_12restTerm__1up	0.0187
JET_EffectiveNP_Statistical6__1down	0.0181
JET_EffectiveNP_Statistical6__1up	0.0175
JET_EffectiveNP_Statistical1__1up	0.0175
TRK_EFF_TIGHT_PP0	0.0169
JET_JER_EffectiveNP_7__1down	0.0169
JET_EffectiveNP_Detector2__1up	0.0158
JET_EtaIntercalibration_NonClosure_negEta__1up	0.0150
JET_BJES_Response__1up	0.0140
JET_EffectiveNP_Statistical5__1up	0.0128
JET_EffectiveNP_Mixed3__1up	0.0123
JET_JER_EffectiveNP_11__1down	0.0123
JET_EffectiveNP_Statistical1__1down	0.0117
JET_EtaIntercalibration_NonClosure_posEta__1up	0.0105

Table 13.2 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for non-minimal dark sector benchmark NMDS3. Part 2 of 3 Tables. Systematics are continued from Table 13.1*

Systematic Uncertainty	Value
TRK_BIAS_Z0_WM	0.00993
JET_EffectiveNP_Mixed3__1down	0.00876
JET_EffectiveNP_Modelling2__1up	0.00817
MET_SoftTrk_ResoPara	0.00817
JET_JER_EffectiveNP_8__1down	0.00759
JET_JER_EffectiveNP_3__1up	0.00759
JET_EffectiveNP_Statistical5__1down	0.00642
JET_EtaIntercalibration_NonClosure_negEta__1down	0.00642
JET_EffectiveNP_Detector2__1down	0.00642
JET_JER_EffectiveNP_11__1up	0.00584
JET_JER_DataVsMC_MC16__1down	0.00584
JET_JER_EffectiveNP_5__1up	0.00526
JET_JER_EffectiveNP_10__1down	0.00467
JET_JER_EffectiveNP_7__1up	0.00409
JET_JER_EffectiveNP_10__1up	0.00350
JET_JER_EffectiveNP_9__1up	0.00350
MET_SoftTrk_ResoPerp	0.00175
JET_JER_EffectiveNP_6__1down	0.00175
TRK_RES_Z0_MEAS	0.00117
JET_JER_EffectiveNP_3__1down	0.00117
JET_BJES_Response__1down	0.00117
MET_SoftTrk_Scale__1up	0.00117
MET_SoftTrk_Scale__1down	0.00117
TRK_RES_D0_MEAS	0.000584
TRK_BIAS_QOVERP_SAGITTA_WM	0
MUON_CB__1up	0
MUON_CB__1down	0
MUON_SAGITTA_DATASTAT__1up	0
MUON_SAGITTA_DATASTAT__1down	0
JET_EtaIntercalibration_NonClosure_highE__1up	0
JET_EtaIntercalibration_NonClosure_highE__1down	0
JET_PunchThrough_MC16__1up	0
JET_PunchThrough_MC16__1down	0
JET_Pileup_RhoTopology__1down	0
TRK_BIAS_D0_WM	0
JET_JER_EffectiveNP_1__1down	0
EG_SCALE_ALL__1up	0
EG_SCALE_ALL__1down	0
MUON_SAGITTA_RESBIAS__1up	0
MUON_SAGITTA_RESBIAS__1down	0
EG_SCALE_AF2__1up	0
EG_SCALE_AF2__1down	0
JET_JER_EffectiveNP_12restTerm__1down	0
MUON_SCALE__1up	0
MUON_SCALE__1down	0
JET_EtaIntercalibration_NonClosure_posEta__1down	0
JET_SingleParticle_HighPt__1up	0
JET_SingleParticle_HighPt__1down	0
EG_RESOLUTION_ALL__1up	0
EG_RESOLUTION_ALL__1down	0
JET_JER_DataVsMC_MC16__1up	0

Table 13.3 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for non-minimal dark sector benchmark NMDS3. Part 3 of 3 Tables. Systematics are contained from Table 13.2*

Systematic Uncertainty	Value
JET_Flavor_Composition__1up	0.463
JET_Pileup_RhoTopology__1up	0.421
JET_Flavor_Composition__1down	0.401
JET_Flavor_Response__1down	0.319
JET_Flavor_Response__1up	0.279
JET_EffectiveNP_Modelling1__1up	0.205
JET_EffectiveNP_Modelling1__1down	0.166
JET_Pileup_OffsetNPV__1up	0.149
JET_EtaIntercalibration_Modelling__1up	0.122
TRK_FAKE_RATE_TIGHT	0.122
JET_Pileup_OffsetMu__1down	0.0947
JET_JER_EffectiveNP_5__1up	0.0701
MET_SoftTrk_ResoPerp	0.0701
JET_Pileup_OffsetNPV__1down	0.0701
MET_SoftTrk_Scale__1down	0.0659
MET_SoftTrk_Scale__1up	0.0653
JET_EtaIntercalibration_NonClosure_2018data__1up	0.0599
JET_EtaIntercalibration_Modelling__1down	0.0539
JET_JER_EffectiveNP_6__1up	0.0515
JET_JER_EffectiveNP_5__1down	0.0479
MET_SoftTrk_ResoPara	0.0461
JET_EtaIntercalibration_NonClosure_2018data__1down	0.0425
JET_JER_EffectiveNP_4__1down	0.0419
JET_Pileup_PtTerm__1up	0.0401
JET_EtaIntercalibration_TotalStat__1up	0.0389
JET_JER_EffectiveNP_12restTerm__1up	0.0360
JET_JER_EffectiveNP_9__1down	0.0348
JET_JER_EffectiveNP_7__1up	0.0318
JET_Pileup_PtTerm__1down	0.0312
JET_EffectiveNP_Statistical2__1down	0.0312
JET_EffectiveNP_Mixed2__1down	0.0276
JET_EffectiveNP_Mixed1__1up	0.0270
JET_BJES_Response__1up	0.0264
JET_JER_EffectiveNP_4__1up	0.0252
JET_EffectiveNP_Detector1__1up	0.0252
JET_JER_EffectiveNP_8__1up	0.0246
JET_JER_EffectiveNP_10__1down	0.0246
JET_EffectiveNP_Statistical4__1up	0.0234
JET_JER_EffectiveNP_10__1up	0.0216
JET_EffectiveNP_Modelling3__1down	0.0204
TRK_EFF_TIGHT_GLOBAL	0.0180
JET_EffectiveNP_Modelling2__1down	0.0174
JET_EffectiveNP_Modelling3__1up	0.0162
TRK_EFF_TIGHT_IBL	0.0144
JET_EffectiveNP_Mixed3__1down	0.0138
JET_EtaIntercalibration_NonClosure_posEta__1down	0.0138
TRK_EFF_TIGHT_PP0	0.0138
JET_EffectiveNP_Statistical6__1up	0.0132
JET_EffectiveNP_Modelling4__1up	0.0132
JET_JER_DataVsMC_MC16__1down	0.0126
JET_EffectiveNP_Statistical1__1down	0.0120

Table 13.4 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for noncommutative black hole benchmark NCBH1. Part 1 of 2 Tables*

Systematic Uncertainty	Value
JET_JER.EffectiveNP_11__1up	0.0120
JET_EffectiveNP_Modelling2__1up	0.0120
TRK_RES_Z0_MEAS	0.0114
JET_EffectiveNP_Mixed3__1up	0.0114
JET_EffectiveNP_Statistical2__1up	0.0114
JET_EtaIntercalibration_TotalStat__1down	0.0108
EG_RESOLUTION_ALL__1up	0.00959
JET_EffectiveNP_Mixed1__1down	0.00959
TRK_RES_D0_MEAS	0.00959
JET_EffectiveNP_Statistical5__1up	0.00899
JET_EffectiveNP_Statistical5__1down	0.00899
JET_EffectiveNP_Statistical3__1down	0.00899
JET_EffectiveNP_Detector2__1up	0.00899
JET_EffectiveNP_Detector2__1down	0.00899
JET_BJES_Response__1down	0.00839
JET_JER.EffectiveNP_8__1down	0.00839
JET_EffectiveNP_Statistical1__1up	0.00839
JET_EffectiveNP_Statistical6__1down	0.00719
TRK_EFF_TIGHT_PHYSMODEL	0.00719
JET_EffectiveNP_Detector1__1down	0.00659
JET_EffectiveNP_Modelling4__1down	0.00659
JET_JER.EffectiveNP_11__1down	0.00479
JET_EffectiveNP_Mixed2__1up	0.00419
JET_EtaIntercalibration_NonClosure_negEta__1down	0.00419
JET_Pileup_OffsetMu__1up	0.00360
JET_JER.EffectiveNP_9__1up	0.00360
JET_EtaIntercalibration_NonClosure_posEta__1up	0.00360
JET_EtaIntercalibration_NonClosure_negEta__1up	0.00300
JET_EffectiveNP_Statistical4__1down	0.00300
JET_EffectiveNP_Statistical3__1up	0.00300
EG_RESOLUTION_ALL__1down	0.00180
JET_JER.EffectiveNP_6__1down	0.00180
TRK_BIAS_D0_WM	0.00120
JET_JER.EffectiveNP_7__1down	0.00120
TRK_BIAS_Z0_WM	0.000600
JET_SingleParticle_HighPt__1up	0
JET_SingleParticle_HighPt__1down	0
MUON_SAGITTA_RESBIAS__1up	0
MUON_SAGITTA_RESBIAS__1down	0
EG_SCALE_AF2__1up	0
EG_SCALE_AF2__1down	0
JET_JER.DataVsMC_MC16__1up	0
TRK_BIAS_QOVERP_SAGITTA_WM	0
MUON_SAGITTA_DATASTAT__1up	0
MUON_SAGITTA_DATASTAT__1down	0
JET_Pileup_RhoTopology__1down	0
MUON_SCALE__1up	0
MUON_SCALE__1down	0
JET_PunchThrough_MC16__1up	0
JET_PunchThrough_MC16__1down	0
JET_EtaIntercalibration_NonClosure_highE__1up	0
JET_EtaIntercalibration_NonClosure_highE__1down	0
MUON_CB__1up	0
MUON_CB__1down	0
MUON_CB__1average	0

Table 13.5 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for noncommutative black hole benchmark NCBH1. Part 2 of 2 Tables, continued from Table 13.4*

Systematic Uncertainty	Value
JET_Flavor_Composition__1down	0.116
JET_Flavor_Composition__1up	0.114
JET_Pileup_RhoTopology__1up	0.112
JET_Flavor_Response__1down	0.0895
JET_Flavor_Response__1up	0.0869
MET_SoftTrk_Resodown	0.0811
MET_SoftTrk_Resoup	0.0801
MET_SoftTrk_Scale__1up	0.0783
MET_SoftTrk_Scale__1down	0.0745
JET_EffectiveNP_Modelling1__1up	0.0720
JET_EffectiveNP_Modelling1__1down	0.0687
JET_EtaIntercalibration_Modelling__1up	0.0505
JET_EtaIntercalibration_Modelling__1down	0.0465
TRK_FAKE_RATE_TIGHTup	0.0381
TRK_FAKE_RATE_TIGHTdown	0.0381
JET_Pileup_OffsetNPV__1up	0.0604
JET_Pileup_OffsetNPV__1down	0.0571
JET_JER_EffectiveNP_4__1down	0.0371
JET_JER_EffectiveNP_2__1up	0.0368
JET_Pileup_OffsetMu__1up	0.0366
JET_EtaIntercalibration_TotalStat__1up	0.0322
JET_Pileup_OffsetMu__1down	0.0309
JET_EtaIntercalibration_NonClosure_2018data__1down	0.0308
JET_EtaIntercalibration_NonClosure_2018data__1up	0.0298
JET_JER_EffectiveNP_1__1up	0.030
JET_EtaIntercalibration_TotalStat__1down	0.0295
JET_JER_EffectiveNP_8__1up	0.0269
JET_JER_EffectiveNP_9__1up	0.0261
JET_JER_EffectiveNP_6__1down	0.0257
JET_EffectiveNP_Statistical2__1down	0.0234
JET_JER_EffectiveNP_3__1down	0.0231
JET_JER_EffectiveNP_12restTerm__1up	0.0231
JET_EffectiveNP_Statistical2__1up	0.0222
JET_EffectiveNP_Mixed2__1down	0.0218

Table 13.6 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for Scalar Cascade Benchmark SC1. Part 1 of 3 Tables*

Systematic Uncertainty	Value
TRK_EFF_TIGHT_GLOBALdown	0.0211
TRK_EFF_TIGHT_GLOBALup	0.0211
JET_EffectiveNP_Mixed2__1up	0.0211
JET_JER_EffectiveNP_10__1down	0.0209
JET_JER_EffectiveNP_11__1up	0.0208
JET_JER_EffectiveNP_5__1up	0.0198
JET_JER_DataVsMC_MC16__1down	0.0194
JET_EffectiveNP_Mixed1__1down	0.0185
JET_Pileup_PtTerm__1up	0.0181
JET_JER_EffectiveNP_5__1down	0.0167
JET_JER_EffectiveNP_7__1down	0.0164
JET_EffectiveNP_Mixed1__1updown	0.0166
JET_EffectiveNP_Detector1__1up	0.0163
JET_EffectiveNP_Mixed1__1up	0.0166
JET_JER_EffectiveNP_7__1up	0.0160
JET_EffectiveNP_Modelling2__1down	0.0157
JET_EffectiveNP_Modelling2__1up	0.0154
JET_JER_EffectiveNP_9__1down	0.0151
JET_EffectiveNP_Detector1__1down	0.0150
JET_Pileup_PtTerm__1down	0.0144
JET_EffectiveNP_Modelling3__1up	0.0134
JET_JER_EffectiveNP_8__1down	0.0131
JET_EffectiveNP_Mixed3__1down	0.0123
JET_JER_EffectiveNP_6__1up	0.0122
JET_EffectiveNP_Modelling3__1down	0.0121
JET_JER_EffectiveNP_11__1down	0.0118
JET_EffectiveNP_Mixed3__1up	0.0114
TRK_EFF_TIGHT_PHYSMODELup	0.0109
TRK_EFF_TIGHT_PHYSMODELdown	0.0109

Table 13.7 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for Scalar Cascade Benchmark SC1. Part 2 of 3 Tables, continued from Table 13.6*

Systematic Uncertainty	Value
JET_EffectiveNP.Statistical6__1down	0.0108
TRK_EFF_TIGHT_IBLdown	0.0126
TRK_EFF_TIGHT_IBLup	0.0126
JET_EffectiveNP.Modelling4__1down	0.00992
EG_RESOLUTION_ALL__1up	0.000945
EG_SCALE_ALL__1down	0.000945
TRK_EFF_TIGHT_PP0down	0.00930
TRK_EFF_TIGHT_PP0up	0.00930
JET_EffectiveNP.Detector2__1up	0.000908
JET_EffectiveNP.Statistical5__1down	0.00880
JET_JER.EffectiveNP__2__1down	0.00879
JET_EtaIntercalibration.NonClosure.negEta__1up	0.00787
JET_EffectiveNP.Statistical1__1down	0.00780
JET_EtaIntercalibration.NonClosure.posEta__1down	0.00742
JET_EffectiveNP.Modelling4__1up	0.00661
JET_EffectiveNP.Statistical6__1up	0.00660
JET_EffectiveNP.Statistical3__1down	0.00640
JET_EffectiveNP.Statistical5__1up	0.00635
JET_EffectiveNP.Statistical1__1up	0.00562
JET_EffectiveNP.Detector2__1downdown	0.00515
JET_EffectiveNP.Detector2__1downup	0.00515
JET_EffectiveNP.Statistical3__1up	0.00436
JET_EffectiveNP.Statistical4__1up	0.00426
JET_EffectiveNP.Statistical4__1down	0.00359
TRK_RES_Z0_MEASdown	0.00352
TRK_RES_Z0_MEASup	0.00352
JET_EtaIntercalibration.NonClosure.negEta__1down	0.00509
JET_JER.EffectiveNP_10__1up	0.004306
JET_JER.EffectiveNP_4__1up	0.00417
JET_EtaIntercalibration.NonClosure.posEta__1up	0.00289
JET_JER.EffectiveNP_3__1up	0.00255
TRK_RES_D0_MEASdown	0.00216
TRK_RES_D0_MEASup	0.00216
TRK_BIAS_Z0_WMdown	0.00212
TRK_BIAS_Z0_WMup	0.00212
JET_BJES_Response__1up	0.00199
JET_BJES_Response__1down	0.00158
TRK_BIAS_D0_WMdown	0.000101
TRK_BIAS_D0_WMup	0.000101
TRK_BIAS_QOVERP_SAGITTA_WMdown	0
TRK_BIAS_QOVERP_SAGITTA_WMup	0
MUON_CB__1down	0
MUON_CB__1up	0
MUON_SAGITTA_DATASTAT__1down	0
MUON_SAGITTA_DATASTAT__1up	0
MUON_SAGITTA_RESBIAS__1down	0
MUON_SAGITTA_RESBIAS__1up	0
MUON_SCALE__1down	0
MUON_SCALE__1up	0
EG_RESOLUTION_ALL__1down	0
EG_SCALE_AF2__1down	0
EG_SCALE_AF2__1up	0
JET_EtaIntercalibration.NonClosure.highE__1updown	0
JET_EtaIntercalibration.NonClosure.highE__1upup	0
JET_EtaIntercalibration.NonClosure.highE__1downdown	0
JET_EtaIntercalibration.NonClosure.highE__1downup	0
JET_JER.DataVsMC.MC16__1up	0
JET_JER.EffectiveNP_12restTerm__1down	0
JET_JER.EffectiveNP_1__1down	0
JET_Pileup_RhoTopology__1down	0
JET_PunchThrough_MC16__1down	0
JET_PunchThrough_MC16__1up	0
JET_SingleParticle.HighPt__1down	0
JET_SingleParticle.HighPt__1up	0

Table 13.8 *Systematic values (in percentages) in order from greatest to lowest (in percentages) for Scalar Cascade Benchmark SC1. Part 3 of 3 Tables, continued from Table 13.7*

13.3 Full names of background MC samples, signal MC samples, and data samples used

Below is a list of the full Background MC Sample names used in the searches in this thesis for Run 2 MC Campaign MC16a. The DSID is highlighted in red.

MC16A

mc16_13TeV.**700323**.Sh.2211_Zmumu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700324**.Sh.2211_Zmumu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700325**.Sh.2211_Zmumu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700326**.Sh.2211_Ztautau_LL_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700327**.Sh.2211_Ztautau_LL_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700328**.Sh.2211_Ztautau_LL_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700329**.Sh.2211_Ztautau_LH_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700330**.Sh.2211_Ztautau_LH_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700331**.Sh.2211_Ztautau_LH_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700332**.Sh.2211_Ztautau_HH_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700333**.Sh.2211_Ztautau_HH_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700334**.Sh.2211_Ztautau_HH_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700335**.Sh.2211_Znnu_pTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700336**.Sh.2211_Znnu_pTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700337**.Sh.2211_Znnu_pTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700338**.Sh.2211_Wenu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700339**.Sh.2211_Wenu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700340**.Sh.2211_Wenu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700341**.Sh.2211_Wmunu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700342**.Sh.2211_Wmunu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700343**.Sh.2211_Wmunu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700344**.Sh.2211_Wtaunu_L_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700345**.Sh.2211_Wtaunu_L_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700346**.Sh.2211_Wtaunu_L_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700347**.Sh.2211_Wtaunu_H_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700348**.Sh.2211_Wtaunu_H_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700349**.Sh.2211_Wtaunu_H_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r9364.p5461
mc16_13TeV.**700439**.Sh.2211_Zqq-ptZ.100.ECMS.deriv.DAOD_EXOT28.e8338.s3126.r9364.p5461
mc16_13TeV.**700441**.Sh.2211_Wqq-ptW.200.ECMS.deriv.DAOD_EXOT28.e8419.s3126.r9364.p5461
mc16_13TeV.**410470**.PhPy8EG_A14.ttbar_hdamp258p75_nonallhad.deriv.DAOD_EXOT28.e6337.s3126.r9364.p5461
mc16_13TeV.**410471**.PhPy8EG_A14.ttbar_hdamp258p75_allhad.deriv.DAOD_EXOT28.e6337.s3126.r9364.p5461
mc16_13TeV.**410644**.PowhegPythia8EvtGen_A14.singletop_schan_lept_top.deriv.DAOD_EXOT28.e6527.s3126.r9364.p5461
mc16_13TeV.**410645**.PowhegPythia8EvtGen_A14.singletop_schan_lept_antitop.deriv.DAOD_EXOT28.e6527.s3126.r9364.p5461
mc16_13TeV.**410646**.PowhegPythia8EvtGen_A14.Wt_DR_inclusive_top.deriv.DAOD_EXOT28.e6552.s3126.r9364.p5461
mc16_13TeV.**410647**.PowhegPythia8EvtGen_A14.Wt_DR_inclusive_antitop.deriv.DAOD_EXOT28.e6552.s3126.r9364.p5461
mc16_13TeV.**410658**.PhPy8EG_A14.tchan.BW50_lept_top.deriv.DAOD_EXOT28.e6671.s3126.r9364.p5461
mc16_13TeV.**410659**.PhPy8EG_A14.tchan.BW50_lept_antitop.deriv.DAOD_EXOT28.e6671.s3126.r9364.p5461
mc16_13TeV.**364702**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364703**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364704**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364705**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ5WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364706**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ6WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364707**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ7WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364708**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ8WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364709**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ9WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364710**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ10WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364711**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ11WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**364712**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ12WithSW.deriv.DAOD_EXOT28.e7142.s3126.r9364.p5461
mc16_13TeV.**361106**.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EXOT28.e3601.s3126.r9364.p5461
mc16_13TeV.**363355**.Sherpa_221_NNPDF30NNLO_ZqqZvv.deriv.DAOD_EXOT28.e5525.s3126.r9364.p5461
mc16_13TeV.**363357**.Sherpa_221_NNPDF30NNLO_WqqZvv.deriv.DAOD_EXOT28.e5525.s3126.r9364.p5461
mc16_13TeV.**363360**.Sherpa_221_NNPDF30NNLO_WplvWmqq.deriv.DAOD_EXOT28.e5983.s3126.r9364.p5461
mc16_13TeV.**363489**.Sherpa_221_NNPDF30NNLO_WlvZqq.deriv.DAOD_EXOT28.e5525.s3126.r9364.p5461
mc16_13TeV.**363359**.Sherpa_221_NNPDF30NNLO_WpeqWmlv.deriv.DAOD_EXOT28.e5583.s3126.r9364.p5461

Below is a list of the full Background MC Sample names used in the searches in this thesis for Run 2 MC Campaign MC16d. The DSID is highlighted in red.

MC16D

mc16_13TeV.**700323**.Sh.2211_Zmumu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700324**.Sh.2211_Zmumu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700325**.Sh.2211_Zmumu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461

mc16_13TeV.**700326**.Sh_2211_Ztautau_LL_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700327**.Sh_2211_Ztautau_LL_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700328**.Sh_2211_Ztautau_LL_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700329**.Sh_2211_Ztautau_LH_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700330**.Sh_2211_Ztautau_LH_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700331**.Sh_2211_Ztautau_LH_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700332**.Sh_2211_Ztautau_HH_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700333**.Sh_2211_Ztautau_HH_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700334**.Sh_2211_Ztautau_HH_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700335**.Sh_2211_Znnu_pTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700336**.Sh_2211_Znnu_pTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700337**.Sh_2211_Znnu_pTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700338**.Sh_2211_Wenu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700339**.Sh_2211_Wenu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700340**.Sh_2211_Wenu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700341**.Sh_2211_Wmunu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700342**.Sh_2211_Wmunu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700343**.Sh_2211_Wmunu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700344**.Sh_2211_Wtaunu_L_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700345**.Sh_2211_Wtaunu_L_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700346**.Sh_2211_Wtaunu_L_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700347**.Sh_2211_Wtaunu_H_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700348**.Sh_2211_Wtaunu_H_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**700349**.Sh_2211_Wtaunu_H_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10201.p5461
mc16_13TeV.**410644**.PowhegPythia8EvtGen_A14_singlelep_schan_lept_top.deriv.DAOD_EXOT28.e6527.s3126.r10201.p5461
mc16_13TeV.**410645**.PowhegPythia8EvtGen_A14_singlelep_schan_lept_antitop.deriv.DAOD_EXOT28.e6527.s3126.r10201.p5461
mc16_13TeV.**410646**.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_top.deriv.DAOD_EXOT28.e6552.s3126.r10201.p5461
mc16_13TeV.**410647**.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_antitop.deriv.DAOD_EXOT28.e6552.s3126.r10201.p5461
mc16_13TeV.**410658**.PhPy8EG_A14_tchan.BW50_lept_top.deriv.DAOD_EXOT28.e6671.s3126.r10201.p5461
mc16_13TeV.**410659**.PhPy8EG_A14_tchan.BW50_lept_antitop.deriv.DAOD_EXOT28.e6671.s3126.r10201.p5461
mc16_13TeV.**410470**.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_EXOT28.e6337.s3126.r10201.p5461
mc16_13TeV.**410471**.PhPy8EG_A14_ttbar_hdamp258p75_allhad.deriv.DAOD_EXOT28.e6337.s3126.r10201.p5461
mc16_13TeV.**364702**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364703**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364704**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364705**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ5WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364706**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ6WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364707**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ7WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364708**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ8WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364709**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ9WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364710**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ10WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364711**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ11WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**364712**.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ12WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10201.p5461
mc16_13TeV.**363355**.Sherpa_221_NNPDF30NNLO_ZqqZvv.deriv.DAOD_EXOT28.e5525.s3126.r10201.p5461
mc16_13TeV.**363357**.Sherpa_221_NNPDF30NNLO_WqqZvv.deriv.DAOD_EXOT28.e5525.s3126.r10201.p5461
mc16_13TeV.**363360**.Sherpa_221_NNPDF30NNLO_WplvWmqq.deriv.DAOD_EXOT28.e5983.s3126.r10201.p5461
mc16_13TeV.**363489**.Sherpa_221_NNPDF30NNLO_WlvZqq.deriv.DAOD_EXOT28.e5525.s3126.r10201.p5461
mc16_13TeV.**363359**.Sherpa_221_NNPDF30NNLO_WpqqWmlv.deriv.DAOD_EXOT28.e5583.s3126.r10201.p5461
mc16_13TeV.**361106**.PowhegPythia8EvtGen_AZNLOC7TEQ6L1_Zee.deriv.DAOD_EXOT28.e3601.s3126.r10201.p5461

Below is a list of the full Background MC Sample names used in the searches in this thesis for Run 2 MC Campaign MC16e. The DSID is highlighted in red.

MC16E

mc16_13TeV.**700323**.Sh_2211_Zmumu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700324**.Sh_2211_Zmumu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700325**.Sh_2211_Zmumu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700326**.Sh_2211_Ztautau_LL_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700327**.Sh_2211_Ztautau_LL_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700328**.Sh_2211_Ztautau_LL_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700329**.Sh_2211_Ztautau_LH_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700330**.Sh_2211_Ztautau_LH_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700331**.Sh_2211_Ztautau_LH_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700332**.Sh_2211_Ztautau_HH_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700333**.Sh_2211_Ztautau_HH_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700334**.Sh_2211_Ztautau_HH_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700335**.Sh_2211_Znnu_pTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700336**.Sh_2211_Znnu_pTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700337**.Sh_2211_Znnu_pTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700338**.Sh_2211_Wenu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700339**.Sh_2211_Wenu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700340**.Sh_2211_Wenu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700341**.Sh_2211_Wmunu_maxHTpTV2_BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700342**.Sh_2211_Wmunu_maxHTpTV2_CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.**700343**.Sh_2211_Wmunu_maxHTpTV2_CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461

mc16_13TeV.700344.Sh.2211.Wtaunu.L_maxHTpTV2.BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.700345.Sh.2211.Wtaunu.L_maxHTpTV2.CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.700346.Sh.2211.Wtaunu.L_maxHTpTV2.CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.700347.Sh.2211.Wtaunu.H_maxHTpTV2.BFilter.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.700348.Sh.2211.Wtaunu.H_maxHTpTV2.CFilterBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.700349.Sh.2211.Wtaunu.H_maxHTpTV2.CVetoBVeto.deriv.DAOD_EXOT28.e8351.s3126.r10724.p5461
mc16_13TeV.700439.Sh.2211.Zqq-ptZ.100.ECMS.deriv.DAOD_EXOT28.e8338.s3126.r10724.p5461
mc16_13TeV.700441.Sh.2211.Wqq-ptW.200.ECMS.deriv.DAOD_EXOT28.e8419.s3126.r10724.p5461
mc16_13TeV.410644.PowhegPythia8EvtGen_A14.singletop_schan_lept_top.deriv.DAOD_EXOT28.e6527.s3126.r10724.p5461
mc16_13TeV.410645.PowhegPythia8EvtGen_A14.singletop_schan_lept_antitop.deriv.DAOD_EXOT28.e6527.s3126.r10724.p5461
mc16_13TeV.410646.PowhegPythia8EvtGen_A14.Wt_DR_inclusive_top.deriv.DAOD_EXOT28.e6552.s3126.r10724.p5461
mc16_13TeV.410647.PowhegPythia8EvtGen_A14.Wt_DR_inclusive_antitop.deriv.DAOD_EXOT28.e6552.s3126.r10724.p5461
mc16_13TeV.410658.PyPy8EG_A14.tchan.BW50.lept_top.deriv.DAOD_EXOT28.e6671.s3126.r10724.p5461
mc16_13TeV.410659.PyPy8EG_A14.tchan.BW50.lept_antitop.deriv.DAOD_EXOT28.e6671.s3126.r10724.p5461
mc16_13TeV.364702.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364703.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364704.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364705.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ5WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364706.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ6WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364707.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ7WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364708.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ8WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364709.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ9WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364710.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ10WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364711.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ11WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.364712.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ12WithSW.deriv.DAOD_EXOT28.e7142.s3126.r10724.p5461
mc16_13TeV.363355.Sherpa_221.NNPDF30NNLO_ZqqZvv.deriv.DAOD_EXOT28.e5525.s3126.r10724.p5461
mc16_13TeV.363357.Sherpa_221.NNPDF30NNLO_WqqZvv.deriv.DAOD_EXOT28.e5525.s3126.r10724.p5461
mc16_13TeV.363359.Sherpa_221.NNPDF30NNLO_WpqqWmlv.deriv.DAOD_EXOT28.e5583.s3126.r10724.p5461
mc16_13TeV.363360.Sherpa_221.NNPDF30NNLO_WplvWmqq.deriv.DAOD_EXOT28.e5983.s3126.r10724.p5461
mc16_13TeV.363489.Sherpa_221.NNPDF30NNLO_WlvZqq.deriv.DAOD_EXOT28.e5525.s3126.r10724.p5461
mc16_13TeV.410470.PyPy8EG_A14.ttbar_hdamp258p75.nonallhad.deriv.DAOD_EXOT28.e6337.s3126.r10724.p5461
mc16_13TeV.410471.PyPy8EG_A14.ttbar_hdamp258p75.allhad.deriv.DAOD_EXOT28.e6337.s3126.r10724.p5461
mc16_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EXOT28.e3601.s3126.r10724.p5461

Below are the list of signal samples used in the searches for Run 2 MC Campaigns MC16a, MC16d, and MC16e. The DSID is highlighted in red.

MC16A

mc16_13TeV.512597.MGPy8EG_A14NNPDF23.NMDS.B3.600.10.3.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512598.MGPy8EG_A14NNPDF23.NMDS.B4.Benchmark_A.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512599.MGPy8EG_A14NNPDF23.NMDS.B5.561.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512601.MGPy8EG_A14NNPDF23.SC.2BStableCas_Phi5.B1.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512602.MGPy8EG_A14NNPDF23.SC.3BStableCas_Phi5.B1.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512603.MGPy8EG_A14NNPDF23.SC.3BStableCas_Phi4.B2.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512604.MGPy8EG_A14NNPDF23.SC.2BUNSTABLECas_Phi6.B1.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.512606.MGPy8EG_A14NNPDF23.SC.3BodyChain_Phi7.B2.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.510217.MGPy8EG_A14NNPDF23.SC.2BStableCascade_Phi7B10.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.510218.MGPy8EG_A14NNPDF23.NMDS.650.String.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.510410.MGPy8EG_A14NNPDF23.NMDS.800.50.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.510409.MGPy8EG_A14NNPDF23.NMDS.1841.String.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.900505.CharybdisPy8EG_A14NNPDF23.65NC55.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461
mc16_13TeV.901951.CharybdisPy8EG_A14NNPDF23.6NC63.deriv.DAOD_EXOT28.e8304.s3126.r9364.p5461

MC16D

mc16_13TeV.512597.MGPy8EG_A14NNPDF23.NMDS.B3.600.10.3.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512598.MGPy8EG_A14NNPDF23.NMDS.B4.Benchmark_A.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512599.MGPy8EG_A14NNPDF23.NMDS.B5.561.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512601.MGPy8EG_A14NNPDF23.SC.2BStableCas_Phi5.B1.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512602.MGPy8EG_A14NNPDF23.SC.3BStableCas_Phi5.B1.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512603.MGPy8EG_A14NNPDF23.SC.3BStableCas_Phi4.B2.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512604.MGPy8EG_A14NNPDF23.SC.2BUNSTABLECas_Phi6.B1.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.512606.MGPy8EG_A14NNPDF23.SC.3BodyChain_Phi7.B2.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.510217.MGPy8EG_A14NNPDF23.SC.2BStableCascade_Phi7B10.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.510218.MGPy8EG_A14NNPDF23.NMDS.650.String.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.510410.MGPy8EG_A14NNPDF23.NMDS.800.50.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.510409.MGPy8EG_A14NNPDF23.NMDS.1841.String.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.900505.CharybdisPy8EG_A14NNPDF23.65NC55.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461
mc16_13TeV.901951.CharybdisPy8EG_A14NNPDF23.6NC63.deriv.DAOD_EXOT28.e8304.s3126.r10201.p5461

MC16E

mc16_13TeV.512597.MGPy8EG_A14NNPDF23.NMDS.B3.600.10.3.deriv.DAOD_EXOT28.e8304.s3126.r10724.p5461
mc16_13TeV.512598.MGPy8EG_A14NNPDF23.NMDS.B4.Benchmark_A.deriv.DAOD_EXOT28.e8304.s3126.r10724.p5461
mc16_13TeV.512599.MGPy8EG_A14NNPDF23.NMDS.B5.561.deriv.DAOD_EXOT28.e8304.s3126.r10724.p5461

mc16_13TeV.512601.MGPy8EG_A14NNPDF23.SC_2BStableCas_Phi5_B1.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.512602.MGPy8EG_A14NNPDF23.SC_3BStableCas_Phi5_B1.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.512603.MGPy8EG_A14NNPDF23.SC_3BStableCas_Phi4_B2.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.512604.MGPy8EG_A14NNPDF23.SC_2BUNSTABLECas_Phi6_B1.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.512606.MGPy8EG_A14NNPDF23.SC_3BodyChain_Phi17_B2.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.510217.MGPy8EG_A14NNPDF23.SC_2BStableCascade_Phi7B10.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.510218.MGPy8EG_A14NNPDF23.NMDS_650_String.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.510410.MGPy8EG_A14NNPDF23.NMDS_800_50.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.510409.MGPy8EG_A14NNPDF23.NMDS_1841_String.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.900505.CharybdisPy8EG_A14NNPDF23_65NC55.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461
mc16_13TeV.901951.CharybdisPy8EG_A14NNPDF23_6NC63.deriv.DAOD_EXOT28.e8304_s3126_r10724_p5461

Below are the data samples used in the searches for this thesis for all data-taking years of Run 2: 2015, 2016, 2017, and 2018

Data 15 data15_13TeV.periodD.physics_Main.PhysCont.DAOD_EXOT28.grp15_v01_p5462
data15_13TeV.periodE.physics_Main.PhysCont.DAOD_EXOT28.grp15_v01_p5462
data15_13TeV.periodF.physics_Main.PhysCont.DAOD_EXOT28.grp15_v01_p5462
data15_13TeV.periodG.physics_Main.PhysCont.DAOD_EXOT28.grp15_v01_p5462
data15_13TeV.periodH.physics_Main.PhysCont.DAOD_EXOT28.grp15_v01_p5462
data15_13TeV.periodJ.physics_Main.PhysCont.DAOD_EXOT28.grp15_v01_p5462

Data 16 data16_13TeV.periodA.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodB.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodC.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodD.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodE.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodF.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodG.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodH.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodI.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462
data16_13TeV.periodJ.physics_Main.PhysCont.DAOD_EXOT28.grp16_v01_p5462

Data 17 data17_13TeV.periodB.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodC.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodD.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodE.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodF.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodG.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodH.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462
data17_13TeV.periodI.physics_Main.PhysCont.DAOD_EXOT28.grp17_v01_p5462

Data 18 data18_13TeV.periodB.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodC.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodD.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodE.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodF.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodG.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodH.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodI.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodJ.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodK.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodL.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodM.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodN.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodO.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462
data18_13TeV.periodP.physics_Main.PhysCont.DAOD_EXOT28.grp18_v01_p5462

13.4 Cross sections, filter efficiencies, and k Factors for MC Background samples used

MC16 DSID	A	Xsec	Filter Efficiency	k Factor
361106		1.900E+03	1	1.026
363355		1.560E+01	2.800E-01	1
363357		6.800E+00	1	1
363359		2.470E+01	1	1
363360		2.470E+01	1	1
363489		1.140E+01	1	1
364702		2.433E+09	9.863E-03	1
364703		2.645E+07	1.166E-02	1
364705		4.550E+03	1.453E-02	1
364706		2.580E+02	9.472E-03	1
364707		1.620E+01	1.110E-02	1
364708		6.250E-01	1.015E-02	1
364709		1.960E-02	1.206E-02	1
364710		1.200E-03	5.893E-03	1
364711		4.226E-05	2.673E-03	1
364712		1.040E-06	4.289E-04	1
410470		7.298E+02	5.438E-01	1
410471		7.298E+02	4.562E-01	1
410644		2.027E+00	1	1.017
410645		1.268E+00	1	1.017
410646		3.790E+01	1	1
410647		3.790E+01	1	0.946
410658		3.699E+01	1	1
410659		2.217E+01	1	1.185
364704		2.546E+05	1.337E-02	1
700323		2.221E+03	2.439E-02	1
700324		2.221E+03	1.300E-01	1
700325		2.221E+03	8.464E-01	1
700326		2.753E+02	2.491E-02	1
700327		2.221E+03	1.254E-01	1
700328		2.753E+02	8.504E-01	1
700329		2.221E+03	2.443E-02	1
700330		2.221E+03	1.255E-01	1
700331		1.013E+03	8.502E-01	1
700332		2.222E+03	2.460E-02	1
700333		9.330E+02	1.248E-01	1
700334		9.325E+02	8.500E-01	1
700335		4.471E+02	8.427E-02	1
700336		4.471E+02	2.024E-01	1
700337		4.471E+02	7.127E-01	1
700338		2.174E+04	9.376E-03	1
700339		2.174E+04	1.490E-01	1
700340		2.174E+04	8.436E-01	1
700341		2.181E+04	9.797E-03	1
700342		2.181E+04	1.460E-01	1
700343		2.181E+04	8.436E-01	1
700344		2.182E+04	9.011E-03	1
700345		7.680E+03	1.460E-01	1
700346		7.680E+03	8.475E-01	1
700347		1.413E+04	8.678E-03	1
700348		1.413E+04	1.434E-01	1
700349		1.413E+04	8.475E-01	1
700439		4.900E+02	1	1
700441		1.447E+02	1	1
510218		1.100E-01	1	1
510217		5.330E-04	1	1
510409		1.330E-04	1	1
510410		2.710E-02	1	1
900505		2.420E-02	1	1
901951		5.067E-02	1	1
512597		1.640E-01	1	1
512598		7.780E-03	1	1
512599		2.150E-01	1	1
512601		1.030E-03	1	1
512602		7.750E-04	1	1
512603		8.350E-04	1	1
512604		3.070E-03	1	1
512605		5.330E-04	1	1
512606		1.200E-03	1	1

Table 13.9 *Cross sections, filter efficiencies, and k Factors for MC background samples corresponding to Run 2 MC Campaign MC16a*

MC16 DSID	D	Xsec	Filter Efficiency	k Factor
361106		1.900E+03	1	1.026
363355		1.560E+01	2.799E-01	1
363357		6.800E+00	1	1
363359		2.470E+01	1	1
363360		2.470E+01	1	1
363489		1.140E+01	1	1
364702		2.433E+09	9.863E-03	1
364703		2.645E+07	1.166E-02	1
364705		4.550E+03	1.453E-02	1
364706		2.580E+02	9.472E-03	1
364707		1.620E+01	1.110E-02	1
364708		6.250E-01	1.015E-02	1
364709		1.960E-02	1.206E-02	1
364710		1.200E-03	5.893E-03	1
364711		4.226E-05	2.673E-03	1
364712		1.040E-06	4.289E-04	1
410470		7.298E+02	5.438E-01	1
410471		7.298E+02	4.562E-01	1
410644		2.027E+00	1	1.017
410645		1.268E+00	1	1.017
410646		3.790E+01	1	1
410647		3.790E+01	1	1
410658		3.699E+01	1	1
410659		2.217E+01	1	1.185
364704		2.546E+05	1.337E-02	1
700323		2221.2	2.439E-02	1
700324		2.221E+03	1.300E-01	1
700325		2.221E+03	8.464E-01	1
700326		2.753E+02	2.491E-02	1
700327		2.221E+03	1.254E-01	1
700328		2.753E+02	8.504E-01	1
700329		2.221E+03	2.443E-02	1
700330		2.221E+03	1.255E-01	1
700331		1.013E+03	8.502E-01	1
700332		2.222E+03	2.460E-02	1
700333		9.330E+02	1.248E-01	1
700334		9.325E+02	8.500E-01	1
700335		4.471E+02	8.427E-02	1
700336		4.471E+02	2.024E-01	1
700337		4.471E+02	7.127E-01	1
700338		2.174E+04	9.376E-03	1
700339		2.174E+04	1.490E-01	1
700340		2.174E+04	8.436E-01	1
700341		2.181E+04	9.797E-03	1
700342		2.181E+04	1.460E-01	1
700343		2.181E+04	8.436E-01	1
700344		2.182E+04	9.011E-03	1
700345		7.680E+03	1.460E-01	1
700346		7.680E+03	8.475E-01	1
700347		1.413E+04	8.678E-03	1
700348		1.413E+04	1.434E-01	1
700349		1.413E+04	8.475E-01	1
700439		4.900E+02	1	1
700441		1.447E+02	1	1
510218		1.100E-01	1	1
510217		5.330E-04	1	1
510409		1.330E-04	1	1
510410		2.710E-02	1	1
900505		2.420E-02	1	1
901951		5.067E-02	1	1
512597		1.640E-01	1	1
512598		7.780E-03	1	1
512599		2.150E-01	1	1
512601		1.030E-03	1	1
512602		7.750E-04	1	1
512603		8.350E-04	1	1
512604		3.070E-03	1	1
512605		5.330E-04	1	1
512606		1.200E-03	1	1

Table 13.10 *Cross sections, filter efficiencies, and k Factors for MC background samples corresponding to Run 2 MC Campaign MC16d*

MC16E DSID	Xsec	Filter Efficiency	k Factor
361106	1.900E+03	1	1
363355	1.560E+01	2.799E-01	1
363357	6.800E+00	1	1
363359	2.470E+01	1	1
363360	2.470E+01	1	1
363489	1.140E+01	1	1
364702	2.433E+09	9.863E-03	1
364703	2.645E+07	1.166E-02	1
364705	4.550E+03	1.453E-02	1
364706	2.580E+02	9.472E-03	1
364707	1.620E+01	1.110E-02	1
364708	6.250E-01	1.015E-02	1
364709	1.960E-02	1.206E-02	1
364710	1.200E-03	5.893E-03	1
364711	4.226E-05	2.673E-03	1
364712	1.040E-06	4.289E-04	1
410470	7.298E+02	5.438E-01	1
410471	7.298E+02	4.562E-01	1
410644	2.027E+00	1	1.017
410645	1.268E+00	1	1.017
410646	3.794E+01	1	0.945
410647	3.791E+01	1	0.946
410658	3.699E+01	1	1
410659	2.217E+01	1	1.185
364704	2.546E+05	1.337E-02	1
700323	2.221E+03	2.439E-02	1
700324	2.221E+03	1.300E-01	1
700325	2.221E+03	8.464E-01	1
700326	2.753E+02	2.491E-02	1
700327	2.221E+03	1.254E-01	1
700328	2.753E+02	8.504E-01	1
700329	2.221E+03	2.443E-02	1
700330	2.221E+03	1.255E-01	1
700331	1.013E+03	8.502E-01	1
700332	2.222E+03	2.460E-02	1
700333	9.330E+02	1.248E-01	1
700334	9.325E+02	8.500E-01	1
700335	4.471E+02	8.427E-02	1
700336	4.471E+02	2.024E-01	1
700337	4.471E+02	7.127E-01	1
700338	2.174E+04	9.376E-03	1
700339	2.174E+04	1.490E-01	1
700340	2.174E+04	8.436E-01	1
700341	2.181E+04	9.797E-03	1
700342	2.181E+04	1.460E-01	1
700343	2.181E+04	8.436E-01	1
700344	2.182E+04	9.011E-03	1
700345	7.680E+03	1.460E-01	1
700346	7.680E+03	8.475E-01	1
700347	1.413E+04	8.678E-03	1
700348	1.413E+04	1.434E-01	1
700349	1.413E+04	8.475E-01	1
700439	4.900E+02	1	1
700441	1.447E+02	1	1
510218	1.100E-01	1	1
510217	5.330E-04	1	1
510409	1.330E-04	1	1
510410	2.710E-02	1	1
900505	2.420E-02	1	1
901951	5.067E-02	1	1
512597	1.640E-01	1	1
512598	7.780E-03	1	1
512599	2.150E-01	1	1
512601	1.030E-03	1	1
512602	7.750E-04	1	1
512603	8.350E-04	1	1
512604	3.070E-03	1	1
512605	5.330E-04	1	1
512606	1.200E-03	1	1

Table 13.11 *Cross sections, filter efficiencies, and k Factors for MC background samples corresponding to Run 2 MC Campaign MC16e*

13.5 Conservative Nonclosure Uncertainties

Cross-checks using most conservative nonclosure uncertainties. The use of these nonclosure uncertainties is unnecessary as the ABCD planes were validated with the data nonclosure uncertainty in the designated nonclosure region and validated each region of the plane, but exist here for comparison.

Using the highest nonclosure uncertainty from data from one of the various validation regions: 9.35%. It is worthwhile to note again that the nonclosure uncertainty used in the analysis validated the region where there was a 9.35% nonclosure uncertainty.

The total data events observed in the signal region is 3703 events. The expected exclusion limit for noncommutative black hole benchmark NCBH1 when incorporating the full statistical uncertainty from the ABCD plane, a luminosity uncertainty of 1.7%, an experimental signal uncertainty of 0.65%, a conservative 5% signal theory uncertainty, and a nonclosure uncertainty of 9.35% is a μ value of 10.11. The cross section of the noncommutative black hole benchmark NCBH1 is 0.022 pb, meaning that an expected exclusion limit can be placed on a cross section of 0.22 pb. The observed exclusion limit is 13.98, meaning that an observed exclusion limit can be placed on a cross section of 0.31 pb. These limits can be seen in Figure 13.2.

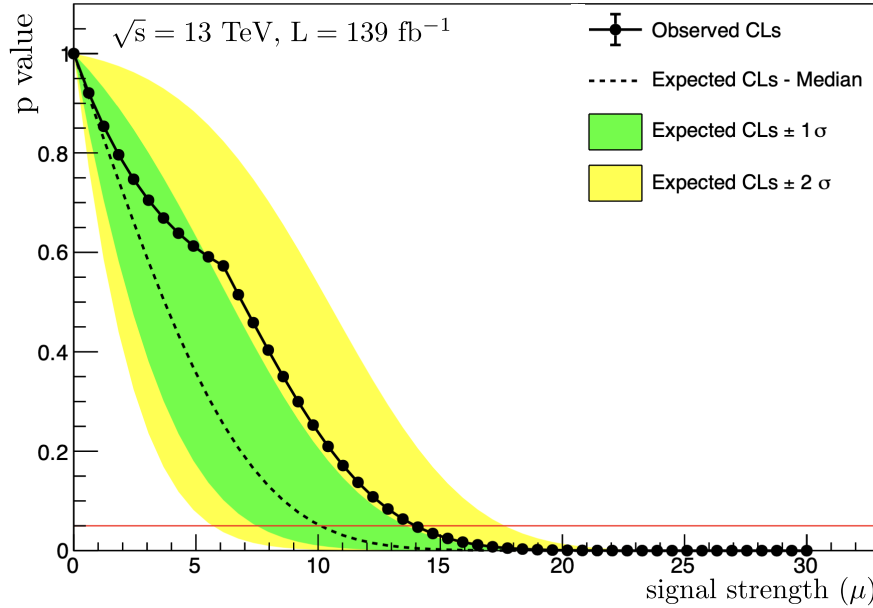


Figure 13.2 *Cross-check rerunning exclusion limits for NCBH Benchmark NCBH1 with nonclosure uncertainty of 9.35%*

This result increases the expected exclusion limit by a factor of 1.54 and increases the observed exclusion limit by a factor of 1.33.

Using the Monte Carlo nonclosure uncertainty of 16.19%:

The total data events observed in the signal region is 3703 events. The expected exclusion limit for noncommutative black hole benchmark NCBH1 when incorporating the full statistical uncertainty from the ABCD plane, a luminosity uncertainty of 1.7%, an experimental signal uncertainty of 0.65%, a conservative 5% signal theory uncertainty, and a nonclosure uncertainty of 16.19% is a μ value of 15.37. The cross section of the noncommutative black hole benchmark NCBH1 is 0.022 pb, meaning that an expected exclusion limit can be placed on a cross section of 0.34 pb. The observed exclusion limit is 18.82, meaning that an observed exclusion limit can be placed on a cross section of 0.41 pb. These limits can be seen in Figure 13.3.

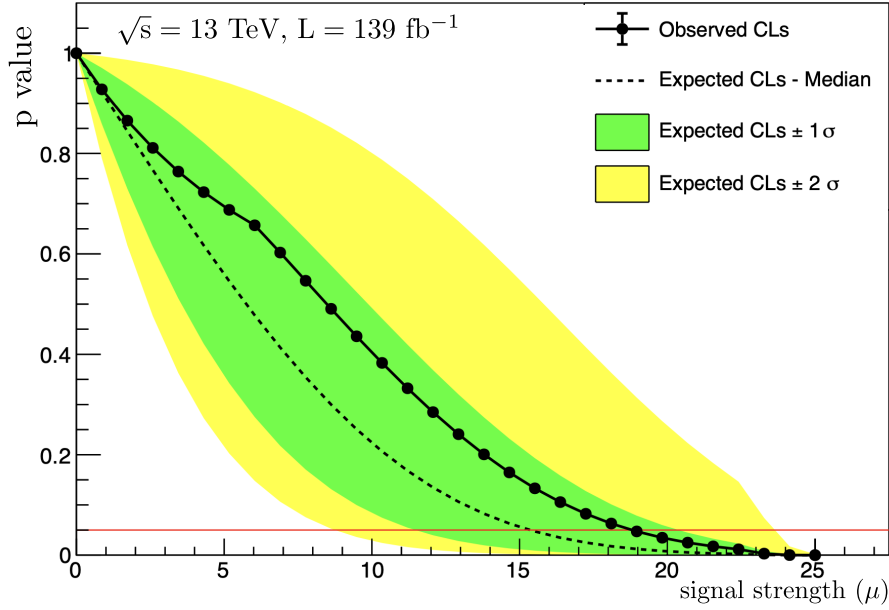


Figure 13.3 *Cross-check rerunning exclusion limits for NCBH Benchmark NCBH1 with nonclosure uncertainty of 16.19%*

This result increases the expected exclusion limit by a factor of 2.34 and increases the observed exclusion limit by a factor of 1.79.

Non-minimal Dark Sectors: Using the highest nonclosure uncertainty from data in one of the validation regions (which the 37% nonclosure uncertainty from the analysis had validated)

The total data events observed in the signal region is 124 events. The expected exclusion limit for NMDS Benchmark NMDS3 when incorporating the full statistical uncertainty from the ABCD plane, a luminosity uncertainty of 1.7%, an experimental signal uncertainty of 1.7%, a conservative theoretical signal uncertainty of 5%, and a nonclosure uncertainty of 52% is a μ value of 103.3. The cross section of the NMDS Benchmark NMDS3 is 1.33

$\times 10^{-4}$ pb, meaning that an expected exclusion limit can be placed on a cross section of 0.0137 pb. The observed exclusion limit is a μ of 107.283, meaning that an observed exclusion limit can be placed on a cross section of 0.0143 pb. These limits can be seen in Figure 13.4.

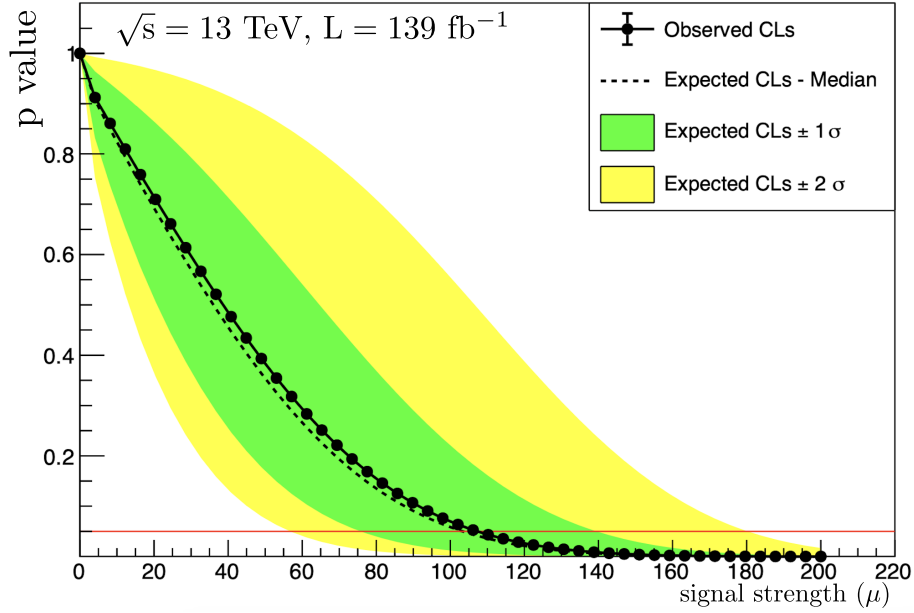


Figure 13.4 *Cross-check rerunning exclusion limits for NMDS Benchmark NMDs3 with nonclosure uncertainty of 52%*

This result increases the expected exclusion limit by a factor of 1.280 and increases the observed exclusion limit by a factor of 1.3.

13.6 Trigger Efficiencies for Signal Monte Carlo

All trigger efficiencies were calculated by the author using code developed by the author to determine the fraction of events in the total Monte Carlo signal sample that pass a given trigger.

Scalar Cascade Benchmark SC4			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.408	HLT_xe90_mht.L1XE50	0.381
HLT_j80_xe80	0.313	HLT_xe110_mht.L1XE50	0.344
HLT_j360	0.0764	HLT_j360_a10_lcw.L1J100	0.095
HLT_ht850.L1J75	0.0302	HLT_j360_a10r.L1J100	0.0904
HLT_4j85	0.00927	HLT_j340	0.085
HLT_5j60	0.00697	HLT_j400_a10_lcw.L1J100	0.0778
HLT_6j45	0.0064	HLT_j400_a10r.L1J100	0.0744
		HLT_420_a10_lcw.L1J100	0.0704
		HLT_j380	0.0692
		HLT_420_a10r.L1J100	0.0677
		HLT_ht850.L1J75	0.0302
		HLT_ht1000.L1J100	0.0195
		HLT_5j55	0.00937
		HLT_4j85	0.00927
mc16d 2017 Trigger	mc16a 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.378	HLT_xe110_pufit_xe65.L1XE50	0.329
HLT_xe100_pufit.L1XE50	0.359	HLT_xe110_pufit_xe70.L1XE50	0.323
HLT_xe100_pufit.L1XE55	0.3535	HLT_xe120_pufit.L1XE50	0.314
HLT_xe110_pufit.L1XE50	0.3373	HLT_j0_perf_ds1.L1J100	0.255
HLT_xe110_pufit.L1XE55	0.334	HLT_j460_a10r.L1J100	0.0586
HLT_j420_a10r.L1J100	0.075	HLT_j460_a10r.L1SC111	0.0593
HLT_j420_a10_lcw_subjes.L1J100	0.0704	HLT_j460_a10_lcw_subjes.L1SC111	0.0569
HLT_j380	0.0695	HLT_j420	0.0556
HLT_j440_a10r.L1J100	0.0677	HLT_j460_a10t_lcw_jes.L1J100	0.0534
HLT_j420_a10_lcw_jes.L1J100	0.0672	HLT_j460_a10t_lcw_jes.L1SC111	0.0534
HLT_j440_a10_lcw_subjes.L1J100	0.0637	HLT_j460_a10t_lcw_jes.L1SC111	0.0534
HLT_j400	0.0635	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.0462
HLT_j460_a10r.L1J100	0.0607	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.0461
HLT_j440_a10t_lcw_jes.L1J100	0.0596	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.046
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.0585	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.0457
HLT_j460_a10_lcw_subjes.L1J100	0.0573	HLT_tau35_medium1_tracktwoEF_xe70.L1XE45	0.0454
HLT_j420	0.05697	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.0447
HLT_tau160_medium1_tracktwo	0.0542	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.0441
HLT_tau160_medium1_tracktwo.L1TAU100	0.0539	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.0441
HLT_j460_a10t_lcw_subjes.L1J100	0.0536		
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.0536		
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.0422		

Table 13.12 *Trigger efficiencies for Scalar Cascade Benchmark SC4 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

Scalar Cascade Benchmark SC2			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.353	HLT_xe90_mht.L1XE50	0.306
HLT_j80_xe80	0.265	HLT_xe110_mht.L1XE50	0.256
HLT_j360	0.0794	HLT_j360_a10_lcw.L1J100	0.105
HLT_ht850.L1J75	0.0518	HLT_j360_a10r.L1J100	0.0995
HLT_4j85	0.0225	HLT_j340	0.0884
HLT_6j45	0.0198	HLT_j400_a10_lcw.L1J100	0.0872
		HLT_j400_a10r.L1J100	0.0822
		HLT_420_a10_lcw.L1J100	0.079
		HLT_420_a10r.L1J100	0.0743
		HLT_j380	0.072
		HLT_ht850.L1J75	0.0518
		HLT_ht1000.L1J100	0.0348
		HLT_5j55	0.0248
mc16d 2017 Trigger	mc16a 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.318	HLT_xe110_pufit_xe65.L1XE50	0.2647
HLT_xe100_pufit.L1XE50	0.295	HLT_xe110_pufit_xe70.L1XE50	0.26
HLT_xe100_pufit.L1XE55	0.0292	HLT_xe120_pufit.L1XE50	0.2505
HLT_xe110_pufit.L1XE50	0.271	HLT_j460_a10r.L1J100	0.0645
HLT_xe110_pufit.L1XE55	0.269	HLT_j460_a10r.L1SC111	0.0645
HLT_j420_a10r.L1J100	0.0811	HLT_j460_a10_lcw_subjes.L1J100	0.0617
HLT_j420_a10_lcw_subjes.L1J100	0.077	HLT_j460_a10_lcw_subjes.L1SC111	0.0617
HLT_440_a10r.L1J100	0.0727	HLT_j420	0.0586
HLT_j380	0.0720	HLT_j460_a10t_lcw_jes.L1J100	0.0585
HLT_420_a10t_lcw_jes.L1J100	0.0714	HLT_j460_a10t_lcw_jes.L1SC111	0.0584
HLT_440_a10_lcw_subjes.L1J100	0.0683	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.0528
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.0669	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.05263
HLT_j460_a10r.L1J100	0.0664	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.0515
HLT_j400	0.0652	HLT_tau35_mediumR1_tracktwoEF_xe70.L1XE45	0.0514
HLT_j440_a10t_lcw_jes.L1J100	0.0648	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.049
HLT_j460_a10_lcw_subjes.L1J100	0.0622	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.0493
HLT_j460_a10t_lcw_jes.L1J100	0.059	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.0489
HLT_j420	0.0585	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.04855
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.0563	HLT_ht1000.L1J100	0.0355
HLT_tau160_medium1_tracktwo	0.0548		
HLT_tau160_medium1_tracktwo.L1TAU100	0.0545		
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.0487		
HLT_ht1000.L1J100	0.0351		

Table 13.13 *Trigger efficiencies for Scalar Cascade Benchmark SC2 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

Scalar Cascade Benchmark SC6			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.439	HLT_xe90_mht.L1XE50	0.417
HLT_j80_xe80	0.341	HLT_xe110_mht.L1XE50	0.382
HLT_j360	0.0872	HLT_j360_a10_lcw.L1J100	0.105
HLT_ht850.L1J75	0.0294	HLT_j360_a10r.L1J100	0.101
HLT_4j85	0.00623	HLT_j340	0.096
HLT_5j60	0.00473	HLT_j400_a10_lcw.L1J100	0.0879
		HLT_j400_a10r.L1J100	0.0842
		HLT_420_a10_lcw.L1J100	0.0804
		HLT_j380	0.0796
		HLT_j420_a10r.L1J100	0.0765
		HLT_ht850.L1J75	0.0294
mc16d 2017 Trigger	mc16a 2017 Trigger Eff.	mc16e 2018	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.407	HLT_xe110_pufit_xe65.L1XE50	0.361
HLT_xe100_pufit.L1XE50	0.389	HLT_xe110_pufit_xe70.L1XE50	0.355
HLT_xe100_pufit.L1XE55	0.382	HLT_xe120_pufit.L1XE50	0.345
HLT_xe110_pufit.L1XE50	0.367	HLT_j460_a10r.L1J100	0.0655
HLT_xe110_pufit.L1XE55	0.364	HLT_j460_a10r.L1SC111	0.0654
HLT_j420_a10r.L1J100	0.0818	HLT_j460_a10_lcw_subjes.L1J100	0.0633
HLT_j420_a10_lcw_subjes.L1J100	0.0783	HLT_j460_a10_lcw_subjes.L1SC111	0.0633
HLT_j380	0.0772	HLT_j420	0.0609
HLT_j440_a10r.L1J100	0.0755	HLT_j460_a10t_lcw_jes.L1J100	0.0595
HLT_420_a10t_lcw_jes.L1J100	0.0743	HLT_j460_a10t_lcw_jes.L1SC111	0.0595
HLT_440_a10_lcw_subjes.L1J100	0.0721	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.0518
HLT_j460_a10r.L1J100	0.0687	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.0515
HLT_j440_a10t_lcw_jes.L1J100	0.0682	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.0510
HLT_j460_a10_lcw_subjes.L1J100	0.066	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.0506
HLT_j420	0.0643	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.0499
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.0637	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.0497
HLT_j460_a10t_lcw_jes.L1J100	0.066	HLT_tau35_mediumR1_tracktwoEF_xe70.L1XE45	0.0453
HLT_tau160_medium1_tracktwo	0.0605	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.0443
		HLT_ht1000.L1J100	0.0182

Table 13.14 *Trigger efficiencies for Scalar Cascade Benchmark SC6 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

Scalar Cascade Benchmark SC5			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.197	HLT_xe90_mht.L1XE50	0.1393
HLT_j80_xe80	0.161	HLT_j360_a10_lcw.L1J100	0.0916
HLT_j360	0.611	HLT_xe110_mht.L1XE50	0.09097
HLT_ht850.L1J75	0.05023	HLT_j360_a10r.L1J100	0.08247
HLT_6j45	0.0243	HLT_j400_a10_lcw.L1J100	0.0711
HLT_4j85	0.0228	HLT_j340	0.0677
HLT_5j60	0.0192	HLT_j400_a10r.L1J100	0.0634
		HLT_420_a10_lcw.L1J100	0.0632
		HLT_j400_a10r.L1J100	0.0569
		HLT_j380	0.0536
		HLT_ht850.L1J75	0.0502
		HLT_ht1000.L1J100	0.0331
		HLT_5j55	0.0277
		HLT_6j45	0.0242
mc16d 2017 Trigger	mc16a 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_puFit.L1XE50	0.185	HLT_xe110_puFit_xe65.L1XE50	0.133
HLT_xe100_puFit.L1XE50	0.161	HLT_xe110_puFit_xe70.L1XE50	0.13
HLT_xe100_puFit.L1XE55	0.159	HLT_xe120_puFit.L1XE50	0.0116
HLT_xe110_puFit.L1XE50	0.139	HLT_j460_a10r.L1SC111	0.0526
HLT_xe110_puFit.L1XE55	0.138	HLT_j460_a10r.L1J100	0.0520
HLT_j420_a10r.L1J100	0.0658	HLT_j460_a10_lcw_subjes.L1SC111	0.052
HLT_j420_a10_lcw_subjes.L1J100	0.0649	HLT_j460_a10_lcw_subjes.L1J100	0.042
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.0606	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.0478
HLT_j440_a10r.L1J100	0.0584	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.0478
HLT_440_a10_lcw_subjes.L1J100	0.0576	HLT_j460_a10t_lcw_jes.L1J100	0.0461
HLT_j380	0.0566	HLT_j460_a10t_lcw_jes.L1SC111	0.0461
HLT_j460_a10r.L1J100	0.0515	HLT_j420	0.0445
HLT_j440_a10t_lcw_jes.L1J100	0.0513	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.0361
HLT_j460_a10_lcw_subjes.L1J100	0.512	HLT_ht1000.L1J100	0.0357
HLT_j440	0.0504	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.0344
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.0462		
HLT_j460_a10t_lcw_jes.L1J100	0.0457		
HLT_j420	0.0444		
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.0409		
HLT_tau160_medium1_tracktwo	0.0402		
HLT_tau160_medium1_tracktwo.L1TAU100	0.0397		
HLT_ht1000.L1J100	0.0354		

Table 13.15 *Trigger efficiencies for Scalar Cascade Benchmark SC5 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

Scalar Cascade Benchmark SC1			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.340	HLT_xe90_mht.L1XE50	0.293
HLT_j80_xe80	0.254	HLT_xe110_mht.L1XE50	0.241
HLT_j360	0.079	HLT_j360_a10_lcw.L1J100	0.106
HLT_ht850.L1J75	0.050	HLT_j360_a10r.L1J100	0.098
HLT_4j85	0.021	HLT_j340	0.088
HLT_6j45	0.020	HLT_j400_a10_lcw.L1J100	0.085
HLT_5j60	0.018	HLT_j400_a10r.L1J100	0.078
		HLT_420_a10_lcw.L1J100	0.077
		HLT_j420_a10r.L1J100	0.071
		HLT_j380	0.071
		HLT_ht850.L1J75	0.050
		HLT_ht1000.L1J100	0.034
		HLT_5j55	0.025
		HLT_6j45	0.021
mc16d 2017 Trigger	mc16a 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_puFit.L1XE50	0.305	HLT_xe110_puFit_xe65.L1XE50	0.252
HLT_xe100_puFit.L1XE50	0.283	HLT_xe110_puFit_xe70.L1XE50	0.247
HLT_xe100_puFit.L1XE55	0.280	HLT_xe120_puFit.L1XE50	0.236
HLT_xe110_puFit.L1XE50	0.260	HLT_j460_a10r.L1J100	0.060
HLT_xe110_puFit.L1XE55	0.259	HLT_j460_a10r.L1SC111	0.060
HLT_j420_a10r.L1J100	0.078	HLT_j460_a10_lcw_subjes.L1J100	0.058
HLT_j420_a10_lcw_subjes.L1J100	0.074	HLT_j460_a10_lcw_subjes.L1SC111	0.058
HLT_420_a10t_lcw_jes.L1J100	0.070	HLT_j420	0.056
HLT_j380	0.070	HLT_j460_a10t_lcw_jes.L1J100	0.055
HLT_j440_a10r.L1J100	0.069	HLT_j460_a10t_lcw_jes.L1SC111	0.055
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.067	HLT_j460_a10t_lcw_jes.L1J100	0.055
HLT_j440_a10_lcw_subjes.L1J100	0.067	HLT_j460_a10t_lcw_jes.L1SC111	0.055
HLT_j400	0.063	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.050
HLT_j460_a10r.L1J100	0.062	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.050
HLT_j440_a10t_lcw_jes.L1J100	0.062	HLT_tau35_mediumR1_tracktwoEF_xe70.L1XE45	0.047
HLT_j420	0.057	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.047
HLT_j460_a10t_lcw_jes.L1J100	0.056	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.045
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.055	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.045
HLT_tau160_medium1_tracktwo	0.054	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.044
HLT_tau160_medium1_tracktwo.L1TAU100	0.054	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.044
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.049	HLT_ht1000.L1J100	0.034
HLT_ht1000.L1J100	0.035	HLT_2j330_a10t_lcw_jes_35smcINF.L1J100	0.023
		HLT_2j330_a10t_lcw_jes_35smcINF.L1SC111	0.023

Table 13.16 *Trigger efficiencies for Scalar Cascade Benchmark SC1 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

Scalar Cascade Benchmark SC3			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.1704	HLT_xe90_mht.L1XE50	0.119
HLT_j80_xe80	0.1308	HLT_xe110_mht.L1XE50	0.0768
HLT_j360	0.0462	HLT_j360_a10_lcw.L1J100	0.0678
HLT_ht850.L1J75	0.0366	HLT_j360_a10r.L1J100	0.061
HLT_4j85	0.0181	HLT_j340	0.0528
HLT_6j45	0.0162	HLT_j400_a10_lcw.L1J100	0.0525
		HLT_j400_a10r.L1J100	0.0478
		HLT_420_a10_lcw.L1J100	0.0465
		HLT_j420_a10r.L1J100	0.0423
		HLT_j380	0.0411
		HLT_ht850.L1J75	0.0366
		HLT_ht1000.L1J100	0.0227
		HLT_5j55	0.0203
		HLT_4j85	0.018
mc16d 2017 Trigger	mc16a 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_puFit.L1XE50	0.153	HLT_xe110_puFit_xe65.L1XE50	0.109
HLT_xe100_puFit.L1XE50	0.133	HLT_xe110_puFit_xe70.L1XE50	0.106
HLT_xe100_puFit.L1XE55	0.131	HLT_xe120_puFit.L1XE50	0.0950
HLT_xe110_puFit.L1XE50	0.113	HLT_j460_a10r.L1J100	0.0372
HLT_xe110_puFit.L1XE55	0.112	HLT_j460_a10r.L1SC111	0.0372
HLT_j420_a10r.L1J100	0.0471	HLT_j460_a10_lcw_subjes.L1J100	0.0354
HLT_j420_a10_lcw_subjes.L1J100	0.0448	HLT_j460_a10_lcw_subjes.L1SC111	0.0354
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.0422	HLT_j420	0.0331
HLT_j440_a10r.L1J100	0.04187	HLT_j460_a10t_lcw_jes.L1J100	0.0328
HLT_j380	0.04183	HLT_j460_a10t_lcw_jes.L1SC111	0.0328
HLT_420_a10t_lcw_jes.L1J100	0.04157	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.0328
HLT_j440_a10_lcw_subjes.L1J100	0.0402	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.0328
HLT_j460_a10r.L1J100	0.0375	HLT_tau35_mediumR1_tracktwoEF_xe70.L1XE45	0.027
HLT_j400	0.0369	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.027
HLT_j440_a10t_lcw_jes.L1J100	0.0368	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.0263
HLT_j460_a10t_lcw_subjes.L1J100	0.0362	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.0263
HLT_j420	0.0327	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.0256
HLT_j460_a10t_lcw_jes.L1J100	0.0327	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.0254
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.0315	HLT_ht1000.L1J100	0.0252
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.0311	HLT_2j330_a10t_lcw_jes_35smcINF.L1J100	0.0223
HLT_tau160_medium1_tracktwo	0.03047	HLT_2j330_a10t_lcw_jes_35smcINF.L1SC111	0.0223
HLT_tau160_medium1_tracktwo.L1TAU100	0.0302		
HLT_ht1000.L1J100	0.0251		
HLT_2j330_a10t_lcw_jes_30smcINF.L1J100	0.0243		

Table 13.17 *Trigger efficiencies for Scalar Cascade Benchmark SC3 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

NCBH1	
mc16a 2015 Trigger	mc16a 2015 Trigger Efficiency
0.938	HLT_xe70_mht
0.912	HLT_j80_xe80
0.828	HLT_j30_xe10_razor170
0.496	HLT_j360
0.339	HLT_ht850_L1J75
0.281	HLT_g200_etcut
0.141	HLT_j225_bloose
0.129	HLT_mu6_2j40_0eta490_inv400_L1MU6_J30.0ETA49_2J20.0ETA49
0.121	HLT_4j85
0.101	HLT_3j175
0.0875	HLT_g120_loose
0.0855	HLT_g40_tight_xe40noL1
0.0735	HLT_5j60
0.0693	HLT_mu4_j40_dr05_3j40_L14J20
0.0651	HLT_mu40
0.063	HLT_tau35_medium1_tracktwo_xe70_L1XE45
0.0579	HLT_j70_bmedium_3j70_L13J25.0ETA23
0.0527	HLT_e24_lhmedium_L1EM20VH
0.0465	HLT_mu20_loose_L1MU15
0.0452	HLT_6j45
0.0445	HLT_e60_lhloose_xe60noL1
0.0408	HLT_tau80_medium1_tracktwo
0.0402	HLT_mu60_0eta105_msonly
0.0398	HLT_e60_lhmedium
0.0391	HLT_tau80_medium1_tracktwo_L1TAU60
mc16a 2016 Trigger	mc16a 2016 Trigger Efficiency
0.930	HLT_xe90_mht_L1XE50
0.917	HLT_xe110_mht_L1XE50
0.828	HLT_j30_xe10_razor170
0.819	HLT_j30_xe10_razor185
0.632	HLT_j360_a10_lcw_L1J100
0.621	HLT_j360_a10r_L1J100
0.565	HLT_j400_a10_lcw_L1J100
0.557	HLT_j400_a10r_L1J100
0.534	HLT_j420_a10_lcw_L1J100
0.533	HLT_j340
0.527	HLT_j420_a10r_L1J100
0.462	HLT_j380
0.339	HLT_ht850_L1J75
0.236	HLT_j80_bmv2c2060_split_xe60_L12J50_XE40
0.234	HLT_ht1000_L1J100
0.166	HLT_j360_bmv2c2085_split
0.137	HLT_j55_bmv2c2060_split_ht500_L14J15
0.133	HLT_j300_bmv2c2077_split
0.127	HLT_g60_loose_xe60noL1
0.122	HLT_e300_etcut
0.121	HLT_g300_etcut
0.121	HLT_4j85
0.118	HLT_g70_loose_xe70noL1
0.117	HLT_j275_bmv2c2070_split
0.113	HLT_j225_bmv2c2060_split

Table 13.18 *Trigger efficiencies for noncommutative black hole benchmark NCBH1 for Run 2 data-taking years: 2015, 2016*

NCBH1	
mc16d 2017 Trigger	mc16d 2017 Trigger Efficiency
0.927	HLT_xe90_pufit_L1XE50
0.920	HLT_xe100_pufit_L1XE50
0.918	HLT_xe100_pufit_L1XE55
0.912	HLT_xe110_pufit_L1XE50
0.911	HLT_xe110_pufit_L1XE55
0.793	HLT_j30_xe60_razor220
0.554	HLT_j420_a10r_L1J100
0.537	HLT_j420_a10_lcw_subjes_L1J100
0.521	HLT_j440_a10r_L1J100
0.506	HLT_j440_a10_lcw_subjes_L1J100
0.503	HLT_j420_a10t_lcw_jes_L1J100
0.492	HLT_j460_a10r_L1J100
0.479	HLT_j225_gsc380_boffperf_split
0.476	HLT_j460_a10_lcw_subjes_L1J100
0.472	HLT_j440_a10t_lcw_jes_L1J100
0.469	HLT_j390_a10t_lcw_jes_30smcINF_L1J100
0.462	HLT_j380
0.443	HLT_j225_gsc400_boffperf_split
0.442	HLT_j460_a10t_lcw_jes_L1J100
0.425	HLT_j400
0.408	HLT_j225_gsc420_boffperf_split
0.400	HLT_j420_a10t_lcw_jes_40smcINF_L1J100
0.393	HLT_j420
0.384	HLT_tau35_medium1_tracktwo_xe70_L1XE45
0.376	HLT_tau160_medium1_tracktwo
0.375	HLT_tau160_medium1_tracktwo_L1TAU100
mc16d 2018 Trigger	mc16a 2018 Trigger Efficiency
0.914	HLT_xe110_pufit_xe65_L1XE50
0.913	HLT_xe110_pufit_xe70_L1XE50
0.905	HLT_xe120_pufit_L1XE50
0.890	HLT_j0_perf_ds1_L1J100
0.491	HLT_j460_a10r_L1SC111
0.491	HLT_j460_a10r_L1J100
0.473	HLT_j460_a10_lcw_subjes_L1J100
0.473	HLT_j460_a10_lcw_subjes_L1SC111
0.443	HLT_j460_a10t_lcw_jes_L1SC111
0.443	HLT_j460_a10t_lcw_jes_L1J100
0.417	HLT_j420_a10t_lcw_jes_35smcINF_L1SC111
0.416	HLT_j420_a10t_lcw_jes_35smcINF_L1J100
0.412	HLT_j225_gsc420_boffperf_split
0.397	HLT_j420
0.358	HLT_tau35_mediumRNN_tracktwoMVA_xe70_L1XE45
0.355	HLT_tau35_medium1_tracktwoEF_xe70_L1XE45
0.341	HLT_tau160_mediumRNN_tracktwoMVA_L1TAU100
0.339	HLT_tau160_medium1_tracktwoEF_L1TAU100
0.337	HLT_tau200_mediumRNN_tracktwoMVA_L1TAU100
0.335	HLT_tau200_medium1_tracktwoEF_L1TAU100
0.237	HLT_ht1000_L1J100
0.209	HLT_mu4_j90_xe90_pufit_2dphi10_L1MU4_XE60
0.186	HLT_mu4_j90_xe90_pufit_2dphi10_L1MU4_J50_XE50_DPHI-J20s2XE30
0.171	HLT_j85_gsc100_bmv2c1050_split_xe85_pufit_xe50_L1XE55
0.140	HLT_j225_gsc360_bhm2c1077_split
0.124	HLT_j45_gsc55_bmv2c1050_split_ht700_L1HT190-J15s5.ETA21

Table 13.19 *Trigger efficiencies for noncommutative black hole benchmark NCBH1 for Run 2 data-taking years: 2017, 2018*

NCBH4	
mc16a 2015 Trigger	mc16a 2015 Trigger Efficiency
0.818	HLT_xe70_mht
0.746	HLT_j80_xe80
0.626	HLT_j30_xe10_razor170
0.288	HLT_j360
0.176	HLT_g200_etcut
0.170	HLT_ht850_L1J75
0.110	HLT_g40_tight_xe40noL1
0.0817	HLT_g120_loose
0.0791	HLT_mu6_2j40_0eta490_invm400_L1MU6_J30.0ETA49_2J20.0ETA49
0.0786	HLT_tau35_medium1_tracktwo_xe70_L1XE45
0.0674	HLT_e24_lhmedium_L1EM20VH
0.0672	HLT_j225_bloose
0.0658	HLT_mu40
0.0624	HLT_mu20_iloose_L1MU15
0.0618	HLT_4j85
0.0567	HLT_tau80_medium1_tracktwo
0.054	HLT_tau80_medium1_tracktwo_L1TAU60
0.053	HLT_e60_lhloose_xe60noL1
0.0502	HLT_e60_lhmedium
0.0467	HLT_5j60
0.0426	HLT_mu60_0eta105_msonly
0.0421	HLT_mu4_j40_dr05_3j40_L14J20
0.0405	HLT_3j175
0.0373	HLT_e120_lhloose
0.032	HLT_6j45
mc16a 2016 Trigger	mc16a 2016 Trigger Efficiency
0.801	HLT_xe90_mht_L1XE50
0.772	HLT_xe110_mht_L1XE50
0.626	HLT_j30_xe10_razor170
0.612	HLT_j30_xe10_razor185
0.365	HLT_j360_a10_lcw_L1J100
0.361	HLT_j360_a10r_L1J100
0.318	HLT_j400_a10_lcw_L1J100
0.314	HLT_j400_a10r_L1J100
0.312	HLT_j340
0.297	HLT_j420_a10_lcw_L1J100
0.293	HLT_j420_a10r_L1J100
0.266	HLT_j380
0.170	HLT_ht850_L1J75
0.125	HLT_g60_loose_xe60noL1
0.115	HLT_ht1000_L1J100
0.114	HLT_g70_loose_xe70noL1
0.113	HLT_j80_bmv2c2060_split_xe60_L12J50_XE40
0.103	HLT_g45_tight_L1EM22VH1_xe45noL1
0.0824	HLT_e300_etcut
0.0823	HLT_g300_etcut
0.0819	HLT_j360_bmv2c2085_split
0.0786	HLT_tau35_medium1_tracktwo_xe70_L1XE45
0.0724	HLT_j55_bmv2c2060_split_ht500_L14J15
0.070	HLT_g140_loose
0.066	HLT_mu40

Table 13.20 *Trigger efficiencies for noncommutative black hole benchmark NCBH4 for all Run 2 data-taking years: 2015, 2016*

NCBH4	
mc16d 2017 Trigger	mc16d 2017 Trigger Efficiency
0.797	HLT_xe90_pufit_L1XE50
0.780	HLT_xe100_pufit_L1XE50
0.774	HLT_xe100_pufit_L1XE55
0.760	HLT_xe110_pufit_L1XE50
0.755	HLT_xe110_pufit_L1XE55
0.577	HLT_j30_xe60_razor220
0.310	HLT_j420_a10r_L1J100
0.300	HLT_j420_a10r_L1J100
0.291	HLT_j440_a10r_L1J100
0.281	HLT_j440_a10_lcw_subjes_L1J100
0.280	HLT_j440_a10_lcw_subjes_L1J100
0.276	HLT_j225_gsc380_boffperf_split
0.272	HLT_j460_a10r_L1J100
0.268	HLT_j380
0.261	HLT_j460_a10_lcw_subjes_L1J100
0.260	HLT_j440_a10t_lcw_jes_L1J100
0.256	HLT_j225_gsc400_boffperf_split
0.253	HLT_tau35_medium1_tracktwo_xe70_L1XE45
0.252	HLT_tau35_medium1_tracktwo_xe70_L1XE45
0.248	HLT_j400
0.241	HLT_j460_a10t_lcw_jes_L1J100
0.238	HLT_j225_gsc420_boffperf_split
0.233	HLT_tau160_medium1_tracktwo
0.232	HLT_tau160_medium1_tracktwo_L1TAU100
0.230	HLT_j420
0.204	HLT_j420_a10t_lcw_jes_40smcINF_L1J100
0.160	HLT_j420_a10t_lcw_jes_40smcINF_L1J100
0.138	HLT_j420_a10t_lcw_jes_40smcINF_L1J100
0.117	HLT_j420_a10t_lcw_jes_40smcINF_L1J100
0.180	HLT_j420_a10t_lcw_jes_40smcINF_L1J100
mc16e 2018 Trigger	mc16e 2018 Trigger Efficiency
0.746	HLT_xe110_pufit_xe65_L1XE50
0.740	HLT_xe110_pufit_xe70_L1XE50
0.723	HLT_xe120_pufit_L1XE50
0.643	HLT_j0_perf_ds1_L1J100
0.273	HLT_j460_a10r_L1SC111
0.273	HLT_j460_a10r_L1J100
0.260	HLT_j460_a10_lcw_subjes_L1SC111
0.260	HLT_j460_a10_lcw_subjes_L1J100
0.2417	HLT_j460_a10t_lcw_jes_L1SC111
0.242	HLT_j460_a10t_lcw_jes_L1J100
0.238	HLT_tau35_mediumRNN_tracktwoMVA_xe70_L1XE45
0.235	HLT_tau35_medium1_tracktwoEF_xe70_L1XE45
0.235	HLT_j225_gsc420_boffperf_split
0.226	HLT_j420
0.216	HLT_j420_a10t_lcw_jes_35smcINF_L1SC111
0.216	HLT_j420_a10t_lcw_jes_35smcINF_L1J100
0.209	HLT_tau160_mediumRNN_tracktwoMVA_L1TAU100
0.208	HLT_tau160_medium1_tracktwoEF_L1TAU100
0.205	HLT_tau200_mediumRNN_tracktwoMVA_L1TAU100
0.203	HLT_tau200_medium1_tracktwoEF_L1TAU100
0.154	HLT_mu4_j90_xe90_pufit_2dphi10_L1MU4_XE60
0.135	HLT_mu4_j90_xe90_pufit_2dphi10_L1MU4_J50_XE50_DPFI-J20s2XE30
0.118	HLT_ht1000_L1J100
0.0983	HLT_g80_loose_xe80noL1
0.0909	HLT_e5_lhloose_nod0_j40_xe70_pufit_2dphi10_L1XE60
0.0861	HLT_j85_gsc100_bmv2c1050_split_xe85_pufit_xe50_L1XE55
0.0815	HLT_e300_etcut
0.0812	HLT_g300_etcut
0.0806	HLT_e5_lhloose_nod0_j50_xe70_pufit_2dphi10_L1J40_XE50_DPFI-J20s2XE50
0.0744	HLT_j225_gsc360_bhmv2c1077_split

Table 13.21 *Trigger efficiencies for noncommutative black hole benchmark NCBH4 for all Run 2 data-taking years: 2017, 2018*

NMDS1			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16a 2016 Trigger Eff.
HLT_xe70_mht	0.736	HLT_xe90_mht.L1XE50	0.696
HLT_j80_xe80	0.574	HLT_xe110_mht.L1XE50	0.633
HLT_j360	0.085	HLT_j360_a10_lcw.L1J100	0.125
HLT_ht850.L1J75	0.063	HLT_j360_a10r.L1J100	0.120
HLT_4j85	0.066	HLT_j340	0.097
HLT_5j60	0.053	HLT_j400_a10_lcw.L1J100	0.096
HLT_6j45	0.052	HLT_j400_a10r.L1J100	0.093
		HLT_j420_a10_lcw.L1J100	0.086
		HLT_j420_a10r.L1J100	0.083
		HLT_j380	0.075
		HLT_5j55	0.074
		HLT_850.L1J75	0.063
		HLT_4j85	0.055
		HLT_6j45	0.052
		HLT_6j45_0eta240	0.041
		HLT_ht1000.L1J100	0.036
mc16d 2017 Trigger	mc16d 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.676	HLT_xe110_pufit_xe65.L1XE50	0.596
HLT_xe100_pufit.L1XE50	0.641	HLT_xe110_pufit_xe70.L1XE50	0.587
HLT_xe100_pufit.L1XE55	0.630	HLT_xe120_pufit.L1XE50	0.563
HLT_xe110_pufit.L1XE50	0.604	HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.085
HLT_xe110_pufit.L1XE55	0.596	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.083
HLT_j420_a10r.L1J100	0.094	HLT_j460_a10r.L1J100	0.712
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.090	HLT_j460_a10r.L1SC111	0.072
HLT_j420_a10_lcw_subjes.L1J100	0.087	HLT_j460_a10_lcw_subjes.L1J100	0.068
HLT_j440_a10r.L1J100	0.080	HLT_j460_a10_lcw_subjes.L1SC111	0.068
HLT_j420_a10t_lcw_subjes.L1J100	0.081	HLT_j460_a10t_lcw_jes.L1J100	0.063
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.080	HLT_j460_a10t_lcw_jes.L1SC111	0.063
HLT_j380	0.078	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.062
HLT_j440_a10t_lcw_jes.L1J100	0.078	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.062
HLT_j460_a10r.L1J100	0.074	HLT_j420	0.060
HLT_j440_a10_lcw_subjes.L1J100	0.072	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.050
HLT_j460_a10_lcw_subjes.L1J100	0.070	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.049
HLT_j420	0.062	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.048
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.060	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.048
HLT_tau160_medium1_tracktwo	0.059	HLT_ht1000.L1J100	0.040
HLT_tau160_medium1_tracktwo.L1TAU100	0.059		
HLT_6j45_0eta240	0.048		
HLT_6j45_0eta240.L1J20	0.041		

Table 13.22 *Trigger efficiencies for NMDS benchmark NMDS1 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

NMDS2			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.897	HLT_xe90_mht_L1XE50	0.873
HLT_j80_xe80	0.083	HLT_xe110_mht_L1XE50	0.836
HLT_j360	0.021	HLT_j360_a10_lcw_L1J100	0.342
HLT_ht850_L1J75	0.200	HLT_j360_a10r_L1J100	0.336
HLT_4j85	0.190	HLT_j400_a10_lcw_L1J100	0.265
HLT_5j60	0.194	HLT_j400_a10r_L1J100	0.257
HLT_6j45	0.154	HLT_j340	0.242
		HLT_j420_a10_lcw_L1J100	0.231
		HLT_j420_a10r_L1J100	0.226
		HLT_5j55	0.223
		HLT_ht850_L1J75	0.200
		HLT_4j85	0.190
		HLT_j380	0.179
		HLT_6j45	0.154
		HLT_ht1000_L1J100	0.118
mc16d 2017 Trigger	mc16d 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit_L1XE50	0.087	HLT_xe110_pufit_xe65_L1XE50	0.821
HLT_xe100_pufit_L1XE50	0.085	HLT_xe110_pufit_xe70_L1XE50	0.817
HLT_xe100_pufit_L1XE55	0.084	HLT_xe120_pufit_L1XE50	0.801
HLT_xe110_pufit_L1XE50	0.824	HLT_j460_a10r_L1J100	0.192
HLT_xe110_pufit_L1XE55	0.821	HLT_j460_a10r_L1SC111	0.192
HLT_j420_a10r_L1J100	0.245	HLT_j460_a10_lcw_subjes_L1J100	0.175
HLT_j420_a10_lcw_subjes_L1J100	0.225	HLT_j460_a10_lcw_subjes_L1SC111	0.175
HLT_j440_a10r_L1J100	0.217	HLT_j420_a10t_lcw_jes_35smcINF_L1J100	0.170
HLT_j390_a10t_lcw_jes_30smcINF_L1J100	0.216	HLT_j420_a10t_lcw_jes_35smcINF_L1SC111	0.170
HLT_j420_a10t_lcw_jes_L1J100	0.201	HLT_j460_a10t_lcw_jes_L1J100	0.156
HLT_j440_a10_lcw_subjes_L1J100	0.198	HLT_j460_a10t_lcw_jes_L1SC111	0.156
HLT_j460_a10r_L1J100	0.189	HLT_tau35_medium1_tracktwo_xe70_L1XE45	0.141
HLT_j380	0.178	HLT_tau35_mediumRNN_tracktwoMVA_xe70_L1XE45	0.140
HLT_j440_a10t_lcw_jes_L1J100	0.175	HLT_j420	0.133
HLT_j460_a10_lcw_subjes_L1J100	0.171	HLT_ht1000_L1J100	0.123
HLT_j420_a10t_lcw_jes_40smcINF_L1J100	0.162	HLT_tau160_mediumRNN_tracktwoMVA_L1TAU100	0.106
HLT_tau35_medium1_tracktwo_xe70_L1XE45	0.157	HLT_tau200_mediumRNN_tracktwoMVA_L1TAU100	0.104
HLT_tau160_medium1_tracktwo	0.123	HLT_tau160_medium1_tracktwoEF_L1TAU100	0.102
		HLT_tau200_medium1_tracktwoEF_L1TAU100	0.100

Table 13.23 *Trigger efficiencies for NMDS benchmark NMDS2 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

NMDS5			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.595	HLT_xe90_mht.L1XE50	0.557
HLT_j80_xe80	0.421	HLT_xe110_mht.L1XE50	0.498
HLT_j360	0.066	HLT_j360_a10_lcw.L1J100	0.090
HLT_ht850.L1J75	0.034	HLT_j360_a10r.L1J100	0.086
HLT_4j85	0.023	HLT_j340	0.074
HLT_5j60	0.021	HLT_j400_a10_lcw.L1J100	0.072
HLT_6j45	0.017	HLT_j400_a10r.L1J100	0.068
		HLT_j420_a10_lcw.L1J100	0.065
		HLT_j420_a10r.L1J100	0.606
		HLT_j380	0.058
		HLT_ht850.L1J75	0.034
		HLT_5j55	0.029
		HLT_4j85	0.023
		HLT_ht1000.L1J100	0.019
		HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.017
		HLT_6j45	0.017
mc16d 2017 Trigger	mc16d 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.538	HLT_xe110_pufit_xe65.L1XE50	0.458
HLT_xe100_pufit.L1XE50	0.504	HLT_xe110_pufit_xe70.L1XE50	0.449
HLT_xe100_pufit.L1XE55	0.493	HLT_xe120_pufit.L1XE50	0.043
HLT_xe110_pufit.L1XE50	0.463	HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.062
HLT_xe110_pufit.L1XE55	0.457	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.060
HLT_j420_a10r.L1J100	0.070	HLT_j460_a10r.L1SC111	0.057
HLT_j420_a10_lcw_subjes.L1J100	0.067	HLT_j460_a10r.L1J100	0.057
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.065	HLT_j460_a10_lcw_subjes.L1J100	0.054
HLT_j440_a10r.L1J100	0.063	HLT_j460_a10_lcw_subjes.L1SC111	0.054
HLT_j420_a10t_lcw_jes.L1J100	0.063	HLT_j460_a10t_lcw_subjes.L1J100	0.050
HLT_j380	0.062	HLT_j460_a10t_lcw_subjes.L1SC111	0.050
HLT_j440_a10_lcw_subjes.L1J100	0.602	HLT_j420	0.049
HLT_j390_a10t_lcw_subjes.L1J100	0.059	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.046
HLT_j460_a10r.L1J100	0.057	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.046
HLT_j420	0.051	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.041
HLT_j460_a10t_lcw_jes.L1j100	0.054	HLT_tau160_medium1_tracktwoEF.L1TAU100	0.041
HLT_tau160_medium1_tracktwo	0.047	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.041
HLT_tau160_medium1_tracktwo.L1TAU100	0.047	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.040
HLT_j420_a10t_lcw_jes_40smcINF.L1J100	0.042	HLT_ht1000.L1J100	0.022

Table 13.24 *Trigger efficiencies for non-minimal dark sector benchmark NMDS5 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

NMDS4			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16a 2016 Trigger Eff.
HLT_xe70_mht	0.094	HLT_xe90_mht_L1XE50	0.922
HLT_j80_xe80	0.903	HLT_xe110_mht_L1XE50	0.090
HLT_ht850_L1J75	0.507	HLT_j360_a10_lcw_L1J100	0.616
HLT_4j85	0.424	HLT_j360_a10r_L1J100	0.608
HLT_j360	0.420	HLT_j400_a10_lcw_L1J100	0.519
HLT_5j60	0.404	HLT_j400_a10r_L1J100	0.514
HLT_6j45	0.359	HLT_ht850_L1J75	0.507
HLT_3j175	0.200	HLT_j420_a10_lcw_L1J100	0.473
HLT_tau35_medium1_tracktwo_xe70_L1XE45	0.050	HLT_j340	0.468
		HLT_5j55	0.463
		HLT_4j85	0.424
		HLT_j380	0.374
		HLT_6j45	0.359
		HLT_h1000_L1J100	0.342
mc16d 2017 Trigger	mc16d 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit_L1XE50	0.919	HLT_xe110_pufit_xe65_L1XE50	0.888
HLT_xe100_pufit_L1XE50	0.906	HLT_xe110_pufit_xe70_L1XE50	0.884
HLT_xe100_pufit_L1XE55	0.904	HLT_xe120_pufit_L1XE50	0.876
HLT_xe110_pufit_L1XE50	0.892	HLT_j460_a10r_L1SC111	0.414
HLT_xe110_pufit_L1XE55	0.891	HLT_j460_a10r_L1J100	0.414
HLT_j420_a10r_L1J100	0.505	HLT_j460_a10_lcw_subjes_L1J100	0.390
HLT_j420_a10_lcw_subjes_L1J100	0.476	HLT_j460_a10_lcw_subjes_L1SC111	0.390
HLT_j440_a10r_L1J100	0.463	HLT_j420_a10t_lcw_jes_35smcINF_L1J100	0.382
HLT_j390_a10t_lcw_jes_30smcINF_L1J100	0.454	HLT_j460_a10t_lcw_jes_L1J100	0.358
HLT_j420_a10t_lcw_jes_L1J100	0.442	HLT_j460_a10t_lcw_jes_L1SC111	0.358
HLT_j460_a10r_L1J100	0.420	HLT_ht1000_L1J100	0.351
HLT_j440_a10t_lcw_jes_L1J100	0.400	HLT_j420	0.291
HLT_460_a10_lcw_subjes_L1J100	0.394		
HLT_j380	0.374		
HLT_j460_a10t_lcw_jes_L1J100	0.394		
HLT_tau160_medium1_tracktwo	0.271		

Table 13.25 *Trigger efficiencies for non-minimal dark sector benchmark NMDS4 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

NMDS6			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe70_mht	0.476	HLT_xe90_mht.L1XE50	0.447
HLT_j80_xe80	0.337	HLT_xe110_mht.L1XE50	0.403
HLT_j360	0.058	HLT_j360_a10_lcw.L1J100	0.077
HLT_ht850.L1J75	0.022	HLT_j360_a10r.L1J100	0.011
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.011	HLT_j400_a10_lcw.L1J100	0.062
HLT_4j85	0.011	HLT_j400_a10r.L1J100	0.057
HLT_5j60	0.010	HLT_j420_a10_lcw.L1J100	0.056
HLT_6j45	0.008	HLT_j420_a10r.L1J100	0.052
		HLT_j380	0.051
		HLT_jt850.L1J75	0.023
mc16d 2017 Trigger	mc16d 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.428	HLT_xe110_pufit_xe65.L1XE50	0.366
HLT_xe100_pufit.L1XE50	0.401	HLT_xe110_pufit_xe70.L1XE50	0.358
HLT_xe100_pufit.L1XE55	0.391	HLT_xe120_pufit.L1XE50	0.342
HLT_xe110_pufit.L1XE50	0.370	HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.044
HLT_xe110_pufit.L1XE55	0.364		
HLT_j420_a10r.L1J100	0.055	HLT_j460_a10r.L1J100	0.044
HLT_j420_a10_lcw_subjes.L1J100	0.053	HLT_j460_a10r.L1SC111	0.044
HLT_j380	0.049	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.043
HLT_j440_a10r.L1J100	0.049	HLT_j460_a10t_lcw_subjes.L1J100	0.042
HLT_j420_a10_lcw_subjes.L1J100	0.049	HLT_j460_a10t_lcw_subjes.L1SC111	0.042
HLT_tau35_medium1_tracktwo_xe70.L1XE45	0.047	HLT_j420	0.039
HLT_j440_a10_lcw_subjes.L1J100	0.046	HLT_j460_a10t_lcw_jes.L1J100	0.039
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.045	HLT_j460_a10t_lcw_jes.L1SC111	0.039
HLT_j400	0.044	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.035
HLT_j460_a10r.L1J100	0.043	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.035
HLT_j440_a10_lcw_subjes.L1J100	0.043	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.033
HLT_460_a10_lcw_subjes.L1J100	0.041	HLT_tau160_mediumRNN_tracktwoEF.L1TAU100	0.032
HLT_j420	0.038	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.032
HLT_tau160_medium1_tracktwo	0.037	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.032

Table 13.26 *Trigger efficiencies for non-minimal dark sector benchmark NMDS6 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

NMDS3			
mc16a 2015 Trigger	mc16a 2015 Trigger Eff.	mc16a 2016 Trigger	mc16d 2016 Trigger Eff.
HLT_xe90_mht.L1XE50	0.974	HLT_xe90_mht.L1XE50	0.974
HLT_j360_a10_lcw.L1J100	0.970	HLT_j360_a10_lcw.L1J100	0.969
HLT_j360_a10r.L1J100	0.968	HLT_j360_a10r.L1J100	0.968
HLT_xe110_mht.L1XE50	0.966	HLT_xe110_mht.L1XE50	0.967
HLT_ht850.L1J75	0.960	HLT_ht850.L1J75	0.960
HLT_j400_a10_lcw.L1J100	0.951	HLT_j400_a10_lcw.L1J100	0.951
HLT_j400_a10r.L1J100	0.949	HLT_j400_a10r.L1J100	0.949
HLT_j30_xe10_razor170	0.948	HLT_j30_xe10_razor170	0.948
HLT_j30_xe10_razor185	0.942	HLT_j30_xe10_razor185	0.942
HLT_j420_a10_lcw.L1J100	0.939	HLT_j420_a10_lcw.L1J100	0.939
HLT_j340	0.937	HLT_j340	0.937
HLT_j420_a10r.L1J100	0.937167	HLT_j420_a10r.L1J100	0.937
HLT_ht1000.L1J100	0.926125	HLT_ht1000.L1J100	0.926
HLT_j380	0.898	HLT_j380	0.898
HLT_4j85	0.837	HLT_4j85	0.837
mc16d 2017 Trigger	mc16d 2017 Trigger Eff.	mc16e 2018 Trigger	mc16e 2018 Trigger Eff.
HLT_xe90_pufit.L1XE50	0.974	HLT_xe110_pufit_xe65.L1XE50	0.963
HLT_xe100_pufit.L1XE50	0.970	HLT_xe110_pufit_xe70.L1XE50	0.962
HLT_xe100_pufit.L1XE55	0.969	HLT_xe120_pufit.L1XE50	0.959
HLT_xe110_pufit.L1XE50	0.965	HLT_h1000.L1J100	0.933
HLT_xe110_pufit.L1XE55	0.965	HLT_j460_a10r.L1J100	0.092
HLT_j420_a10r.L1J100	0.947	HLT_j460_a10r.L1SC111	0.092
HLT_j420_a10_lcw_subjes.L1J100	0.940	HLT_j460_a10t_lcw_subjes.L1J100	0.913
HLT_j440_a10_lcw_subjes.L1J100	0.926	HLT_j460_a10t_lcw_subjes.L1SC111	0.913
HLT_h1000.L1J100	0.933	HLT_j420_a10t_lcw_jes_35smcINF.L1J100	0.900
HLT_j420_a10t_lcw_jes.L1J100	0.932	HLT_j420_a10t_lcw_jes_35smcINF.L1SC111	0.900
HLT_j440_a10_lcw_subjes.L1J100	0.926	HLT_j420	0.850
HLT_j460_a10r.L1J100	0.919	HLT_tau160_mediumRNN_tracktwoMVA.L1TAU100	0.774
HLT_j440_a10_lcw_subjes.L1J100	0.916	HLT_tau200_mediumRNN_tracktwoMVA.L1TAU100	0.774
HLT_460_a10_lcw_subjes.L1J100	0.910	HLT_tau160_mediumRNN_tracktwoEF.L1TAU100	0.769
HLT_j380	0.899	HLT_tau200_medium1_tracktwoEF.L1TAU100	0.769
HLT_j460_a10t_lcw_jes.L1J100	0.898	HLT_tau35_mediumRNN_tracktwoMVA_xe70.L1XE45	0.767
HLT_j390_a10t_lcw_jes_30smcINF.L1J100	0.894	HLT_tau35_medium1_tracktwoEF_xe70.L1XE45	0.764
HLT_tau160_medium1_tracktwo	0.821		

Table 13.27 *Trigger efficiencies for non-minimal dark sector benchmark NMDS3 for all Run 2 data-taking years: 2015, 2016, 2017, 2018*

Bibliography

- [1] G. Aad, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012).
- [2] CMS, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) 30–61,
<http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- [3] R. T. D’Agnolo and M. Low, *Disorder and mimesis at hadron colliders*, Journal of High Energy Physics **2019** (2019),
<https://doi.org/10.1007%2Fjhep08%282019%29163>.
- [4] ATLAS, *Search for new resonances in mass distributions of jet pairs using 139 fb1 of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Journal of High Energy Physics **2020** (2020), [http://dx.doi.org/10.1007/JHEP03\(2020\)145](http://dx.doi.org/10.1007/JHEP03(2020)145).
- [5] G. Albouy et al., *Theory, phenomenology, and experimental avenues for dark showers: a Snowmass 2021 report*, The European Physical Journal C **82** (2022),
<http://dx.doi.org/10.1140/epjc/s10052-022-11048-8>.
- [6] T. Han, Z. Si, K. M. Zurek, and M. J. Strassler, *Phenomenology of hidden valleys at hadron colliders*, JHEP **07** (2008) 008, [arXiv:0712.2041](https://arxiv.org/abs/0712.2041) [hep-ph].
- [7] C. ATLAS, *Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions with $\sqrt{s} = 13$ TeV with the ATLAS detector*, Physical Review D **103** (2021),
<https://doi.org/10.1103%2Fphysrevd.103.112006>.
- [8] C. D0, *Search for Large Extra Dimensions via Single Photon plus Missing Energy Final States*, Physical Review Letters **101** (2008),
<http://dx.doi.org/10.1103/PhysRevLett.101.011601>.
- [9] C. ATLAS, *Search for new phenomena in final states with large jet multiplicities and missing transverse momentum using $\sqrt{s} = 13$ TeV proton-proton collisions recorded by ATLAS in Run 2 of the LHC*, Journal of High Energy Physics **2020** (2020), <https://doi.org/10.1007%2Fjhep10%282020%29062>.
- [10] B. Hobbs, *The Standard Model of particle physics is brilliant and completely flawed*, <https://www.abc.net.au/news/science/2017-07-15/the-standard-model-of-particle-physics-explained/7670338>, 2017.
Accessed: 2024-05-8.

- [11] J. E. Augustin, *Discovery of a Narrow Resonance in $e^+ - e^-$ Annihilation*, Phys.Rev.Lett **33** (1974).
- [12] J. J. Aubert, *Experimental Observation of a Heavy Particle J* , Phys.Rev.Lett **33** (1974).
- [13] Kodama, *Observation of tau neutrino interactions*, Phys. Lett. B **504** (2001).
- [14] S. W. Herb, *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV ProtonNucleus Collisions*, Phys.Rev.Lett **39** (1977).
- [15] G. Arnison, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, Phys. Lett. B **126** (1983).
- [16] P. Bagnaia, *Evidence for $Z_0 \rightarrow e^+ e^-$ at the CERN pp Collider*, Phys. Lett. B **129** (1983).
- [17] F. Abe, *Observation of top quark production in pp collisions*, Phys.Rev.Lett **74** (1995).
- [18] S. Abachi, *Observation of top quark production in pp collisions*, Phys.Rev.Lett **74** (1995).
- [19] Y. Ashie, *Evidence for an oscillatory signature in atmospheric neutrino oscillation*, Phys. Rev. Lett. **93** (2004).
- [20] K. G. Begeman, *Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics*, Mon. Not. Roy. Astron. Soc **249** (1991).
- [21] T. G. Rizzo, *Noncommutative inspired black holes in extra dimensions*, Journal of High Energy Physics **2006** (2006) 021–021,
<https://doi.org/10.1088%2F1126-6708%2F2006%2F09%2F021>.
- [22] A. Patil, *Neural Networks / Artificial Neural Networks*,
<https://pdg.lbl.gov/2019/reviews/rpp2019-rev-extra-dimensions.pdf>,
2020. Accessed: 2023-07-7.
- [23] R. Giambò, *Gravitational collapse of a spherical scalar field*, 2023.
- [24] R. Casadio, O. Micu, and F. Scardigli, *Quantum hoop conjecture: Black hole formation by particle collisions*, Physics Letters B **732** (2014) 105–109,
<https://doi.org/10.1016%2Fj.physletb.2014.03.037>.
- [25] G. Dvali and S. Sibiryakov, *Creating semiclassical black holes in collider experiments and keeping them on a string*, JHEP **03** (2008) 007, arXiv:0801.3808,
<https://cds.cern.ch/record/1082579>.
- [26] D. M. Gingrich and K. Martell, *Study of highly excited string states at the Large Hadron Collider*, Physical Review D **78** (2008),
<http://dx.doi.org/10.1103/PhysRevD.78.115009>.
- [27] D. M. Gingrich, *Quantum black holes with charge, color and spin at the LHC*, Journal of Physics G: Nuclear and Particle Physics **37** (2010) 105008,
<http://dx.doi.org/10.1088/0954-3899/37/10/105008>.

- [28] *Search for strong gravity in multijet final states produced in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Journal of High Energy Physics **2016** (2016), [http://dx.doi.org/10.1007/JHEP03\(2016\)026](http://dx.doi.org/10.1007/JHEP03(2016)026).
- [29] CMS, *Search for black holes and other new phenomena in high-multiplicity final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Physics Letters B **774** (2017) 279–307, <http://dx.doi.org/10.1016/j.physletb.2017.09.053>.
- [30] ATLAS, *Search for TeV-scale gravity signatures in high-mass final states with leptons and jets with the ATLAS detector at $\sqrt{s} = 13$ TeV*, Physics Letters B **760** (2016) 520–537, <http://dx.doi.org/10.1016/j.physletb.2016.07.030>.
- [31] ATLAS Collaboration Collaboration,, *Search for quantum black hole production in lepton + jet final states using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **109** (2024) 032010, <https://link.aps.org/doi/10.1103/PhysRevD.109.032010>.
- [32] ATLAS Collaboration, G. Aad et al., *Search for lepton-flavour violation in high-mass dilepton final states using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **23** (2020) 082, [arXiv:2307.08567](https://arxiv.org/abs/2307.08567) [hep-ex].
- [33] ATLAS, *Search for new phenomena with photon + jet events in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Journal of High Energy Physics **2016**, [http://dx.doi.org/10.1007/JHEP03\(2016\)041](http://dx.doi.org/10.1007/JHEP03(2016)041).
- [34] CMS, *Search for heavy resonances and quantum black holes in $e\mu$, $e\tau$, and $\mu\tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2023**, [http://dx.doi.org/10.1007/JHEP05\(2023\)227](http://dx.doi.org/10.1007/JHEP05(2023)227).
- [35] CMS Collaboration, A. Tumasyan et al., *Search for resonances in events with photon and jet final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **12** (2023) 189, [arXiv:2305.07998](https://arxiv.org/abs/2305.07998) [hep-ex].
- [36] CMS, *Search for new physics with dijet angular distributions in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2017** (2017), [http://dx.doi.org/10.1007/JHEP07\(2017\)013](http://dx.doi.org/10.1007/JHEP07(2017)013).
- [37] D.-C. Dai, C. Issever, E. Rizvi, G. Starkman, D. Stojkovic, and J. Tseng, *Manual of BlackMax. A black-hole event generator with rotation, recoil, split branes, and brane tension. Version 2.02*, Comput. Phys. Commun. **236** (2019) 285–301, [arXiv:0902.3577](https://arxiv.org/abs/0902.3577) [hep-ph].
- [38] J. A. Frost, J. R. Gaunt, M. O. Sampaio, M. Casals, S. R. Dolan, M. A. Parker, and B. R. Webber, *Phenomenology of production and decay of spinning extra-dimensional black holes at hadron colliders*, Journal of High Energy Physics **2009** (2009) 014–014, <https://doi.org/10.1088/1126-6708/2009/10/014>.
- [39] S. Ansoldi, P. Nicolini, A. Smailagic, and E. Spallucci, *Non-commutative geometry inspired charged black holes*, Physics Letters B **645** (2007) 261–266, <https://doi.org/10.1016/j.physletb.2006.12.020>.

- [40] CMS, *Search for dark photons in decays of Higgs bosons produced in association with Z bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2019** (2019), [http://dx.doi.org/10.1007/JHEP10\(2019\)139](http://dx.doi.org/10.1007/JHEP10(2019)139).
- [41] ATLAS, *Search for dark photons from Higgs boson decays via ZH production with a photon plus missing transverse momentum signature from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Journal of High Energy Physics **2023** (2023), [http://dx.doi.org/10.1007/JHEP07\(2023\)133](http://dx.doi.org/10.1007/JHEP07(2023)133).
- [42] CMS, *Search for dark photons in Higgs boson production via vector boson fusion in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2021** (2021), [http://dx.doi.org/10.1007/JHEP03\(2021\)011](http://dx.doi.org/10.1007/JHEP03(2021)011).
- [43] ATLAS, *Observation of electroweak production of two jets in association with an isolated photon and missing transverse momentum, and search for a Higgs boson decaying into invisible particles at 13 TeV with the ATLAS detector*, The European Physical Journal C **82** (2022), <http://dx.doi.org/10.1140/epjc/s10052-021-09878-z>.
- [44] ATLAS Collaboration,, *Constraining the Dark Sector with the monojet signature in the ATLAS experiment*, tech. rep., CERN, Geneva, 2021.
<https://cds.cern.ch/record/2772627>. All figures including auxiliary figures are available at
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-020>.
- [45] CMS, *Search for long-lived particles decaying into muon pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV collected with a dedicated high-rate data stream*, Journal of High Energy Physics **2022** (2022), [http://dx.doi.org/10.1007/JHEP04\(2022\)062](http://dx.doi.org/10.1007/JHEP04(2022)062).
- [46] CMS, *Search for long-lived particles decaying to a pair of muons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2023** (2023), [http://dx.doi.org/10.1007/JHEP05\(2023\)228](http://dx.doi.org/10.1007/JHEP05(2023)228).
- [47] ATLAS, *Search for light long-lived neutral particles that decay to collimated pairs of leptons or light hadrons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Journal of High Energy Physics **2023** (2023), [http://dx.doi.org/10.1007/JHEP06\(2023\)153](http://dx.doi.org/10.1007/JHEP06(2023)153).
- [48] ATLAS Collaboration,, *Search for light long-lived neutral particles from Higgs boson decays via vector-boson-fusion production from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep., CERN, Geneva, 2023.
<https://cds.cern.ch/record/2870215>. All figures including auxiliary figures are available at
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2023-051>.
- [49] CMS, *Search for direct production of GeV-scale resonances decaying to a pair of muons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2023** (2023), [http://dx.doi.org/10.1007/JHEP12\(2023\)070](http://dx.doi.org/10.1007/JHEP12(2023)070).

- [50] CMS, *A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV*, Physics Letters B **796** (2019) 131–154, <http://dx.doi.org/10.1016/j.physletb.2019.07.013>.
- [51] ATLAS, *A search for prompt lepton-jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Journal of High Energy Physics **2016** (2016), [http://dx.doi.org/10.1007/JHEP02\(2016\)062](http://dx.doi.org/10.1007/JHEP02(2016)062).
- [52] ATLAS, *Search for resonant production of dark quarks in the dijet final state with the ATLAS detector*, Journal of High Energy Physics **2024** (2024), [http://dx.doi.org/10.1007/JHEP02\(2024\)128](http://dx.doi.org/10.1007/JHEP02(2024)128).
- [53] ATLAS, *Search for non-resonant production of semi-visible jets using Run 2 data in ATLAS*, Physics Letters B **848** (2024) 138324, <http://dx.doi.org/10.1016/j.physletb.2023.138324>.
- [54] CMS, *Search for resonant production of strongly coupled dark matter in proton-proton collisions at 13 TeV*, Journal of High Energy Physics **2022** (2022), [http://dx.doi.org/10.1007/JHEP06\(2022\)156](http://dx.doi.org/10.1007/JHEP06(2022)156).
- [55] CMS, *Search for new particles decaying to a jet and an emerging jet*, Journal of High Energy Physics **2019** (2019), [http://dx.doi.org/10.1007/JHEP02\(2019\)179](http://dx.doi.org/10.1007/JHEP02(2019)179).
- [56] CMS Collaboration, *Search for long-lived particles decaying to a pair of muons in pp collisions at $\sqrt{s} = 13.6$ TeV with 2022 data*, tech. rep., CERN, Geneva, 2023. <https://cds.cern.ch/record/2868338>.
- [57] M. Low, *Communications with Scalar Cascade Theorists*, Personal communication.
- [58] K. R. Dienes, D. Kim, H. Song, S. Su, B. Thomas, and D. Yaylali, *Nonminimal dark sectors: Mediator-induced decay chains and multijet collider signatures*, Physical Review D **101** (2020), <https://doi.org/10.1103/PhysRevD.101.075024>.
- [59] K. Dickerson, *This 17-mile circular machine is filled with proton beams moving at nearly the speed of light — and it just got a big upgrade*, <https://www.businessinsider.com/tour-of-large-hadron-collider-lhc-particle-accelerator-2015-4>, 2015. Accessed: 2023-02-4.
- [60] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. <https://cds.cern.ch/record/823808>.
- [61] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. <https://cds.cern.ch/record/782076>.
- [62] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. <https://cds.cern.ch/record/815187>.

- [63] E. Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, <https://cds.cern.ch/record/2636343>, General Photo.
- [64] ATLAS Collaboration, M. Capeans, G. Darbo, K. Einsweiler, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010. <https://cds.cern.ch/record/1291633>.
- [65] W. Lukas, *ATLAS inner tracking detectors: Run 1 performance and developments for Run 2*, tech. rep., CERN, Geneva, 2016. <https://cds.cern.ch/record/1956718>.
- [66] N. Garelli, *Performance of the ATLAS Detector in Run-2*, EPJ Web Conf. **164** (2017) 01021, <https://cds.cern.ch/record/2310929>.
- [67] J. G. Panduro Vazquez, *The ATLAS Data Acquisition System: from Run 1 to Run 2*, tech. rep., CERN, Geneva, 2016. <https://cds.cern.ch/record/1954156>.
- [68] C. ATLAS, *Public ATLAS Luminosity Results for Run-2 of the LHC*, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [69] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003, <http://cds.cern.ch/record/1129811>, Also published by CERN Geneva in 2010.
- [70] F. Kuger, *Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel*, 2017. <https://cds.cern.ch/record/2277011>. Presented 31 Jul 2017.
- [71] ATLAS Collaboration, G. Aad et al., *Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons*, The European Physical Journal C **70** (2010) 755–785, <http://dx.doi.org/10.1140/epjc/s10052-010-1403-6>.
- [72] ATLAS Collaboration, G. Aad et al., *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996. <https://cds.cern.ch/record/331061>.
- [73] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. <https://cds.cern.ch/record/782076>.
- [74] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS trigger system in 2015*, The European Physical Journal C **77** (2017), <http://dx.doi.org/10.1140/epjc/s10052-017-4852-3>.
- [75] Cornelissen, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, <https://cds.cern.ch/record/1020106> (2007).
- [76] C. ATLAS, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC*, The European Physics Journal C (2017).

- [77] C. ATLAS, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, The European Physics Journal C (2017), <https://link.springer.com/article/10.1140/epjc/s10052-017-5004-5>.
- [78] ATLAS, *Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data*, JINST 14 (2019).
- [79] ATLAS, *“Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, The European Physics Journal C (2016).
- [80] ATLAS, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, The European Physics Journal C (2018), <https://cds.cern.ch/record/2305380>.
- [81] G. S. Matteo Cacciari, Gavin P. Salam, *The anti- k_t jet clustering algorithm*, Journal of High Energy Physics (2008), <https://iopscience.iop.org/article/10.1088/1126-6708/2008/04/063>.
- [82] D. E. S. Stephen D. Ellis, *Successive combination jet algorithm for hadron collisions*, Physics Review D **48** (1993), <https://journals.aps.org/prd/abstract/10.1103/PhysRevD.48.3160>.
- [83] ATLAS, *“Tagging and suppression of pileup jets with the ATLAS detector”*, Technical Report ATLAS-CONF-2014-018 (2014).
- [84] C. ATLAS, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, The European Physics Journal C (2017), <https://arxiv.org/pdf/1703.10485>.
- [85] Thaler, *Identifying Boosted Objects with N-subjettiness*, JHEP (2011), <https://doi.org/10.48550/arXiv.1011.2268>.
- [86] ATLAS, *The ATLAS trigger system for LHC Run 3 and trigger performance in 2022*, Journal of Instrumentation **19** (2024) P06029, <http://dx.doi.org/10.1088/1748-0221/19/06/P06029>.
- [87] M. C. Aleksa, W. P. Cleland, Y. T. Enari, M. V. Fincke-Keeler, L. C. Hervas, F. B. Lanni, S. O. Majewski, C. V. Marino, and I. L. Wingerter-Seez, *ATLAS Liquid Argon Calorimeter Phase-I Upgrade: Technical Design Report*, tech. rep., 2013. <https://cds.cern.ch/record/1602230>. Final version presented to December 2013 LHCC.
- [88] M. Weirich et al., *Development of the jet Feature EXtractor (jFEX) for the ATLAS Level 1 calorimeter trigger upgrade at the LHC*, PoS **TWEPP-17** (2017) 133.
- [89] K. Kohler, *Neural Networks / Artificial Neural Networks*, https://indico.cern.ch/event/711895/contributions/3132810/attachments/1712089/2760482/TDAQ_jFEX_algorithms.pdf, 2020. Accessed: 2019-05-24.
- [90] Jeff Dandoy, *IsolatedJetTree*, <https://gitlab.cern.ch/atlas-jetetmiss-jesjer/MCcalibrations/IsolatedJetTree>.

- [91] D. M. Gingrich, *Noncommutative geometry inspired black holes in higher dimensions at the LHC*, Journal of High Energy Physics **2010** (2010), [http://dx.doi.org/10.1007/JHEP05\(2010\)022](http://dx.doi.org/10.1007/JHEP05(2010)022).
- [92] ATLAS, *Search for contact interactions and large extra dimensions in the dilepton channel using proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, The European Physical Journal C **74** (2014), <http://dx.doi.org/10.1140/epjc/s10052-014-3134-6>.
- [93] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, Journal of High Energy Physics **2014** (2014), [http://dx.doi.org/10.1007/JHEP07\(2014\)079](http://dx.doi.org/10.1007/JHEP07(2014)079).
- [94] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, *UFO - The Universal FeynRules Output*, Comput. Phys. Commun. **183** (2012) 1201–1214, [arXiv:1108.2040](https://arxiv.org/abs/1108.2040) [hep-ph].
- [95] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, and C. Wanotayaroj, *Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, Journal of High Energy Physics **2015** (2015), [http://dx.doi.org/10.1007/JHEP02\(2015\)075](http://dx.doi.org/10.1007/JHEP02(2015)075).
- [96] Jeff Dandoy, *xAODAnaHelpers*, <https://github.com/UCATLAS/xAODAnaHelpers>.
- [97] ATLAS, *Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TRIG-2019-01/>, 2024. Accessed: 2024-02-09.
- [98] Agostinelli, “*GEANT4 – a simulation toolkit*”, Instrum. Meth. (2003).
- [99] Hoeche, “*Sherpa Manual Version 2.2.11*”, <http://sherpa.hepforge.org> (2021).
- [100] Alioli, “*The POWHEG BOX*”, <https://powhegbox.mib.infn.it/> (2011).
- [101] Bierlich, “*A comprehensive guide to the physics and usage of PYTHIA 8.3.2022*”, [arXiv:2203.11601](https://arxiv.org/abs/2203.11601) (2022).
- [102] ATLAS, “*Track Reconstruction Performance of the ATLAS Inner Detector*”, ATLAS-PHY-PUB-2015-018 (2015).
- [103] ATLAS, “*Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s}=13$ TeV*”, Eur. Phys. J. C (2021).
- [104] ATLAS, “*Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s}=13$ TeV*”, Eur. Phys. J. C (2019).
- [105] ATLAS, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, Technical report ATLAS-CONF-2015-029 (2015).

- [106] A. Patil, *Neural Networks / Artificial Neural Networks*, <https://aniketpatilvashi.medium.com/neural-networks-artificial-neural-networks-70a17ce096e1>, 2020. Accessed: 2023-07-7.
- [107] A. Connolly, *Deep Learning: Classifying Astronomical Images*, https://www.astroml.org/astroML-notebooks/chapter9/astroml_chapter9_Deep_Learning_Classifying_Astronomical_Images.html, 2023. Accessed: 2023-07-10.
- [108] P. Baheti, *Activation Functions in Neural Networks [12 Types & Use Cases]*, <https://www.v7labs.com/blog/neural-networks-activation-functions>, 2021. Accessed: 2023-07-21.
- [109] D. Godoy, *Understanding binary cross-entropy / log loss: a visual explanation*, <https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-a-visual-explanation-a3ac6025181a>, 2021. Accessed: 2023-07-25.
- [110] S. Ruder, *An overview of gradient descent optimization algorithms*, [https://www.ruder.io/optimizing-gradient-descent/#:~:text=Stochastic%20gradient%20descent,-Stochastic%20gradient%20descent&text=%CE%B8%3D%CE%B8%E2%88%92%CE%B7%E2%8B%85%E2%88%87%CE%B8J\(%CE%B8%3B,%3B%20y%20\(%20i%20\)%20\)%20](https://www.ruder.io/optimizing-gradient-descent/#:~:text=Stochastic%20gradient%20descent,-Stochastic%20gradient%20descent&text=%CE%B8%3D%CE%B8%E2%88%92%CE%B7%E2%8B%85%E2%88%87%CE%B8J(%CE%B8%3B,%3B%20y%20(%20i%20)%20)%20), 2016. Accessed: 2023-07-13.
- [111] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Póczos, R. Salakhutdinov, and A. Smola, *Deep Sets*, 2018.
- [112] P. T. Komiske, E. M. Metodiev, and J. Thaler, *Energy flow networks: deep sets for particle jets*, *Journal of High Energy Physics* **2019** (2019), <https://doi.org/10.1007/2Fjhep01%282019%29121>.
- [113] G. Kasieczka and D. Shih, *Robust Jet Classifiers through Distance Correlation*, *Physical Review Letters* **125** (2020), <http://dx.doi.org/10.1103/PhysRevLett.125.122001>.
- [114] Ben Nachman, *DisCo*, <https://github.com/bnachman/DisCo/blob/master/DisCo.ipynb>.
- [115] G. Kasieczka, B. Nachman, M. D. Schwartz, and D. Shih, *Automating the ABCD method with machine learning*, *Physical Review D* **103** (2021), <https://doi.org/10.1103/2Fphysrevd.103.035021>.
- [116] ROOT Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., New York, 2012. <https://cds.cern.ch/record/1456844>.
- [117] K. Cranmer, *Practical Statistics for the LHC*, 2015.
- [118] Iacopo Longarini, *ABCD*, <https://gitlab.cern.ch/atlas-phys/exot/ueh/displaced-lj/abcdfit/-/tree/master>.
- [119] ATLAS, “*Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*”, Technical report ATLAS-CONF-2014-018 (2019).